Visualization of Hyperspectral Images using Bilateral Filtering

Ketan Kotwal, and Subhasis Chaudhuri, *Senior Member, IEEE*

Abstract

This paper presents a new approach for the hyperspectral image visualization. A bilateral filtering based approach is presented for hyperspectral image fusion to generate an appropriate resultant image. The proposed approach retains even the minor details that exist in individual image bands, by exploiting the edge preserving characteristics of a bilateral filter. It does not introduce visible artifacts in the fused image. A hierarchical fusion scheme has also been proposed for implementation purposes to accommodate a large number of hyperspectral image bands. The proposed scheme provides computational and storage efficiency without affecting quality and performance of the fusion. It also facilitates the midband visualization of a subset of the hyperspectral image cube. Quantitative performance results are presented to indicate the effectiveness of the proposed method.

Index Terms

Image fusion, hyperspectral imaging, bilateral filter, visualization, hierarchical fusion.

I. Introduction

With the advancement in the remote sensing technology, several hundred narrowband images of the same area on the earth are available. The hyperspectral sensors provide high resolution spectral information covering the visible as well as the infrared wavelength spectra. This information is collected over contiguous, narrow spectral bands by measuring the reflectance response of the surface to different wavelengths. The hyperspectral image applications have been extensively researched in the areas of remote sensing, environment monitoring, geological surveying and surveillance due to their distinct advantages in classification and object identification. Recently, hyperspectral imaging has also been found to be extremely useful in surveillance for coastal as well as cross-boarder material transport [1]–[3]. However, analysis of hyperspectral data faces following challenges-

High computational cost: The computational costs of processing are very high because of the large number of image bands (usually 200-250).

Storage requirement: A single hyperspectral image dataset contains a large number of image bands of the scene. The storage of such data requires a few hundreds of megabytes of memory.

Redundancy: The nearby image bands in the hyperspectral data cube exhibit a high degree of spatial correlation among them due to the contiguous and narrow nature of the hyperspectral sensors. This redundancy of the data needs to be removed for the efficient processing and reduction in the computational overheads.

The hyperspectral data presents abundant multidimensional information which contains far more image bands than those can be displayed on the standard tristimulus display. Therefore, efficient and appropriate means of visualization of the hyperspectral data is the need of the hour [4].

In order to overcome the above mentioned problems and to exploit this rich source of information, efficient data fusion could be a primary key step in the hyperspectral image visualization which deals with combining the data from multiple sensors. Various image fusion techniques are being developed for application oriented visualization, effective interpretation, extraction of useful features and to provide a better description of the scene using reduced datasets.

In this paper, we propose a very generic but extremely powerful scheme for the fusion of the hyperspectral image cube using bilateral filtering for visualization purpose. We propose a hierarchical approach to spectral fusion to accommodate a large number of image bands. This makes the approach computationally very efficient. Our scheme can be used to generate an RGB image for the standard tristimulus display, as well as to generate a single grayscale image from the given hyperspectral image cube.

The remainder of the paper is organized as follows: Section II provides a review of existing hyperspectral image fusion and visualization techniques. We describe the proposed approach for the hyperspectral image fusion using a bilateral filter in Section III. Section IV describes the actual implementation of our algorithm. The results and quality assessment are discussed in Section V. Conclusions are drawn in Section VI.

II. Related Work

The research in the remote sensing image fusion can be broadly divided into two categories- *i*) multispectral (MS) and panchromatic (pan) image fusion, and *ii*) multiband image fusion. In the former category, the multispectral images with higher spectral resolution but lower spatial resolution are fused with the panchromatic images of higher spatial but lower spectral resolution. Various methods for achieving this task, known as pan-sharpening have been discussed along with the comparison of these algorithms in [5], [6]. Several pan-sharpening approaches have been reported which include IHS (Intensity-Hue-Saturation) transformation [7], component substitution framework [8], specific image formation model [9], etc. Recently, wavelet packet transform based fusion [10] and contourlet transform based fusion [11] techniques have also been suggested for pan-sharpening which depend on multi-resolution decomposition of spectral bands. The pan-sharpening algorithms involve extraction of high resolution spatial data from panchromatic image and merging it with multispectral images for sharpening.

The multiband image fusion deals with combining several images of the same spatial resolution to form a single image. The pan-sharpening techniques do not prove to be suitable for the multiband image fusion, where all the images have same spatial as well as spectral resolution.

The multiband image fusion techniques can be classified into three categories based on the level at which the fusion operation takes place namely, pixel based, feature based, and classification based. Different techniques for fusing several images into a single image have been developed, however the extension of these schemes to hyperspectral image fusion problem is non trivial. The classification based methods classify the hyperspectral data

into various classes and these classes are assigned pseudo-colors for the visualization purpose. These methods suffer from following major drawbacks- limited training data for classification, high computational costs and non-natural appearance of the resultant image. The pixel or feature based techniques aim at selecting relative importance of the pixels across various bands and accordingly designing a fusion strategy. The simplest method of hyperspectral image fusion is based on the averaging of spectral bands. This method assigns equal weights to each spectral band and produces the result equivalent to integration of spectral response at each pixel. As the information across bands is not uniformly distributed, this technique loses high amount of information. Guo *et al.* [12] proposed assignment of unequal weights to spectral bands depending upon the application and purpose of visualization; therefore application specific kernels were used to assign spectral weights to the bands. These weights depend upon the application specific information contained by the band, hence helping in retaining the useful contents in order to achieve better classification.

Various multi-resolution based methods for fusing images have recently been extended to hyperspectral image fusion. In [13], three multi-resolution techniques have been analyzed for the performance. These algorithms include decomposition of the base images using a Gaussian pyramid to generate multi-resolution representations. These pyramidal decompositions are fused at each level using a predefined fusion rule to generate the representation of final fused image at that pyramidal level, which are appropriately synthesized to obtain the fused image. The fusion rules are based on saliency of the pixel, which is a measure of "relative importance" of the pixel calculated from its neighborhood. The most commonly used fusion rule for the lowest resolution is averaging, while rules for fusion at the other levels include use of energy, gradient and high frequency contents in the neighborhood of the pixel in order to compare the relative amount of information among different images at the corresponding pixel location. The image fusion rules involving selection of a pixel in a particular band with higher information and discarding the corresponding pixels in other bands are not advisable for multispectral fusion, however fusion rules involving a weighted average can be extended to multispectral fusion, where the sum of weights is normalized to one. The wavelet based techniques have been proved superior to traditional image fusion for their ability to provide directional information at each decomposition level [14]. The hyperspectral detectors cause line-to-line stripping and scanline noise due to unequal detector sensitivities and electronic factors [4]. The wavelet based fusion methods do not yield good results with this kind of noise, as spatial distortions in the form of ringing or aliasing effects, or blurring of contours and textures may take place during the fusion process [6].

