Peer reviewed version

Link to publication record in Explore Bristol Research
PDF-document

Region-Based Multimodal Image Fusion Using ICA Bases

Nedeljko Cvejic, David Bull, and Nishan Canagarajah

Abstract—**In this paper, we present a novel multimodal image fusion algorithm in the independent component analysis (ICA) domain. Region-based fusion of ICA coefficients is implemented, where segmentation is performed in the spatial domain and ICA coefficients from separate regions are fused separately. The ICA coefficients from given regions are consequently weighted using the Piella fusion metric in order to maximize the quality of the fused image. The proposed method exhibits significantly higher performance than the basic ICA algorithm and also shows improvement over other state-of-the-art algorithms.**

Index Terms—**Fusion metrics, image fusion, independent component analysis (ICA), region-based fusion.**

I. INTRODUCTION

RECENT advances in sensor technology, microelectronics, and wireless communications have enabled the development of low-cost, multifunctional sensor nodes with sensing, data processing, memory, battery, and communication components [1], [2]. These have naturally evolved into sensor networks. Simultaneously, the mentioned advances have aroused a need for processing techniques that efficiently combine the information from different sensors into a single composite for interpretation [3].

Multisensor data often presents complementary information about the region surveyed, so data fusion provides an effective method to enable comparison and analysis of such data [4]. Image and video fusion is a sub area of the more general topic of data fusion, dealing with image and video data. 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 in applications such as remote sensing and medical imaging.

The fusion of multimodality images and video sources is becoming increasingly important for surveillance purposes, navigation, and object tracking applications. The main reason for combining visible and infrared (IR) sensors is that a fused image, constructed by combination of features, enables improved detection and unambiguous localization of a target (represented in the IR 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.* [7] 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. This is known as independent component analysis (ICA) [8]. Recently, several algorithms have been proposed [9], in which ICA and bases are used for transform domain image fusion.

In this paper, we refine the approach by a novel multimodal image fusion algorithm in the ICA domain. This uses separate training subsets for visible and IR images to determine the most important regions in the input images and, consequently, fuses the ICA coefficients using fusion metrics to maximize the quality of the fused image.

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 $N \times N$, centered around the pixel (m_o, n_0) The result of windowing is an "image patch" which is defined as [7]

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

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 $v_1, v_2, \ldots, v_M$ stand for the projections of the original image patch on the basis patch, i.e., $v_i = \langle p(m, n), b_i(m, n) \rangle$.

The authors are with the Centre for Communications Research, University of Bristol, Bristol BS8 1UB, U.K. (e-mail: n.cvejic@bristol.ac.uk; Dave.Bull@bristol.ac.uk; Nishan.Canagarajah@bristol.ac.uk).

A 2-D representation of the image patches can be simplified to a 1-D 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 the be expressed as follows:

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

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

$$\underline{p}(t) = B\underline{v}(t) \quad (4)$$
$$\underline{v}(t) = B^{-1}\underline{p}(t) = A\underline{p}(t). \quad (5)$$

Thus, $B = [\underline{b}_1 \underline{b}_2 \ldots \underline{b}_M]$ represents an unknown mixing matrix (analysis kernel) and $A = [\underline{a}_1 \underline{a}_2 \ldots \underline{a}_M]^T$ the unmixing matrix (synthesis kernel). This transform projects the observed signal $\underline{p}(t)$ 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 [9]. 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 optimization of the negentropy. The following rule defines a FastICA approach that optimizes negentropy, as proposed in [8]:

$$\underline{a}_i^+ \leftarrow \varepsilon\left\{\underline{a}_i \phi\left(\underline{a}_i^T \underline{z}\right)\right\} - \varepsilon\left\{\phi'(\underline{a}_i^T \underline{z})\right\} \underline{a}_i \quad 1 \le i \le K \quad (7)$$
$$A \leftarrow A(A^T A)^{-0.5} \quad (8)$$

where $\phi(x) = -\partial G(x)/\partial x$ defines the statistical properties $G(x) = \log p(x)$ of the signals in the transform domain [8]. In our implementation, we used

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

where α and β are constants and ζ is a small constant to prevent numerical instability, in the case that $x \rightarrow 0$ [8]. 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 $v_f(t)$ is constructed in the ICA domain, we can move back to the spatial domain, using the synthesis kernel A, and synthesize the image $i_f(x, y)$.

