AN IMAGE FUSION APPROACH BASED ON FUZZY GIBBS RANDOM FIELDS WITH LOCAL LEVEL PROCESSING

[1]Mohit Nayak, [2]Krishna Kant Nayak
Department of Electronics & Communication,
Bansal Institute of Science & Technology, Bhopal, India
Mohitnayak1788@gmail.com

Abstract—**Fuzzy Gibbs Random Field (FGRF) models with local level processing is a powerful tools to model image characteristics accurately and been successfully applied to a large number of image processing applications. In This paper we investigate the problem of fusion of many types of images, e.g., multispectral image fusion, based on FGRF models with local level processing incorporates the contextual constraints via FGRF models with local level processing into the fusion model. This algorithm is applicable to both multiscale decomposition (MD)-based image fusion and non-MD-based image fusion. Experimental results are provided to demonstrate the improvement of fusion performance by our algorithms.**

Keywords — **Fuzzy Gibbs Random Field, local level processing, Multiresolution decomposition, Multispectral image fusion.**

I. INTRODUCTION

Image fusion is having importance in many image analysis works in which image data is obtained from multiple sources e.g. multi focus images, multi illuminated images, in medical field different types of images having very less part of important information can be fused with different category image to get more information from single fused image. As general the purpose of image fusion is to combine relevant information from two or more source images into one single image such that the single image contains as much information from all the source images as possible. The source images involved in such applications can be taken at different times and/or using different sensors. On the analysis of result, some source images may hiding some part of image and source images from different sensors show different physical features. Thus, the fused image is to have a more accurate description of the scene and is, therefore, more useful for human visual or machine perception [1].

In remote sensing applications, there have been a few studies on fusing high-resolution panchromatic images and low-resolution multispectral images to improve the spatial resolution [2], [3]. In this paper, we focus on the fusion of images having the same resolution, e.g., multispectral image fusion. A multispectral band covers only a narrow spectral range [3], and different bands represent different aspects of the scene. Multispectral image fusion involves the fusion of several bands in order to improve spectral resolution. The existing image fusion approaches can be classified into three categories: pixel-level, feature-level, and decision-level [4].This paper is focused on the pixel-level fusion approach.

Before image fusion, an image registration algorithm usually needs to be applied in order to align the source images [5].

The basic pixel-level fusion rule includes two steps:

1) Firstly, we need to find out whether a source image contributes to the fused image for each pixel.

2) Secondly, the intensity of the pixel in the fused image is obtained from all the contributing source images.

Among the pixel-level fusion rules, two traditional algorithms to fusion are to average the pixel intensities from all the source images or take the maximal pixel intensity among all the source images [25]. The averaging approach is effective in removing the Gaussian noise and increases the signal-to-noise ratio (SNR) but makes the image smoother and results in the loss of contrast information. The maximizing approach can produce the fused image at full contrast but is sensitive to sensor noise [7]. To overcome the limitations of the averaging and maximizing approaches, Sharma *et al.* [7] proposed a Bayesian image fusion approach and related it to local principal component analysis.

In recent years, multi scale decomposition (MD)-based techniques have been successfully applied to image fusion for different applications such as concealed weapon detection [8] and hyper spectral image fusion [9]. Different MD methods including pyramid transform and discrete wavelet transform have been applied to image fusion. The performances of these MD-based image fusion approaches are evaluated in [6] for a digital camera application. The MD-based image fusion approaches consist of three steps:

1) The source images are first decomposed into several scale levels using a pyramid transform or a wavelet transform.

2) Fusion is then applied at each level of the source images.
3) Finally, we invert the transform to synthesize the fused image [25].

The MD-based image fusion approach provides both spatial and frequency domain localization and achieves much better performance while the use of the transform increases the computational complexity [25]. So one can choose or not to employ transforms on images depending on different applications. For the MD-based fusion approaches, the basic fusion rule is applied to MD representations of the images at each resolution level. For the non-MD-based fusion approach,

the basic fusion rule is directly applied to the source images. Generally, the main drawback of the pixel-level fusion rule is that the decision on whether a source image contributes to the fused image is made pixel by pixel and, therefore, may cause spatial distortion in the fused image, which affects further processing, such as classification and detection. As we know, the pixels in an image are spatially correlated. Thus, for a source image, if one of its pixels contributes to the fused image, its neighbors are also likely to contribute to the fused image. It implies that the decision making during the first step of the fusion process should exploit the property of spatial correlation. Thus, it is important to incorporate spatial correlation into the fusion model, and the use of such a fusion model is expected to improve fusion performance [25].

A straightforward approach to make use of spatial correlation is to use a window- or region-based method [7], [10]–[13]. The idea is to estimate the intensity of a fused pixel from that of the source images in a small window. Yang and Blum [11] assumed that the decision making of pixels within a small window is constant and developed an expectation-maximization algorithm by employing a Gaussian mixture image model to adaptively find the fusion result. Burt and Kolczynski [10] proposed a weighted average algorithm to estimate the fused image in a pyramid transform domain. The weights are measured based on a local energy or variance (called "salience") within a small window. Lozci *et al.* [12] modified the weighted average algorithm by incorporating a generalized Gaussian statistical model. Lallier and Farooq [13] designed a weighted average scheme for fusing IR and visual images in a surveillance scenario. In their algorithm, larger weights are assigned to either the warmer or cooler pixels for the IR image and to the pixels having larger local variance for the visual image. The aforementioned algorithms [10]–[13] are used in the MD-based fusion approach.

However, the application of FGRF with local level processing models for pixel-level image fusion on images with the same resolution has not been considered. In this paper, we propose the fusion algorithm by incorporating the contextual constraints via FGRF with local level processing models into the fusion model. The algorithm models the decision making at the first step of the fusion rule as a FGRF, and then the algorithm models both the decision making and the true image as FGRFs. Also, the algorithm is applicable for both the MD-based fusion approach and the non-MD-based fusion approach.

This paper is organized as follows. In Section II, The Literature Review as formulate the image fusion problem based on a statistical model. Then, Work By Min Xu, Hao Chen And Pramod K. Varshney [25] the MRF-based image fusion approach is presented in Section III. In Section IV, present proposed fusion approach and section V Result shows the comparison with MRF fusion approaches via some experiments. Finally, some concluding remarks are provided in Section VI.

