OPEN ACCESS

University of BRISTOL

Link to published version (if available):
10.1109/ICIF.2007.4407977

Link to publication record in Explore Bristol Research
PDF-document

University of Bristol - Explore Bristol Research

General rights

Statistical Model-based Fusion of Noisy Multi-band Images in the Wavelet Domain

Artur Łoza, Alin Achim, David Bull and Nishan Canagarajah
Department of Electrical and Electronic Engineering
University of Bristol, UK
Email: [artur.loza, alin.achim, dave.bull, nishan.canagarajah]@bristol.ac.uk

Abstract— A new method for multimodal image fusion, based on statistical modelling of wavelet coefficients, is proposed in this paper. The algorithm draws from the Weighted Average scheme, but incorporates Laplacian bivariate parent–child statistical dependencies. The interscale dependency is brought in the form of shrinkage functions. The proposed method has been shown to perform very well with noisy datasets, outperforming other conventional methods in terms of fusion quality and noise reduction in the fused output.

Keywords: Image fusion, statistical modelling, multimodal, denoising.

I. Introduction

The purpose of image fusion is to combine information from multiple images of the same scene into a single image that ideally contains all the important features from each of the original images. In this work, a scenario is considered, when sets of multimodal images are being collected for purposes such as surveillance, monitoring, tracking, detection and recognition. In all these applications, the use of complementary information from multiple modalities is beneficial in terms of the improved performance and reduction of the information overload. However, multimodal fusion also poses problems related to the statistical diversity of the data and increased or accumulated noise content caused by the use of cheap sensors or low light conditions. It is expected that the application of an appropriate statistical model will benefit the image fusion both in terms of quality and robustness to signal corruption or changing environment. The aim of this paper is to investigate ways of employing appropriate statistical models that exploit wavelet coefficient dependencies across spatial locations and adjacent scales, in order to efficiently fuse multimodal images that may be corrupted with noise.

The majority of early image fusion approaches, although effective, have not been based on strict mathematical foundations. Only in recent years have more rigorous approaches been proposed, including those based on estimation theory [1]. A Bayesian fusion method based on Gaussian image model has been proposed in [2]. A general approach, allowing modelling both Gaussian and non-Gaussian distortions to the input images, has been proposed in [3], where a Hidden Markov Model has been used to describe the correlations between the wavelet coefficients across scales.

Recent work on non-Gaussian modelling for image fusion has been proposed in [4], where the image fusion proto-type method [5], combining images based on the "match and salience" measure (variance and correlation), has been modified and applied to images modelled by symmetric α-stable distributions. Our present research follows this direction, however, the underlying distribution is described by the Laplacian model. The advantage of the Laplacian model lies in the availability of analytical expressions for their probability density functions (pdf) as well as in simple and efficient parameter estimators. Additionally, the use of bivariate shrinkage functions, allows for noise reduction in the fused output, as well as the spatial and interscale dependencies of wavelet coefficients to be taken into account.
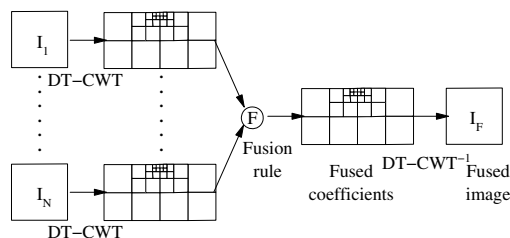

Fig. 1. Pixel-based image fusion scheme using the DT-CWT.

As is the case with many recently proposed techniques, our developments are made using the wavelet transform, which constitutes a powerful framework for implementing image fusion algorithms [4], [6]. Specifically, methods based on multiscale decompositions consist of three main steps: first, the set of images to be fused is analysed by means of the wavelet transform, then the resulting wavelet coefficients are fused through an appropriately designed rule, and finally, the fused image is synthesized from the processed wavelet coefficients through the inverse wavelet transform. This process is depicted in Fig. 1. The developments presented in this paper can be categorised as fusion rule design.

The remaining part of the paper is organised as follows. The univariate and bivariate Laplacian image distribution models are described in Section II. In Section III, a new approach to image fusion is proposed, based on bivariate Laplacian modelling of wavelet coefficients. The performance of the method is evaluated on sets of multimodal image sequences and is compared with other standard techniques in Section IV. Finally, Section V presents the conclusions of the study and suggests areas for future work.

