

Sensor noise effects on signal-level image fusion performance

Vladimir S. Petrović^{a,b,*}, Costas S. Xydeas^a

^a Department of Communication Systems, Lancaster University, Bailrigg, Lancaster, LA1 4YR, UK

^b Visaris, Bore Kostića 5 11077 Novi Beograd, Yugoslavia

Received 30 April 2002; received in revised form 21 August 2002; accepted 17 September 2002

Abstract

The aim of this paper is twofold: (i) to define appropriate metrics which measure the effects of input sensor noise on the performance of signal-level image fusion systems and (ii) to employ these metrics in a comparative study of the robustness of typical image fusion schemes whose inputs are corrupted with noise. Thus system performance metrics for measuring both absolute and relative degradation in fused image quality are proposed when fusing noisy input modalities. A third metric, which considers fusion of noise patterns, is also developed and used to evaluate the perceptual effect of noise corrupting homogenous image regions (i.e. areas with no salient features). These metrics are employed to compare the performance of different image fusion methodologies and feature selection/information fusion strategies operating under noisy input conditions. Altogether, the performance of seventeen fusion schemes is examined and their robustness to noise considered at various input signal-to-noise ratio values for three types of sensor noise characteristics.

© 2003 Elsevier B.V. All rights reserved.

Keywords: Image fusion; Fusion performance evaluation; Sensor noise effect

1. Introduction

Significant research effort has been directed in recent years into the development of multisensor signal-level image fusion (MSL-IF) schemes. This is due to the key role MSL-IF plays in determining the performance of advanced multisensor imaging display/visual information processing systems, in applications such as avionics, medical imaging, earth observation and security/surveillance. Within these application areas, where there is a tendency to employ several sensing modalities under a wide range of operating conditions, the prospect arises of fusing input images of low visual quality. As a result, “noisy” input information associated with individual sensors may affect significantly fusion system performance. This is because input “noise” may be treated by the fusion system as valid information and transferred to the fused output image. Furthermore input image noise may affect the selection/fusion process of the MSL-IF system in a way that introduces additional unwanted artefacts and distortion into the fused image [1–6].

The performance characteristics of image fusion algorithms, operating in noise free conditions, are considered in a number of papers. Zhang and Blum [6] present a thorough investigation into several multiresolution fusion methodologies for a digital camera application. Fusion schemes are categorised according to their basic multiresolution/pyramid image representation approach and mechanisms for pyramid coefficient fusion. Pohl and van Genderen [27] provide a comprehensive review of fusion techniques as applied to the field of remote sensing. However, despite the unquestionable usefulness of such contributions almost no effort has been allocated in measuring and comparing the effects of input noise on the performance of image fusion systems. Furthermore, no independent assessment is available of image fusion algorithms specifically designed to operate robustly under noisy input conditions [3,5].

This paper presents an investigation into the effects of sensor noise on the performance of signal-level image fusion systems. Its aim is twofold:

- (i) to develop appropriate metrics which measure the detrimental effect of input sensor noise on the fusion performance of MSL-IF schemes and

* Corresponding author.

E-mail addresses: pera@absolutok.net (V.S. Petrović), c.xydeas@lancaster.ac.uk (C.S. Xydeas).

- (ii) to employ these metrics in a comparative study of the robustness of several published image fusion schemes operating under noisy input conditions.

Three different experimental scenarios for fusing images are considered, using a two-sensors suite that is formed from visible light, infrared or image enhanced (low light) cameras. In the first scenario, only one input image suffers from sensor noise. This is analogous to the practical situation when “low contrast” conditions exist for one of the sensors, e.g. a visible light sensor operating in foggy conditions with contrast boosting applied which in turn amplifies sensor noise. Here uncorrupted, “true” scene information is still available from the noise free output of the second sensor. In the second scenario, both sensors suffer from noticeable levels of sensor noise and the expected degradation in the quality of the fused image is greater. In the third noisy fusion scenario there is no “true” scene information. Instead, the fusion of noise patterns is examined in order to investigate the behaviour and performance of MSL-IF algorithms operating on homogenous image regions that are corrupted by sensor noise.

Thus novel “noisy fusion” metrics are developed and used, in the first two scenarios, to measure the effects of additive sensor noise on the performance of several signal-level image fusion algorithms operating across a range of input signal-to-noise ratio (SNR) values. These metrics evaluate: (i) the absolute degradation of information in the output fused image, with respect to true scene information and (ii) the relative degradation of information in the fused image, with respect to the noisy input image(s). The third scenario considers the effect different fusion methods have on the characteristics of fused image noise. Overall, the performance of seventeen different fusion schemes, operating under varying noisy input conditions, is examined. These schemes were specifically chosen to cover a broad range of image fusion methodologies and feature fusion (information fusion) techniques. Furthermore, two MSL-IF schemes with built-in noise suppression procedures are also investigated.

The next section refers briefly to sensor noise models and the generation of noisy input image data. In Section 3, several novel noisy fusion performance metrics are derived using the objective image fusion performance evaluation framework of Xydeas and Petrović [13,14]. Noisy fusion performance results are presented, compared and discussed in Section 4. Finally, concluding remarks are offered in Section 5.

2. Sensor noise modelling and noisy image data

Passive multisensor imaging arrays often include (i) visible light sensors that measure scene illumination in

the visible spectrum (0.45–0.7 μm) (ii) infrared sensors that measure the thermal radiance of scene objects in the infrared part of the spectrum (1.5–15 μm) and (iii) low light or image enhanced cameras [25]. In general, sensor noise is the result of several processes associated with the underlying physics of recording an observation [8–12]. Typically however, additive noise is the predominant component of noise encountered in such devices. Additive noise is modelled as a random signal that is simply added to the original signal. In the practical model used in this investigation, an input image A is corrupted to yield image A_n by (i) generating a noise-seed signal N according to particular sensor noise characteristics, (ii) scaling N by an appropriate factor k_n to produce a desired signal-to-noise ratio “ n ” and (iii) adding the scaled noise signal to A to yield $A_n = A + k_n N$. When several, say M , input images are to be corrupted, M different N_1 to N_M seeds are generated. The process of generating noisy images is also illustrated graphically in the lower part of Fig. 3 at the end of this section.

Sensor noise models that were used to create noise seed signals, relate to widely used charge coupled device (CCD) cameras [7,15]. These sensors measure the charge induced by incident photons on a mesh of electro-optical elements. The more photons hit a particular element, the larger the amount of charge stored, and the lighter the pixel (sensed by that element). Stored charges are read out electronically from the electro-optic array in a pre-defined pattern. Within this CCD sensing framework there are two significant sources of noise: (i) the dark (leakage) current noise which is sensitive to temperature and (ii) photon generated sampling noise which is dependant on the number of electrons (incident photons) and is relatively insensitive to temperature [12]. In visible and infrared CCD sensors both the dark current and photon noise components are present under normal operating conditions. In low to medium cost sensors, which form the majority of sensors employed in practice, the two components combine and at room temperatures the noise can be modelled as Gaussian [11,16]. Low-light (LL) or image intensified devices operate in a slightly different manner to ordinary visual and infrared sensors. They are required to boost the range of the input signal, which also affects photon/shot noise contained in the signal and this type of noise becomes dominant. Photon noise arises due to the effect that not all of the incident photons free an electron. As the probability of a photon generating an electron–hole pair in the photoelectric layer remains constant, the amplitude of noise generated in an array element (pixel) depends on the number of incident photons (illumination). In this case, noise variance is directly proportional to illumination (pixel value) and photon noise is modelled by a Poisson distribution [8,11].

These two models were used to generate the noise seed signals N . For visible and infrared range imagery,

N is obtained using a zero mean, unit variance, Gaussian random number generator [9,11,16]. For LL imagery, noise seeds are zero mean, Poisson distributed random variables, see Eq. (1) where U_i is a uniformly distributed random variable and λ is signal dependent and given by Eq. (2) [17]. The scale factor s determines the level of dependence of the noise variance (power) on the signal value $A(n, m)$. In our experiments $s = 18$.

$$N = \min \left\{ n : \prod_{i=1}^n U_i < e^{-\lambda} \right\} - \lambda \quad (1)$$

$$\lambda(k, l) = \frac{A(k, l)}{s} \quad (2)$$

2.1. Input data

Input image data is divided into three sets (I, II and III) each consisting of 10 pairs of multispectral images. In each set, one image in each pair belongs to a specific sensor modality, which is “visual” for set I and “infrared” and “low-light” for sets II and III, respectively. The second image in each input pair corresponds to one of the other two modalities. Table 1 gives the modality composition for the three input sets. Note that image content varied considerably including outdoor, industrial and urban area scenes.

Input images were corrupted at six different SNR levels: ∞ (noise-free), 30, 20, 10, 3 and 1 dB. Fig. 1 shows an example of a visual range image (set I) corrupted by a single Gaussian noise seed scaled to obtain

Table 1
Sensor modality composition of input data sets

	Input pair composition		
	Visual sensor	Infrared sensor	Low-light visual/IR sensor
<i>Input</i>			
Set I	10	10	0
Set II	7	10	3
Set III	2	8	10

the desired SNR. The noise-free image is shown in Fig. 1a, whereas images shown in Fig. 1b,c,d,e and f correspond to SNR values of 30, 20, 10, 3 and 1 dB respectively. Notice that initially two different noisy fusion experimental scenarios were considered. In the first, only the first image (same modality) is corrupted in each input pair. Also, the first two image sets are corrupted with Gaussian and the third with Poisson distributed noise. In the second scenario however, both input images are corrupted, with a matching SNR using the appropriate noise model and all input image pairs are used as a single input set. Overall, for every fusion method examined, 180 input image pairs (30 noise-free pairs \times 6 SNR values) were employed.