II. LITERATURE REVIEW

Image fusion is essentially an estimation problem. The objective is to estimate the underlying scene, assuming that each source image contains a good view of only part of the scene [1]. Blum [1] has proposed a statistical model for the image fusion problem. Assume that there are *N* source images to fuse. Each source image can be modeled as

$$y_i(\boldsymbol{r}) = H_i(\boldsymbol{r})\, x(\boldsymbol{r}) + w_i(\boldsymbol{r}), i = 1, \ldots, N \qquad (1)$$

Where, **r** indicates the spatial coordinates of a pixel, $y_i(\boldsymbol{r})$ is the intensity of the i^{th} source image at **r**, $x(\mathbf{r})$ is the intensity of the true scene at **r** to be estimated, $w_i(\boldsymbol{r})$ is the noise, and $H_i(\boldsymbol{r})$ is the sensor selectivity coefficient, taking on values from $\Theta = \{q1, q2, \ldots\}$ representing the percentage of the true scene contributing to the i^{th} source image [7]. In our work, we use $\Theta = \{0, 1\}$, which determines if the true scene contributes to the *i*th source image or not [1]. In the following, for simplicity of notation, "(**r**)" is omitted. Note that (1) represents the relationship between the source images and the true scene. According to this model, if the true scene contributes to the source image, the source image is modeled as a true scene plus a Gaussian noise. If the true scene does not contribute to the source image, the source image is modeled as Gaussian noise.

In practice, particularly in multiple sensor applications and multifocus applications, this model has some limitations. The source images obtained from different sensors sense different aspects of the true scene, and this model may be a coarse approximation in this case. The image fusion problem essentially involves the estimation of H_i and *x*. The two traditional algorithms, namely, the averaging and maximizing algorithms, can also be expressed using this model. For the averaging algorithm, $H_i = 1$ *for all i*. For the maximizing algorithm,

$$H_i = 1; i = max\ i\ \{yi\};\ H_i = 0, \text{otherwise.}$$

When H_i is given, the pixel intensity of the fused image can be easily calculated by a Least Squares (LS) technique as [16]

$$\hat{x} = (H^{\mathrm{T}}H)^{-1}H^{\mathrm{T}}Y \qquad (2)$$

Where *H* denotes the vector $[H_1, H_2 \ldots H_N]^{\mathrm{T}}$ and *Y* denotes the vector $[y_1, y_2, \ldots, y_N]^{\mathrm{T}}$. In practice, we only have the source images available without any prior information and the coefficient *H* is usually unknown. According to the LS technique, from the set of all possible values that the coefficient H can take, the one which produces the highest energy should be selected, i.e.,

$$\begin{aligned} \widehat{H} &= \min_H\{(Y - H_{\hat{x}})^T\,(Y - H_{\hat{x}})\} \\ &= \min_H\{Y^TY - (Y^TH)(H^TH)^{-1}((H^TY)\} \\ &= \max_H\{(Y^TH)(H^TH)^{-1}((H^TY)\} \end{aligned} \qquad (3)$$

Note, since $H_i \in \{0, 1\}$ H has 2^N possible values. Once H is available, the intensity of the fused image at pixel r i.e., is obtained by an LS approach [16], which is

$$\hat{x} = (\hat{H}^T \hat{H})^{-1} \hat{H}^T Y. \quad (4)$$

In the aforementioned model, both the coefficient *H* and the intensity of the fused image *x* at each pixel are estimated pixel by pixel, and therefore, it is very sensitive to sensor noise. Furthermore, since the estimation of the fused image is based on the estimation of the coefficients, the estimation of the coefficient *H* plays an important role in the fusion process. The estimation accuracy of the coefficients directly influences the estimation of the fused image. Since the coefficient *H* of a pixel is likely to be similar to the coefficients corresponding to other pixels in its neighborhood due to spatial correlation, we can get better estimates of *H* by utilizing spatial correlation.

A straightforward and simple approach is to assume that the coefficients of pixels within a small window are constant and then select the coefficient which produces the highest energy of pixels within a small window. This strategy has been used in [11]. However, the goal of the LS approach is to minimize the data error $||y - \hat{y}||^2$, which does not necessarily lead to a small estimation error for either *H* or *x*. A popular strategy for improving the estimation error of LS is to incorporate prior information on *H* or *x* [17]. Motivated by this fact and the fact that the MRF model for local level processing in the form of prior Gibbs distributions is currently the most effective way to describe the local behavior of both the intensity field and the discontinuity field [20].

We propose to employ an FGRF model to estimate the coefficients at Global level and then MRF model for local level processing. It is expected to improve the estimation accuracy of the coefficients *H*, thereby leading to improved fusion results.

II.I MRF Model for Image Fusion

The image fusion problem is to estimate the true scene *x*. However, before the estimation of *x*, there is a need to obtain an accurate estimate of *H*, which represents the decision whether the true scene is present in the source image, i.e., whether the source image contributes to the fused image. In the previous section, we considered the estimation of *x* and *H* on a pixel level. In this section, there are two MRF-based image fusion approaches, which design the estimator by incorporating the spatial correlation through the prior probability density function (pdf) of *H* and *x*. The intensity of a fused pixel then depends not only on the intensities of the pixel in the source images but also on that of the neighboring pixels [25]. In the first algorithm, only the coefficients are modeled using an MRF, denoted as MRF_H. In the second algorithm, both the coefficients and the fused image are modeled using MRFs, denoted as MRF_HX.

Some notations used in this paper are listed as follows:

- X: the whole true scene (fused image);
- H_i: the coefficients of the *i*th source image;
- Y_i: the intensities of the *i*th source image;
- H: the coefficients of source images, where H(r, i) = H_i;
- Y: the intensities of source images, where Y(r, i) = Y_i.

The maximum $a\ posteriori$ (MAP) criterion is used to find the optimal solution for the estimation problem. The estimation procedure based on the MAP criterion chooses the most likely values of coefficients and the fused image among all possible values given the observed source images. The resulting probability of error is minimal among all other detectors [16]. This criterion is expressed as

$$\{\hat{H}, \hat{X}\} = \arg\{\max_{H,X}[P(X, H|Y)]\}. \quad (5)$$

However, due to high computational complexity, it is difficult to directly obtain the final solution. Note that $P(X, H|Y) = P(X|H, Y)P(H|Y) = P(H|X, Y)P(X|Y)$. Thus, a suboptimal method is adopted by MRF model. To decompose the problem (5) into two sub problems and iteratively solve the two sub problems

$$\{\hat{H}^{n+1} = \arg\{\max_{H}[P(H|Y, \hat{X}^n)]\}.$$
$$\{\hat{X}^{n+1} = \arg\{\max_{X}[P(X|Y, \hat{H}^n)]\}. \quad (6)$$

Where $\hat{H}^n$ denotes the *n*th update of the estimate of H and $\hat{X}^n$ denotes the *n*th update of the estimate of X. The below equation shows iteratively update estimates of H and X as:

$$P(X^{n+1}, H^{n+1}|Y) \geq P(X^n, H^n|Y) \quad (7)$$

And, therefore, achieve the optimum at the end.

