Infrared Physics & Technology 85 (2017) 478–501

Contents lists available at ScienceDirect

Infrared Physics & Technology

journal homepage: www.elsevier.com/locate/infrared

Review

A survey of infrared and visual image fusion methods

Xin Jin [a], Qian Jiang [a], Shaowen Yao [b,*], Dongming Zhou [a,*], Rencan Nie [a], Jinjin Hai [c], Kangjian He [a]

[a] Yunnan University, School of Information, Kunming 650091, China
[b] Yunnan University, School of Software, Kunming 650091, China
[c] National Digital Switching System Engineering & Technological R&D Center, Zhengzhou 450002, China

HIGHLIGHTS

- This paper surveyed the methods of IR and VI image fusion in recent years.
- Different classification strategies for IR and VI image fusion were reported.
- This paper provided a comprehensive display of diverse applications.
- A detailed description of the development in this field was represented.
- This paper concluded the corresponding challenges and tendencies in this field.

ARTICLE INFO

Article history:
Received 25 January 2017
Revised 15 June 2017
Accepted 10 July 2017
Available online 12 July 2017

Keywords:
Image fusion
Infrared and visual image
Image representation
Feature extraction
Image quality assessment

ABSTRACT

Infrared (IR) and visual (VI) image fusion is designed to fuse multiple source images into a comprehensive image to boost imaging quality and reduce redundancy information, which is widely used in various imaging equipment to improve the visual ability of human and robot. The accurate, reliable and complementary descriptions of the scene in fused images make these techniques be widely used in various fields. In recent years, a large number of fusion methods for IR and VI images have been proposed due to the ever-growing demands and the progress of image representation methods; however, there has not been published an integrated survey paper about this field in last several years. Therefore, we make a survey to report the algorithmic developments of IR and VI image fusion. In this paper, we first characterize the IR and VI image fusion based applications to represent an overview of the research status. Then we present a synthesize survey of the state of the art. Thirdly, the frequently-used image fusion quality measures are introduced. Fourthly, we perform some experiments of typical methods and make corresponding analysis. At last, we summarize the corresponding tendencies and challenges in IR and VI image fusion. This survey concludes that although various IR and VI image fusion methods have been proposed, there still exist further improvements or potential research directions in different applications of IR and VI image fusion.

Contents

* Corresponding author.
E-mail addresses: yaosw@ynu.edu.cn (S. Yao), zhoudm@ynu.edu.cn (D. Zhou).

http://dx.doi.org/10.1016/j.infrared.2017.07.010
1350-4495/

1. Introduction

In the natural environment, objects will radiate electromagnetic wave with different frequencies, called thermal radiation, which cannot be seen by human eyes [1,2]. IR images are taken by infrared sensor to record the thermal radiation of different objects, these could be used in ground object identification and surface parameters inversion, such as hidden targets and identification of camouflage [1,3]. The characteristics of IR images would reduce the influence of external environment, such as sunlight, smog and other condition factors [1,4]. IR images are sensitive to the objects and areas with obviously infrared thermal characteristics which could not be represented in VI images [5,6]. VI images are taken to record the visibly reflective properties of spectrum information of the objects, they contain numerous visible edges and details of objects, which could provide a perceptual scene description for human eyes and are consistent with human visual characteristics [1,5]. The goal of IR and VI image fusion is to obtain a complementary fused image with abundant detail information in VI images and effective target areas in IR images [5]. As a result, IR and VI image fusion techniques are widely used in night-vision imaging equipment which could improve the nocturnal ability of human and robot [7]. The rapid expansions of computer imaging technologies and sensor technologies will help to improve the quality and effect of night-vision imaging equipment, which often lead to robust information processing for robot vision, and reveal information that is invisible to human eyes [8]. The accurate, reliable and complementary descriptions of the scene in fused IR and VI images make these techniques be widely applied to military surveillance [9,10,11], agricultural automation [12,13], industrial applications [14,15,16], biometric identification [17–35], medical imaging [36], remote sensing [37–44], electronics testing [2], space exploration [2,47], the generation of thermographic 3D model of the building [50] and the examination of ancient documents [48,49].

In recent years, image fusion gradually becomes a hot research field, and IR and VI image fusion techniques are the important part of image fusion field. Depending on different applications, various IR and VI image fusion methods have been developed, and these methods could be generally divided into three levels: pixel level, feature level and decision level (symbol level) [45]. Pixel-level based method is the popular research tendency for the whole image fusion field, because it has minimum artifacts in the fused image whose pixels are determined from a set of image pixels or other forms of image parameters at the lowest physical level [51]. Besides, the higher fusion levels, such as feature level or symbol level, are the combinations of the information in the form of image feature descriptors and probabilistic variables [52].

Firstly, the pixel level based IR and VI image fusion methods could be roughly divided into spatial domain based algorithms and transform domain based algorithms. The typical spatial fusion algorithms are average weighted and block based methods; the methods based on subspace analysis also are used in this field, such as PCA [104] and ICA [138]. The transform based IR and VI fusion algorithms are pyramids [2,13,77], wavelet [26,28,31,35], curvelet [41,129] and contourlet [40,104]; besides the novel multi-scale decomposition methods are NSCT [84,7], NSST [1,93,95] and other decomposition methods [5,28,30]. In addition to the methods mentioned above, there are many others image representation methods used in IR and VI image fusion, such as compressive sensing (CS) [3,42] and sparse representation (SR) [111]. In recent years, Markov random fields (MRFs) [99,100] and robust principal component analysis (RPCA) [133,135] have been used in IR and VI images fusion field. Secondly, feature level based image fusion methods could deal with the comprehensive feature information and structural characteristics of image, such as edges, contours, textures and corners, to enhance the image and get the segmentation or region target distribution information of the image [15,26,32,55]. And then, the information from each source image will be separately extracted and combined [57]. The typical methods are based on region target detection [58], edge extraction [60], regional saliency extraction [5] and region segmentation [128,131]. Thirdly, decision-level fusion is the highest level of the three levels, the information is determined and identified from different source images; and then decisions will be taken to fuse the extracted information according to specific criteria [57,60,78]. However, the researches of decision level image fusion is the least, and mainly are used in face recognition [79–82]; the fusion strategies are generally based on neural classifiers [79], confidence measure [80] and D-S evidence theory [82]. Each level of IR and VI image fusion methods constitutes a different set of image processing methods; and the methods are applied to different domains by using different arithmetic operations [57]. Some researchers also adopt the combination of two level methods to accomplish the IR and VI image fusion task [32,60]. The combinations of different level methods could integrate the corresponding advantages together, which can provide more possible to obtain good fusion effect.

In keeping with the most of literatures in IR and VI image fusion, these methods also can be divided into two categories: pixel-based fusion and region-based fusion [45,101]. Pixel-based IR and VI image fusion methods mainly rely on image representation methods; the image detailed information will be analyzed at pixel level, then the fused image will be obtained according to some fusion strategies. IR is sensitive to the objects whose infrared radiation is more obvious, the objects will have obviously salient characteristics; as a result, the region-based IR and VI image fusion methods would get the salient region or saliency map by segmentation or image saliency techniques. The salient information (such as, salient region or saliency map) of IR image could be directly fused in the final image to insure it retain as much infrared information as possible. The most critical step of region-based fusion method is how to accurately extract the salient region in IR image. Region-based method is an effective way to fuse the detailed information of VI image and the target regions of IR image into the final result [5].

No matter which taxonomy it is, it should consider the following two problems which were described by Wang et al. [111]. (i) How to effectively extract the image information from the input source images, which is the key factor to image fusion quality. (ii) How to reasonably fuse the information from multiple information sources into the finally fused image. In order to solve above two problems, and by reference to Piella's thought [45], we think the fusion method should have the following abilities: (i) The method should be able to extract the complementary information of different images, such as the thermal characteristics information of IR images and the detailed information of VI images, which mainly relies on the resolution of the image representation algorithm for image structure information; (ii) The method should be able to accurately fuse the complementary information into the final image from the multiple source images, which mainly relies on the fusion strategies; (iii) The whole process should do not introduce any artifacts or inconsistencies which can distract or mislead a human observer or any subsequent image processing steps [45].

Many advanced image representation methods and fusion strategies have been proposed to accomplish IR and VI image fusion task [111]. As a result, the fusion effect has been rapidly improved. However, we find that there are still many problems which are not fully studied or solved. On one hand, the existing IR and VI image fusion technologies would not be able to meet the rapid development of the computer and sensor technologies; meanwhile, and the previously impossible to implement fusion technologies would be got satisfied. The advanced computing devices and sensors with more superior performance are continuously developed, which would further drive the development of IR and VI image fusion technologies. On the other hand, there will be raised more and more new demands with the people's life way keep changing. For example, biological recognition techniques make people's lives be more convenient and safer, and the IR and VI image fusion technologies are widely applied to these techniques. Certainly, the important applications of IR and VI image fusion are military field, for example, airborne infrared detection equipment work in the air, so they have special requirements and continuously keep putting forward new demands, especially, they need lighter and efficient equipment on unmanned aircraft. For another example, the portable night vision equipment need faster and better image quality to adapt soldiers to complex battlefield environment.

In recent years, there have been proposed a large number of approaches in image fusion field due to the rapid development of the imaging techniques [45,112]. The main applications of image fusion are medical image fusion [8,113] and remote sensing image fusion [85,114–116] which have been well studied, and many scholars have published corresponding review papers in these areas or technosphere [4,8,23,113,116,117]. However, another common image fusion field of IR and VI image has not published an integrated review paper about the last several years. As a result, this paper represents an overview of IR and VI image fusion techniques and applications, and the product assessment methods are also presented, besides future improvements and potential research directions are discussed as well. This survey concludes that there still exists some challenges and chances in different IR and VI image fusion applications, although there are various IR and VI image fusion methods have been proposed; especially, the drive of new applications, the development of devices and image analysis theories, the requirement of low cost and easy-use, the performance of anti-noise and so on. Because there is not a complete survey of IR and VI image fusion techniques about the recent several years, this paper tries to provide a detailed description of the current development and represent a comprehensive display of various applications, which hope to describe a panoramically perspective view to the current research status of this field.

This paper is arranged as follows. In Section 2, we will provide a detailed description of the applications and technospheres to represent the demands of IR and VI image fusion technique. Section 3 will provide a taxonomical overview of the different families of image fusion methods, especially different image representation methods. Section 4 will investigate common IR and VI image fusion assessments metrics. Section 5 will introduce some experiments of typical and important methods. Section 6 will present the main future trends and challenges. Section 7 concludes this paper.

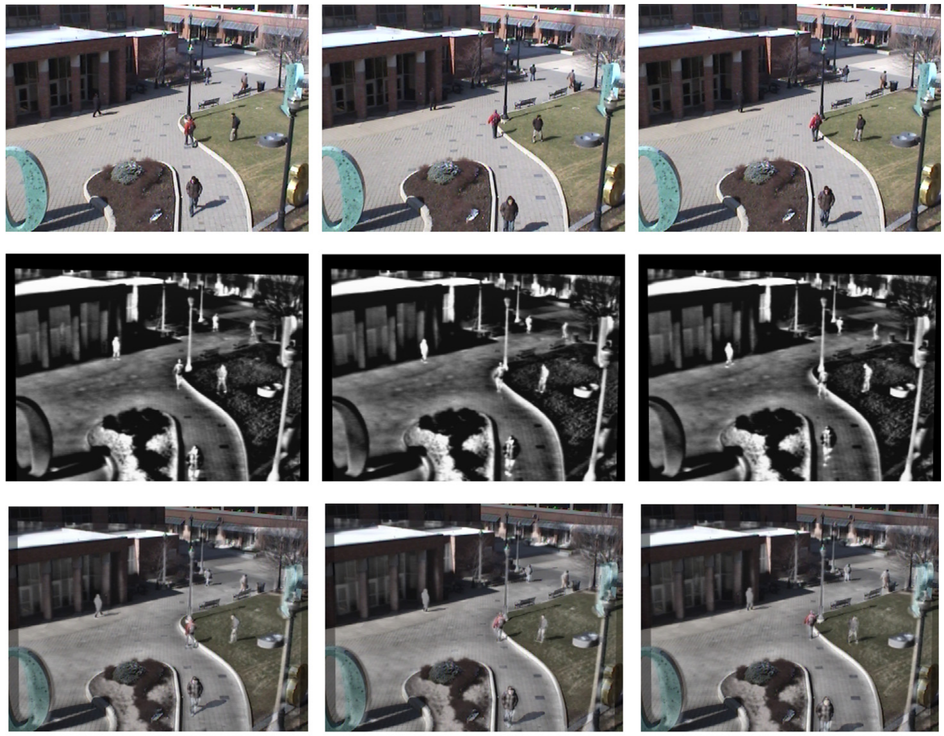
2. Main application fields and technospheres

In addition to the traditional applications that are introduced in other papers, we find that the IR and VI image fusion techniques also can be used as the part of numerous practical techniques, such as video fusion, night color vision, biological recognition and other applications. In this section, the main application fields and technospheres of IR and VI image fusion will be introduced.

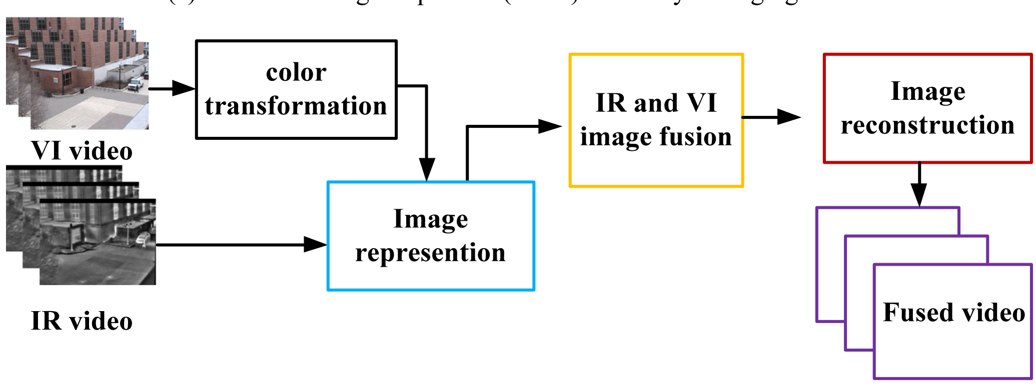
2.1. Infrared and visible video fusion

Most of practical applications need to fuse IR and VI videos or dynamic images, such as surveillance, agricultural automation, industrial applications, electronics testing and so on, which are important to video techniques; however, the study of this field was rare [41,53]. IR and VI video fusion approaches can effectively fuse the complementary information from multiple source videos into a fused video, and it can improve the robustness and accuracy of the video [21,41,53], the examples of IR and VI image video (sequences) fusion and a generally basic fusion flow diagram are shown in Fig. 1. Compared with general IR and VI image fusion techniques, the video fusion methods need to have the characteristics of high computational efficiency and low space requirements [53]. Beyond that, with the rapid development of technologies and equipment, we think that video processing techniques would gradually become an interesting area and attract more attentions. In this section, we will summarize recent video fusion researches.

By considering the stability and consistency of the information in spatial and temporal dimension, two IR and VI video fusion techniques were proposed by Xu et al. [41,53], one of the methods was based on uniform discrete curvelet transform and spatial-temporal information [41]; another spatiotemporal video fusion algorithm was based on motion compensation in the wavelet-transform domain [53].



(a) IR and VI image sequences (video) fusion by averaging method



(b) A general IR and VI image sequences (video) fusion flow diagram

Fig. 1. IR and VI image video (sequences) fusion.

Denman et al. investigated many methods for fusing visual and thermal images for person tracking and abandoned object detection, and they proposed a modified condensation filter to track and aid in the fusion of the modalities, and obtained interesting results in the field of target tracking [54]. Chan et al. performed 13 pixel-based image fusion algorithms and examined corresponding performances for the detection and tracking by a given target tracker, and they identified five fusion methods that produced significantly better performance, three of them also managed to achieve a relatively high efficiency [55]. Zhang et al. made similar study [56].

Liu et al. employed target detection method to improve the performance of the feature-based IR and VI dynamic images fusion technique based on dual-tree complex wavelet transform (DT-CWT) which was used to decompose all the source image sequences [58]. Bennett et al. presented a technique for enhancing underexposed visible-spectrum video by fusing it with simultaneously captured video from sensors in nonvisible spectra. They showed that the RGB and IR video streams could be captured by using same optical path, which could be fused into an enhanced version of RGB video [59]. MRF is used to fuse IR and VI video source, in Han's research, the main idea of saliency detection was to generate a saliency map for IR image, and the saliency of the region drives the fusion procedure [101]. This method achieved a near real-time performance (3–4 pairs of frames per second) according to the computing time of the videos with different resolutions.

