http://ansinet.com/itj

ITJ

ISSN 1812-5638

INFORMATION TECHNOLOGY JOURNAL

ANSInet
Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Information Technology Journal 10 (11): 2220-2225, 2011
ISSN 1812-5638 / DOI: 10.3923/itj.2011.2220.2225

Wavelet-domain Hidden Markov Tree Model Approach to Fusion of Multispectral and Panchromatic Images

[1]Jianjun Yin, [2]Ming Gu, [3]Jiang Wang and [3]Jianqiu Zhang
[1]Department of Electronic Engineering, Fudan University, China
[2]Suzhou-CAS Semiconductor Integrated Technology Co., Ltd., China
[3]Departmemnt of Electronic Engineering, Fudan University, China

Abstract: We propose a wavelet-domain Hidden Markov Tree (HMT) model-based multi-spectral and panchromatic images fusion algorithm in this study. Our algorithm exploits the wavelet-domain HMT model learnt from the high-resolution panchromatic image to perform super-resolution operation to the low-resolution multispectral image. In this way, the desired high-resolution multispectral image is obtained. The experimental results showed that the proposed algorithm can produce sharper images as well as retaining good color. Moreover, as a result of the insensitivity of the wavelet coefficients' statistical information to the noises, our algorithm exhibits stronger robustness to the noises.

Key words: Image fusion, multispectral image, panchromatic image, wavelet, hidden Markov tree model

INTRODUCTION

The fusion of Panchromatic (PAN) and Multispectral (MS) images is the process of combining the PAN and MS images to produce images characterized by both high spatial and spectral resolutions. With the fusion of different images, we can overcome the limitations of information obtained from individual sources.

It has been noted by Joshi *et al.* (2006) that the fusion of MS and PAN images is an ill-posed problem. To solve the ill-posed problems, we need regularizations. Accordingly, the model-based fusion approach is proposed. First of all, we compute the statistical prior information from the PAN image. Then, to obtain the high-resolution MS image, we can perform super-resolution to the low-resolution MS image with the prior information. The most important part of the model-based fusion methods is how to obtain the appropriate prior information from the PAN images. Recently, various models have been proposed for the model-based fusion, including auto-regressive model (Joshi *et al.*, 2006), Gaussian Model (Hardie *et al.*, 2004), inhomogeneous Gaussian-Markov model (Joshi and Jalobeanu, 2010) and total variational model (Kumar and Dass, 2009).

However, these models are all computed in the spatial domain. Although, the spatial-domain statistical models generally work well when there is little noise, the performance degenerates considerably when the noise increases. Wavelet-domain information is well-known for its robustness to noise (Eismann and Hardie, 2005) and is therefore, exploited here to obtain a more robust prior model. We propose a MS and PAN images fusion method using wavelet-domain hidden Markov tree model. This model can model both the statistical characteristics of the wavelet coefficients and the relationship across wavelet coefficients hierarchy. It is widely used in image denoising, image super-resolution and compressive sensing (Romberg *et al.*, 2001). It provides useful prior information for model-based image fusion methods. Because of the robustness of the wavelet-domain statistical model, the proposed algorithm is less sensitive to noise. This characteristic will be validated by our experiment in present study.

Wavelet-domain hidden Markov tree model: An image can be decomposed as (Mallat, 1999):

$$f(x)\in L^2(R^2)$$

$$f(s)=\sum_{k\in Z^2} u_{j_0,k}\phi^{LL}_{J_0,k}(s)+\sum_{B}\sum_{j=-\infty}^{J_0}\sum_{k\in Z^2} w^{B}_{j,k}\Psi^{B}_{j,k}(s) \quad (1)$$

where,

$$B\in\{LH,HL,HH\}$$

$$u_{j_0,k}=\iint_{s\in R^2} f(s)\phi^{LL}_{j_0,k}{}^{*}(s)ds$$

Corresponding Author: Jianjun Yin, Department of Electronic Engineering, Fudan University, China