In addition, a third noisy fusion scenario was also considered that examines the fusion of noise signals in homogenous image regions and uses, as input images, noise “patterns”. Twenty pairs of such random noise patterns were created and arranged in two sets. In the first set, 40 different random Gaussian distributed

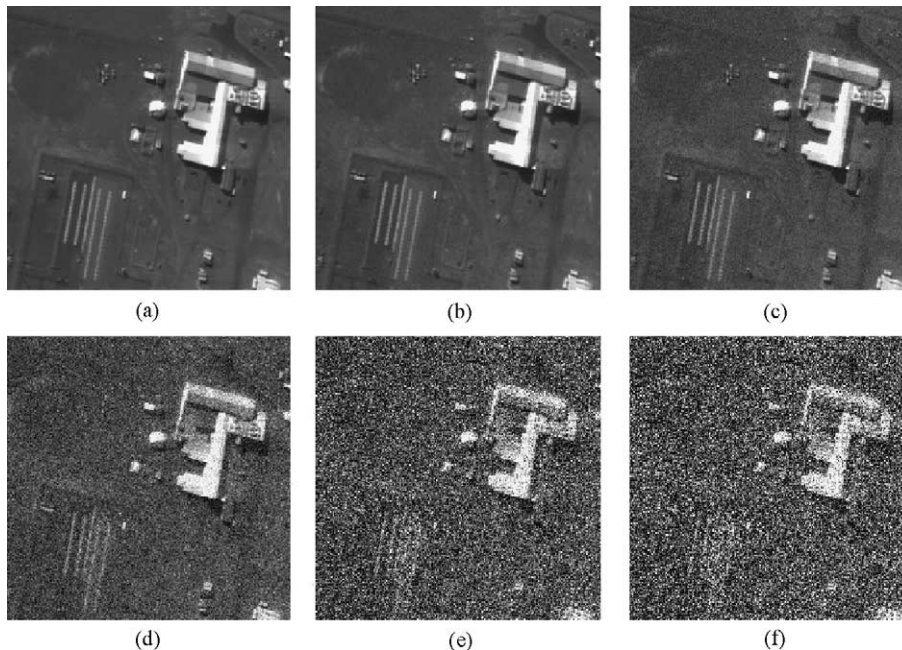


Fig. 1. Input images: (a) noise-free and (b), (c), (d), (e) and (f) with SNR values of 30, 20, 10, 3 and 1 dB, respectively.

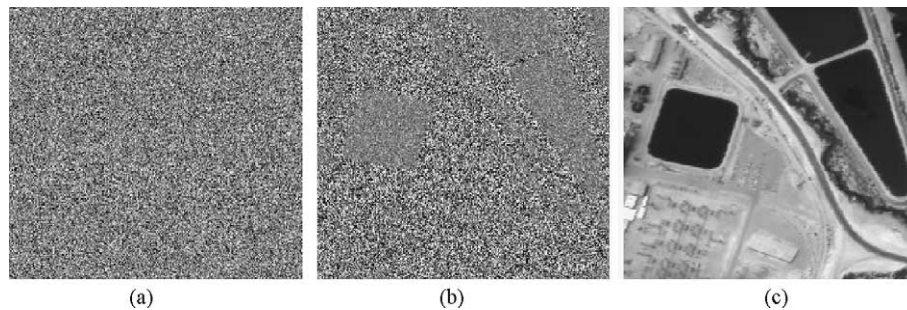


Fig. 2. Noise pattern inputs: (a) Gaussian, (b) Poisson and (c) reference input image used to produce image (b).

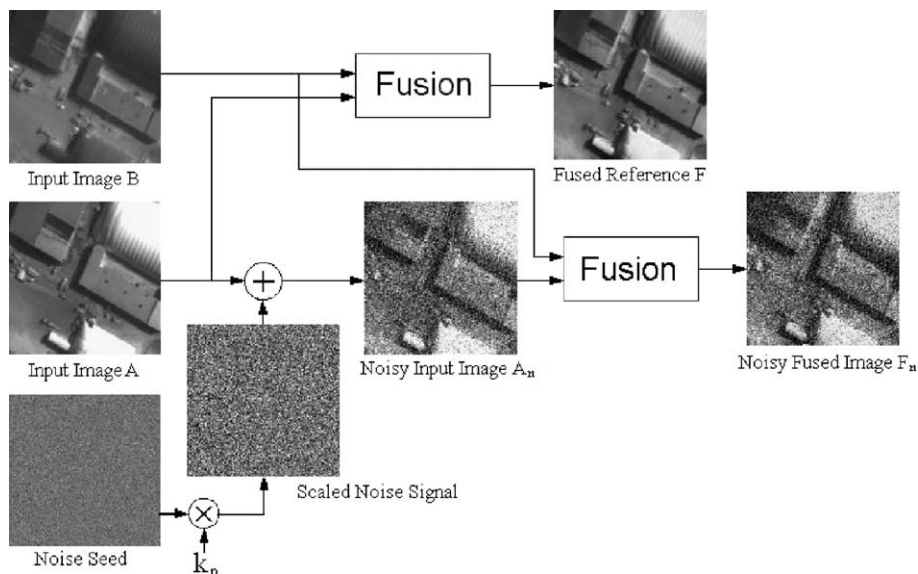


Fig. 3. Single noisy input fusion and corresponding noise-free fusion processes.

patterns, of size 512×512 pixels, produced twenty input pairs (visual–infrared sensor array images). In the second set, 20 Poisson and 20 Gaussian distributed patterns formed the input data (for the visual/infrared—low-light sensor case). In the generation of Poisson noise data, the input images from the low-light input set III were used to provide the value of λ (see Eqs. (1) and (2)). Furthermore, all input noise patterns are translated to the middle of the grey level range (128) and scaled to have a standard deviation of 20. An example of an input pair of noise patterns is shown in Fig. 2. The signal in Fig. 2b (middle) represents a Poisson distributed noise signal produced using the reference image of Fig. 2c (right). Signal dependence is noticeable in the large-scale structure of the noise signal power, which is low where the original image is dark. In contrast, Gaussian distributed noise in Fig. 2a is of “similar” power everywhere in the signal.

Notice that during experimentation in the first two scenarios, a true, “noise-less” fused reference image $F = \Phi_p(A, B)$ is also produced for each input pair, where $\Phi_p(A, B)$ represents fusion of images A and B using fusion method p . Thus in the single noisy input scenario,

fusion is performed using corrupted versions of the input image A_n , at five SNR levels i.e. $n = 30, 20, 10, 3,$ and 1dB . This produces five noisy fused images $F_n = \Phi_p(A_n, B)$. Fig. 3 shows diagrammatically these two fusion processes. The same approach is taken when both input images are corrupted. This time however, noisy fused output images are produced while fusing corrupted inputs at matching SNR levels i.e. $F_n = \Phi_p(A_n, B_n)$. Finally, in the last scenario, random noise patterns N_1 and N_2 are fused to produce a fused noise pattern $N_f = \Phi_p(N_1, N_2)$.

3. Noisy image fusion performance evaluation

Signal-level image fusion performance evaluation has only recently received attention from the research community and certain MSL-IF evaluation methods have been proposed. Noise-free fusion system performance metrics such as the root mean square error (RMSE), mutual information and the percentage of correct decisions of Zhang and Blum [6], the standard

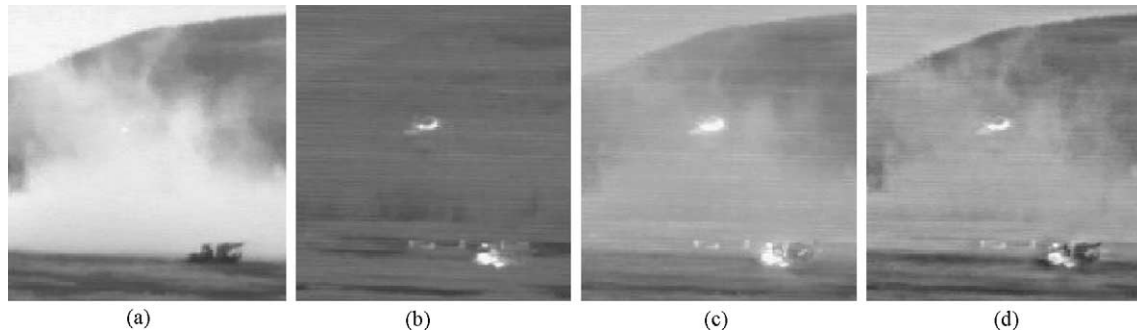


Fig. 4. Multisensor input images (a) and (b) and two fused images (c) with $Q_p^{AB/F_n} = 0.5699$ and (d) with $Q_p^{AB/F_n} = 0.6277$.

deviation of Li et al. [2] and various similar schemes outlined in Pohl and van Genderen [27], compare the output fused image with a reference image and form a distance between the two, e.g. square of the difference in the RMSE metric of Zhang and Blum [6]. These metrics, which potentially are also applicable to the case of noisy fusion, measure “relative” fusion performance (with respect to a noise free fused reference).

Both the “relative” and “absolute” image fusion performance of MSL-IF systems operating under noisy input conditions can be objectively measured using the subjectively meaningful fusion performance evaluation framework of Xydeas and Petrović [13,14] (see Appendix A). In this approach, visual information is associated with “edge” related information that is measured at each image pixel. An image fusion process that succeeds in transferring all of the visual information from any number of inputs into the fused image is said to have achieved ideal fusion. Thus, by evaluating the relative amount of “edge” information that is transferred from the inputs into the output fused image, a Q_p^{AB/F_n} measure of fusion performance is obtained.

An application example of the Q_p^{AB/F_n} fusion performance metric is given in Fig. 4 where the two multisensor input images, shown in Fig. 4a and b, are fused using the RoLP and DWT methods (see Section 4.1) to yield images 4c and 4d, respectively. Fused image 4d is a good representation of the scene with all the important features easily visible and recognisable. However, the fused image shown in Fig. 4c exhibits a general loss of contrast and also smearing of the boundaries of important objects. These image quality observations are reflected in the corresponding Q_p^{AB/F_n} performance scores.

Furthermore, statistical methods for evaluating fusion performance in the case when “homogenous” input test images are corrupted by noise are also proposed in this section.

3.1. Overall degradation of performance in noisy fusion

Evaluating absolute fusion performance requires the understanding of what constitutes ideal fusion. Recall

that the main objective of a signal-level image fusion system is to transfer, the content of input images into a fused output image as faithfully as possible. This implies a possible “loss” of information as well as the introduction of “artefacts” in the fused image. With this in mind, the $Q_p^{AB/F}$ measure [13,14], see Eq. (A.6) in Appendix A, is a fusion performance metric whose value increases toward unity as the amount of information loss and artefacts in F decreases. However, when input images are corrupted by noise, loss of input “noise information” in the output image is an advantageous characteristic of the underlying fusion process. This means that $Q_p^{A_n B_n / F_n}$ (estimated for a fusion system p operating on noisy input images, A_n and B_n to produce an output image F_n) is not subjectively meaningful since loss of “noise information” in F_n will yield a reduction in $Q_p^{A_n B_n / F_n}$.