II.II Fusion Approach: MRF Modeling for Coefficients H (MRF_H)

Motivated by the fact that the coefficients of the source images exhibit spatial correlation, the model shows the coefficient H by an MRF model. Let *S* be a set of sites in an image and $\Lambda \in \{0, 1, \ldots, L-1\}$ be the phase space. We assume that the coefficients $H(S) \in \Lambda^S$ follow MRF properties with the Gibbs potential $U_c(H)$. The marginal pdf for H is written as [14]

$$P_H(H) = \frac{1}{Z_H} exp\left[\frac{-1}{T}\sum_{c \subset S} U_c(H)\right] \quad (8)$$

Where $Z_{\mathbf{H}}$ is normalization constant given by

$$Z_H = \sum_{H \in \Lambda^s} \exp\left[\frac{-1}{T}\sum_{c \subset S} U_c(H)\right]. \quad (9)$$

The estimate of H is given by

$$\hat{H}^{n+1} = \arg\{\max_{H}[P(H|Y, \hat{X}^n)]\}. \quad (10)$$

We apply Bays' rule, which provides the following result:

$$P\left(H\middle|Y, \hat{X}^n\right) = \frac{P(H, Y|\hat{X}^n)}{P(Y|\hat{X}^n)} = \frac{P(Y|H, \hat{X}^n)P(H)}{P(Y|\hat{X}^n)} \quad (11)$$

And because $P(Y|\hat{X})$ is a constant for all the values of H, (10) can be written as

$$\widehat{H}^{n+1} = \arg\{\max_H [P(Y|H,\widehat{X}^n)P(H)]\}. \quad (12)$$

In the model given in (1), the noise of each source pixel is assumed to be an independent and identically distributed $(i.i.d.)$ Gaussian noise with zero mean and variance of σ^2, and therefore, the conditional pdf of the source image **Y** given **H** and $\widehat{X}^n$ is given by

$$P(Y|H,\widehat{X}^n) = \frac{\exp\left[-\frac{\sum_i (Y_{i-} H_i\widehat{X}^n)^T (Y_{i-} H_i\widehat{X}^n)}{2\sigma^2}\right]}{(2\pi\sigma^2)^{\frac{M}{2}}} \quad (13)$$

Where variable M represents total number of pixels. Then, substituting (8) and (13) into (12) and taking the constant term out, we obtain

$$\widehat{H}^{n+1} = \arg\{\max_H [\exp(-E(H))]\}. \quad (14)$$

Where

$$\mathrm{E(H)} = \frac{\sum_i (\mathrm{Y_{i-}\ H_i\widehat{X}^n})^{\mathrm{T}} (\mathrm{Y_{i-}\ H_i\widehat{X}^n})}{2\sigma^2} + \sum_{c\subset S} \mathrm{U_c\,(H)}. \quad (15)$$

According to the aforementioned result, we observe that maximization in (14) is equivalent to minimization of E (**H**). Thus, the optimal estimate for H can be expressed as

$$\widehat{H}^{n+1} = \arg\{\min_H (E(H))\}. \quad (16)$$

Note that, for two source images with size $300 * 300$, H has a total of 490000 possible configurations. Thus, in practice, due to the large search space on H, the solution of (16) cannot be obtained directly, and therefore, the simulated annealing (SA) algorithm [18] is applied here to search for the optimal solution of (16). The solution for the second sub-problem, i.e., the estimate for X, is obtained by (4).

The iterative algorithm is n described in terms of the following steps:

1) Start with an initial estimate of H and X. Estimate the initial parameters (noise variance and some parameters in the pdf of H) and set the initial temperature.

2) At each iteration, obtain a new estimate of H based on the Gibbs pdf given in (8) with the Gibbs potential *E (H)* using a Gibbs sampling procedure [14].
3) Update the fused image using (4).
4) Reduce the temperature using a predetermined schedule and repeat 2) and 3) until convergence.

Here, the temperature is a parameter which is used to control the randomness of the coefficient generator, and we consider that the algorithm converges when the two consecutive updates are within tolerance of each other. At steps 2) and 3), we visit each pixel from left to right and from top to bottom when we update the coefficients and the fused image. Eventually, the resulting coefficient will converge to the solution of (16), and the fused image is simultaneously obtained. Compared with the maximizing approach, the averaging approach, and the LS approach, the solution of this algorithm is obtained through an optimization algorithm, and therefore, it increases the computation time.

II.III Fusion Approach: MRF Modeling for Coefficients H and Fused Image X (MRF_HX)

In the aforementioned algorithm, the assumption is that the coefficients H follow an MRF model. Then, the intensity of the fused pixel is estimated by an LS technique. In practice, the fused image also has the property of high spatial correlation. Thus, assume that the fused image also follows an MRF model with a Gibbs potential $Vc(X)$. Hence, the marginal pdf for X is written as [14]

$$P_X(X) = \frac{1}{Z_X} exp\left[\frac{-1}{T}\sum_{c\subset S} U_c(X)\right] \quad (17)$$

Where Z_X is normalization constant given by

$$Z_X = \sum_X exp\left[-\frac{1}{T}\sum_{c\subset S} V_c(X)\right] \quad (18)$$

Under this assumption, the MAP criterion to obtain the optimal X is written as

$$\widehat{X}^{n+1} = \arg\{\max_X [P(X|Y,\widehat{H}^n)]\}. \quad (19)$$

in a similar manner as for the estimation of $\widehat{H}$, (19) reduces to

$$\widehat{X}^{n+1} = \arg\{\min_X (\Delta(X))\}. \quad (20)$$

Where

$$\Delta(X) = \frac{1}{2\sigma^2}\sum_i \left(Y_{i-} \widehat{H_i}^n X\right)^T (Y_{i-} \widehat{H_i}^n) + \sum_{c\subset S} V_c(X) \quad (21)$$

II.IV Extension to the MD-Based Fusion Framework

Here, the applicability of the two proposed algorithms to the MD-based fusion approach is discussed. For the non-MD based fusion approach, Y and H in the data model (1) denote the intensities of the source images and their corresponding coefficients and X denotes the intensity of the fused image. This data model (1) is also applicable for the data after the MD process [1], [7]. Thus, if the MD-based fusion approach is employed, we assume that the MD transform is applied to the source images, Y refers to the values of the MD representations of the source images at some resolution level, H refers to the corresponding coefficients, and X refers to the values of the MD representations of the fused image at the same resolution level. Thus, instead of directly applying the image fusion model (1) on the source images, one can perform the MD on the source images and then apply the image fusion model (1) on the MD representations at each resolution level. By using multi-resolution transforms such as discrete wavelet transform, the source image is decomposed into different frequency bands,

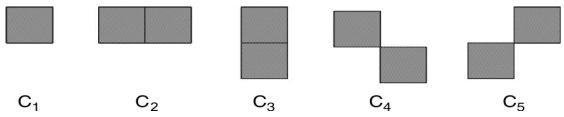


Figure1 Cliques considered in the eight-neighborhood system [25].