Researchers have proposed a number of dimensionality reduction based approaches for the image fusion problem. Principal component analysis (PCA) is used as a classical technique of dimensionality reduction [15]. This technique finds out the basis vectors in the direction of maximum data variance, and projects image data onto those basis vectors to create the resultant image. In [16], the first three principle component vectors have been mapped to RGB or HSI color space. In [17], three different schemes of partitioning the hyperspectral dataset have been suggested, which facilitate efficient fusion to generate an RGB image. Since the PCA based fusion schemes are entirely dependant on image data statistics, same grayscale values in different fused images are not consistent in the visual information. Additionally, this method involves a significant amount of computation. To avoid the data dependency

of PCA based methods, Jacobson *et al.* [18] introduced linear projection onto fixed basis functions which are based on certain optimizing criteria for a standard device color space (sRGB) and perceptual color space (CIELab). They have also extended this approach by assigning higher weights to the bands having higher SNR values when SNR values for hyperspectral image bands are known. An image fusion approach involving classification as an intermediate stage employing two layers to integrate the mixture information is reported in [19]. Demir *et al.* [20] proposed a low complexity approach to generate a color display using the method of one bit transform (1BT) to determine three suitable image bands with more information. However this technique loses the information available in the rest of the hyperspectral image cube. A visualization technique based on convex optimization for preservation of spectral distances is proposed in [21]. This technique allows interactive visualization to meet the needs of human analyst. In [22], an analysis of RGB color composition for hyperspectral image visualization has been carried out for the approaches involving data transformation and data classification.

In this paper we propose a very different type of fusion strategy that tries to define the fused image as a weighted sum of individual components, where the weight reflects the locally dominating features of the pixel in a given band. Thus, there is no loss of information from individual spectral bands. The method is fully automatic and is made computationally and memory-wise very efficient through a hierarchical implementation of the proposed scheme.

The problem of quality assessment of fused grayscale (or RGB) image still does not have a universally accepted solution. Several measures for image as well as fusion quality have been suggested when the reference image or ground truth is available. However such performance evaluation in the absence of a reference image is still an open problem. In [23], some goals for visualization of hyperspectral images are discussed, which mainly include-consistent rendering, edge preservation, computational ease, equal energy white point and natural palette for fused RGB image. Zhu *et al.* [24] suggested use of entropy and correlation coefficient as quantitative measures for fusion performance. Notwithstanding above, we have used the existing methods of performance evaluation to compare the performance of the proposed scheme of fusion with other methods.

III. Proposed Method

The process of image fusion aims at formation of a single image which selectively merges maximum possible features from the source images. In case of hyperspectral image bands, the images are obtained by sampling a continuous spectrum at narrow wavelength intervals (e.g. each channel of AVIRIS instrument has a nominal bandwidth of 10 nm). As the spectral response of the scene does not vary much over the adjacent bands, extracting specific information contained by a particular band is the key to fusion process. We calculate the value of each pixel of the fused image by a normalized weighted average of the pixels among different bands at the corresponding location. The critical part of the algorithm lies in choosing appropriate weights representing the subtle information at each location along the available spatial and spectral ranges, which indicates selecting different weights for every pixel in each of the image band. We define the weight of the particular pixel in the image from the relative importance of the feature with respect to its neighborhood. A smoothing 2-D filter removes the slowly varying

features from the image and therefore subtracting the filtered image from original image gives the important local details in the image. The magnitude of these local details decides the weight of the pixel in fusing the set of pixels at the same location along the wavelength axis. However, any ordinary smoothing filter tends to blur the edges and then the difference image contains artifacts in the neighborhood of any edge pixel. This introduces visible artifacts in the fused image. Hence an appropriate smoothing filter must be used which does not smooth strong edges while removing minor variations in the image. We propose a bilateral filtering based approach to accomplish the above task and achieve an efficient hyperspectral image fusion.

In the following subsections we first discuss preliminaries of the bilateral filter. We describe our approach of the multiband fusion and the scheme of hierarchical fusion for computational efficiency in the subsequent subsections.

A. Bilateral Filter

Bilateral filter is a nonlinear filtering technique introduced by Tomasi and Manduchi [25] which encompasses the combined effects of spatial domain filtering and range filtering to achieve image smoothing while preserving sharp edges. This technique combines the neighborhood pixels based on their spatial as well as photometric closeness, to assign pixel values which satisfy proximity in both the domains. Mathematically the bilateral filter can be represented as following [25]:

$$\mathbf{h}(\mathbf{x}) = \frac{1}{k(\mathbf{x})} \int_{-\infty}^{\infty} \mathbf{f}(\boldsymbol{\xi})\, c(\boldsymbol{\xi}, \mathbf{x})\, \mathcal{S}\big(\mathbf{f}(\boldsymbol{\xi}), \mathbf{f}(\mathbf{x})\big)\, d\boldsymbol{\xi} \tag{1}$$

where $\mathbf{f}(\mathbf{x})$ is the input image and $\mathbf{h}(\mathbf{x})$ is the corresponding bilaterally filtered image. The function $c(\boldsymbol{\xi}, \mathbf{x})$ measures spatial closeness of neighborhood center $\mathbf{x}$ and a surrounding point $\boldsymbol{\xi}$. The photometric similarity (range closeness) between the neighborhood pixel $\mathbf{x}$ and surrounding point $\boldsymbol{\xi}$ is measured by $\mathcal{S}\big(\mathbf{f}(\boldsymbol{\xi}), \mathbf{f}(\mathbf{x})\big)$. The normalization factor k is given by

$$k(\mathbf{x}) = \int_{-\infty}^{\infty} c(\boldsymbol{\xi}, \mathbf{x})\, \mathcal{S}\big(\mathbf{f}(\boldsymbol{\xi}), \mathbf{f}(\mathbf{x})\big)\, d\boldsymbol{\xi}. \tag{2}$$