Measuring only the representation of “true” scene information in the fused image solves the problem of taking into account correctly the loss of noise information. This true information is contained in the noise free input images, A and B , and meaningful performance assessment is achieved using the Q_p^{AB/F_n} metric. The value of Q_p^{AB/F_n} increases when the fused output image is a more accurate representation of the noise free A and B images, i.e. when there is a reduction of noise in F_n . Q_p^{AB/F_n} measures the overall success, of the fusion scheme p , in representing true scene information in the noisy fused image F_n , and can be used to determine fusion performance under noisy input conditions. Thus for decreasing SNR values “ n ”, Q_p^{AB/F_n} describes the absolute degradation of fusion performance with increasing noise. Notice that Q_p^{AB/F_n} takes into account the effects of artefacts and distortions introduced by the fusion process itself.

3.2. Relative noisy fusion performance

An interesting aspect of system performance is the “relative” degradation of information in fused images with increasing input image noise. This represents the “robustness” of the system to input noise and is

measured by evaluating the degradation in F_n with respect to F (clean fused image). Thus, Q_p^{FF/F_n} , formed as in Appendix A, measures the degradation in the fused image F_n obtained when the input(s) images have an SNR of n dB, as compared to the noise free fused output F . This time however, any distortions introduced by the fusion algorithm are not taken into consideration since they appear in both F and F_n . Robust fusion methods are undisturbed by noise, and information in F_n closely resembles that in F . Q_p^{FF/F_n} for such systems remains relatively high, whereas for noise sensitive methods Q_p^{FF/F_n} decreases rapidly with decreasing SNR values.

Another metric of “relative” noisy fusion performance is the change in degradation found in the fused image with respect to that of input images. This metric considers the fact that robust fusion algorithms may suppress noise effects and, as a result, fused images may degrade less than input images for corresponding levels of noise. Conversely, fused images obtained via noise sensitive algorithms either degrade in the same manner as input images or contain amplified noise. By comparing the degradation of information in input images corrupted by noise, with that of fused images, a relative performance (information) gain achieved by the fusion process can be measured. Again, using the objective measure framework of Appendix A, a measure of relative fusion “gain” can be defined when one or two inputs are corrupted by noise. In particular, in the single noisy input case, Q_p^{AA/A_n} measures the loss of visual information as a result of corrupting the noise-free image A with n dB noise to obtain A_n . Also, Q_p^{FF/F_n} measures the loss of information at the output of the fusion system, due to the input being corrupted by noise. The relative effect of noise on the visual information observed in the input and output fused images can then be defined as the relative difference between the degradation of the fused image Q_p^{FF/F_n} and the degradation of the input image Q_p^{AA/A_n} . For the single noisy input case then, the relative noisy fusion gain D_n^p , is defined as:

$$D_n^p = \frac{Q_p^{FF/F_n} - Q_p^{AA/A_n}}{Q_p^{AA/A_n}} \quad (3)$$

Note that most fusion methods treat input noise patterns as valid information and fuse them directly into F_n . This transfer of noise information is not desirable and for a decreasing SNR, a robust fusion method preserves the true information in the fused image and the rate of decrease of Q_p^{FF/F_n} is slower than that of Q_p^{AA/A_n} resulting in a positive value of D_n^p . In general, D_n^p is negative when input noise is “amplified” by the fusion process, i.e. further loss of information in F_n is caused by the fusion process. Conversely, a positive value indicates that the fusion system p possesses noise suppression characteristics. Note also that when

SNR = ∞ dB, $A_n = A$, $F_n = F$, i.e. there is no degradation in the input(s) or fused images, Q_p^{AA/A_n} and Q_p^{FF/F_n} are unity, and $D_n^p = 0$.

When both input images suffer from sensor noise, degradation in the second input image, measured through Q_p^{BB/B_n} , is also taken into account. The relative performance gain of the fused image is evaluated as the “cumulative” gain of Q_p^{FF/F_n} over the degradation of the two inputs, i.e.:

$$D_n^p = \frac{2Q_p^{FF/F_n} - Q_p^{AA/A_n} - Q_p^{BB/B_n}}{Q_p^{AA/A_n} + Q_p^{BB/B_n}} \quad (4)$$

D_n^p values have the same meaning as in the previous case, and high values are desirable. In general, when compared to the single noisy input case (Eq. (3)), lower D_n^p values are expected. Furthermore, high D_n^p values in this case also identify fusion algorithms with improved information selection/fusion and verification properties.

3.3. Noise pattern fusion performance measures

The Q_p^{AB/F_n} , Q_p^{FF/F_n} and D_n^p metrics described above, concentrate on measuring the “preservation” of significant (edge) information from the uncorrupted input images into the fused image. Under noisy input conditions however, homogenous regions with no salient features become important since distracting noise patterns are easily perceived in them due to the lack of visual masking effects. Noise in these areas, which may cover a large part of the fused image, has a distracting effect on the observer and reduces perceived fused image quality. When both input images are corrupted by noise, the properties of the noise patterns in the fused image depend on the fusion algorithm. Ideally, robust image fusion algorithms should discern between real scene information and noise, and suppress the latter. In reality, most fusion algorithms treat noise as valid scene information and transfer it into the fused output image.

The effect of different fusion approaches on output noise introduced in homogenous regions is evaluated here while fusing signals containing only random noise patterns and examining the statistics of the resulting fused signal. Relevant statistics in this case, are the variance, the mean value and the probability density function (PDF) of the fused output noise signal. Fused noise power, represented by the variance, relates directly to the “detectability” of the noise pattern in the fused image. Relative change in the noise power, produced by a fusion scheme p is measured as the noise power gain S_p , defined in Eq. (5), where σ_{0p}^2 and σ_{0i}^2 are the fused and input noise variances respectively. Fusion algorithms, which amplify input noise, have $S_p > 1$, while noise suppression in fused image pro-

duces $S_p < 1$. In our experiments, input noise power was set to 400 (σ_{oi} of 20) for both inputs and only σ_{0p}^2 required measurement.

$$\frac{S_p = \sigma_{0p}^2}{\sigma_{oi}^2} \quad (5)$$

Another parameter that characterises a homogenous region in an image is the local mean μ . When both input images have the same mean, (for example, $\mu_1 = \mu_2 = 128$ is used in our experiments) the fused signal should exhibit the same value for μ_f . Any deviation from this value indicates a failure by the fusion algorithm. Finally, the third indicator of the effect fusion algorithms have on output noise is the PDF of the fused noise signal. Input noise signals exhibit either Gaussian or Poisson distributions and any change in the shape of the output PDF indicates a change in the appearance of the noise signal. PDF shapes closely concentrated around zero amplitude indicate better noisy fusion performance since such noise is harder to detect and has less effect on true information. Conversely, more noise pixels of larger amplitudes indicate annoying speckle noise in the fused signal.

4. MSL-IF system performance results

The noisy fusion performance metrics defined in the previous section were employed in a comparative study of seventeen different fusion schemes operating on the noisy input image sets described in Section 2. The overall Q_p^{AB/F_n} , Q_p^{FF/F_n} and D_n^p scores for each input set were obtained as average values calculated over all input image pairs in that set. Thus for example, Q_p^{AB/F_n} is estimated as

$$Q_p^{AB/F_n} = \frac{1}{M} \sum_{i=1}^M Q_p^{A_i B_i / F_{i,n}} \quad (6)$$

where i indicates the i th input image pair and M is the total number of pairs in an input set.

Section 4.1, examines the noisy fusion performance of seven basic fusion methodologies. The effect which different feature selection/information fusion strategies have on noisy fusion performance is considered in Section 4.2. Finally, the performance of two “robust” fusion systems with built-in noise suppression characteristics is discussed in Section 4.3. Results are given for the three previously mentioned fusion scenarios, i.e. single noisy input, both noisy inputs and noisy pattern fusion. Notice however that presenting the complete set of simulation results is unrealistic, due to paper size restrictions, and only the most important and representative results are shown and discussed.

4.1. Fusion system performance in the presence of input noise

Initially, seven MSL-IF methodologies were examined in the context of noisy input conditions. These broadly represent the spectrum of image fusion approaches available today. The simplest of these techniques is arithmetic fusion, where each pixel in the fused image is obtained as an arithmetic combination of corresponding input pixels. Although computationally efficient, this approach generally achieves modest performance. Arithmetic fusion is described by

$$F(n, m) = k_A A(n, m) + k_B B(n, m) + C \quad (7)$$

and in this study image averaging Scheme 1 uses $[k_A, k_B, C] = [1/2, 1/2, 0]$. Arithmetic fusion is also the basis of Scheme 2. However this time principal component analysis (PCA) is used to determine the optimal weighting coefficients (k_A and k_B) according to the eigenvalues and eigenvectors of the covariance matrix formed between the two input images [25].

The simple bi-scale fusion method of Scheme 3 [18] decomposes input images into two levels of scale using adaptive averaging templates. Small-scale details are fused using the pixel-based “select max” approach of Eq. (8) where D_A , D_B and D_F represent the two inputs and the fused small-scale signals, respectively and large-scale information is fused using arithmetic fusion.

$$D_F(n, m) = \left\{ \begin{array}{ll} D_A(n, m), & \text{if } |D_A(n, m)| \geq |D_B(n, m)| \\ D_B(n, m), & \text{otherwise} \end{array} \right\} \quad (8)$$

The “select max” coefficient approach is also used as the feature selection/pyramid fusion (FS/PF) mechanism in the next four schemes which employ different multiresolution image representations. In these representations, also known as image pyramids, image signals are decomposed into several lower resolution sub-bands that contain varying size information at different resolutions. This enables the fusion process to operate effectively at different scale ranges [1–6, 19–22]. The selection process of Eq. (8), with “ D ” now representing the values of the input and fused pyramid coefficients, compares the absolute value of corresponding input coefficients and transfers the most significant input coefficients into the fused pyramid.