II. GENERALISED GAUSSIAN STATISTICAL MODEL

A. Univariate Model

As a starting point towards understanding the basic properties of images, we will consider the marginal models of their wavelet transform. Although Symmetric Alpha-Stable ($S\alpha S$) modelling have been already applied to image fusion successfully [4], it was decided to investigate an alternative approach, the Generalised Gaussian Distribution (GGD), with its particular case, the Laplacian distribution. The advantage of the Laplacian model is the availability of analytical expressions and simple parameter estimators that can be used in image fusion and denoising.

A GGD family of distributions can be written as [7]

$$p(x) = K(\alpha, \beta) \exp\left(-\left|\frac{x}{\alpha}\right|^{\beta}\right), \qquad (1)$$

where α and β are the distribution parameters (scale and shape parameter, respectively), $K(\alpha, \beta) = \beta/(2\alpha\Gamma(1/\beta))$ is a normalisation constant, and $\Gamma(t) = \int_0^\infty e^{-u} u^{t-1} du$ is the gamma function. Throughout this work, only signals with zero mean are considered. Gaussian and Laplace distributions are special cases of GGD for $\beta = 2$ and $\beta = 1$, respectively. It is sometimes convenient to express α in terms of the standard deviation σ:

$$\alpha = \sigma \left(\frac{\Gamma\left(\frac{1}{\beta}\right)}{\Gamma\left(\frac{3}{\beta}\right)}\right)^{\beta/2}. \qquad (2)$$

The marginal Laplacian pdf is then written as follows

$$p(x) = \frac{1}{\sigma\sqrt{2}} \exp\left(-\frac{\sqrt{2}|x|}{\sigma}\right) \qquad (3)$$

whereas the marginal Gaussian pdf is written

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{x^2}{2\sigma^2}\right). \qquad (4)$$

Examples of Generalized Gaussian family of distributions are shown in Figure 2. Note different cusp and tail behaviour depending on the value of the shape parameter.
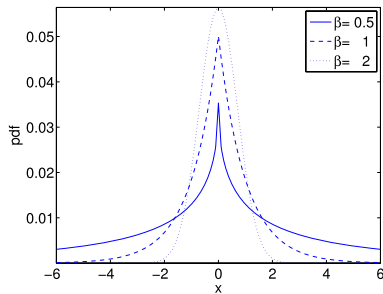

Fig. 2. Examples of Generalized Gaussian family of distributions for different values of β
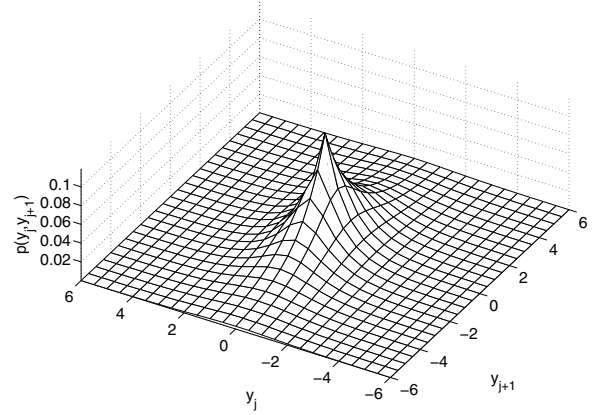

Fig. 3. Bivariate joint Laplacian distribution

B. Bivariate Model

Univariate marginal histograms characterise distributions of wavelet coefficients without taking into account the interscale dependencies. However, the work of many authors, for example, [8]–[10], shows strong evidence that coefficients of the multiscale decompositions of images exhibit significant dependencies on subsequent locations, scales and across different orientations. In this study, we consider the problem of modelling interscale dependencies between child coefficients (y_j, fine resolution) and corresponding parent coefficients (y_{j+1}, coarse resolution). A bivariate dependent Laplacian model proposed in [9] is used in this work:

$$p(y_j, y_{j+1}) = \frac{3}{2\pi\sigma_j\sigma_{j+1}} \exp\left(-\sqrt{3}\sqrt{\left(\frac{y_j}{\sigma_j}\right)^2 + \left(\frac{y_{j+1}}{\sigma_{j+1}}\right)^2}\right) \qquad (5)$$

This anisotropic model includes the isotropic case when $\sigma_j = \sigma_{j+1}$ is assumed. An example of a joint parent–child pdf is shown in Figure 3.

