A WAVELET-BASED NOISE-AWARE METHOD FOR FUSING NOISY IMAGERY

Xiaohui Yuan and Bill P. Buckles

{xyuan, bbuckles}@cse.unt.edu
Department of Computer Science and Engineering
University of North Texas, Denton, TX 76203

ABSTRACT

Fusion of images in the presence of noise is a challenging problem. Conventional fusion methods focus on aggregating prominent image features, which usually result in noise enhancement. To address this problem, we developed a wavelet-based, noise-aware fusion method that distinguishes signal and noise coefficients on-the-fly and fuses them with weighted averaging and majority voting respectively. Our method retains coefficients that reconstruct salient features, whereas noise components are discarded. The performance is evaluated in terms of noise removal and feature retention. The comparisons with five state-of-the-art fusion methods and a combination with denoising method demonstrated that our method significantly outperformed the existing techniques with noisy inputs.

Index Terms— Image Fusion, Wavelet Transforms, Noise

1. INTRODUCTION

Image fusion integrates multiple images to improve quality for tasks such as visualization, segmentation, object identification, etc. Fusion approaches described in the literature focus on maximizing the feature preservation ([1, 2, 3] and references therein). During image acquisition and data transformation, noise is unavoidable and is usually intensified at many stages. Fusing noisy images usually results in heavily distorted composites. Conventionally, noise removal methods are employed, which, however, introduces the problem of suppressing valuable features.

To address this issue, noise-aware fusion methods have been developed. Burt and Kolczynski [4] used selection and averaging rules to merge Laplacian coefficients with significant difference and high similarity, respectively. Alternatively, Pavel et al. [5] developed a method to integrate synthetic templates (noise-free) with real sensor images. The fusion was achieved with the weighted sum of the two according to the local variability in the Laplacian pyramid subbands. More often, in conjunction with multi-resolution-based fusion methods, consistency verification is used. It is a method to overcome isolated noise and coefficient selection artifacts and assumes the coefficient selection integrity within a local neighborhood [3]. In these methods, a universal, preselected threshold was used for fusion. Recognizing that the amount of noise present in images varies, we developed an adaptive method to combine image features according to the noise strength in the wavelet subbands.

Inspired by the success of wavelet shrinkage [6, 7], we developed a wavelet-based, noise-aware fusion method that classifies wavelet coefficients on-the-fly and fuses the signal coefficients by minimizing the mean square error. The noise coefficients, on the other hand, are suppressed to zero to avoid distortion unless an agreement is found in all corresponding subbands.

The rest of this article is organized as follows: Section 2 describes our fusion method in detail. Section 3 presents the experimental results from both feature retaining and noise suppression aspects. Section 4 gives the concluding remarks and future work.

2. WAVELET-BASED NOISE-AWARE FUSION

In wavelet-based image fusion approaches, source images are transformed into the wavelet domain, where discontinuity features are separated from the background information. The combination strategy then selects appropriate coefficients and produces a fused set of wavelet subbands. In the presence of noise, wavelet subbands consist of a mixture of signal and noise coefficients. Distinguishing these two types of coefficients is critical for an effective fusion method to preserve image features and avoid integrating noise.

In wavelet shrinkage-based noise removal, a carefully chosen threshold, denoted by λ, separates signal and noise coefficients [6, 7].

$$I(\lambda) = \{I_s(x,y)\} \cup \{I_n(x,y)\} \tag{1}$$

where I denotes a wavelet subband. $I_s(x,y)$ are the coefficients greater than the threshold λ. These coefficients are essentially signal coefficients that encode prominent image features. Due to their large magnitude, they are less likely to be altered by noise. That is, noise can change only a small

This work was partially supported by National Science Foundation grant IIS-0722106.

portion of the magnitude of a signal coefficient. $I_n(x, y)$ are the coefficients less than λ. If there is no noise imposed to the image, these coefficients are mostly zeros or very close to zero. The presence of noise causes their values to depart from nil. Based on the above model, we can design fusing strategies to combine signal and noise coefficients. The details of our method are described in the following three sections.

2.1. Fusion of Signal Coefficients

When fusing signal coefficients, we need to maximize the image feature retention while minimizing the error covariance. Let J denote the coefficients in the fused subband and I denote the input subbands. The fusion result is achieved with the following equation:

$$J = \arg\min_J E[(I - J)^T (I - J)] \tag{2}$$

where $E[\cdot]$ denotes the expectation. Assume the fused subband is the weighted sum of the inputs

$$J = \sum_{k=1}^{K} \Phi_k I_k \tag{3}$$

where K is the number of source images. The estimation error is orthogonal to the observed subband, hence, we have