By choosing the range and spatial filtering functions as the Gaussian functions, the bilateral filtering extends the concept of the Gaussian blurring in the spatial domain with additional range weight contributed by the intensity domain. Let I be the image to be processed using bilateral filter. Let G_{σ_S} be the Gaussian spatial kernel, similar to traditional Gaussian filter. The value of σ_S decides the spatial extent of the kernel or the neighborhood under consideration. Let G_{σ_R} be the Gaussian range kernel, where σ_R decides "amplitude" of the edge and its corresponding weight [25]. The new value of pixel (x, y) of the image I is obtained from pixels $(\tilde{x}, \tilde{y})$ in the neighborhood of the corresponding pixel as shown below,

$$I_{BF}(x, y) = \frac{1}{W(x, y)} \sum_{\tilde{x}} \sum_{\tilde{y}} \Big\{ G_{\sigma_S}(x - \tilde{x}, y - \tilde{y})\; G_{\sigma_R}(I(x, y) - I(\tilde{x}, \tilde{y}))\; I(\tilde{x}, \tilde{y}) \Big\} \tag{3}$$

$$G_{\sigma_S}(x, y) = e^{-\frac{x^2 + y^2}{2\sigma_S^2}} \tag{4}$$

$$G_{\sigma_R}(a) = e^{-\frac{a^2}{2\sigma_R^2}} \tag{5}$$

where $W(x, y)$ is the normalization factor similar to $k(\mathbf{x})$ in the discrete form, and I_{BF} is the output of the bilateral filter.

The ideal bilateral filtering process involves the product of two Gaussian functions and therefore it is computationally time consuming. However faster and computationally simpler approximations of discrete bilateral filter have been proposed in the literature [26], [27]. Bilateral filter has been used in various applications of computational photography, like image denoising, tone mapping and high dynamic range imaging [28]–[30]. In a recent paper Bennett *et al.* [31] have also used a modified bilateral filter for the fusion of visual and IR video, yielding enhanced visual band output. However, their methodology is very different from that of proposed method. They use bilateral filter for noise removal and subsequent noise free edge extraction from IR data. These edges are simply added to the visual band data through an appropriate but heuristically chosen weight. In the proposed method the output of the bilateral filter is directly used in calculating the weights for fusion of various bands. Hence the use of bilateral filter in the proposed technique is more in the spirit of defining α-mattes [32], [33], rather than details extraction.

B. Image Fusion

The objective of fusing a subset of the hyperspectral data cube is to obtain a single image which retains the features from all the constituent bands as far as possible. The hyperspectral images are generated as the reflectance response of the surface, which mainly depends on the properties and composition of the material in the scene. Certain substances exhibit stronger response over a given wavelength range, and form strong and sharp features in the corresponding subset of the image cube. On the other hand, certain substances have peak reflectance response over a very narrow wavelength range. These substances form weak features, i.e. edges and textures, which are seen only in few images. We need to preserve particularly these weak features in the correct proportion during the fusion.

The bilateral filtering operation is well known for the characteristics of edge preservation and removal of fine textures. We propose bilateral filtering technique to separate the weak edges and fine textures by subtracting the bilateral filtered image from the original image from the subset of hyperspectral image cube. We use magnitudes of these features as the weights during the fusion after normalizing them to one across the wavelength. This approach is similar to the compositing of low dynamic intensity photographs with variable exposures as discussed in [34]. Here we handle the problem of fusing remote sensing images with variable reflectance of various regions contained in the scene spread over a wide spectrum.

Let $I(x, y, \lambda_1)$ to $I(x, y, \lambda_M)$ be the subset of hyperspectral image cube, containing M images from consecutive wavelength bands λ_1 to λ_M. We calculate weight at each pixel (x, y) for each image, w_1 to w_M using the bilateral filter as shown below,

$$w_i(x, y) = \frac{|I(x, y, \lambda_i) - I_{BF}(x, y, \lambda_i)| + K}{\sum_{i=1}^{M} \left(|I(x, y, \lambda_i) - I_{BF}(x, y, \lambda_i)| + K\right)} \tag{6}$$

where I_{BF} is the corresponding bilateral filtered image. K is a positive real number that allows flexibility in the fusion process by increasing or decreasing the effect of actual weight components as well as prevents numerical

instability at homogenous regions. Here the weights are proportional to the difference images which mainly contain the weak features and edges. The constant K also helps in retaining the strong edges in the fused image which are removed in the difference image. The denominator in the Equation 6 guarantees that the sum of all weights equals unity, $\sum_{i=1}^{M} w_i(x,y) = 1$. The fused image of the hyperspectral cube subset, I_F is given by

$$I_F(x,y) = \sum_{i=1}^{M} w_i(x,y)\, I(x,y,\lambda_i). \tag{7}$$

C. Hierarchical Fusion

A typical hyperspectral image dataset in remote sensing contains a few hundred images to be fused into a single image (for grayscale) or three images (for RGB). Combining all the image bands of the given data set together results in assigning very small fractional weights to the locations in the each of the image bands. In this procedure, some of the weights are comparable to the truncation errors and hence some of the minor details may wash out during fusion. Furthermore, the data along the spectral dimension is needed for computation of such fractional weights. Therefore the procedure requires the entire hyperspectral cube to be read into memory. Considering the huge size of a hyperspectral image cube, the memory requirement goes over a few hundreds of megabytes.

We suggest a process of hierarchical fusion to overcome these issues and also to provide a better image analysis. For the hyperspectral image cube of dimensions $(X \times Y \times N)$, containing N image bands, we apply the bilateral filter based fusion across a contiguous subset of dimensions $(X \times Y \times M)$ to generate $P = \lceil N/M \rceil$ different fused images at the first stage.

These P images represent the results of the first stage of the fusion. These images form the base images for the second stage fusion, which is carried out using the same algorithm over newly formed image cube of dimensions $(X \times Y \times P)$. We repeat the procedure of such a fusion hierarchically to generate the final fused image within, say, J steps. By generating three fused images at the pre-final stage and assigning them to appropriate color channels, we can obtain RGB representation of the hyperspectral image cube for the tristimulus visualization.

This technique requires only M number of contiguous image bands for the fusion at the corresponding stage. Here, the memory requirement reduces significantly as only M images out of N are read into memory, which are typically less than 10 % of the original number of image bands. The resultant fused images at the intermediate stages facilitate the visualization and analysis of midband reflectance response of the scene. Thus, if δ is the bandwidth of the individual narrow bands of the hyperspectral image band, then $M\delta$ is the bandwidth of each of the fused images in the first stage. The images contain the response over a wider bandwidth as one moves to the higher levels. Each of the images at the prefinal stage encompasses approximately one third of the bandwidth of the total range of hyperspectral bandwidth. The final fused image represents the fusion over the entire bandwidth of the available data.

