International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-7 Issue-4s, November 2018

# A Comprehensive Survey of Multimodal Image Fusion Schemes

**Bhavna Bharath, Suganthi N**

***Abstract—Multimodal images are scenes with anatomy details that are captured using two different devices. Different imaging techniques give complementary details about what is visualized. Infrared and visual images are examples of multimodal images that are fused together in order to obtain a single comprehensive fused image. Combining multimodal images yield enhanced features for image analysis, feature extraction and detection. Infrared and Visual image fusion will fuse the source images into single extensive image to raise image quality .This will in turn decrease the redundancy in image data. This is broadly used in different applications to improve the perception of the scene. The reliability, accuracy and complementary details of the scene in the resultant fused image makes these approaches be used in multiple areas. Recently, many fusion methods have been formulated due to the sprouting demands & advancement of image depiction schemes. However, a unified survey paper about this field has not been published in a few years. Consequently, we make a survey report to record the methodical advancements of visual and infrared image fusion. In this paper, firstly the overview of applications of IR and VI image fusion is represented. Secondly, we present the existing state of the art fusion techniques. Finally, image quality metrics are discussed to measure the efficiency of the fusion algorithm. Although, this survey halts with various fusion methods that have been proposed earlier there is still room for improvement in research in the field of multimodal image fusion.***

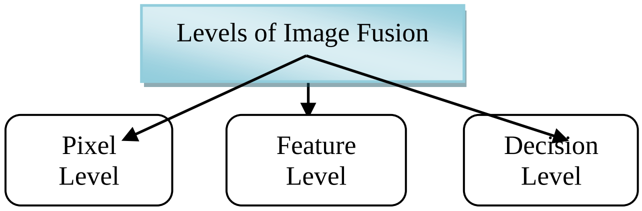
***Keywords: Multimodal , Image Analysis, Image Fusion.***

## I. INTRODUCTION

In general, all objects emit or radiate infrared or heat energy with different frequencies called thermal radiation which are not visible to naked eyes[1]. The infrared sensor detects and records the thermal radiation thus creating an image based on that detail. Thermal imaging is a night vision technology that enables a method that improves the visibility of objects in dark environment. It works in environments even when there is no illumination or without ambient light[3]. It can penetrate though fog, haze and smoke.

The attributes of infrared images would decrease the external environment influence such as sunlight, smog and other conditions. Infrared images are receptive to areas and objects with thermal characteristics that could not be depicted in visual images[1].Visual images are captured to note the spectral information of the objects in the scene which contain corner and interesting points, lines, visible edges which provide perceptual scene description to naked eyes[5]. The objective of thermal and visual image fusion is to obtain a correlated fused image with ample detailed information in visual image and adequate target areas in infrared images[4]. The reliable, accurate and complementary details of the components of the scene in the fused image make these methods be majorly applied to applications such as military surveillance[27], industrial application[2], biometric identification[6], medical imaging[7], space exploration[27], remote sensing[3,27].,etc.

Different infrared and visual image fusion methods have been developed depending upon various applications. They are divided into three levels namely; pixel level feature level and decision level.




Pixel level fusion method is a popular and widely used research strategy for the image fusion field on the whole. This fusion level introduces minimum artefacts in the final image after fusion[13]. Other fusion levels such as feature level or decision level fusion are the composition of image data in the form of probabilistic variables and image feature details[16]. The pixel level based infrared and visual image fusion technique could be in turn classified into spatial domain and transform domain based algorithm.

A spatial domain based fusion algorithm use block based or weighted averaging principle; PCA[12] and ICA[13] are the methods used that are based on subspace analysis. Some of the transform domain based image fusion algorithms are wavelet[16], pyramids[5], curvelet[26] and contourlet[14] ; Some of the novel multiscale decomposition methods are NSST[18], NSCT[20] and other techniques of decomposition[7]. In addition, some of the image representation methods used in infrared and visual image fusion are sparse representation(SR)[26] and compressive sensing(CS)[26].