IR and VI video fusion should simultaneously take into account both temporal and spatial dimensions of the video, and it is a form of concrete image fusion application [41,53]; however, traditional static image fusion technologies cannot be directly used for IR and VI video fusion due to the limitation of timeliness and computing resources. As a result, the most important points of IR and VI video fusion are speed and quality; however, it is very difficult to balance these two. At present, most of the works mainly focus on the fusion quality, and have not paid enough attention to the speed. Generally speaking, it could provide well vision perception for human when the processing speed of video reaches about 24

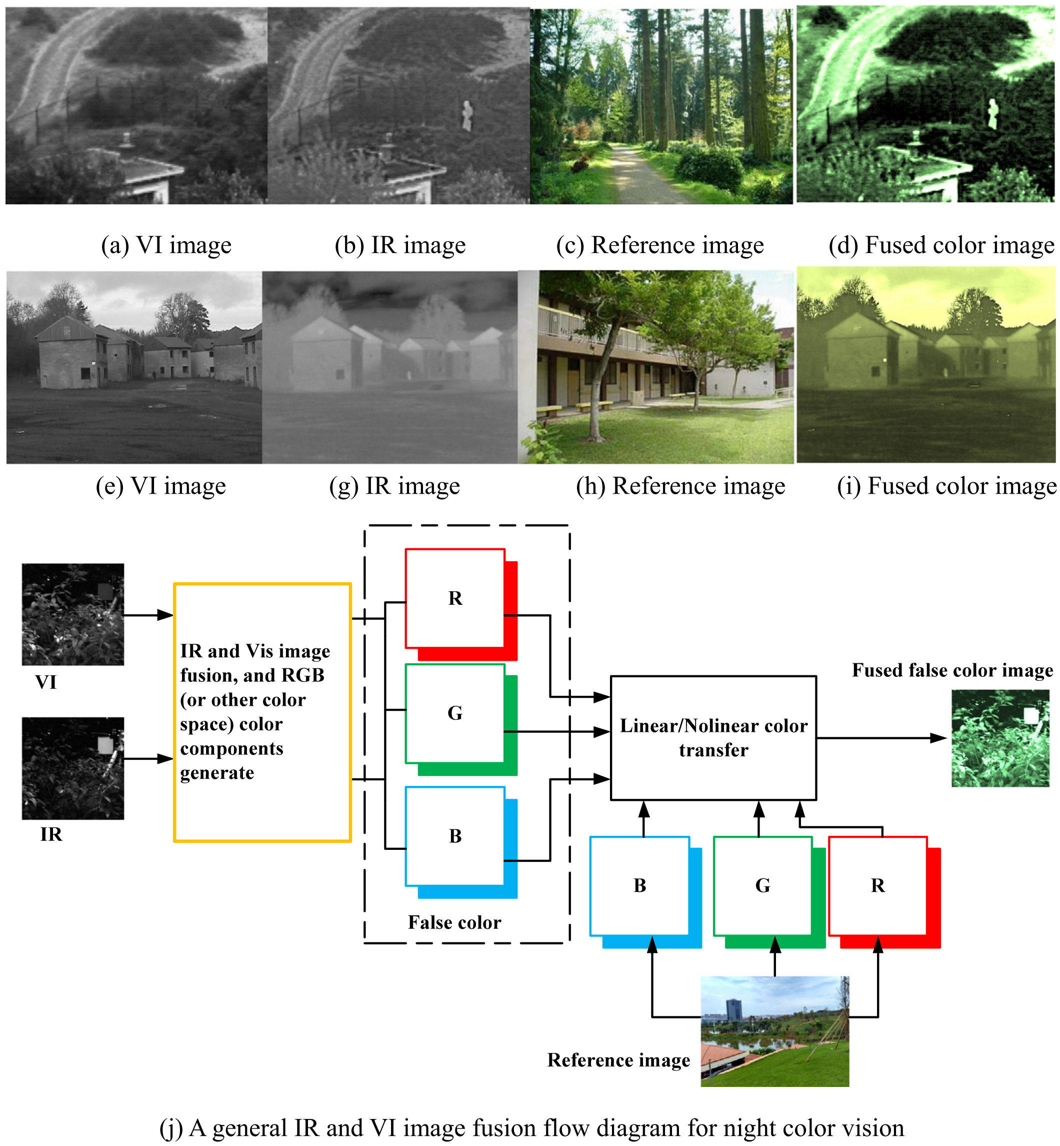


Fig. 2. IR and VI image fusion for night color vision.

frames per second (FPS); however, most of the IR and VI video fusion methods could not reach this standard. More unfortunately, few researchers used FPS as the evaluation metrics of IR and VI fusion. The slow processing speed lead to its limitation which is not easy used for real-time video surveillance applications. The advanced transform based IR and VI video fusion algorithm usually suffer the high computational complexity, in spite of high video quality [21,54]. In addition to this, different color space of IR and VI video fusion methods also would have different performances, such as color, artifact and so on, which also need to be considered.

2.2. Infrared and visible image fusion for night color vision

IR and VI image fusion for night color representation is the combination approach for fusing night-time IR with VI imagery [61,62]. The final obtained color images possess a natural day-time color appearance due to the application of color transfer technique which will help the observer to understand the image by making scene interpretation more intuitive [61], two examples of IR and VI image fusion for night color vision and a general fusion flow diagram are shown in Fig. 2. The important impetus for this kind of research is that human visual system is more sensitive to color image than gray image; specifically, the human eye can discern several thousand colors, whereas it can only distinguish about 100 shades of gray at any instance [61,62,63], as a results, gray image colorization has been got many attentions from scholars, especially the color contrast enhanced IR and VI image fusion methods were widely applied in military equipment [9]. Besides, the rapid development of multi-band infrared and visual night-vision systems has led to an increased interest in color fused ergonomic representations of multiple image sensor signals [61,62,63,64,65,66,67].

For a long time, there was not an appropriately scheme for false color rendering of night-time imagery, until Reinhard et al. introduced a method to transfer one image's color characteristics to another [66,68]. Subsequently, Toet et al. proposed many methods to complete this task [61,66]. From then on, similar color transfer methods were designed by many researchers [63,64,65]. Perceptual evaluation problems for the quality of IR and VI color image fusion were also discussed in Ref. [66]. And an evaluation for objectively assessing the quality of IR and VI color fusion image was proposed as well [67].

Generally speaking, a fused false color RGB image of IR and VI image will be produced by mapping three individual bands of a multiband night-vision system to the respective channels of a RGB image, as shown in Fig. 2(j). Toet's method first got the channels of a RGB image according to IR and VI image; and then the color transfer methods were adapted to transform the color of reference RGB image into the new RGB image by XYZ, LMS and Lab color space [61]. In the research of Tsagaris' et al., non-negative matrix factorization (NMF) based fusion approach was proposed. It provided an additive part-based representation of source image, and then color transfer was achieved by the transformation of XYZ, LMS and Lab color space [62]. Davis et al. presented a background-subtraction technique by fusing contours from thermal and visible image for persistent object detection in urban settings [64]. Hogervorst et al. introduced a simple and fast method to consistently apply natural daytime colors to multi-band night-vision image [70]. Another simple, fast and easily deployed lookup-table based method that gave multi-band night-time image a consistent natural daytime color appearance also was proposed [71].

The idea of image segmentation was often used in this field [73]. For example, Zheng et al. presented a "local-coloring" method that was used to render the night-vision image segment-by-segment by taking advantages of image segmentation, segment recognition, histogram matching and image fusion techniques [72]. Yin et al. focused on the color transfer step, and introduced a ratio of local to global divergence of the IR image to improve color contrast. The enhancement ratios for both hot and cold targets were larger than one, while it tended to one for the background. Thus both hot and cold targets could be enhanced to keep the details of the background [69]. Qian et al. discussed a simple fusion method of IR and VI image for night color vision, local histogram equalization was introduced to improve IR and VI images' quality before fusion operation, an amended global color transform method was adopted to get the natural colors and improve the target detection ability, the targets were popped out and meanwhile the backgrounds were rendered with natural colors by stretch factor [74]. And another method also was introduced in Ref. [75].

Besides, Guo et al. focused on the influence of color characteristics of typical scenes and different sizes of color regions, and they proposed a color-combination harmony models for color fusion images of three typical scenes (sea and sky, plants, and towns and buildings) [76]. Yu et al. proposed a false color image fusion method for merging IR and VI images in transform domain based on Laplacian pyramid transform and YC_BC_R color space, which

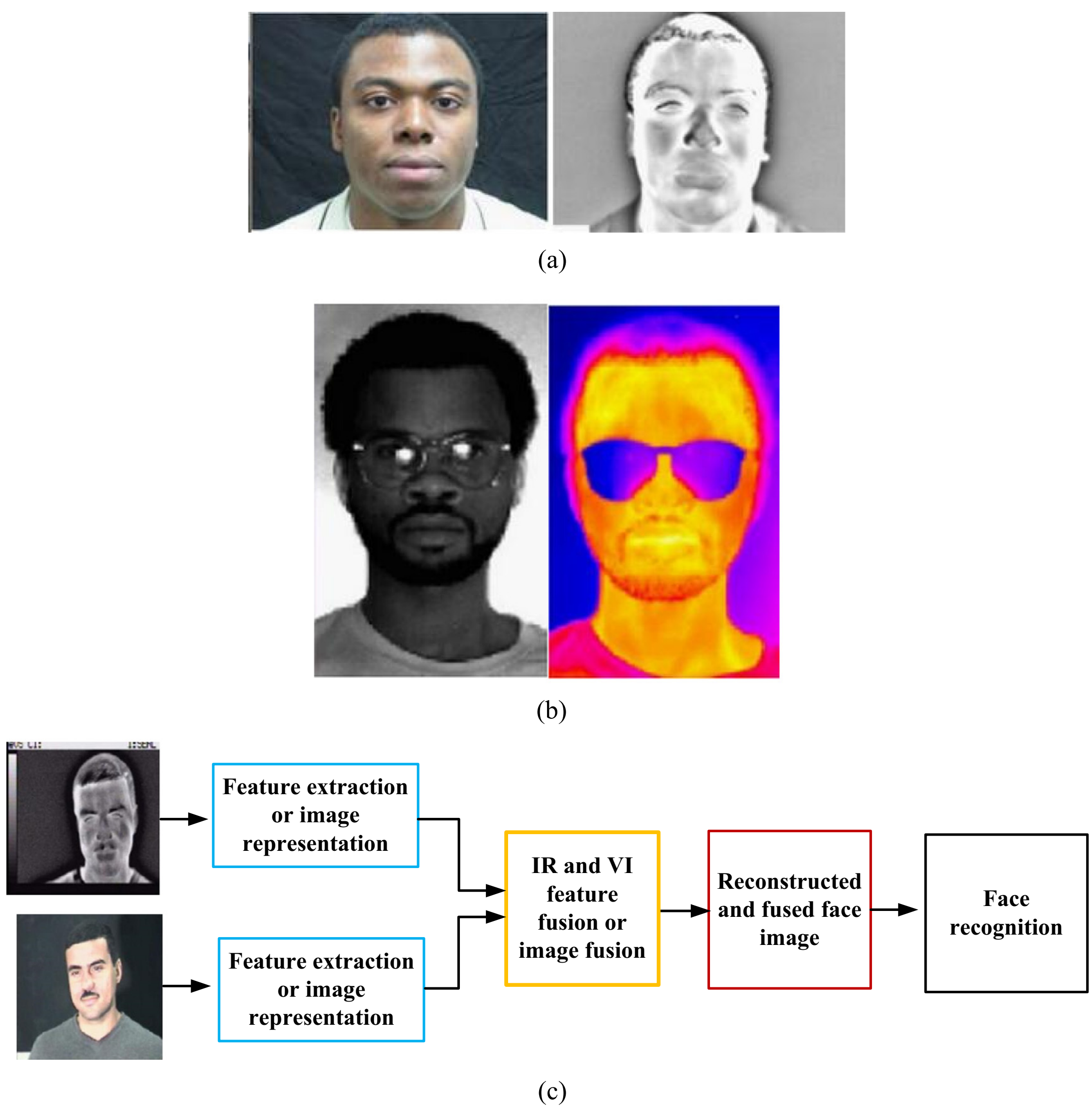


Fig. 3. Face recognition of IR and VI image fusion. (a) A corresponding pair of VI and IR face images from the IRIS-M3 database [117]. (b) A corresponding pair of VI and false color thermal images [117,118]. (c) A general face recognition flow diagram based on IR and VI image fusion.

proposed a new way to complete this task, however, it might not have a high computational efficiency [77].

Color and detailed information of the fused image are the most important points for this field, and the color mainly rely on color transfer method and reference color image. The detailed information of the fused image depends on the image represent method and fusion rule. At present, all the IR and VI image fusion methods for night color representation need a reference image, and the quality of the final fused color image are seriously dependent on it. Each reference image is corresponded to the scene of fused IR and VI images, which lead to the poor flexibility and adaptability of the image fusion algorithm. Therefore, the dependence of reference image should be reduced in the fusion methods of IR and VI image for night color representation, and it is a problem demanding prompt solution. Besides, as with other image processing fields, there is not a single uniform metric for IR and VI color fusion image. Due to the particularity of this image fusion field, more reasonable and effective metrics methods to evaluate the quality of color image fusion need to be studied. And the transform based method achieved a very good effect, and may be a promising research direction. On account of the fused color image could provide more information than the gray one and more suitable for human eye vision, the researches in this field are valuable for many practical situations.

2.3. Infrared and visible image fusion used in biological recognition

Face recognition is a rapidly growing research area due to the increasing demands for the security in commercial and law enforcement applications [23]. It is well known that face recognition techniques of visual images have reached a significant level of maturity with many practical successes [23]. However, the recognition rate of visual face based method may degrade under poor illumination conditions or for subjects of various skin colors [23]. While the thermal IR based face recognition techniques will have a good performance when there is no control over illumination or for detecting disguised faces [23]. The examples of face images of IR and VI and a general face recognition flow diagram based on IR and VI image fusion are shown in Fig. 3.

In IR and VI image fusion field, the bionic optimization algorithm is often used for face recognition to increase the recognition rate; the key objective of the optimization is to capture most salient features from each spectrum in different kinds of images [24]. Bebis et al. investigated two different fusion schemes in their study and employed a simple and general framework based on genetic algorithms (GAs) to find an optimum fusion strategy in face recognition [24]. Hermosilla proposed an IR and VI image fusion system by combining the visible and thermal features which were obtained from the most current local matching descriptors, and the method maximized face recognition rates through the use of GAs [27]. In addition to the GAs, particle swarm optimization based IR and VI image fusion method also was designed to solve the problem of face verification [35].

Good match exploration for thermal infrared face recognition method was proposed by Bai et al., the method was based on (Y-styled Window Filter) YWF-(Scale Invariant Feature Transform) SIFT with multi-scale fusion [28]. In Singh's research, two face images from different spectrums were fused by using DWT-based fusion algorithm; the amplitude and phase features were extracted from the fused image using 2D log polar Gabor wavelet at image level fusion [32]. The fusion of visible and long wave infrared face images was performed by using 2-granular SVM; it used multiple SVMs to learn both the local and global properties of the multispectral face images at different granularity levels and resolution [33]. Ma et al. focused on the registration of VI and thermal IR face images for fusion-based face recognition which had acquired a good effect [34]. Performance of different IR and VI fusion methods were also analyzed in Hizem's face recognition study [25].

In addition to face recognition technologies, there are some other applications of biological recognition based on IR and VI image fusion techniques. A ear-based human identification method was researched by Abaza et al. [18]. The idea of the image fusion was used for periocular region-based person identification [20]. Iris recognition based feature fusion method was proposed by the combination of 1D log-Gabor wavelet and 1D Haar wavelet [26]. As a result, we have reason to believe that more and more biometric technologies would be invented.

VI image based face recognition techniques have been well studied, but there still some problems needed to be solved in practical applications. In real scenario, illumination, shooting angle, facial expression variation, decorations, background and so on, which will have seriously impacts on the recognition effect. While, IR image will provide complementary information for face recognition, this information will be not revealed in VI image. Due to the characteristics of IR image, the IR and VI image fusion based face recognition techniques will have a good performance when there is no control over illumination or for detecting disguised faces [23]. In this field, we find that the lightweight transform based IR and VI image fusion algorithms are more suitable than the complex one due to the biological recognition techniques usually need a fast computing speed to overcome mass calculation. In recent years, bionic optimization algorithm based IR and VI image fusion methods are often used in biological recognition, and this kind of method could improve recognition accuracy by additional computational expense; besides, advanced classifier also would contribute to the recognition rate. Due to IR and VI image fusion techniques could provide more complementary information for biological recognition, we think it would receive abiding concern and development.

2.4. Infrared and visible image fusion used in other applications

In addition to above mentioned applications, IR and VI image fusion techniques are widely used in many specific fields. The applications of IR and VI image fusion technique are shown in Fig. 4. In this section, we will introduce some other applications of this kind of technologies.

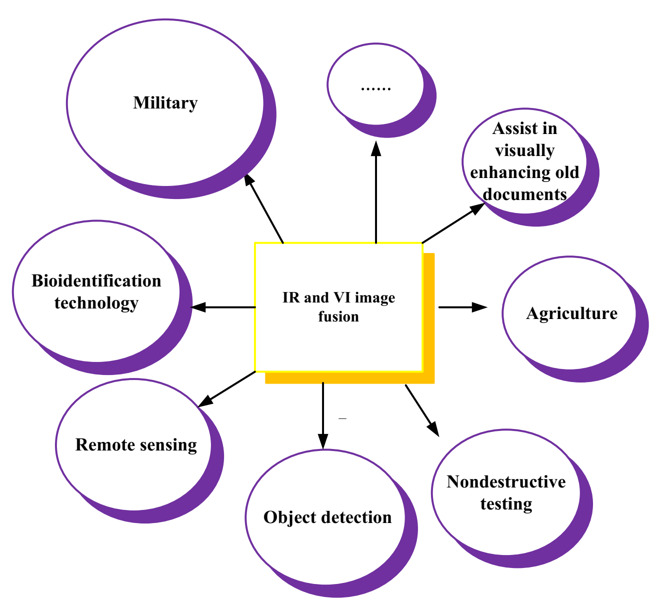


Fig. 4. The applications of IR and VI image fusion technique.

In the field of object detection, IR and VI image has very good characteristics of the complementarity; therefore, it also got some attentions from scholars. For instance, Elguebaly et al. proposed an object detection technique based on IR and VI image fusion method [11], and an fusion method of VI and thermal images for fruit detection was researched by Bulanon et al. [13]. Besides, González et al. proposed a fusion approach to combine the features from both modalities in pedestrian detection [19]. Many IR and VI image fusion methods were discussed in the research of task-based scan-path assessment in complex scenarios [21], and particle filter tracking of camouflaged targets by adaptive fusion of IR and VI image data was proposed by Talha et al. [22].

IR and VI image fusion was widely used in remote sensing field as well. Typical examples as follows: urban object detection was researched by Eslami et al. [37], and geostationary meteorological satellite IR and VI images fusion method based on thermal physical properties was researched by Lei [39]. In Chang's research, a multi-contourlet-based adaptive fusion of IR and VI remote sensing images was proposed [40], similarly, a remote sensing image classification of long-wave IR hyperspectral and VI images was proposed by Lu et al. [43]. Besides, an improved IR and VI image fusion for astronomical images also was researched by Ahmad et al. [47].