Fig. 1 shows the schematic representation of the proposed hierarchical fusion. The scheme easily accommodates any increase in the number of bands without affecting the performance of the fusion scheme and the quality of the fused image.

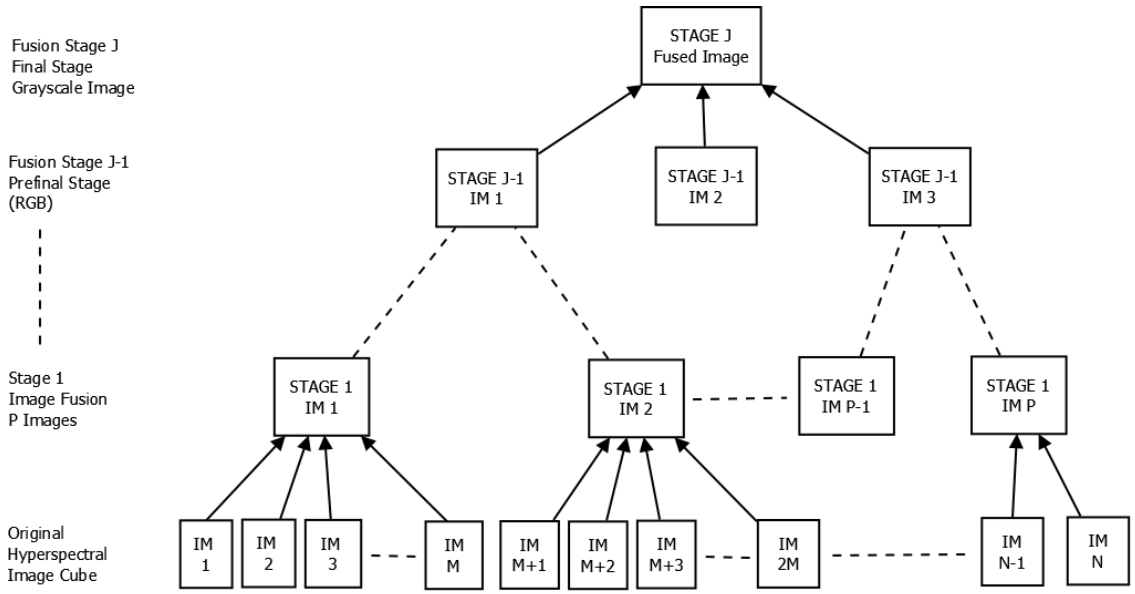


Fig. 1: Proposed scheme of the hierarchical fusion for hyperspectral image cube.

It may be noted that in literature, some of the multi-resolution based fusion methods are termed as hierarchical fusion [13]. The proposed technique is different from these multi-resolution based approaches. We do not perform any subband decomposition of the image cube.

IV. Implementation

The fusion process is controlled by 3 parameters- σ_S, σ_R and K. Choice of appropriate values of these parameters is necessary in order to achieve a better fusion quality. One can set these parameters depending upon the contents of the given hyperspectral data and by evaluating the quality of fusion. This approach, although more accurate, becomes data dependent and it cannot be directly applied to different datasets. In order to automate the fusion process without any degradation in quality, we adopt the following guidelines as discussed in [28]. The value of range kernel parameter σ_R defines the minimum amplitude of the intensity value in the image that can be considered as an "edge". We calculate the difference between maximum and minimum values of intensity of a given image dataset to find the actual intensity range in the image. We choose the value of σ_R as the fraction of the intensity range to define the edge measure.

$$\sigma_R = \alpha\big[max(I(x,y,\lambda)) - min(I(x,y,\lambda))\big] \tag{8}$$

where α is set to 0.02 during experiments for all data sets,

The choice of σ_S decides the size of spatial details retained during fusion, which is related to ground projected instantaneous field of view *(GIFOV)* in remote sensing. Generally the value of σ_S is fixed as the percentage of the image dimensions. Thus, for an image of dimensions $(X \times Y)$, we define spatial kernel parameter as-

$$\sigma_S = \beta\big[min(X,Y)\big]. \tag{9}$$

In all test experiments, we have used $\beta = 0.50$.

The constant K can be selectively chosen to adjust the needs of an image analyst. Lower values of K result in boosting the finer details, which may result in over sharpening. On the other hand, excessively higher values of K result in an "average-like" fusion tending to smooth out the fused image. For the 16 bit experimental data, we selected the value of K to be equal to 50 after experimenting over the range from 10 to 1,000.

The discrete approximation of bilateral filtering enables a faster computation. Weiss [27] suggested a Box kernel approach for faster bilateral filtering; Paris and Durand [26] implemented a 3D kernel for fast approximation of the bilateral filter. These algorithms speed up the bilateral filtering process due to ease of computation. We have used the 3D kernel approximation of bilateral filter by Paris and Durand [26], which has a complexity of $\mathcal{O}\big(|\mathcal{S}| + \frac{\mathcal{S}}{\sigma_S^2}\frac{\mathcal{R}}{\sigma_R}\big)$ where $\mathcal{S}$ is the image size and $\mathcal{R}$ denotes range (intensity) domain [26]. Therefore for the given image domain $\mathcal{R}$ and fusion parameters σ_S and σ_R, the fusion process has linear time complexity with respect to the number of images to be fused. Very recently Yang *et al.* [35] have proposed a very fast, constant time complexity method of computing the bilateral filter. This can be used to further speed up the computation.

V. Results and Discussions

To test our algorithm, we have used hyperspectral urban data provided by the Hyperion imaging sensor used in the EO-1 spacecraft for Earth observation.[1] The data consists of 242 bands (0.4 μm to 2.5 μm) with 30 meter spatial resolution. We have used the terrain corrected dataset (designated as *Level G1*), which is provided as 16 bit radiance values. The selected dataset depicts the urban region of Palo Alto, California (Latitude: 37.4761^0 N, Longitude: 122.1341^0 W). The dimension of the Hyperion data cube is $(256 \times 512 \times 242)$. We have also used the Moffett Field data provided by the Airborne Visible Infrared Imaging Spectrometer (AVIRIS) operated by the National Aeronautics and Space Administration (NASA)/Jet Propulsion Laboratory (JPL).[2] This data is provided as 16 bit reflectance values in 224 bands, where each image band has the size of (614×512). The selected dataset depicts the Moffett Field in California covering vegetation, water resources and urban area (Latitude: 37.4496^0 N, Longitude: 122.0143^0 W). This dataset is designated as 'Moffett02'.