$$w_{j,k}^{B} = \iint_{s\in R^2} f(s)\Psi_{j,k}^{B\,*}(s)ds$$

$$\{\phi_{J_0,k}^{LL}, \Psi_{j,k}^{B}, j \le J_0, k \in Z^2\}$$

form an orthonormal basis for $L^2(R)$, while This results in a natural quad-tree structuring of the three subbands. It is known that most of the energy concentrates in a small fraction of wavelet coefficients. This sparsity property is desirable because it enables an image to be represented with a small number of wavelet coefficients. The wavelet coefficients of natural images have two primary properties: non-Gaussianity and persistency. That is why, the wavelet coefficients have peaky, heavy-tailed marginal distributions and large/small values of wavelet coefficients tend to propagate through the scales of the quad-trees. The HMT approximates the non-Gaussian marginal pdf as an independent Gaussian mixture and employs Markov chain to capture the interscale persistence property.

WAVELET-DOMAIN HMT BASED MULTISPECTRAL IMAGE FUSION

Multispectral image observation model: The multispectral image fusion problem can be cast in a high-resolution restoration framework (Joshi *et al*., 2006), where an appropriate prior model learnt using the high spatial resolution PAN observation is exploited to regularize the solution. Figure 1 shows the block scheme of the model that represents the relationship between the available low-resolution MS-image and its high-resolution version for a single spectral channel. In particular, the observed low-resolution MS images are models as decimated and noisy versions of their high-resolution versions.

Let Z_1 represent the available high-resolution PAN image and represent the lexicographically ordered high-resolution MS images (each $N^2\times1$ of pixels) and are the corresponding vectors ($M^2\times1$) containing pixels from the low-resolution observations, then we can write:

$$Z_m, m=2,3,...,p$$

$$y_m, m=2,3,...,p,$$

$$Y_m = Dz_m + \upsilon_m \quad (2)$$

where, D the decimation matrix of size $M^2\times N^2$ and v_m is the $M_2\times1$ noise vector that assume to be zero-mean i.i.d. process with variance δ^2. Although, there are numerous methods to estimate the decimation matrix, in this study, we assume the decimation is fixed in this study, because the estimation of the decimation matrix is not the key point

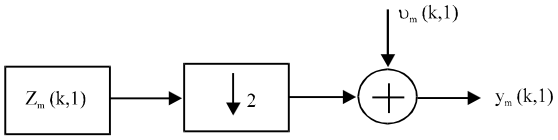

Fig. 1: Low-resolution image formation model

of this study. An example for a decimation factor of q = 2 and with z_m of the size 16×1, can be written as:

$$D = \frac{1}{4}\begin{pmatrix} 1100110000000000 \\ 0011001100000000 \\ 0000000011001100 \\ 0000000000110011 \end{pmatrix} \quad (3)$$

Thus, the low-resolution pixel is the average of high resolution pixels over a neighborhood of q^2 corrupted with additive noise. The decimation process is represented by the matrix D.

Our problem is now reduced to estimating z_m given y_m, m = 2,3,…, p and z_1. The problem is an ill-posed inverse problem, because the matrix D is not invertible. Therefore, obtaining a robust solution requires a reasonable assumption about the nature of the true image, which is called the prior information. Once the prior model for the true image is chosen, the properties of the model only depend on the model parameters. However, the parameters of the prior model are unknown as the true high resolution MS images are unavailable and have to be estimated. In this study, the parameters are learnt from the available high-resolution PAN observations. In other words, we assume the spatial correlation for the unknown high-resolution MS images can be estimated from the available high-resolution PAN data. In this study, we propose a prior model using wavelet-domain hidden Markov tree.