Which makes the model (1) more closely fit the MD representations. However, the use of multiresolution transforms increases the complexity of the algorithm. It is noted that, since the multiresolution transform may result in the loss of locality in MRF models for the MD image representations [28], i.e., the local MRF property may not hold on X, it is not suggested to use the algorithm MRF_HX with the MD-based fusion approach. The coefficient H for each pixel represents whether the true scene X (fused image) contributes to the source image. Pixels in a large area may all contribute to the true scene; however, all the pixels in the area may not contain the same intensities. Thus, the coefficient H has more spatial correlation over a larger area than the intensity of the true scene X. After MD transformation, coefficients H may still exhibit spatial correlation while MRF property may not hold for X. Thus, only the algorithm MRF_H is applied in the MD-based fusion approach. In the next section, the proposed algorithm is discussed.

III. PROPOSED ALGORITHM

Before The Fuzzy Gibbs Random Field provides the global variable processing to overcome the redundant information field from the source image firstly and provide the space for more informative field to spread over the entire range. In this way FGRF model is a useful tool in image processing and after that we apply the MRF model for local level processing this combination provides the better result as compared to the MRF based image fusion model. Here first we describe the Fuzzy Gibbs Random Field. Than we use this FGRF model for image processing. After applying the FGRF model we apply The MRF model of image fusion and compare the results for both the methods.

A set of random variables F is said to be a Fuzzy Gibbs Random Field (FGRF) on S with respect to N if and only if its configurations obey a Gibbs distribution A Gibbs distribution takes the following form

$$P(f) = Z^{-1}\exp\left(\frac{-1}{T}U(f)\right) \quad (22)$$

Where

$$Z = \sum_{f\in F}\exp\left(\frac{-1}{T}U(f)\right) \quad (23)$$

is a normalizing constant called the partition function, **T** is a constant called the temperature which shall be assumed to be 1 unless otherwise stated, and $U(f)$ is the energy function. The energy

$$U(f) = \sum_{c\in C} V_c(f) \quad (24)$$

is a sum of clique potentials $V_c(f)$ over all possible cliques C. The value of $V_c(f)$ depends on the local configuration on the clique **c**. obviously, the Gaussian distribution is a special member of this Gibbs distribution family. A FGRF is said to be homogeneous if $V_c(f)$ is independent of the relative position of the clique **c** in **S**. It is said to be isotropic if V_c is independent of the orientation of c. It is considerably simpler to specify a FGRF distribution if it is homogeneous or isotropic than one without such properties. The homogeneity is assumed in most MRF vision models for mathematical and computational convenience. The isotropy is a property of direction-independent blob-like regions.

To calculate a Gibbs distribution, it is necessary to evaluate the partition function Z which is the sum over all possible configurations in F. Since there are a combinatorial number of elements in F for a discrete L, the evaluation is prohibitive even for problems of moderate sizes. Several Approximation methods exist for solving this problem. $P(f)$ Measures the probability of the occurrence of a particular configuration, or ``pattern'', f. The more probable configurations are those with lower energies. The temperature **T** controls the sharpness of the distribution. When the temperature is high, all configurations tend to be equally distributed. Near the zero temperature, the distribution concentrates around the global energy minima. Given **T** and $U(f)$, we can generate a class of ``patterns'' by sampling the configuration space F according to $P(f)$. For discrete labelling problems, a clique potential $V_c(f)$ can be specified by a number of parameters. For example, letting $f_c = (f_i f_{i'} f_{i''})$ be the local configuration on a triple-clique $c = (i', i', i'')$, f_c takes a finite number of states and therefore $V_c(f)$ takes a finite number of values. The continuous labelling problems f_c can vary continuously. In this case, $V_c(f)$ is a (possibly piecewise) continuous function of f_c .

Sometimes, it may be convenient to express the energy of a Gibbs distribution as the sum of several terms, each ascribed to cliques of a certain size, that is,

$$U(f) = \sum_{[i]\in C_1} V_1(f_i) + \sum_{[i,i\prime]\in C_2} V_1(f_i, f_{i\prime}) + \sum_{[i,i',i'']\in C_3} V_1(f_i, f_{i\prime}, f_{i''}) + . \quad (25)$$

The above implies a homogeneous Gibbs distribution because V_1, V_2 and V_3 are independent of the locations of $i,\ i'\ and\ i''$ for non-homogeneous Gibbs distributions, the clique functions should be written as $V_1\ (i, f_i)$, $V_2\ (i, i', f_i), f_{i\prime})$ and so on.

An important special case is when only cliques of size up to two are considered. In this case, the energy can also be written as

$$U(f) = \sum_{i\in S} V_1(f_i) + \sum_{i\in S}\sum_{i\prime\in N_i} V_2(f_i, f_{i\prime}) \quad (26)$$

Note that in the second term on the RHS, {i, i'} and {i', i} are two distinct cliques in C_2 because the sites in a clique are ordered. The Conditional probability can be written as

$$P\left(f_i \middle| f_{N_i}\right) = \frac{e^{-\left[V_1(f_i)+\sum_{i\prime\in N_i} V_2(f_i, f_{i\prime})\right]}}{\sum_{f_i\in L} e^{-\left[V_1(f_i)+\sum_{i\prime\in N_i} V_2(f_i, f_{i\prime})\right]}} \quad (27)$$

III.I Choice of MRF Models for local level processing

We provide the examples to evaluate the fusion performance of our fusion algorithms. For the MRF-based fusion algorithms, MRF_HX, used in the experiments, we consider

five clique types in the eight neighborhood system: *C*1, *C*2, *C*3, *C*4, *C*5, associated with the singleton, vertical pairs, horizontal pairs, left-diagonal pairs, and right-diagonal pairs, respectively. They are shown in Fig. 1. The Gibbs energy function of the coefficient of the source image is defined by an auto logistic function, given by [14]

$$\sum_{c\subset S} U_c(H) = a^T L \quad (28)$$

where a= $[a_2, \dots\dots a_5]$ is the paramter vector of the coefficient H and

$$L = \left[\sum_{(s,t)\in C_2} I[H(s),H(t)], \dots, \sum_{(s,t)\in C_5} I[H(s),H(t)]\right] \quad (29)$$

$$I(a,b) = -1, \text{ if a=b} \quad (30)$$

I(a,b) = 1, otherwise.

