1. **INTRODUCTION**

Image fusion is process of combining multiple input images into a single output image which contains better description of the scene than the one provided by any of the individual input images. The fused image provides detailed information about the scene which is more useful for human vision perception and machine perception or further image-processing tasks such as segmentation, feature extraction and object recognition. A feature closely related to image quality is focus. Sharp images provide better information than blurry images. However, in some situations it is not possible to obtain totally focused images in just one single camera shot, since some regions appear to be blurred due to variations in the depth of the scene and of the camera lenses focus. This means that if the camera is focused at one specific object, another region of the scene can be out of focus. An interesting solution is to take more pictures of the desired landscape in the same position, but with focus centered in different elements of the scenery. Then, using the image fusion concept, all source images are combined, creating a single image that contains all the best focused regions. Image fusion is becoming very popular in digital image processing.

Image fusion may cover many diverse scenarios and purposes ranging from characterization of constituents in 3D hyper spectral images or in sets of related 2D images, process modeling, quantitative analysis, or simply taking advantage of the complementary spectroscopic information of different spectroscopic platforms. Image fusion may cover many diverse scenarios and purposes ranging from characterization of constituents in 3D hyper spectral images or in sets of related 2D images, process modeling, quantitative analysis, or simply taking advantage of the complementary spectroscopic information of different spectroscopic platforms. In satellite imaging, two types of images are available. The panchromatic image acquired by satellites is transmitted with the maximum resolution available and the multispectral data are transmitted with coarser resolution. This will usually be two or four times lower. At the receiver station, the panchromatic image is merged with the multispectral data to convey more information. This merging process is known as image fusion.

* 1. **Objectives**

The purpose of image fusion is not only to reduce the amount of data but also to construct images that are more appropriate and understandable for the human and machine perception. Multi sensor data fusion has become a discipline which demands more general formal solutions to a number of application cases. Several situations in image processing require both high spatial and high spectral information in a single image. This is important in remote sensing. However, the instruments are not capable of providing such information either by design or because of observational constraints.

Panchromatic images are the images collected in the broad visual wavelength range but rendered in black and white and Multispectral images are the images optically acquired in more than one spectral or wavelength interval. Each individual image is usually of the same physical area and scale but of a different spectral band. The purpose of the project is to obtain an image where the spatial resolution of a hyper spectral image and the spectral resolution of a multispectral image are featured into a single image thus presenting more data in a single image rather than preserving multiple images to serve the same purpose.

Therefore, the objectives of a good image fusion technique must include:

1. Importing all the necessary toolboxes to read the images.

2. Accurate transformation of input images.

3. Efficient algorithm to carry out the fusion process.

4. Minimal mean square error and a high correlation coefficient.

5. Overall redundancy reduction in the dataset size.

* 1. **Methodology**

The overall flow of the development process follows the order of input image data collection, transforming the images into numerical transformations, usage of appropriate fusion rules to combine the essential features of the input images ,and finally to re-transform the images into a perceptible one. Comparative analysis of image fusion methods demonstrates that different metrics support different user needs, sensitive to different image fusion methods, and need to be tailored to the application. Quality of the fused image depends on the application. Various quality measures have been discussed to perform quantitative comparison of these methods.

Input Images

Final Output Image

Inverted Transform

Image Fusion

Application of various transforms

Figure 1.1 Overall Methodology for fusion process

* 1. **Organization of the project**

The chapter one of the project deals with introduction of growing importance and need for filtering out the highly voluminous datasets of images required for remote sensing and other scientific applications in order to yield better, simpler, accurate and efficient computational results and also images that are more appropriate and understandable for the human and machine perception.

The second chapter contains theoretical analysis of the project which elaborates about the existing systems, their shortcomings or disadvantages and how the proposed solution is advantageous and more efficient in carrying out the process of image fusion.

The chapter three contains the architecture of the project, the underlying algorithms and their working, different modules and their description along with the detailed algorithm used in developing the model.

