Belief Propagation Based Image Fusion

Paul Hill, *Member, IEEE*, David Bull, *Fellow, IEEE*

***Abstract*—This letter describes the application of belief propagation methods to image fusion within a complex wavelet decomposition (the Dual Tree Complex Wavelet). Belief propagation within each transform subband iterates through a lattice based Bayesian belief network. This leads to precisely controlled spatial coherence of subband coefficient fusion through the definition of belief graph probabilities. This results in a significant improvement in quantitatively measured fusion performance for a large database of over 160 fusion image pairs from a range of fusion applications including remote sensing, multi-focus and multi-modal sources. Improvements in qualitative image fusion performance is also demonstrated.**

***Index Terms*—Image Fusion, Wavelets, Bayesian Belief Networks, Bayesian Belief Propagation**

I. Introduction

IMAGE fusion is the process of combining two or more images to produce a single fused output that combines important perceptual elements of the inputs. Effective fusion offers significant benefits for scene understanding, visualisation, target location / recognition and situational awareness within such diverse domains as remote sensing, surveillance, medicine and military applications. Image fusion can be implemented for just a single image pair, multiple images, images of different modalities and resolutions, or by extension, video sequences. This letter focuses on the fusion of two registered images of the same resolution.

A large and diverse number of methods have been used for image fusion. These include simple pixel based methods and those derived from Principle Component Analysis (PCA) on a pixel and transform basis [1]. Transform based methods have also been used extensively. These transform based methods include those based on the following transformations: the Intensity-Hue-Saturation (IHS) transform [2], multiscale transforms such as pyramid decompositions [3], [4], wavelet, complex wavelet, undecimated wavelet [5]–[7] and total variational methods [8].

Image fusion using the Dual Tree Complex Wavelet Transform (DT-DWT [9]) has been recognised as being highly effective due to improved directional selectivity, approximate shift invariance and complex subband analysis of the DT-CWT. This letter has focused on using this framework (introduced by Hill et al. [10]). A significant issue with any wavelet fusion is the spatial incoherence of the subband based fusion process (spatial coherence is how consistent the choice of subband coefficient, between the two input image coefficient choices, is within small spatial areas). Spatial coherence has been previously addressed using simple weighting and majority filters [3], [11], [12]. However, the contribution of this paper is the application of Belief Propagation techniques in order to flexibly, optimally and globally control spatial coherence and therefore improve quantitative and qualitative results.

This work was funded by the EPSRC Platform grant EP/M000885/1.
P. Hill and D. Bull are with the Department of Electrical and Electronic Engineering, The University of Bristol, Bristol, BS8 1UB, UK e-mail: (paul.hill@bristol.ac.uk) Digital Object Identifier:

II. Review

A. Wavelet Based Image Fusion

Early fusion methods based on pyramid decompositions have now largely been superseded by Discrete Wavelet Transform (DWT)-based methods [13]–[15]. The fusion of two sources utilising the DWT can be defined in terms of the two registered input sources $\boldsymbol{I_0}$ and $\boldsymbol{I_1}$, the wavelet transform itself $\boldsymbol{\omega}$ and a fusion rule $\boldsymbol{\theta}$, defined to combine co-located coefficients within the transform domain. The fused wavelet coefficients are then inverted using an inverse wavelet transform $\boldsymbol{\omega^{-1}}$ to produce the resulting fused image $\boldsymbol{F}$, thus:

$$\boldsymbol{F} = \boldsymbol{\omega^{-1}}(\boldsymbol{\theta}(\boldsymbol{\omega}(\boldsymbol{I_0}), \boldsymbol{\omega}(\boldsymbol{I_1}))). \tag{1}$$