Due to its simplicity, this class of MRF model has been extensively used in [14], [20]–[22] for modelling a wide variety of images, both as region formation models and as texture models. Furthermore, we use a Gaussian MRF model to represent the Gibbs energy function of the fused image, which is given by

$$\sum_{(c\in S} V_c(X) = \frac{1}{2}(X(s) - P^T G)^2 \quad (31)$$

Where **p** = $[p2, \dots, p5]$ is the parameter vector of the image model and its potential vector **G** is defined as

$$G = \left[\sum_{(s,t)\in C_2} X(t), \dots, \sum_{(s,t)\in C_5} X(t)\right]. \quad (32)$$

For simplicity, we choose **p** = [0.333, 0.333, 0.1667, 0.1667] in our experiments. The Gaussian MRF model is widely used for modeling image texture [14]. Under this model, the analytical solution for (20) can be easily derived by:

$$\frac{\partial \Delta(X)}{\partial X} = 0. \quad (33)$$

Substituting (13) and (25) into (27), it yields

$$\hat{X}^{n+1} = \left(1 + \sum_i H^{n}{}_i^T H_i^n/\sigma^2\right)^{-1}\left(P^T G + \sum_i H^{n}{}_i^T Y_i/\sigma^2\right) \quad (34)$$

The estimate given by (28) for one pixel involves vector multiplication, which has the computational complexity O(N). Thus, the estimation of the whole fused image has the computational complexity $O(M * N)$.

III.II Parameter estimation

Modelling the Markov pdf parametrically involves the data driven optimal estimation of the parameters associated with the potential functionsV_c. The model parameters must be estimated for each data set as part of the image processing algorithm. In our algorithms, the noise variance $\sigma 2$ in (13) and the parameter **a** in the coefficient MRF pdf in (22) are unknown. Thus, we need to estimate these parameters in our algorithms. Because we assume that the noise in the fusion model is a Gaussian noise, it is straightforward to estimate the noise variance by the maximum likelihood (ML) criterion. It is given by

$$\sigma^2 = \quad argmax P(Y|H,X,\sigma^2)$$

$$= \frac{1}{MN}\sum_i (Y_{i-} H_i X)^T (Y_{i-} H_i X). \quad (35)$$

The direct ML estimation of the parameters associated with the pdf of H is known to be a difficult problem [32]. The ML estimate of **a** is

$$\hat{a} = arg\max_a(H,a) = \arg\min_a V_c(H,a) - lnZ_H \quad (36)$$

The potential function $Vc\,(H,a)$ can be simply computed. However, the normalization term Z_{H} involves a summation over all possible configurations of **H**, which is practically impossible due to the large computation time. Note that, for two source images with size 300 * 300, **H** has a total of 490000 possible configurations. An alternative method for approximation to ML estimation is maximum pseudo likelihood (MPL) estimation, which was proposed by Besag [23]. The MPL estimation method is a suboptimal method, which is given by

$$\hat{a} = arg\max_a \prod_s P(H(s),a)$$
$$= arg\min_a \sum_s V_c(H(s),a) - lnZ_{H(s)}. \quad (37)$$

The differences among the fused results are usually difficult to be measured only based on observation, particularly when the fused images are multiband. Objective and quantitative analysis can benefit to a comprehensive evaluation. Various image quality indices have been developed for the purpose of image fusion [28]–[21]. Some of these indices validate the spatial resolution, while others focus on the spectral properties of the obtained fused result. In this paper, we employ three such indices.

1) SNR: The SNR in decibels, is a direct index to compare the fused image to the reference one [26]. For multiband images, it can be calculated band-by-band and also globally averaged
$$SNR\left(Z,\hat{Z}\right) = 10log_{10}\frac{\sum Z^2}{\sum(Z-\hat{Z})^2} \quad (38)$$
2) Spectral angle Mapper (SAM): The averaged SAM is used as a measurement of spectral distortion between the fused and reference HS images [26]. The SAM of two spectral vectors (**Z***n* and ˆ**Z***n*) at a given spatial position *n* is defined as
$$SAM\left(Z_n,\hat{Z}_n\right) = \arccos\left(\frac{\langle Z_n,\hat{Z}_n\rangle}{||Z_n||_2\cdot||\hat{Z}_n||_2}\right) \quad (39)$$

It can be measured in either degrees or radians. This value is averaged over all pixels of the image. The smaller the value of SAM, lesser is the spectral distortion. In this paper, we measure the SAM in degrees.

1) Universal image quality index (UIQI): A UIQI [22] has been widely used for image similarity evaluation and was also applied to validate fusion techniques [21]. UIQI of two images (*A* and *B*) is defined as
$$Q = \frac{4\sigma_{AB}\mu_A\mu_B}{(\sigma_A^2+\sigma_B^2)(\mu_A^2+\mu_B^2)} = \frac{\sigma_{AB}}{\sigma_A\sigma_B}\cdot\frac{2\mu_A\mu_B}{\mu_A^2+\mu_B^2}\cdot\frac{2\sigma_A\sigma_B}{\sigma_A^2+\sigma_B^2}. \quad (40)$$

This quality index models any distortion as a combination of three different factors: loss of correlation, luminance distortion, and contrast distortion. The dynamic range of Q is $[-1, 1]$, and the best value 1 is obtained if $A = B$. When

applying this index to a multiband image, it is applied band-by-band and averaged over all bands. [26]

III.III Match and silence measure

In pattern-selective fusion the composite image is assembled from selected component patterns of the source images. In the pyramid implementation, the pyramid basis functions serve as the component patterns. Here we define two distinct modes of combination: selection and averaging. At sample locations where the source images are distinctly different, the combination process selects the most salient component pattern from the source pyramids and copies it to the composite pyramid, while discarding less salient patterns. But at sample locations where the source images are similar, the process averages the source patterns. Again, selection avoids double exposure artefacts in the composite. Averaging reduces noise and provides stability where source images contain the same pattern information. Gradient pyramid basis functions. Pattern selective image fusion is guided by two measures: a match measure that determines the mode of combination at each sample position (selection or averaging), and salience measures that determine which source pattern is chosen in the selection mode.