III. PROPOSED FUSION METHOD USING ICA

A. Separated Training Sets

In the proposed method, training images are separated in two groups prior to the training process, IR and visible. The training uses a set of IR images and a set of visible images, with content comparable to the test set. A number of rectangular patches used for training was randomly selected from the training images. The introduction of separate training subsets provides us with two sets of ICA bases. The first set is used to decompose the IR input image patches $v_i(t) = A_i p_i(t)$ and the second subset to transform the visible input image patches to ICA domain $v_v(t) = A_v p_v(t)$.

Separate ICA basis sets for decomposition of the input images are more specifically trained to capture statistical properties of the specific modality of the input (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, and which results in an "average" ICA basis set that is not able to take the full advantage of ICA decomposition. It is important to note that before the reconstruction of the fused image in the pixel domain, it is necessary to normalize the energy of the two ICA bases subsets. The normalization provides the necessary amplitude balance between the ICA coefficients.

B. Region-Based Fusion of ICA Coefficients

Several features can be employed in the estimation of the contribution of each input image to the fused output image. For example, the authors in [9] use the mean absolute value of each ICA coefficient as an activity indicator in each input image

$$E_i(t) = |\underline{v}_i(t)|, \quad E_v(t) = |\underline{v}_v(t)|. \quad (10)$$

As the ICA bases tend to focus on the edge information, large values for $E_k(t)$ $(k \in \{i, v\})$ correspond to increased activity in the patch, e.g., the existence of edges or a specific texture. Based on this observation, the standard ICA image fusion method divides the ICA domain coefficients in two groups [9]. The first group consists of the regions that contain details ($E_k(t)$ larger then a threshold) and the second group contains the region with background information ($E_k(t)$ smaller then a threshold). The threshold that determines whether a region is "active" or "nonactive" is set heuristically. As a result, the segmentation map $s_i(t)$ is created for the IR input image

$$s_i(t) = \begin{cases} 1, & \text{if } E_i(t) > 2 \cdot \text{mean}_t\{E_i(t)\} \\ 0, & \text{otherwise} \end{cases} \quad (11)$$

as well as for the visible input image $s_v(t)$. The segmentation maps of input images are combined to form a single segmentation map, using the logical OR operator

$$s(t) = \text{OR}\{s_i(t), s_v(t)\}. \quad (12)$$

After the input images are segmented into active and nonactive regions, different fusion rules can be used for fusion of each group of regions. However, the threshold in (11) that determines the "activity" of a region is set heuristically. Thus, the regions obtained by thresholding of the ICA coefficients do not always correspond to objects in the images to be fused. Our experiments showed that important objects in the IR input images
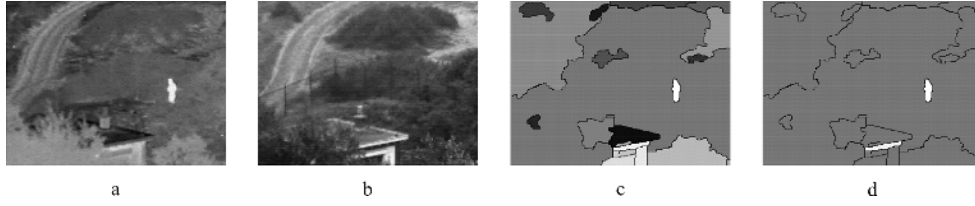
Fig. 1. Segmentation and region selection prior to fusion. (a) IR input image. (b) Visible input image. (c) Regions obtained by joint segmentation of the input images. (d) The image mask: white from IR, gray from visible.

(e.g., a person or a smaller object) are often masked by textured high-energy background in the visual image. In this case, the important objects from the IR image become blurred or, in extreme cases, completely masked. Therefore, we perform segmentation in the spatial domain and then fuse patches from separate regions separately. This differs from the methods in [9] and [10] where the fusion was performed on a more general, pixel level.