Multispectral image fusion using wavelet-domain hidden markov tree: According to maximum a posterior principle $\hat{z}_m$, can be estimated by maximizing $P(Z_m,/y_m)$. Using Bayes theorem, we have:

$$P(z_m / y_m) = P(y_m / z_m)P(z_m) \quad (4)$$

Taking the logarithm of (4), $\hat{z}_m$ can be obtained via:

$$\hat{z}_m = \arg\max_{z_m}[\log P(y_m / z_m) + \log P(z_m)] \quad (5)$$

As Eq. 2 shows, can be written as:

$$\log P(y_m / z_m)$$

$$\log P(y_m / z_m) \propto -\frac{\| y_m - Dz_m \|}{2\delta^2} \quad (6)$$

$\log P(z_m)$ is characterized by hidden Markov tree model:

$$\log P(z_m) \propto \log P(w_m) = \sum_{i=1}^{N} \log \sum_{t=1}^{2} f(w_{im} \mid s_{im} = t) p(s_{im} = t) \quad (7)$$

where, w_m is the corresponding wavelet coefficient of z_m w_{im} and s_{im} are ith wavelet coefficient and state variable of z_m. Due to the complexity of Eq. 7, some simplification will be made. First let consider when $s_{im} = t$:

$$-\log f(w_{im} \mid s_{im} = t) \propto \frac{w_{im}^2}{2\delta_{im,t}^2} \quad (8)$$

where, $\delta_{i,m}$ is the variance of wavelet coefficient w_{im} when $s_{im} = t$. If $s_{im} = 1$ denotes that the wavelet coefficient is small, then. In other words, when $s_{im} = 1$, the value of the wavelet coefficients are suppressed. We can simplify Eq. 7 as follows:

$$\delta_{im,1}^2 \ll \delta_{im,2}^2$$

$$\log P(z_m) \approx -\sum_{i=1}^{N} \frac{w_{im}^2}{p_{s_{im}}(1)\,\delta_{im,1}^2 + p_{s_{im}}(1)\delta_{im,2}^2} \quad (9)$$

Finally, in order to avoid the ringing effect, we regularize the restoration with Gaussian-Markov random field similar to (Joshi and Jalobeanu, 2010), whose energy function is:

$$U(z_m) = \sum_i \sum_j U_x + U_y + U_u + U_v \quad (10)$$

where,

$$U_x = b_{i,j}^x [z_m(i,j) - z_m(i-1,j)]^2 \quad (11)$$

$$U_y = b_{i,j}^y [z_m(i,j) - z_m(i,j-1)]^2 \quad (12)$$

$$U_u = b_{i,j}^u [z_m(i,j) - z_m(i-1,j-1)]^2 \quad (13)$$

$$U_v = b_{i,j}^v [z_m(i,j) - z_m(i-1,j+1)]^2 \quad (14)$$

$z_m(i,j)$ denote the value of pixel at coordinate (i, j) of the image z_m, $b_{i,j}^x, b_{i,j}^y, b_{i,j}^u, b_{i,j}^v$ is the parameters learnt from z_l:

$$\hat{b}_{i,j}^x = \frac{1}{\max\{8[z_l(i-1,j) - z_l(i,j)]^2, 8\}} \quad (15)$$

$$\hat{b}_{i,j}^y = \frac{1}{\max\{8[z_l(i,j-1) - z_l(i,j)]^2, 8\}} \quad (16)$$

$$\hat{b}_{i,j}^u = \frac{1}{\max\{8[z_l(i-1,j-1) - z_l(i,j)]^2, 8\}} \quad (17)$$

$$\hat{b}_{i,j}^v = \frac{1}{\max\{8[z_l(i-1,j+1) - z_l(i,j)]^2, 8\}} \quad (18)$$

Combining Eq. 6, 9 and 10 the optimization formula can be written as:

$$\hat{z}_m = \underset{z_m}{\arg\min} \sum_{i=1}^{N} \frac{w_{im}^2}{p_{s_{im}}(1)\delta_{im,1}^2 + p_{s_{im}}(1)\delta_{im,2}^2} + \lambda_1 \| y_m - Dz_m \| + \lambda_2 U(\hat{z}_m)$$

$$\text{st.: } w_m = Wz_m \quad (19)$$

W is the wavelet transform matrix. Due to the linearity of the wavelet transform Eq. 19 is a quadratic programming problem. Because the size of this system is extremely large, an iterative approach called conjugate gradient algorithm (Shewchuk, 1994) is used to solve this problem. In addition, due to the scale of the system, the direct product the wavelet transform matrix W and decimation matrix D with a vector is too slow and time-consuming. In this study, we employ fast wavelet transform (Mallat, 1999) to calculate the product W of W^T or\with any vector and compute the product of D with a vector with convolution.