IR and VI also was applied in agricultural field, in the study of Mendoza et al., four nondestructive sensing systems were evaluated and combined for predicting apple fruit firmness and soluble solids content based on fused technique [12]. Another representative study was the research of the registration of thermal and visible light images of diseased plants using silhouette extraction in wavelet domain [14].

Computer technology has been widely applied in the protection and analysis of cultural relics [156]. Some researchers devoted themselves to the study of IR and VI image fusion method to assist in visually enhancing old documents. Gargano et al. fused the text which extracted from the IR image with the VI image according to the idea of image enhancement. In this method, a perceptual based approach was proposed to fuse VI and near infrared images in the examination of ancient documents, which was integrated by the combination of automatic color equalization algorithm and a perceptual-based enhancement technique [48]. One of the main contribution of Kim et al. was the way to detect the regions in the visible-spectrum images that could be enhanced by using the

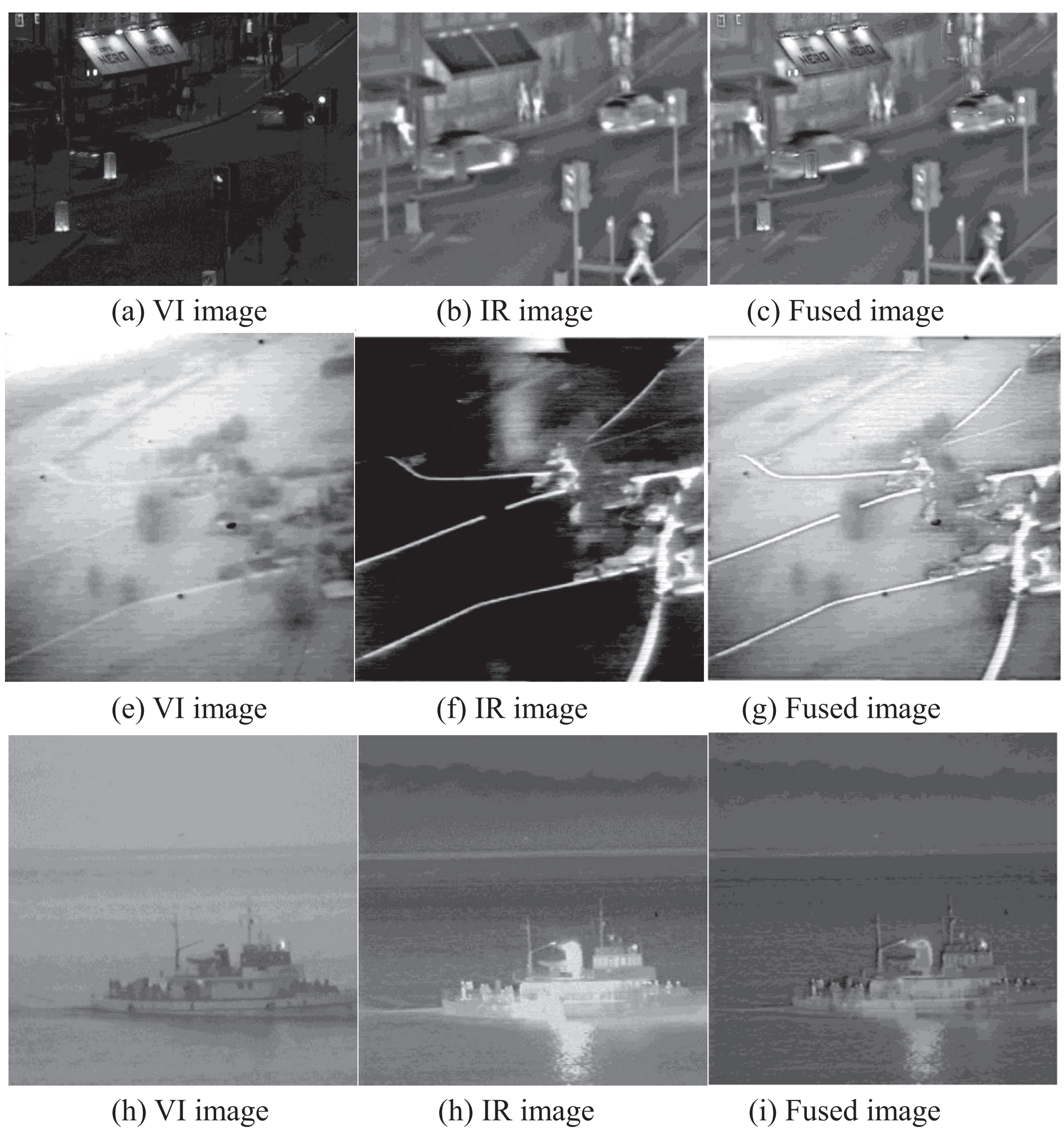


Fig. 5. Three examples of IR and VI images and the fused images.

Table 1
The statistics of IR and VI image fusion.

Families of fusion method	Image representation methods	Fusion strategies
Pyramid transforms	Contrast pyramid [2,106–108]	Optimized weighted coefficients by teaching learning based optimization [2], immune clonal selection algorithm [106], multi-objective evolutionary algorithm [107], Otsu method and morphology [108]
	Laplacian pyramid [13,77]	Fuzzy logic and region of interes [13]
	Steerable pyramid [109]	Laplacian pyramid-based procedure [109]
Common multi-scale transform (MSA)	Dual-tree DWT [105,134]	Fuzzy logic and population-based optimization [105], particle swarm optimized [134]
	DWT [3,8,44,26,28,31,35,53,101,103,135], and stationary wavelet transform [14]	Improved entropy weighted fusion rule and max-abs-based fusion rule [3], target regions [8,103], averaging [31], spatiotemporal energy-based [53], robust principal component analysis (RPCA) and regional variance estimation [135]
	Spectral graph wavelet transform [29]	Bilateral filter [29], saliency map [29]
	Curvelet transform [41,129]	Spatial-temporal energy [41], local variance weighted and fourth-order correlation coefficient match [129]
	Contourlet transform [40,104]	Golden section algorithm and local energy [40], mean gradient and PCA [104]
Nonsubsampled multi-scale and multi-direction geometrical transform	NSCT [7,84,122,123,124,125,130,131,132,133]	Fractal dimension phase and congruency [7], pixel information estimation [84], Gaussian fuzzy logic and PCNN [122], neighborhood energy and neighborhood characteristic regional process [123], regional visual characteristics and cross-gradient [124], adaptive regional average energy rule and maximum absolute selection rule [125], saliency regions detection [130], region segmentation [131], corresponding activity measures [132], maximum absolution selection [133]
	NSST [1,93,95,96,97,98]	Saliency analysis [1], fast non-negative matrix factorization [93], SF and PCNN [95], fast non-classical receptive field and local directional contrast [96], region average energy and local directional contrast [97], spiking cortical model [98]
Other multi-scale transform	Edge-preserving filters [5,144], Y-styled window filter-scale invariant feature transform [28], directional bilateral filter [30], morphological theory [127], anisotropic diffusion and karhunen-loeve transform [137], Gaussian and bilateral filters [141], weighted least squares filter [142], directional nonlocal means filter [143], rolling guidance filter and Gaussian filter [145], mean filter and median filter based [146], directionlets transform [150], BEMD and NSDFB [151]	Regional saliency map [5,146], correlation coefficient and contrast enlargement strategy [127], absolute maximum and the proposed Gaussian based weighting coefficients methods [141], saliency map [142,144], guided filtering [144], visual saliency map and weighted least square [145], mean and local variance maximum principle [150], fuzzy logic [151], regional energy and regional clarity [151], sparse representation based and activity level measurement [158]
Compressive sensing and sparse representation	Compressive sensing [3,42,87,88,89,125]	Improved entropy weighted fusion rule and a max-abs-based fusion rule [3], self-adaptive weighted average fusion scheme based on SD [87], maximum absolute of entry [89]
	Sparse representation [111,119,120,121,158]	Salient detect [111], PCNN [119,121], energy of edge and novel modified Laplacian [119], kernel density estimation clustering method and singular value decomposition [121], SR-based fusion approach and activity level measurement [158]
Other methods	Markov random fields [99,100,101]	Salient structures of the source images [99], visibility area selection [100], saliency detection [101]
	Optimization-based method [148]	Global entropy and gradient constrained regularization [148]
	Fuzzy theory [105,125,152]	Support vector transform [152]
	Morphological center operator [127]	Utilizing the contrast enlargement strategy [127]
	Local-window-based frequency-tuned method [147]	Saliency maps [147]
	Higher order singular value decomposition [102]	
	PCA [104], RPCA [133,135]	
	ICA [138]	
	Contour detection [149]	
	Total variation [153,154]	

information from near infrared images; another contribution was to extract a single map containing information which was fused from all the near infrared images [49].

In other fields, we also find some applications of IR and VI image fusion. There are a few examples, such as visible-infrared fusion schemes for road obstacle classification [15], and advanced assessment in nondestructive testing scheme [16]. Besides, an eyeglasses removal method for thermal image based on the fusion of visible and thermal image was proposed by Wong et al. [17]. In the field of medicine, real-time fusion of IR and visible range video images using a tunable pattern selective image fusion (PSF) approach was applied in intraoperative assessment of critical biliary structures [36]. IR and VI image fusion also was used in construction technology [50].

In addition to traditional applications, many new demands are raised in response to proper time and conditions as the mentioned in this section. The new demands will benefit from the advanced image representation methods and fusion strategies for different applications. We think that these new demands and applications would push to the development of IR and VI image fusion technologies. Of course, the challenges would also appear constantly.

3. Infrared and visual image fusion methods

In general, the detailed feature information of IR and VI images would be described by some image representation methods which provide the possibilities to fuse more information of different IR and VI images. The performance of image representation method

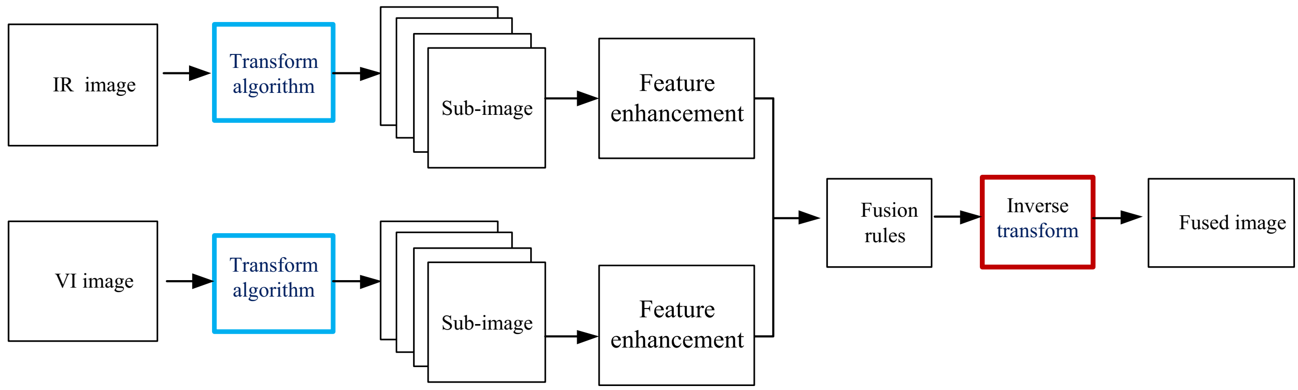


Fig. 6. A General transform domain image fusion flow diagram.

determines the quality of fused image at some level. Certainly, the subsequent processing and fusion rule will be other key factors for the quality of fused image. Three examples of IR and VI images, and the fused images are shown in Fig. 5. In this section, we will show the detailed survey of IR and VI image fusion methods. And Table 1 shows the summary of major IR and VI image fusion methods including, Families of fusion method, Image representation methods, and Fusion strategies.

3.1. Multi-scale analysis

In recent decades, multi-scale analysis (MSA) theories are well developed, which make image analysis tools be more and more effective; therefore, MSA-based image fusion methods become diverse as well [91]. To this day, MSA-based image fusion methods have been widely studied, so we can infer that lots of advanced IR and VI image fusion techniques would be proposed in succession [45,92,93]. The first step of MSA-based techniques is image transform which is used to get the different sub-band coefficients of source images; and then the recombined coefficients of the transformed image would be obtained according to different fusion strategies; at last, the fused image can be obtained by inverse transform. Generally, these techniques could be primarily classified into several mainstreamed categories according to the different decomposition or transformation ways. A general transform domain IR and VI image fusion scheme is shown in Fig. 6.

3.1.1. Pyramid transforms

In IR and VI image fusion field, various pyramid algorithms were adopted to decompose source images, such as steerable pyramid [109], Laplacian pyramid [77,130] and contrast pyramid [106,107,140]. This kind of method usually uses pyramid algorithms (PT) for image decomposition to extract its detailed information, which could help to improve the fused image quality.

Steerable pyramid was a multi-scale, multi-orientation and self-inverting image decomposition method which was applied in the research of Liu et al., the image was first divided into collection of sub-bands localized in scale and orientation, a robust recursive fusion strategy was designed by applying a Laplacian pyramid to combine the coefficients [109]. In Yu's research, Laplacian pyramid was used to decompose the source images for false color image fusion [77]. Besides, two layers Laplacian pyramid decomposition image fusion method also was used by Bulanon et al. for fruit detection technique [13].

There were three papers published by Jin et al., which about the application of contrast pyramid (CP) in IR and VI image fusion field. In the early research of Jin et al., CP was first used for the source images to get corresponding sub-band coefficients; then directional filter banks were constructed for filtering each sub-band image. In this scheme, the immune clonal selection algorithm was employed to optimize the fusion coefficients for better fused image [106]. In 2014, Jin et al. proposed an image fusion method through the combination of CP and teaching learning based optimization (TLBO) for IR and VI image fusion. The paper's innovation came in the weighted coefficients which could be automatically adjusted by the fitness function of TLBO according to the fused image quality [2]. In another paper, it first used CP to decompose the source images into sub-band coefficients; then multi-objective evolutionary algorithm based method was introduced to optimize fusion coefficients, thus the weighted coefficients could be adjusted automatically according to fitness function [107].

PT based IR and VI image fusion methods have the advantages of lower computational complexity and less memory space; however, the disadvantage of PT is that it would smooth some image details [142]. And PT fails to introduce spatial information into the decomposition course, and always leads to blocking phenomena [93]. As a result, PT alone cannot achieve good IR and VI image fusion quality, generically. Due to the limitation of CP, these kinds of methods are often used together with other methods to overcome their shortcomings, such as bionic optimization algorithms and other image representation methods [2,106,107,109]. When bionic optimization algorithms were applied to improve the quality of the fused image of PT, computational complexity would be also increased. Need to point that the transform domain based IR and VI color image fusion technique was a new idea to improve the quality of fused color images, such as Laplacian pyramid algorithm.

3.1.2. Common multi-scale analysis (transform)

Multi-scale analysis (MSA) approaches are very popular in IR and VI image fusion methods; especially wavelet transform (WT), curvelet transform, contourlet transform. Most researches are based on this kind of method, and it has become a very important research area for IR and VI image fusion due to its advantages [142]. MSA image fusion schemes would decompose IR and VI images into several components according to different scales; each of these components could capture the presented information of the given scale. Gemma Piella thought that there were many reasons for using MSA in image fusion method. Specifically, (i) the image of real-world object is usually composed of various structures information at different scales [45]; (ii) there is strong evidence that the human visual system processes information in a multiresolution fashion [45]; (iii) multiresolution methods offer computational advantages, moreover, appear to be robust [45].

WT is the most representative method among the numerous MSA methods [46]. WT utilizes the ideas of localization of short time Fourier transform (STFT) that can provide us with the frequency of images and the space associated to those frequencies, which makes it become a good frequency analysis tool. WT can automatically adapt to the requirement of time-frequency signal analysis, which makes it be able to focus on the arbitrary signal details and overcome the shortcomings of Fourier transform. After, many wavelet-based methods were developed (e.g. lifting wavelet and spectral graph wavelet, etc.), and these methods have good direction and nonredundancy, but wavelets cannot efficiently represent and process multidimensional information [130]. Geometry-based methods (e.g. curvelet, contourlet, etc.) are efficient anisotropy analysis tools in high dimensions. Above mentioned methods are based on the structure information of the image itself, while do not take account into visual attention of human [130].

The discrete wavelet transform (DWT) technique is applied to image fusion of IR and VI image whose sparse image coefficients would be got by MSA fusion schemes [3,26,28,35,101]. Aran et al. employed wavelet-modified maximum average correlation height (Wave MACH) filter for automatic target recognition applications in the fused IR and VI images [31]. Niu et al. used DWT to research the fusion concerning target detection and perception [103]. Yan et al. proposed an IR and VI image fusion method based on bilateral filter and spectral graph wavelet transform (SGWT) which had the advantage of spectral characteristics that could decompose images in graph domain [29]. Multiwavelet decomposition was applied to image fusion in Han's research whose coefficients contained a low-frequency approximation component and multiple high-frequency detailed components which represented sharp edges and details at various scales [44]. A motion-compensated wavelet transform (MCWT) model was used to fuse video frames by Xu et al. [53]. In Saeedi's scheme, the source images were decomposed by dual-tree DWT (DT-DWT) to get different scales sub-band coefficients [105]. Madheswari et al. used particle swarm optimization to optimize weighting factor of the subband of DT-DWT by maximized entropy and minimized root mean square error [134]. Wang et al. also presented an improved image-fusion algorithm based on the lifting wavelet transform (LWT) [135].

Curvelet transform can decompose the image into different subbands at different scales and various directions, it was used to fuse IR and VI images in Shao's research, the low frequency subbands were fused by the local variance weighted strategy, and the high-frequency subbands were fused by the fourth-order correlation coefficient match strategy [129]. In Li's study, uniform discrete curvelet transform (UDCT) was employed to decompose the source videos, and a local spatial-temporal energy based fusion rules were employed for the coefficients [41]. As a kind of multi-scale representation tool, UDCT has lower redundancy ratio, a hierarchical data structure, shift-invariant and compute efficiency, which is desirable properties [41].

