Objective Image Fusion Performance Characterisation.

Conference Paper *in* Proceedings / IEEE International Conference on Computer Vision. IEEE International Conference on Computer Vision · January 2005

DOI: 10.1109/ICCV.2005.175 · Source: DBLP

2 authors:

Vladimir Petrovic
The University of Manchester

C. Xydeas
Lancaster University

Objective Image Fusion Performance Characterisation

Vladimir Petrović [1], Costas Xydeas [2]
[1] *Imaging Science and Biomedical Engineering, University of Manchester, UK*
[2] *Dept. of Communication Systems, University of Lancaster, UK*
v.petrovic@manchester.ac.uk

Abstract

Image fusion as a way of combining multiple image signals into a single fused image has in recent years been extensively researched for a variety of multisensor applications. Choosing an optimal fusion approach for each application from the plethora of algorithms available however, remains a largely open issue. A small number of metrics proposed so far provide only a rough, numerical estimate of fusion performance with limited understanding of the relative merits of different fusion schemes. This paper proposes a method for comprehensive, objective, image fusion performance characterisation using a fusion evaluation framework based on gradient information representation. The method provides an in-depth analysis of fusion performance by quantifying: information contributions by each sensor, fusion gain, fusion information loss and fusion artifacts (artificial information created). It is demonstrated on the evaluation of an extensive dataset of multisensor images fused with a wide range of established image fusion algorithms. The results demonstrate and quantify a number of well known issues concerning the performance of these schemes and provide a useful insight into a number of more subtle yet important fusion performance effects not immediately accessible to an observer.

1. Introduction

Multisensor image fusion has become an active filed of research as more and more applications such as medical imaging, security, avionics, surveillance and night vision utilise multisensor imaging arrays. Such arrays provide a wider spectral coverage and reliable information even in adverse environmental conditions at a price of a considerable increase in the amount of data. Image fusion deals with the data overload by combining visual information from multiple image signals into a single fused image with the direct aim of preserving the full content value of the multisensor information.

So far, a plethora of image fusion algorithms have been proposed [1-7] that exhibit a variety of multisensor information preservation properties. Additionally, considerable data reduction involved in the fusion process often means some algorithms introduce specific artifacts, false information into the fused image [2,4,5]. A reliable method for choosing an optimal fusion algorithm for each particular application however, largely remains an open issue. Only a handful of blind, fusion evaluation metrics [8-11] that are of interest as they require no prior ground truth (such as an ideally fused image, available only in special cases [3,4]) have been proposed so far. The gradient representation metric of Xydeas and Petrović [8] is based on the idea of measuring localised preservation of input gradient information in the fused image. Another aproach is to use mutual information, accepted as reliable when dealing with multisensor information, between the inputs and the fused image, Qu *et. al.* [9]. Such a global approach however is susceptible to noise and a more localised approach based on comparing similarity between the inputs and the fused image through local statistics was proposed by Piella and Hejmans [10]. Finally, the concept of visible differences evaluating fusion performance as the total area of visible differences between the inputs and the fused image was also considered in [11].

All existing metrics provide only a rough, numerical estimate of fusion performance usually in terms of the global similarity between the inputs and the fused image [9-11]. As a result their scores provide only limited understanding on the relative merits of different fusion schemes in each application. Generally, fusion is a complex process of information transfer and information representation and in some applications certain aspects of this process are more critical and may require special consideration. For example, a loss of a vital feature during the process of medical image fusion may lead to a wrong diagnosis and more serious consequences. In military avionics, a fusion artifact introduced into the fused image by the fusion process could lead to a benign object being classified as a threat or a valid target. In order to consider such application specific aspects of fusion performance in a robust manner, a much more comprehensive fusion evaluation is required.

This paper proposes a method for comprehensive, objective image fusion performance characterisation based on quantification of significant information components involved in image fusion: information contributions by each sensor, fusion gain, fusion information loss and fusion artifacts (artificial information created during fusion). The proposed approach is based on a powerful gradient based fusion evaluation framework which is described in section 2. Fusion performance characterisation is introduced and described in section 3. This is then used to evaluate a wide range of established image fusion algorithms on an extensive dataset of multisensor images and provide a useful insight into a number of subtle yet important aspects of their performance in section 4. The work is summarised and concluded in section 5.

