A Novel Statistical Image Fusion Rule for Noisy Source Images

Tamanna Howlader[1,4], Fatema Tuz Jhohura[2] and S. M. Mahbubur Rahman[3]

[1]ISRT, Dhaka University, Dhaka, BANGLADESH
[2]BRAC Center, Dhaka, BANGLADESH
[3]Department of EEE, BUET, Dhaka, BANGLADESH
[4]Corresponding author: Tamanna Howlader, e-mail: tamanna@isrt.ac.bd

Abstract

Image fusion is an emerging area of statistical application that deals with the integration of images captured by multiple sensors or different modes of imaging systems to obtain an image having higher information content than the individual images. A practical limitation of existing fusion algorithms is that the input images are often assumed to be noise-free. This paper presents a new statistical image fusion technique using the coefficients of discrete wavelet transform (DWT) of noisy source images. The main feature of the proposed method is that it takes into account the statistical dependency between the wavelet coefficients of noisy source images as well as the noise-free fused image using a locally adaptive joint probability density function. The density function is then used in the Bayesian maximum a posteriori (MAP) estimation technique to derive a closed-form estimator for the noise-free wavelet coefficients of the fused image. To alleviate the problem of shift-variance property of traditional decimated DWT in a computationally efficient way, the proposed algorithm is implemented using the technique of cycle-spinning. Results are presented for the experiments carried out on a large number of test images to evaluate the performance of the proposed MAP-based method as compared to commonly-used fusion methods. Comparisons are made with respect to standard performance metrics such as the structural similarity, peak signal-to-noise ratio, and cross-entropy.

Keywords: Bayesian MAP estimation, Discrete wavelet transform, Locally-adaptive joint PDF.

1. Introduction

The two dimensional (2D) visual signal called 'image' is a very powerful medium of communication. This can best be described by the familiar saying, 'A picture is worth a thousand words'. In the digital era, the application of image signals has increased manifold. Examples include the use of X-ray images, positron emission tomography, computed tomography (CT), magnetic resonance imaging (MRI) and ultrasound images in medical diagnosis, synthetic aperture radar (SAR) images in remote sensing, cDNA microarray images in studies of differential gene expression, images from smart CCD cameras in industrial process control, infrared or thermal images in night vision, and so on. Being an easy, compact, and readily available way to represent the world that surrounds us, the image is omnipresent in our everyday lives. Thus, today's youth lives in what may be called an 'image society'.

An important area of statistical application is image fusion, which is the process of combining two or more images acquired using multiple sensors or different imaging modalities, into a single image that contains more information than any of the individual source images. The objective of image fusion is twofold: to provide better clarity for human visual perception, and to compress huge volume of data of the source images for easy storage and transmission. It is being used in areas such as remote sensing, medical diagnosis, target recognition, as well as agricultural, environmental and maritime monitoring.

In recent times, the transform-based pixel-level image fusion techniques are widely used because the transformed data exhibit certain properties that enable the fusion task to be

performed more efficiently. For instance, the discrete wavelet transform (DWT) has been very successful in image fusion owing to its space-frequency localization property (Roy et al. 2011). Details on the theory and implementation of the DWT may be found in Mallat (1999). The traditional decimated DWT, however, suffers from one significant disadvantage. It lacks shift-invariance property, as a result of which, visual artifacts appear in the reconstructed image. To overcome this problem in a computationally efficient way, Coifman and Donoho (1995) suggested cycle-spinning while adopting DWT for an algorithm. Some of the popular DWT-based fusion rules are those given by Li et al. (1995), Wu et al. (2004), Pu and Ni (2000) and Yunhao et al. (2006). A practical limitation of these methods is that they assume the input images are noise-free. Unfortunately, noise is inherent in all imaging systems being caused by factors such as, atmospheric abberations, sensor noise and quantization noise that may be introduced at various stages in acquisition or transmission (Petrović, 2001). Thus, the performance of these methods may be seriously compromised when the source images are of poor quality.