$$E[(\tilde{J} - \tilde{\Phi}\tilde{I})\tilde{I}^T] = 0 \tag{4}$$

where $\tilde{J}$ and $\tilde{I}$ are the demeaned subbands. By expanding the expectation and combining with Equation 3, the fused subband $\tilde{J}$ is computed as follows:

$$\tilde{J} = \sum_k C_{\tilde{J}\tilde{I}_k} C^{-1}_{\tilde{I}_k \tilde{I}_k} \tilde{I}_k \tag{5}$$

where $C_{\tilde{J}\tilde{I}_k}$ and $C_{\tilde{I}_k \tilde{I}_k}$ are the covariances. Assume that noise has zero mean and variance $\sigma^2_{n,k}$ and is independent of J. The fusion rule becomes the combination of the observed subbands according to the subband variance, $\sigma^2_{I_k}$, and the noise variance, which is normalized:

$$J = \sum_k \frac{\sigma^2_{I_k} - \sigma^2_{n,k}}{\sigma^2_{I_k}} I_k \tag{6}$$

2.2. Fusion of Noise Coefficients

Noise coefficients are suppressed to zero unless there is an agreement on the existence of a feature. That is, if coefficients from input subbands are consistent with little variance, we then take the average of these coefficients; otherwise, suppression is applied.

Let τ, $\tau \in [0, 1]$, be a threshold. A set of coefficients, denoted by $\mathcal{E}_\tau$, is selected and the set size is expressed as τK. In our method, variance is used as the consistency metric,

$$\mathcal{C}(\mathcal{E}_\tau) = \frac{1}{\tau K - 1} \sum_{j=1}^{\tau K} (I_j(x, y) - \bar{I}(x, y))^2 \tag{7}$$

where $\bar{I}(x, y)$ is the mean of set $\mathcal{E}_\tau$. The coefficients are consistent upon satisfying the following two criteria.

$$\tau \geq 0.5 \tag{8}$$

$$\mathcal{C}(\mathcal{E}_\tau) < \epsilon \tag{9}$$

where ϵ is chosen manually.

Assume $\tau = 0.5$ and we have a consistent set of coefficients $\mathcal{E}_\tau$. The median value of $\mathcal{E}_\tau$ can be proven to be one of the consistent coefficients. The consistency evaluation starts with the median value and iteratively recruits the coefficients from $I_1(x, y), I_2(x, y), \ldots, I_K(x, y)$ that have the least effect on the current mean of the subset. The recruitment is constrained by the above criteria and terminates once Eq. 9 becomes invalid. The presence of a consistent subset is determined by Eq. 8. The fusion of noise coefficients is as follows:

$$J(x, y) = \begin{cases} E[\mathcal{E}_\tau] & \text{consistent } \mathcal{E}_\tau \text{ exists} \\ 0 & \text{otherwise} \end{cases} \tag{10}$$

One issue of the coefficient classification is the coexistence of signal and noise coefficients at the same subband location (x, y). Such classification disagreement typically occurs in two cases. (1) noise strength is variant, and (2) complementary features are in the input images. When the disagreement occurs, the coefficient in the fused composite is determined solely on signal coefficients, whereas noise coefficients are neglected.

2.3. Noise Estimation and Threshold Selection

In our method, accurate subband noise estimation is a critical step. To estimate the subband noise strength, Donoho [6] used the following empirical formula that employs a median filter of the high-high subband on the first decomposition level, HH_1.

$$\hat{\sigma}_n = \frac{\text{Median}(|HH_1|)}{0.6745}$$

This method is insensitive to outliers of high magnitude but gives only a rough estimation.

In [8], subband noise was estimated by finding the maximum in the local coefficient variance function. For each coefficient, a window was used to estimate the local variance, and a histogram of the local variance is computed for every subband. The noise variance is chosen to be the peak in this histogram. This estimation tends to underestimate the noise variance in the case where the image has very few edges.

Yuan and Buckles [9] developed a method to overcome the bias caused by sample size. This method takes advantage of the sparsity property of the wavelet subband. The idea is to identify the coefficients that were distorted away from zero and calculate the variance, which closely approximates the noise distribution if the following two conditions are satisfied: 1) The noise variance is much greater than that of the underlying clean coefficients used for variance computation, 2) The number of such coefficients is statistically large enough.

The noise estimation is achieved by computing subband variance as a function of magnitude threshold, denoted by ν, and identifying the inflection point of its first derivative

$$\nu = arg\max \frac{\partial \sigma^2(\nu)}{\partial \nu} \tag{11}$$

Hence, the noise variance is estimated as follows:

$$\sigma_n^2(\nu) = \frac{1}{N-1}\sum^{N}(I-\bar{I})^2 \tag{12}$$

Given the noise variance, the optimal threshold can be determined by minimizing the risk function

$$\lambda = arg\min_{\lambda} E(\hat{I}_s(\lambda) - I_s)^2 \tag{13}$$

where $\hat{I}$ is the denoised result. The closed form solution derived with numerical method is as follows [7]:

$$\lambda = \frac{\sigma_n^2(\nu)}{\sqrt{\sigma_I^2 - \sigma_n^2(\nu)}} \tag{14}$$

Table 1. SNR of fusion results. Noise STD=25.