Schemes 4–6 are based on derivatives of the Gaussian low-pass pyramid [1, 19, 20]. The ratio of low pass (RoLP) pyramid [19] in Scheme 4, is formed by evaluating the ratio between each Gaussian pyramid level and an expanded (interpolated) version of the next, coarser level. Values in the RoLP pyramid directly represent local luminance contrast. The alternative, Laplacian pyramid representation, used in Scheme 5 is formed by *subtracting* from each Gaussian pyramid level the expanded version of the next level [20]. In Scheme 6 a

Gradient pyramid representation is used, which is obtained by filtering each level of the Laplacian pyramid with four gradient filters (horizontal, vertical and two diagonals). This process generates for each input image four pyramids according to orientation [1]. This introduces *orientation sensitivity* to the fusion process whereby information of different orientations is fused independently.

The final multiresolution fusion Scheme 7 uses the wavelet (DWT) pyramid representation [21]. This approach decomposes an input image, using a series of 1D filter banks, into a composite pyramid having three sub-bands images of different orientation at each resolution level. The absolute values of the DWT pyramid coefficients reflect local area (rather than direct pixel-level) image saliency at corresponding scales. As with all other multiresolution methodologies, the fused output image is obtained from the fused pyramid by applying the appropriate multiresolution reconstruction process [1,19–21].

Fusion performance results are shown in Fig. 5 when the above schemes operate with only one noisy input image. Fig. 5a displays absolute performance Q_p^{AB/F_n} values for the corrupted visible light input data Set I (results obtained using the other two sets of input images are very similar). All schemes exhibit decreasing

absolute performance with decreasing SNR values, a behaviour that levels off below 5 dB. Notice that image averaging, (Scheme 1 *), performs poorly at high SNR values, but outperforms the other schemes for $SNR \leq 20$ dB. This is due to destructive superposition, which reduces true scene contrast at high SNR but also reduces noise power at low SNR where true scene information is already severely corrupted. Above 10 dB, the Laplacian pyramid fusion (Scheme 5 ∇), performs slightly better than the other pyramid schemes indicating relative insensitivity to small, random changes in the input signal. PCA fusion (Scheme 2 +) performs worst.

Relative fusion performance degradation Q_p^{FF/F_n} results are shown (for set I) in Fig. 5b. Image averaging (Scheme 1 *) is least affected by the increase in input noise, while PCA (2 +) fusion is the most sensitive. Below 10 dB, the Q_p^{FF/F_n} values of the other schemes are similar.

Noisy fusion gain D_n^p results are shown (for input sets I and II) in Fig. 5c and d. Positive values indicate a tendency of the fusion process to minimise the noise corruption effect which input noise can have on the output fused image. Only PCA fusion (Scheme 2 +) exhibits $D_n^p < 0$ values which indicate noise amplification. This is due to the assignment of “unbounded” weighting factors (>1) in the arithmetic fusion equation

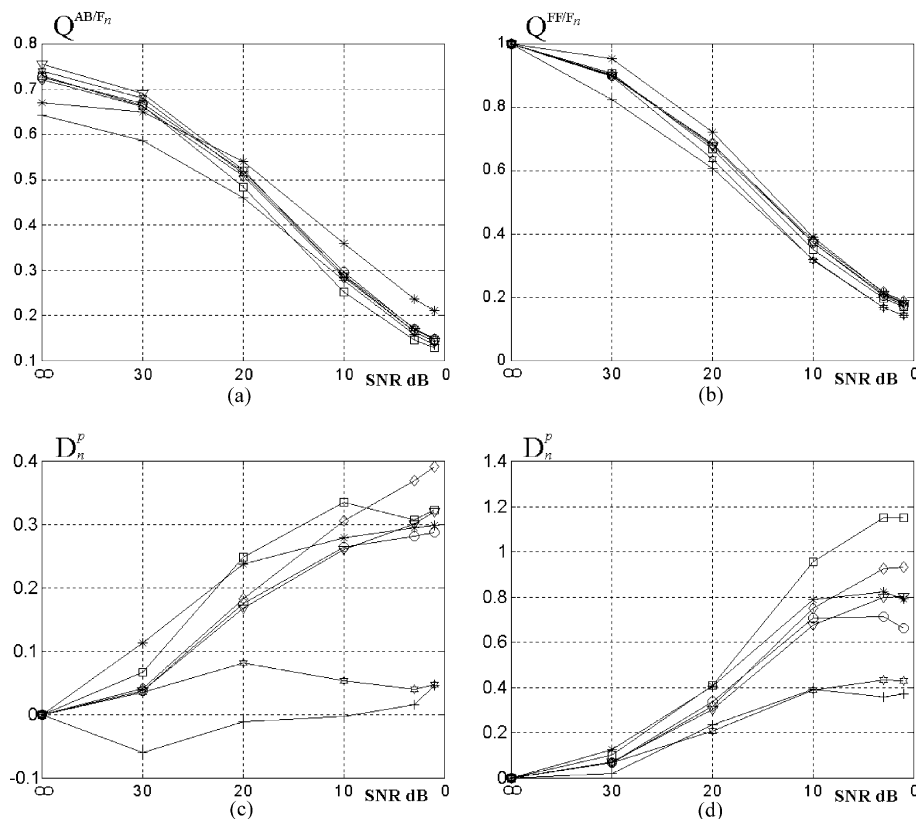


Fig. 5. Noisy fusion performance against input SNR: (a) Q_p^{AB/F_n} and (b) Q_p^{FF/F_n} for set I, (c) and (d) D_n^p for sets I and II; (*) image averaging Scheme 1, (+) PCA Scheme 2, (O) bi-level Scheme 3, (□) RoLP Scheme 4, (∇) Laplacian Scheme 5, (\otimes) gradient Scheme 6 and (\diamond) Scheme DWT 7.

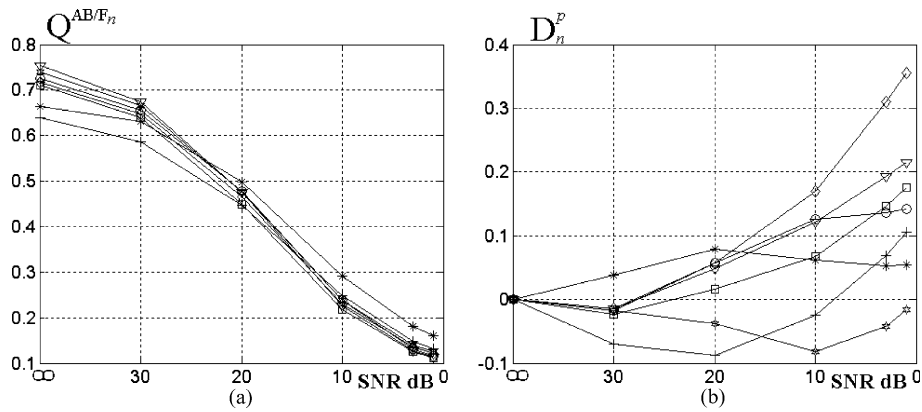


Fig. 6. Fusion of two noisy inputs (a) Q_p^{AB/F_n} and (b) D_n^p ; line markings are the same as in Fig. 5.

(3) to the modality with the greater local variance (e.g. the noisy input in this case). Such factors effectively produce an amplified noise signal in the fused image. In general, D_n^p scores for all schemes are close to zero at high SNR (≥ 30 dB) values and increase as SNR decreases. At SNR < 20 dB, Q_p^{FF/F_n} and Q_p^{AA/A_n} values decrease and, for a given difference between these two quantities, the lower Q_p^{AA/A_n} is, the larger the D_n^p ratio becomes. The relative difference between D_n^p scores obtained from sets I and II (see Fig. 5c and d) can be attributed to the fusion process selecting more information from the visible light sensor input images of Set I. At high noise levels, much of small-scale detail information in the visible light range sensor images is lost, and D_n^p values are relatively low.

In addition to the measurements illustrated in Fig. 5, the same image data was subjected to informal subjective evaluation. Subjects were able to identify correctly the performance differences indicated by the Q_p^{AB/F_n} scores (see Fig. 5a) of the Averaging (Scheme 1 *) and PCA (Scheme 2 +) systems. Furthermore, the small spread of Q_p^{AB/F_n} values obtained from the remaining systems Schemes 3–7) was also reflected by the inability of subjects to firmly discriminate between the performances of these systems.

Performance results obtained when both input images were noisy are shown in Fig. 6 (all input sets combined) with Q_p^{AB/F_n} scores in Fig. 6a and D_n^p values in Fig. 6b. The trend of decreasing Q_p^{AB/F_n} with increasing noise power, which levels off below 5 dB, is very similar to that obtained previously in the case of a single noisy input image. This indicates that the Q_p^{AB/F_n} performance of all schemes is affected in the same manner when sensor noise is present in one or both inputs. The effect of the second noise signal is evident in the reduced scale of D_n^p —compare Figs. 5d and 6b. Notice that fusion schemes, which rely on feature selection mechanisms for fusion, i.e. Schemes 3–7, have low D_n^p scores for high SNR. The reason is confusion in the selection process caused by input noise, which appears as significant

information and influences selection decisions. Also at low SNR values, the better spectral decomposition properties of the DWT analysis process confines noise more effectively within frequency sub-bands, as compared to the other schemes, and thus the DWT-based fusion system (Scheme 7 \diamond) performs best. In terms of D_n^p performance and at high SNR values (> 20 dB), image averaging, (Scheme 1 *) outperforms all other schemes.

An interesting result shown in Fig. 6b is the sensitivity to noise of the gradient pyramid fusion (Scheme 6 \star). The rather relaxed, orientation sensitive, spectral image decomposition used in this scheme, coupled with a direct representation in the pyramid coefficients of “local” pixel image contrast [1], means that input noise patterns produce in all orientation pyramids significant values which in turn upset the operation of the FS/PF mechanism. Although orientation sensitivity is also a feature of the DWT-based fusion system (Scheme 7 \diamond), DWT pyramid coefficient values represent “local area” pixel image contrast, which effectively limits the influence of input noise on pyramid values. This provides an improved D_n^p performance at low SNR values (see Fig. 6b).