Different from curvelet transform which is first developed in continuous domain and then is discretized for sampled data, contourlet transform starts with a discrete-domain construction [136]. Contourlet transform was proposed based on WT, which was regarded as a "real" two-dimensional images representation method, also called pyramidal directional filter bank (PDFB) [104]. Compared with WT, it has a better sparse expression characteristic which is good at describing contours and directional texture information of the image [104]. Li et al. proposed a contourlet transform based IR and VI image fusion scheme whose low-frequency images were fused by PCA, and high-frequency coefficients were fused by mean gradient strategy. A pixel-level remote sensing image fusion method based on multi-contourlet transform was proposed by Chang et al., the method had good direction selectivity and energy convergence. In their research, golden section algorithm was adopted to fuse the low frequency coefficients; the local energy feature was adopted high-frequency directional coefficients [40].

Common MSA-based method for IR and VI image fusion could achieve good effective, but there still have some weaknesses [93], such as, the DWT and DTCWT cannot capture curves and edges of the images very well. DWT is only sensitive to the point-wise singularities, but it cannot capture other types of salient features, which often causes artifacts and Gibbs effects in the fused images. Similarly, curvelet transform also cannot effectively represent the detailed spatial information; as a result it does not have obvious property of spatial sampling. Because the distinct mechanism of contourlet transform made it lack of the shift-invariance property. Hence, the contourlet transform based image fusion results also easily appear to Gibbs phenomena. Thus, lots of improved algorithms are proposed based on these common MSA to increase fusion performance according to the characteristics of IR and VI image. But these still could not completely overcome the inherent imperfections.

3.1.3. Nonsubsampled multi-scale and multi-direction geometrical analysis (transform)

Non-subsampled contourlet transform (NSCT) is constructed as an effective image representation framework; it is based on common multi-scale geometrical transform theories [83]. NSCT is designed to overcome the pseudo-Gibbs phenomena around singularities, and it is a shift-invariant version of contourlet transform [83,84]. Compared with the common MSA techniques, NSCT is characterized with much better fusion performance in terms of both subjective and objective evaluations [93,7,123,130,131,132,133]. Zhao et al. designed a fast NSCT scheme for IR and VI images fusion [84]. He et al. developed a NSCT-based method by combining pulse coupled neural network and Gaussian fuzzy logic [122]. Bhatnagar et al. proposed a framework which applied NSCT to decompose IR and VI images into different frequency subband coefficients to effectively fuse images by different rules [7], and the similar schemes can be found in the research Chen et al. [123], Adu et al. [124], Li et al. [132] and Liu et al. [131]. Besides, there were some studies based on the combination of NSCT and other image representation methods; NSCT was combined with compressed sensing in the research of Zhang et al. [125], Chen et al. proposed an attention-based hierarchical fusion for IR and VI image based on NSCT and Gaussian pyramid [130], and Fu et al. integrated RPCA with NSCT to fuse IR and VI image [133].

Although NSCT has many advantages, the high computational complexity of NSCT limits its application, especially the real-time system [93]. The absence of shift-invariance in ST easily tends to appear Gibbs' phenomena. In order to overcome the above problems, the theory of non-subsampled shearlet transform (NSST) was proposed [94]. NSST was proposed with two advantages of good image representation performance and low computational cost; it not only has the advantages of common MSA tools, but provides a superior mathematical structure and flexibility [93]. Besides, the advantages of feature capturing and representation capabilities make NSST become a popular technique in IR and VI image fusion method. An IR and VI image fusion method based on saliency analysis and NSST was proposed by Zhang et al., NSST was used to select the fusion coefficients of the background, and saliency analysis method was used to extract the salient regions which would be fused in the final result [1]. This approach combined pixel based fusion method and region based fusion method. Besides, Kong et al. made lots of IR and VI image fusion researches with NSST. In [93], NSST was used to decompose the source image into to a low-frequency coefficients and a series of high-frequency coefficients, the former was fused by fast non-negative matrix factorization (FNMF), and the later were fused by the modifications FNMF. In [95], Kong et al. presented an IR and VI image fusion method based on NSST-SF-PCNN, NSST was used to decompose the source image, and spatial frequency was calculated as the linking strength of PCNN to determine coefficients of the source images. Afterwards, IR and VI image fusion scheme was designed by the combination of NSST with spiking cortical model (SCM), the method aimed to achieve good fusion performance and simultaneously reduce computational complexity [98]. In another two studies of Kong et al., the sub-image coefficients of IR and VI image were generated by NSST and were fused by different methods, one by the improved fast non-classical receptive field and the local directional contrast (LDC) [96], another one by region average energy (RAE) and LDC [97].

Non-subsampled transform tools have the advantages of multi-scale and multi-directions geometrical decomposition which are good at describing contours and directional texture information of the source images. As a result, these methods could effectively decompose the features of IR and VI image into several subbands; however, how to further extract the regional characteristics of thermal radiation in IR image is the key point to improve the performance of image fusion method. Therefore, most researchers focus on the techniques to accurately extract the regional features of IR image, and it also is the key research issue in non-subsampled transform based method for IR and VI image fusion. Generally speaking, non-subsampled transform based IR and VI image fusion methods have some limitations because of higher computational complexity and heavier memory; however, these methods usually could obtain better fusion quality. Therefore, non-subsampled transform tools become the most popular technologies in IR and VI image fusion in recent years, and they could achieve very good fusion effect; however, the high computational complexity limits the application in real-time applications in some scenarios.

3.1.4. Other multi-scale analysis methods

In addition to the above mentioned classics MSA tools, there are some specific MSA methods, and most the design principles are based on filter technologies which are applied to decompose the features of IR and VI images. The smoothing operations of filter could extract the detailed information of the image according to the difference between original image and filtered image, which provide feasibility for IR and VI image fusion.

Hu et al. proposed a MSA algorithm, named multiscale directional bilateral filter (MDBF) which was a multiscale, multidirectional and shift-invariant image decomposition scheme, and it was used to decompose the source image into directional detail subbands and approximation subbands [30]. In the research of Yan et al., multiscale directional nonlocal means filter (MDNLM) was designed as a multiscale, multidirectional, and shift-invariant image decomposition scheme to represent the geometric structure information of the source images [143]. Similarly, Zhou et al. proposed a hybrid multi-scale decomposition method by using multi-scale Gaussian and bilateral filters to get three scale levels: small-scale levels, the large-scale levels and the base level; then absolute maximum and the proposed Gaussian based weighting coefficients methods to fuse the images [141]. Y-styled window filter (YWF)-scale invariant feature transform (SIFT) based multi-scale image fusion scheme was used in face recognition field by Bai et al. [28].

Saliency testing techniques are often combined with filter techniques to fuse IR and VI image. In Cui's research, the image smoothing framework based on edge-preserving image smoothing method by using L_1 fidelity with L_0 gradient was adopted as MSA scheme for IR and VI images; the saliency map of each decomposition layer were extracted by frequency-tuned saliency map extraction algorithm; the final fused image was reconstructed by synthesizing different levels with proper weight [5]. In the study of Gan et al., another multi-scale edge-preserving filter decomposition scheme was designed to fuse IR and VI image fusion by combining with saliency maps [144]. The weighted least squares filter based multi-scale decomposition was designed to decompose IR and VI image, and the pixel value based saliency map was adopted for image fusion in different decomposition level [142]. Ma et al. designed a rolling guidance filter (RGF) and Gaussian filter based MSA method to decompose the IR and VI images into base and detail layers, then an improved visual saliency map based method was used to fuse the base layers, and a weighted least square optimization was used to fuse the detail layers [145]. In Bavirisetti's research, a median filter and mean filter based MSA scheme was designed to decompose the source image; a special weight map construction process based on visual saliency also was proposed for IR and VI image fusion. This method had the characteristic of low computational complexity [146]. In other research of Bavirisetti et al., anisotropic diffusion and Karhunen-Loeve transform based IR and VI image fusion scheme was proposed [137].

Except filter based multi-scale transform scheme, there are some other multi-scale transform schemes. In Bai's research, multi-scale morphological theory based MSA was proposed, it first adopted the morphological center or anti-center operator to extract the features of IR and VI images to get the correlation coefficients; then the multi-scale morphological theory was used to extract the multi-scale features through the correlation coefficients to fuse the features by contrast enlargement strategy [127]. A directionlets transform based IR and VI image fusion scheme was proposed by Zhou et al., which the low-frequency was fused by mean rule and high-frequency coefficients were fused by local variance maximum principle [150]. Zhu et al. proposed an IR and VI image fusion method based on bidimensional empirical mode decomposition (BEMD) and flexible directional expansion of non-subsampled directional filter banks (NSDFB) [151].

Most of these kinds of methods are based on the filter decomposition of the source images, and have achieved good results in IR and VI image fusion field. The obvious regional features of thermal infrared characteristics in IR images provide the possibility. In general, the methods could decompose these image features by different filters according to the characteristic of IR and VI images; then the decomposed sub-images would be fused by successive processing and fusion rules; finally, the fused result is reconstructed by synthesizing different scales with synthetic weights of filter based MSA. These technologies are very popular in IR and VI image fusion in recent years, and the advantages of these technologies are easy to implement and calculate. Unfortunately, the technologies have a common problem that the fixed filter parameters would cause its poor flexibility. More unfortunately, the quality of the fused image seriously depends on the effect of filter. Given the limitations of different filters, it cannot guarantee the effectiveness of the algorithm for all IR and VI images, such as, thermal infrared characteristics of IR image is not obvious, and the effect of noise. However, we have sufficient reasons to believe that such technologies will be better developed through the efforts of researchers.

3.2. Compressive sensing and sparse representation

In this section, we will introduce compressive sensing (CS) and sparse representation (SR) based IR and VI image fusion methods. And these two technologies are becoming hot research field in recent years.

The advantages of CS and SR based IR and VI image fusion methods are that they can reduce the computational complexity and improve processing speed to a certain extent. And what needs to be pointed out is that the performances of this kind of IR and VI image fusion methods rely largely on the compressive sampling matched pursuit algorithm and overcomplete dictionary construction. At present, many implemented methods are being put forward by researchers for IR and VI image fusion; however, the valid image compression and sparse coefficient representation, the accurate image recovery techniques are still needed further study. IR and VI images contain abundant details by two-dimensional data matrix, but the information loss of these methods has some restrictions in their applications. The dependence for compressive sampling matched pursuit algorithm and overcomplete dictionary construction is another limitation as well. As a result, CS and SR are often combined with other image representation methods to overcome the problems, such as MSA, PCNN. And the relevant researches in this field would be introduced in the following two sections.

3.2.1. Compressive sensing

Compressed sensing theory, which has been received widespread attentions after it was proposed; it can effectively reduce the computational complexity and increase the operating rate of image processing schemes due to the low signal sampling and compression method, which is based on the sparsity of signal under a certain transformation [3]. The theories of CS prove that a sparse or compressible signal can be exactly reconstructed from a small number of nonadaptive linear projections, which are far fewer than the number of samples if the signal were sampled at the Nyquist rate [86,87]. When it is applied to the image fusion field, only a part of sparse coefficients are needed to be fused and simultaneously enhance the quality of the fused image [3].

In recent years, many scholars do lots of researches in CS based image fusion technologies. In Liu's research, DWT was adopted to get the sparse coefficients of IR and VI images; low frequency coefficients were fused by an improved entropy weighted fusion rule; high frequency coefficients were fused by a max-absolute-based fusion rule; the fused image was reconstructed by a compressive sampling matched pursuit algorithm after local linear projection using a random Gaussian matrix [3]. A simultaneous images super-resolution and fusion method based on compressed sensing and multiscale dictionaries learning scheme was designed by Ren et al., the source images were reduced to a task of signal recovery from the compressive measurements based on the sparse prior of image patches and the framework of compressed sensing, and multiscale dictionaries learning technology was introduced to fuse the high-resolution image at each scale [42]. Another CS principle based IR and VI image fusion scheme was studied by Li et al.; the recovery tool of total variation optimization was used in the special domain of CS, a self-adaptive weighted average fusion scheme based on standard deviation of measurements was developed to merge IR and VI images [87]. In Jameel's IR and VI image fusion method, the numbers of compressive measurements were adaptively adjusted by the amount of information [88]. Ding et al. proposed an IR and VI image fusion method via CS, the paper selected three trained overcomplete dictionaries by K-means singular value decomposition (K-SVD), and the dictionary of the patches respectively came from three sources, such as IR images, VI images and the combined patches; besides, two sparse vectors approximations containing orthogonal matching pursuit and polytope faces pursuit algorithms and two fusion rules were proposed [89]. A hybrid scheme of adaptive-Gaussian (AG) fuzzy membership method, CS technique and total variation (TV) based gradient descent reconstruction algorithm, was proposed to solve the problem of IR and VI image fusion by Zhang et al. [125].

3.2.2. Sparse representation

Sparse representation (SR) also received wide attentions in signal processing field in the past decades, which was a very interesting research field in IR and VI image fusion as well. The purpose of SR is to use as few information components as possible to represent the source signal based on the overcomplete dictionary. As a result, the quality of the overcomplete dictionary determines the signal representation ability of sparse coding [158]. Generally speaking, there are two methods of offline approaches to obtain a dictionary; one is to use the analytical models such as discrete cosine transform (DCT) and curvelet transform (CT); the other one is to apply the machine learning technique to obtain the dictionary from a large number of training image patches [110,158]. In the field of image processing, the image signals can be represented as a linear combination of a "few" atoms from a overcomplete dictionary, and the sparse coefficients are regarded as the salient features of source images [111]. Due to the inherent advantages of SR, it is widely used in IR and VI image fusion field. In Wang's research, non-negative sparse representation method was used to decompose the IR and VI images into pure additive and sparse coefficients, which can effectively extract the features of source images; the fusion guide vector was obtained by regional consistency rule to fuse them [111]. Target separation and SR based IR and VI image fusion scheme was proposed by Lu, the infrared target was detected and separated from the background to fuse them into the final image, and the rest of background image was fused by SR based method [121]. In addition to the above-mentioned method, SR based image fusion scheme also can be found in [119,120,158].

3.3. Other infrared and visual image fusion methods

Markov random fields (MRF) were also used in the field of IR and VI image fusion by some researchers. In study of Sun et al., a gradient domain image fusion scheme by the combination of Poisson fusion framework with MRF based structural fusion model was proposed [99]. In Shibata's research, the visible areas were selected by minimizing the MRF energy evaluated, which based on both local visibility and measured inconsistencies [100]. MRFs-based saliency detections method also was designed to fuse IR and VI image by Han et al. [101].

In the research of Yang et al., support value transform (SVT) was used to decompose the IR and VI image into low-frequency and high frequency components which were combined by fuzzy logic combination rule [152]. Fuzzy theory also can be found in other research of IR and VI image fusion [151], and an adaptive-Gaussian fuzzy membership scheme was used in Zhang' s research [125]; Saeedi et al. also adopted fuzzy based method to fuse high frequency DT-DWT in IR and VI image fusion [105].

In recent years, robust principal component analysis (RPCA) got many attentions. In Fu's research, IR and VI images were decomposed by RPCA to get the corresponding sparse feature, meanwhile the source images were also decomposed into low frequency coefficients and high-frequency coefficients by NSCT; the sparse matrixes of RPCA were used to guide the fusion rule of low frequency coefficients and high frequency coefficients [133]. Lifting wavelet transform and RPCA based IR and VI image fusion method also was researched by Wang et al. [135].

Artificial intelligence information processing methods have a very good application prospect [90], one of the representative method is pulse coupled neural networks (PCNN) based image processing applications. PCNN's similarity group neurons will issue synchronous pulses under the effect of mutual coupling pulses; it shows superior performance in image processing field [91,92,121]. In the field of IR and VI image fusion, PCNN and spiking cortical model (SCM) were used in Kong's researches by integrating with NSST [95,98]. Yin et al. proposed a method by the combination of shift-invariant dual-tree complex shearlet transform, SR and adaptive dual-channel PCNN [119].

Last but not least important studies were made by some researchers in IR and VI image fusion field. Ma et al. focused on the application of total variation in IR and VI image processing [153], and another IR and VI image fusion method was proposed based on gradient transfer and total variation minimization [154]. In independent component analysis (ICA) domain a multimodal image fusion algorithm was designed by Cvejic et al., and it adopted segmentation algorithm to extract the critical regions of IR and VI images, which were used to fuse the ICA coefficients [138]. Zhao et al. proposed a multi-window visual saliency extraction method for IR and VI image fusion based on the idea of local-window-based frequency-tuned method [147], and Zhao et al. designed an optimization-based scheme by the combination of global entropy and gradient constrained regularization [148]. In addition to these, three object's contour detection based IR and VI image fusion schemes were proposed by Zhao et al. [149], and a higher order singular value decomposition of tensors for images fusion scheme was proposed by Thomason et al. [102].

These kinds of IR and VI image fusion methods are not very popular, but the methods could provide many new perspectives to complete this task. MRF is often applied as saliency detection operation to extract the areas with obviously thermal radiation in IR image; besides, saliency extraction is widely used in IR and VI image fusion, as a result, window (region) based techniques also could be employed to extract the thermal-radiating features of IR images. However, if thermal-radiating of IR image is weak and only have a small difference with VI image, the saliency extraction would become difficult. Fuzzy theory could be used to measure the detailed information of RI and VI images by membership; analogously, PCNN also could be applied to extract this information according to its specific mechanism. Besides, RPCA, ICA, total variation and higher order singular value decomposition are often used to decompose the features of IR and VI image into different coefficients for easily fuse them, which are similar to other MSA based methods. Most of these tools are lightweight and computable methods, although their detail capture capability may be not excellent. We think that some of these methods have good innovative ideas which would provide new thinking for IR and VI image fusion.

4. Assessment of fused image

In order to evaluate the quality of the fused IR and VI image, image fusion assessments metrics are widely adopted by most researchers [55], which can be divided into two groups: subjective evaluations and objective evaluation metrics. The purposes of image quality assessment methods are generally used to measure the contribution of the source images in fused image, and also could be applied to find the optimal setting of parameters for a specific fusion algorithm [45]. Besides, the fused image evaluation methods could also be adopted to evaluate the effectiveness of image fusion approaches [45,157].

