



Cvejic, N., Bull, D. R., & Canagarajah, C. N. (2007). Improving fusion of surveillance images in sensor networks using independent component analysis. *IEEE Transactions on Consumer Electronics*, 53(3), 1029 - 1035. 10.1109/TCE.2007.4341582

Link to published version (if available):  
[10.1109/TCE.2007.4341582](https://doi.org/10.1109/TCE.2007.4341582)

[Link to publication record in Explore Bristol Research](#)  
PDF-document

## University of Bristol - Explore Bristol Research

### General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:  
<http://www.bristol.ac.uk/pure/about/ebr-terms.html>

### Take down policy

Explore Bristol Research is a digital archive and the intention is that deposited content should not be removed. However, if you believe that this version of the work breaches copyright law please contact [open-access@bristol.ac.uk](mailto:open-access@bristol.ac.uk) and include the following information in your message:

- Your contact details
- Bibliographic details for the item, including a URL
- An outline of the nature of the complaint

On receipt of your message the Open Access Team will immediately investigate your claim, make an initial judgement of the validity of the claim and, where appropriate, withdraw the item in question from public view.

# Improving Fusion of Surveillance Images in Sensor Networks Using Independent Component Analysis

Nedeljko Cvejic<sup>1</sup>, *Member, IEEE*, David Bull<sup>1</sup>, *Senior Member, IEEE*,  
Nishan Canagarajah<sup>1</sup>, *Member, IEEE*

**Abstract** — *In this paper we present a novel algorithm for fusion of multimodal surveillance images, based on ICA, which has an improved performance over sensor networks. Improvements have been demonstrated through separate training process for different modalities and the use of a fusion metric to maximise the quality of the fused image. Sparse coding of the coefficients in ICA domain is used to minimize noise transferred from input images into the fused output. Experimental results confirm that the proposed method outperforms other state-of-the-art methods in the sensor network environment, characterized by JPEG 2000 compression and data packetisation.*

**Index Terms** — Image fusion, Independent Component Analysis, fusion metrics, sensor networks, JPEG 2000.

## I. INTRODUCTION

As the size and cost of sensors decrease, sensor networks are increasingly becoming an attractive method to collect information in a given area [1]. However, there are still many technical challenges, mainly related to fusing the individual sensor data through an intelligent decision making process while reducing errors and compression noise [2]. Multi-sensor data often presents complementary information about the region surveyed and data fusion provides an effective method to enable comparison, interpretation and analysis of such data. Image and video fusion is a sub area of the more general topic of data fusion, dealing with image and video data [3]. The aim of image fusion, apart from reducing the amount of data, is to create new images that are more suitable for the purposes of human/machine perception, and for further image-processing tasks such as segmentation, object detection or target recognition [4].

A relatively lower level of interest in infrared imagery, compared to visible imagery, has been due to high cost of thermal sensors, lower image resolution, higher image noise and lack of widely available data sets. However, these drawbacks are becoming less relevant as infrared imaging advances, making the technology important for applications such as video surveillance and navigation and object tracking. Night vision cameras, which produce images in multiple spectral bands, e.g. thermal and visible, also became available.

<sup>1</sup> This work has been funded by the UK Data and Information Fusion Defence Technology Centre (DIF DTC). Authors are with the Centre for Communications Research, Department of Electrical and Electronic Engineering, University of Bristol, Merchant Venturers Building, Woodland Road, Bristol BS8 1UB, United Kingdom. Corresponding author's email: n.cvejic@bristol.ac.uk.

These different bands provide complementary information since they represent different characteristics of a scene

Fusion of visible and infrared (IR) images and video sources is becoming increasingly important for surveillance purposes. The main reason is that a fused image, constructed by combination of features of visible and infrared inputs, enables improved detection and unambiguous localisation of a target (represented in the thermal image) with respect to its background (represented in the visible image) [5]. A human operator using a suitably fused representation of visible and IR images may therefore be able to construct a more complete and accurate mental representation of the perceived scene, resulting in a larger degree of situation awareness [6].

The image fusion process can be performed at different levels of information representation: signal, pixel, feature and symbolic level. Nikolov et al [4] proposed a classification of image fusion algorithms into spatial domain and transform domain techniques. The transform domain image fusion consists of performing a transform on each input image and, following specific rules, combining them into a composite transform domain representation. The composite image is obtained by applying the inverse transform on this composite transform domain representation.

Instead of using a standard bases system, such as the DFT, the mother wavelet or cosine bases of the DCT, one can train a set of bases that are suitable for a specific type of image. A training set of image patches, which are acquired randomly from images of similar content, can be used to train a set of statistically independent bases. Independent Component Analysis (ICA) is a widely used method that is able to identify statistically independent basis vectors in a linear generative model [7]. Recently, several algorithms have been proposed [8,9], in which ICA and bases are used for transform domain image fusion.

In this paper, we present a novel algorithm for fusion of multimodal images based on the ICA. It was tested in a sensors network environment and it has exhibited an improvement in performance in fusion of infrared (IR) and visible images over other state-of-the-art methods.