III. PROPOSED FUSION METHOD

A. Weighted average, WA

We reformulate and modify the Weighted Average (WA) method [5] by considering a more appropriate statistical model (Laplacian) and interscale dependencies. The WA method has been chosen as a starting point for several reasons: it is based on statistical Gaussian-like modelling, it allows easy modification and it covers a range of selection rules, from maximum, weighted average, to average.

Although in [5] wavelet coefficients variances are used as saliency measures, appropriate for the Gaussian model, alternative methods for computing WA parameters can be used. A modification of WA has been proposed in [4], based on α-stable image modelling, with the use of the distribution dispersions and symmetric covariation coefficients in place of the variance and covariance. In this correspondence, two main modifications of the WA fusion are proposed: firstly, the model describing the wavelet coefficient is assumed to be Laplacian; secondly, by incorporating wavelet shrinkage

functions into averaging, both noise corruption and the parent–child dependencies are accounted for. Both modifications are described in details in the following sections. For the completeness of the presentation we recall the original method below (based on [5] and [4]):

1) Decompose each input image into subbands.
2) For each highpass subband pair X, Y:
a) Compute saliency measures, σ_x and σ_y.
b) Compute matching coefficient

$$M = \frac{2\sigma_{xy}}{\sigma_x^2 + \sigma_y^2}, \qquad (6)$$

where σ_{xy} stands for covariance between X and Y.
c) Calculate the fused coefficients using the formula $Z = W_x X + W_y Y$ as follows:
- if $M > T$ ($T = 0.75$) then $W_{\min} = 0.5\left(1 - \frac{1-M}{1-T}\right)$ and $W_{\max} = 1 - W_{\min}$ (weighted average mode, including mean mode for $M = 1$),
- else $W_{\min} = 0$ & $W_{\max} = 1$ (selection mode),
- if $\sigma_x > \sigma_y$ $W_x = W_{\max}$ and $W_y = W_{\min}$, else $W_x = W_{\min}$ and $W_y = W_{\max}$.
3) Average coefficients in lowpass residual.
4) Reconstruct the fused image from the processed subbands and the lowpass residual.

B. Saliency and Matching Measures

We will show below how the saliency and matching measures in (6) can be computed for samples coming from the assumed distribution. The maximum likelihood estimate of variance for Laplacian distribution is

$$\sigma_x = \frac{\sqrt{2}}{K} \sum_{k=1}^{K} |x_k|. \qquad (7)$$

The required covariance measure, can be estimated as

$$\sigma_{xy} = \left(\frac{\sqrt{2}}{K} \sum_{k=1}^{K} |x_k y_k|^{\frac{1}{2}}\right)^2. \qquad (8)$$

The corresponding original WA measures, when x is assumed to be Gaussian, are the familiar variance and covariance estimates, respectively:

$$\sigma_x^2 = \frac{1}{K} \sum_{k=1}^{K} x_k^2, \quad \sigma_{xy} = \frac{1}{K} \sum_{k=1}^{K} x_k y_k. \qquad (9)$$
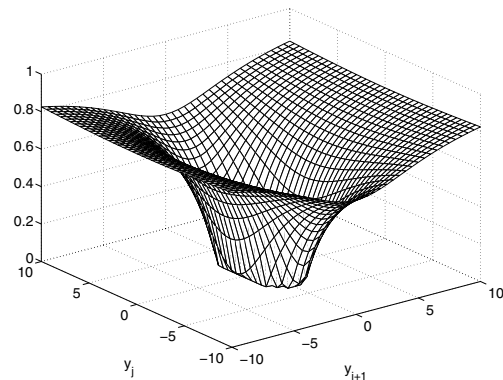

Fig. 4.   Illustration of wavelet shrinkage shrinkage function $A = \hat{w}_1/y_1$.

