Survey on image fusion techniques used in remote sensing

Kalyani S. Ulabhaje
ulabhajekalyani@gmail.com
G.H. Raisoni College of Engineering,
Nagpur, Maharashtra

Sujata More
sujatamore123@gmail.com
G.H. Raisoni College of Engineering,
Nagpur, Maharashtra

ABSTRACT

Image fusion in remote sensing has emerged as a sought-after protocol because it has proven beneficial in many areas, especially in studies of agriculture, environment, and related fields. Simply put, image fusion involves garnering all pivotal data from many images and then merging them in fewer images, ideally into a solitary image. This is because this one fused image packs all the pertinent information and is more correct than any picture extracted from one solitary source. It also includes all the data that is required. Additional advantages are: it lessens the amount of data and it creates images that are appropriate and that can be understood by humans and machines. This paper reviews the three image fusion processing levels, which include feature level, decision level, and pixel level. This paper also dwells upon image fusion methods that fall under four classes: MRA, CS, model-based solutions and hybrid and shows how each class has some distinct advantages as well as drawbacks.

Keywords—Image fusion, Remote sensing, High resolution images, PAN images, MS images

1. INTRODUCTION

Remote sensing means securing data about an object from calculations performed at a distance away from the object and without the sensor really having any contact with it. Remote sensing, essentially, is employed for the observation and study of our planet. Remote sensing obtains data pertaining to the earth’s surface structure and what is on its contents, through gathering and analysing spectral calculations performed at a distance. This is the chief aim of remote sensing. The earth emits scattered/reflected or self-produced electromagnetic energy in varying wavelengths bands. This forms the basis for the recording. Remotely sensed data offers information in spatial, spectral, temporal and radiometric areas. Each area contains a resolution that is pertinent to the information collected [1][2][3]. In these areas, different sensors carry varied resolutions [4][5][6]. Also, in different parts of the electromagnetic spectrum, pixels are evident in remote sensing systems. Hence, remotely sensed images contain varying spatial, spectral and temporal resolutions. The finer the resolution, the lesser is the total ground area that is observed.

There are high multispectral sensors (MS), which have a lower spatial resolution. And there are panchromatic sensors (PAN), which have a higher spatial resolution and a wide spectral bandwidth. However, there exists an exciting process in which these varied images can be merged. With the appropriate blending of the data obtained, we can produce an image with the best characteristics of both, namely high spatial and high spectral resolution. This process, a sort of multi-sensor data fusion, is known as pan-sharpening of MS image. Multi-sensor image fusion techniques combine two or more geometrically registered images of the same scene into a single image. The fused images may provide enhanced interpretation possibilities and more dependable outcomes [7]. Many commercial optical satellites can deliver MS and PAN images. Among the satellites are Landsat, IKONOS, Quick Bird, SPOT, Pleiades, IRS, Leica ADS40, GeoEye, OrbView and WorldView.

For the blending of images, there are laid-down protocols. W. Wenbo, Y. Jing et al. [8] in their work have examined and studied the merging of remote sensing images and its application in image classification. Depending on the stage at which the fusion of data is accomplished, the merging techniques are executed at three different levels of processing: feature level, decision level, and pixel level. This survey paper highlights these three sophisticated levels of fusion techniques: in section 2 the feature level, in section 3 the decision level and in section 4 the pixel level.

2. IMAGE FUSION METHODS

Feature level, Decision level, and Pixel level fusion are the main three types of fusion methods used in remote sensing Descriptive study of all these three methods are as follows.