## II. IMAGE ANALYSIS USING ICA

In order to obtain a set of statistically independent bases for image fusion in the ICA domain, training is performed with a predefined set of images. Training images are selected in such a way that the content and statistical properties are similar for the training images and the images to be fused. An input image  $i(x,y)$  is randomly windowed using a rectangular

window  $w$  of size  $N \times N$ , centred around the pixel  $(m_0, n_0)$ . The result of windowing is an “image patch” which is defined as [8]:

$$p(m, n) = w(m, n) \cdot i(m_0 - N/2 + m, n_0 - N/2 + n) \quad (1)$$

where  $m$  and  $n$  take integer values from the interval  $[0, N-1]$ . Each image patch  $p(m, n)$  can be represented by a linear combination of a set of  $M$  basis patches  $b_i(m, n)$ :

$$p(m, n) = \sum_{i=1}^M v_i b_i(m, n) \quad (2)$$

where  $v_1, v_2, \dots, v_M$  stand for the projections of the original image patch on the basis patch, i.e.  $v_i = \langle p(m, n), b_i(m, n) \rangle$ . A 2D representation of the image patches can be simplified to a 1D representation, using lexicographic ordering. This implies that an image patch  $p(m, n)$  is reshaped into a vector  $\underline{p}$ , mapping all the elements from the image patch matrix to the vector in a row-wise fashion. Decomposition of image patches into a linear combination of basis patches can be expressed as follows:

$$\underline{p}(t) = \sum_{i=1}^M v_i(t) \underline{b}_i = [\underline{b}_1 \underline{b}_2 \dots \underline{b}_M] \cdot [v_1(t) v_2(t) \dots v_M(t)]^T \quad (3)$$

where  $t$  represents the image patch index. If we denote  $B = [\underline{b}_1 \underline{b}_2 \dots \underline{b}_M]$  and  $v(t) = [v_1(t) v_2(t) \dots v_M(t)]^T$ , then equation (3) reduces to:

$$\underline{p}(t) = B v(t) \quad (4)$$

$$v(t) = B^{-1} \underline{p}(t) = A \underline{p}(t) \quad (5)$$

Thus,  $B = [\underline{b}_1 \underline{b}_2 \dots \underline{b}_M]^T$  represents an unknown mixing matrix (analysis kernel) and  $A = [a_1 a_2 \dots a_M]^T$  the unmixing matrix (synthesis kernel). This transform projects the observed signal  $p(t)$  on a set of basis vectors. The aim is to estimate a finite set of  $K < N^2$  basis vectors that will be capable of capturing most of the input image properties and structure.

In the first stage of basis estimation, Principal Component Analysis (PCA) is used for dimensionality reduction. This is obtained by eigenvalue decomposition of the data correlation matrix  $C = E\{\underline{p}\underline{p}^T\}$ . The eigenvalues of the correlation matrix illustrate the significance of their corresponding basis vector [8]. If  $V$  is the obtained  $K \times N^2$  PCA matrix, the input image patches are transformed by:

$$\underline{z}(t) = V \underline{p}(t) \quad (6)$$

After the PCA preprocessing step we select the statistically independent basis vectors using the optimisation of the negentropy. The following rule defines a FastICA approach that optimises negentropy, as proposed in [7]:

$$\underline{a}_i \leftarrow \varepsilon \{ \underline{a}_i \phi(\underline{a}_i^T \underline{z}) \} - \varepsilon \{ \phi'(\underline{a}_i^T \underline{z}) \} \underline{a}_i \quad 1 \leq i \leq K \quad (9)$$

$$A \leftarrow A(A^T A)^{-0.5} \quad (8)$$

where  $\phi(x) = -\partial G(x)/\partial x$  defines the statistical properties  $G(x) = \log p(x)$  of the signals in the transform domain [7]. In our implementation we used:

$$G(x) = \alpha \sqrt{\zeta + x} + \beta \quad (9)$$

where  $\alpha$  and  $\beta$  are constants and  $\zeta$  is a small constant to prevent numerical instability, in the case that  $x \rightarrow 0$  [7].

After the input image patches  $\underline{p}(t)$  are transformed to their ICA domain representations  $\underline{v}_k(t)$ , we can perform image

fusion in the ICA domain in the same manner as it is performed in, for example, the wavelet domain. The equivalent vectors  $\underline{v}_k(t)$  from each image are combined in the ICA domain to obtain a new image  $\underline{v}_f(t)$ . The method that combines the coefficients in the ICA domain is called the “fusion rule”. After the composite image  $\underline{v}_f(t)$  is constructed in the ICA domain, we can move back to the spatial domain, using the synthesis kernel  $A$ , and synthesise the image  $i_f(x, y)$ .