A feature level based image fusion technique generally deals with core comprehensive feature details and characteristics of image related to structure such as corners, edges, lines, points, contours, textures, to enhance the clarity of the image, after which the image information from each source image will be individually extracted and the most pertinent features will be combined. Decision level fusion is the integration of multiple or different level methods will combine the respective advantages of the methods together which provides better fusion results.

**Manuscript published on 30 November 2018.**
***Correspondence Author(s)**
**Bhavna Bharath,** Assistant Professor, Department of Computer Science & Engineering, Kumaraguru College of Technology, Coimbatore,India,
**Dr. Suganthi N ,** Assistant Professor, Department of Computer Science & Engineering, Kumaraguru College of Technology, Coimbatore, India,

Wang et al. [27] described two issues that need to be considered while performing image fusion which are:

- How efficiently the image data is extracted from the source images which is the fundamental or pivotal aspect to the quality of image fusion.
- How to satisfactorily fuse the image data from numerous information sources into the final fused image.

The fusion effect has been rapidly improved as many kinds of fusion strategies & advanced image representation methods have been introduced to achieve infrared and visual image fusion. Nonetheless, we deduce that there are still many issues with regard to fusion methods that are not studied properly and solved. It is important to note that the existing visual and thermal image fusion strategies will not be able to deal with and meet the accelerated advancement of computer sensor technologies . On the other hand, the development of advanced sensors and computing devices with exceptional performance are carried out which will further stimulate the development of visual and infrared image fusion technologies.

Recently, number of image fusion approaches have been proposed due to the increased development of imaging standards and techniques[27]. There are multiple applications in this field such as remote sensing[6], medical field[7] etc,. These applications have been well analysed and studied and many review papers have been published by research scholars in these techno spheres[19-25]. However, an integrated paper consisting of common infrared and visual image fusion techniques have not been published in the last few years. Therefore, this paper constitutes a panorama of visual and infrared image fusion methodologies, application and objective assessment.

This survey concludes that there are still many challenges, new applications though many fusion methods exist; there is still room for rapid development of sensor technology, devices and analysis theories. This paper intends to provide a brief explanation of current trend and serve as a representation of different applications which would give a viewpoint to current status of research in this field.

This paper is organized as follows: In Section II, a brief description of applications are provided to give out the demands of sudden technique. Section III will provide an overview of multimodal image fusion methodologies. Section IV will examine the image using quantitative assessment metrics. Section V will cover the estimated observations from typical methods. Section VI concludes the paper.

## II. Major Application Areas

Though there are other traditional applications mentioned in other papers, this visual and infrared image fusion technique can also be made use of for multiple practical applications such as night vision, video fusion, biological recognition, surveillance and many such applications. This section mainly covers the major application areas in which fusion is incorporated.

### *A.* Video Fusion

Practical applications such as industrial applications, surveillance and so on need to fuse the dynamic images or visual and infrared videos. The fusion methods can combine the complementary image data which improves the accuracy and robustness of the video. The flow diagram for video fusion is shown in Fig.1

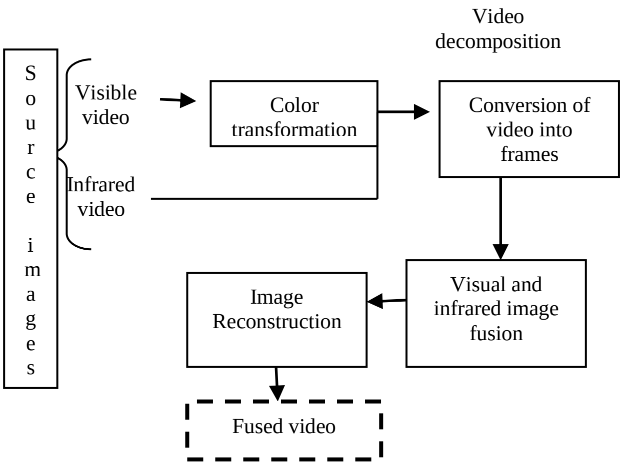



**Fig.1. Video fusion flow diagram.**

Compared to other visual and infrared image fusion methods, when video fusion is to be performed it must possess high computational efficiency with very low space requirements. With the fast development of sensor technology and imaging equipments, multi-resolution video fusion has attracted the attention of many researchers. Beyond that, when stability and consistency of image data in temporal and spatial dimensions are considered, two types of visual and infrared video fusion strategies were stated by Chang et al. [27] , where one method is on spatial temporal information and uniform discrete curvelet transform and another method is a video fusion algorithm based on motion estimation and compensation in the wavelet domain.