The chapter four contains details with the description of the technology used, advantages of the technology being implemented and different Unified Model Language(UML) diagrams like use case diagram, activity diagram, State chart diagrams.

The chapter five has discussion on the key aspects of the implementation, the results obtained and screenshots of the output and the detailed description and presentation of the performance comparisons.

The sixth chapter has conclusion and scope for the future which includes the application scope and domain breadth of the project how the project can be enhanced for the future. Conclusion is a brief description of the project and scope, consists of how it is useful and its scalability soon.

1. **ANALYSIS OF IMAGE FUSION TECHNIQUES**

**2.1 Existing Systems**

**Intensity Hue Saturation (IHV) Based Transformation**

HSL (hue, saturation, lightness) and HSV (hue, saturation, value) are alternative representations of the RGB color model, designed in the 1970s by computer graphics researchers to more closely align with the way human vision perceives color-making attributes. In these models, colors of each hue are arranged in a radial slice, around a central axis of neutral colors which ranges from black at the bottom to white at the top.

The HSV representation models the way paints of different colors mix together, with the saturation dimension resembling various tints of brightly colored paint, and the value dimension resembling the mixture of those paints with varying amounts of black or white paint. The HSL model attempts to resemble more perceptual color models such as the Natural Color System (NCS) or Munsell color system, placing fully saturated colors around a circle at a lightness value of ​1⁄2, where a lightness value of 0 or 1 is fully black or white, respectively**.** It is important, therefore, that the features of interest can be distinguished in the color dimensions used. Because the *R*, *G*, and *B* components of an object's color in a digital image are all correlated with the amount of light hitting the object, and therefore with each other, image descriptions in terms of those components make object discrimination difficult. Descriptions in terms of hue/lightness/chroma or hue/lightness/saturation are often more relevant.

Transformations like HSV or HSI were used as a compromise between effectiveness for segmentation and computational complexity. They can be thought of as similar in approach and intent to the neural processing used by human color vision, without agreeing in particulars: if the goal is object detection, roughly separating hue, lightness, and chroma or saturation is effective, but there is no particular reason to strictly mimic human color response.

**Principle Component Analysis (PCA) Based Image Fusion**

Principal Component Analysis is a quantitatively rigorous method for achieving simplification. Often, its operation can be thought of as revealing the internal structure of the data in a way which best explains the variance in the data. The method generates a new set of variables called 61 Principal Components (PC). Each principal component is a linear combination of the original variables and all the PCs are orthogonal to each other, and as a whole form an orthogonal basis for the space of the data; thereby removing redundant information, if any.

Principal component analysis (PCA) is a statistical procedure that uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components.

Principal component analysis is carried out which aims at reducing a large set of variables to a small set that still containing most of the information that was available in the large set. The technique of principal component analysis enables us to create and use a reduced set of variables, which are called principal factors. A reduced set is much easier to analyze and interpret. This transformation is defined in such a way that the first principal component has the largest possible variance (that is, accounts for as much of the variability in the data as possible), and each succeeding component in turn has the highest variance possible under the constraint that it is orthogonal to the preceding components. The resulting vectors (each being a linear combination of the variables and containing *n* observations) are an uncorrelated orthogonal basis set. PCA is sensitive to the relative scaling of the original variables.

The steps involved in the PCA transform are

1. Calculate the covariance matrix or the correlation matrix of the data sets to be transformed. The covariance matrix is used in the case of the unstandardized PCA, while the standardized PCA uses the correlation matrix.

2. Calculate the eigenvalues and the eigenvectors from the correlation matrix.

3. Principal Components of the given data set are the eigenvectors of the covariance matrix of the input.

The direction of the maximum variance is used to compute the first principal component. The second principal component is forced to be situated in the subspace vertical (perpendicular) of the first. Inside this subspace, this component points the direction of maximum variance. The third principal component is in the maximum variance direction in the Subspace vertical to the first two and so on. PCA is also known as Hotelling Transform or Karhunen-Loeve transforms. The fusion is accomplished by weighted average of images to be fused. Eigen vector related to the largest Eigen value of the covariance matrices of each source are used to obtain weights for each source image. It computes a compress and best description of the data set. The PCA basis vectors like FFT, DCT and wavelet are changing rapidly and its basis vectors depend on the data set. The most important advantage of PCA is, when its data size is compressed as well as if dimensions are altered then there is no much loss of information at the output image.