### III. PROPOSED FUSION METHOD

#### A. Separated Training Sets

In the proposed method, images used for training the ICA bases are separated in two groups prior to the training process. Namely, all IR training images are grouped into a separate training subset, whereas all the visible training images constitute the second training subset. Introduction of separate training subsets provides us with two sets of ICA bases. The first ICA basis set is used to decompose the IR input image patches  $v_i(t) = A p_i(t)$  and the second subset to transform the visible input image patches to ICA domain  $v_v(t) = A_v p_v(t)$ .

TABLE I

FUSION PERFORMANCE MEASURED BY PETROVIC METRIC; THE FIGURES REPRESENT THE MEAN VALUE OF THE METRIC OVER 25 FRAMES OF THE SURVEILLANCE SEQUENCE “UN CAMP”.

| Fusion method | Number of ICA training patches |                |        |                |                |
|---------------|--------------------------------|----------------|--------|----------------|----------------|
|               | 100                            | 200            | 400    | 1000           | 2000           |
| Standard ICA  | 0.472                          | 0.533          | 0.572  | 0.581          | 0.584          |
| Proposed ICA  | 0.320                          | 0.355          | 0.387  | 0.396          | 0.401          |
| Fusion method | Number of ICA training patches |                |        |                |                |
|               | $4 \cdot 10^3$                 | $8 \cdot 10^3$ | $10^4$ | $2 \cdot 10^4$ | $4 \cdot 10^4$ |
| Standard ICA  | 0.587                          | 0.585          | 0.588  | 0.587          | 0.592          |
| Proposed ICA  | 0.405                          | 0.406          | 0.406  | 0.406          | 0.406          |

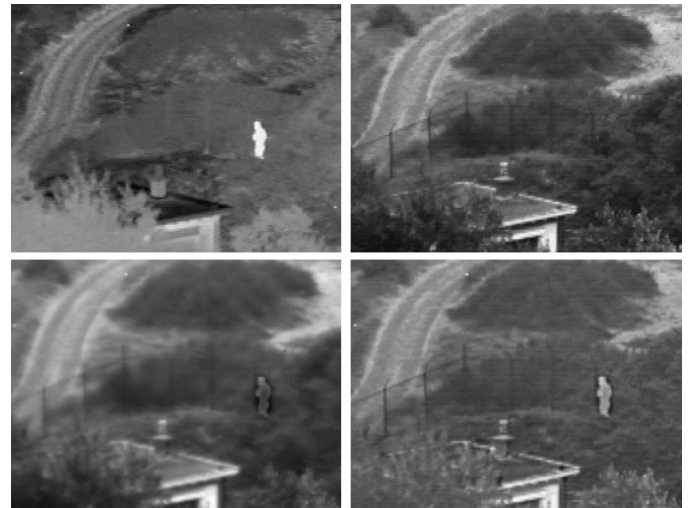


Fig. 1. Impact of the number of training patches on the subjective quality of the fused images. Top: input IR image (left) and input visible image (right). Bottom: fused image using standard ICA fusion and 100 training patches (left) and fused image using the proposed method and 100 training patches (right).

Separate ICA basis sets for decomposition of input images are more specifically trained to capture statistical properties of the specific modality of the input images (IR/visual). This

enables the proposed method to outperform the standard method [9], in which images of both IR and visible modality are used for training which results in an "average" ICA bases set that is not able to take the full advantage of ICA decomposition.

Fig.1 confirms that when two separate training sets are used, the subjective quality of the fused image is increased considerably; e.g. fence detail is far more visible and person walking is brighter and less blurred. In Table I, subjective results are reinforced by values obtained by the Petrovic image fusion metric [10]. It is clear that significantly higher metric values are obtained using separate training sets. Table I also shows that performance of the ICA fusion algorithm does not improve significantly when the number of training patches exceeds  $10^3$ . Thus, the number of training patches has been fixed to  $10^3$  in order to make a trade-off between performance and computational complexity of the algorithm.

### B. Nonlinear shrinkage of coefficients in ICA domain

In the case when the images to be fused and possibly also the set of training images are corrupted with noise, it is crucial to determine the ICA coefficients to be used in the reconstruction of the fused image so that the noise transferred from input images into the fused output is minimized. An approach similar to image denoising in the ICA domain [11] has been used to reduce noise in the fused image. Assume that we observe an  $N$ -dimensional vector  $x$  as:

$$x = s + n \quad (10)$$

where  $s$  is the vector of the original signal and  $n$  is Gaussian white noise. The goal of signal denoising is to find  $s' = g(x)$  such that  $s'$  is close to  $s$  in some well-defined sense. In order to use the standard ICA domain method to denoise noisy images [12], one needs to employ the fixed-point algorithm on the noise-free data to get the ICA transformation matrix, and then to use maximum likelihood to estimate parameters for the shrinkage scheme. The method [12] works as follows:

1. Estimate the orthogonal ICA transformation matrix  $W$  using a set of noise-free representative data  $z$ .
2. For  $i=1,2,\dots,N$  estimate a density model which approximates the actual distribution of the variable  $s_i = w_i^T z$  (where  $w_i$  is the  $i$ -th column of  $W$ ). Based on the estimated model and the variance of  $n$ , determine the nonlinear shrinkage function  $g_i$ .
3. For each observed  $x$ , ICA transform is performed ( $y=Wx$ ), nonlinear shrinkage derived ( $s_i' = g_i(y_i)$ ) and finally reverse transform calculated.