2. Gradient based fusion performance

For the fusion characterisation proposed in this paper the gradient based, objective fusion evaluation proposed in [8] is used. The $Q^{AB/F}$ framework [8] associates important visual information with gradient information and assesses fusion by evaluating the success of gradient information transfer from the inputs to the fused image. Fusion algorithms that transfer more input gradient information into the fused image more accurately are said to perform better. 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** and orientation **α** ($\in[0,\pi]$) information for each input and fused image pixel. Using these parameters, relative strength and orientation "change" factors G and A, between each input and the fused image, are derived, e.g.:

$$G^{AF}_{n,m} = \begin{cases} \dfrac{g^F_{n,m}}{g^A_{n,m}}, & if \quad g^A_{n,m} > g^F_{n,m} \\ \dfrac{g^A_{n,m}}{g^F_{n,m}}, & otherwise \end{cases} \qquad (1)$$

$$A^{AF}_{n,m} = 2\pi^{-1}\left|\left|\alpha^A_{n,m} - \alpha^F_{n,m}\right| - \pi/2\right| \qquad (2)$$

These factors are the basis of the edge information preservation measure Q^{AF} obtained by sigmoidal mapping of strength and orientation change factors. This quantity models the perceptual loss of input information in the fused image and constants Γ, κ_g, σ_g, κ_α, σ_α determine the exact shape of the sigmoid mappings:

$$Q^{AF}_{n,m} = \Gamma(1+e^{\kappa_g(G^{AF}_{n,m}-\sigma_g)})^{-1}(1+e^{\kappa_\alpha(A^{AF}_{n,m}-\sigma_\alpha)})^{-1} \qquad (3)$$

Total fusion performance $Q^{AB/F}$ is evaluated as a weighted sum of edge information preservation values for both input images Q^{AF} and Q^{BF} where the weights factors w^A and w^B represent perceptual importance of each input image pixel. The range is $0 = Q^{AB/F} = 1$, where 0 means complete loss of input information has occurred and $Q^{AB/F}=1$ indicates "ideal fusion" with no loss of input information. In their simplest form, the perceptual weights w^A and w^B take the values of the corresponding gradient strength parameters g_A and g_B.

$$Q^{AB/F} = \frac{\sum_{\forall n,m} Q^{AF}_{n,m} w^A_{n,m} + Q^{AF}_{n,m} w^B_{n,m}}{\sum_{\forall n,m} w^A_{n,m} + w^B_{n,m}} \qquad (4)$$

3. Fusion performance characterisation

In its base context, the process of information fusion can be seen as an information transfer problem in which two or more information sets are combined into a new one that should contain all the information from the original sets. This is illustrated graphically using a simple Venn diagram in Figure 1. In the case of image signal fusion information is linked to a spatial location that can each be considered an individual information set. During the process of fusion, input images A and B are combined into a new fused image F by transferring, ideally all of their information into F. The information successfully transferred, shown in Figure 1 as the shaded intersection between A and B on one side and F, is precisely what the gradient fusion evaluation framework $Q^{AB/F}$ measures, relative to the total size of A∪B.

In practice however, not all of the available input information is necessarily transferred into the fused image and some loss of information from both inputs occurs. At the same time the fusion process itself occasionally creates additional, false information, known as *fusion artifacts* in the fused image. In order to provide a comprehensive characterisation of the information fusion process it is imperative to quantify these effects as well as provide a breakdown of the contributions in the successfully fused information provided by each sensor as well as information that is common and exclusive to either of them.