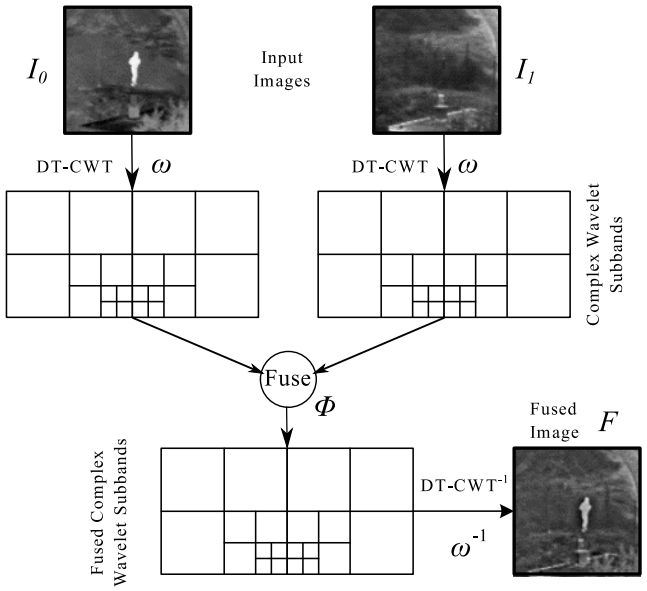

Fig. 1: Fusion of two images using the DT-CWT

We have used the DT-CWT for our experiments as it has been found to give excellent results due to its approximate shift invariance and improved directional selectivity compared to the DWT [10]. Figure 1 illustrates the utilised two image fusion scenario using the DT-CWT.

The most basic fusion rule for wavelet domain fusion is the "choose maximum" rule (where the largest magnitude coefficient from either image is retained in the fused output). Although giving good results, this fusion rule can lead to isolated coefficients and general spatial incoherence.

Previously defined wavelet based image fusion methods have integrated spatial consistency methods [11], [12]. For example Li's method [3] uses a local majority filter (chooses the current coefficient based on the largest number maximum choices in a local window). To obtain the output coefficient $v(k,l)$ (at spatial position k,l), the maximum count $n(k,l)$ within a local window R can be expressed as

$$n\left(k,l\right)=\sum_{(i,j)\in R}\mathbb{1}_{\mathbb{R}_{\geq 0}}\left(S_0(i+k,j+l)-S_1(i+k,j+l)\right) \tag{2}$$

where S_0 and S_1 are the current subbands from each image and $\mathbb{1}$ is the indicator function. The output coefficient $v(k,l)$ is chosen as

$$v(k,l)=\begin{cases}S_0(k,l) & n_{(k,l)}>L/2\\ S_1(k,l) & \text{otherwise}\end{cases} \tag{3}$$

where L is the number coefficients in the local region R. This, and associated methods are very rudimentary as they don't globally optimise the control of fusion coherence.

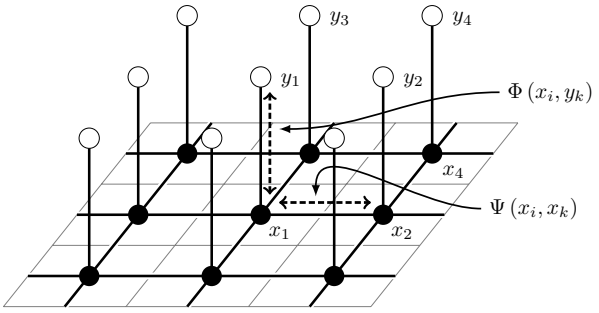

Fig. 2: Bayesian Graph Model for Belief Propagation within a single wavelet subband (the black dots represent the hidden states (x) and the white dots represent the observations (y)).

III. Approach

A. Belief Propagation for Image Fusion

Belief propagation is implemented on a subband by subband basis in order to accurately control the spatial coherence of the fusion process within the wavelet transform. Figure 2 shows the graphical model structure for the belief propagation system. One graphical model is defined for each wavelet subband. Each of the observed and hidden nodes within the model are associated with a single coefficient within the considered subband. The hidden nodes (denoted x_k) are 4-connected across the entire subband and co-located with associated subband coefficients at index position k. Additionally, one observed node (denoted y_k) is associated (and singly connected) to each hidden node (x_k). For the fusion of two images, each hidden node can be in one of two states; one state associated with each of the input image coefficients. The observed state relates to the observations of the coefficient magnitudes from each image.

Although this network contains loops, it has been found that stable results can still be obtained with looped networks though belief propagation [16].

The probability of the choice of one image coefficient (out of the two possible) is proportional to the product of all sets of compatibility matrices Ψ and vectors Φ [16], [17]:

$$P(x|y)=\frac{1}{Z}\prod_{(i,j)}\Psi_{ij}\left(x_i,x_j\right)\prod_{i}\Phi_i\left(x_i,y_i\right), \tag{4}$$

where the first product is over spatial neighbours i and j and Z is a normalising constant. $\Psi_{ij}\left(x_i,x_j\right)$ and $\Phi_i\left(x_i,y_i\right)$ are the pairwise compatibility functions (also shown in figure 2).