Comparing with wavelet based methods, the ICA method needs additional noise-free data to estimate the transformation matrix  $W$  and shrinkage nonlinearities. In some cases, as fusing surveillance images obtained using a sensor network, this might be difficult to obtain. Therefore, it is desirable to develop methods for estimating transformation matrix  $W$  and shrinkage nonlinearities  $g_i$  directly from noisy data. If the

method [12] is implemented on  $8 \times 8$  frames in the input images, there are  $2 \times 63$  parameters to estimate. The variance of noise in the corrupted signal must be known as well in order to facilitate the basis selection.

In order to keep the computational complexity of the algorithm low, we decided to use a shrinkage scheme in [11], in which only one control parameter is required. The experiments in [11] showed that the method with this new shrinkage scheme produces similar performance in image denoising as the method in [12], while keeping the computational load significantly lower. The shrinkage scheme described in [11] is defined as follows:

$$\bar{y}_i = \begin{cases} 0 & \text{if } |y_i| < \rho E(|y_i|), \\ y_i & \text{otherwise} \end{cases} \quad (11)$$

where  $\rho$  is a control parameter. It is clear that  $y_i$  depends on two variables:

$$y_i = (Ws)_i + (Wn)_i \quad (12)$$

In most cases  $(Ws)_i$  should be a very sparse matrix. As an illustration, let the distribution of  $(Ws)_i$  be as follows:

$$\text{Pr ob}[(Ws)_i = 0] = a \quad \text{and} \quad (13)$$

$$\text{Pr ob}[(Ws)_i = -1] = \text{Pr ob}[(Ws)_i = 1] = \frac{1-a}{2} \quad (14)$$

Because of the sparseness,  $a$  should be close to 1, for instance let  $a=0.8$  and  $(Wn)_i$  be Gaussian noise with a standard deviation 0.2. Thus,

$$E(|y_i|) \approx 0.2 \quad (15)$$

If we set  $\rho=2$  and based on the property of the Gaussian distribution, we have

$$E[\bar{y}_i - (Ws)_i]^2 \approx 0.2 \times 0.04 = 0.008 \quad (16)$$

That implies that the scheme [11] can efficiently reduce the noise in the images. It should also be pointed out that for the above approximation still holds for  $2 < \rho < 4$ . It means that this scheme is considerably robust with respect to selection of the control parameter  $\rho$ .

### C. Reconstruction of the Fused Image Using Fusion Metrics

Here, we propose a novel method for reconstruction of the fused image using statistical properties of both the input images. In the standard ICA method, reconstruction of the fused image is performed on the patch-per-patch base [8]:

$$p_f(t) = U(t) + M_i(t) + M_v(t) \quad (17)$$

where  $p_f(t)$  represents the  $t$ -th patch of the fused image  $i_f(x,y)$  and  $U(t)$  is the  $t$ -th frame obtained by inverse transform of the fused ICA coefficients.  $M_i(t)$  is the mean value of the corresponding frame from the IR input image  $i_i(x,y)$  and  $M_v(t)$  is the mean value of the corresponding frame from the visual input image  $i_v(x,y)$ . We propose a new approach for reconstruction of the fused image [9]:

$$p_f(t) = U(t) + M_i(t) \cdot w_i + M_v(t) \cdot w_v \quad (18)$$

Weights  $w_i \in [0,1]$  and  $w_v (=1-w_i) \in [0,1]$  are used to balance the contributions from both visual and IR images in the synthesis of the fused image. Weighting coefficients are set to

a predefined value (e.g.  $w_i=1$  and  $w_v=0$ ) and then gradually increased/decreased. The Petrovic metric [10] is calculated at each step and when the maximum value of a fusion performance metric is reached, the process stops and reconstruction of the fused image is performed with the calculated weights. In that sense, the weighting coefficients are chosen so that the quality of the fused image is maximized.

#### IV. IMPACT OF ERRORS ON PERFORMANCE OF IMAGE FUSION ALGORITHMS

To advance the work on image fusion without incurring the overhead of communications system simulation, we have investigated the possibilities for abstracting the behaviour of the underlying paths between sensors and sinks.

In the case of image fusion applications, it is relatively easy to study the effect of packet losses on the quality of received and fused images. A simple way of doing this would be to apply a random loss parameter to the image data to simulate the loss of a certain percentage of the transmitted packets and then see what the output quality of the image is. This method is only good for studying the effect of random packet losses or delays of a certain probability, i.e. as if the packets were going through a network that exhibited stationary impairment behaviour, meaning that errors are mutually independent.