3. FEATURE LEVEL FUSION METHOD

In this method, the process of culling the features happens and then, employing the advanced technique, the fusion procedure is accomplished. For instance, in feature level fusion, it is imperative to detach objects through the process of parceling them down in parts. Features are akin to attributes, like shape, boundaries, nearby areas of the observed environment; they are taken out from the original first obtained images. Objects from varying sources that are ditto in nature are linked to each other and then merged, for further determination. As compared to methods of pixel level, information is processed at much higher stages in methods of feature level fusion.
There are some fusion methods at feature level that are listed below:

3.1 Retina based multi-resolution
The human retina works much more than an image-capturing camera. It transforms the optical data into electrical signals. Moreover, it also processes visual signals before sending them to levels that are higher in the visual system. Taking a cue from this, image fusion can use the processing tenets of the human optical system. To this very effect, H. Ghassemian et al.[9] have advanced a method of fusion of multi-resolution data based on retinal visual channels decomposition. It is inspired by the analytical outcomes stemming from the analysis of retina based image: the energy-filled spectral features are scattered in the lower frequency sub bands and the spatial features are disbursted in the higher frequency sub bands. The high-scale spatial attributes that are culled from the PAN image, are appended to the low-scale spatial characteristics from thematic mapper (TM) images. Consequently, the MS images are improved by the procedure of visual channels. H. Ghassemian et al.[9] have exhibited a computer retina inspired model.

It is not essential to resample the image in the retina related fusion method. This becomes a benefit when compared to other methods, such as Brovey, wavelet, PCA, and IHS. Between PAN image and pixels of MS images, it is able to function in any aspect ratio. In any image blending method, resampling tends to weaken the spectral attributes of the MS image. Therefore, the resampling process should be shunned.

3.2 Softmax regression-based feature fusion
X. Bai et al. [10] have propounded a SoftMax regression-based feature fusion method by knowing singular weights for different attributes. This method involves gauging the object-to-class sameness steps and the conditional probabilities that each object belongs to different classes. Additionally, a relative method is used to estimate the class-to-class sameness between various classes. And at the end, utilising the secured fusion and sameness data, a vector machine classifier is constructed.

3.3 Multirate filter banks
This method relies on multirate filter banks image synthesizer: the energy-filled spectral attributes are scattered in the lower frequency sub bands, and the spatial characteristics, edges, are disbursted in the higher frequency sub bands [11].

3.4 A blend of shape, texture and spectral data
To enhance the accuracy of classification, F. Mirzapour et al. [12] in their work employed a blend of shape features and texture from HS images, as also spectral data.

3.5 Fusing hierarchical broken down outcomes into MRF
M. Golipour et al.[13] propounded a new method based on spectral-spatial classification for HS images. This involves infusing first in the Bayesian system the hierarchical segmentation outcomes into Markov random field spatial.

3.6 Learning-based superresolution fusion
H. Song et al.[14] have put forth a learning-propped up super resolution fusion method. It merges the swath width and spectral elements of Landsat TM/Enhanced TM Plus (ETM+) and the spatial resolution of SPOT5.

4. DECISION LEVEL FUSION METHODS
Decision fusion (or clarification level) is the most elevated handling level. It is the way toward blending data from diverse individual information sources after every information source has been put to initial categorisation. In the choice level combination, the obtained outcomes from various local classifiers are joined to decide the ultimate choice. In short, the input pictures are sorted and processed singly to cull data. Thereafter, the decision guidelines are utilized to consolidate removed data to fortify basic elucidation and resolve contrasts and deliver a superior reading of the scrutinised objects. The input choices are represented as symbols or labels with various degrees of certainty.

Some fusion methods at decision level are listed below:
In the systems of parallel and distributed processing decision level fusion enjoys incredible uses. In diverse applications like verification of fingerprints and classification of remote sensing images, fusion methods are widely employed. S. Helfroush et al. [15] exhibited the concept of decision level blending with respect to texture, orientation and spectral attributes of the fingerprint image.

J. Atl Benediktsson et al.[16] have considered hybrid classification methods based on unanimous findings derived from numerous sources of information. Their focus dwells on the neural and statistical network methods and the blend of those methodologies. At the outset, each source of information is handled independently and displayed employing statistical approaches. Thereafter, weighting systems are utilised to control the impact of every datum source in the joined classification. The weights are streamlined to enhance the correctness of the joined classification.

In several approaches like voting [17], rank-based, Bayesian inference, and Dempster-Shafer methods [18][19], decision fusion methods can be used gainfully.