Test Case	I	II	III	IV
Input 1	-0.6	5.8	3.4	5.4
Input 2	2.1	4	3.3	2
WNAF	**11.3**	**11.6**	**14.3**	**10.2**
WDS	10.5	9.4	12.5	9
WDR	9.6	8.8	12.1	8.5
WF	1.9	6.7	5.2	4
PCA	2	9.1	8.9	7.1
LapF	3.5	6.4	5	4.3
CPF	0.3	6.8	4.5	4.1
GPF	2.4	6.7	5.6	4.2

3. EXPERIMENTS AND DISCUSSIONS

We evaluated our method from two aspects: noise suppression and feature preservation. References were created by fusing the clean images and noisy inputs were also created from the same clean images by adding Gaussian noise. We applied Wiener denoising method to images before and after the fusion process.

Figure 1 illustrates results of fusing remote sensing images. The noise standard deviation (STD) in this experiment was 25. Given the noisy inputs, Figure 1(a) and (b), a conventional wavelet-based fusion[1] resulted in a more distorted outcome, Figure 1(c). Figure 1(d) and (e) are the results produced with the same wavelet-based fusion method, and Wiener denoising was applied to the source images and the fusion result respectively. The noise was dampened but fine features were removed as well as some artificial edges were generated. The result of our method is shown in (f), which clearly shows the improvement in noise removal. Also, the fine features became recognizable.

[1]In this wavelet-based fusion, we applied select-max to all the high-pass subbands and average to the low-pass subband.

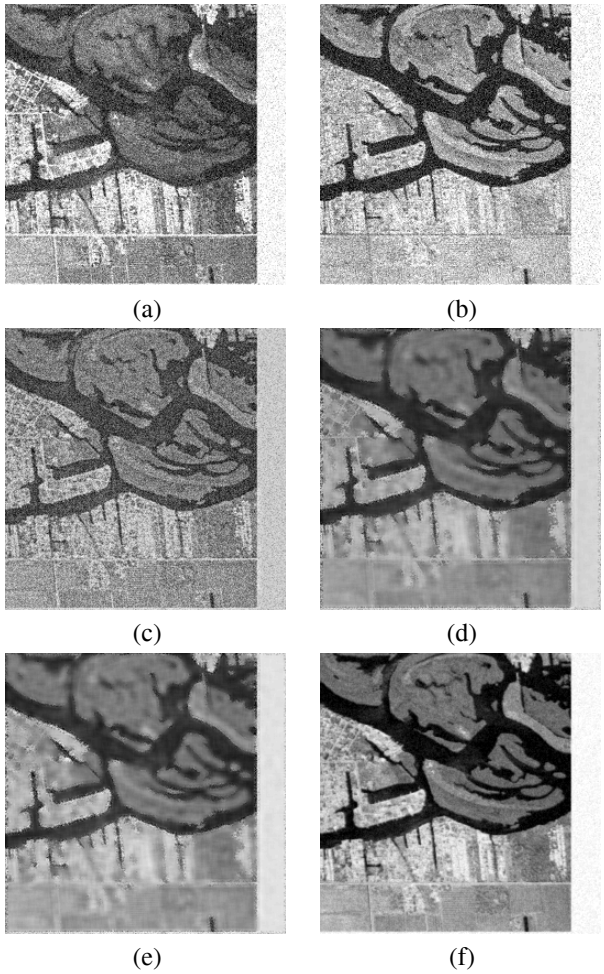

Fig. 1. Fusion of multi-spectral imagery. (a), (b): noisy inputs. (c)-(f): fusion results with conventional wavelet-based method, wavelet-based method with denoised inputs, wavelet-based method with denoised result, and our method.

Quantitative noise suppression was evaluated with signal-to-noise ratio (SNR). Table 1 lists the average SNR over five trials with our method (WNAF), wavelet-based fusion with denoised source images (WDS), wavelet-based fusion with denoised fusion result (WDR), wavelet fusion (WF), PCA-based fusion (PCA), Laplacian-based fusion (LapF), contrast pyramid-based fusion (CPF), and Gaussian pyramid-based fusion (GPF). Without denoising, conventional methods resulted in poor SNR performance. In many cases, e.g., case I with CPF, PCA, and WF, the fusion results had even smaller SNRs than the inputs. Our method outperformed others as shown in the table with best SNRs high-lighted.

To evaluate feature retaining performance, we used the

Table 2. False Negative Ratio of the fusion results given various noise STDs.