However, it is rare that any network would reveal such stationary behaviour. In reality, network behaviour can be highly dynamic and correlated, with many factors influencing the presented impairments, such as congestion, buffer overflows, traffic patterns and route reliability. Sensor networks especially add further factors such as unpredictable radio conditions, mobility, energy levels and node loss possibilities.

A quick assessment of these aspects of the wireless sensor networks reveals that they can potentially exhibit a dynamic and rapidly changing response to any traffic. Application testing for these types of networks thus becomes a more complex and non-trivial task. It is desirable to find a way of accurately modelling the impairment behaviour of these networks and for these models to be used for better application testing.

##### A. Packet Loss

In the set of simulations described below, each JPEG frame was assumed to be 50kB in size and the frame data was accordingly fragmented into number packets, where packet length ranges from 64 to 1024 bytes – note that this assumption eliminates residual variable-length frames. The loss data from the trace files can be used to derive which parts of each JPEG frame were lost; the remaining packets can then be pieced together to see what effect on the image the lost packets have had (within an image fusion context).

The trace files are sorted according to the arrival time of each packet at the destination node. This means that packets are displayed as they arrive, i.e. the packets from each traffic flow are combined. A sort of the impairment data based on

source, destination, port number and sequence number (in that order) will separate the flows contained within each trace file so that they can be examined individually and the impairment mapping on the real application data applied. Passing real application data through the generated network impairments will have the effect of emulating the flow of the traffic through a wireless sensor network, thus providing results approximating a real system.

##### B. Packet Delay

In a system with real-time constraints (e.g. real-time image fusion performed during a military operation), network impairments associated with packet delay become important as there is a limited time window within which each fusion task can be performed. Packets arriving at a fusion node from different sources will not necessarily be aligned in time, so the fusion algorithm must cater for this in its processing of source images. A way to incorporate delay characteristics in the evaluation of a fusion algorithm could be to treat packets received after a certain time delay as lost packets and then performing the fusion task only with currently available data. For example, consider a real time image fusion system that needs half a second to perform a fusion task and is receiving images at the rate of one every second. In such a case, the system can only afford to use packets that have been received within the first half-second following the previous fusion task so that the current fusion task will be completed on time.

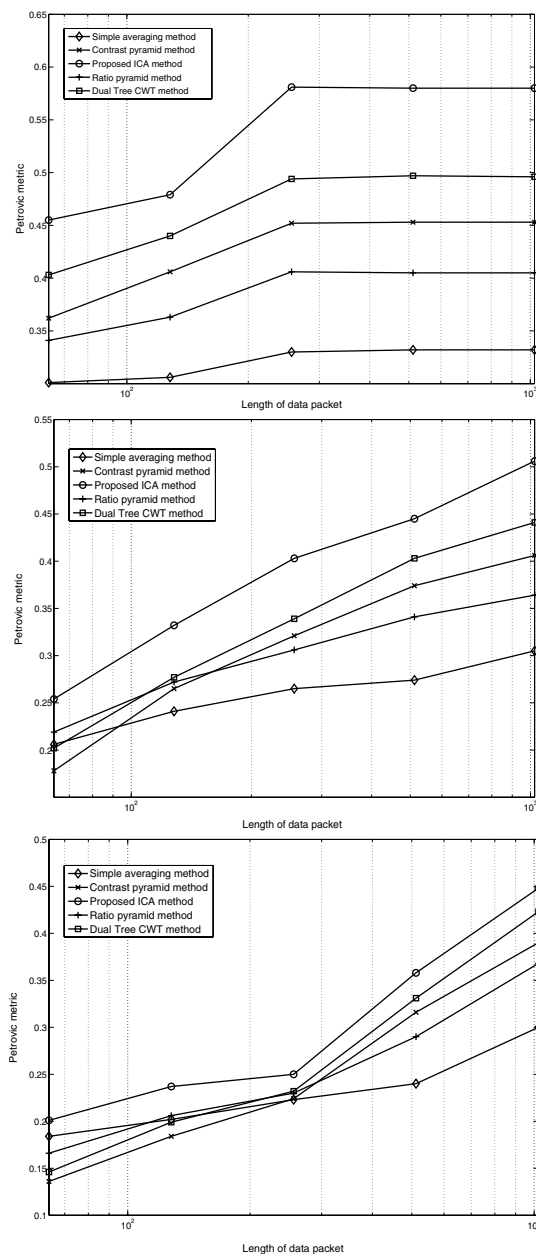
##### C. Markov modelling of network impairments

There are many stochastic processes in the area of wireless communications that exhibit analytical intractability, and other methods have to be used to estimate their performance. One such method which has been employed with success in a wide variety of problems is the notion of Markov chains (MC) [13]. With the help of existing efficient statistical tools, MC can often be applied to the specifics of different random stochastic problems in an accurate manner.

Markov models have been used extensively in areas such as wireless channel characterisation [14], for the characterisation of packet loss and delay at the radio [15], and at the MAC level [16]. Of course, the higher one progresses into the system, composing services layer by layer, to investigate the dynamics that occur, the more complex it is to describe the behaviour, since more parameters and relationships become important. Markov models exist in many forms (e.g. Multiple Markov chains, Monte Carlo based models, etc.) and there are potentially many parameters to configure without any default approach being available. Hence, intuition and experience play an important role in the ability to produce an accurate model using MC.