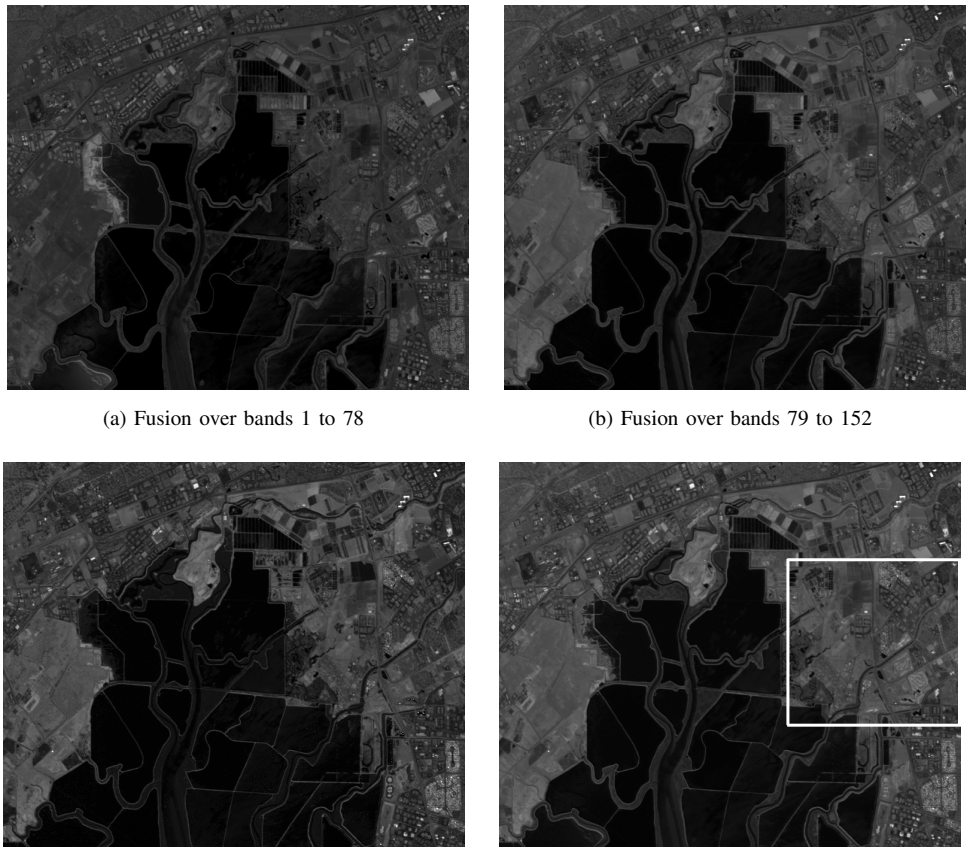
Another two hyperspectral datasets, one each from the Hyperion and the AVIRIS were also used for testing the algorithm. The second Hyperion dataset depicts the area located at latitude 37.4758^0 N and longitude 117.4916^0 E containing geological features. The second dataset from AVIRIS represents the Moffett Field area, adjacent to the first AVIRIS dataset, Moffett02. This dataset is designated as 'Moffett03'.

Some of the bands from the hyperspectral image sets have near zero response as the molecular absorption bands of water and carbon dioxide block the transmission of radiation in the certain wavelength bands [4]. After removing these bands, the number of available bands reduces to about 195 to 200.

We have employed a three stage hierarchical fusion for both the Hyperion and the AVIRIS datasets. Each image band in the AVIRIS dataset represents the spectral response of the scene over 10 nm on average. The first stage

[1]The Hyperion data downloaded from http://eo1.usgs.gov

[2]The AVIRIS data downloaded from http://aviris.jpl.nasa.gov



(a) Fusion over bands 1 to 78

(b) Fusion over bands 79 to 152

(c) Fusion over bands 153 to 224

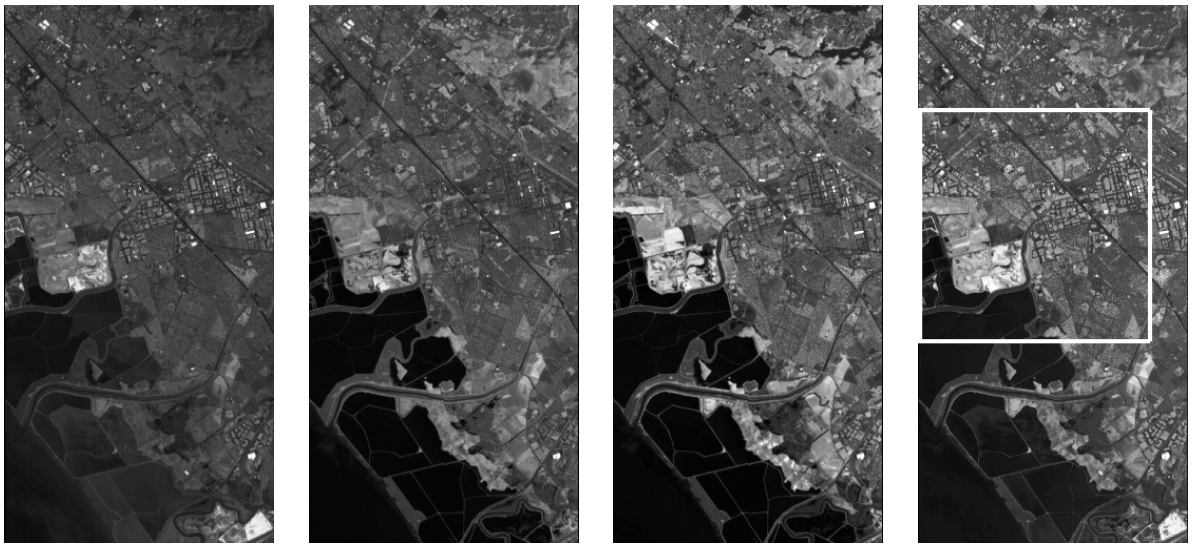
(d) Final fused image

Fig. 2: (a-c) Results of second stage fusion of Moffett field image cube (Moffett02) from AVIRIS data, and (d) the final fused image.

fusion generated 18 fused images where each of the fused image has the bandwidth of approximately 110 - 120 nm. These 18 images were used for the second stage fusion forming 3 images. The first fused image (Fig. 2a) represents the fusion from bands 1 to 78 covering the bandwidth of 735.190 nm (from 365.045 nm to 1100.235 nm). The second fused image covers the bandwidth of 706.875 nm (from 1100.255 nm to 1807.130 nm) where bands from 79 to 152 are fused using the bilateral filter technique. The last fused image (Fig. 2c) at the pre-final stage represents the fusion of bands from 153 to 224 from the original hyperspectral image cube. The bandwidth of fused image is 705.140 nm (from 1806.685 nm to 2511.825 nm). The band numbers in each fused image are not exactly one third of the number of bands in the original hyperspectral image cube as few image bands were removed during preprocessing, as mentioned earlier. These fused images of the dataset Moffett Field (Moffett02) by the AVIRIS are shown in Figs. 2(a-c). Naturally, one can display these three fused images under a tristimulus (pseudo color) display by assigning them blue, green and red channels respectively. We generate the overall fused image from these three images for the final level fusion and is shown in Fig. 2d. The final fused image of the Moffett Field has the bandwidth of 2146.780 nm as the collective spectral response of the scene integrated using the bilateral

filtering technique.