A graphic example of image fusion with two noise-corrupted inputs is shown in Fig. 7. Input A corrupted at the SNR values of 30 and 1 dB is given in Fig. 1b and f (see page 8) respectively, while input image B is shown in Fig. 7a and d at corresponding SNR values. Fused output images obtained using image averaging (Scheme 1) are shown in Fig. 7b and e, while fused images produced by the RoLP-based approach (Scheme 4) are shown in Fig. 7c and f, respectively. At 30 dB, the effect of noise on the input and fused images is hardly noticeable and RoLP fusion (Scheme 4 Fig. 7c) performs better ($Q_4^{AB/F_{30}} = 0.636$, $Q_1^{AB/F_{30}} = 0.628$). However, at the extreme case of 1 dB SNR, the definition of the main objects in the scene is better in the fused image produced by averaging. Again this is consistent with the performance scores in Fig. 6a ($Q_4^{AB/F_1} = 0.105$, $Q_1^{AB/F_1} = 0.151$).

The performance of these seven schemes operating on random noise patterns was also evaluated using

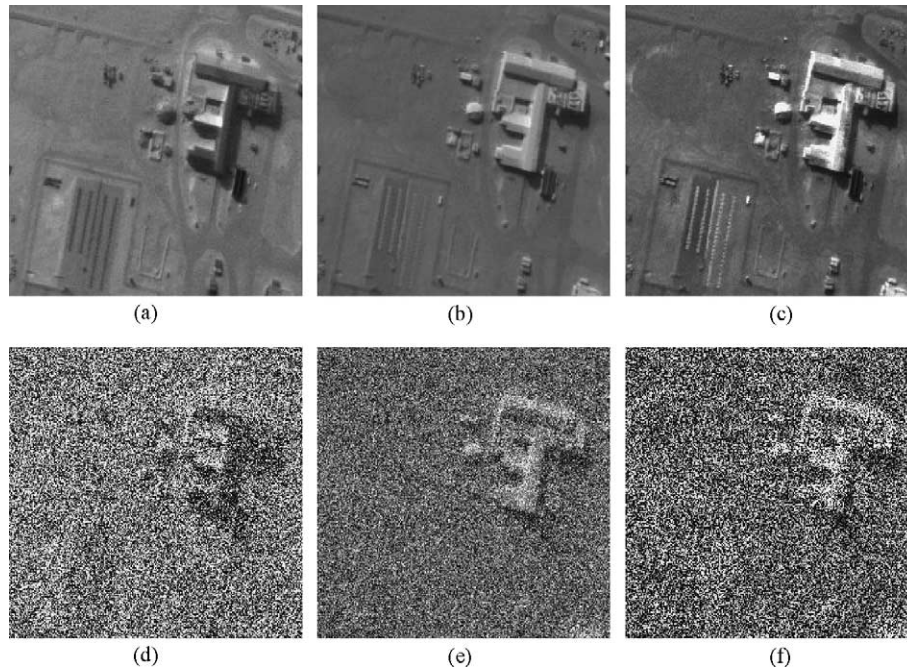


Fig. 7. Noisy image fusion: input image B at (a) 30 dB and (d) 1 dB SNR. Fused images, averaging (Scheme 1) (b) and (e) and RoLP (Scheme 4) (c) and (f) for corresponding SNR values.

the S_p noise power gain metric. S_p is calculated as an average over 20 pairs of input noise signals and results are given in Table 2 for both Gaussian–Gaussian (G–G) and Poisson–Gaussian (P–G) inputs. Image averaging (Scheme 1) is the only approach that achieves noise power reduction with $S_1 = 0.5$, all other schemes increase noise power. As expected from previous performance results, the PCA-based fusion system yields the largest noise amplification with $S_2 = 2.85$ and 4.45 for G–G and P–G inputs, respectively. The difference in the two scores is due to the structure of the Poisson noise which leads into very large weighting coefficients in the arithmetic fusion process. The noise power gain of other methods is similar, $S_p \approx 1.64$, with the exception of gradient pyramid fusion (Scheme 6) which performs better, with $S_6 = 1.35$ and 1.38. The increase in output noise power for these schemes can be attributed to the feature selection process, which chooses and transfers into the fused

image the most salient input noise. In contrast, the averaging of two zero mean noise signals used in Scheme 1 reduces output power through destructive superposition. These differences in the power of fused noise can be seen in the appearance of fused output noise signals in images (e) and (f) of Fig. 7, obtained from the Averaging and RoLP-based fusion schemes, respectively.

In terms of the “mean” performance indicator (see Section 3.3) and with both input noise signals having the same mean value, the mean μ_F of all the fused output noise patterns is within 0.5%, of the input mean values ($\mu_1 = \mu_2 = 128$). However when μ_1 and μ_2 assume different values, the overall mean of the fused output noise signal is close to the arithmetic mean of μ_1 and μ_2 for all systems except PCA (Scheme 2) and RoLP (Scheme 4). The mean μ_F of PCA fused noise signals varies considerably from input pair to input pair ($\sigma_0 > 40$, $E[\mu_F] \approx (\mu_1 + \mu_2)$) whereas RoLP-based fusion produces

Table 2
Noise power gain results

	Noise type	Fusion method							
		Input, σ_{0i}^2	Scheme 1, averaging	Scheme 2, PCA	Scheme 3, multi-scale	Scheme 4, RoLP	Scheme 5, Laplacian	Scheme 6, Gradient	Scheme 7, Wavelet
G–G	σ_{0p}^2	400	200	1141	664	649	648	540	655
	S_p	–	0.50	2.85	1.66	1.62	1.62	1.35	1.64
P–G	σ_{0p}^2	400	200	1781	671	654	654	553	660
	S_p	–	0.50	4.45	1.68	1.64	1.63	1.38	1.65

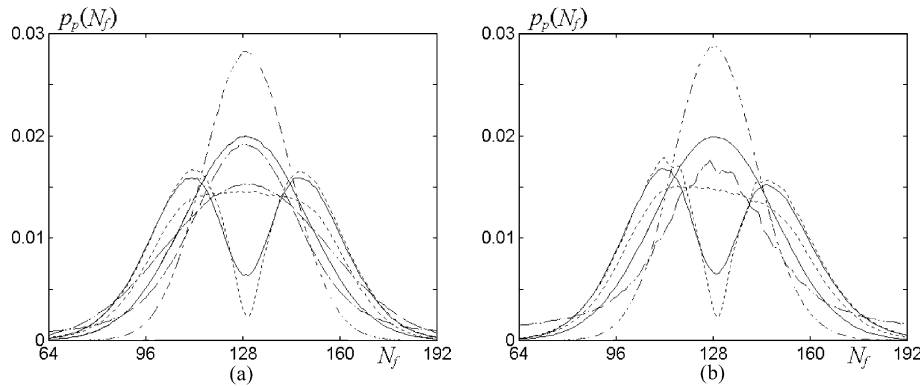


Fig. 8. PDFs for (a) Gaussian–Gaussian, and (b) Poisson–Gaussian inputs.

mean values that are closer to the larger of the two input means i.e. $\mu_F \rightarrow \mu_{\max}$, $\mu_{\max} = \max(\mu_A, \mu_B)$.

Output fused noise is characterised by its PDF. Fig. 8a and b shows the PDFs, $p_p(N_f)$ of the fused noise signals produced by Schemes 1–7, using Gaussian–Gaussian and Poisson–Gaussian inputs respectively. From these plots it is evident that no two schemes have the same effect on the nature of the noise signal (the reference input PDF is the middle solid Gaussian line). Noise fused by image averaging, PCA and DWT (Scheme 1, 2 and 7) maintain the Gaussian shape, but with different variances. The PDF of Scheme 1 is narrow and tall (dashed line) which indicates low intensity, less noticeable noise. The PDF of PCA Scheme 2 (dot-dashed Gaussian) remains considerably high even at the extremes of the range shown in Fig. 8, which indicates the presence of many high amplitude noise pixels. The maximisation of *local area* contrast observed in DWT fusion (Scheme 7, bottom dot-dashed Gaussian) dilutes noise hot-spots and the result is a less speckly (less noticeable) noise signal. Bi-level fusion, RoLP and Laplacian pyramid fusion (3,4 and 5) exhibit interesting “two-hump” distributions that suggest a speckled appearance for the output fused noise (dotted line Scheme 3, solid line Schemes 4 and 5); compare Fig. 7f (RoLP) and e (averaging). This is due to *pixel* contrast maximisation tendency, which underlines these pyramid fusion methodologies. Finally, the dotted flattop single-hump shape belongs to Gradient pyramid fusion (Scheme 6) and suggests an evenly distributed noise signal.

4.2. FS/PF strategy and noisy fusion performance

The effect that different feature selection/pyramid fusion strategies have on noisy-input fusion performance was examined experimentally, using systems which are based on the DWT and Laplacian pyramid image representations. Thus six further fusion schemes were examined, which employ the DWT representation

and differ in their basic *selection/fusion* approach (i.e. coefficient selection or arithmetic combination) as well as in the size of the selection template and in the selection criterion. Pixel-based “select-max” (see Eq. (8)) DWT-based Scheme 7 was used as a *selection* type fusion reference system, having an effective template (window) size of 1×1 . Scheme 8 employs the same “select-max” coefficient selection, using however a 3×3 size template on the coefficients of the input pyramids. An even larger 5×5 template is used in the same manner in Scheme 9. Here however the above basic coefficient selection is also followed by a consistency verification process that takes the form of “majority filtering” of the selection map. This additional step is aimed at reducing the undesirable effects on the selection process of input noise “hot-spots” [2]. The area-based fusion mechanism proposed in Ref. [1] is used in Scheme 10 where a measure of local saliency and a measure of similarity is calculated between the inputs at every pyramid pixel, using a 3×3 window. When low similarity is observed between corresponding windows in the input pyramids, the pixel with the larger saliency found within the windows is copied into the fused pyramid. Otherwise, the fused value is produced as the weighted sum of inputs with the weights evaluated from the saliency and match measures [1]. The hybrid arithmetic/selective fusion process used in the cross-band algorithm of Ref. [22], is employed in Scheme 11. This method integrates information from several pyramid levels and sub-bands, in a process also known as single- and multi-scale grouping [6], in order to select/fuse input pyramid values using averaging, addition or direct coefficient selection [22].