Equation (4) is difficult to evaluate for any non trivial case. However, it can be evaluated using an iterative update method known as belief-propagation (BP). BP uses a message-passing system that updates "messages" $\boldsymbol{m_{ij}}$ from hidden node x_i to x_j. These "messages" are two dimensional vectors (for the two image fusion case). Using $m_{ij}\left(x_j\right)$ to denote a component of $\boldsymbol{m_{ij}}$ associated with x_j (a possible subband coefficient) this message can be updated using [16], [18]:

$$m_{ij}\left(x_j\right)=\sum_{x_i}\Psi_{ij}\left(x_i,x_j\right)\prod_{k\neq j}m_{ki}\left(x_i\right)\Phi_i\left(x_i,y_i\right). \tag{5}$$

The summation is over all the possible image coefficients x_i at node i (2 in this case). The product is over all 4-connected neighbours of the node i (with the exception of node j). When this iterative update has converged, the BP estimate of the marginal probability vector $\boldsymbol{b_i}$ can be found using:

$$b_i\left(x_i\right)=\prod_{k}m_{ki}\left(x_i\right)\Phi_i\left(x_i,y_i\right), \tag{6}$$

where $b_i\left(x_i\right)$ is the component of $\boldsymbol{b_i}$ associated with image coefficient x_i. The MAP estimate for the output coefficient x_{jMAP} can be chosen as the maximum component within $\boldsymbol{b_i}$

$$x_{jMAP}=\text{argmax}_{x_j}b_i\left(x_i\right), \tag{7}$$

B. Compatibility Functions

Φ is vector valued with each component representing the compatibility between each hidden state and the observations. The components of Φ reflect the compatibility of observed coefficient magnitudes (for each image) to an ideal output coefficient. Each component of the Φ vector (associated with each image) is therefore set to:

$$\Phi\left(x_k,y_k\right)=\exp\left(-\frac{d(x_k,y_k)}{2\sigma^2}\right) \tag{8}$$

where $d(x_k,y_k)$ is a distance measure between the hidden state x_k and its associated observation y_k. This is defined as $d(x_k,y_k)=S_{max}-|c_k|$ where $|c_k|$ is the magnitude of the subband coefficient c_k and S_{max} is the maximum of $|c_k|$ for both image subbands. The subband magnitudes are normalised to unit range and therefore $S_{max}=1$. The larger the observed coefficient the more "compatible" the coefficient

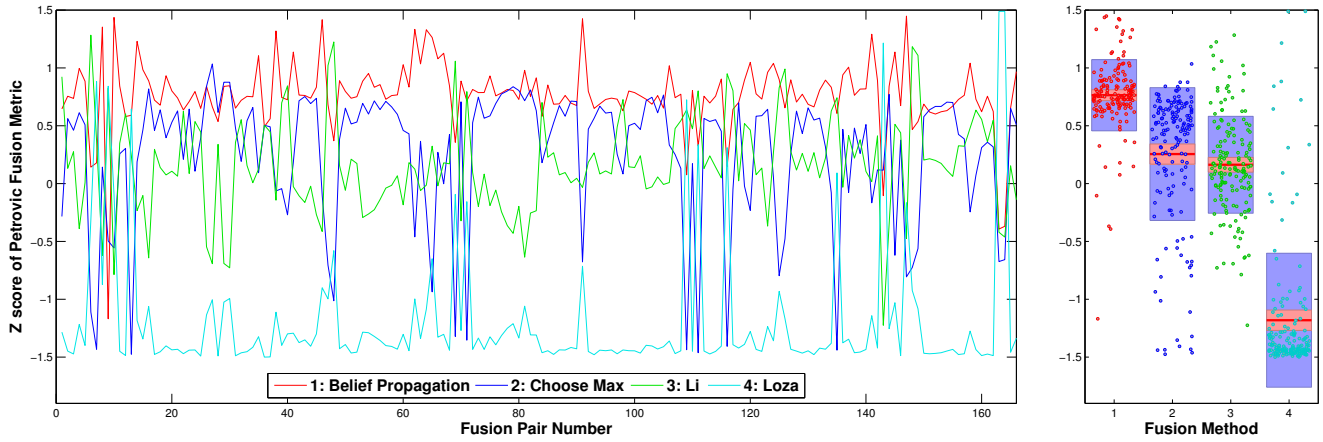

Fig. 3: Z score results for Petrovic Metric. Left hand side shows the results for all 166 image fusion pairs (indexed on the x axis). Right hand side, the same datapoints are overlaid on a 1.96 SEM (95%) confidence interval in red and a SD interval in blue.

is with the ideal fused value and therefore the larger the value of $\Phi(x_k, y_k)$ (according to (8)).