For instance, voting methods face a major problem: they are good enough for local classification outcomes for the class of local winners in its delineated pixel. This leads to a plethora of inaccuracies of decision fusion outcomes for the secured information that is class correlated. While the method based on rank focuses more on information than on voting approaches. It tends to utilise the outcomes of local classification for a delineated pixel but in all the classes. As a matter of fact, however, here the outcomes of local classification should incorporate the rank or values of the classification measure. This leads to a massive surge in the volume of information of local classifiers’ outputs, the receiving decision fusion station, and system of communication between fusion station and local classifier.

On the other hand, the Bayesian method does not take into account the factor of unpredictability and might display inaccuracies and multiplicity in the calculations of probabilities. The Bayesian method’s estimation of probabilities can turn out to be precarious in the event when the number of known motions or propositions is less than the number of unknown ones. In contrast, though, the Dempster-Shafer method can handle unpredictability and surmount some of the above faults because it can be employed without prior distributions of probability.

In their work, A. J. Rashidi et al.[20] advanced the combined measures method as a potent tool at the decision level to produce a high achieving design of multi-sensor image fusion. In this method, the dispersed multiple sensors deliver the images in diverse spectral bands such as microwave, infrared, visible and thermal. Then, they cull the mathematical elements of the outcomes of multi-sensor local classification and utilise them for the display of the classifier conduct through two steps:
believability and exactitude. Then they set up the believability and exactitude distribution vectors matrices for initiating two enhancements of the Dempster-Shafer method. Thereafter, utilising these two factors they put forth the combined measures decision fusion method. This combined measures approach is able to overcome any decision fusion faults with respect to the outcomes of unpredictable local classifiers as also clear local classifiers.

M. Fauvel et al.[21] in their work advanced a method for blending of multiple classifiers, which offers correlative or superfluous outcomes. In this, at the outset, each classifier processes information independently, and membership degrees for the concerned classes are offered, for each pixel. Next, the outcomes offered by the algorithms as per the capabilities of the classifiers are combined utilising a fuzzy decision rule. Then, what is advanced is a general system, relying on the definition of two measures of exactitude, for joining data from numerous independent classifiers in multiclass classification. The first measure calculates, for each pixel, the dependability of the data given by each classifier. And the second measure evaluates the global correctness of each classifier. Finally, the outcomes are combined with a flexible fuzzy operator governed by these two measures.

In their work, F. Tabib Mahmoudi et al.[22] employed, in an object recognition scheme, decision level fusion in multi-views imagery. This method incorporates two steps: single view and multi-views. Firstly, on single, individual views, an object related evaluation is done. Secondly, through a decision level blending based on the data of the scene context, the classified objects of all views are merged together. This is meant to refine the classification outcomes.

5. PIXEL LEVEL FUSION METHODS

In layman’s terms, image fusion means collecting all crucial data from many images and then merging them in fewer images, preferably a single image. Why? Because this one fused image contains all the relevant information and is more correct than any image derived from a single source, and also has all the needed information. Image fusion has a two-fold purpose: to lessen the amount of information and to create images that are more apt and fathomable for the perception of humans and machines.

In specialized terms, it implies that multisensory image fusion consolidates at least two geometrically enlisted pictures of a similar scene into a solitary picture that is more effectively deciphered than any of the firsts. The point is to get data of more noteworthy quality where the correct meaning of more prominent quality will rely on the applications [23]. It could mean change in additional handling [24], better visual appearance [25], or improvement in the visual translation of information [26].

The blend of MS pictures at a higher spatial resolution through the use of a PAN picture is generally called pan-sharpening of MS picture. An MS image can have up to 8 bands of spectral resolution, secured in close-infrared and wavelengths that can be seen. Whereas a PAN image can have a spatial resolution that is less than half a meter. So, by merging MS and PAN pictures, images with both high spatial and spectral resolutions can be obtained. The strategy utilized for the merging of individual sources must consider the physical properties of every methodology. At the pixel level, the algorithms that achieve fusion are categorized into 4 classes. They are multiresolution analysis (MRA), hybrid methods (a combination of CS and MRA), model-based algorithms and component substitution (CS).