Arithmetic pyramid fusion (see Eq. (7)) is employed in Schemes 12 and 13. In Scheme 12, the fused pyramid is the direct sum of the two input pyramids $[k_A, k_B, C] = [1, 1, 0]$. Scheme 13 uses a weighted sum (see Eq. (9)) which assigns proportionally larger weights to more salient features that produce larger pyramid values. A review and methodical evaluation of the performance

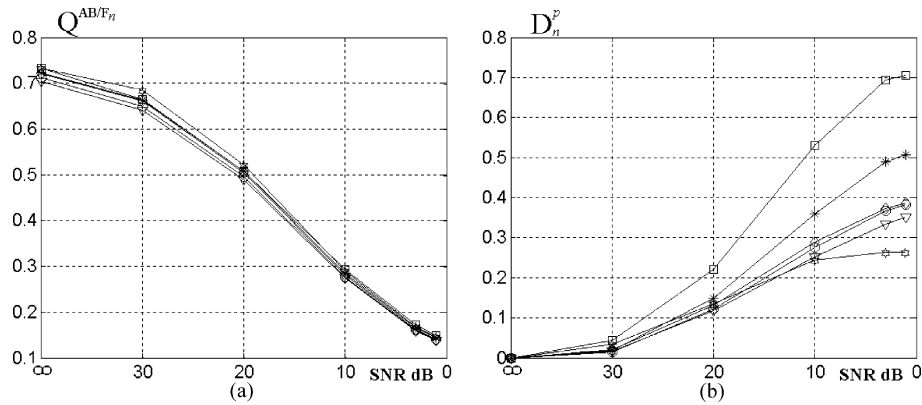


Fig. 9. FS/PF performance with a single noisy input: (a) Q_p^{AB/F_n} for input set I, and (b) D_n^p for input data set III; Schemes: 7 (*), 8 (O), 9 (∇), 11 (\diamond), 12 (\square), 13 (\star).

of a number of feature selection/pyramid fusion strategies, when operating under noise free input conditions, can be found in [6].

$$D_l^F(n, m) = \frac{|D_l^A(n, m)|D_l^A(n, m) + |D_l^B(n, m)|D_l^B(n, m)}{|D_l^A(n, m) + D_l^B(n, m)|} \quad (9)$$

Furthermore, in order to examine the relative importance of FS/PF mechanisms, when used in association with other pyramidal image representations, two further Laplacian pyramid fusion-based Schemes 14 and 15 were implemented and tested. These systems use the 3×3 window, area-based selection/fusion method of Scheme 8, and the Burt and Kolczynski [1] combined selection/arithmetic fusion of Scheme 10, respectively.

Q_p^{AB/F_n} and D_n^p results obtained from the DWT pyramid-based systems operating in a single noisy-input scenario are shown in Fig. 9. The true scene information preservation Q_p^{AB/F_n} values, shown for these systems in Fig. 9a, differ little at high SNR values and even less at lower SNRs. However, D_n^p results exhibit more variation (see Fig. 9b) where direct sum arithmetic fusion (Scheme 12 \square) consistently performs best and the weighted sum (Scheme 13 \star) performance is the worst. This is especially so for low input SNR values where more weight is assigned to noise than true scene information. Amongst the *selection* fusion methods, pixel-based (Scheme 7 *) performs the best and the 5×5 area selection (Scheme 9 ∇) produces the worst results. D_n^p values obtained from the 33 area selection-based Schemes 8 O and 10 (not shown), are in-between those of Schemes 7 and 9. Results obtained from the B&K [1] 3×3 area-based system (Scheme 10) are not shown as they are almost identical to those of Scheme 8 which employs a simple 3×3 area selection process.

Fig. 9 suggests that when noise power in one input image increases, the relative performance of multiresolution pyramid fusion schemes is less affected as the size of the selection template is reduced. This behaviour can be attributed to an increased number of erroneous co-

efficient selections in favour of noise “hot-spots”, when larger selection templates are used.

As an illustration of the relevance of the D_n^p results shown in Fig. 9, examples of fused images obtained using the pixel-based “select-max” (Schemes 7 *) and the 5×5 area selection (Scheme 9 ∇) are shown in Fig. 10. When the input images are noise free, the image on the left, 10a (Schemes 7), is only slightly better than image 10b (Schemes 9) which has less pronounced contrast. However, when a significant amount of noise is added to one of the inputs (SNR = 10 dB), the output fused images are of noticeably different quality. The pixel-based “select max” fused image of Scheme 7, (shown in Fig. 10c) although corrupted, still contains all significant features in an easily detectable form. The 5×5 area selection-based Scheme 9 output image

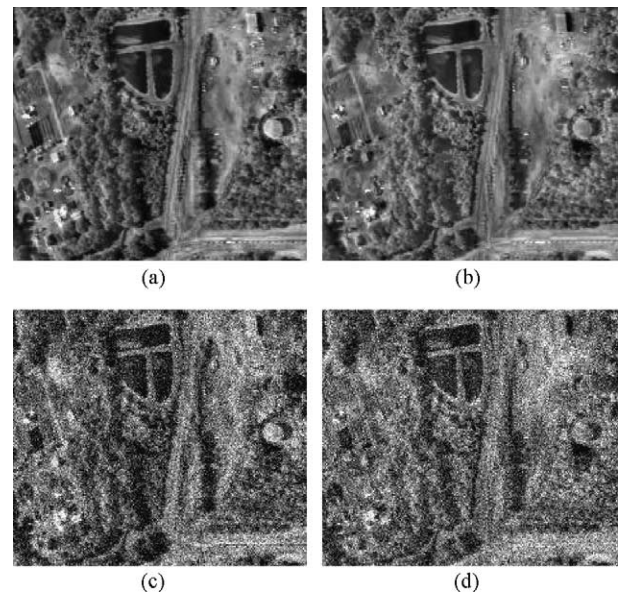


Fig. 10. Pixel-based select-max fusion (Scheme 7) at ∞ dB (a) and 10 dB (c); 5×5 area selection fusion (Scheme 9) at ∞ dB (b) and 10 dB (d).

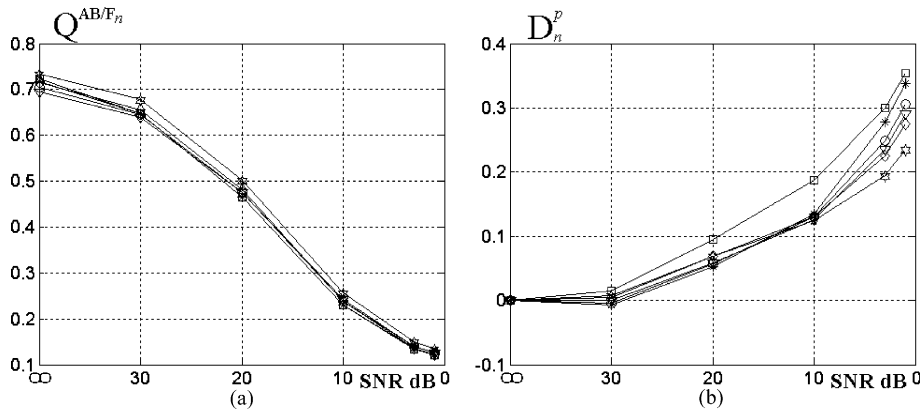


Fig. 11. System performance when both inputs are noisy: (a) Q_p^{AB/F_n} and (b) D_n^p using input data set I, marking is the same as in Fig. 9.

however, shown in Fig. 10d, exhibits serious degradations in object boundaries and a significant loss of small-scale details. The corresponding D_n^p scores of $D_{10\text{ dB}}^7 = 0.209$ and $D_{10\text{ dB}}^9 = 0.113$ reflect correctly this subjective difference in image fusion quality.

Similar system fusion performance results are obtained when both inputs are corrupted (see Fig. 11). In terms of absolute fusion performance Q_p^{AB/F_n} , Scheme 13 (\star) has a more noticeable advantage over the other schemes. Also, in terms of D_n^p (Fig. 11b), Scheme 12 (\square) still performs better than the other systems, albeit with a smaller advantage.

A final comparison between these systems, in terms of noise pattern fusion performance, is given in Table 3. Arithmetic fusion methods perform as their arithmetic fusion coefficients $[k_A, k_B, C]$ would suggest, that is the direct sum doubles output noise power ($S_{12} = 2$), and the weighted sum preserves the input power level ($S_{13} \cong 1$). For selective fusion methods, a strong correlation exists between the size of the selection template and the power of the fused noise signal. The larger the selection window is, the smaller is the output noise power. Thus, the 5×5 area selection (Scheme 9) shows a small increase in noise power with $S_{12} = 1.11$ and 1.18 . In contrast, pixel-based (1×1) “select max” (Scheme 7) exhibits a considerably higher noise power gain with $S_7 = 1.64$ and 1.65 . These results appear to be in contradiction with the noisy fusion system performance conclusions derived from Figs. 9 and 11, with respect to selection template

sizes. However, as stated earlier, the Q_p^{AB/F_n} , Q_p^{FF/F_n} and D_n^p metrics capture the subjective effect of noise on “coherent” edge related true scene visual information. S_p , on the other hand, is only associated with noise power values (in “uniform” image regions) which are subjectively less meaningful since the human visual system filters out much of the noise [23] (people see through noise), particularly in homogeneous image regions. The overall message is therefore that, in the presence of input noise, smaller selection template sizes preserve better “true” scene information.

The above fusion system performance findings were also observed using the Laplacian pyramid fusion representation and FS/PF schemes. Thus fusion performance results obtained from Schemes 5, 14 and 15 were in agreement with those obtained from the corresponding DWT-based systems.