Fused Image

PCA Based weighted average fusion

PCA

MS Image

PAN Image

Figure 2.1 Principal Component Analysis based Image Fusion Process.

The PCA-based weighted fusion involves separately fusing the high frequency (HF) and the low frequency (LF) parts of an image. The two frequency components are obtained by a filtering mechanism and finally the fused components are added together to get the resultant fused output.

**2.2** **Disadvantages of Existing Systems for Fusion**

* Only a limited number of bands are involved (i.e, only three bands are involved).
* Sensitive to the area being sharpened and the produced fusion image may vary depending on the selected input image subset.
* Probability of the alteration of spectral information in the image is higher.
* Poor directional selectivity for diagonal features.
* The spectral degradation which reduces the contrast levels in the fused image and hence could affect negatively the classification results.
* The special quality can be degraded in the resulting image.
* Gray scale images are in focus in most of the existing work instead of color images.
* It has been observed that the problem of the uneven illuminate has not been taken into account in majority of existing techniques.
* The majority of the existing approaches rely upon transformations therefore some color artefacts may occur which may decrease the performance of the transform based vision fusion techniques.

**2.3 Advantages of the Proposed System**

Non Subsampled Contourlet Transform (NSCT) based image fusion overcomes all the above mentioned disadvantages and shortcomings to a great extent. Additionally, there are a few added advantages too. They are:

* The NSCT has a high image detail capture ability, especially for the line capture.
* This algorithm can supply binary representation of the line local scale, position and direction.
* It can easily realize precise positioning for line.
* The algorithm has low time and space complexity.
* Highly shift invariant.

**2.4 Performance measures of a Good Image Fusion Technique**

Certain quantitative measures are deployed in order to monitor the quality or the degree to which the features have been extracted and the level of pixel accuracy maintained. They are:

**Root Mean Square Error**

1. **IMPLEMENTATION OF THE IMAGE FUSION TECHNIQUES**
   1. **Discrete Wavelet Transform (DWT) based Image Fusion**

Wavelet Transform (WT) is a mathematical tool that converts original signal or image into different domains for analysis and processing. Recently, WT has gained substantial attention in many fields and applications such as engineering, physics, medical signal and medical image processing. Moreover, WT allow signal or image to be stored more efficiently than Fourier transform. Proposed scheme uses WT as preprocessing tool to remove the unwanted data. The (discrete) wavelet transform maps an image onto yet another basis, defined by a “special” matrix B. This transform:

•Captures scale,

•Is invertible, orthogonal and square

•Is image independent (not all my images have to be faces, or eyes).

Wavelet transforms are linear transforms whose basis functions are called wavelets. In all wavelet based image fusion schemes the wavelet transforms W of the two registered input images I1(x, y) and I2(x, y) are computed and these transforms are combined using some kind of fusion rule. Then the inverse wavelet transform W−1is computed and the fused image I(x, y) is reconstructed.

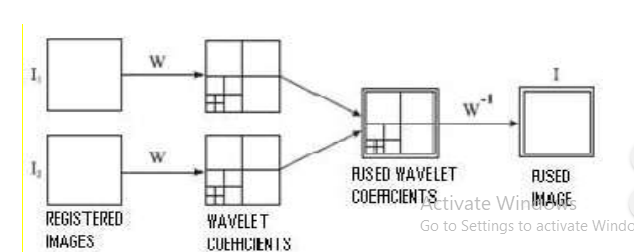
****

Figure 3.1.1 DWT Based Image Fusion

  The multi resolution image fusion technique based on wavelet transform is a method by which the wavelet coefficients of Panchromatic image is calculated in order to extract the spatial information and then it is combined with the corresponding coefficients of Multispectral image. The coefficients are combined using different fusion rule like substitution, maximum selection, averaging etc. Two aspects that determine the efficiency of image fusion are, firstly how effective the spatial information contained in the PAN image can be extracted and secondly, how effectively the extracted information can be injected into the original multispectral image without introducing spectral distortion, i.e. the effectiveness of fusion rule.