Salience measure: The salience of a particular component pattern is high if that pattern plays a role in representing important information in a scene. Salience is low if the pattern represents unimportant information, or, particularly, if it represents corrupted image data. Specific measures of salience may be based on criteria for the vision task at hand. In general a pattern may be expected to be important if it is relatively prominent in the image. Thus the amplitude of a pattern can be taken as a generic measure of its salience. Alternatively, the contrast of the component pattern with neighbouring patterns can provide that measure [27]. We define salience at sample $\hat{m}$ as a local energy, or variance, within neighborhood p:

$$S_I(m,n,k,l) = \sum_{\hat{m},\hat{n}} p(\hat{m},\hat{n}) D_I(m+\hat{m}, n+\hat{n}, k, l)^2 \quad (41)$$

In practice the neighborhood p is small, typically including only the sample itself (point case) or a **3** by **3** or 5 by 5 array of samples centred on the sample (area case).

Match measure: The match measure is used to determine which of the two combination modes to use at each sample position, selection or averaging. The relative amplitudes of corresponding patterns in the two source pyramids can be used as a measure of their similarity, or match. Alternatively, the correlation between pyramids in the neighborhood of the source patterns can provide that measure [27]. Here we define the match at sample $\hat{m}$ as a local normalized correlation within neighborhood p:

$$M_{AB(\vec{m})} = \frac{2\sum_{\acute{m},\acute{n}} p(\acute{m},\acute{n}) D_A(\vec{m}+\acute{m}\acute{n}) D(\vec{m}+\acute{m}\acute{n})}{S_A(\vec{m})+S_B(\vec{m})} \quad (42)$$

Again, the neighborhood p may include only the given component pattern (point case) or it may include a local array **of** components (area case). MAE has value 1 for identical patterns, value -1 for patterns that are identical except that they have opposite sign, and a value between -1 and 1 for all other patters. Unlike the usual definition of normalized correlation, it has a value less than 1 for patterns that are identical except for a scale factor. In the next section, some examples are provided for illustration.

IV. RESULT

We have presented a general approach to image fusion and have shown that it can be applied to diverse fusion tasks. Fusion is performed in a pyramid transform domain. The implementation described here extends prior algorithms in two important respects. First, a measure of pattern match has been introduced to control the mode used in image combination, selection or averaging. Second, both this match measure and the salience measures of past implementations are now defined as functions of the neighborhood of each pyramid sample rather than functions of only the sample itself. The neighborhood size can be small, 3 by 3 in the examples shown here. These modifications address problems that are encountered with past implementations of pyramid-based fusion.

Example-1: MULTI EXPOSURE IMAGE FUSION:

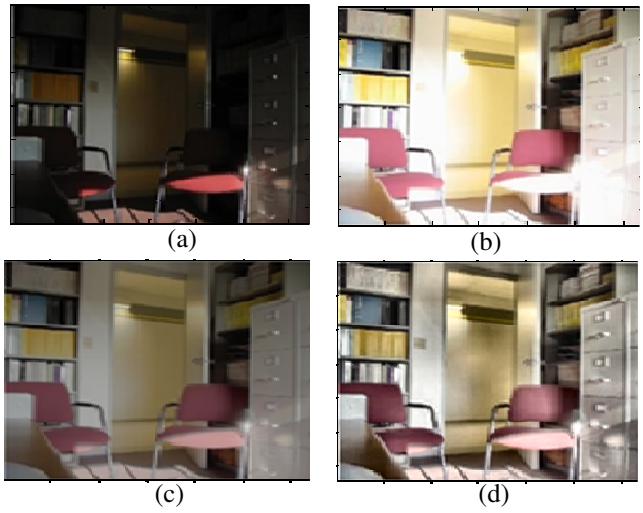
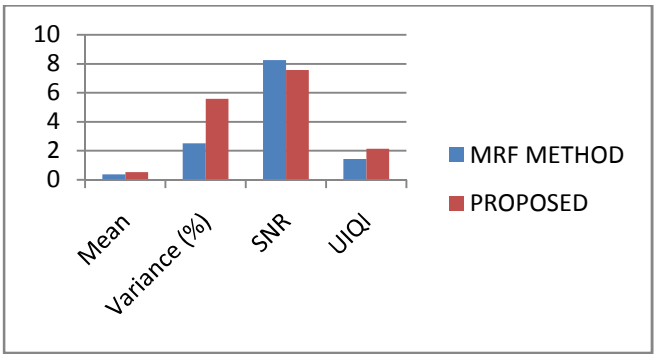


Figure: 2. (a) Original Image of Office with less exposure (b) Original image of Office with High Exposure (c) Fused Image by MRF method (d) Fused Image by Proposed Method.

TABLE-1: FOR MULTI EXPOSURE IMAGES FUSION

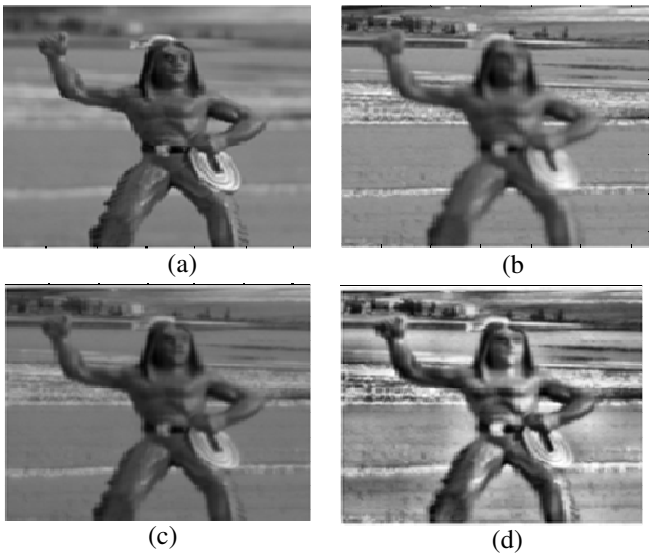
Method	Mean	Variance (%)	SNR	UIQI
MRF METHOD	0.3781	2.51	8.2547	1.4299
PROPOSED METHOD	0.5223	5.59	7.5762	2.1336



Graph-1: Graph Shows parameter for image of Office

In particular they provide greater shift invariance and immunity to video noise, and they provide at least a partial solution to the problem of combining components that have roughly equal salience but opposite contrast: mismatched patterns always are handled through selection, never averaging. The fusion algorithm was found to perform well (based on visual inspection) for a range of tasks without requiring adjustment of the algorithm parameters. In fact, results were remarkably insensitive to changes in these image fused by using MRF model and our Proposed FGRF with images, medical field images by different sensors and also take general images to show the fusion mechanism. This comparison is showing the results give better visual quality in our proposed model.