C. Bivariate Shrinkage Functions

In the following, we propose to incorporate shrinkage functions into the fusion rule to achieve a degree of noise reduction in the fused image. Shrinkage of wavelet coefficients has became a very efficient tool for denoising images [9]–[11]. Recently, the one-dimensional 'soft-' and 'hard-thresholding' functions [11] have been modified to allow modelling of interscale dependencies using Laplacian [9] distributions. The shrinkage functions are derived as maximum a posteriori estimators $\hat{w}_j$ of noisy wavelet coefficient y_j corrupted by Gaussian white noise n: $y_j = w_j + n$, $n : N(0, \sigma_n)$. In particular, for the model (5) used in this study, the following shrinkage relationship was derived in [9]:

$$\hat{w}_j = Ay_j = \frac{\left(r - \frac{\sqrt{3}\sigma_n^2}{\sigma_{y_j}}\right)_+}{r} y_j \qquad (10)$$

where $r = \sqrt{y_j^2 + y_{j+1}^2}$ and $(x)_+$ is a thresholding operator, setting negative values to zero. The shrinkage function A above has been limited to an isotropic case (equal variances of parent and child coefficients assumed) based on the recommendations of authors of [9], who claim that there is a negligible improvement achieved from the use of the anisotropic distribution. The input-output relationships between noisy and estimated coefficients, and their interscale dependency is evident when the shrinkage function (also called transfer function), $A = \hat{w}_j/y_j$, is analysed (see Figure 4). For example, it can be seen that some small values of y_j that would normally be set to zero in univariate shrinkage, are not zeroed but only shrunk, if the corresponding parent coefficient y_{j+1} is large. Also, large child coefficients are scaled depending not only on their original value but also on their parent coefficient value.

The noise standard deviation σ_n required to compute the shrinkage functions, is obtained from the observed data by computing the Median Absolute Deviation (MAD) of coefficients at the first level of an wavelet decomposition [11]

$$\hat{\sigma}_n = \frac{MAD(y_1)}{0.6745}. \qquad (11)$$

D. Joint Image Fusion and Denoising

In previous paragraphs, we have presented the components of the proposed image fusion method. The complete algorithm can be summarised as follows. The appropriate statistical model has been incorporated into the WA scheme by replacing in (6) the original Gaussian saliency measures (9) with their Laplacian counterparts (7–8). Then, the bivariate shrinkage functions, derived from the assumed model, have been applied to wavelet coefficients of the input images.

Let us finally illustrate the significance of using shrinkage functions into the WA scheme by rewriting the modified matching measure (6):

$$M = \frac{2A_x A_y \sigma_{xy} + \varepsilon}{A_x^2 \sigma_x^2 + A_y^2 \sigma_y^2 + \varepsilon},$$
(12)

where A_x and A_y denote the shrinkage functions calculated for a subband of an input image and ε is a small constant to stabilise the measure when both nominator and denominator are 0. In the proposed method, the saliency measures $A_x^2 \sigma_x^2$ and $A_y^2 \sigma_y^2$ are used when deciding which coefficient should be selected or assigned a greater weight to.

IV. RESULTS

In this section we show results obtained using the proposed method. The results are then compared to some common image fusion methods operating in the wavelet domain. Fusion performance is assessed by means of quality metrics, noise contents and visual inspection.

A. Fusion Methods

All the methods included in the comparison use the Dual-Tree Complex Wavelet Transform (DT-CWT) [12], as implemented in [13]. This type of wavelet transform has been shown to be nearly shift-invariant and have better directional selectivity compared to conventional Discrete Wavelet Transform [12]. It was found experimentally, that *AntonB* filters with 5 decomposition levels gave the best results. The fusion methods used in the comparison are briefly described below.

a) Maximum selection, MAX: The simplest and most popular fusion method consisting of building the fused image by picking up the coefficient with maximum absolute value among all inputs at each location in the corresponding sub-bands of the wavelet decomposition (see [6], for example).

b) Weighted Average with Gaussian model, WA: An implementation based on [5] (described also in Section III-A), with a 5 pixel (4-connected) window for computing local statistics of the DT-CWT coefficients.

c) Weighted average with Laplacian model, LAP: The method we propose is based on WA, however, since the underlying model is Laplacian, different saliency and matching measure have been used (see Section III-D).

d) Fusion with bivariate shrinkage, SHR: The method described in Section III-D combining Laplacian modelling as in LAP, but with bivariate shrinkage function (10), applied to wavelet coefficients prior to parameter estimation and fusion.

e) Fusion with 'partial' shrinkage, SHP: Here we also propose fusion of not only 'completely' (as in SHR) but also 'partially' denoised images. This somewhat heuristic, option may be used in circumstances when lower degree of denoising is required, for example in low-noise scenarios. In such cases denoising could remove useful visual information, readable for a human observer or computer algorithm despite of noise presence. We propose to denoise only those coefficients that fall into the 'selection mode' category, i.e. when the matching measure value is small. The motivation for not denoising coefficient in weighted average mode is twofold: the high matching measure (and thus a high degree of similarity between the inputs) indicates low noise contents, also, the operation of averaging reduces the noise level to some degree, so that applying shrinkage may not be necessary.