There are 3 parameters that the fusion process must fulfill to achieve a properly fused image. The process must first conserve all important data; second, remove all unnecessary data and noise, and third, curtail discrepancies and by-products and artifacts. The picture’s quality can be affected by noise generated by image sensors. In their work, W. Zhao et al. [27] advanced a scheme for the merging of multi-spectral image and removing noise in the gradient area. The gradients can mirror the alterations of local details of pictures.

Before merging the images, in order to regulate the images, an algorithm pertaining to registration is required to be put in place [28]. In short, all the pictures being handled have to be referenced geographically and co-registered. MS images with low spatial resolution have to be converted into new, fresh images with ditto resolution being applied to PAN pictures. In the process of fusion, if the errors with respect to registration are not attended to, that can adversely impact the quality of the merging. In their work, P. Cheng et al. [29] have cited methods for correction of radiometric and geometric aspects of remote sensing information.

There is a process for conversion of the MS image. The conversion is achieved through the use of symmetric digital filtering kernels that are of uneven lengths. Employing cubic or linear functions of interpolations, these kernels use local polynomials in individual pieces. Then, because their ratio of scale is formed out of an even number, both commercial PAN and MS image products are moved by half pixels bearing an odd number. Finally, filters of even lengths can be applied, in order to balance the movements of half pixels between the PAN and MS sampling grids.

5.1 MRA Methods

Of late, image fusion process has been successfully integrating approaches based on multi-scale decomposition (MD) in varied applications like merging of the hyperspectral image [30]. For instance, diverse MD approaches like discrete wavelet transform and pyramid transform have been incorporated in image blending. In image merging based on MD approaches there are usually the following steps: firstly, through the use of wavelet or pyramid transform, the initial images are made to decompose into myriad scale levels. Then secondly, at every level of the initial images, the merging process is executed. And thirdly, the fused image is obtained through the inversion of the transform. It is true that the utilization of the transform makes the computational process more complex. But the merging approach based on MD definitely offers frequency and spatial area localization and accomplishes far better outcomes. Thus, according to the applications that are in play, the decision to opt or not to utilize transforms can be taken. In the merging methods based on MD, the fundamental fusion rule gets applied to MD representations of pictures at each resolution stage. On the other hand, though, in fusion methods not based on MD, the fundamental fusion rule gets applied to the initially sourced images directly. J. Zhou, J. Nunez et al. [31] have cited a quantitative comparison involving both spatial and spectral attributes, to assess the wavelet transform and other conventional algorithms. Meanwhile, M.V. Joshi et al. [32] have presented a wavelet à trous ("with holes") algorithm.

For the CS approaches, interpolation is more important than in the MRA approaches. This is because the later stage in MRA
might reposition the data details culled from PAN if linear non-zero phase filters are employed [33]. G. Vivone et al. [34] have stated that an injection scheme can be used to model both the MRA and CS approaches.

Majority of the approaches based on multiresolution analysis utilize a variety of transforms, such as contourlet, curvelet, and wavelet. Similarly, several transforms are used in MRA to cull spatial details. Among them are; curvelet [35][36], contourlet [37][38], Laplacian Pyramid [39], undecimated wavelet transform [40], and decimated wavelet transform [41]. Spatial resolution and picture size shrink from one level to the next in pyramid methodology. On the other hand, spatial resolution lessens from one level to another in the wavelet à trous algorithm, but the picture size stays the same at all levels.

For 2-D objects, a tensor product of two 1-D wavelet transforms is utilized. This product corrects the vertical and horizontal irregularities in 1-D. Also, the wavelet transform can represent objects having 1-D irregularities. By employing a rule of parabolic scaling contourlet and curvelet transform can correct 2-D irregularities around smooth curves. Hence, J. Xu et al.[42] suggested the ripplet transform, which injects the parameters of degree and support, which can display irregularities around shaped curves. The ripplet transform has another use, as a potent tool for culling spatial data from PAN pictures that are high spatial and low spectral in attributes.