Lamb et al. examined many methodologies for fusion of infrared and thermal images for deserted object detection and human tracking for which the researchers proposed a modified condensation filter to record and support in the fusion of different modalities and achieved intriguing results in the area of target tracking. Jing et al. worked on target detection in order to improve the efficiency and performance of feature level multimodal dynamic image fusion method based on dual-tree complex wavelet transform(DT-CWT) algorithm. This algorithm was employed to decompose the source images.

Visual and infrared video must concurrently take both spatial dimensions and temporal dimensions of the video so that it can be used in distinct and factual image fusion applications[4].The advancement in transform domain video fusion provided high grade video quality but suffered high computational complexity[27]. In video fusion apart from performance, the artifacts , color and so on also need to be taken into consideration.


Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-7 Issue-4s, November 2018

### *B.* Night Vision

Visual and infrared image fusion for night colour representation is a composite way for merging night time thermal or infrared with visual imagery[10]. The resultant colour image has a natural illuminated appearance after colour transfer technique is applied in order to make the observer understand the scene in the image by making the perception of the scene clear[6].

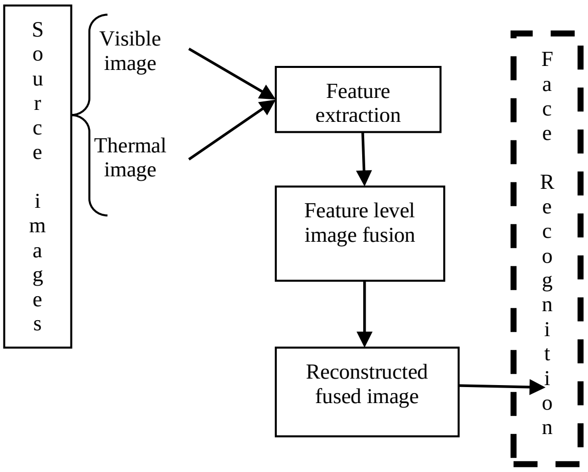



**Fig.2. Flow of colour transfer technique in night vision for face recognition.**

The core motivation for night vision based research is that human visual system is less sensitive to gray image than it is with colour image because human eye can detect huge number of colours but it can discriminate about hundred shades of gray at a time[6]. This is the reason gray scale image colorization has got the attention of researchers. It is also used in colour contrast enhancement of fused image for military equipment applications[10]. Perceptual evaluation issues for quality of visual and infrared colour image fusion is mentioned in the paper [11]. The objective assessment of quality of visual and thermal image fusion was proposed by Yuan et al.[25]. The detailed information and colour of the final fused image forms the most significant part of this area. The colour depends on the reference image(visual image) and the colour transfer technique. The fusion rule and image representation method play a vital role where the image details in the fused image is concerned. The need for reference image in visual and infrared image fusion is essential because the quality of resultant fused image depends on it. Though there is flexibility in using a fusion algorithm, the dependency of using a reference image should be decreased since the reference image is similar to the scene represented by fused image. This fused image gives more detailed information captured to gray scale image which makes it suitable for human vision and can be applied to many practical scenarios.

### *C.* Biological Recognition

Due to the growing demands of security related application, face recognition is used in many applications. There have been many practical success stories with respect to face recognition technologies. But, the rate at which the visual face gets recognized depends upon the illumination conditions or skin tone of the subject[16]. In thermal infrared based recognition methods, it is found to have good performance since there is no need for appropriate illumination conditions or for finding camouflaged faces[16].