B. Datasets

In order to evaluate the fusion methods, datasets extracted from two different sources have been used. *Tropical* dataset, contains stills from visible and infrared light videos recorded during Eden data gathering by University of Bristol [14]. This imagery is characterised by low light conditions and frequent occlusion of the objects of interest, affecting the visible light sequences most.

A second dataset, *Aviris*, contains images selected from the public AVIRIS 92AV3C hyperspectral database collected over a test site called Indian Pine in northwestern Indiana [15]. In this work, we fuse pairs of manually selected bands. Extension of our method to multiple bands, possibly automatically selected, is the subject of ongoing investigation.

C. Performance Metrics

Two computational metrics were used to evaluate the quality of fusion: a quality index measuring similarity (in terms of illuminance, contrast and structure) between the input images and the fused images [16] (Q_1); and the Petrovic metric [17] (Q_2) measuring the amount of edge information transferred from the source images to the fused image. In order to fully evaluate the performance of the methods, we have also shown the improvement in terms of Signal to Noise Ratio (SNR), which has been calculated as follows:

$$\Delta\text{SNR} = 10 \log \frac{\sigma_n^2}{\sigma_{n,f}^2}.$$

Here $\sigma_{n,f}$ is the noise variance in the fused output. The noise variance has been estimated according to (11). The results presented below are averaged over 10 image subsets drawn from the datasets.

D. Fusion of Noisy Images

To evaluate the performance of the fusion methods in a noisy environment the input images have been corrupted with Gaussian white noise, with standard deviation σ_n scaled so that the input SNR is 10 dB. Table I shows average metric values of fused images, calculated between fused output and original clean input images. As can be seen from Table I and

TABLE I
METRIC-BASED RANKINGS OF FUSION METHODS, NOISY IMAGES (SNR=10 DB)