Since the image is a 2D random signal (Gonzalez and Woods, 2011), the probability density function (PDF) is a useful tool in the development of efficient methods for many types of image processing problems. For example, in image denoising, the Laplacian, generalized Gaussian, symmetric normal inverse Gaussian, scale mixtures of Gaussian (SMG) and quasi-Cauchy PDFs have been used to model the DWT coefficients of an entire subband of images (Roy et al. 2011), while the Gaussian PDF has been used to model local neighboring DWT coefficients (Cai and Silverman 2001). Generally, noise is also random and PDFs such as Gaussian (Howlader and Chaubey 2009), Poisson (Le et al. 2007), etc. have been used to describe its nature. Among the various noise models, the additive white Gaussian noise model (AWGN) is of interest because it occurs frequently and other common noise models can be mathematically remodelled as AWGN (Rangayyan et al. 1998). A limited number of studies have tried to address the problem of noisy image fusion using a statistical approach. For instance, Loza et al. (2007) and Yin et al. (2011) have used the bivariate Laplacian distribution to derive computationally expensive fusion rules. In Rahman et al. (2010), noisy images are fused using a method based on a locally-adaptive bivariate Gaussian PDF, while the generalized Gaussian and alpha-stable distributions have been used in a method by Loza et al. (2010). In this article, a new locally-adaptive multivariate PDF is derived, which takes into account the statistical dependencies between the DWT coefficients of the noisy source images and the noise-free fused image. This PDF is then used in Bayesian maximum a posteriori (MAP) estimation to obtain a closed-form expression for the noise-free DWT coefficients of the fused image. To obtain a better performance as compared to the decimated DWT, the cycle-spinning technique is used for fusing the images. The new fusion rule is evaluated through experiments conducted on a large number of data sets using commonly-used metrics, viz., structural similarity (SSIM), peak-signal-to-noise ratio (PSNR) and cross entropy (CEN) (Roy et al. 2011) as well as in terms of visual quality.

The paper is organized as follows. In Section 2, the proposed DWT-based image fusion rule for noisy source images is described. The simulation results are given in Section 3. Finally, Section 4 provides the conclusion.

2. Fusion of noisy source images in DWT domain

In this section, the proposed pixel-level image fusion technique is described for source images corrupted by AWGN. Let $x_1^c(k_1, k_2)$ and $x_2^c(k_1, k_2)$ denote the DWT coefficients of the noise-free source images 1 and 2, respectively, at spatial location (k_1, k_2) of a given subband c ($c \in A, O$). Here, A and O represent approximate and detailed subbands, respectively. Then the noisy

coefficients of the images at (k_1, k_2) can be written as

$$y_1^c(k_1,k_2) = x_1^c(k_1,k_2) + n_1^c(k_1,k_2) \tag{2.1}$$
$$y_2^c(k_1,k_2) = x_2^c(k_1,k_2) + n_2^c(k_1,k_2) \tag{2.2}$$

where $n_1^c(k_1,k_2)$ and $n_2^c(k_1,k_2)$ are the noise coefficients of source images 1 and 2, respectively, and are distributed as i.i.d $\mathcal{N}(0,\sigma_{nn})$. It is assumed that σ_{nn} is known and signal is independent of the noise. If σ_{nn} is unknown, it may be estimated using the median-absolute-deviation method given in Donoho and Johnston (1995). Since the wavelet coefficients of images in a subband are spatially non-stationary (Liu and Moulin, 2001), the random variables of the coefficients of the noise-free and noisy images are index dependent. Let $x_1^c(k_1,k_2)$ and $x_2^c(k_1,k_2)$ represent observed values of the random variables $X_1^c(k_1,k_2)$ and $X_2^c(k_1,k_2)$, respectively. Similarly, we define random variables $Y_1^c(k_1,k_2)$ and $Y_2^c(k_1,k_2)$. On the other hand, wavelet coefficients of noise are spatially stationary and subband independent for which the corresponding random variable is denoted by N. The design of an efficient fusion method may be formulated as a statistical estimation problem for $X_f^c(k_1,k_2)$, where the latter denotes the random variable for the DWT coefficient of the fused image at the same spatial location. The aim is to obtain an estimate of the detailed DWT coefficient of the fused image, $\hat{x}_f^O(k_1,k_2)$, using prior information regarding random variables $Y_1^O(k_1,k_2)$, $Y_2^O(k_1,k_2)$ and $X_f^O(k_1,k_2)$. We adopt the Bayesian MAP estimation approach, in which, the mode of the posterior density $p_{X_f^O|Y_1^O,Y_2^O}\big(X_f^O|Y_1^O,Y_2^O\big)$ is used as an estimate for $X_f^O(k_1,k_2)$. For notational simplicity, the superscript O and indices (k_1,k_2) are suppressed in the remainder of the paper, unless stated otherwise. Thus, the fused DWT coefficients of the detailed subbands of a given decomposition level can be obtained as