A similar procedure has been applied to the Hyperion data of Palo Alto region (designated as urban dataset), where each original image band measures the spectral response over the bandwidth of 10 nm. The fused images in the first stage comprise the bandwidth of approximately 110 - 120 nm, which are generated by fusing 12 original contiguous bands. The second stage fusion generated 3 images from 18 resultant images of the first stage fusion. The first fused image (Fig. 3a) at this stage results from fusing the bands from 1 to 74, corresponding to the wavelengths from 350.505 nm to 887.235 nm. The second fused image of pre-final stage represents the fusion over bands 75 to 154. The bandwidth of this fused image of the Urban data scene is 807.531 nm (from 887.240 nm to 1694.771 nm). The last fused image at this stage has the bandwidth of 888.508 nm, which represents the fusion of bands 155 to 242 of the original hyperspectral image cube, corresponding to the wavelengths from 1693.765 nm to 2582.273 nm. Figs. 3(a-c) show these fused images of Palo Alto region collected by the Hyperion imaging sensor after the second stage. We obtain the final grayscale fused image of the urban dataset of the Hyperion data by fusing these three pre-final stage images, which has the bandwidth of 2231.768 nm and is shown in Fig. 3d.



(a) Fusion over bands 1 to 74 (b) Fusion over bands 75 to 154 (c) Fusion over bands 155 to 242 (d) Final fused image

Fig. 3: (a-c) Results of second stage fusion of urban image cube (Palo Alto) from Hyperion data, and (d) the final fused image. These images are linearly scaled to the range [0, 255] for display purposes.

It is clear from these results that the fused images are able to retain the textural details in the constituent bands very well. This is specifically so in the urban areas where the need to preserve the spatial details is very crucial. Further, ringing artifacts and blurring of textures that appear in the fused images using wavelet based approaches [6] are totally absent in the results of the proposed method.

Objective assessment of fusion quality for the multiband image fusion in the absence of ground truth still does not have a universally accepted solution. The performance parameters for multispectral and pan image fusion have

been reported in the literature [36]–[38], but these cannot be applied to the fusion schemes where reference image is unavailable. We present the performance evaluation of the proposed scheme on the basis of statistical assessment parameters given in [9], [39]. The entropy of the image describes the total amount of information contained in the image, which is calculated as-

$$H_x = -\sum_x p_X(x) \ln p_X(x) \tag{10}$$

where p_X is the probability density of the intensity level x in the image. However image entropy tends to be higher when the image contains noise. Therefore entropy cannot be a reliable performance measure in case of noisy images. The average gradient of an image, $\bar{g}$ is the measure of image sharpness in terms of gradient values. If I_x denotes the difference in the x-axis direction and I_y denotes the difference in the y-axis direction for the image size N, the average gradient is given by-

$$\bar{g} = \frac{1}{N} \sum_x \sum_y \sqrt{{I_x}^2 + {I_y}^2}. \tag{11}$$

This parameter has also been reported as *definition* of an image in [39].

We compare our results with some of the recent techniques- the three-band representation [18], use of piecewise linear functions [18], technique of the color matching functions [18], and one bit transform (1BT) based method [20]. The visualization results for the proposed approach and the above mentioned approaches are shown in finer details in Fig. 4 for a block of the AVIRIS hyperspectral image cube to visually inspect the quality and sharpness of the fused images. It can be seen that the result of the proposed method not only provides sharp features but also preserves the finer details with accuracy. At the bottom right, the rectangular regions are seen to have some linear textural pattern, which is not properly visible in results of visualization using other methods. The output in Fig. 4b appears to be quite saturated at various locations. However the results of the proposed method do not suffer from the saturation problem. Similar comparative results for a block of the Hyperion hyperspectral image cube are provided in Fig. 5, where it can be observed that the proposed method retains the minor details from its constituent image bands. At the bottom middle portion of the image, more details can be clearly seen in the proposed method. Other methods do not yield so much details in the fused image. We also provide the statistical assessment results of the fused images in Tables I and II. It can be seen that the proposed approach yields results with a higher variance, representing a better contrast. Higher entropy of the bilateral filter based fused image indicates a higher information content over the other methods. Similarly better results can also be seen for the average gradient ($\bar{g}$) parameter, which measures the sharpness of the image. The mean of the image, representing the average gray value of the image pixels, suggests that the proposed method does not necessarily prefer the pixels with higher intensity values.

VI. Conclusion

In this paper, we propose a new approach for the hyperspectral image fusion for the purpose of visualization using bilateral filtering. Apart from yielding better results, the proposed scheme provides flexibility in the process of fusion by modifying the bilateral filter parameters and the tuning parameter K, to match the needs of the observer. The bilateral filter being a non-iterative filter and with efficient implementation scheme being available, the fusion

(a) (b)