EXPERIMENTAL RESULTS

Experimentation on synthetically generated image: Here, we consider the verification of our model for the proposed fusion method by considering the synthetically generated data. To start with, we generate a checkerboard image and treat it as a high-spatial-resolution image. The observed low-resolution image is then formed by decimation and addition of an i.i.d Gaussian noise of the variance 0.0004. The high-resolution image is then estimated from the degraded version. Fig. 2a shows the original high-resolution image. In Fig. 2b, we show the decimated and noisy version of the same image. A decimation factor of two was used and the decimation is obtained via the decimation matrix similar to Eq. 3. Figure 2c shows the estimated high-resolution image using the

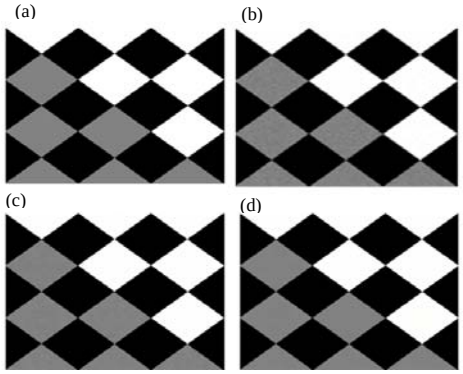

Fig. 2 (a-d): Fusion results on synthetically generated checkerbo ard image, (a) 256×256 pixels high-resolution checkerboard image, (b) Decimated and noisy image, (c) Reconstructed high-resolutionimage using IGMRF model and (d) Estimated image with the proposed method

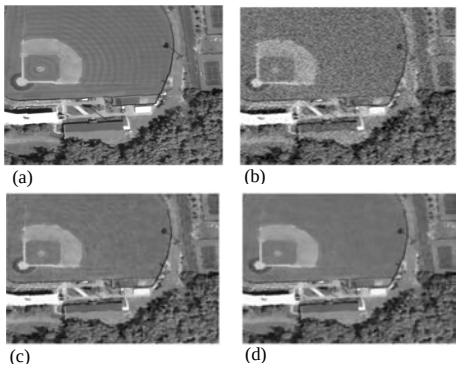

Fig. 3 (a-d): The experiment using satellite 256×256 images. (a) Pixels satellite image, (b) The observed decimated and noisy image, (c) The reconstructed image using IGMRF-based method and (d) Estimated image with the proposed method

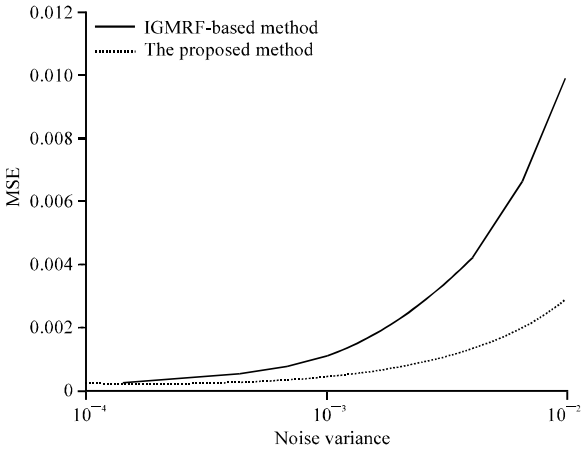

Fig. 4: The comparison of the robustness for the proposed method and IGMRF-based method

IGMRF-based approach. It can be seen that, although, the estimated image exhibits lower noise, there is still considerable artifacts in the center of each box. The MSE between the reconstructed image and the real image is 7.24×0^{-5}. Figure 2d is the reconstructed image using the proposed algorithm. The fused image is free of ringing effect or noticeable noise. The MSE between the fused image and the real image is 5.0×0^{-5}. The experiment demonstrates the working of our model and also clarifies that the method is capable of performing image fusion. Moreover, our algorithm is proved to be robust to noise.