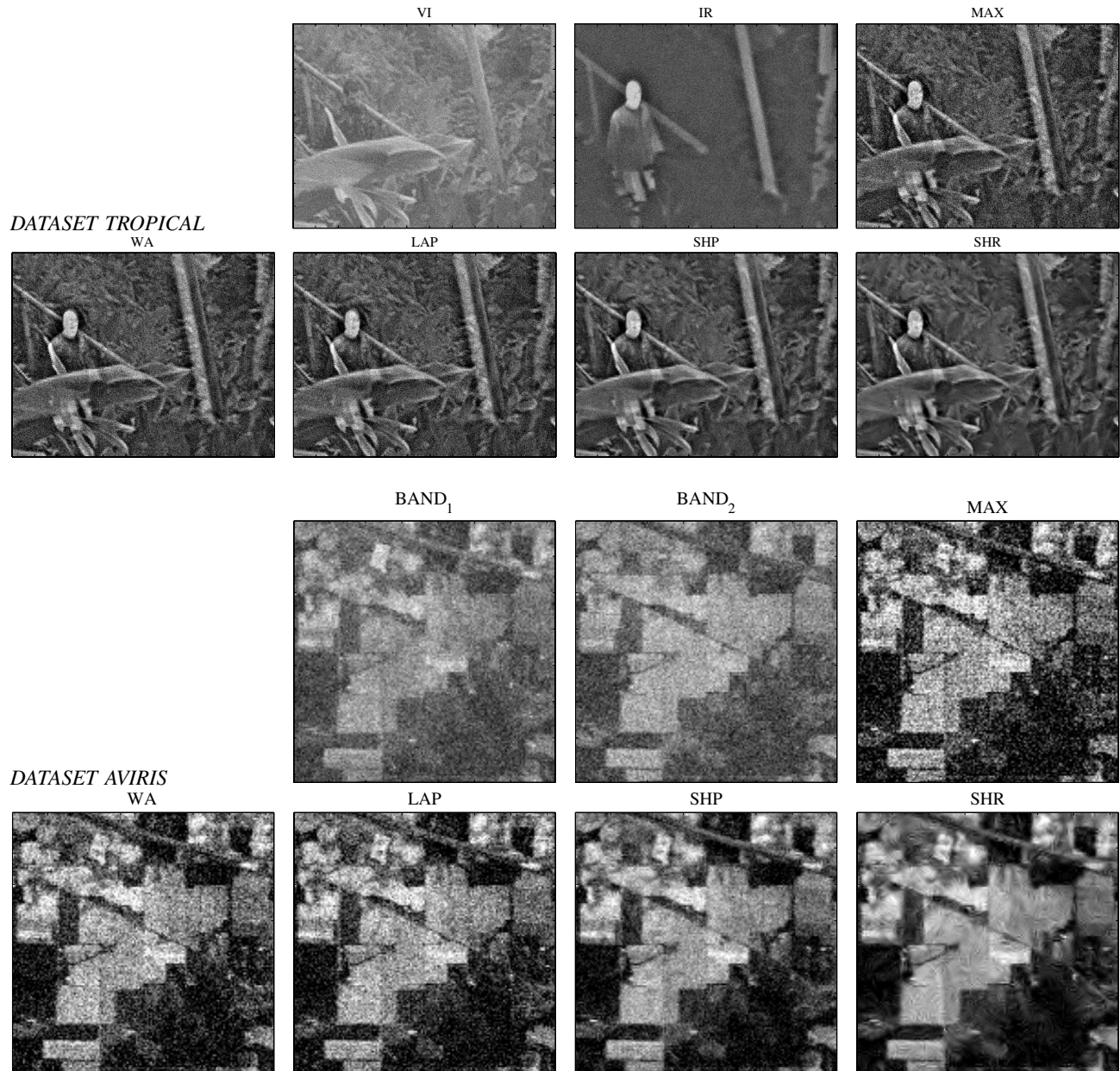| Tropical						Aviris					
Q_1		Q_2		Δ SNR, dB		Q_1		Q_2		Δ SNR, dB	
SHR	0.516	SHR	0.290	SHR	340.9	SHR	0.696	SHR	0.496	SHR	344.3
SHP	0.398	SHP	0.242	SHP	3.5	SHP	0.617	SHP	0.470	SHP	4.3
LAP	0.274	LAP	0.179	LAP	-2.1	LAP	0.506	LAP	0.399	LAP	-1.0
WA	0.266	WA	0.174	WA	-3.0	WA	0.488	WA	0.385	WA	-2.1
MAX	0.240	MAX	0.156	MAX	-4.8	MAX	0.462	MAX	0.364	MAX	-3.9

Fig. 5.   Examples of fused noisy images (SNR=10 dB)

Figure 5, highly consistent results have been achieved both in terms of the quality measures and visual perception. According to the quality measures shown, all three proposed methods (LAP, SHP and SHR) outperform significantly WA and MAX fusion. In all cases, some improvement has been achieved by using Laplacian modelling alone, without even applying shrinkage. It can be also observed that MAX fusion appears to be very sensitive to the presence of noise. By looking at the image examples shown in Figure 5, it can be confirmed that the denoised (SHR, SHP) images look the most pleasing for the *Tropical* sequence. For *Aviris* dataset, it seems, however, that SHR 'oversmoothed' some visually important features and thus the SHP and LAP methods are favoured as convincing trade-off between noise reduction and information loss.

V. CONCLUSIONS AND FUTURE WORK

In this paper we introduced a modification of WA fusion, that has produced an improvement in quality and noise robustness of image fusion. The improvement has been achieved in two ways: i) by using a more appropriate Laplacian statistical model for wavelet coefficients, and ii) by incorporating bivariate shrinkage functions into the weighting and saliency measures. By doing so, the images are combined based both on their local saliency and the noise levels, with interscale dependencies taken into account. The proposed method has been shown to perform very well with noisy datasets, outperforming the conventional WA and MAX algorithms. The method has also been shown to reduce significantly the noise variance in the fused output images.