The general image fusion scheme of this study involves decomposition of the input images, calculation of the wavelet transform modulus maxima, fusion strategy, and reconstruction for new fusion image as shown. The basic idea of all multi resolution fusion schemes is motivated by the human visual system being primarily sensitive to local contrast changes, e.g. the edges or corners. In the case of wavelet transform fusion all respective wavelet coefficients from the input images are combined using the fusion rules. Since wavelet coefficients having large absolute values contain the information about the salient features of the images such as edges and lines, a good fusion rule is to take the maximum of the [absolute values of the] corresponding wavelet coefficients. The maximum absolute value within a window is used as an activity measure of the central pixel of the window. A binary decision map of the same size as the DWT is constructed to record the selection results based on a maximum selection rule. Rather than using a binary decision, the resulting coefficients are given by a weighted average based on the local activity levels in each of the images’ sub bands.

This transform decomposes the signal into mutually orthogonal set of wavelets, which is the main difference from the continuous wavelet transform (CWT), or its implementation for the discrete time series sometimes called discrete time continuous wavelet transform (DTCWT). It must be pointed out that the main idea of wavelet-based image fusion is motivated by the factor that the sub band images constitute the details and the features of the original images. The larger absolute transform values in these sub bands correspond to sharper intensity changes and thus to salient features in the image such as edges, lines, and region boundaries. Therefore a good integration rule is to select the larger absolute value of the two coefficients of the wavelet transform at each point. Low-frequency sub bands related to the coarse part of the images, while high-frequency corresponds to the region boundaries or edges. As mentioned above, the task of image fusion is to combine the complementary information carried by multimodality images.