Experimental illustration with a single image: The objective of this section is to test the efficacy of the proposed method when we consider a real satellite images as the test image and try to recover this image from its degraded version. We choose a single panchromatic image of the Quickbird image of the Missisipi University. The satellite image is shown in Fig. 3a. The observed image Fig. 3b is obtained via decimation and adding the noise with variance 0.0025. Figure 3c indicates the image obtained by IGMRF-based algorithm. There is noticeable noise in the image. The MSE between the fused image and real image is 0.0012. Figure 3d is the image reconstructed with the proposed algorithm. The image has less noise and has MSE with the real image 7.60×10^{-4}. However, the image fused with our algorithm lose some detailed information in the baseball court.

In addition, we compare the robustness of the proposed algorithm with IGMRF-based method. Adding noise with variance from 0.0001 to 0.01 to the image, we obtain the relationship between the MSE of the estimated image with the real image and the variance of the noise for both method in Fig. 4. It can be seen that the proposed algorithm exhibit better robustness to noise than IGMRF-based algorithm does.

Experimentations on quickbird images: Here, we compare the performance of various methods for multispectral fusion. We employ both the subjective and objective evaluation criteria. Each evaluation criteria has their own advantage. The objective evaluation criteria provide normalized and consistent metric for the image fusion performance. However, the human eyes have unparalleled sense of perception to color. Therefore, this

Table 1: The subjective criteria of the fused image obtain with various algorithm

Method		Cubic interpolation	IHS-based method	IGMRF-based method	The proposed algorithm
CC	R	0.9881	0.8192	0.9939	0.9953
	G	0.9900	0.7987	0.9948	0.9961
	B	0.9934	0.8202	0.9957	0.9972
	Average	0.9905	0.8127	0.9948	0.9962
D	R	2.8817	15.2124	1.8741	1.8630
	G	2.8443	16.4384	1.8786	1.8379
	B	2.4638	16.4433	1.8792	1.6557
	Average	2.4260	16.0314	1.8773	1.7855
SSD	R	18.7892	52.0999	12.2129	11.6082
	G	18.3321	51.8046	12.2382	11.2296
	B	16.9835	51.8359	12.2389	9.4040
	Average	18.0349	51.9135	12.2300	10.7472

comparison is performed both visually and quantitatively using the following indicators.

Standard Deviation of Difference image (SDD): The SSD is the standard variance of the difference image, relative to the mean of the original image, defined as follows:

$$SDD = \sqrt{\frac{1}{MN}\sum_{x}\sum_{y}[(F(x,y)-E(x,y)-(\bar{F}-\bar{E})]^2} \quad (20)$$

The smaller the SSD is, the better the quality of the merged image.

Correlation coefficient (CC): CC is the correlation coefficient between the original and the merged images. It should be as close to 1 as possible:

$$CC(f,g) = \frac{\sum[(f(i,j)-\bar{f})\times g(i,j)-\bar{g})]}{\sqrt{\sum[(f(i,j)-\bar{f})^2]\times\sum[(g(i,j)-\bar{g})^2]}} \quad (21)$$

where, f is the merged image, $\bar{f}$ is its mean value, g is the original high-resolution multispectral image, $\bar{g}$ represents for the mean of the multispectral image.

Difference coefficient (D): Difference coefficient is the difference between the means of the original and the merged image (bias), in radiance as well as its value relative to the mean of the original image (RM). The smaller these differences are, the better the spectral quality of the merged image:

$$D(k) = \sum_{i,j}|(fuse(i,j,k)-ms(i,j,k))|/(m\times n) \quad (22)$$

where, fuse is the merged image, ms is the original high-resolution multispectral image, m, n are the width and height of the merged image, respectively.