Ψ is matrix valued with the elements representing the compatibility of a hidden state x_i with its neighbour x_j. In the case of a two image fusion system this is a 2×2 matrix. These values can be calculated from the coefficients. However, our aim is the control of the spatial coherence of the fusion process. We therefore set the compatibility of neighbouring nodes with the same hidden states to be a constant (ρ) and the compatibility of different hidden neighbouring states to be zero:

$$\Psi = \begin{bmatrix} \Psi_{11} & 0 \\ 0 & \Psi_{22} \end{bmatrix}, \tag{9}$$

where $\Psi_{11} = \Psi_{22} = \rho$. Ψ can be defined separately for each of the 4-connected directions according to application requirements. However, they are defined as being equal within our two image fusion case.

C. Implementation and Experimental Conditions

The implementation of equations (5) and (6) requires a large number of products. Through the common transformation of likelihoods, compatibilities and messages to the log domain, equations (5) and (6) can be implemented using iterative additions saving considerable computation. Equations (4)-(7) are transformed to (10)-(13).

$$\Phi(x_k, y_k) = \frac{d(x_k, y_k)}{2\sigma^2}, \qquad \Psi = \ln(\rho I_2) \tag{10}$$

$$m_{ij}(x_j) = \sum_{x_i} \Psi_{ij}(x_i, x_j) + \sum_{k \neq j} \left(m_{ki}(x_i) + \Phi_i(x_i, y_i) \right), \tag{11}$$

$$b_i(x_i) = \sum_{k} \left(m_{ki}(x_i) + \Phi_i(x_i, y_i) \right), \tag{12}$$

$$x_{jMAP} = \text{argmin}_{x_j} b_i(x_i), \tag{13}$$

setting ρ to be 0.3679, Ψ is now the identity matrix I_2. σ is set to 0.1342 for all the experiments.

1) Weighted Update: It was found that local minima were prevented and the overall control of belief propagation more precisely controllable using a smoothed update of the messages from one iteration to the next. Specifically, defining a message $m_{t,ij}(x_j)$ at iteration, $t+1$ equation (11) is transformed to a weighed sum of the new and old messages

$$m_{t+1,ij}(x_j) = (1-\alpha)\, m_{t,ij} + \alpha \left(\sum_{x_i} \Psi_{ij}(x_i, x_j) + \sum_{k \neq j} \left(m_{t,ki}(x_i) + \Phi_i(x_i, y_i) \right) \right), \tag{14}$$

where α is the weighting parameter (set to 0.75 for all experiments). Also, then equation (12) becomes

$$b_i(x_i) = \sum_{k} \left(m_{T,ki}(x_i) + \Phi_i(x_i, y_i) \right), \tag{15}$$

where T is the index of the last iteration.

IV. Results

A. Dataset

We employed the extensive dataset used by Petrovic in [19]. It is noted that the majority of previous papers on image fusion have used a very small dataset. In contrast here, we present an in-depth statistical evaluation of over 166 image fusion pairs.