$$\hat{x}_f = \arg\max_{x_f} p_{X_f|Y_1,Y_2}\big(x_f|y_1,y_2\big) = \arg\max_{x_f} p_{Y_1,Y_2,X_f}\big(y_1,y_2,x_f\big) \tag{2.3}$$

In a recent paper (Roy et al. 2011), justification was provided for the zero-mean trivariate Gaussian PDF as a reasonable model for the joint PDF of X_1, X_2, and X_f. By considering this model, and the linear additive noise models in (2.1) and (2.2), the joint PDF for Y_1, Y_2, and X_f can be derived by transformation of variables. The resulting model is

$$\begin{aligned} p_{Y_1,Y_2,X_f}\big(y_1,y_2,x_f\big) =& \frac{\lambda_1}{\sqrt{8\pi^3\sigma_{11}\sigma_{22}\sigma_{ff}\gamma}} \exp\Bigg[-\frac{1}{2\gamma}\Big\{\alpha_1 y_1^2 + \alpha_2 y_2^2 + \alpha_3 x_f^2 - 2\beta_1 y_1 y_2 \\ & - 2\beta_2 y_1 x_f - 2\beta_3 y_2 x_f\Big\} + \frac{\lambda_2^2}{2\lambda_3}\Bigg]; \ -\infty < (y_1,y_2,x_f) < \infty \end{aligned} \tag{2.4}$$

where

$$\lambda_1 = \sqrt{\frac{\gamma}{(\alpha_1+\alpha_2-2\beta_1)\sigma_{nn}+\gamma}}, \quad \lambda_2 = \frac{1}{\gamma}\Bigg[y_1(\alpha_1-\beta_1) + y_2(\alpha_2-\beta_1) - x_f(\beta_2-\beta_3)\Bigg],$$
$$\lambda_3 = \frac{1}{\gamma}\Bigg[\alpha_1+\alpha_2-2\beta_1\Bigg] + \frac{1}{\sigma_{nn}}, \quad \alpha_1 = \frac{1-\rho_{2f}^2}{\sigma_{11}}, \quad \alpha_2 = \frac{1-\rho_{1f}^2}{\sigma_{22}}, \quad \alpha_3 = \frac{1-\rho_{12}^2}{\sigma_{ff}}$$
$$\beta_1 = \frac{\rho_{12}-\rho_{1f}\rho_{2f}}{\sqrt{\sigma_{11}}\sqrt{\sigma_{22}}} \quad \beta_2 = \frac{\rho_{1f}-\rho_{12}\rho_{2f}}{\sqrt{\sigma_{11}}\sqrt{\sigma_{ff}}} \quad \beta_3 = \frac{\rho_{2f}-\rho_{12}\rho_{1f}}{\sqrt{\sigma_{22}}\sqrt{\sigma_{ff}}} \quad \text{and}$$
$$\gamma = 1 + 2\rho_{12}\rho_{1f}\rho_{2f} - \rho_{12}^2 - \rho_{1f}^2 - \rho_{2f}^2$$

Here, σ_{uu} and ρ_{uv}, for $u,v \in (1,2,f)$, $u \neq v$, are the variance and correlation parameters respectively, that are estimated using the local neighboring DWT coefficients of the noise-free

images. Then the MAP estimator $\hat{x}_f$ is obtained by substituting (2.4) into (2.3). This is equivalent to solving the equation $\frac{\partial}{\partial x_f}\Big[\ln p_{Y_1,Y_2,X_f(y_1,y_2,x_f)}\Big] = 0$, which yields

$$\hat{x}_f = \frac{1}{\alpha_3 - \omega_1\omega_2^2}\left[\Big\{\beta_2 - (\alpha_1 - \beta_1)\omega_1\omega_2\Big\}y_1 + \Big\{\beta_3 - (\alpha_2 - \beta_1)\omega_1\omega_2\Big\}y_2\right] \tag{2.5}$$

$$\text{where} \quad \omega_1 = \frac{\sigma_{nn}}{(\lambda_3 - 1/\sigma_{nn})\gamma\sigma_{nn} + \gamma} \quad \text{and} \quad \omega_2 = \sqrt{\sigma_{ff}}(\beta_2 + \beta_3)$$