4.1. Subjective evaluation methods

Subjective evaluation methods are used to evaluate the quality of the fused image according to human vision perspective, which is the most popular and direct image quality assessment methods. This is because the ultimate user or interpreter of the fused image is human, which cause subjective evaluation methods to be paramount importance for IR and VI image fusion [45,66,157]. The subjective evaluation can be further divided into absolute evaluation and relative evaluation, which implemented by widely-recognized five-level quality scale and obstacle scale. The key criterions probably include apparent image sharpness, edge definition, image distortion, the degree to maintain the source images detail information and so on.

4.2. Objective evaluation metrics

Mathematical calculation is a kind of popular IR and VI image fusion quality evaluation method; it is defined by some computable mathematical formulas. Different metrics (evaluation methods) are used to represent the fused image quality in different aspects, therefore many different evaluation methods would be simultaneously adopted in IR and VI image fusion field. This kind of evaluation methods is usually regarded as the objective standards for evaluating the fused image. The frequently-used evaluation metrics are shown in Table 2. And we will introduce some of them in this section.

1. Information entropy (IE)

$$H = -\sum_{i=0}^{L-1} P_i \log_2 P_i, \tag{1}$$

where P_i is the probability of the gray level i in the image, L is the gray level of the image from 0 to 255.

Table 2
The statistics of evaluation metrics and references.

Evaluation metrics	References
Information entropy (IE)	[2,5,13,29,39,40,42,44,97,98,95,96,103,104,106,107,108,121,123,124,126,130,131,133,134,141,142,145,147,148,150,151,153,154,158]
Average gradient (AG)	[2,5,29,39,42,44,7,96,97,106,107,39,108,119,135,150,151]
Standard deviation (STD)	[2,13,14,40,42,29,88,96,97,106,107,108,119,129,130,134,150,151,152,158]
Spatial frequency (SF)	[29,95,98,129,104,107,108,122,127,134,148,151]
Mutual information (MI)	[1,8,29,30,40,42,44,84,7,87,88,89,95,96,98,99,101,103,104,111,119,121,122,123,129,131,132,135,144,151,153,154]
Edge based on similarity measure (Q^{ABF})	[1,30,39,44,84,7,89,99,103,111,119,121,122,123,132,133,144,146,152]
Q_W	[1,30,89,111,126,158]
Visual information fidelity (VIF)	[30,84,88,122,132,146]
Structural similarity index measure (SSIM)	[8,84,7,96,97,102,121,133]
Peak signal to noise ratio (PNSR)	[10,7,104,135,133]
Q_E	[1,30,40,44,84,111]
Mean gradient (MG)	[104,127]
Q_0	[1,30,111,146]
Run time	[2,30,55,101,106]
FMI	[153,154]
Q	[84]
CC	[135]
UIQI	[42]

According to Shannon's principle of information theory, IE could reflect the average information and represent the texture richness in the fused image. The larger the IE is, the more plentiful the information amount of the fused image is. And IE is one of the most commonly used as image quality evaluation indicator. However, the value of IE would seriously increase if there were artifacts and noises in the fused image, which cannot represent the quality of the final image. Especially there will be lots of noises in IR image. As a result, we think IE only can be used as an auxiliary evaluation metric in IR and VI image fusion field.

2. Average gradient (AG)

$$AG = \frac{1}{M \times N}\sum_{i=1}^{M}\sum_{j=1}^{N}\sqrt{\frac{1}{2}((F(i,j) - F(i+1,j))^2 + (F(i,j) - F(i,j+1))^2)}, \tag{2}$$

where M is the row of the image, and N is the column of the image, $F(i, j)$ is gray level of the image F at pixel (i, j).

The AG can represent the presentation abilities of the fused image for details and textures, which can be used to assess the image sharpness. Generally speaking, the larger the average gradient value is, the richer information the fused image contains, the better the fused result is [5]. AG is an often-used evaluation metric in IR and VI image fusion, and it can evaluate the fusion effect only by the fused IR and VI image, which is independent of the standard reference image.

3. Standard deviation (STD)

The standard deviation (STD) of an image is defined by (3).

$$STD = \sqrt{\frac{1}{M \times N}\sum\nolimits_{i=1}^{M}\sum\nolimits_{j=1}^{N}(F(i,j) - \mu)^2}. \tag{3}$$

where $F(i,j)$ is the pixel value of the fused image at the position (i,j) and μ is the mean value of the image. And the mean value (MV) of the image is defined by (4).

$$\mu = \frac{1}{M \times N}\sum_{i=1}^{M}\sum_{j=1}^{N}F(i,j). \tag{4}$$

where $F(i,j)$ is the pixel value of the fused image at the position (i,j).

MV describes the brightness of the fused IR and VI image, and it also is a fundamental evaluation metric. STD represents the statistical distribution and contrast characteristics of the fused image. The larger the STD is, the more dispersed distribution of gray level in fused image and the greater the contrast is, the better the visualization of the fused image is [104]. STD also is a common evaluation metric in IR and VI image fusion, as shown in Table 2, and it is used by lots of scholars to evaluate the fused images.

4. Spatial frequency (SF)

SF represents the overall activity of image in spatial domain; it can be divided into spatial row frequency (RF) and spatial column frequency (CF), whose mathematical expressions are defined as follows.

$$SF = \sqrt{RF^2 + CF^2}, \tag{5}$$

$$RF = \sqrt{\frac{1}{M \times N} \quad \sum_{i=1}^{M}\sum_{j=2}^{N}[F(i,j) - F(i,j-1)]^2}, \tag{6}$$

$$CF = \sqrt{\frac{1}{M \times N} \quad \sum_{i=1}^{N}\sum_{j=2}^{M}[F(i,j) - F(i,j-1)]^2}. \tag{7}$$

where M is the row of the image, and N is the column of the image, $F(i, j)$ is gray level of the image F at pixel (i, j).

SF represents the clarity of the details and the space change of image. The bigger the SF is, the richer the textures and edges are. And it also is independent of the reference image. The unwanted artifacts in the fused IR and VI image will multiply the value of SF. In this case, the SF cannot correctly reflect the quality of fused image.

5. Mutual information (MI)

The MI is defined by (8), which is between the source image A-B and the fused image F [161].

$$MI = \sum_{i=0}^{L-1}\sum_{j=0}^{L-1}\sum_{k=0}^{L-1}P_{ABF}(i,j,k)\log_2\frac{P_{ABF}(i,j,k)}{P_{AB}(i,j)P_F(k)}, \tag{8}$$

where $P_{ABF}(i,j,k)$ is the normalization union gray level histogram of image A, B and F, $P_{AB}(i,j)$ is the normalization union gray level histogram of image A and B, $P_F(k)$ is the normalization gray level histogram of F, L is the gray level of the image, (i,j,k) represent the pixel value of the images A, B and F, respectively.

MI describes the amount of information of the source images, which is fused in the final image. The greater MI is, the more information is fused from the source image to the final image, which represents the effectiveness of the image fusion method [91,104]. MI reflects the statistical dependence of two random variables from information theory viewpoint; it measures the similarity of image intensity distribution of the corresponding image pair [136]. MI is one of the most useful evaluation metric in image fusion field, and it is a very important and frequently-used in IR and VI image fusion, as shown in Table 2.

6. Edge based on similarity measure (Q^{ABF})

The Q^{ABF} metric evaluates the sum of edge information preservation values and is defined by (9) [162].

$$Q^{ABF} = \left(\sum_{i=1}^{M}\sum_{j=1}^{N}(Q^{AF}(i,j) \times \omega_A(i,j) + Q^{BF}(i,j) \times \omega_B(i,j))\right) \times \left(\sum_{i=1}^{M}\sum_{j=1}^{N}(\omega_A(i,j) + \omega_B(i,j))\right)^{-1}, \tag{9}$$

where $Q^{AF}(i,j) = Q_\beta^{AF}(i,j)Q_\alpha^{AF}(i,j)$, $Q_\beta^{AF}(i,j)$, $Q_\beta^{AF}(i,j)$ and $Q_\alpha^{AF}(i,j)$ are the edge strength and orientation preservation values, respectively; $Q^{BF}(i,j)$ is similar to $Q_\beta^{AF}(i,j)$ and $\omega_A(i,j)$ and $\omega_B(i,j)$ are weights to measure the importance of $Q^{AF}(i,j)$ and $Q^{BF}(i,j)$, respectively. The dynamic range Q^{ABF} is [0, 1], and it should be close to 1 as far as possible for the best fusion result, as $Q^{ABF} = 1$. In addition, (i,j) represents the pixel location, M and N are the size of images, respectively.

Q^{ABF} is a measurement of how much the edge information is fused from the source images to the final image. The larger value of Q^{ABF} is, the more edge information be fused into the final image and the better fusion result is achieved [39,93]. The value of Q^{ABF} ranges from 0 to 1, which is closer to 0, the more edge information is lost; to the contrary, which is closer to 1, the more information is kept [162]. As MI, Q^{ABF} also is a widely recognized evaluation metric in IR and VI image fusion, as shown in Table 2. Need to point that the calculation of Q^{ABF} needs source images and the fused image.

7. Structural similarity index measure (SSIM)

Mathematically, SSIM between two variables U and V defined as (10), and the code can be found [165].

$$SSIM(U,V) = \frac{\sigma_{UV}}{\sigma_U\sigma_V}\frac{2\mu_U\mu_V}{\mu_U^2+\mu_V^2}\frac{2\sigma_U\sigma_V}{\sigma_U^2+\sigma_V^2}. \tag{10}$$

where μ_U, μ_V are mean intensities and σ_U, σ_V, σ_{UV} are the variances and covariance, respectively.SSIM is designed by modeling any image distortion as the combination of loss correlation, radiometric and contrast distortion, which combines the luminance distortion, contrast distortion and structure distortion between the fused image and the source images [84,166].

8. Q_W

$$Q_W = \sum_{w\in W} c(w)(\lambda(w)Q_0(A,F|w) + (1-\lambda(W))Q_0(B,F|w)). \tag{11}$$

where $Q_0(A,F|w)$ and $Q_0(B,F|w)$ are calculated by using the method in [164] by a local sliding window w. The saliency weight $\lambda(w)$ calculated by (12).

$$\lambda(w) = \frac{s(A|w)}{s(A|w)+s(B|w)}. \tag{12}$$

where the salience measures $s(A|w)$ and $s(B|w)$ are calculated with the variance of A and B in window w, respectively. The $c(w)$ is the normalized salience of w among all the local windows. The $c(w)$ obtained by (13).

$$c(w) = \frac{\max(s(A|w), s(B|w))}{\sum_{w'\in W}\max(s(A|w'), s(B|w'))}. \tag{13}$$

The loss of correlation, luminance distortion, and contrast distortion are combined by the metric Q_0 to evaluate the degree of distortion of the fused image, and the metric Q_W further takes the salience of information into account to reflect the quality of the fused image [136].

9. Visual information fidelity (VIF)

The VIF quantifies the distortions of the images including additive noises, blurs, and global or local contrast changes [167]. And literature [84] assesses image quality by modeling human visual system, natural scenes and image distortion [163].

10. Peak signal to noise ratio (PNSR)

Peak signal to noise ratio (PNSR) represents the ratio between the maximum possible power of a signal and the power of distorting noise that affects the quality of its effectiveness, which is used to measure the proximity of the source image and the final image [7,104,135]. Because the lack of the standard reference image in IR and VI image fusion, the calculation of PNSR can only use the source images.

11. Other image evaluation methods

Run time is used to test the time complexity of image fusion algorithm; it is an important index for algorithm performance evaluation.

On the basis of MI, a non-reference objective image fusion metric was proposed by Haghighata et al., named FMI, which represents the amount of information fused from the source images to the final image [153,154,155]. Besides, the mean gradient (MG) represents the changes of the details and textures of the fused image. The greater the MG is, the clearer the details are, and the better qualities of fused results are [104,127].

Correlation coefficient (CC) is used to represent the correlation (degree of linear coherence) between the source and the final images [42,135]. The metric Q is used to evaluate the overall distortion of the loss of correlation, luminance distortion and contrast distortion [84,159]. In addition, there are some less commonly used image quality assessment methods, such as Q_E [160], UIQI [42], Q_0 [30,146].

Human is the ultimate user or interpreter of the fused IR and VI image, and human quality measure seems to be very reasonable. However, the shortcomings of subjective evaluation method are obvious, such as poor consistency, time consuming, difficult to reproduce and verify, etc [45]. Objective evaluation metrics mainly rely on mathematical calculation to describe the amount of the information in the fused image from the source IR and VI images, and this is a kind of reliable indicators, because the performance of information preservation can be effectively evaluated. The lack of standard reference image in IR and VI image fusion task cause its quality evaluation to be very difficult because it only can rely on the input source images. The quality evaluation task of the fused IR and VI image is still an open problem [96]. Due to the development of mathematical theories, more image fusion assessment methods would be proposed; therefore, we think that it would cause the assessments for IR and VI image fusion to be more and more reasonable and effective. Because there may be some artifacts and noises in the fused IR and VI image, which will lead to the value of some evaluation metrics incorrectly increase; however, it is not the promotion of final image quality. It is obvious that one evaluation metric alone cannot effectively represent the quality of final result; as a result, different objective evaluation metrics would be used together to measure how much the information is fused from the source images to the final image.

5. Experiments for typical methods

In this section, we implemented some experiments of typical fusion methods for several groups of often-used IR and VI images, as shown in Figs. 7–9(a) and (b), which came from different scenes with various details. These images have different thermal-radiating feature and distribution, which could test the information extraction ability of IR and VI image fusion method. These typical image fusion methods using in this paper were: principal component analysis (PCA), Laplacian pyramid (LP); contrast pyramid (CP); wavelet transform (WT); nonsubsampled contourlet transform (NSCT); nonsubsampled shearlet transform (NSST); mean filter and median filter (MF-MF). All the subband coefficients of above mentioned methods are fused by the simplest way of absolute value maximum.

In order to evaluate the performance of different image fusion methods, we adopted the most frequently-used image fusion evaluation metrics for the experiments, which were mutual information (MI), entropy (EN), average gradient (AG), standard deviation (STD), space frequency (SF), mean value (MV) and edge based on similarity measure (Q^{ABF}).

The first group of source images and the fused images with different methods are shown in Fig. 7. PCA only could complete this task, but the performance is not very well. The fused images from NSST, NSCT and MF-MF, are brighter than others, and the detail information are more than others as well. Table 3 shows the objective evaluation metrics with different methods in this pair of IR and VI image. Overall, the fused images from NSST and NSCT are better than others.

The second group of source images and the fused images with different methods are shown in Fig. 8. CP and PCA could complete

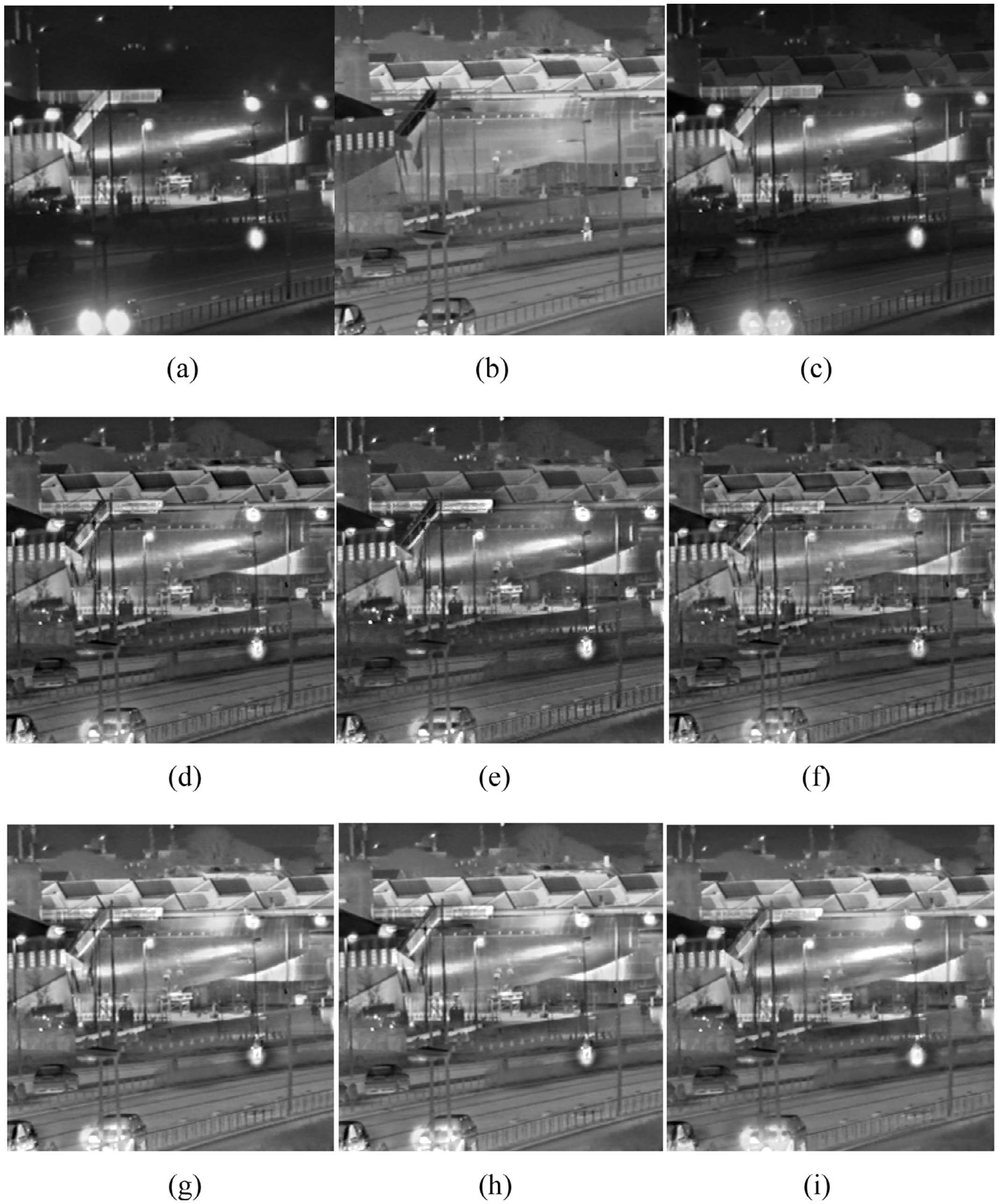


Fig. 7. Source images S1 and the fused images with different methods. (a) VI. (b) IR. (c) PCA. (d) LP. (e) CP. (f) WT. (g) NSST. (h) NSCT. (i) MF-MF.

this task, but the performances are not very well. The region with obviously thermal radiation would be fused into the final images by NSST and NSCT, which followed by MF-MF. And the detail information of NSST, NSCT and MF-MF are also better than others. Table 4 shows the objective evaluation metrics with different methods in this pair of IR and VI image. The MV of NSST, NSCT and MF-MF is larger than others. However, the most evaluation metrics of the methods are similar to each other. This is because the difference of IR and VI image is very small, and only the infrared region of a person is obvious in the IR image.