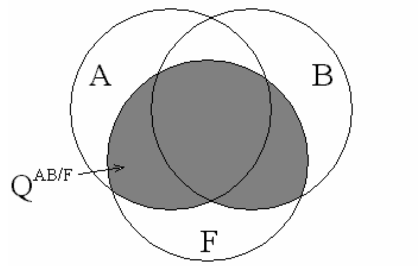

Figure 1: Graphical representation of the image information fusion process

3.1. Individual sensor contribution

Contributions in information towards the fused image made by each individual sensor (input image) in the fusion process are explicitly evaluated in the $Q^{AB/F}$ framework described in Section 2. Gradient information preservation estimates Q^{AF} and Q^{BF} given by equation (3), represent the amount of information transferred from A and B respectively into F at each location (n,m). Total information contributions of each input towards F are then evaluated according to equation (5) (for either A or B) as the sum of all local contributions modulated by the measure of local perceptual importance (w^A and w^B [8]). The contributions of individual inputs into the fusion process are also illustrated on Figure 2.

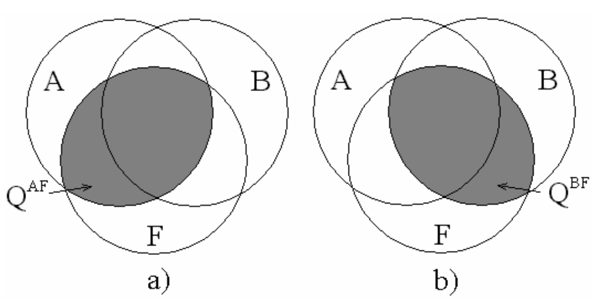

Figure 2: Information contribution from A into F , Q^{AF} – a), and B into F Q^{BF} – b)

$$Q^{\{A,B\}F} = \frac{\sum_{\forall n,m} Q^{\{A,B\}F}_{n,m} w^{\{A,B\}}_{n,m}}{\sum_{\forall n.m} w^A_{n,m} + w^B_{n,m}} \qquad (5)$$

3.2. Fusion gain: exclusive information in F

Contributions of individual inputs in the fusion process Q^{AF} and Q^{BF} consist of both the information that is exclusive to each input and information that is common across them. The former category is of particular importance as it represents *fusion gain*, or the advantage of applying fusion, as this information would not be available to an observer having only individual inputs. This quantity is illustrated on Figure 3a and consists of two distinctive contributions of exclusive information from each of the inputs. The evaluation of fusion gain from the adopted framework is not straightforward and requires a full breakdown of the contributions made by the input images.

Initially, local exclusive information in F, Q^{Δ} is found as the absolute difference of the gradient information preservation coefficients Q^{AF} and Q^{BF}, equation (6). This quantifies the total amount of local exclusive information across the fused image. For locations with strong correlation between the inputs Q^{Δ} will be small or zero, indicating no exclusive information. Conversely, in areas where one of the inputs provides a meaningful feature that is not present in the other this quantity will tend towards 1.

$$Q^{\Delta}_{n,m} = \left| \; Q^{AF}_{n,m} \; - \; Q^{BF}_{n,m} \; \right| \qquad (6)$$

The common information component for all locations across the fused image Q^C is then evaluated as a difference between the total useful information in F, given as the sum of Q^{AF} and Q^{BF}, and total exclusive information given by Q^{Δ}, equation (7). Additionally, a factor of ½ is introduced as common information is contained in both Q^{AF} and Q^{BF}, see Figure 2.

$$Q^{C}_{n,m} = \frac{Q^{AF}_{n,m} + Q^{BF}_{n,m} - Q^{\Delta}_{n,m}}{2} \qquad (7)$$

Local estimates of exclusive information components of each input, see Figure 3a, can now be derived using the input contributions Q^{AF} and Q^{BF} and the estimate of their common information in F, Q^C. Consequently, $Q^{\Delta A/F}$ is the proportion of useful information fused in F that exists only in A, while $Q^{\Delta B/F}$ is the equivalent exclusive information from B, equations (8,9).

