Objective Image Fusion Performance Measure

C. S. Xydeas, V. Petrović

A measure for objectively assessing pixel level fusion performance is defined. The proposed metric reflects the quality of visual information obtained from the fusion of input images and can be used to compare the performance of different image fusion algorithms. Experimental results clearly indicate that this metric is perceptually meaningful.

Introduction: The recent availability of multisensor systems in key image application areas, such as remote and airborne sensing, motivated researchers to work on image fusion in general and pixel level image fusion in particular. Thus a plethora of pixel level fusion algorithms have been developed [1,2,3] with different performance and complexity characteristics. Fusion performance is mainly assessed using informal subjective preference tests and, so far, little if any effort has been directed towards the development of objective image fusion performance metrics.

Given the input and single fused output images, this letter addresses the problem of measuring fusion performance objectively. Thus a performance measurement framework is defined in the next section which quantifies the fusion process and is subsequently used to compare the performance of different pixel level fusion systems. Furthermore, experimental results of this metric are presented in the *Results* section and are shown to be in agreement with preference scores obtained from informal subjective tests. This clearly indicates that the proposed fusion measure is perceptually meaningful. Concluding remarks are provided in the final section.

Fusion Measure: The goal in pixel level image fusion is to combine and preserve in a single output image all the “important” visual information that is present in a number of input images. Thus an objective fusion measure should i) extract all the perceptually important information that exists in the input images and ii) measure the ability of the fusion process to transfer as accurately as possible this information into the output image.

In this work we associate important visual information with the “edge” information that is present in each pixel of an image. Notice that this visual to edge information association is supported by Human Visual System [4] studies and is extensively used in image analysis and compression systems. Furthermore, by evaluating the amount of edge information that is transferred from input images to the fused image, a measure of fusion performance can be obtained.

Specifically, consider two input images A and B, and a resulting fused image F. Note that the following methodology can be easily applied to more than two input images. A Sobel edge operator is applied to yield the edge strength $g(n,m)$ and orientation $\alpha(n,m)$ information for each pixel $p(n,m)$, $1 \leq n \leq N$ and $1 \leq m \leq M$. Thus for an input image A:

$$g_A(n,m) = \sqrt{s_A^x(n,m)^2 + s_A^y(n,m)^2} \qquad \ldots \qquad (1)$$

$$\alpha_A(n,m) = \tan^{-1}\left(\frac{s_A^y(n,m)}{s_A^x(n,m)}\right) \qquad \ldots \qquad (2)$$

where $s_A^x(n,m)$ and $s_A^y(n,m)$ are the output of the horizontal and vertical Sobel templates centred on pixel $p_A(n,m)$ and convolved with the corresponding pixels of image A.

The relative strength and orientation values of $G^{AF}(n,m)$ and $A^{AF}(n,m)$ of an input image A with respect to F are formed as

$$G^{AF}(n,m) = \begin{cases} \dfrac{g_F(n,m)}{g_A(n,m)}, & if \quad g_A(n,m) > g_F(n,m) \\ \dfrac{g_A(n,m)}{g_F(n,m)}, & otherwise \end{cases} \qquad \ldots \qquad (3)$$

$$A^{AF}(n,m) = \frac{\left| \left| \alpha_A(n,m) - \alpha_F(n,m) \right| - \pi/2 \right|}{\pi/2} \qquad \ldots \qquad (4)$$

These are used to derive the edge strength and orientation preservation values

$$Q_g^{AF}(n,m) = \frac{\Gamma_g}{1 + e^{\kappa_g (G^{AF}(n,m) - \sigma_g)}} \qquad \ldots \qquad (5)$$

$$Q_\alpha^{AF}(n,m) = \frac{\Gamma_\alpha}{1 + e^{\kappa_\alpha (A^{AF}(n,m) - \sigma_\alpha)}} \qquad \ldots \qquad (6)$$

$Q_g^{AF}(n,m)$ and $Q_\alpha^{AF}(n,m)$ model perceptual loss of information in F, in terms of how well the strength and orientation values of a pixel $p(n,m)$ in A are represented in the fused image. The constants Γ_g, κ_g, σ_g and Γ_α, κ_α, σ_α determine the exact shape of the sigmoid functions used to form the edge strength and orientation preservation values, see equations *(5)* and *(6)*. Edge information preservation values are then defined as

$$Q^{AF}(n,m) = Q_g^{AF}(n,m)\ Q_\alpha^{AF}(n,m) \qquad \ldots \qquad (7)$$

with $0 \le Q^{AF}(n,m) \le 1$. A value of 0 corresponds to the complete loss of edge information, at location (n,m), as transferred from A into F. $Q^{AF}(n,m)=1$ indicates "fusion" from A to F with no loss of information.

Having $Q^{AF}(n,m)$ and $Q^{BF}(n,m)$ for $N*M$ size images, a normalised weighted performance metric $Q_P^{AB/F}$ of a given fusion process P that operates on images A and B, and produces F is obtained as follows:

$$Q_P^{AB/F} = \frac{\sum_{n=1}^{N} \sum_{m=1}^{M} Q^{AF}(n,m) w^A(n,m) + Q^{BF}(n,m) w^B(n,m)}{\sum_{i=1}^{N} \sum_{j=1}^{M} (w^A(i,j) + w^B(i,j))} \qquad \ldots \qquad (8)$$

Notice that the edge preservation values, $Q^{AF}(n,m)$ and $Q^{BF}(n,m)$, are weighted by $w^A(n,m)$ and $w^B(n,m)$ respectively. In general edge preservation values which correspond to pixels with high edge strength, should influence $Q_P^{AB/F}$ more than those of relatively low edge strength. Thus, $w^A(n,m)=[g_A(n,m)]^L$ and $w^B(n,m)=[g_B(n,m)]^L$ where L is a constant. Also note that $0 \le Q_P^{AB/F}(n,m) \le 1$.