5.2 Model-Based Methods
Q. Wei et al.[43] in their work propounded an image fusion approach involving a fast-multi-band, which comprise the probabilities found in the recordings and observations. In this fast algorithm, it is suggested, augmenting the probabilities can end in a solution for the Sylvester equation. The solution is secured by using the circulant’s properties and down sampling the patterns linked with the problem of fusion. This approach can be made general by including advanced information for the fusion problem, permitting a Bayesian calculator.

H. Ghassemian et al.[44] have put forth a stratified Bayesian model to merge several multi-band pictures having varied spatial and spectral resolutions. This involves the introduction of an apt advanced distribution and sampling of posterior distribution with Hamiltonian Monte Carlo algorithm.

To achieve image fusion, M. Guo et al.[45] propounded an online coupled dictionary learning (OCDL) scheme. In this, the OCDL fully utilizes the accessible lower spatial resolution of MS picture and high spatial resolution of PAN picture to lessen the spectral misrepresentation and conserves the spatial data of the MS picture. In the OCDL approach, a superimposition method is interjected to create two transitional pictures for the creation of the coupled dictionary for every band. The coupled dictionaries are updated through a method of continual updating. The coupled dictionaries can be labeled as the online process of dictionary learning.

5.3 CS Methods
The CS related methods involve beaming an MS image into a different space. This is done by utilizing a transform that detaches the spatial format from the spectral data in varying parts. Thereafter, the segment containing the spectral format is substituted with the PAN picture. It is found that if the correlation is higher between the PAN image and the substituted segment, it creates less misrepresentation. Therefore, it necessitates execution of matching of histogram between the PAN picture and the chosen segment prior to its substitution. The PAN image that is histogram-conformed should have the same variance and mean of the substituted segment. In the end, the data is brought back to the original space by using the inverse transformation. Thus, is finished the process of pan sharpening.

5.4 Hybrid Methods
Hybrid methods, combining the elements of MRA and CS approaches, reap the benefits of both. In their study, F. Chen et al.[46] put forth an enhanced merging method that employs a wavelet decomposition to cull comprehensive data of PAN images. Meanwhile, M. Ghahremani et al.[47] have propounded an ICA and curvelet related image blending approach. Additionally, a remote sensing image merging approach that utilizes a combination of curvelet transform and IHS has been advanced by A. Valizadeh et al.[48]. Other image fusion strategies are based on various wavelet-related pan sharpening approaches. For instance, in these methods, from the PAN image with high spatial resolution, high frequency detail coefficients are procured, which are then joined with spectral data collected from the MS picture using a combination methodology. Y. Luo et al.[49] have produced a technique involving PCA transformation and additive wavelet decomposition.

6. CONCLUSION
Finally, image fusion techniques in remote sensing have garnered widespread acceptance. The majority of satellites deliver both PAN images having high spatial resolution and MS images having a lower spatial resolution. And scores of applications that rely on remote sensing have to have high spatial as well as high spectral resolutions. Therefore, some potent image fusion techniques come in very handy in creating images that deliver the needed specifications. One properly fused image contains all the relevant information and is more correct than any image derived from a single source, and also has all the needed information. Image fusion has a two-fold purpose: to lessen the amount of information and to create images that are more apt and fathomable for the perception of humans and machines. This paper has briefly examined the three image fusing processing levels, which comprise, feature level, decision level, and pixel level. Their methods fall under four varied classes: MRA, CS, model-based solutions and hybrid. Each class has some distinct advantages and drawbacks. For instance, MRA approaches tend to deliver higher spatial misrepresentations but vastly better spectral regularity. In contrast, CS approaches, though fast and easy to put in place, might create substantial spectral distortions. And hybrid approaches are utilized to get the best of both the CS and MRA methods.

7. REFERENCES