The third group of source images and the fused images with different methods are shown in Fig. 9. The performance of PCA is very poor. The ground detail of LP and CP is clear than others, but the celestial detail of NSST, NSCT and MF-MF is better than others. Most of these methods could fuse the lightspot of the house into the final image. Table 5 shows the objective evaluation metrics with different methods in this pair of IR and VI image. The MI, SF and STD of PCA are much larger than others; however, its visual effect is the worst. Other indicators have their merits.

In general, the brightness of NSST, NSCT and MF-MF is better than others; and the visible detailed information of VI image and infrared target areas of IR image could be fused into the final result. Although the fusion approaches based on the complex MSA tools could achieve better image fusion quality by combining all of the coefficients, and they usually suffer high computational complexity. On the other hand, the simple IR and VI image fusion approaches may not achieve very good performance, but they have low computational complexity. The balance of computational complexity and fusion effect would be important for IR and VI image fusion; therefore further researches are required for various practical applications and demands. But the evaluation metrics are very dynamic, and it proves that some reasons make the value of some evaluation metrics incorrectly increase; however, it is not the promotion of final image quality. There are many IR and VI image

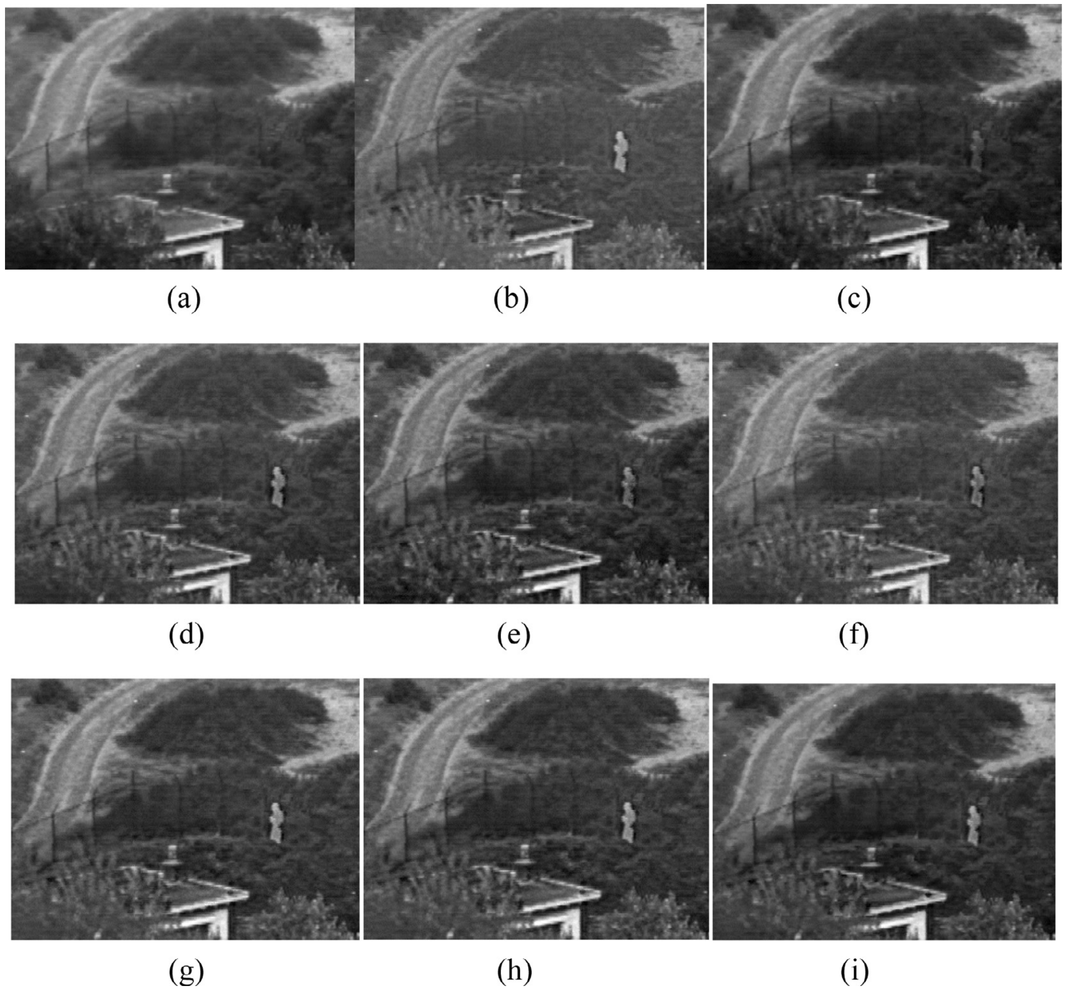


Fig. 8. Source images S2 and the fused images with different methods. (a) VI. (b) IR. (c) PCA. (d) LP. (e) CP. (f) WT. (g) NSST. (h) NSCT. (i) MF-MF.

fusion evaluation methods, however, none of them is universally believed to be always more reasonable and effective than others in IR and VI image fusion field. The experiments reveal that the quality evaluation of IR and VI image fusion is a difficult and challenging problem due to its special properties. And most of the evaluation metrics only could reflect the fusion quality to some extent. Therefore, more effective and comprehensive assessments for IR and VI image fusion should be researched.

6. Future trends and limitations

IR and VI image fusion play an important role in various applications, and it gets a lot of attentions in recent years. The corresponding technologies have developed very well. However, the rapid expansion of computer imaging technologies and sensor technologies would provide many new possibilities to improve IR and VI image fusion performance, and the emerging demands of various scenarios would offer more impetus, simultaneously. As a result, there are still many problems which are not fully studied or solved, and future improvements and researches are needed as well.

6.1. The combination of different image representation methods

In recent years, it is an obvious tendency that the integration of two or more image representation methods for future researches because the development of the image representation theories and the technologies for fusion strategies, and we could call these kinds of categories as hybrid techniques, such as the integration of NSST and NMF [93], the combination of NSST and PCNN [95] or (ICM [98]), SR and SVD [121], contourlet transform and PCA [104], etc. Different image representation methods have their special advantages as well as some shortcomings. By the combining of different image representation methods, the shortcomings of single image fusion method will be retained; simultaneously, the advantages of different methods will be integrated together to get better fusion products. Various experiments and researches have been shown that the hybrid IR and VI image fusion techniques could perform better than the individual image representation methods [93,121]. The combination of different image representation methods will provide more possibilities to further improve the effect of IR and VI image fusion.

6.2. Learning based image fusion methods

In general, the lack of IR and VI image fusion samples makes learning based method be not easily applied in this field. In recent years, the theories and applications of deep learning got a lot of attentions and had a very good development. But beyond that, data learning and model training methods have been greatly improved as well. Two convolutional neural networks (CNN) based remote sensing image fusion methods were proposed by Li et al. [168] and Zhang et al. [169], respectively; besides, a multi-focus image

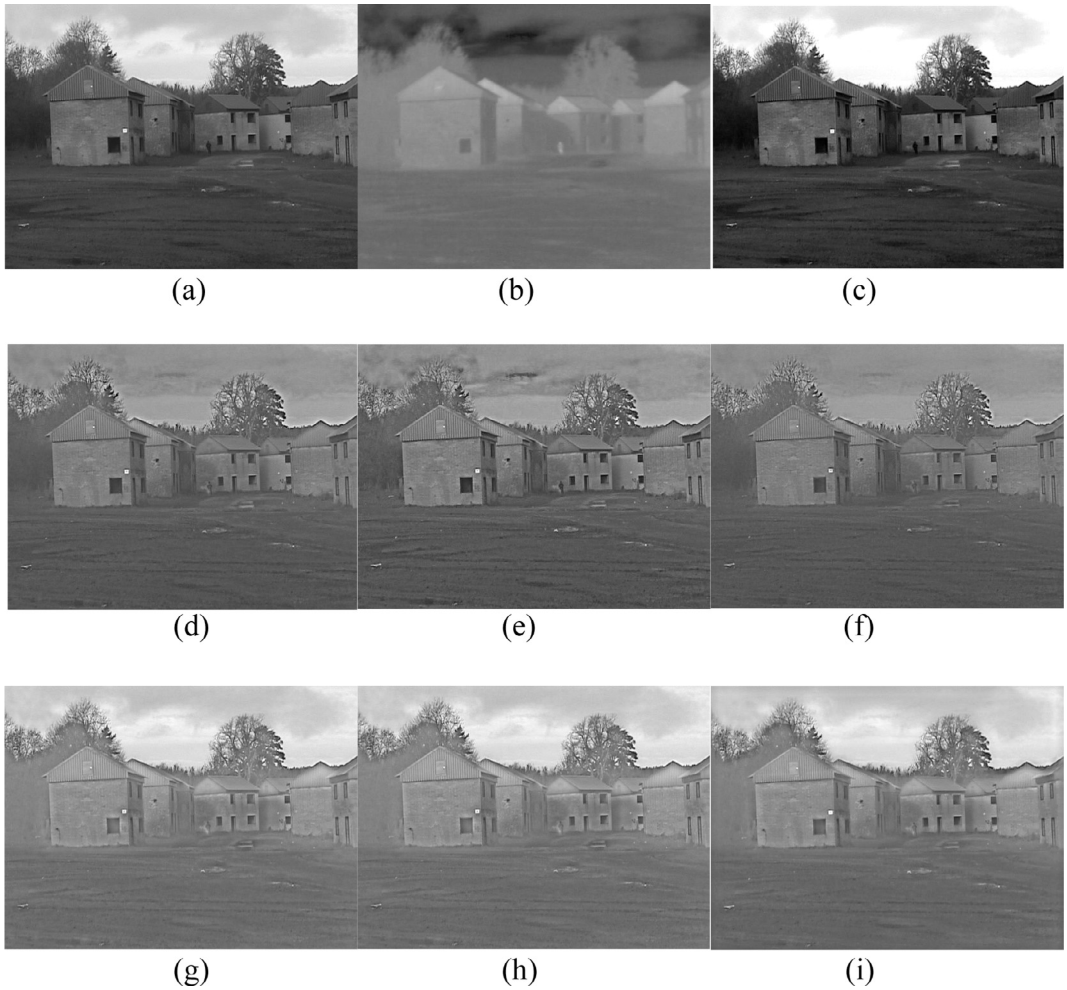


Fig. 9. Source images S3 and the fused images with different methods. (a) VI. (b) IR. (c) PCA. (d) LP. (e) CP. (f) WT. (g) NSST. (h) NSCT. (i) MF-MF.

Table 3
Objective evaluation metrics with different methods for Fig. 7.

Methods	MI	EN	AG	STD	SF	MV	Q^{ABF}
PCA	3.3834	6.8412	5.6072	41.4233	13.6770	74.7700	0.4093
LP	2.6012	7.2852	10.8110	44.1305	22.9021	89.8115	0.6137
CP	2.4392	7.1225	10.1225	42.3625	22.6402	86.5486	0.5166
WT	2.3973	7.1850	10.6726	39.5146	22.1715	88.3827	0.5420
NSST	2.9085	7.3735	10.6252	45.6077	22.2748	116.8013	0.5880
NSCT	3.0206	7.3680	10.5038	45.4740	22.1052	116.8092	0.6023
MF-MF	2.8766	7.2999	8.8197	44.0832	19.2454	117.2252	0.5496

Table 4
Objective evaluation metrics with different methods for Fig. 8.

Methods	MI	EN	AG	STD	SF	MV	Q^{ABF}
PCA	4.8779	6.7585	5.3901	32.1367	10.1297	86.8509	0.6814
LP	3.2057	6.7577	7.0462	31.8885	12.6753	87.5854	0.6935
CP	3.2074	6.7995	7.3444	32.2340	13.2969	87.0938	0.6796
WT	3.1671	6.7152	7.1632	30.9175	12.6791	87.7755	0.6706
NSST	3.3284	6.7455	7.1602	31.7984	12.8248	97.5898	0.6853
NSCT	3.3589	6.7425	7.0896	31.7681	12.7282	97.5913	0.6940
MF-MF	3.3600	6.6795	5.9460	30.7780	11.2382	97.1069	0.6698

fusion also was proposed by Liu et al. [170]. In IR and VI image fusion field, with more and more source image data being available, learning based methods may be another interesting direction, such as deep learning based techniques.

6.3. New applications

The IR and VI image fusion methods in the application of face recognition were widely studied and obtained good effects;

Table 5
Objective evaluation metrics with different methods for Fig. 7.

Methods	MI	EN	AG	STD	SF	MV	Q^{ABF}
PCA	5.1979	6.9176	6.2193	79.1938	17.1633	103.9673	0.6811
LP	2.9147	6.4933	5.2367	28.9499	13.6714	110.3706	0.6665
CP	2.7715	6.6107	6.0323	29.3041	14.5412	109.7787	0.6617
WT	2.8809	6.4328	5.1799	27.9548	13.4985	110.3001	0.6189
NSST	3.3108	6.7948	5.1967	42.0412	13.6592	142.3288	0.6406
NSCT	3.3214	6.7927	5.1656	42.0381	13.6408	142.3297	0.6438
MF-MF	3.2175	6.7444	4.1497	41.4337	11.4990	142.5417	0.5766

beyond that, some scholars also pay attentions to other biological recognition fields, such as ear-based human identification [18], periocular region-based person identification [20], iris recognition [26], which show us some new prospects for the applications of IR and VI image fusion techniques. Similarity, except the traditional applications of military surveillance, there may be some new situations which also need this kind of technology, such as road obstacle classification [15], task-based scanpath assessment in complex scenarios [21], and tracking of camouflaged target [22].

Many new demands are raised in response to the proper time and conditions. For examples, there are some important applications in agriculture, such as, predicting apple fruit firmness and soluble solids content [12], fruit detection [13], and diseased plants using silhouette extraction [14]. Possible applications, such as advanced assessment in nondestructive testing scheme [16], assistant medical treatment [36], assessments for visually enhancing for old documents [48,49] and construction technology [50], also will be got more attentions. All these applications will push to the development of IR and VI image fusion technology.

6.4. Computation complexity and availability

IR and VI image fusion methods are generally integrated in an independent work platform which has limited computing resource and storage space; as a result, it requires the fusion methods have the characteristics of high computational efficiency and low space requirement to make the equipment as easy-use as possible; especially, in military portable devices and airborne equipment, which are the main applied fields of the IR and VI image fusion technologies.

At present, most of IR and VI image fusion algorithms tend to be complicated; it is well known that various complicated algorithms are proposed to achieve good fusion effect. But quite a number of researchers do not consider the requirements of practical applications; the complex algorithm is not easy to be used in equipment development. And most researchers pay little attention to the hardware realization. In addition, because the operation time is long and the computing resources consuming is huge, which caused some adverse impacts for the actual application of IR and VI image fusion algorithm. There were some researchers who have been proposed some low-cost and quick image fusion algorithm to fulfil this kind of tasks which should get more attentions of scholars [28,70,101,146]. In view of this, we think some lightweight and effective IR and VI fusion scheme is very encouraging. Fortunately, some researchers have been noticed this kind of problems and begun to reduce the complexity and accelerated the computational speed of IR and VI image fusion algorithms [10,84,96,146]. And the IR and VI image fusion algorithm which was proposed by Bavirisetti et al. made a good start for us [146], which only two simple filter and some easily mathematical operation were adopted to their scheme.

6.5. Anti-noise performance

In general, most of researches only take the performance of IR and VI image fusion methods into consideration, and ignore the anti-noise performance of image fusion technologies. It is well known that IR image contains a lot of noise; as a result, it will be very useful for IR and VI image fusion methods if the anti-noise ability can be taken into account by researchers. We think that the anti-noise performance of this kind of algorithm is one of the most important problems in further IR and VI image fusion study; and it should be paid more attentions by researchers. At present, some researchers have begun to take this problem into account in their studies [8,139]; and we think the fusion algorithm would be more suitable in practical application by considering the anti-noise problem. However, there are still no correspondingly detailed research to this problem, and it also does not exist a completely convincing scheme that proposed by researchers.

6.6. IR and VI image fusion quality evaluation

In practice, it does not exist in reference image for the quality evaluation of the IR and VI image fusion method, and the common mathematical calculation methods are not necessarily suitable for image quality assessment. These reasons will make the evaluation system be untrusted. As a result, it is not an easy task to quantitatively evaluate the fused IR and VI image quality [158]. Many scholars proposed lots of IR and VI image fusion evaluation methods in recent years [160,162,167], however, none of them is universally believed to be always more reasonable and effective than others in IR and VI image fusion field. Thus, the researchers synchronously apply several metrics to make a comprehensive evaluation [158]. Besides, recent advances in image fusion methods have necessitated the creation of new ways of assessing fused images, which have previously focused on the use of subjective quality ratings combined with computational metric assessment [21].

The image evaluation methods based on deep learning are worthy of being studied [171,172,173,174], because the mathematics based objective evaluation index system has various shortcomings at present. Since the special nature of IR and VI images [175,176], the conventional image fusion quality evaluation is not fully adapted to this field. Besides, current objective image evaluation system does not completely suitable for human visual characteristics, and the higher cognitive level image evaluation method is scarce. As a result, we think that the analysis of the fused image quality assessment method for IR and VI image should take more human's visual characteristics into consideration, and it needs to be further studied. The validity of tone mapping operators is analyzed by the psychophysical experiments in the research of Ledda et al., and it reveals that the psychophysical method is effective and feasible for the evaluation of image quality and also shows that it could be used to evaluate the fused quality of IR and VI image [177].