Thus, the estimate of the fused coefficient at a given location is obtained as a linear combination of the noisy coefficients of the source images at that same location, where the linear weights are determined by the noise variance σ_{nn} as well as the parameters σ_{uu} and ρ_{uv}, for $u, v \in (1, 2, f)$, $u \neq v$. The fused DWT coefficients of the approximate subband at the highest-level of decomposition are obtained as

$$\hat{x}_f^A = \frac{1}{2}(y_1^A + y_2^A) \tag{2.6}$$

At each spatial index (k_1, k_2), the unknown parameters, $\sigma_{uu}(k_1, k_2)$ and $\rho_{uv}(k_1, k_2)$, are estimated from the M coefficients in a square-shaped local neighborhood $\mathcal{S}(k_1, k_2)$ centered at $x_f(k_1, k_2)$ or $y_r(k_1, k_2)$, $r \in (1, 2)$. For the purpose of estimation, the coefficients within $\mathcal{S}(k_1, k_2)$ are assumed to be i.i.d. Then the maximum likelihood estimates of the parameters are obtained as

$$\hat{\sigma}_{rr}(k_1, k_2) = \max\left(\frac{1}{M}\sum_{\mathcal{S}(k_1,k_2)}[y_r(w_1, w_2)]^2 - \sigma_{nn}, 0\right) \tag{2.7}$$

$$\hat{\sigma}_{ff}(k_1, k_2) = \frac{1}{M}\sum_{\mathcal{S}(k_1,k_2)}\left[\hat{x}_f^0(w_1, w_2)\right]^2 \tag{2.8}$$

$$\hat{\rho}_{rf}(k_1, k_2) = \max\left(\min\left(\frac{\frac{1}{M}\sum_{\mathcal{S}(k_1,k_2)} y_r(w_1, w_2)\hat{x}_f^0(w_1, w_2)}{\sqrt{\hat{\sigma}_{rr}(k_1, k_2)\hat{\sigma}_{ff}(k_1, k_2)}}, 1\right), -1\right) \tag{2.9}$$

$$\hat{\rho}_{12}(k_1, k_2) = \max\left(\min\left(\frac{\frac{1}{M}\sum_{\mathcal{S}(k_1,k_2)} y_1(w_1, w_2)y_2(w_1, w_2) - \sigma_{nn}}{\sqrt{\hat{\sigma}_{11}(k_1, k_2)\hat{\sigma}_{22}(k_1, k_2)}}, 1\right), -1\right) \tag{2.10}$$

where $\hat{x}_f^0$ is an initial estimate for x_f, which may be obtained using any simple fusion rule. The determination of $\hat{x}_f^0$ requires the denoised wavelet coefficients of source images that may be obtained using any crude but computationally fast denoising algorithm.

3. Experimental results

In this section, representative results are presented for experiments conducted on a large number of test images to evaluate the performance of the proposed MAP-based method relative to four popular fusion methods, namely, the wavelet maxima (Li et al. 1995), variance feature (Wu et al. 2004), contrast measure (Pu and Ni 2000) and adjustable parameter (Yunhao et al. 2006) methods. Results are given for three sets of two-source images: $Clock$, SAR and $CT - MRI$. These images represent the multifocus, multimodal, and multisensored images, respectively. Noisy images are created synthetically by adding WGN to the noise-free images considering several noise levels. The DWT coefficients are obtained by employing a 4-level

Table 1: Performance of the DWT-based fusion methods using cycle-spinning.