In reality, the lighting conditions(illumination), the face expression , focus angle, foreground and background are factors responsible for accurate face recognition. Recently, the bionic optimization algorithm based face recognition is the most used technique for biological recognition wherein it improves the accuracy of recognition. The complementary information given by infrared and visual image fusion technique for recognition is more and it is improving[30,31].

### *D.* Other areas of Application

Apart from the mentioned applications there are many more specific areas where visual and infrared image fusion is applied. Used in agriculture, non-destructive testing, object detection, remote sensing, military and so on. Object detection got the attention of researchers since the fused image provided better results of detailed complementary image information. Nizar et al. proposed a novel object detection methodology for fruit detection[11]. Researchers are also working on capturing multiple features from different modalities for pedestrian detection applications[10]. Talha et al. proposed an adaptive fusion strategy using particle filter for tracking the hidden targets[10]. Remote sensing based applications use infrared and visual imaging fusion technology where geostationary objects can be identified. Infrared and visual imagery was employed in agricultural area to predict the firmness of fruits and to determine the soluble solids content. A study using silhouette extraction performed in the wavelet domain can find out diseased plants.

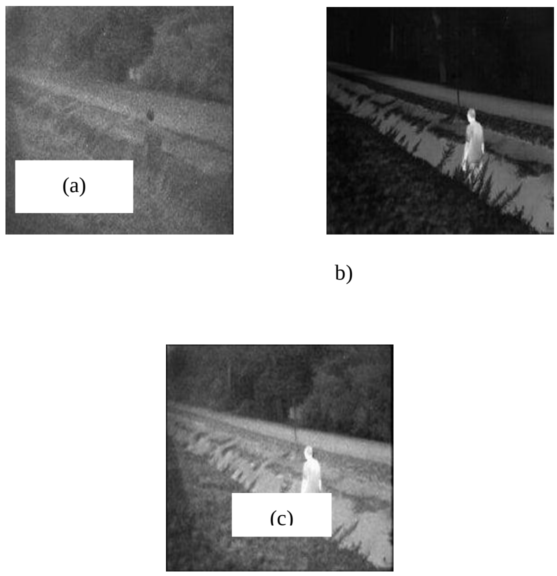



**Fig.3. (a) Visual Image**

**(b) Infrared Image**

**(c) Fused Image**


Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

Image fusion framework is also used in road obstacle classification systems, construction and so on. In addition to these applications, many new demands are put into place for various other applications for which new image representation methods are emerging. With these improvements in the field of imaging, we feel that it would lead to the development of fusion of images with multiple modalities.

## III. EXISTING MULTIMODAL FUSION METHODOLOGIES

| S.No | Image Fusion Algorithm | Advantages |
|---|---|---|
| 1 | Optimal region based fusion in spatial domain. | Optimized sharper images are obtained. |
| 2 | Discrete wavelet transform. | Increases the visual effect and preserves spectral details of source images. |
| 3 | Joint bilateral filter and wavelet thresholding. | Removes noise and preserves textures, edges and curves. |
| 4 | Discrete cosine transform. | Improves the quality of output image. |
| 5 | Statistical property of neighbourhood. | Reduces loss of information and noise. |
| 6 | PSO. | Enhances features like edge and contrast and also reduces speckle. |
| 7 | Curvelet transform. | Preserves spectral characteristics. |
| 8 | Orthogonal transform. | Captured high frequency content such as edges and slant texture. |
| 9 | Variance & spatial frequency based fusion. | Increases the perception of scene but suffers side effects like blurring artefacts and reduces quality. |

**Table 1 shows the commonly used fusion strategies** with their advantages. Generally, the intricate details and features of visual and infrared images will be interpreted by a few image representation approaches which gives prospects to fuse and combine more details of different source images.