Results: In order to establish the subjective relevance of the proposed methodology in assessing the performance of pixel level fusion systems, $Q_P^{AB/F}$ was estimated in relation to three fusion algorithms. The first, *Scheme I*, is a conventional multiresolution fusion system that employs an "area" type subband pixel selection approach during pyramid fusion [3]. *Scheme II* employs the same conventional Quadrature Mirror Filter (QMF) decomposition approach with an advanced cross band selection technique for pyramid fusion [1]. *Scheme III* is a computationally efficient system based on a background/foreground decomposition and fusion process.

Thus informal subjective tests were performed using pairs of input images and the corresponding fused output images produced by two different fusion algorithms. Subjects were asked to vote in favour of one of the two systems or indicate that the algorithms perform equally well. A preference for a particular fused image assigned one point to the system used to produce it, whereas half a point was given to both systems in the case of equal preference. An average Subjective Score (SS) was therefore obtained for each fusion system, using the above hard decision process, when this test was applied over a set of *K* pairs of input images.

The same hard decision preference and corresponding point allocation was also employed using the $Q_P^{AB/F}$ objective measure. That is, for a particular input pair:

$Q_I^{AB/F} > Q_{II}^{AB/F}$ *, 1 point assigned to Scheme I*

$Q_I^{AB/F} < Q_{II}^{AB/F}$ *, 1 point assigned to Scheme II*

$Q_I^{AB/F} = Q_{II}^{AB/F}$ *, ½ point assigned to both Scheme I and Scheme II*

This process yielded an average Objective Score (OS).

Figure 1 shows the SS and OS values obtained from two experiments. The first assessed schemes *I* and *II* and involved $K=10$ pairs of input images and 11 subjects, see Figure 1 (a). The SS and OS values for Schemes *I* and *III* obtained for $K=12$ and 9 subjects are shown in Figure 1 (b). The pairs of input images used in these experiments are aerial visible and infrared registered imagery [1,2]. These were selected to represent a wide range of imaging content as captured using hyperspectral sensors. Notice that the level of agreement between subjective and objective results is particularly high. At the same time, the size of the set of input pairs, *K*, used in these experiments is relatively small.

Conclusion: A novel objective pixel level image fusion assessment framework is presented in this brief paper and has been used to compare the performance of different fusion algorithms. Preliminary experiments show that the hard decision fusion performance assessment obtained using the proposed objective measure agrees remarkably well with that obtained from informal subjective tests.

Acknowledgements: The authors gratefully acknowledge the financial and technical support provided throughout the project by the BAe Military Aircraft and Aerostructures Mission Systems Group.

References

[1] V Petrović, C Xydeas, "Multiresolution image fusion using cross band feature selection", *Proc. SPIE,* Vol. 3719, 1999, pp 319-326

[2] V Petrović, C Xydeas, "Computationally Efficient Pixel-level Image Fusion", *Proc. Eurofusion99*, October 1999., pp 177-184

[3] H Li, S Munjanath, S Mitra, "Multisensor Image Fusion Using the Wavelet Transform", *Graphical Models and Image Proc.,* Vol. 57, No. 3, 1995, pp 235-245

[4] W Handee, P Wells, "The Perception of Visual Information", *Springer*, New York, 1997

Author's affiliations: Prof. C. Xydeas, Department of Communication Systems University of Lancaster, Bailrigg, Lancaster, LA1 4YR, UK and V. Petrović, Department of Electrical Engineering, University of Manchester, Oxford Road, Manchester, M13 9PL, UK, vladimir@spec16.ee.man.ac.uk

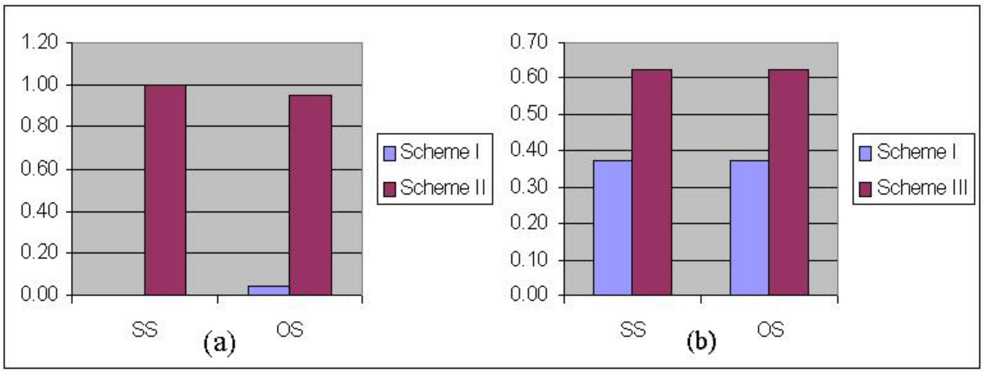

Figure 1: *Hard decision Subjective and Objective performance assessment with* $L=1.5$, $\Gamma_g=0.9994$, $\kappa_g=15$, $\sigma_g=0.5$ and $\Gamma_\alpha=0.9879$, $\kappa_\alpha=22$, $\sigma_\alpha=0.8$