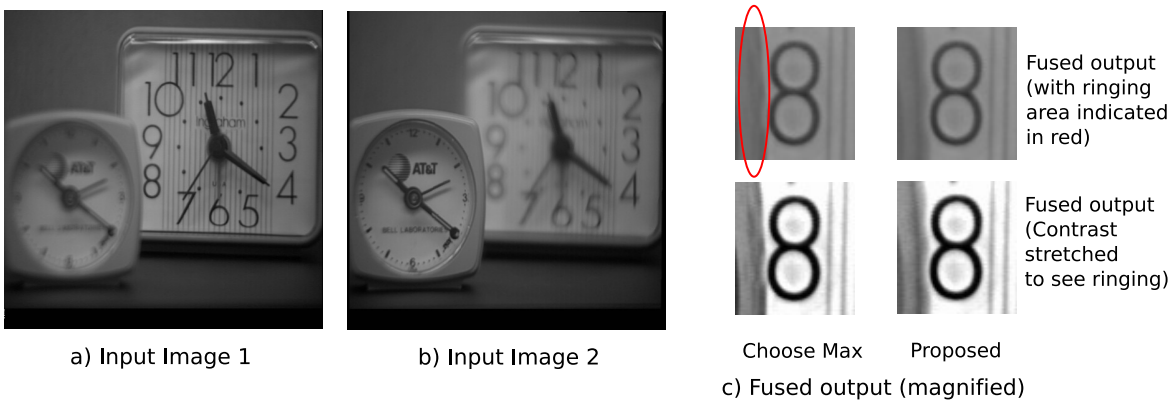

Fig. 4: (a) and (b) Multifocus image set (image pair index 39 within Petrovic dataset [19]). (c) Proposed and choose maximum method fused result (enlarged). The lower images are contrast stretched to better visualise the ringing artefacts of the choose max method compared to the proposed method. Ringing artefacts can be seen to the left of the letter 8 in the choose maximum fused image compared to the proposed fused image.

B. Image Fusion Metrics

Numerous image fusion metrics have been proposed. However, within this letter we focus on reference free metrics due to the lack of ground truth within our dataset.

The two universally adopted image fusion metrics are those developed by Petrovic [19], [20] and Piella [21]. We exclusively use the Petrovic metric here because: (i) Petrovic provided an extensive study of the correlation of his metric with subjective tests using the same dataset of 166 images we have used and (ii) the Piella metric uses the luminance aspects of the SSIM metric [22] that are not appropriate for all the images in the dataset.

C. Z scores

The Petrovic image fusion metric [20] is used to compare: the proposed method, the choose maximum method, Li's method [3] and Loza's method [12] for all of the 166 fusion image pairs within the image database [19]. All these methods use the DT-CWT and therefore give a fair comparison. When calculating the mean and variance values across all methods for each image pair there are significant variations across the entire dataset. The presentation of the raw results without compensation for these variations would therefore not show the relative performance between the considered methods across the dataset. The results of each method and image pair are therefore converted to Z-scores. The left hand side of figure 3 shows the results across the entire dataset for the four methods. The right hand side of the same figure shows the same datapoints grouped together with each method where datapoints are laid over a 1.96 SEM (95% confidence interval) in red and a 1 SD (in blue).

To quantify these results, three hypotheses are defined: the mean score of the proposed method is higher than the mean of each of the other methods (i.e. using the results shown in figure 3).

In order to test these hypotheses, a right-tailed, unpaired, t-test was done on the entire dataset comparing the Z scores of the proposed method to the other three methods. The resulting P-value was calculated for each image (the probability that the null hypothesis was true).

- 0.28555: P-value for Proposed vs Choose maximum [10]
- 0.16438: P-value for Proposed vs Li [3]
- 0.00044: P-value for Proposed vs Loza [12]

Figure 4 shows qualitative improvements in the fusion results of an example pair of multi-focus images (image pair index 39 within the database). There is significantly less ringing on the left of the figure "8" within the proposed method results compared to the nearest performing alternative (choose maximum). Although the maximum P value is fairly large, its value depends on the number and type of methods. Additionally, our proposed method gives the best overall results compared to all other methods for a very large dataset including a large number of modalities, domains and applications. Improved results can be anticipated through the precise application tuning of parameters not available for other methods.

V. Conclusion

This letter presents an effective new structure for image fusion using Bayesian belief propagation within each wavelet subband of the DT-CWT. The use and definition of belief graph network probabilities gives precise and application-specific tunable control of spatial coherence throughout the fusion process. No such control is available for previously developed methods. The structure also gives the potential to integrate cross scale, orientation and temporal (in the case of video sequences) graph connections that will be able to exploit correlations in these directions. Future work will focus on these areas. The method demonstrates qualitative and quantitative improvements over comparable methods for a database of over 160 image pairs. Previous studies do not commonly use such large databases.

References

[1] U. Patil and U. Mudengudi, "Image fusion using hierarchical PCA." in *image Information Processing (ICIIP), 2011 International Conference on*. IEEE, 2011, pp. 1–6.

[2] Y. Leung, J. Liu, and J. Zhang, "An improved adaptive intensity–hue–saturation method for the fusion of remote sensing images," *IEEE Geoscience and Remote Sensing Letters*, vol. 11, no. 5, pp. 985–989, 2014.