7. Conclusion

IR and VI image fusion is a popular research field of information fusion. In recent years, the continually proposed IR and VI image fusion methods make this kind of technology rapidly develop and gradually become perfection; however, there still exist some further improvements and potential directions in different applications of IR and VI image fusion. In this paper, we surveyed the techniques of IR and VI image fusion and provided a summarization in this field; erenow, the Introduction of this survey detailedly describes the classification strategies of IR and VI image fusion technologies. With the rapid improvement of technologies and the powerfully promotion of various applications, more specific and effective IR and VI image fusion techniques would be needed; and the corresponding image fusion quality measures also need further development. There are some possibilities to improve the image fusion quality by concluding this survey: (i) the resolution capabilities of the advanced image representation algorithm for image structure information will further improve the effectiveness of the fused image; (ii) the advanced image fusion strategies will provide better guarantee for image fusion quality. In addition, there are several possible research directions: (i) the improvement of anti-noise performance of IR and VI image fusion technologies; (ii) the reduction of complexity and the improvement of efficiency for the image fusion methods; (iii) the suitable method of the fused IR and VI image evaluation for human perception. In this paper, we conclude the corresponding challenges and tendencies of IR and VI image fusion field, which show that various applications and demands would further promote the development of this kind of technique.

Conflict of interest

The authors declare that there is no conflict of interest in this article.

Acknowledgements

The authors thank the editors and the anonymous reviewers for their careful works and valuable suggestions for this study. This study is supported by the National Natural Science Foundation of China (No. 61365001, No. 61463052 and No. 61640306). We thank to the support of Scientific Research Fund of Education Department of Yunnan Province in China (No. 2017YJS108) and Doctoral Candidate Academic Award of Yunnan Province in China. We also thank Dr. Jingyu Hou and Dr. Shin-Jye Lee for their valuable advises.

References

[1] Zhang Baohua, Lu Xiaoqi, Pei Haiquan, Zhao Ying, A fusion algorithm for infrared and visible images based on saliency analysis and non-subsampled Shearlet transform, Infrared Phys. Technol. 73 (2015) 286–297.
[2] Haiyan Jin, Yanyan Wang, A fusion method for visible and infrared images based on contrast pyramid with teaching learning based optimization, Infrared Phys. Technol. 64 (2014) 134–142.
[3] Zhaodong Liu, Hongpeng Yin, Bin Fang, Yi Chai, A novel fusion scheme for visible and infrared images based on compressive sensing, Optics Commun. 335 (2015) 168–177.
[4] Shutao Li, Xudong Kang, Leyuan Fang, Jianwen Hub, Haitao Yin, Pixel-level image fusion: a survey of the state of the art, Inform. Fusion 33 (2017) 100–112.
[5] Guangmang Cui, Huajun Feng, Zhihai Xu, Qi Li, Yueting Chen, Detail preserved fusion of visible and infrared images using regional saliency extraction and multi-scale image decomposition, Opt. Commun. 341 (2015) 199–209.
[6] Xinnan Fan, Pengfei Shi, Jianjun Ni, Min Li, A thermal infrared and visible images fusion based approach for multitarget detection under complex environment, Math. Problems Eng. 2015 (2015), 750708:1–750708:11.
[7] G. Bhatnagar, Z. Liu, A novel image fusion framework for night-vision navigation and surveillance, SIViP 9 (1) (2015) 165–175.
[8] Alex Pappachen James, Belur V. Dasarathy, Medical image fusion: a survey of the state of the art, Inform. Fusion 19 (2014) 4–19.
[9] Amanda C. Muller, S. Narayanan, Cognitively-engineered multisensor image fusion for military applications, Inform. Fusion 10 (2009) 137–149.
[10] Z.J. Feng, X.L. Zhang, L.Y. Yuan, et al., Infrared target detection and location for visual surveillance using fusion scheme of visible and infrared images, Math. Problems Eng. 3 (2013) 831–842.
[11] Tarek Elguebaly, Nizar Bouguil, Finite asymmetric generalized Gaussian mixture models learning for infrared object detection, Comput. Vis. Image Underst. 117 (2013) 1659–1671.
[12] Fernando Mendozaa, Renfu Lu, Haiyan Cen, Comparison and fusion of four nondestructive sensors for predicting apple fruit firmness and soluble solids content, Postharvest Biol. Technol. 73 (2012) 89–98.
[13] D.M. Bulanon, T.F. Burks, V. Alchanatis, Image fusion of visible and thermal images for fruit detection, Biosys. Eng. 103 (1) (2009) 12–22.
[14] Shan-e-Ahmed Raza, Victor Sanchez, Gillian Prince, John P. Clarkson, Nasir M. Rajpoot, Registration of thermal and visible light images of diseased plants using silhouette extraction in the wavelet domain, Pattern Recogn. 48 (2015) 2119–2128.
[15] Anca Apatean, Alexandrina Rogozan, Abdelaziz Bensrhair, Visible-infrared fusion schemes for road obstacle classification, Transp. Res. Part C 35 (2013) 180–192.
[16] K. Eisler, C. Homma, M. Goldammer, M. Rothenfusser, W. Arnold, Fusion of visual and infrared thermography images for advanced assessment in nondestructive testing, Rev. Sci. Instrum. 84 (2013), 064902:1-5.
[17] W.K. Wong, Haitao Zhao, Eyeglasses removal of thermal image based on visible information, Inform. Fusion 14 (2013) 163–176.
[18] Ayman Abaza, Thirimachos Bourlai, On ear-based human identification in the mid-wave infrared spectrum, Image Vis. Comput. 31 (2013) 640–648.
[19] A. González, Z. Fang, Y. Socarras, et al., Pedestrian detection at Day/Night time with visible and FIR cameras: a comparison 820 Sensors 16 (6) (2016), :1-820:11.
[20] Muhammad Uzairn, Arif Mahmood, Ajmal Mian, Chris McDonald, Periocular region-based person identification in the visible, infrared and hyperspectral imager, Neurocomputing 149 (2015) 854–867.
[21] T.D. Dixon, S.G. Nikolov, J.J. Lewis, et al., Task-based scanpath assessment of multi-sensor video fusion in complex scenarios, Inform. Fusion 11 (1) (2010) 51–65.
[22] Mohammed Talha, Rustam Stolkin, Particle filter tracking of camouflaged targets by adaptive fusion of thermal and visible spectra camera data, IEEE Sens. J. 14 (1) (2014) 159–166.
[23] Seong G. Kong, Jingu Heo, Besma R. Abidi, Joonki Paik, Mongi A. Abidi, Recent advances in visual and infrared face recognition-a review, Comput. Vis. Image Underst. 97 (2005) 103–135.
[24] George Bebis, Aglika Gyaourova, Saurabh Singh, Ioannis Pavlidis, Face recognition by fusing thermal infrared and visible imagery, Image Vis. Comput. 24 (2006) 727–742.
[25] W. Hizem, L. Allano, A. Mellakh, B. Dorizzi, Face recognition from synchronised visible and near-infrared images, IET Signal Proc. 3 (4) (2009) 282–288.
[26] Faranak Shamsafar, Hadi Seyedarabi, Ali Aghagolzadeh, Fusing the information in visible light and near-infrared images for iris recognition, Mach. Vis. Appl. 25 (2014) 881–899.
[27] G. Hermosilla, F. Gallardo, G. Farias, et al., Fusion of visible and thermal descriptors using genetic algorithms for face recognition systems, Sensors 15 (8) (2015) 17944–17962.
[28] Junfeng Bai, Yong Ma, Jing Li, Hao Li, Yu Fang, Rui Wang, Hongyuan Wang, Good match exploration for thermal infrared face recognition based on YWF-SIFT with multi-scale fusion, Infrared Phys. Technol. 67 (2014) 91–97.
[29] X. Yan, H. Qin, J. Li, et al., Infrared and visible image fusion with spectral graph wavelet transform, J. Opt. Soc. Am. A 32 (9) (2015) 1643–1652.
[30] Jianwen Hu, Shutao Li, The multiscale directional bilateral filter and its application to multisensory image fusion, Inform. Fusion 13 (2012) 196–206.
[31] Amit Aran, Soumika Munshi, Vinod K. Beri, Arun K. Gupta, Spectral band invariant wavelet-modified MACH filter, Opt. Lasers Eng. 46 (2008) 656–665.
[32] Richa Singh, Mayank Vatsa, Afzel Noore, Hierarchical fusion of multi-spectral face images for improved recognition performance, Inform. Fusion 9 (2008) 200–210.
[33] Richa Singh, Mayank Vatsa, Afzel Noore, Integrated multilevel image fusion and match score fusion of visible and infrared face images for robust face recognition, Pattern Recogn. 41 (2008) 880–893.
[34] Jiayi Ma, Ji Zhao, Yong Ma, Jinwen Tian, Non-rigid visible and infrared face registration via regularized Gaussian fields criterio, Pattern Recogn. 48 (2015) 772–784.
[35] R. Raghavendra, Bernadette Dorizzi, Ashok Rao, G. Hemantha Kumar, Particle swarm optimization based fusion of near infrared and visible images for improved face verification, Pattern Recogn. 44 (2011) 401–411.
[36] B.V. Hanna, A.M. Gorbach, F.A. Gage, et al., Intraoperative assessment of critical biliary structures with visible range/infrared image fusion, J. Am. Coll. Surg. 206 (3) (2008) 1227–1231.
[37] Mehrdad Eslami, Ali Mohammadzadeh, Developing a spectral-based strategy for urban object detection from airborne hyperspectral TIR and visible data, IEEE J. Select. Top. Appl. Earth Observat. Remote Sens. 9 (5) (2016) 1808–1816.
[38] Stephen R. Lach, John P. Kerekes, Xiaofeng Fan, Fusion of multiple image types for the creation of radiometrically-accurate synthetic scenes, J. Appl. Remote Sens. 3 (2009) 033501.
[39] Lei Han, Buzha Wulie, Yiling Yang, Hongqing Wang, Direct fusion of geostationary meteorological satellite visible and infrared images based on thermal physical properties, Sensors 15 (1) (2015) 703–714.

[40] X. Chang, L. Jiao, F. Liu, et al., Multicontourlet-based adaptive fusion of infrared and visible remote sensing images, IEEE Geosci. Remote Sens. Lett. 7 (3) (2010) 549–553.
[41] D.J.L. Qingping, X.U. Liang, Visible and infrared video fusion using uniform discrete curvelet transform and spatial-temp oral information, Chin. J. Electron. 24 (4) (2015) 761–766.
[42] Kan Ren, Fuyuan Xu, Super-resolution images fusion via compressed sensing and low-rank matrix decomposition, Infrared Phys. Technol. 68 (2015) 61–68.
[43] X. Lu, J. Zhang, T. Li, et al., Synergetic classification of long-wave infrared hyperspectral and visible images, IEEE J. Select. Top. Appl. Earth Observat. Remote Sens. 8 (7) (2015) 3546–3557.
[44] L. Han, L. Shi, Y. Yang, et al., Thermal physical property-based fusion of geostationary meteorological satellite visible and infrared channel images, Sensors 14 (6) (2014) 10187–10202.
[45] Gemma Piella, A general framework for multiresolution image fusion: from pixels to regions, Inform. Fusion 4 (2003) 259–280.
[46] P.S.P. Wang, Jianwei Yang, A review of wavelet-based edge detection methods, Int. J. Pattern Recognit Artif Intell. 26 (7) (2013) 965–1004.
[47] Attiq Ahmad, Muhammad Mohsin Riaz, Abdul Ghafoor, Tahir Zaidi, An Improved Infrared/Visible Fusion for Astronomical Images, Advances in Astronomy 2015 (2015), 203872:1–8.
[48] Marco Gargano, Duilio Bertani, Marinella Greco, John Cupitt, Davide Gadia, Alessandro Rizzi, A perceptual approach to the fusion of visible and NIR images in the examination of ancient documents, J. Cult. Heritage 16 (4) (2014) S15–S17.
[49] Seon Joo Kim, Fanbo Deng, Michael S. Brown, Visual enhancement of old documents with hyperspectral imaging, Pattern Recogn. 44 (2011) 1461–1469.
[50] S. Lagüela, J. Armesto, P. Arias, J. Herráez, Automation of thermographic 3D modelling through image fusion and image matching techniques, Autom. Construct. 27 (2012) 24–31.
[51] M. Kumar, S. Dass, A total variation-based algorithm for pixel-level image fusion, IEEE Trans. Image Process. 18 (9) (2009) 2137–2143.
[52] Petrović, S. Valdimir, Xydeas, S. Costas, Gradient-based multiresolution image fusion, IEEE Trans. Image Process. 13 (2) (2004) 228–237.
[53] Liang Xu, Junping Du, Zhenhong Zhang, Infrared-visible video fusion based on motion-compensated wavelet transforms, Iet Image Process. 9 (4) (2014) 318–328.
[54] Simon Denman, Todd Lamb, Clinton Fookes, Vinod Chandran, Sridha Sridharan, Multi-spectral fusion for surveillance systems, Comput. Electr. Eng. 36 (2010) 643–663.
[55] A.L. Chan, Fusing concurrent visible and infrared videos for improved tracking performance, Opt. Eng. 52 (1) (2013) 177–182.
[56] X.W. Zhang, Y.N. Zhang, Z. Guo, et al., Advances and perspective on motion detection fusion in visual and thermal framework, J. Infrar. Millimeter Waves 30 (4) (2012) 354–360.
[57] Yi Zhou, Abedalroof Mayyas, Ala Qattawi, Mohammed Omar, Feature-level and pixel-level fusion routines when coupled to infrared night-vision tracking scheme, Infrar. Phys. Technol. 53 (2010) 43–49.
[58] Congyi Liu, Zhongliang Jing, Gang Xiao, Bo Yang, Feature-based fusion of infrared and visible dynamic images using target detection, Chin. Opt. Lett. 5 (5) (2007) 274–277.
[59] E.P. Bennett, J.L. Mason, L. Mcmillan, Multispectral bilateral video fusion, IEEE Transact. Image Process. Publ. IEEE Signal Process. Soc. 16 (5) (2007) 1185–1194.
[60] W. He, W. Feng, Y. Peng, et al., Multi-level image fusion and enhancement for target detection, Optik-International Journal for Light and Electron Optics 126 (11–12) (2015) 1203–1208.
[61] A. Toet, Natural colour mapping for multiband nightvision imagery, Inform. Fusion 4 (3) (2003) 155–166.
[62] V. Tsagaris, V. Anastassopoulos, Fusion of visible and infrared imagery for night color vision, Displays 26 (4–5) (2005) 191–196.
[63] A. Toet, M.A. Hogervorst, Progress in color night vision, Opt. Eng. 51 (1) (2012) 010901.
[64] J.W. Davis, V. Sharma, Background-subtraction using contour-based fusion of thermal and visible imagery, Comput. Vis. Image Underst. 106 (2–3) (2007) 162–182.
[65] Z. Cheng, Q. Yang, B. Sheng, Deep Colorization, 2015 IEEE International Conference on Computer Vision (ICCV). IEEE, 2015, pp. 415–423.
[66] A. Toet, E.M. Franken, Perceptual evaluation of different image fusion schemes, Displays 24 (1) (2003) 25–37.
[67] Y. Yuan, J. Zhang, B. Chang, Objective quality evaluation of visible and infrared color fusion image, Opt. Eng. 50 (3) (2011) 103–108.
[68] E. Reinhard, M. Ashikhmin, B. Gooch, P. Shirley, Color transfer between images, IEEE Comput. Graphics Appl. 21 (5) (2001) 34–41.
[69] S. Yin, L. Cao, Y. Ling, et al., One color contrast enhanced infrared and visible image fusion method, Infrared Phys. Technol. 53 (2) (2010) 146–150.
[70] M.A. Hogervorst, A. Toet, Fast natural color mapping for night-time imagery, Inform. Fusion 11 (2) (2010) 69–77.
[71] M.A. Hogervorst, A. Toet, F.L. Kooi, TNO Defense Security and Safety, Method and System for Converting at Least one First-spectrum Image into a Second-spectrum Image, 2006.
[72] Y. Zheng, E.A. Essock, A local-coloring method for night-vision colorization utilizing image analysis and fusion, Inform. Fusion 9 (2) (2008) 186–199.
[73] Marom Levinera, Masha Maltz, A new multi-spectral feature level image fusion method for human interpretation, Infrared Phys. Technol. 52 (2009) 79–88.
[74] X. Qian, Y. Wang, B. Wang, Effective contrast enhancement method for color night vision, Infrared Phys. Technol. 55 (1) (2012) 130–136.
[75] X. Qian, Y. Wang, B. Wang, Fast color contrast enhancement method for color night vision, Infrared Phys. Technol. 55 (1) (2012) 122–129.
[76] Shaoshu Guo, W. Jin, L. Wang, Objective color harmony assessment for visible and infrared color fusion images of typical scenes, Opt. Eng. 51 (11) (2012) 1371–1379.
[77] X. Yu, J. Ren, Q. Chen, et al., A false color image fusion method based on multi-resolution color transfer in normalization YCBCR space, Optik-Int. J. Light Electron Opt. 125 (20) (2014) 6010–6016.
[78] R.S. Wang, T. Cai, Symbol-based image fusion for object recognition, in: Proceedings of the International Conference on Multisource-Multisensor Information Fusion (FUSION 98), 1998, pp. 509–514.
[79] V.E. Neagoe, A.D. Ropot, A.C. Mugioiu, Real time face recognition using decision fusion of neural classifiers in the visible and thermal infrared spectrum, Advanced Video and Signal Based Surveillance, 2007. AVSS 2007. IEEE Conference on. IEEE, 2007, pp. 301–306.
[80] Y.F. Zhao, Y.X. Yin, Decision-level fusion of infrared and visible images for face recognition, Control & Decision Conference IEEE, 2008, pp. 2411–2414.
[81] S. Yeom, Multi-frame decision level fusion for face classification based on a photon-counting linear discriminant analysis, Society of Photo-Optical Instrumentation Engineers, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 2011.
[82] Aili Wang, Jinna Jiang, Haoye Zhang, Multi-sensor Image Decision Level Fusion Detection Algorithm Based On D-S Evidence Theory, 2014 Fourth International Conference on Instrumentation and Measurement, Computer, Communication and Control, IEEE Conference on. IEEE, 2014, pp. 620–623.
[83] A.L. Cunha, J. Zhou, M.N. Do, The nonsubsampled contourlet transform: theory, design, and applications, IEEE Trans. Image Process 15 (10) (2006) 3089–3101.
[84] Chunhui Zhao, Yunting Guo, Yulei Wang, A fast fusion scheme for infrared and visible light images in NSCT domain, Infrared Phys. Technol. 72 (2015) 266–275.
[85] Hassan Ghassemian, A review of remote sensing image fusion methods, Inform. Fusion 32 (2016) 75–89.
[86] J. Romberg, Imaging via compressive sampling, Signal Process. Magazine IEEE 25 (2) (2008) 14–20.
[87] 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.
[88] Amina Jameel, Abdul Ghafoor, Muhammad Mohsin Riaz, Adaptive compressive fusion for visible/IR sensors, IEEE Sens. J. 14 (7) (2014) 2230–2231.
[89] Meng Ding, Li Wei, Bangfeng Wang, Research on fusion method for infrared and visible images via compressive sensing, Infrared Phys. Technol. 57 (2013) 56–67.
[90] X. Jin, R. Nie, D. Zhou, S. Yao, Y. Chen, J. Yu, et al., A novel DNA sequence similarity calculation based on simplified pulse-coupled neural network and Huffman coding, Phys. A Stat. Mech. Its Appl. 461 (2016) 325–338.
[91] X. Jin, R. Nie, D. Zhou, et al., Multifocus color image fusion based on NSST and PCNN, J. Sens. (2016), 8359602:1-8359602:12.
[92] X. Jin, D. Zhou, S. Yao, et al., Remote sensing image fusion method in CIELab color space using nonsubsampled shearlet transform and pulse coupled neural networks, J. Appl. Remote Sens. 10 (2) (2016), 025023:1-025023:18.
[93] Weiwei Kong, Yang Lei, Huaixun 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.
[94] G. Easley, D. Labate, W.Q. Lim, Sparse directional image representation using the discrete shearlet transforms, Appl. Comput. Harm. Anal. 25 (1) (2008) 25–46.
[95] Weiwei Kong, Longjun Zhang, Yang Lei, Novel fusion method for visible light and infrared images based on NSST-SF-PCNN, Infrared Phys. Technol. 65 (2014) 103–112.
[96] Weiwei Kong, Jianping Liu, Technique for image fusion based on NSST domain improved fast non-classical RF, Infrared Phys. Technol. 61 (2013) 27–36.
[97] Weiwei Kong, Technique for gray-scale visual light and infrared image fusion based on non-subsampled shearlet transform, Infrared Phys. Technol. 63 (2014) 110–118.
[98] Weiwei Kong, Binghe Wang, Yang Lei, Technique for infrared and visible image fusion based on non-subsampled shearlet transform and spiking cortical model, Infrared Phys. Technol. 71 (2015) 87–98.
[99] Jian Sun, Hongyan Zhu, Zongben Xu, Chongzhao Han, Poisson image fusion based on Markov random field fusion model, Inform. Fusion 14 (2013) 241–254.
[100] Takashi Shibata, Masayuki Tanaka, Masatoshi Okutomi, Versatile visible and near-infrared image fusion based on high visibility area selection, J. Electron. Imaging 25 (1) (2016) 013016.
[101] Jungong Han, Eric J. Pauwels, Paul de Zeeuw, Fast saliency-aware multi-modality image fusion, Neurocomputing 111 (2013) 70–80.
[102] G. Michael, Thomason, Jens Gregor, Higher order singular value decomposition of tensors for fusion of registered images, J. Electron. Imaging 20 (1) (2011) 013023.
[103] Yifeng Niu, Shengtao Xu, Lizhen Wu, Weidong Hu, Airborne infrared and visible image fusion for target perception based on target region segmentation and discrete wavelet transform, Math. Problems Eng. 275138 (2012) 1–10.