$$Q^{\Delta A/F}_{n,m} = Q^{AF}_{n,m} - Q^{C}_{n,m} \qquad (8)$$

$$Q^{\Delta B/F}_{n,m} = Q^{BF}_{n,m} - Q^{C}_{n,m} \qquad (9)$$

These quantities represent effectively, local fusion gain achieved by fusing A and B with respect to each individual {A, B}. This can be generalised into fusion gain estimates with respect to each input for the whole fusion process by a perceptually weighted [8] integration of the local estimates across the fused image, equation (10). Furthermore, $Q^{\Delta A/F}$ and $Q^{\Delta B/F}$ make up the total exclusive information fused into F, or the *total fusion gain* $Q_\Delta^{AB/F}$, Figure 3a. Total fusion gain $Q_\Delta^{AB/F}$ is thus directly evaluated by corresponding weighted integration of across F, equation (11).

$$Q^{\Delta\{A,B\}/F} = \frac{\sum_{\forall n,m} Q_{n,m}^{\Delta\{A,B\}/F} w_{n,m}^{\{A,B\}}}{\sum_{\forall n.m} w_{n,m}^{A} + w_{n,m}^{B}} \qquad (10)$$

$$Q_\Delta^{AB/F} = \frac{\sum_{\forall n.m} Q_{n,m}^{\Delta A/F} w_{n,m}^{A} + Q_{n,m}^{\Delta B/F} w_{n,m}^{B}}{\sum_{\forall n,m} w_{n,m}^{A} + w_{n,m}^{B}} \qquad (11)$$

From equations (10) and (11) it is easy to see that the total fusion gain of a fusion process is the sum of the individual gains with respect to each input:

$$Q_\Delta^{AB/F} = Q^{\Delta A/F} + Q^{\Delta B/F} \qquad (12)$$

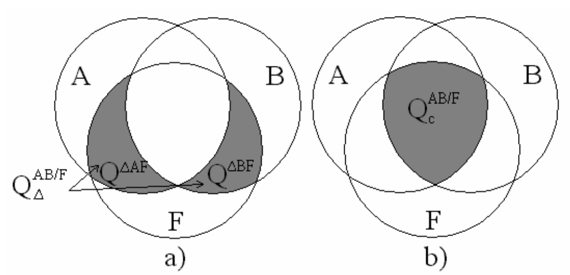

Figure 3: Exclusive information fused from both inputs $Q^{\Delta A/F}$ and $Q^{\Delta B/F}$ - total fusion gain $Q_\Delta^{AB/F}$ a), and common information fused into F $Q_c^{AB/F}$ b)

3.3. Common information in F

Local estimates of the information common to all inputs, equation (7) $Q^C_{n,m}$ are integrated into the total common information contribution $Q_C^{AB/F}$. $Q_C^{AB/F}$ shown graphically in Figure 3b, is evaluated by integrating a weighted $Q^C_{n,m}$ across the fused image, equation (13).

$$Q_C^{AB/F} = \frac{\sum_{\forall n,m} Q_{n,m}^{C} (w_{n,m}^{A} + w_{n,m}^{B})}{\sum_{\forall n,m} w_{n,m}^{A} + w_{n,m}^{B}} \qquad (13)$$

Common information contribution $Q_C^{AB/F}$ is complimentary to the exclusive information $Q_\Delta^{AB/F}$ and they make up the total input information representation by the fused image, $Q^{AB/F} = Q_C^{AB/F} + Q_\Delta^{AB/F}$. Substituting from (12) a full breakdown of the information transfer from the inputs is available that provides a trully comprahensive characterisation of the information fusion process A,B $\rightarrow$ F, equation (14).

$$Q^{AB/F} = Q_C^{AB/F} + Q^{\Delta A/F} + Q^{\Delta B/F} \qquad (14)$$

3.4. Fusion loss

Fusion loss $L^{AB/F}$ is a measure of the information *lost* during the fusion process. This is information available in the input images but not in the fused image. Fusion loss is illustrated graphically in Figure 4a. A direct loss of information is identified as Q^{AF} and Q^{BF} values lower than 1, however in order to evaluate fusion loss properly one must be able to distinguish it from fusion artifacts, see 3.5, that also result in Q^{AF} and $Q^{BF} < 1$, The $Q^{AB/F}$ framework differentiates using relative gradient strengths in the inputs and fused images, equation (1). Hence, a following classification is made for each location: if gradient strength in F is larger than that in the inputs, F contains artifacts; conversely, a weaker gradient in F indicates a loss of input information. The total fusion information loss is thus evaluated as a perceptually weighted local fusion loss, given as $1-Q^{AF}$ and $1-Q^{BF}$ for both inputs A and B, equation (15), integrated over locations where signal gradient is stronger in the inputs compared to the fused image, i.e. where $r_{n,m}$ flag is 1, equation (16).