### *A.* Multiscale Analysis based Image Fusion

Recently, the development of theories related to multiscale analysis has increased which lead to the increase in effectiveness of image analysis tools which in turn makes multi-scale analysis based fusion methodologies more diverse. Firstly, in multiscale based techniques, a decomposition algorithm(image transform) is used to obtain different frequency coefficients of source images, the coefficients are recombined using certain fusion rules and finally inverse transform is applied to get the final fused image.

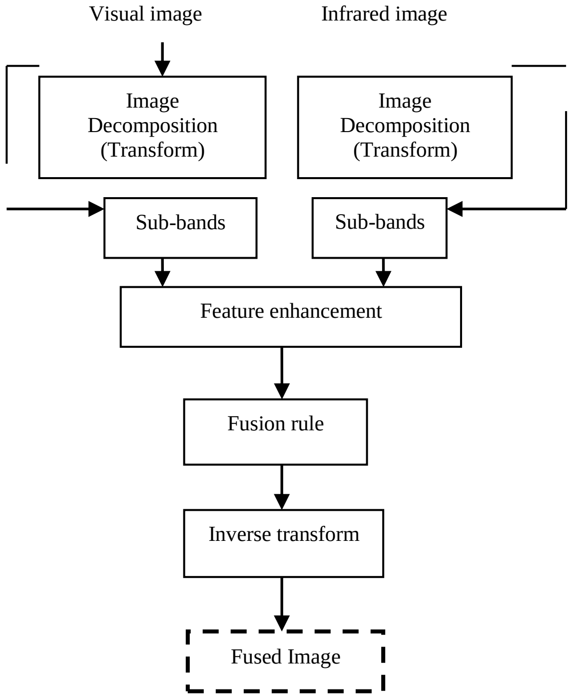



**Fig.4. Image fusion Process**

### *B.* Pyramid Transform based Image Fusion

Different pyramid methods were applied to source images in order to decompose them into frequency sub-bands such as laplacian pyramid, steerable pyramid and contrast pyramid[16]. The pyramid transform algorithm is used for performing image decomposition in order to derive intricate information which increases the fusion quality. Steerable pyramid is a multi-orientation decomposition technique which when applied to source images divides them into different sub-bands, with scale and orientation localized . Laplacian pyramid is a dynamic recursive fusion framework to reconstruct the coefficients[27]. It is also used in false color image fusion and fruit detection. The main benefit of using pyramid transform based image fusion of different modalities is its low computational complexity and memory space. Like every other fusion algorithm even this algorithm has a setback; where it smoothes some part of the image content and affects clarity[3] which causes blocking effect since it fails to integrate spatial information. Therefore, Pyramid transform cannot achieve optimal fusion quality. But when these algorithms are used together,such discrepancies and be avoided

### *C.* Compressive Sensing & Sparse Representation based Image Fusion

The main advantage of compressive sensing and sparse representation are that they tend to minimize the computational complexity thereby increasing the processing speed. Presently, many methods are implemented by researchers in fusing visual and infrared images where accurate image recovery, sparse coefficient representation and valid compression of image still need more study.


Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

Compressive sensing theory can efficiently minimize the computational complexity & maximize the image processing operating schemes since it has very less signal sampling. Liu et al. adopted DWT to obtain sparse coefficients of multimodal images in which the low frequency components are merged using weighted averaging fusion rule and high frequency components are merged by maximum selection fusion rule.

Sparse representation was another such area which received huge attention among researchers in image and signal processing. Sparse representation takes few image details to represent the source descriptor. These sparse coefficients are nothing but image signals or in other terms most salient features of input images. Since sparse representation has many such inherent pros, this scheme is widely used in fusion of multimodal images. Many such sparse representation based fusion is mentioned in[26].

## IV. QUANTITATIVE ASSESMENT OF RESULTANT FUSED IMAGE

The quality of fused visual and infrared images is evaluated based on certain image assessment metrics which is being incorporated by many researchers to quantify their results. The evaluation methods are broadly classified into two groups which are objective evaluation metrics and subjective evaluation metrics.