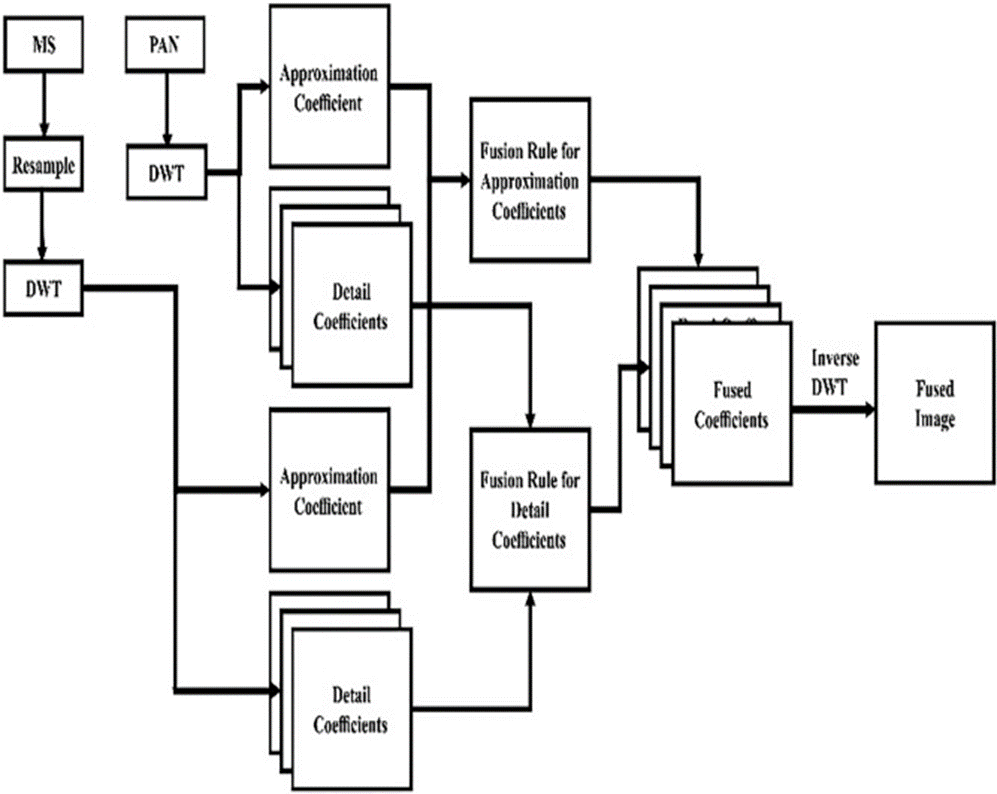
****

Figure 3.1.2 A detailed DWT based fusion process depiction.

**More properties of the wavelet transform**:

•No need for more pixels

•Explicit multi-scale representation

•Invertible

•Linear

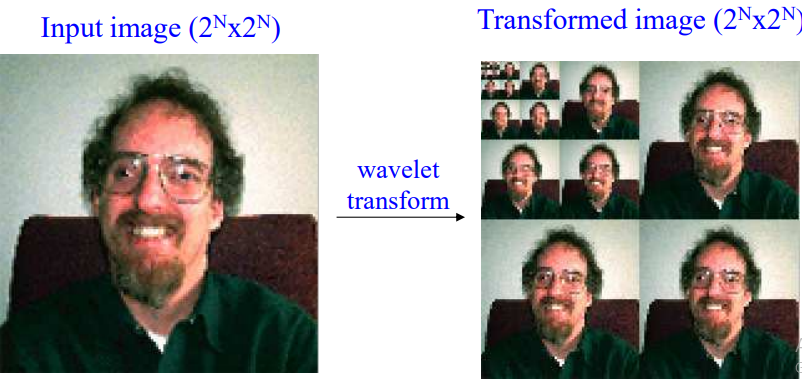
****

Figure 3.1.2 Example depicting the pixel decomposition in DWT

In the case of images, the DWT is performed first row by row and then column by column. As a result, four sub images are generated:

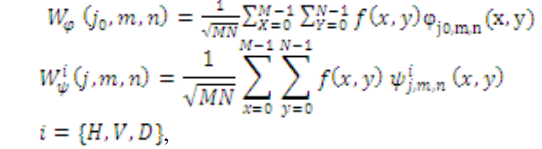
* The top left sub image contains an approximation of the original image.
* The top right sub image contains the horizontal details.
* The bottom left sub image contains the vertical details.

The bottom right sub image contains the diagonal details.

In most cases, the filters used in DWT have floating point coefficients and therefore, floating point output value. When the input data consists of sequences of integers, which in the case of images, the resulting filtered outputs no longer consist of integers. Therefore, these coefficients have to be rounded prior to the entropy coding step.

Mallat’s algorithm enabled the application of two dimensional DWT using one dimensional filter banks. Its general structure is very similar to that of the Gaussian pyramid based approach. Input signals are transformed using the wavelet decomposition process into the wavelet pyramid representation. Contrary to Gaussian pyramid based methods, high pass information is also separated into different sub band signals according to orientation as well as scale. The scale structure remains logarithmic, i.e. for every new pyramid level the scale is reduced by a factor of 2 in both directions. The wavelet pyramid representation has three different sub band signals containing information in the horizontal, vertical and diagonal orientation at each pyramid level. The size of the pyramid coefficients corresponds to “contrast” at that particular scale in the original signal, and can therefore, be used directly as a representation of saliency. In addition, wavelet representation is compact, i.e. the overall size of all sub-band signals in the pyramid is the same as the size of the original image.

The discrete wavelet transform in two dimensions of functions f(x, y) of size M\*N is



Here the function W is an approximation of f(x, y) and j is an arbitrary starting scale and i is a superscript that assumes the values H, V, and D.

As this method relies upon transformations therefore some color artifacts may occur which may decrease the performance of the transform based vision fusion techniques. In wavelet transform technique not much importance is given to edge detection is given, which plays an important role in identifying objects from the normal scenario.