The high-resolution panchromatic image is shown in Fig. 5a and the low-resolution multispectral image is presented in Fig. 5b. We have performed registration before fusion. The image reconstructed with cubic interpolation is illustrated in Fig. 5c. It can be seen that

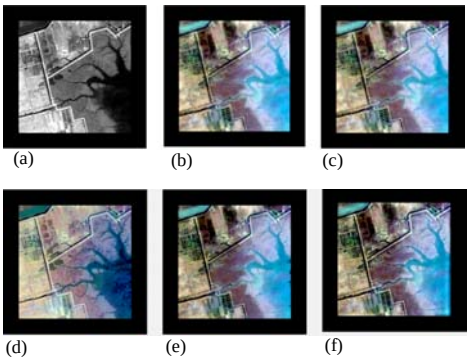

Fig. 5 (a-f): Real satellite image experiment (a) high-resolution panchromatic image, (b) low-resolution multispectral image, (c) high resolution image restored with cubic interpolation, (d) image fused with HSV approach (e) image fused with IGMRF-based method and (f) image reconstructed with the proposed algorithm

the merged image is very blurry. The image merged with IHS-based method in Fig. 5d suffers from significant color distortion. Figure 5e is the image fused with IGMRF model-based approach. Although, the merged image's color is relatively better, the border of the image has distortion. In addition, the details such as the branch of the river are blurred. In comparison, the image in Fig. 5f, merged with the proposed method has the best quality. Although, the fused image has small distortion at the border, it preserves color and detail information very well. Table 1 compares the quantitative quality metrics of various fusion approaches. It can be seen that the proposed method produces images with the best spectral and spatial quality.

CONCLUSIONS

This study proposes a multispectral image fusion algorithm using wavelet-domain Markov tree model. The wavelet-domain Markov tree model and the IGMRF model

learnt from the high-resolution PAN image are used as prior model to regularize the super-resolution restoration of the multispectral image. The experiments show that our algorithm is more robust to additive noise than the algorithm using IGMRF as prior model. Moreover, the fused images are shaper and retain better color than traditional algorithms.

ACKNOWLEDGMENTS

The study reported in this study was funded under the National Natural Science Foundation of China under Grant No. 60872059 and the National Basic Research Program of China (2011CB302900).

REFERENCES

Eismann, M.T. and R.C. Hardie, 2005. Hyperspectral resolution enhancement using high-resolution multispectral imagery with arbitrary response functions. IEEE Trans. Geosci. Remote Sens., 43: 455-465.

Hardie, R.C., M.T. Eismann and G.L. Wilson, 2004. MAP estimation for hyperspectral image resolution enhancement using an auxiliary sensor. IEEE Trans. Image Process., 13: 1174-1184.

Joshi, M. and A. Jalobeanu, 2010. MAP estimation for multiresolution fusion in remotely sensed images using an IGMRF prior model. IEEE Trans. Geosci. Remote Sens., 48: 1245-1255.

Joshi, M.V., L. Bruzzone and S. Chaudhuri, 2006. A model-based approach to multiresolution fusion in remotely sensed images. IEEE Trans. Geosci. Remote Sens., 44: 2549-2562.

Kumar, M. and S. Dass, 2009. A total variation-based algorithm for pixel-level image fusion. IEEE Trans. Image Process., 18: 2137-2143.

Mallat, S., 1999. A Wavelet Tour of Signal Processing. 2nd Edn., Academic Press, London, ISBN-10: 012466606X.

Romberg, J.K., C. Hyeokho and R.G. Baraniuk, 2001. Bayesian tree-structured image modeling using wavelet-domain hidden Markov models. IEEE Trans. Image Process., 10: 1056-1068.

Shewchuk, J.R., 1994. An introduction to the conjugate gradient method without the agonizing pain. Technical Report CMUCS-TR-94-125, Carnegie Mellon University. http://www.cs.cmu.edu/~quake-papers/painless-conjugate-gradient.pdf