In the experiments presented in Section V we modelled the impairments that the network impacts upon the traffic it generates by the time it reaches its final destination within the network. This comes mainly in the form of packet delay and loss that any type of application traffic may be subjected to whilst relaying its packets through the sensor network. Our focus is aimed specifically at modelling the packet irregularities

within a sensor network used for image fusion applications. We have aimed to be generic in approach, anticipating that this technique can easily be ported to other application types.



**Fig. 2. Fusion performance in a simulated wireless sensor network, measured by Petrovic metric. Values represent the mean value of the metric over 32 fused images in the image sequence “UN Camp”. Top: Fusion performance for probability of packet loss equal to  $10^{-3}$ . Middle: Fusion performance for probability of packet loss equal to  $10^{-2}$ . Bottom: Fusion performance for probability of packet loss equal to  $10^{-1}$ .**

It was decided to use one of the simplest of models to see how well it models the loss behaviour. A sequence of 2500 packets (amongst other traffic) was sent over a wireless ad-hoc network (simulated according to the parameters in [17]) with the received packets being processed to result in a binary sequence showing the temporal loss, i.e. which packets were dropped. This binary sequence was then used to train a 2-state Markov

model. Each of the two states corresponded to the status of the network for the current packet (i.e. successful packet delivery or packet loss). State 0 is a packet loss and state 1 is a successful packet delivery. The resulting state transition and emission matrices were then used as the model parameters to generate a new binary sequence, corresponding to the temporal loss produced by the Markov modelling process.

## V. EXPERIMENTAL RESULTS

The proposed image fusion method was tested in a surveillance scenario with two input images: infrared and visible. The images used in our experiments are surveillance images from TNO Human Factors and Octec Ltd., publicly available at the Image Fusion web site [18]. We compared the proposed method with a simple averaging method, the contrast pyramid (CP) method [3], ratio pyramid (RP) method [3] and the dual-tree complex wavelet transform (DT-CWT) [19]. CP, RP and DT-CWT methods have been chosen for comparison because they have been previously reported to obtain excellent performance in multimodal image fusion [3,11]. Before performing image fusion using the proposed algorithm, the ICA bases were trained using a set of images with content comparable to the test set. The number of rectangular patches ( $N=8$ ) used for training was 1000, randomly selected from the training set. Obtained ICA coefficients are combined using the principle described in Section 3, while reconstruction of the fused image was done using optimisation based on the Petrovic metric [11].



**Fig. 3. Subjective fusion results, image 1827, UN camp sequence, packet length 128 bytes, probability of packet loss  $10^{-1}$ . Top: input IR image (left), input visible image (right). Middle: fused image using averaging (left) and contrast pyramid (right). Bottom: fused image using DT-CWT (left) and the proposed ICA method (right)**

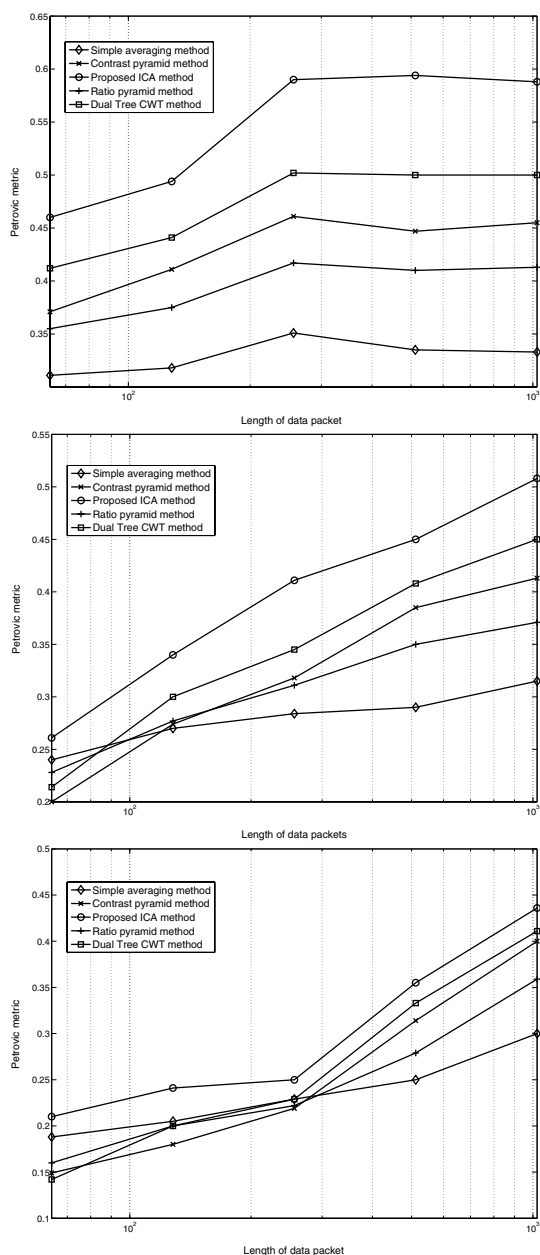


Fig. 4. Fusion performance in a simulated wireless sensor network, measured by Petrovic metric. Values represent the mean value of the metric over 19 fused images in the image sequence "Trees". Top: Fusion performance for probability of packet loss equal to  $10^{-3}$ . Middle: Fusion performance for probability of packet loss equal to  $10^{-2}$ . Bottom: Fusion performance for probability of packet loss equal to  $10^{-1}$ .