4.3. Noise suppression and signal-level image fusion

In this section, two specific fusion systems with built-in noise suppression mechanisms are considered. Both schemes are based on the DWT representation using the pixel-based “select max” fusion approach and include thresholding of lower (high-resolution) pyramid levels. Scheme 16 uses soft thresholding [5], whereby all sub-band signal values smaller than a threshold are set to zero, while the remaining values are reduced by the threshold value. Scheme 17 uses hard thresholding [3]

Table 3
Noise power gain results of different FS/PF schemes based on the DWT representation

	Noise type	Fusion scheme							
		Input σ_{0i}^2	Scheme 7, 1×1	Scheme 8, 3×3	Scheme 9, 5×5 [2]	Scheme 10, 3×3 B&K [1]	Scheme 11, \times -band	Scheme 12, direct sum	Scheme 13, weighted sum
G-G	σ_{0p}^2	400	655	492	546	445	507	801	400
	S_p	–	1.64	1.23	1.37	1.11	1.27	2.00	1.00
P-G	σ_{0p}^2	400	660	511	558	472	524	799	408
	S_p	–	1.65	1.28	1.40	1.18	1.31	2.00	1.02

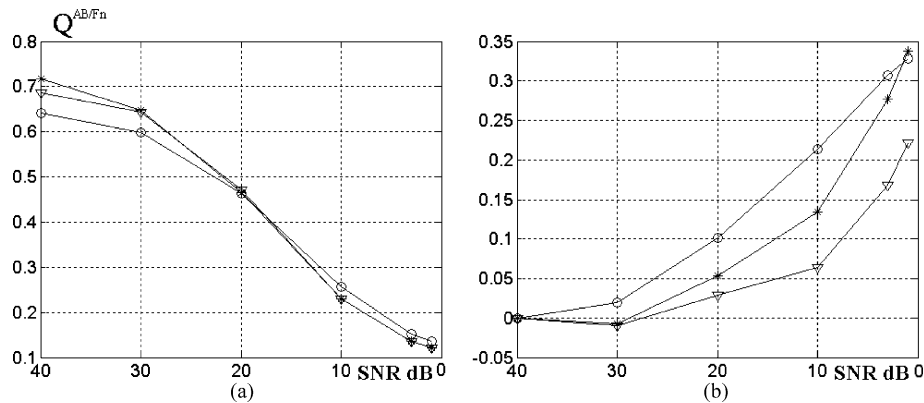


Fig. 12. Noise suppression: Q_p^{AB/F_n} (a) and D_n^p (b); Schemes 7 (*), 16 (O) and 17 (∇).

where only values smaller than the threshold are set to zero. In both cases, threshold values can be fixed or adaptive. Both schemes are compared with the basic DWT pixel-based “select-max” Scheme 7 and any performance advantage they display over this system can be attributed directly to the noise suppression mechanism.

In our investigations, different values/settings for fixed and adaptive thresholds and numbers of resolution levels to which thresholding is applied, were examined. In both cases, the best performance was obtained with an adaptive threshold T_{sb} , evaluated individually for each sub-band signal x_{sb} , according to:

$$T_{sb} = k * \sqrt{E[x_{sb}^2]} \quad (10)$$

The optimum value of k was experimentally determined as $k = 0.75$ and system performance was maximised when thresholding only the two lowest pyramid levels. In general, thresholding of higher pyramid levels degrades fusion performance at high SNR values, without a significant improvement at low SNR values. Furthermore, no significant advantage in system performance is obtained when thresholding is only applied to the lowest pyramid level. A similar behaviour is observed with k values. Values $k < 0.75$ provide no performance gains, whereas for higher values system performance deteriorates at high SNR values but improves modestly at low SNR values.

Fig. 12 shows Q_p^{AB/F_n} and D_n^p results for these three schemes. In general, noise suppression should (i) reduce absolute fusion performance Q_p^{AB/F_n} at high input SNR values and (ii) improve Q_p^{AB/F_n} at the lower end of the SNR scale. Soft thresholding (Scheme 16 O) performs in this manner showing gains at low SNR values, when compared to the basic pixel-based selection fusion (Scheme 7 *). System performance is significantly reduced at high SNRs, when compared to that of Scheme 7 because of a general reduction in contrast introduced into the input images. In terms of D_n^p performance, the advantage of using soft thresholding (Scheme 16 O) is again obvious (see Fig. 12b).

The noise pattern fusion performance of these schemes is given in Table 4. Soft thresholding clearly shows better noise suppression characteristics when compared to the other two schemes. The system exhibits almost no increase in overall fused noise power, ($S_{16} = 1.04$) due to its design that reduces the value of all pyramid coefficients, not just those whose value is smaller than the threshold. Hard thresholding (17) on the other hand, achieves only a modest improvement over Scheme 7 ($S_{17} = 1.56$ and 1.57). Note that the probability density functions of the fused noise signals produced by Schemes 16 and 17 are of Gaussian shape.

A graphic example of fusion performance with built-in noise suppression is demonstrated in Fig. 13. Noise-free input images, shown in Fig. 4(a) and (b), are

Table 4
Noise fusion gain, S_p , of Schemes 7, 16 and 17

	Noise type	Fusion scheme			
		Input, σ_{oi}^2	Scheme 7, selection	Scheme 16, soft thresh.	Scheme 17, hard thresh.
G-G	σ_0^2	400	655	416	625
	S_p	–	1.64	1.04	1.56
P-G	σ_0^2	400	660	417	630
	S_p	–	1.65	1.04	1.57

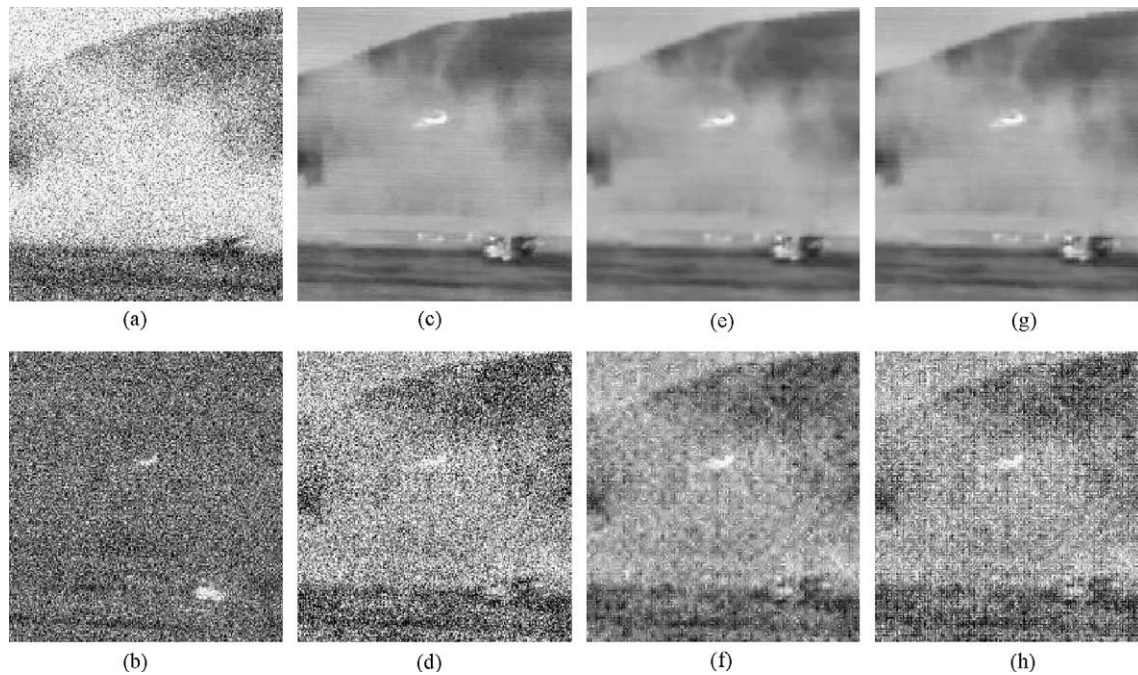


Fig. 13. Fusion noise removal, input images at 10 dB, (a) and (b); DWT direct fusion (Scheme 7) at ∞ and 10 dB, (c) and (d); soft thresholding (Scheme 16), (e) and (f); hard thresholding (Scheme 17), (g) and (h) at matching SNR values.

corrupted at an SNR value of 10 dB (see Fig. 13a and b) and fused using Schemes 7, 16 and 17. Direct DWT fusion, Scheme 7 produced the noise free fused image shown in Fig. 13c), which scores $Q_7^{AB/F} = 0.643$ with the objective measure. Schemes 16 and 17 operating on noise free inputs produce images that are less sharp (see Fig. 13e and g) and so score less, i.e. $Q_{16}^{AB/F_\infty} = 0.579$ and $Q_{17}^{AB/F_\infty} = 0.587$, respectively. When significant noise corruption is present at the input images, direct DWT fusion performance (Schemes 7) suffers considerably (see Fig. 13d) where the main objects in the input images are hardly visible. This is reflected in the low absolute fusion performance score of $Q_7^{AB/F_{10}} = 0.105$. Noise suppression using soft thresholding (Schemes 16) however provides a much clearer scene representation (see Fig. 13f) and a significantly higher $Q_{16}^{AB/F_{10}} = 0.213$ score. Hard thresholding provides some advantage (see Fig. 13h) and achieves a $Q_{17}^{AB/F_{10}} = 0.150$ score.

5. Summary

The performance of several signal-level image fusion systems was examined, for the case when their input signals are corrupted at varying levels of sensor noise. This was done using novel noisy fusion performance metrics developed via the subjectively meaningful, objective image fusion performance evaluation framework recently established by Xydeas and Petrović [13,14]. The proposed metrics measure absolute fusion

performance, relative degradation of fusion performance (robustness) and fusion gain and were used to evaluate a wide range of image fusion methodologies and different feature selection/fusion techniques including two systems with built-in noise suppression techniques.

It was found that multiresolution fusion systems generally preserve or in some cases even increase noise power in fused signals. Arithmetic fusion meanwhile trades-off absolute fusion performance at high input SNR values for improved performance at significant levels of input noise (SNR < 20 dB). Furthermore, in terms of relative fusion performance robustness, DWT fused images degrade least for increasing levels of input noise (especially towards low SNR levels). At this juncture, it was also noted that fusion performance is more significantly affected when it is the visible range input that suffers from sensor noise (compared to infrared or low light imagery).

In terms of the feature selection/pyramid fusion strategy used in multiresolution fusion, it was shown that under noisy input conditions arithmetic pyramid selection/fusion and in particular the “weighted sum method”, provides the best absolute fusion performance even though such arithmetic FS/PF methods exhibit performance limitations when input images are free of noise. When using selection-based fusion, performance is found to depend strongly on the size of the selection template. Small template sizes were found to better preserve true scene information from significantly corrupted inputs while larger templates produce a more

pleasing (less powerful) noise in homogenous regions of the fused image. Fusion performance robustness with respect to noise was also found to increase as the size of the selection template decreases.

Finally, the detrimental effect on the output fused image of high levels of input sensor noise can be reduced via the “soft thresholding” of pyramid coefficients. This is achieved at the cost of reduced fusion performance at high input SNR values. No such performance improvements are obtained from systems employing “hard thresholding” of pyramid coefficients.