C. The Segmentation Algorithm

The quality of the segmentation algorithm is of vital importance to the fusion process. An adapted version of the combined morphological-spectral unsupervised image segmentation algorithm is used, which is described in [11], enabling it to handle multimodal images. The algorithm works in two stages. The first stage produces an initial segmentation by using both textured and non-textured regions. The detail coefficients of the dual-tree complex wavelet transform (DT-CWT) are used to process texture. The gradient function is applied to all levels and orientations of the DT-CWT coefficients and up-sampled to be combined with the gradient of the intensity information to give a perceptual gradient. The larger gradients indicate possible edge locations. The watershed transform of the perceptual gradient gives an initial segmentation. The second stage uses these primitive regions to produce a graph representation of the image which is processed using a spectral clustering technique.

The method can use either intensity information or textural information or both to obtain the segmentation map. This flexibility is useful for multimodal fusion where some *a priori* information of the sensor types is known. For example, IR images tend to lack textural information with most features having a similar intensity value throughout the region. Therefore, we have used an intensity only segmentation map, as it gives better results than a texture-based segmentation.

The segmentation can be performed either separately or jointly. For separate segmentation, each of the input images generates an independent segmentation map for each image

$$S_1 = \sigma(i_1, D_1), \ldots, S_N = \sigma(i_N, D_N) \qquad (13)$$

where D_n represent detail coefficients of the DT-CWT used in segmentation. Alternatively, information from all images could be used to produce a joint segmentation map

$$S_{\text{joint}} = \sigma(i_1 \cdots i_N, D_1 \cdots D_N). \qquad (14)$$

In general, jointly segmented images work better for fusion [12]. This is because the segmentation map will contain a minimum number of regions to represent all the features in the scene most efficiently. A problem can occur for separately segmented im-

ages, where different images have different features or features which appear as slightly different sizes in different modalities. Where regions partially overlap, if the overlapped region is incorrectly dealt with, artefacts will be introduced and the extra regions created to deal with the overlap will increase the time taken to fuse the images.

D. Calculation of Priority and Fusion Rules

After the images are jointly segmented, it is essential to determine the importance of regions in each of the input images. We have decided to use the normalized Shannon entropy of a region as the priority. Thus, the priority $P(r_{t_n})$ is given as

$$P(r_{t_n}) = \frac{1}{|r_{t_n}|} \sum_{\forall \theta, \forall l, (x,y) \in r_{t_n}} d_{n(\theta,l)}^2(x,y) \log d_{n(\theta,l)}^2(x,y) \qquad (15)$$

with the convention $0 \log(0) = 0$, where $|r_{t_n}|$ is the size of the region r_{t_n} in input image n and $d_{n(\theta,l)}(x,y) \in D_{n(\theta,l)}$ are detail coefficients of the DT-CWT used in segmentation. Finally, a mask M is generated that determines which image each region should come from in the fused image. An example of the IR input image, visual input image, performed joint segmentation and the image fusion mask is given in Fig. 1.

E. Reconstruction of the Fused Image Using Fusion Metrics

In addition, we implement a novel method for reconstruction of the fused image, using statistical properties of the both input images. In the standard ICA method, reconstruction of the fused image is performed on the patch-per-patch base [9]

$$p_f(t) = U(t) + \frac{1}{2}(M_i(t) + M_v(t)) \qquad (16)$$

where $p_f(t)$ represents the tth patch of the fused image $i_f(x,y)$ and $U(t)$ is the tth 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