	Noise STD=15				Noise STD=25				Noise STD=35			
Test Case	I	II	III	IV	I	II	III	IV	I	II	III	IV
WNAF	**.0149**	.1120	**.1421**	.0738	**.0152**	**.1370**	**.1843**	.0860	**.0187**	**.1555**	**.2321**	**.0915**
WDS	.0242	.1942	.4911	.1530	.0386	.2030	.5376	.1593	.0337	.2122	.5671	.1603
WDR	.0314	.2102	.5203	.1585	.0385	.2201	.5641	.1667	.0430	.2274	.5739	.1677
WF	.0195	.1095	.2608	.0697	.0275	.1486	.3294	.0881	.0328	.1714	.3656	.0984
PCA	.0422	.1148	.3356	.0825	.0445	.1472	.3759	.0933	.0464	.1669	.3965	.1003
LapF	.0179	**.1064**	.2262	**.0638**	.0261	.1430	.2922	**.0828**	.0321	.1662	.3408	.0938
CPF	.0445	.1159	.2221	.0664	.0487	.1507	.2854	.0866	.0511	.1730	.3319	.0974
GPF	.0294	.1307	.3271	.0859	.0349	.1650	.3730	.0998	.0396	.1860	.4034	.1073

multi-resolution False Negative Ratio (FNR):

$$\text{FNR} = \frac{\sum_{\psi^* - \psi_j > 0} (\psi^* - \psi_j)}{\sum \psi^*} \tag{15}$$

where ψ^* and ψ are the binary edge map reconstructed from wavelet subbands of reference image and fusion result, where edge pixels are marked with 1, and 0 elsewhere. It measures the diminished image features in the fused image. Smaller FNR value represents better feature retaining performance. Tables 2 lists the average FNR results over five trials with best results high-lighted. The same test images were used but three noise STDs were evaluated. Our method gave the best feature preservation performance in nine our of twelve cases. In the three cases, its average FNRs were very close to the best. The combination of fusion and denoising obviously is less effective at recovering image features, which is also illustrated in Figure 1(d) and (e). The results were blurry in smooth regions or regions that contain fine details.

4. CONCLUSION

In this paper, we described a wavelet-based, noise-aware image fusion method to address the problem of fusing noisy inputs. Our method distinguishes signal and noise coefficients in wavelet subbands on-the-fly by computing the noise variance for every subband. Due to the nature of signal coefficients, they are fused with weighted averaging to maximize the feature retention. For noise coefficients, we avoid integration by keeping possible features that are supported by all input instances. Our method was evaluated from feature retention and noise removal aspects and experimental results demonstrated that our method greatly improved the fusion performance in the presence of noise.

With very low noise or fine texture that resembles noise, our method may perform less optimally. We believe this is a problem that originates from noise estimation. To overcome this problem, Gaussianity of subband histograms can be evaluated. If the distribution is super-Gaussian, i.e. little or no noise presence, a conventional method can be used. Otherwise, a noise-aware method is applied.

5. REFERENCES

[1] L. J. Chipman, T. M. Orr, and L. N. Graham, "Wavelets and image fusion," in *Proceeding of SPIE*, 1995, vol. 2569, pp. 208–219.

[2] H. Li, B. S. Manjunath, and S. K. Mitra, "Multisensor image fusion using the wavelet transform," *Graphical Models and Image Processing*, vol. 57, pp. 235–245, May 1995.

[3] Z. Zhang and R. S. Blum, "A categorization of multiscale-decomposition-based image fusion schemes with a performance study for a digital camera application," in *Proceeding of SPIE*, 1999, vol. 87.

[4] P. J. Burt and R. J. Kolczynski, "Enhanced image capture through fusion," in *Proceeding of the 4th International Conference on Computer Vision*, Berlin, Germany, May 1993, pp. 173–182.

[5] M. Pavel, J. Larimer, and A. Ahumada, "Sensor fusion for synthetic vision," in *SID International Symposium, Boston, Digest of Technical Papers*, Playa del Rey, CA, 1992.

[6] D. L Donoho, "De-noising by soft thresholding," *IEEE Transactions on Information Theory*, vol. 41, no. 3, pp. 613–627, May 1995.

[7] S. G. Chang, B. Yu, and M. Vetterli, "Adaptive wavelet thresholding for image denoising and compression," *IEEE Transactions on Image Processing*, vol. 9, no. 9, pp. 1532–1546, Sept. 2000.

[8] R. K. Sharma, *Probabilistic Model-based Multisensor Image Fusion*, Ph.D. thesis, Oregon Graduate Institute of Science and Technology, Oct. 1999.

[9] X. Yuan and B. P. Buckles, "Subband noise estimation for adaptive wavelet shrinkage," in *Proceedings of International Conference of Pattern Recognition*, Cambridge, UK, Aug. 2004.