Acknowledgements

The authors gratefully acknowledge BAE Systems for their financial support during the early stages of this project and the Defence Research Establishment Valcartier, Canada and the US government’s AMPS programme for some of the imagery used in this research.

Appendix A. Objective evaluation of signal-level image fusion performance

Performance metrics used to evaluate the noisy fused images in this investigation were derived from the subjectively meaningful, image fusion performance evaluation framework proposed by Xydeas and Petrović [13,14]. In this framework, important visual information is associated with edge information measured for each image pixel. An image fusion process that transfers all of the important visual (edge) information from any number of inputs into the fused image is said to have achieved ideal fusion. Correspondingly, by evaluating the relative amount of edge information that is transferred from input images to the output, fused image, a measure of fusion performance is obtained. General structure of this objective fusion performance evaluation framework is shown in Fig. 14.

Specifically, assuming two input images A and B and a resulting fused image F , a Sobel edge operator is applied to yield the strength $g(n, m)$ and orientation $\alpha(n, m)$ information [26] for each pixel $p(n, m)$, $1 \leq n \leq N$ and $1 \leq m \leq M$. Using this, the relative strength and orientation “change” values, $G^{AF}(n, m)$ and $A^{AF}(n, m)$, of an input image A with respect to F are formed as:

$$G^{AF}(n, m) = \left\{ \begin{array}{ll} \frac{g_F(n, m)}{g_A(n, m)}, & \text{if } g_A(n, m) > g_F(n, m) \\ \frac{g_A(n, m)}{g_F(n, m)}, & \text{otherwise} \end{array} \right\} \quad (\text{A.1})$$

$$A^{AF}(n, m) = \frac{||\alpha_A(n, m) - \alpha_F(n, m)|| - \pi/2|}{\pi/2} \quad (\text{A.2})$$

Strength and orientation “change” measures are then used to derive the edge strength and orientation preservation values, $Q_g^{AF}(n, m)$ and $Q_\alpha^{AF}(n, m)$:

$$Q_g^{AF}(n, m) = \frac{\Gamma_g}{1 + e^{\kappa_g(G^{AF}(n, m) - \sigma_g)}} \quad (\text{A.3})$$

$$Q_\alpha^{AF}(n, m) = \frac{\Gamma_\alpha}{1 + e^{\kappa_\alpha(A^{AF}(n, m) - \sigma_\alpha)}} \quad (\text{A.4})$$

These quantities model the perceptual loss of information in the fused image, in terms of how well the edge strength and orientation values of a pixel $A(n, m)$ are represented at $F(n, m)$. The constants Γ_g , κ_g , σ_g and Γ_α , κ_α , σ_α determine the exact shape of the sigmoid nonlinearities used to form the edge strength and orientation preservation values from which the overall edge information preservation values are then defined:

$$Q^{AF}(n, m) = Q_g^{AF}(n, m)Q_\alpha^{AF}(n, m) \quad (\text{A.5})$$

$0 \leq Q^{AF}(n, m) \leq 1$, where 0 corresponds to the complete loss of edge information, at location $F(n, m)$, and $Q^{AF}(n, m) = 1$ indicates ideal “fusion” from A to F with no loss of information. Having $Q^{AF}(n, m)$ and $Q^{BF}(n, m)$, a normalised weighted performance metric $Q_p^{AB/F}$ of a given process p that fuses A and B into F is given as:

$$Q_p^{AB/F} = \frac{\sum_{n=1}^N \sum_{m=1}^M Q^{AF}(n, m)w_A(n, m) + Q^{BF}(n, m)w_B(n, m)}{\sum_{i=1}^N \sum_{j=1}^M (w_A(i, j) + w_B(i, j))} \quad (\text{A.6})$$

The overall objective image fusion performance measure given by Eq. (A.6), also takes into account the relative perceptual importance of the visual information found in the input images. This allows for the fact that not all edges in the input images influence the overall perceived fusion performance in the same manner. Edge preservation values, $Q^{AF}(n, m)$ and $Q^{BF}(n, m)$, are weighted by coefficients, $w_A(n, m)$ and $w_B(n, m)$, which reflect the perceptual importance of the corresponding edge elements within the input images. In this way more salient edges that would be assigned more importance by an

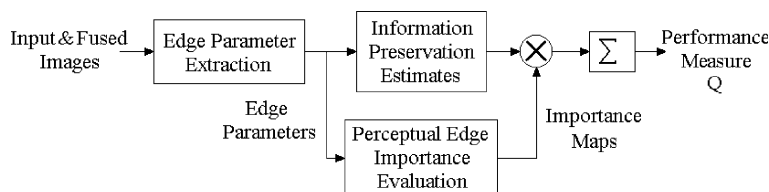


Fig. 14. General structure of objective image fusion performance evaluation.

observer influence the fusion performance result more than the less salient or undetectable ones. The perceptual importance weights, $w_A(n, m)$ and $w_B(n, m)$, are evaluated using a number of parameters associated with the edge at pixel (n, m) according to:

$$w_A(n, m) = d_A(n, m)C_A(n, m)P_A(n, m)g_A(n, m)^L \quad (\text{A.7})$$

In Eq. (9), $d_A(n, m)$ is the binary detectability flag, which is 0 if the edgel at $A(n, m)$ is below the detection threshold of the observer and 1 otherwise. $C_A(n, m)$ is the edge orientation coherence parameter, which is close to 1 if the edgel is in a neighbourhood of edgels of similar orientation, and is closer to 0 as edgels in the neighbourhood become more random. $P_A(n, m)$ introduces the positional importance, which, depending on the fusion application, might be 1 for the pixels in the centre of the input image and decreases linearly as the distance increases. The parameters mentioned so far can be viewed as “modulation” of the last term in Eq. (A.7) which is dependant on the edge strength $g_A(n, m)$ at pixel $A(n, m)$, where L is a constant. More detailed descriptions of the mechanism for assigning perceptual importance to visual information and of the objective performance measure framework, as well as details on its subjective validation can be found in Xydeas and Petrović [13,14]. Notice finally, that the resulting range of the fusion measure is also $0 \leq Q_p^{A/B/F}(n, m) \leq 1$, where 0 indicates total loss of information from the input images and 1 ideal fusion.

References

- [1] P. Burt, R. Kolczynski, Enhanced image capture through fusion, Proceedings of the Fourth International Conference on Computer Vision, Berlin 1993, pp. 173–182.
- [2] H. Li, B. Munjanath, S. Mitra, Multisensor image fusion using the wavelet transform, Graphical Models and Image Processing 57 (3) (1995) 235–245.
- [3] L. Chipman, T. Orr, L. Graham, Wavelets and image fusion, Proceedings of SPIE 2569 (1995) 208–219.
- [4] L. Ramac, M. Uner, P. Varshney, M. Alford, D. Ferris, Morphological filters and wavelet based image fusion for concealed weapons detection, Proceedings of SPIE 3376 (1998) 110–119.
- [5] L. Grewe, R. Brooks, Atmospheric attenuation reduction through multi-sensor fusion, Proceedings of SPIE 3376 (1998) 102–109.
- [6] Z. Zhang, R. Blum, A categorization of multiscale-decomposition-based image fusion schemes with a performance study for a digital camera application, Proceedings of the IEEE 87 (8) (1999) 1315–1326.
- [7] G. Waldman, J. Wootton, Electro-optical systems performance modelling, Artech House, Norwood, 1993.
- [8] A. Frenkel, M. Sartor, M. Wlodawski, Photon-noise-limited operation of intensified CCD cameras, Applied Optics 36 (22) (1997) 5288–5297.
- [9] G. Healey, R. Kondepudy, Radiometric CCD camera calibration and noise estimation, IEEE Transactions of PAMI 16 (3) (1994) 267–276.
- [10] M. Nelson, J. Johnson, T. Lomheim, General noise processes in hybrid infrared focal plane arrays, Optical Engineering 39 (11) (1991) 1682–1700.
- [11] R. Boie, I. Cox, An analysis of camera noise, IEEE Transactions of PAMI 14 (6) (1992) 671–674.
- [12] A. Prabala, CCD vs. CMOS imagers for Sci/industrial cameras, Advanced Imaging (January) (2000) 16–42.
- [13] C. Xydeas, V. Petrović, Objective image fusion performance measure, Electronic Letters 36 (4) (2000) 308–309.
- [14] C. Xydeas, V. Petrović, Objective pixel-level image fusion performance measure, Proceedings of SPIE 4051 (2000) 89–99.
- [15] J. Beynon, D. Lamb, Charge-Coupled Devices and their Applications, McGraw-Hill, London, 1980.
- [16] A. Fairhurst, A. Lettington, Method of predicting the probability of human observers recognising targets in simulated thermal images, Optical Engineering 37 (3) (1998) 744–751.
- [17] S. Ross, A First Course in Probability, MacMillan, New York, 1994.
- [18] V. Petrović, C. Xydeas, Computationally Efficient Pixel-level Image Fusion, Proceedings of Eurofusion99, Stratford-upon-Avon, October 1999, pp. 177–184.
- [19] A. Toet, Hierarchical image fusion, Machine Vision and Applications 3 (1990) 3–11.
- [20] A. Akerman, Pyramid techniques for multisensor fusion, Proceedings of SPIE 1828 (1992) 124–131.
- [21] S. Mallat, A theory for multiresolution signal decomposition: the wavelet representation, IEEE Transactions of PAMI 11 (7) (1989) 674–693.
- [22] V. Petrović, C. Xydeas, Multiresolution image fusion using cross band feature selection, Proceedings of SPIE 3719 (1999) 319–326.
- [23] W. Handee, P. Wells, The Perception of Visual Information, Springer, New York, 1997.
- [24] R. Sharma, M. Pavel, T. Leen, Multi-stream video fusion using local principal components analysis, Proceedings of SPIE 3436 (1998) 717–725.
- [25] M. Sonka, V. Hlavac, R. Boyle, Image Processing, Analysis and Machine Vision, PWS Publishing, Boston, 1999.
- [26] C. Pohl, J. van Genderen, Multisensor image fusion in remote sensing: concepts, methods and applications, International Journal of Remote Sensing 19 (5) (1998) 823–854.