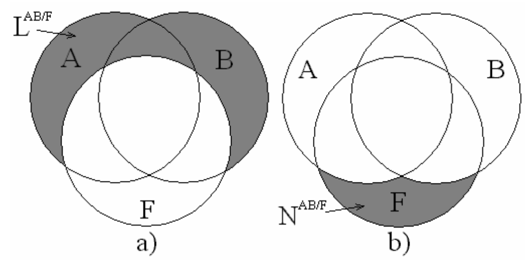

Figure 4: Fusion information loss $L^{AB/F}$ b) and fusion artifacts, $N^{AB/F}$ b)

$$L^{AB/F} = \frac{\sum_{\forall n,m} r_{n,m} [(1-Q_{n,m}^{AF}) w_{n,m}^{A} + (1-Q_{n,m}^{BF}) w_{n,m}^{B}]}{\sum_{\forall n,m} w_{n,m}^{A} + w_{n,m}^{B}} \qquad (15)$$

$$r_{n,m} = \begin{cases} 1 \ , \ if \ \ g_{n,m}^{F} < g_{n,m}^{A} \ or \ g_{n,m}^{F} < g_{n,m}^{B} \\ 0 \ , \ otherwise \end{cases} \qquad (16)$$

3.5. Fusion artifacts

Fusion artifacts represent visual information introduced into the fused image by the fusion process that has no corresponding features in any of the inputs. Fusion artifacts are essentially false information that directly detracts from the usefulness of the fused image, and can have serious consequences in certain fusion applications. Fusion artifacts are identified graphically as the information exclusive to F in Figure

4b. Within the adopted framework fusion artifacts can be evaluated as gradient information that exists in F but not in any of the input images.

Presently, local estimates of fusion artifacts or *fusion noise* as they are sometimes called, $N_{n,m}$ are evaluated as fusion loss at locations where fused gradients are stronger than input, equation (17). Total fusion artifacts for the fusion process A,B $\rightarrow$ F are evaluated as a perceptually weighted integration of the fusion noise estimates over the entire fused image, equation (18).

$$N_{n,m} = \begin{cases} 2 - Q^{AF}_{n,m} - Q^{BF}_{n,m}, \; if \;\; g^F_{n,m} > (g^A_{n,m} \,\&\, g^F_{n,m}) \\ 0 \qquad\qquad\qquad\qquad , \; otherwise \end{cases} \quad (17)$$

$$N^{AB/F} = \frac{\sum_{\forall n,m} N_{n,m}(w^A_{n,m} + w^B_{n,m})}{\sum_{\forall n,m}(w^A_{n,m} + w^B_{n,m})} \qquad (18)$$

4. Results

The proposed fusion performance characterisation approach is demonstrated on the evaluation of a comprehensive range of multiresolution and multiscale image fusion algorithms long considered one of the most effective methods of merging two images [1-6]. Initially, input images are decomposed into a so called *pyramid* representation by itterative resolution reduction and extraction of high frequency components into pyramid sub-bands. A new image pyramid is then constructed from the values of the input pyramids using a pre-defined fusion criterion. In this manner it is possible to consider features of differenct scales separately even if they occupy overlapping areas in the image [1,2]. Some multiresolution representations also offer *orientation sensitivity* additionally separating information according to orientation [3-6]. Once a new, fused pyramid is constructed, multiresolution reconstruction is applied that produces the fused image. Multiscale approach is similar, but uses no resolution reduction resulting in a vastly redundant pyramid representation [3,7]. A number of such multiresolution fusion approaches were considered.