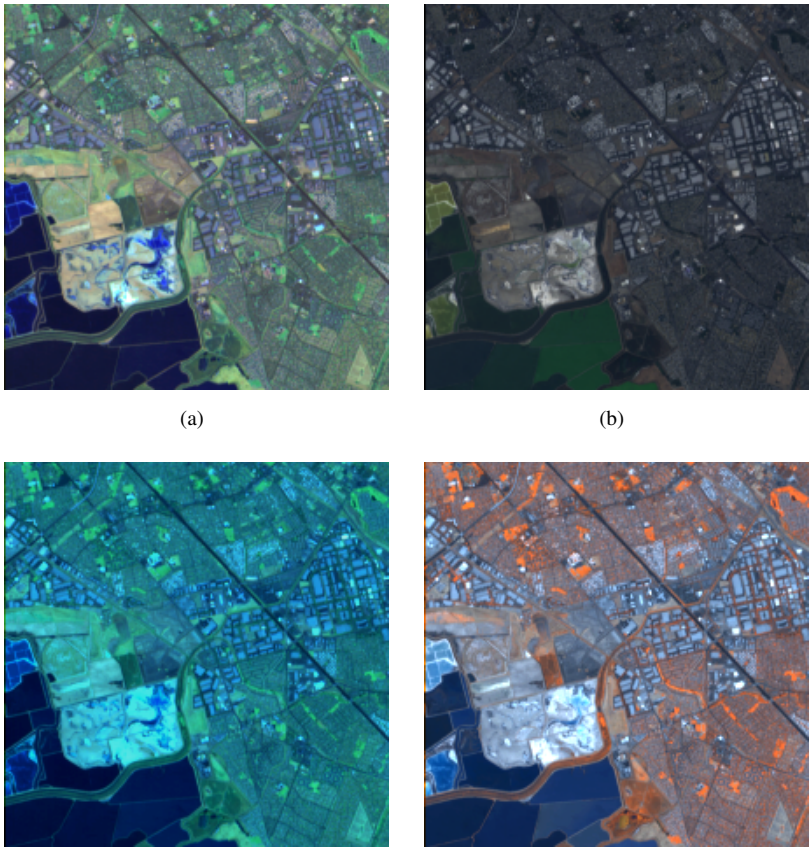
(c) (d)

Fig. 4: Visual comparison of performances of various image fusion approaches over a block of the Moffett field hyperspectral image cube (Moffett02) shown in Fig. 2d. Results of fusion using (a) proposed method after tristimulus display of fused images shown in Figs. 2(a-c), (b) three band selection approach, (c) piecewise linear function approach, and (d) color matching function approach. Gamma is set to 1.5 for all color displays.

TABLE I: PERFORMANCE EVALUATION OF THE PROPOSED APPROACH FOR THE AVIRIS IMAGES. M02 STANDS FOR MOFFETT02 AND M03 STANDS FOR MOFFETT03 DATA.

Fusion Method	Mean μ		Variance σ^2		Entropy H		Avg Gradient $\bar{g}$	
	M02	M03	M02	M03	M02	M03	M02	M03
Three Band Selection	24.43	21.73	121.45	86.64	4.89	4.75	2.82	2.21
Piecewise Linear Function	30.53	23.95	383.46	282.05	5.84	5.55	3.89	3.05
Color Matching Function	15.81	24.88	131.12	157.50	5.01	5.19	2.95	2.59
One Bit Transform	32.39	15.83	481.41	102.37	6.19	4.82	4.43	2.42
Proposed method	33.39	23.16	491.86	391.29	6.02	5.60	4.32	3.35

process is faster and computationally simpler. The weight for each of the pixel in the entire hyperspectral cube is calculated separately, depending upon the relative contribution of the corresponding pixel in the scene locally.



(a) (b)

(c) (d)

Fig. 5: Visual comparison of performances of various image fusion approaches over a block of Urban hyperspectral image cube (Palo Alto) shown in Fig. 3d. Results of fusion using (a) proposed method after tristimulus display of fused images shown in Figs. 3(a-c), (b) three band selection approach, (c) piecewise linear function approach, and (d) color matching function approach.

TABLE II: PERFORMANCE EVALUATION OF THE PROPOSED APPROACH FOR THE HYPERION IMAGES. URB STANDS FOR URBAN AND GEO STANDS FOR GEOLOGICAL DATA.

Fusion Method	Mean μ		Variance σ^2		Entropy H		Avg Gradient $\bar{g}$	
	Urb	Geo	Urb	Geo	Urb	Geo	Urb	Geo
Three Band Selection	52.42	96.67	312.47	349.03	5.70	6.17	5.32	5.38
Piecewise Linear Function	41.19	99.02	282.54	405.05	5.80	6.16	4.22	6.03
Color Matching Function	56.16	90.78	260.91	374.70	5.78	6.31	4.41	5.72
One Bit Transform	48.71	114.70	477.04	337.07	6.24	6.10	6.41	5.19
Proposed method	49.21	91.65	619.50	417.27	6.28	6.35	6.37	6.14

Thus the fusion process exploits all available information at all locations in all bands, unlike the methods where the entire band is assigned the same weight.

We also propose a hierarchical implementation scheme for fusion as it can accommodate any increase in the number of bands without performance degradation. This scheme also enables analysis of intermediate band fused images as the result of multiband image fusion. This allows fusion of images upto any given spectral bandwidth suitable for mid-band visualization. The proposed technique does not require construction of multi-resolution pyramids, and therefore it is computationally simpler and faster. The algorithm does not involve any training phase as used in classification based fusion techniques.

ACKNOWLEDGMENT

The authors are thankful to the reviewers for their constructive suggestions and Bharti Centre for Communication for logistic support. Funding under JC Bose Fellowship is gratefully acknowledged.

REFERENCES

[1] D. Stein, J. Schoonmaker, and E. Coolbaugh, "Hyperspectral Imaging for Intelligence, Surveillance, and Reconnaissance," *Space and Naval Warfare Systems Center,San Diego,CA*, Aug. 2001.

[2] H. Ren and C. Chang, "Automatic spectral target recognition in hyperspectral imagery," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 39, pp. 1232–1249, Oct. 2003.

[3] J. Ardouin, J. Lávesque, and T. Rea, "A demonstration of hyperspectral image exploitation for military applications," in *Proceedings of the 10th International Conference on Information Fusion*, (Québec, Canada), pp. 1–8, Jul. 2007.

[4] R. A. Schowengerdt, *Remote Sensing - Models and Methods for Image Processing*. Academic Press, Inc., 3rd ed., 2007.

[5] Z. Wang, D. Ziou, C. Armenakis, D. Li, and Q. Li, "A comparative analysis of image fusion methods," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, pp. 1391–1402, Jun. 2005.

[6] L. Alparone, L. Wald, J. Chanussot, C. Thomas, P. Gamba, and L. Bruce, "Comparison of Pansharpening Algorithms: Outcome of the 2006 GRS-S Data-Fusion Contest," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, pp. 3012–3021, Oct. 2007.

[7] L. Alparone, L. Facheris, S. Baronti, A. Garzelli, and F. Nencini, "Fusion of Multispectral and SAR Images by Intensity Modulation," in *Proceedings of the 7th International Conference on Information Fusion*, (Stockholm, Sweden), pp. 637–643, 2004.

[8] W. Dou, Y. Chen, X. Li, and D. Z. Sui, "A general framework for component substitution image fusion: An implementation using the fast image fusion method," *Comput. Geosci.*, vol. 33, no. 2, pp. 219–228, 2007.