[3] H. Li, B.S. Manjunath and S.K. Mitra, "Multi-sensor image fusion using the wavelet transform," *IEEE International Conference on Image Processing (ICIP)*, vol. 1, pp. 51–55, November 1994.

[4] T. Mertens, J. Kautz, and F. Van Reeth, "Exposure fusion: A simple and practical alternative to high dynamic range photography," in *Computer Graphics Forum*, vol. 28, no. 1. Wiley Online Library, 2009, pp. 161–171.

[5] B. Yu, B. Jia, L. Ding, Z. Cai, Q. Wu, R. Law, J. Huang, L. Song, and S. Fu, "Hybrid dual-tree complex wavelet transform and support vector machine for digital multi-focus image fusion," *Neurocomputing*, vol. 182, pp. 1–9, 2016.

[6] A. Ellmauthaler, C. L. Pagliari, and E. A. da Silva, "Multiscale image fusion using the undecimated wavelet transform with spectral factorization and nonorthogonal filter banks," *IEEE Transactions on image processing*, vol. 22, no. 3, pp. 1005–1017, 2013.

[7] P. Hill, M. Al-Mualla, and D. Bull, "Perceptual image fusion using wavelets." *IEEE transactions on image processing: a publication of the IEEE Signal Processing Society*, 2017.

[8] J. Ma, C. Chen, C. Li, and J. Huang, "Infrared and visible image fusion via gradient transfer and total variation minimization," *Information Fusion*, vol. 31, pp. 100–109, 2016.

[9] I.W. Selesnick, R. G. Baraniuk and N. Kingsbury, "The dual-tree complex wavelet transform - A coherent framework for multiscale signal and image processing," *IEEE Signal Processing Magazine*, vol. 22, no. 6, pp. 123–151, November 2005.

[10] P.R. Hill, C.N. Canagarajah and David R. Bull, "Image Fusion Using Complex Wavelets," *Procceeding of the 13th British Machine Vision Conference*, pp. 487–496, 2002.

[11] P.J. Burt and R.J. Kolczynski, "Enhanced image capture through fusion," *Proceedingsof the 4th International Conference on Computer Vision*, pp. 173–182, 1993.

[12] A. Loza, D. Bull, N. Canagarajah, and A. Achim, "Non-Gaussian model-based fusion of noisy images in the wavelet domain," *Computer Vision and Image Understanding*, vol. 114, no. 1, pp. 54–65, 2015.

[13] S. Nikolov, P.R. Hill, D.R. Bull, C.N. Canagarajah, "Wavelets for image fusion," in *Wavelets in Signal and Image Analysis, from Theory to Practice*, A. Petrosian and F. Meyer, Eds. Kluwer Academic Publishers, 2001.

[14] P. Hill, "Wavelet Based Texture Analysis and Segmentation for Image Retrieval and Fusion," *PhD, Department of Electrical and Electronic Engineering, University of Bristol*, 2002.

[15] L. Chipman, T. Orr and L. Graham, "Wavelets and image fusion," *Wavelet Applications in Signal and Image Processing III*, vol. 2569, pp. 208–219, November 1995.

[16] W. T. Freeman, E. C. Pasztor, and O. T. Carmichael, "Learning low-level vision," *International journal of computer vision*, vol. 40, no. 1, pp. 25–47, 2000.

[17] J. Besag, "Spatial interaction and the statistical analysis of lattice systems," *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 192–236, 1974.

[18] J. Pearl, *Probabilistic reasoning in intelligent systems: networks of plausible inference*. Morgan Kaufmann, 2014.

[19] V. Petrović, "Subjective tests for image fusion evaluation and objective metric validation," *Information Fusion*, vol. 8, pp. 208–216, 2007.

[20] C. Xydeas and V. Petrovic, "Objective image fusion performance measure," *Electronics Letters*, vol. 36, no. 4, pp. 308–309, 2000.

[21] G. Piella and H. Heijmans, "A New Quality Metric for Image Fusion," *IEEE International Conference on Image Processing (ICIP)*, pp. 176–203, 2003.

[22] Z. Wang, A. C. Bovik, H. R. Sheikh and E. P. Simoncelli, "Wavelets for Image Image quality assessment: From error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, April 2004.