In order to evaluate performance of the image fusion algorithms in the sensor network environment, input images were first compressed using JPEG2000. Sensor network transmission was simulated by dividing the image into data packets. These packets were transmitted over a simulated sensor network in which the packet loss was modelled by the 2-state Markov model, obtained as described in Section IV. After the recovered packets at the receiver side were recomposed into images, these images were used as inputs for the fusion algorithms. The performance of methods was measured using the Petrovic metric and the results are given in Fig. 2. and Fig. 4.

Results show that the proposed algorithm performs significantly better in the sensor network environment than the other state-of-the-art methods, for all data packet lengths and probability of packet loss. The subjective quality of the fused images using the aforementioned fusion methods is presented in Fig. 3 and Fig. 5. When fused outputs are inspected visually, it can be seen that the fused image obtained using the proposed algorithm incorporates less noise from each of the input images. Compared to the multiresolutional methods, the noise in the fused image obtained using the proposed method is visually less annoying, while the important detail is still adequately represented.

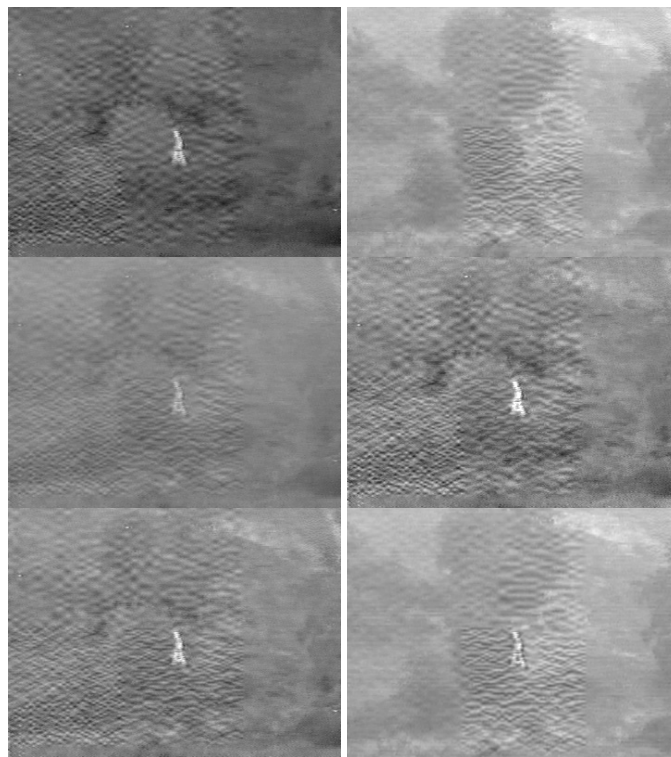


Fig. 5. Subjective fusion results, image 4917, Trees sequence, packet length 128 bytes, probability of packet loss  $10^{-1}$ . Top: input IR image (left), input visible image (right). Middle: fused image using averaging (left) and contrast pyramid (right). Bottom: fused image using DT-CWT (left) and the proposed ICA method (right)

In general, the proposed method is able to more successfully suppress the noise in the input images in different modalities, reducing significantly the distortion in the fused image, both visually and objectively. It also outperforms other state-of-the-art algorithms, in terms of the Petrovic fusion performance metric.

## REFERENCES

- [1] Q. Wang, Y. Zhu and L. Cheng, "Reprogramming wireless sensor networks: challenges and approaches," *IEEE Network*, Vol. 20, No. 3, pp. 48-55.
- [2] R. Jiang and B. Chen, "Fusion of censored decisions in wireless sensor networks," *IEEE Transactions on Wireless Communications*, Vol. 4, No. 6, pp. 2668-2673.
- [3] R. Blum and Z. Liu, *Multi-sensor Image Fusion and Its Applications*, CRC Press, London, United Kingdom, 2005.