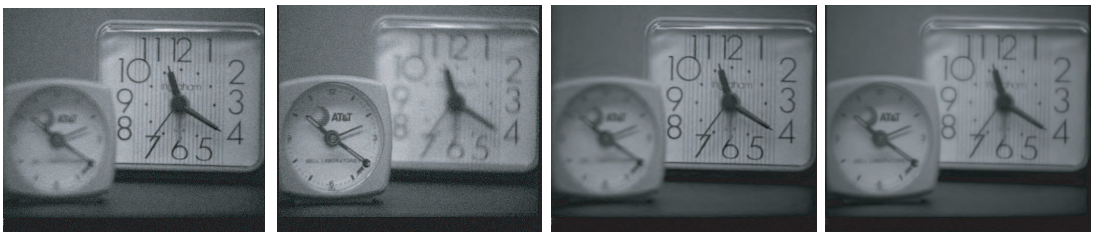
Images	Methods†	$\sigma_{nn} = 10^2$			$\sigma_{nn} = 20^2$		
		SSIM	CEN	PSNR	SSIM	CEN	PSNR
CT-MRI	Wavelet maxima	0.4133	6.91	47.65	0.3884	7.05	47.42
	Variance feature	0.3981	6.91	47.99	0.3766	7.02	47.75
	Contrast measure	0.4372	6.72	48.59	0.4118	6.90	48.32
	Adjustable parameter	0.3506	6.76	48.32	0.3361	6.87	48.08
	Proposed MAP	**0.4419**	**6.06**	**49.56**	**0.4273**	**6.25**	**49.43**
SAR	Wavelet maxima	0.8150	7.02	66.99	0.7638	7.24	64.69
	Variance feature	0.8067	7.00	66.95	0.7605	7.21	64.80
	Contrast measure	0.8337	6.91	68.15	0.7807	7.16	65.55
	Adjustable parameter	0.7934	6.94	67.35	0.7490	7.15	65.19
	Proposed MAP	**0.8379**	**6.66**	**68.82**	**0.7943**	**6.83**	**67.11**
Clock	Wavelet maxima	0.8947	8.06	70.49	0.8527	8.33	67.77
	Variance feature	0.8943	8.04	70.68	0.8553	8.31	68.04
	Contrast measure	0.9037	8.01	70.99	0.8649	8.29	68.30
	Adjustable parameter	0.8875	7.97	73.00	0.8551	8.23	69.20
	Proposed MAP	**0.9288**	**7.76**	**73.57**	**0.9074**	**7.88**	**72.02**

† Parameters of the locally-adaptive methods are estimated using a 3×3 window.

DWT, wherein the orthogonal Symlet filter of length 8 is used. Cycle-spinning has been implemented by considering the left, right, up, and down-shifting each by 4 pixels of the images. Objective comparisons are made based on three standard performance metrics, viz., the SSIM, CEN and PSNR. A better fusion method should provide higher values for SSIM and PSNR, and a lower value for CEN. Table 1 gives the values of these indices for the methods being compared when $\sigma_{nn} = 10^2$ and 20^2. For each method, preprocessing of the noisy source images was performed using the SureShrink denoising method (Donoho and Johnstone 1995). For the proposed method, initial estimates of the fused coefficients were obtained by averaging the DWT coefficients of the estimated noise-free source images. From the table, it is evident that the proposed method consistently provides the best fusion results at both the noise levels. Similar results were obtained for other noise levels and using other denoising algorithms such as the VisuShrink (Donoho and Johnstone 1994) and NeighCoeff (Cai and Silverman 2001). Visual comparisons were also made between the fused images obtained using the proposed method and other methods for different image types. Due to space limitations, only the fused images of the test set *Clock* are shown in Figure 1 for the adjustable parameter and the proposed MAP-based method. The fused image for the adjustable parameter method is shown since the results of Table 1 indicate that it performs best among the other methods. The superiority of the proposed method is easily discerned if one notices that the small clock on the left of the fused image for the adjustable parameter method is blurred, whereas all objects are in focus for the image generated by the proposed method. Visual assessments made from other test sets lead to similar conclusions of superiority of the proposed method as compared to the others.

4. Conclusion

In this paper, a new statistical image fusion technique has been proposed for two noisy source images. The method is based on a multivariate PDF for the DWT coefficients of the noisy source images and the noise-free fused image. To obtain a good fusion performance, cycle-spinning is performed to suppress the artifacts due to shift-variance property of the decimated DWT. The strength of the proposed method is the integrated approach of the fusion and denoising, and thus reducing overall computational complexity. The method may be easily



(a) (b) (c) (d)

Figure 1: Results of image fusion for noisy *Clock* images ($\sigma_{nn} = 10^2$). Source images are focused on (a) right and (b) left. Fused images obtained using (c) the adjustable parameter and (d) the proposed MAP-based methods.

extended to an arbitrary number of source images available in many real-life applications. In conclusion, this paper has provided an interesting application of statistics in image processing, which we hope, will motivate young statisticians to conduct further research in this emerging area.