[104] He Li, Lei Liu, Wei Huang, Chao Yue, An improved fusion algorithm for infrared and visible images based on multi-scale transform, Infrared Phys. Technol. 74 (2016) 28–37.
[105] Jamal Saeedi, Karim Faez, Infrared and visible image fusion using fuzzy logic and population-based optimization, Appl. Soft Comput. 12 (2012) 1041–1054.
[106] Haiyan Jin, Licheng Jiao, Fang Liu, Yutao Qi, Fusion of infrared and visual images based on contrast pyramid directional filter banks using clonal selection optimizing, Opt. Eng. 47 (2) (2008) 027002.
[107] Haiyan Jin, Qian Xi, Yanyan Wang, Xinhong Hei, Fusion of visible and infrared images using multiobjective evolutionary algorithm based on decomposition, Infrar. Phys. Technol. 71 (2015) 151–158.
[108] H. Xu, Y. Wang, Y. Wu, et al., Infrared and multi-type images fusion algorithm based on contrast pyramid transform, Infrared Phys. Technol. 78 (2016) 133–146.
[109] Z. Liu, K. Tsukada, K. Hanasaki, et al., Image fusion by using steerable pyramid, Pattern Recogn. Lett. 22 (9) (2001) 929–939.
[110] Y. Niu, B. Wang, Hyperspectral Target Detection Using Learned Dictionary, IEEE Geosci. Remote Sens. Lett. 12 (7) (2015) 1–5.
[111] Jun Wang, Jinye Peng, Xiaoyi Feng, Guiqing He, Jianping Fan, Fusion method for infrared and visible images by using non-negative sparse representation, Infrared Phys. Technol. 67 (2014) 477–489.
[112] Z. Wang, S. Wang, Y. Zhu, et al., Review of image fusion based on pulse-coupled neural network, Arch. Comput. Methods Eng. (2015) 1–13.
[113] J. Du, W. Li, K. Lu, et al., An overview of multi-modal medical image fusion, Neurocomputing 215 (2016) 3–20.
[114] Luyi Bai, Changming Xu, Cong Wang, A review of fusion methods of multi-spectral image, Optik 126 (2015) 4804–4807.
[115] Hassan Ghassemian, A review of remote sensing image fusion methods, Inform. Fusion 32 (2016) 75–89.
[116] B. Yang, Z. Jing, H. Zhao, Review of pixel-level image fusion, J. Shanghai Jiaotong Univ. (Science) 15 (1) (2010) 6–12, http://dx.doi.org/10.1007/s12204-010-7186-y.
[117] R.S. Ghiass, O. Arandjelović, A. Bendada, et al., Infrared face recognition: a comprehensive review of methodologies and databases, Pattern Recogn. 47 (9) (2014) 2807–2824.
[118] S. Wu, Z. Fang, Z. Xie, W. Liang, Recent Advances in Face Recognition-Chapter 13: Blood Perfusion Models for Infrared Face Recognition, InTech, 2008.
[119] M. Yin, P. Duan, W. Liu, et al., A novel infrared and visible image fusion algorithm based on shift-invariant dual-tree complex shearlet transform and sparse representation, Neurocomputing 226 (2017) 182–192.
[120] Haitao Yin, Shutao Li, Leyuan Fang, Simultaneous image fusion and super-resolution using sparse representation, Inform. Fusion 14 (2013) 229–240.
[121] Lu Xiaoqi, Zhang Baohua, Zhao Ying, Liu He, Pei Haiquan, The infrared and visible image fusion algorithm based on target separation and sparse representation, Infrared Phys. Technol. 67 (2014) 397–407.
[122] Kangjian He, Dongming Zhou, Xuejie Zhang, Rencan Nie, Quan Wang, Xin Jin, Infrared and visible image fusion based on target extraction in the nonsubsampled contourlet transform domain, J. Appl. Remote Sens. 11 (1) (2017) 015011.
[123] Yong Chen, Jie Xiong, Huan-lin Liu, Qiang Fan, Fusion method of infrared and visible images based on neighborhood characteristic and regionalization in NSCT domain, Optik 125 (2014) 4980–4984.
[124] Jianhua Adu, Jianhong Gan, Yan Wang, Jian Huang, Image fusion based on nonsubsampled contourlet transform for infrared and visible light image, Infrared Phys. Technol. 61 (2013) 94–100.
[125] Qiong Zhang, Xavier Maldague, An adaptive fusion approach for infrared and visible images based on NSCT and compressed sensing, Infrared Phys. Technol. 74 (2016) 11–20.
[126] Din-Chang Tseng, Yu-Shuo Liu, Chang-Min Chou, Image fusion with contrast improving and feature preserving, Math. Problems Eng. 2015 (603195) (2015) 1–14.
[127] Xiangzhi Bai, Morphological center operator based infrared and visible image fusion through correlation coefficient, Infrared Phys. Technol. 76 (2016) 546–554.
[128] P. Zhao, Image-fusion-based contour extraction scheme, Opt. Eng. 44 (12) (2005), 127005:1-127005:7.
[129] Shao Zhenfeng, Liu Jun, Cheng Qimin, Fusion of infrared and visible images based on focusmeasure operators in the curvelet domain, Appl. Opt. 51 (12) (2012) 1910–1921.
[130] Yanfei Chen, Nong Sang, Attention-based hierarchical fusion of visible and infrared images, Optik 126 (2015) 4243–4248.
[131] Liu Kun, Guo Lei, Li Huihui, Chen Jingsong, Fusion of infrared and visible light images based on region segmentation, Chin. J. Aeronaut. 22 (2009) 75–80.
[132] Huafeng Li, Hongmei Qiu, Zhengtao Yu, Yafei Zhang, Infrared and visible image fusion scheme based on NSCT and low-level visual features, Infrared Phys. Technol. 76 (2016) 174–184.
[133] Zhizhong Fu, Xue Wang, Jin Xu, Ning Zhou, Yufei Zhao, Infrared and visible images fusion based on RPCA and NSCT, Infrared Phys. Technol. 77 (2016) 114–123.
[134] K. Madheswari, N. Venkateswaran, Swarm intelligence based optimization in thermal image fusion using dual tree discrete wavelet transform, Quantitat. InfraRed Thermogr. J. 14 (2017) 24–43.
[135] Zhuozheng Wang, J.R. Deller Jr., Blair D. Fleet, Pixel-level multisensor image fusion based on matrix completion and robust principal component analysis, J. Electron. Imaging 25 (1) (2016) 013007.
[136] S. Li, B. Yang, J. Hu, Performance comparison of different multi-resolution transforms for image fusion, Inform. Fusion 12 (2) (2011) 74–84.
[137] Durga Prasad Bavirisetti, Ravindra Dhuli, Fusion of infrared and visible sensor images based on anisotropic diffusion and Karhunen-Loeve transform, IEEE Sens. J. 14 (10) (2015) 203–209.
[138] Nedeljko Cvejic, David Bull, Nishan Canagarajah, Region-based multimodal image fusion using ICA bases, IEEE Sens. J. 7 (5) (2007) 743–751.
[139] Songfeng Yin, Liangcai Cao, Yongshun Ling, Guofan Jin, Fusion of noisy infrared and visible images based on anisotropic bivariate shrinkage, Infrared Phys. Technol. 54 (2011) 13–20.
[140] A. Toet, J.M. Valeton, L.J. van Ruyven, Merging thermal and visual images by a contrast pyramid, Opt. Eng. 28 (7) (1989) 789–792.
[141] Z. Zhou, B. Wang, S. Li, et al., Perceptual fusion of infrared and visible images through a hybrid multi-scale decomposition with Gaussian and bilateral filters, Inform. Fusion 30 (2016) 15–26.
[142] Jufeng Zhao, Qiang Zhou, Yueting Chen, Huajun Feng, Zhihai Xu, Qi Li, Fusion of visible and infrared images using saliency analysis and detail preserving based image decomposition, Infrared Phys. Technol. 56 (2013) 93–99.
[143] X. Yan, H. Qin, J. Li, et al., Infrared and visible image fusion using multiscale directional nonlocal means filter, Appl. Opt. 54 (13) (2015) 4299–4308.
[144] Wei Gan, Xiaohong Wu, Wei Wu, Xiaomin Yang, Chao Ren, Xiaohai He, Kai Liu, Infrared and visible image fusion with the use of multi-scale edge-preserving decomposition and guided image filter, Infrar. Phys. Technol. 72 (2015) 37–51.
[145] J. Ma, Z. Zhou, B. Wang, et al., Infrared and visible image fusion based on visual saliency map and weighted least square optimization, Infrar. Phys. Technol. 82 (2017) 8–17.
[146] Durga Prasad Bavirisetti, Ravindra Dhuli, Two-scale image fusion of visible and infrared images using saliency detection, Infrared Phys. Technol. 76 (2016) 52–64.
[147] Jufeng Zhao, Xiumin Gao, Yueting Chen, Huajun Feng, Daodang Wang, Multi-window visual saliency extraction for fusion of visible and infrared images, Infrared Phys. Technol. 76 (2016) 295–302.
[148] J. Zhao, G. Cui, X. Gong, et al., Fusion of visible and infrared images using global entropy and gradient constrained regularization, Infrared Phys. Technol. 81 (2017) 201–209.
[149] Peng Zhao, Ting-Fa Xu, Guo-Qiang Ni, Object's contour detection scheme based on image fusion, Int. J. Comput. Math. 87 (13) (2010) 2857–2870.
[150] Xin Zhou, Xin Yin, Rui-An Liu, Wei Wang, Infrared and visible image fusion technology based on directionlets transform, EURASIP J. Wireless Commun. Network. (2013) 42, 2013.
[151] P. Zhu, Z. Huang, H. Lei, Fusion of infrared and visible images based on BEMD and NSDFB, Infrared Phys. Technol. 77 (2016) 82–93.
[152] Fengbao Yang, Hong Wei, Fusion of infrared polarization and intensity images using support value transform and fuzzy combination rules, Infrared Phys. Technol. 60 (2013) 235–243.
[153] Yong Ma, Jun Chen, Chen Chen, Fan Fan, Jiayi Ma, Infrared and visible image fusion using total variation model, Neurocomputing 202 (2016) 12–19.
[154] Jiayi Ma, Chen Chen, Chang Li, Jun Huan, Infrared and visible image fusion via gradient transfer and total variation minimization, Inform. Fusion 31 (2016) 100–109.
[155] Mohammad Bagher Akbari Haghighat, Ali Aghagolzadeh, Hadi Seyedarabi, A non-reference image fusion metric based on mutual information of image features, Comput. Electr. Eng. 37 (2011) 744–756.
[156] H. Li, J. Zhang, J. Sun, A visual analytics approach for deterioration risk analysis of ancient frescoes, J. Visualizat. 1–14 (2016).
[157] A. Toet, J.K. Ijspeert, A.M. Waxman, et al., Fusion of visible and thermal imagery improves situational awareness, Displays 18 (2) (1997) 85–95.
[158] Yu Liu, Shuping Liu, Zengfu Wang, A general framework for image fusion based on multi-scale transform and sparse representation, Inform. Fusion 24 (2015) 147–164.
[159] Z. Wang, A.C. Bovik, A universal image quality index, IEEE Signal Process. Lett. 9 (3) (2002) 81–84.
[160] G. Piella, H. Heijmans, A new quality metric for image fusion, in: Proceedings of the Tenth International Conference on Image Processing, Barcelona, Spain, 2003, pp. 173–176.
[161] C.S. Xydeas, V.S. Petrovic, Objective image fusion performance measure, Electron. Lett. 36 (4) (2000) 308–309.
[162] G. Qu, D. Zhang, P. Yan, Information measure for performance of image fusion, Electron. Lett. 38 (7) (2002) 313–315.
[163] H.R. Sheikh, A.C. Bovik, Image information and visual quality, in: Proceedings-ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing, vol. 3, 2004, iii-709-12.
[164] Z. Wang, A.C. Bovik, H.R. Sheikh, E.P. Simoncelli, Image quality assessment: from error visibility to structural similarity, IEEE Trans. Image Process. 13 (2004) 600–612.
[165] http://www.ece.uwaterloo.ca/~z70wang/research/ssim.
[166] J. Zhao, R. Laganiere, Z. Liu, Performance assessment of combinative pixel-level image fusion based on an absolute feature measurement, Int. J. Innovative Comput. Inf. Control 6 (A3) (2007) 1433–1447.
[167] H.R. Sheikh, A.C. Bovik, Image information and visual quality, IEEE Trans. Image Process. 15 (2) (2006) 430–444.
[168] Li Hong, Liu Fang, Yang Shuyuan, Zhang Kai, Remote sensing image fusion based on deep support value learning networks, Chin. J. Comput. 39 (8) (2016) 1583–1596.

[169] J. Zhong, B. Yang, G. Huang, et al., Remote sensing image fusion with convolutional neural network, Sens. Imag. 17 (1) (2016) 1–16.
[170] Y. Liu, X. Chen, H. Peng, et al., Multi-focus image fusion with a deep convolutional neural network, Inform. Fusion 36 (2016) 191–207.
[171] W. Hou, X. Gao, D. Tao, et al., Blind image quality assessment via deep learning, IEEE Transact. Neural Networks Learn. Syst. 26 (6) (2015) 1275–1286.
[172] S. Bianco, L. Celona, P. Napoletano, et al. On the Use of Deep Learning for Blind Image Quality Assessment, 2016. https://arxiv.org/abs/1602.05531.
[173] W. Hou, X. Gao, D. Tao, et al., Blind image quality assessment via deep learning, IEEE Transact. Neural Networks Learn. Syst. 26 (6) (2017) 1275–1286.
[174] L. Kang, P. Ye, Y. Li, et al., Convolutional neural networks for no-reference image quality assessment, CVPR IEEE (2014) 1733–1740.
[175] Shaoshu Gao, Weiqi Jin, Lingxue Wang, Quality assessment for visible and infrared color fusion images of typical scenes, Chin. Opt. Lett. 8 (2012) 18–22.
[176] X. Yu, Q. Chen, G. Gu, et al., Perceived assessment metrics for visible and infrared color fused image quality without reference image, Opt. Rev. 22 (1) (2015) 109–122.
[177] P. Ledda, A. Chalmers, T. Troscianko, et al., Evaluation of tone mapping operators using a High Dynamic Range display, ACM Transact. Graph. 24 (3) (2005) 640–648.