### A. *Subjective Evaluation Schemes*

In this scheme, the quality of resultant fused image is determined or evaluated based on the perspective of human vision which is a direct quality evaluation method. This is due to the fact that finally the interpreter or observer of the resultant image is human. The criteria that are involved in subjective evaluation method are edge definition, sharpness, distortion in image, noise, image details, features and so on.

### B. *Objective Evaluation Schemes*

This is done through mathematical calculation using some computable formulas to evaluate fusion quality. Different metrics cover different aspects which gives scope for performing different evaluation methodologies to be adopted concurrently.
The most used metrics for evaluation is listed below.

- *Information Entropy*
  It defines the amount of information content transferred to the fused image from the source images. Mathematically, it is defined as,

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

  where $P_i$ is the gray level probability i in image and L ranges from 0 to 255 which defines the gray level of image.
- *Spatial Frequency(SF)*
  It defines clarity of details in an image.

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

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

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

  where M is the row and N is the column and F(i,j) is the gray level of fused image F at Position (i,j) , RF is the row frequency and CF is the column frequency.. Larger SF implies well defined edges and textures.
- *Mutual Information(MI)*
  It gives the amount of details of input image which is fused in the resultant image. If MI value is more, then more image content is fused from input images to resultant image.
  MI is defined by ,

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

- *Edge based Similarity Measure($Q^{ABF}$)*
  The amount of edge details preserved in the final image is given by this metric.
  It is denoted by $Q^{ABF}$.
  If this value is large, more edge details are transferred to final image which shows better fusion result.

## V. ESTIMATED OBSERVATIONS FROM TYPICAL METHODS

Some typical fusion techniques are shown which are laplacian , gradient and shift invariant discrete wavelet transform. So, to determine the efficiency of the fusion algorithm, some evaluation metrics are used.

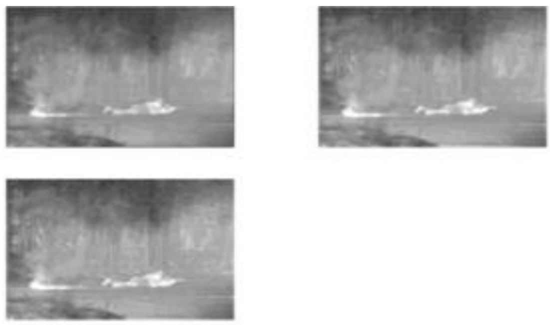


**Fig.4. Image fusion using (a)Lap (b) Grad (c) SIDWT**

| Method | MI | Entropy | SF | $Q^{ABF}$ |
|---|---|---|---|---|
| Laplacian | 3.3934 | 7.2853 | 22.9121 | 0.6237 |
| Gradient | 2.6013 | 7.1235 | 22.6301 | 0.5265 |
| SIDWT | 2.4492 | 6.8512 | 13.6871 | 0.4091 |

## VI. CONCLUSION

Image fusion using multiple modalities is an emerging research area and it is rapidly developing to achieve perfection. But there is a need for more improvement in potential streams in varied applications. In this paper,

we have summarized the techniques of visual and infrared image fusion , major applications and fusion technologies. More specific and effective fusion schemes will be needed with the rapid improvement in technologies. There are some potential directions of research which can guarantee better fusion results such as anti noise parameter, complexity reduction which in turn increases human visual perception. This paper concludes with various challenges and other applications with demands that need further improvement.


Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

## REFERENCES


1. Zhang Bahua, Zhao Ying ," A fusion algorithm for infrared and visual based saliency analysis and non-sub sampled shearlet transform", Infrared Phys. Tech. 73(2015).
2. Zhadong Liu, Yi Chai , "A novel fusion scheme for visual and infrared images based on compressive sensing", Optics Comm. 335(2015).
3. Xudong Kang, Haitao Yin , "Pixel level image fusion: a survey of the state of the art", Information fusion 33(2017).
4. Huajun Feng, Qi Li "Detailed preserved fusion of visible and infrared images using regional saliency extraction and multiscale image decomposition", Optical Comm 341(2015).
5. Xinan Fan, Min Li ,"A thermal infrared and visible image fusion based approach for multitarget detection under complex environment", Math. Problem Engg(2015).
6. Bhatnagar, Liu, "A novel image fusion framework for night-vision navigation and surveillance.9(2015).
7. Alex James, Belur ,"Medical image fusion: a survey of the state of the art", Inform. fusion 19(2014).
8. Amanda Muler, Narayan," Cognitively engineered multisensor image fusion for military applications,Inform, fusion 10(2009).
9. X.Zhang, L.Yuan, "Infrared target detection and location for visual surveillance using fusion scheme of visible and infrared images", Math problems engg .3(2013).
10. Tarek, Nizar, "Finite asymmetric generalized gaussian mixture models learning for infrared object detection", Comp.Vis.Image Underst.117(2013).
11. V.Alcanatis, T.Burks, "Image fusion of visible and thermal images for fruit detection",Biosys.Eng.103(2009).
12. C.Homa, W.Arnold, "Fusion of visual and infrared thermography images for advanced assessment in non destructive testing",Rev.Sci.Inst 84(2013).
13. [.Gonzalez, Z.Fang, "Pedestriam detection at day/night time with visible and FIR cameras: a comparison", Sensors 16(2016).
14. D.Dixon, J.Lewis, "Task based scan path assessment of multisensor video fusion in complex scenarios",Inform .fusion.11(2010).
15. Mohammad Talha, Rustam Stokin, "Particle filter tracking of camouflaged targets by adaptive fusion of thermal and visible spectra camera data",IEEE sensors.14(2014).
16. Seong Kong, Jing Heo, "Recent advances in visual and infrared face recognition-a review",Comput.Vis.Image Underst.97(2005).
17. George Bebi, Aglika Gyaourov ," Face recognition by fusing thermal infrared and visible imagery",Image Vis.Comput.24(2006).
18. Hermosilla, G.Faras, "Fusion of visible and thermal descriptors using GA for face recognition", Sensors 15(2015).
19. [R.Ragavendran , Ashok Rao," Pso based fusion of near infrared and visible images for improved face verification",Pattern regn.44(2011).
20. Buzha Wuuile, Yilng Yang,"Direct fusion of geostationary meteorological satellite visible and infrared images based on thermal physical properties",Sensors(2015).
21. F. Liu, L.Jiao, "Multi contourlet based adaptive fusion of infrared and visinle remote sensing images",IEEE geosci.Remote sensing Letters 7(2010).
22. X.U.Lang, D.J.L.Quing, "Visible and infrared video fusion using unifrom curvelt and saptio temporal informatiom(2015).
23. Fuyan Xu, Kan Ren,"Super resolution images fusion via compressed sensing and low rank matrix decomposition,Infrared tech.68(2015).
24. [Gema Piella ," A general framework for multi resolution image fusion from pixels to regions,"Inform fusion 4 (2003).
25. Y.Yan, J.Zhang,"Objective quality evaluation of visible and infrared color image fusion",Opt.Engg.50(2011).
26. Jinping Fan, Xiayoi Feng,"Fusion method for infrared and visible images by using non-negative sparse representation",Inf. Phys.Tech.67(2014).
27. Qan Jiang, Rencan Nie ,Xin Jin,"A survey of infrared and visual image fusion methods",Infrared Phy Tech.85(2017).
28. L Latha, S Thangasamy," On improving the Performance of Multimodal biometric authentication through Ant colony optimization", Wseas Transactions on Information Science and Applications, Issue 12, Volume 8, December 2011.
29. L Latha, M Pabitha, S Thangasamy," Effectual human authentication for critical security applications using retinal images", ICTACT Journal on Image and Video Processing, 2010.
30. L Latha, S Thangasamy, "Providing multimodal biometric authentication using five competent traits" , The Imaging Science Journal, 2013.
31. L Latha, S Thangasamy ," Robust Way of Multimodal Biometric Score Normalization",Journal of Applied Security Research, 2012.