References

Cai, T. and Silverman, B. (2001) Incorporating Information On Neighboring Coefficients Into Wavelet Estimation. *Sankhya: The Indian Journal of Statistics*, 63, 127-148.

R. Coifman, R. and Donoho, D. L. (1995) Translation-invariant de-noising, in Wavelets and Statistics, A. Antoniadis and G. Oppenheim, Eds. Berlin, Germany: Springer-Verlag

Donoho, D. L. and Johnstone, I. M. (1994) "Ideal Spatial Adaptation Via Wavelet Shrinkage", *Biometrika*, 81, 425-455.

Donoho, D. L. and Johnstone, I. M. (1995). "Adapting To Unknown Smoothness Via Wavelet Shrinkage", *Journal of the American Statistical Association*, 90, 1200-1224.

Gonzales, R.C. and Woods, R.E. (2011) *Digital Image Processing*, 3rd Edition, Pearson Prentice Hall.

Howlader, T. and Chaubey, Y. P. (2009) "Wavelet-Based Noise Reduction By Joint Statistical Modeling of cDNA Microarray Images," *Journal of Statistical Theory and Practice*, 3, 349–370.

Johnstone, I. M. and Silverman, B. W. (2005) "Empirical Bayes Selection of Wavelet Thresholds," *Annals of Statistics*, 33, 1700–1752.

Le, T., Chartrand, R. and Asaki, T. J. (2007) "A Variational Approach to Reconstructing Images Corrupted by Poisson Noise", *Journal of Mathematical Imaging and Vision*, 27,257-263.

Li, H., Manjunath, B. S. and Mitra, S. K. (1995) "Multisensor Image Fusion Using the Wavelet Transform," *Graphical Models and Image Processing*, 57, 235–245.

Loza, A., Achim, A., Bull, D., and Canagarajah, N. (2007) "Statistical Model-based Fusion of Noisy Multi-band Images in the Wavelet Domain", *10th International Conference on Information Fusion*,1-6.

Loza, A., Bull, D., Canagarajah, N., and Achim, A. (2010). "Non-Gaussian Model-Based Fusion of Noisy Images in the Wavelet Domain, *Computer Vision and Image Understanding*", 114, 54-65.

Mallat, S. (1999) *A Wavelet Tour of Signal Processing*, Second Edition, CA: Academic Press, San Diego.

Petrovič, V.S. (2001) *Multisensor Pixel-level Image Fusion*, Ph.D. thesis, University of Manchester.

Pu, T. and Ni, G. Q. (2000) "Contrast-Based Image Fusion Using the Discrete Wavelet Transform," *Optical Engineering*, 39, 2075–2082.

Rangayyan, R. M., Ciuc, M. and Faghih, F. (1998) "Adaptive-Neighborhood Filtering of Images Corrupted By Signal-Dependent Noise" *Applied Optics*, 37, 4477-4487.

Roy, S., Howlader, T. and Rahman S.M.M. (2013) "Image Fusion Technique Using Multivariate Statistical Model For Wavelet Coefficients", *Signal, Image and Video Processing*, 7, 355-365.

Wu, J., Liu, J., Tian, J. and Huang, H. (2004) "Multi-Scale Image Data Fusion Based on Local Deviation of Wavelet Transform," Proceedings of IEEE International Conference on Intelligent Mechatronics and Automation, Chengdu, China, 677–680.

Yin, S., Cao, L., Ling, Y. and Jin, G. (2011) "Fusion of Noisy Infrared And Visible Images Based On Anisotropic Bivariate Shrinkage", *Infrared Physics & Technology*, 54, 1320.

Yunhao, C., Lei, D., Jing, L., Xiaobing, L. and Peijun, S. (2006) "A New Wavelet-Based Image Fusion Method For Remotely Sensed Data," *International Journal of Remote Sensing*, 27, 1465–1476.