[9] M. Joshi, L. Bruzzone, and S. Chaudhuri, "A Model-Based Approach to Multiresolution Fusion in Remotely Sensed Images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, pp. 2549–2562, Sep. 2006.

[10] W. Zhang and J. Kang, *Lecture Notes in Control and Information Sciences*, vol. 344, ch. QuickBird Panchromatic and Multi-Spectral Image Fusion Using Wavelet Packet Transform, pp. 976–981. Springer Berlin, 2006.

[11] V. Shah, N. Younan, and R. King, "An Efficient Pan-Sharpening Method via a Combined Adaptive PCA Approach and Contourlets," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 46, pp. 1323–1335, May 2008.

[12] B. Guo, S. Gunn, B. Damper, and J. Nelson, "Hyperspectral image fusion using spectrally weighted kernels," in *Proceedings of the 8th International Conference on Information Fusion*, vol. 1, (PA, USA), pp. 7 pp.–, Jul. 2005.

[13] T. Wilson, S. Rogers, and M. Kabrisky, "Perceptual-based image fusion for hyperspectral data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 35, pp. 1007–1017, Jul. 1997.

[14] J. J. Lewis, R. J. O'Callaghan, S. G. Nikolov, D. R. Bull, and N. Canagarajah, "Pixel- and region-based image fusion with complex wavelets," *Information Fusion*, vol. 8, no. 2, pp. 119 – 130, 2007.

[15] P. Ready and P. Wintz, "Information Extraction, SNR Improvement, and Data Compression in Multispectral Imagery," *IEEE Transactions on Communications*, vol. 21, pp. 1123–1131, Oct. 1973.

[16] J. Tyo, A. Konsolakis, D. Diersen, and R. Olsen, "Principal-components-based display strategy for spectral imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 41, pp. 708–718, Mar. 2003.

[17] V. Tsagaris, V. Anastassopoulos, and G. Lampropoulos, "Fusion of hyperspectral data using segmented PCT for color representation and classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, pp. 2365–2375, Oct. 2005.

[18] N. Jacobson, M. Gupta, and J. Cole, "Linear Fusion of Image Sets for Display," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, pp. 3277–3288, Oct. 2007.

[19] S. Cai, Q. Du, and R. Moorhead, "Hyperspectral Imagery Visualization Using Double Layers," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, pp. 3028–3036, Oct. 2007.

[20] B. Demir, A. Çelebi, and S. Ertürk, "A Low-Complexity Approach for the Color Display of Hyperspectral Remote-Sensing Images Using One-Bit-Transform-Based Band Selection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, pp. 97–105, Jan. 2009.

[21] M. Cui, A. Razdan, J. Hu, and P. Wonka, "Interactive Hyperspectral Image Visualization Using Convex Optimization," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, pp. 1673–1684, Jun. 2009.

[22] Q. Du, N. Raksuntorn, S. Cai, and R. Moorhead, "Color Display for Hyperspectral Imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 46, pp. 1858–1866, Jun. 2008.

[23] N. Jacobson and M. Gupta, "Design goals and solutions for display of hyperspectral images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, pp. 2684–2692, Nov. 2005.

[24] Y. Zhu, P. Varshney, and H. Chen, "Evaluation of ICA based fusion of hyperspectral images for color display," in *Proceedings of the 10th International Conference on Information Fusion*, (Québec, Canada), pp. 1–7, Jul. 2007.

[25] C. Tomasi and R. Manduchi, "Bilateral Filtering for Gray and Color Images," in *ICCV '98: Proceedings of the Sixth International Conference on Computer Vision*, (Bombay, India), p. 839, 1998.

[26] P. Sylvain and D. Frédo, "A Fast Approximation of the Bilateral Filter using a Signal Processing Approach," in *ECCV '06: Proceedings of the European Conference on Computer Vision*, (Graz, Austria), May 2006.

[27] B. Weiss, "Fast median and bilateral filtering," *ACM Trans. Graph.*, vol. 25, no. 3, pp. 519–526, 2006.

[28] S. Paris, P. Kornprobst, J. Tumblin, and F. Durand, "A gentle introduction to bilateral filtering and its applications," in *ACM SIGGRAPH classes*, (Los Angeles,California), pp. 1–50, ACM, 2008.

[29] S. Bae, S. Paris, and F. Durand, "Two-scale tone management for photographic look," *ACM Trans. Graph.*, vol. 25, no. 3, pp. 637–645, 2006.

[30] E. P. Bennett and L. McMillan, "Fine feature preservation in HDR tone mapping," in *ACM SIGGRAPH Sketches*, (Boston, Massachusetts), p. 125, ACM, 2006.

[31] E. Bennett, J. Mason, and L. McMillan, "Multispectral Bilateral Video Fusion," *IEEE Transactions on Image Processing*, vol. 16, pp. 1185–1194, May 2007.

[32] R. Brinkmann, *The Art and Science of Digital Compositing*. Morgan Kaufmann Publishers, 2nd ed., 1999.

[33] J. Blinn, "Compositing. 1. Theory," *Computer Graphics and Applications, IEEE*, vol. 14, pp. 83–87, Sep. 1994.

[34] S. Raman and S. Chaudhuri, "Bilateral Filter Based Compositing for Variable Exposure Photography," in *Eurographics '09: Proceedings of the Eurographics Conference*, (Munich, Germany), Mar. 2009.

[35] Q. Yang, K.-H. Tan, and N. Ahuja, "Real-Time O(1) Bilateral Filtering," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, (Miami, Florida), Jun. 2009.

[36] C. Xydeas and V. Petrovic, "Objective image fusion performance measure," *Electronics Letters*, vol. 36, pp. 308–309, Feb. 2000.

[37] L. Wald, "Quality of High Resolution Synthesized Images: Is there a simple criterion?," in *Proceedings of International Conference on Fusion of Earth Data*, vol. 1, (Sophia Antipolis,France), pp. 99–105, Jan. 2000.

[38] Z. Wang and A. Bovik, "A Universal Image Quality Index," *IEEE Signal Processing Letters*, vol. 9, pp. 81–84, Mar. 2002.

[39] W. Cao, B. Li, and Y. Zhang, "A remote sensing image fusion method based on PCA transform and wavelet packet transform," in *Proceedings of the International Conference on Neural Networks and Signal Processing*, vol. 2, (Nanjing, China), pp. 976–981 Vol.2, Dec. 2003.