Laplacian modelling, although based on a simplification ($\beta = 1$ assumed), appears to offer a good complexity–accuracy trade-off compared to sophisticated shape parameter estimation. The precision of such a model can be improved by using other, for example noise robust, estimators of the salience measures. Alternatively, if simplicity is not a priority, a more complex model, such as GGD can be employed instead of the Laplacian.

Although the bivariate shrinkage functions rely on parent–child dependencies, their distributions could be modelled more explicitly. For example, an image fusion method could be based on conditional saliency measures derived from such distributions. Also, the multidimensional extension of our method needs to be derived if multiple inputs are to be fused.

Additional benefits of the developed method should also be verified experimentally, for example by performing object tracking or detection in fused images or videos.

REFERENCES

[1] R. S. Blum, "On multisensor image fusion performance limits from an estimation theoryperspective," *Information Fusion*, vol. 7, no. 3, pp. 250–263, Sep 2006.
[2] R. Sharma and M. Pavel, "Adaptive and statistical image fusion," *Society for Information Display Digest*, vol. 17, no. 5, pp. 969–972, May 1996.
[3] J. Yang and R. S. Blum, "A Statistical Signal Processing Approach to Image Fusion using Hidden MarkovModels," in *Multi-Sensor Image Fusion and Its Applications.* Marcel Dekker/CRC, 2005.
[4] A. M. Achim, C. N. Canagarajah, and D. R. Bull, "Complex wavelet domain image fusion based on fractional lower order moments," in *Proc. of the 8th International Conference on Information Fusion, Philadelphia PA, USA, 25–29 July*, 2005.
[5] P. Burt and R. Kolczynski, "Enhanced image capture through fusion," in *Proc. 4th International Conference on Computer Vision*, Berlin 1993, pp. 173–182.
[6] S. G. Nikolov, P. Hill, D. Bull, and N. Canagarajah, "Wavelets for image fusion," in *Wavelets in Signal and Image Analysis*, A. Petrosian and F. Meyer, Eds. Kluwer Academic Publishers, 2001, pp. 213–244.
[7] S. G. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 7, pp. 674–693, July 1989.
[8] E. P. Simoncelli, "Modeling the joint statistics of images in the wavelet domain," in *Proceedings of SPIE 44th Annual Meeting*, vol. 3813, Denver, CO, USA, Jul 1999, pp. 188–195.
[9] L. Sendur and I. Selesnick, "Bivariate shrinkage functions for wavelet-based denoising exploiting interscaledependency," *IEEE Transactions on Signal Processing*, vol. 50, no. 11, pp. 2744– 2756, Nov 2002.
[10] A. Achim and E. Kuruoglu, "Image denoising using bivariate α-stable distributions in the complexwavelet domain," *IEEE Signal Processing Letters*, vol. 12, no. 1, pp. 17–20, Jan 2005.
[11] R. Coifman and D. Donoho, "Time-invariant wavelet denoising," in *Wavelet and Statistics, vol. 103, Lecture Notes in Statistics*, A. Antoniadis and G. Oppenheim, Eds. Springer-Verlag, 1995, pp. 125–150.
[12] N. Kingsbury, "Complex wavelets and shift invariance," Time-Scale Representations and Time-Frequency Analysis and Applications, IEE Electronics and Communications Seminar, London, February 2000.
[13] S. Cai and K. Li, "Matlab implementation of wavelet transforms," 2003, available at http://taco.poly.edu/WaveletSoftware.
[14] J. J. Lewis, S. G. Nikolov, A. Łoza, E. F. Canga, N. Cvejic, J. Li, A. Cardinali, C. N. Canagarajah, D. R. Bull, T. Riley, D. Hickman, and M. I. Smith, "The Eden Project Multi-Sensor Data Set," in *Technical Report TR-UoB-WS-Eden-Project-Data-Set*, Available at http://www.imagefusion.org, 2006.
[15] Online:, "Airborne visible/infrared imaging spectrometer," available at http://aviris.jpl.nasa.gov/.
[16] N. Cvejic, A. Łoza, C. N. Canagarajah, and D. R. Bull, "A similarity metric for assessment of image fusion algorithms," *International Journal Of Signal Processing*, vol. 2, no. 2, pp. 178–182, 2005.
[17] V. S. Petrovic and C. S. Xydeas, "Sensor noise effects on signal-level image fusion performance." *Information Fusion*, vol. 4, pp. 167–183, 2003.