Contrast [1] and Laplacian [2] pyramids are derived from the Gaussian low-pass pyramid, as is the gradient pyramid that also offers orientation sensitivity [6]. Discrete wavelet transform was considered both in the form of a multiresolution DWT and multiscale DWF fusion framework [2-5]. A simple, block multiscale fusion algorithm based on two scales only BMIF [7] was also considered along with the simplest method of image fusion, signal averaging included for reference. All methods (bar averaging) were implemented using the select absolute max pyramid fusion strategy and equivalent decomposition depth (5 resolution levels). The algorithms were comprehensively evaluated on a dataset of 166 different multisensor image pairs, covering a wide range of fusion application scenarios. The performance of the different fusion algorithms is characterised using the proposed framework through a selection of fusion performance parameters averaged over the 166 input image pairs in Table 1.

Table 1: Performance characterisation results for the tested fusion schemes

Fusion scheme	$Q^{AB/F}$	Q_C	Q_Δ	$L^{AB/F}$	$N^{AB/F}$
Averaging	0.449	0.294	0.155	0.551	0
Contrast	0.539	0.193	0.346	0.26	0.201
Laplacian	0.621	0.195	0.425	0.226	0.153
Gradient	0.554	0.247	0.307	0.42	0.026
DWT mres	0.574	0.181	0.393	0.233	0.193
DWF msc	0.643	0.206	0.437	0.224	0.194
BMIF	0.562	0.186	0.377	0.257	0.133

Image averaging as a special case provides a good validation for the proposed framework. As expected it has the weakest overall fusion performance, however performs optimally on common information, again expected as averaging of two identical signals produces a perfect reconstruction. Conversely, it performs poorly on exclusive information often obliterated through destructive superposition, which is also reflected in a very high fusion loss. No inherent signal boosting results in no visible fusion artifacts and hence $N^{AB/F}$=0. Note that fusion loss $L^{AB/F}$, fusion artifacts $N^{AB/F}$ and fused information score $Q^{AB/F}$ are complimentary indicating a comprehensive evaluation the information fusion process, i.e. $Q^{AB/F} + L^{AB/F} + N^{AB/F} = 1$.

Of the other systems wavelet multiscale approach performs the best ($Q^{AB/F}$ =0.643) mainly due to fewer artifacts ($N^{AB/F}$=0.133) compared to the multiresolution system ($N^{AB/F}$=0.233), also identified in [4]. This is also illustrated on image fusion of two input images in Figure 5a and 5b with four different fusion algorithms. In practice, DWT fusion suffers from reconstruction errors or *ringing artifacts* [5] visible on the water surfaces in image Figure 5c. The artifacts are not visible in the DWF fused image 5d. Image averaging 5e produces an image with much of the contrast and details lost explaining the high information loss of this scheme ($L^{AB/F}$ =0.551).