- [4] S. Nikolov, P. Hill, D. Bull and N. Canagarajah, "Wavelets for image fusion", chapter in *Wavelets in Signal and Image Analysis*, Kluwer, Dordrecht, The Netherlands, 2001.
- [5] A. Toet, J. K. Ijspeert, A. M. Waxman and M. Aguilar, "Perceptual evaluation of different image fusion schemes," *Displays*, Vol. 24, pp. 25–37, 2003.
- [6] A. Toet and E. M. Franken, "Fusion of visible and thermal imagery improves situational awareness," *Displays*, Vol. 18, pp. 85–95, 1997.
- [7] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent Component Analysis*, John Wiley and Sons, 2001.
- [8] N. Mitianoudis, T. Stathaki, "Pixel-based and Region-based Image Fusion schemes using ICA bases," *Information Fusion*, to appear.
- [9] N. Cvejic, J. Lewis, D. R. Bull, and C. N. Canagarajah "Region-Based Multimodal Image Fusion Using ICA Bases," *Proc. IEEE International Conference on Image Processing*, Atlanta, GA, 1801-1804, 2006.
- [10] V. Petrovic, and C. Xydeas, "Objective evaluation of signal-level image fusion performance," *Optical Engineering*, Vol. 44, No. 8, 08703, 2005.
- [11] Q. Zhang, H. Yin, and N. Allinson, "A simplified ICA based denoising method," *Proc. IEEE-INNS-ENNS International Joint Conference on Neural Networks*, Como, Italy, pp. 479–482, 2000.
- [12] A. Hyvärinen, "Sparse code shrinkage: Denoising of non-Gaussian data by maximum likelihood estimation," *Neural Computation*, Vol. 11, No. 7, pp. 1739–1768, 1999.
- [13] L. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proceedings of the IEEE*, Vol. 77, No. 2, pp. 257–286, 1999.
- [14] W. Turin, and R. van Nobelen, "Hidden Markov modelling of flat fading channels," *IEEE Journal on Selected Areas in Communications*, Vol. 16, No. 9, pp. 1809–1817, 1998.
- [15] M. Hassan, M. Krunz, and I. Matta, "Markov-based channel characterization for tractable performance analysis in wireless packet networks," *IEEE Transactions on Wireless Communications*, Vol. 3, No. 3, pp. 821–831, 2004.
- [16] G. Bianchi, "Performance analysis of the IEEE 802.11 distributed coordination function," *IEEE Journal on Selected Areas in Communications*, Vol. 18, No. 3, pp. 535–547, 2000.
- [17] S. Tan, C. Efthymiou, D. Kothris, A. Munro, A. Nix, and D. Bull, "Joint application and communications modelling objectives, metrics and solutions," *Data and Information Fusion Defence Technology Centre Report*, UOB-DIF-DTC-PROJ202-TR09, University of Bristol, 2006.
- [18] ImageFusion web site, <http://imagefusion.org>
- [19] J. Lewis, R.J. O'Callaghan, S. Nikolov, D. Bull, and N. Canagarajah. "Pixel- and region-based image fusion with complex wavelets," *Information Fusion*, to appear.



**Nedeljko Cvejic** received the Dipl.-Ing. degree in electrical engineering from the University of Belgrade, Serbia, in 2000 and the Dr. Tech. degree from the University of Oulu, Finland, in 2004. From 2001 to 2004, he was a Research Scientist at the Department of Electrical and Information Engineering, University of Oulu, Finland.

He is currently a Research Associate with the Department of Electrical and Electronic

Engineering of the University of Bristol, United Kingdom. He has published more than 50 papers and a book. His research interests include image and video fusion, sensor networks, and digital watermarking.



**David Bull** is Professor of Signal Processing and Head of the Electrical and Electronic Engineering Department at the University of Bristol. Prior to his current appointment he has been a Systems Engineer at Rolls Royce and subsequently a Lecturer at Cardiff University. He leads the Signal Processing activities within the Centre for Communications Research where he is Deputy Director. He has been a member of the UK

Foresight Panel and the Steering Group for the DTI/EPSRC LINK programme in Broadcast Technology. He is a past director of the VCE in Digital Broadcasting and Multimedia Technology and is currently Chairman and Technical Director of ProVision Communication Technologies Ltd., specialising in wireless multimedia communications. Between 2003 and 2005 he was also a member of the Science and Technology Board of the UK Defence Technology Centre in Data and Information Fusion.

David Bull has worked widely in the fields of 1 and 2-D signal processing and has published over 250 papers, various articles and 2 books. He has won two IEE Premium awards for this work. His current research is focused on the problems of image and video communications for low bit rate wireless, internet and broadcast applications. He is widely supported in these areas by both industry, Europe, MoD and EPSRC.



**Nishan Canagarajah** is currently a Professor of Multimedia Signal Processing at Bristol. Prior to this he was an RA and lecturer at Bristol investigating DSP aspects of mobile radio receivers. He has BA (Hons) and a PhD in DSP Techniques for Speech Enhancement both from the University of Cambridge. His research interests include image and video coding, image segmentation, content based video retrieval, 3D video and image fusion.

He is widely supported in these areas by industry, EU and the EPSRC. He has been involved in a number of EU FP5 and FP6 projects where the team has been developing novel image/video processing algorithms. He has published more than 160 papers and two books. He is a member of the Executive Team of the IEE PN on Multimedia Communications and a member of EPSRC Peer Review College.