Example-2: MULTI FOCUS IMAGE FUSION



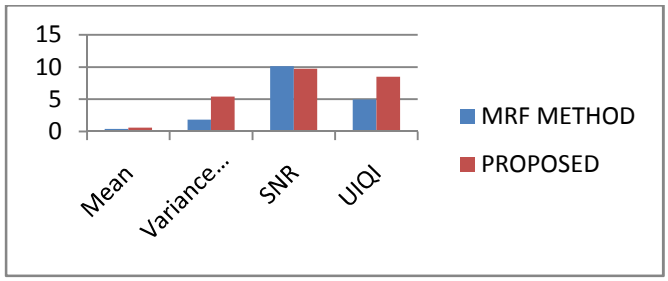
(a) (b

(c) (d)

Figure: 3. (a) Original Image with Focus on human body (b) Original image with Focus on background (c) Fused Image by MRF method (d) Fused Image by Proposed Method.

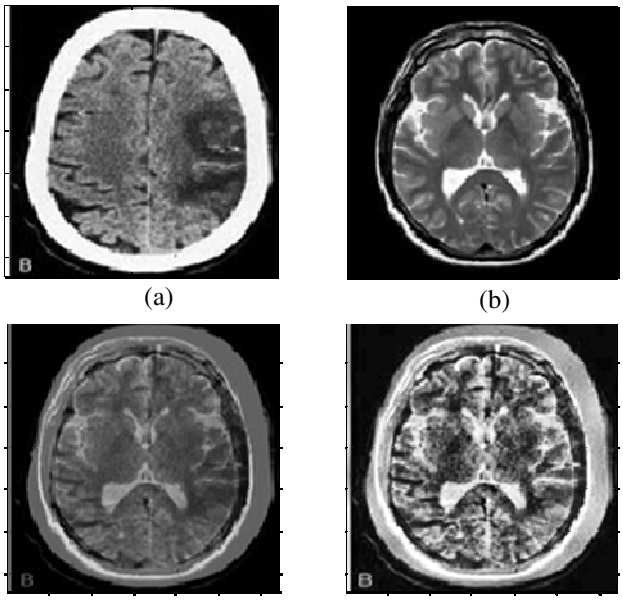
TABLE-2: FOR MULTI EXPOSURE IMAGE FOCUSED IMAGES

Method	Mean	Variance (%)	SNR	UIQI
MRF METHOD	0.4142	1.84	10.1351	4.9480
PROPOSED	0.5775	5.41	9.7291	8.4762



Graph-2: Graph Shows parameter variation for image of Human body

Example-3: CT SCANE AND MRI IMAGE FUSION



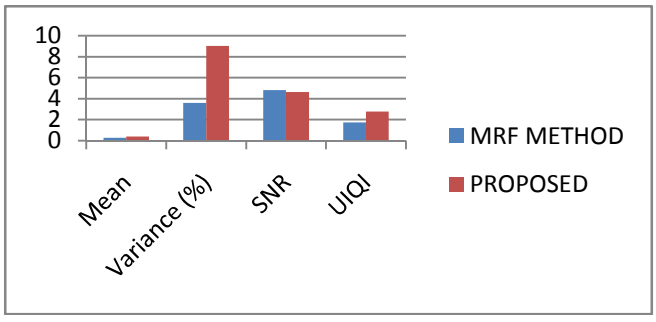
(a) (b)

(c) (d)

Figure: 4. (a) Original CT Scan Image of human Brain top view (b) Original MRI Image of human Brain (same part) top view (c) Fused Image by MRF method (d) Fused Image by Proposed Method.

TABLE-3: FOR CT SCANE AND MRI IMAGE FUSION

Method	Mean	Variance (%)	SNR	UIQI
MRF METHOD	0.2713	3.61	4.8277	1.7528
PROPOSED	0.4089	9.01	4.6262	2.7690



Graph-3: Graph Shows parameter variation for Image of human Brain

Example-4: DIFFERENT IMAGE FUSION CONTAINING DIFFERENT INFORMATION

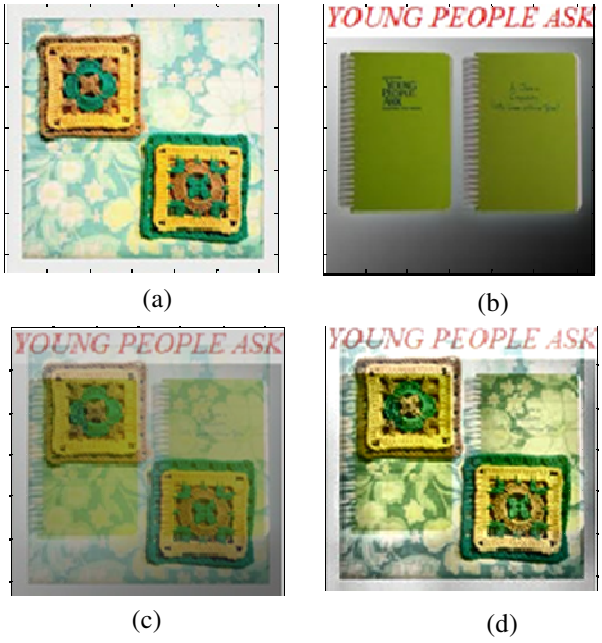
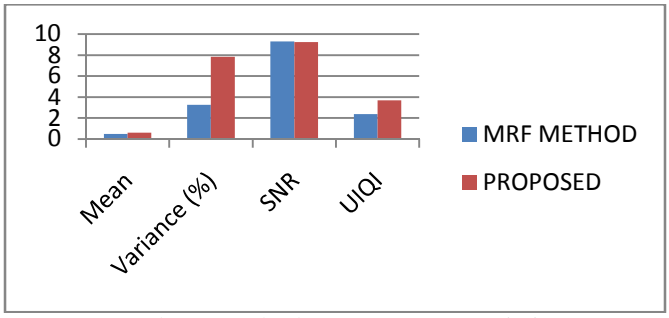


Figure: 5. (a) Original Image of carpet (b) Original Image of files (c) Fused Image by MRF method (d) Fused Image by Proposed Method.

TABLE-4: For different image fusion containing different information.