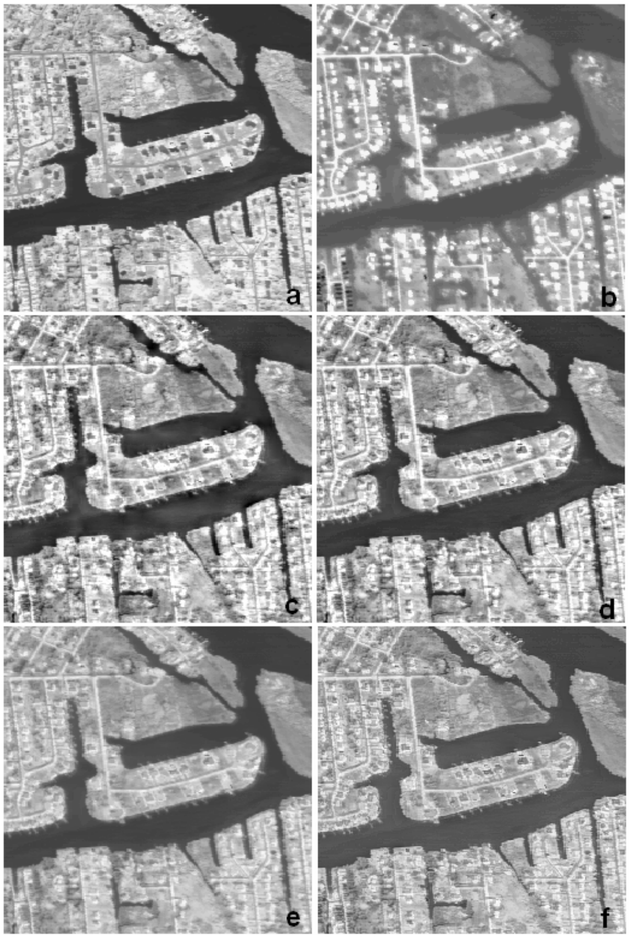

Figure 5: Fusion example: input images a) and b); fused images: DWT c), DWF d), image averaging e), gradient pyramid f)

An interesting result is that for gradient pyramid fusion that shows a very high fusion loss ($L^{AB/F}$ =0.42) and low fusion artifacts ($N^{AB/F}$=0.026) clearly seen on the gradient pyramid fused image in Figure 5f. The difference compared to Laplacian fusion can be explained by the introduction of orientation sensitivity which is the only difference between these two approaches [2,6].

5. Conclusions

This paper proposed a comprehensive approach to objective image fusion performance characterisation that provides a much deeper insight into the advantages and disadvantages of fusion algorithms than has been possible so far. Based on a gradient information representation approach, the method provides a quantification of all important aspects of information transfer involved in image fusion. The system robustly evaluates fusion gain as the advantage of applying fusion, individual sensor contributions, information loss in fusion and fusion artifacts. The approach is demonstrated on the evaluation of a wide range of multiresolution image fusion algorithms whose performance is presented in an as yet unconsidered detail. The results validate performance trends such as that redundant multiscale representations result in fewer fusion artifacts than non-redundant, previously available only through extensive observation tests.

Further work on the proposed concepts will include the introduction of the spatial dimension to the evaluation that would enable to accurately identify performance aspects within the reference of the actual image signals. A comprehensive evaluation of a wider range of fusion algorithms will also be undertaken to robustly characterise their performance.

References

[1] A Toet, L van Ruyven, J Velaton, “Merging thermal and visual images by a contrast pyramid”, *Optical Engineering*, 1989, Vol. 28(7), pp 789-792

[2] Z Zhang, R Blum, A Categorization of Multiscale-Decomposition-Based Image Fusion Schemes with a Performance Study for a Digital Camera Application, *Proc. of the IEEE*, Vol. 87(8), 1999, pp1315-1326

[3] H Li, B Munjanath, S Mitra, Multisensor Image Fusion Using the Wavelet Transform, *Graphical Models and Image Proc.*, Vol. 57(3), 1995, pp 235-245

[4] Y Chibani, A Houacine, Redundant versus orthogonal wavelet decomposition for multisensor image fusion, *Pattern Recognition*, Vol. 36, 2003, pp 879-887

[5] L Chipman, T Orr, L Graham, "Wavelets and Image Fusion", *Proc. SPIE*, Vol. 2569, 1995, pp 208-219

[6] P Burt, R Kolczynski, "Enhanced Image Capture Through Fusion", *Proc. 4th International Conference on Computer Vision*, Berlin 1993, pp 173-182

[7] V Petrović, C Xydeas, Computationally Efficient Pixel-level Image Fusion, *Proceedings of Eurofusion99*, Stratford-upon-Avon, October 1999, pp 177-184

[8] C Xydeas, V Petrović, Objective Image Fusion Performance Measure, *Electronics Letters*, Vol. 36, No.4, February 2000, pp 308-309

[9] G Qu, D Zhang, P Yan, Information measure for performance of image fusion, *Electronics Letters*, Vol. 38(7), pp 313-315

[10] G Piella, H Heijmans, A New Quality Metric for Image Fusion, *Proceedings International Conference on Image Processing*, 2003, Vol. 3, pp 173-176

[11] V Petrović, C Xydeas, Evaluation of Image Fusion Performance with Visible Differences, *Proceedings of ECCV*, Prague 2004, LNCS Vol. 3023, pp 380

[12] Burt P, Adelson E, The Laplacian pyramid as a compact image code, *IEEE Transactions on Communications*, COM-31(4), 532-5