Method	Mean	Variance (%)	SNR	UIQI
MRF METHOD	0.4933	3.25	9.2891	2.3730
PROPOSED	0.6107	7.85	9.2463	3.6888



Graph-4: Graph Shows parameter variation

V. CONCLUSION

In this paper, we have studied the image fusion problem based on a statistical model. We utilized the fact that decision making in the fusion process has significant correlation within its neighborhood and assumed that it can be modelled as a FGRF with local level processing. Based on that, a new statistical fusion algorithm, namely, FGRF_HX has been proposed. This approach is applicable for MD-based fusion approaches. In particular, when the raw source images are directly used for fusion without pre-processing, the fused image can also be modelled as an FGRF with local level processing, and then, the fusion result can be obtained using the MAP criterion incorporating the *a priori* Gibbs distribution of the fused image. Visual inspection and quantitative performance evaluation both demonstrate that the employment of the FGRF with local level processing model in the fusion approaches resulted in a better fusion performance than the traditional fusion approaches. In our proposed image fusion algorithms, we assumed a simple relationship between each source image and the true scene, i.e., a source image either contributes to the fused image or does not contribute to the fused image. Thus, it results in a mismatch between the fusion model and the real image data set. To improve this, one can assume that the coefficient in the data model can take any real value, which may increase the accuracy of the fusion algorithms [25]. In addition, in the developed image fusion algorithms, we assumed that the noise in the source image is an $i.i.d.$ Gaussian noise. Since this is a rather limiting assumption, if we can build the noise model to include non-Gaussian distortion or possibly correlated Gaussian mixture distortion, this model should be closer to realistic sensor images and the estimation of fused image may improve. However, the FGRF with local level processing modelling of the coefficient in the image model is a good model to describe the fusion process, which improves the fusion performance. In recent years, other optimization algorithms such as the graph-cut-based approach [19] have become very popular, and they can find the solution in a more computationally efficient manner than the SA optimization algorithm. The use of optimization algorithms such as the graph-cut-based optimization approach will be investigated in the future to improve the efficiency of the fusion algorithm.

REFERENCES

[1] R. S. Blum, "On multisensor image fusion performance limits from an estimation theory perspective," *Inf. Fusion*, vol. 7, no. 3, pp. 250–263, Sep. 2006.

[2] Z. Wang, D. Ziou, C. Armenakis, D. Li, and Q. Li, "A comparative analysis of image fusion methods," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 6, pp. 1391–1402, Jun. 2005.

[3] C. Thomas, T. Ranchin, L. Wald, and J. Chanussot, "Synthesis of multispectral images to high spatial resolution: A critical review of fusion methods based on remote sensing physics," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 5, pp. 1301–1312, May 2008.

[4] C. Pohl and J. van Genderen, "Multisensor image fusion in remote sensing: Concepts, methods, and applications," *Int. J. Remote Sens.*, vol. 19, no. 5, pp. 823–854, 1998.

[5] P. K. Varshney, B. Kumar, M. Xu, A. Drozd, and I. Kasperovich, "Image registration: A tutorial," in *Proc. NATO ASI*, Albena, Bulgaria, 2005.

[6] Z. Zhang and R. S. Blum, "A categorization of multiscale-decompositionbased image fusion schemes with a performance study for a digital camera application," *Proc. IEEE*, vol. 87, no. 8, pp. 1315–1326, Aug. 1999.

[7] R. K. Sharma, T. K. Leen, and M. Pavel, "Probabilistic image sensor fusion," in *Proc. Adv. Neural Inf. Process. Syst. 11*, 1999, pp. 824–830.

[8] H.-M. Chen, S. Lee, R. Rao, M.-A. Slamani, and P. Varshney, "Imaging for concealed weapon detection: A tutorial overview of development in imaging sensors and processing," *IEEE Signal Process. Mag.*, vol. 22, no. 2, pp. 52–61, Mar. 2005.

[9] Y. Zhang, S. De Backer, and P. Scheunders, "Noise-resistant waveletbased Bayesian fusion of multispectral and hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 11, pp. 3834–3843, Nov. 2009.

[10] P. Burt and R. Kolczynski, "Enhanced image capture through fusion," in *Proc. 4th Int. Conf. Comput. Vis.*, 1993, pp. 173–182.

[11] J. Yang and R. Blum, "A statistical signal processing approach to image fusion for concealed weapon detection," in *Proc. IEEE Int. Conf. Image Process.*, 2002, pp. 513–516.

[12] A. Lozci, A. Achim, D. Bull, and N. Canagarajah, "Statistical image fusion with generalized Gaussian and Alpha-Stable distributions," in *Proc. 15th Int. Conf. Digital Signal Process.*, 2007, pp. 268–271.

[13] E. Lallier and M. Farooq, "A real time pixel-level based image fusion via adaptive weight averaging," in *Proc. 3rd Int. Conf. Inf. Fusion*, 2000, pp. WEC3/3–WEC313.

[14] S. Z. Li, *Markov Random Field Modeling in Computer Vision*. New York: Spinger-Verlag, 2001.

[15] L. Bedini, A. Tonazzini, and S. Minutoli, "Unsupervised edge-preserving image restoration via a saddle point approximation," *Image Vis. Comput.*, vol. 17, no. 11, pp. 779–793, Sep. 1999.

[16] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Upper Saddle River, NJ: Prentice-Hall, 1993.

[17] Y. C. Eldar, A. Beck, and M. Teboulle, "Bounded error estimation: A Chebyshev center approach," in *Proc. 2nd IEEE Int. Workshop Comput. Adv. Multi-Sensor Adapt. Process.*, 2007, pp. 205–208.

[18] S. Lakshmanan and H. Derin, "Simultaneous parameter estimation and segmentation of Fuzzy Gibbs Random Fields using simulated annealing," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, no. 8, pp. 799–813, Aug. 1989.

[19] V. Kolmogorov and R. Zabin, "What energy functions can be minimized via graph cuts?" *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 2, pp. 147–159, Feb. 2004.

[20] P. Bremaud, *Markov Chains, Gibbs Fields, Monte Carlo Simulation, and Queues*. New York: Springer-Verlag, 1999.

[21] H. Derin and H. Elliott, "Modeling and segmentation of noisy and textured images using Fuzzy Gibbs Random Fields," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-9, no. 1, pp. 39–55, Jan. 1987.

[22] T. Kasetkasem, "Image analysis methods based on Markov random field models," Ph.D. dissertation, Syracuse Univ., Syracuse, NY, Dec., 2002.

[23] J. Besag, "On the statistical analysis of dirty pictures," *J. R. Stat. Soc.*, vol. 48, no. 3, pp. 259–302, 1986.

[24] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*. Upper Saddle River, NJ: Prentice-Hall, 2008.

[25] Min Xu,, Hao Chen, and Pramod K. Varshney, *″* An Image Fusion Approach Based on Markov Random Fields*″ IEEE Transactions* On Geoscience And Remote Sensing, Vol. 49, No. 12, Dec. 2011.

[26] Yifan Zhang, Steve De Backer, and Paul Scheunders," Noise-Resistant Wavelet-Based Bayesian Fusion of Multispectral and Hyperspectral Images" *IEEE TRANSACTIONS* ON GEOSCIENCE AND REMOTE SENSING, VOL. 47, NO. 11, NOVEMBER 2009

[27] Peter J. Burt and Raymond J. Kolczynski " Enhanced Image Capture Through Fusion" David Sarnoff Research Center, Princeton NJ, 08543-5300.