$$p_f(t) = U(t) + M_i(t) \cdot w_i + M_v(t) \cdot w_v. \qquad (17)$$

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. One of the fusion performance metrics [13], [14] is calculated at each step. We decided to exploy
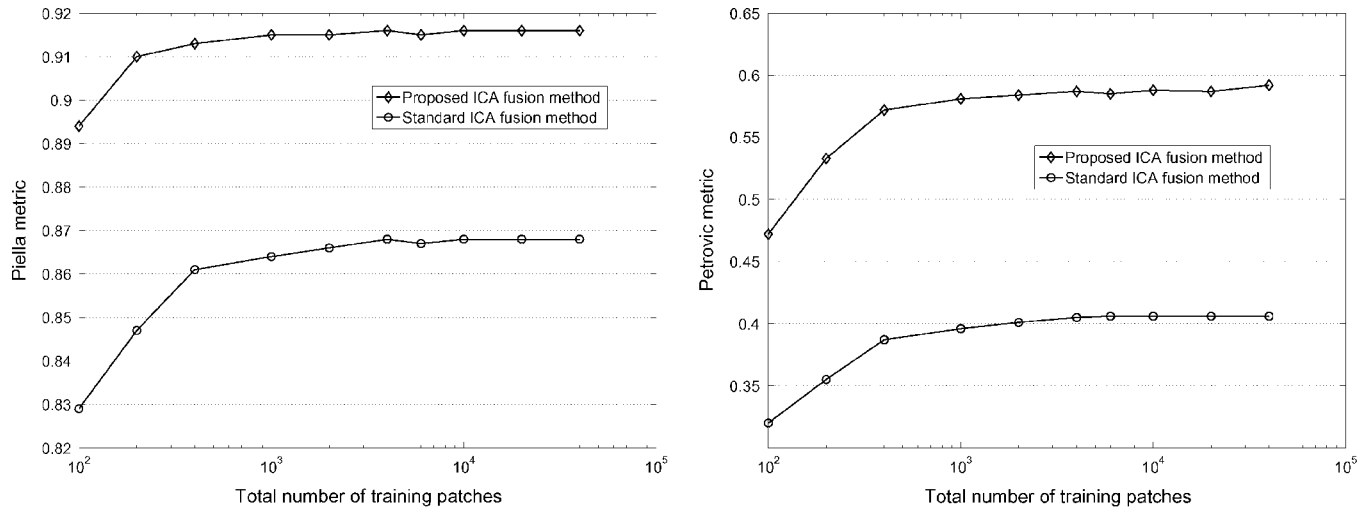
Fig. 2. Fusion performance for "UN Camp" image sequence. Comparison of fusion performance versus number of training patches for the proposed and standard ICA fusion method. (left) Piella metric. (right) Petrovic metric.

the Piella metric [13] and Petrovic metric [14] because these are the most widespread tools for evaluation of image fusion algorithms. 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. EXPERIMENTAL RESULTS

The proposed image fusion method was tested in different surveillance scenarios with two modalities: IR and visible. In order to make a comparison between the proposed method and the standard ICA method, the images were fused using the approach described in [9]. We compared these results with a simple averaging method, the ratio method [15], the Laplace transform (LT) [16], and the DT-CWT [12]. Before performing image fusion, the ICA bases were trained using a set of five IR images and five visible images, with content comparable to the test set. A number of rectangular patches $(N = 8)$ used for training was randomly selected from the training set. Lexicographic ordering was applied to the image patches and then PCA performed. Following this, a number of the most important bases were selected, according to the eigenvalues corresponding to these bases. After that, the ICA update rule in (7) was iterated until convergence. ICA coefficients were obtained using the principle described in Section II, while reconstruction of the fused image was performed using optimization based on the Piella fusion performance metric [13].

A. Comparison With the Standard ICA Image Fusion Method

Initially, experiments were focused on evaluating the performance of the proposed algorithm and comparison with the standard ICA fusion algorithm. In order to compare how the fusion algorithm's performance depends on the training process, the impact of the number of training patches taken from the training set was tested. The number of images in the training set was fixed at ten for the standard ICA fusion method and at five images per subset of training images in the IR and visible domain, for the proposed method. The number of training patches

taken from the training set (subsets) was then varied from 100 to 40 000 in order to evaluate both algorithms' performance with different number of training patches. Size of training patches was 8×8 $(N = 8)$ and 32 of the most significant bases obtained by training are selected using the PCA algorithm.

The results in Figs. 2 and 4 show that the proposed algorithm significantly outperforms the standard ICA fusion algorithm for the UN Camp and Octec surveillance image sequences, with constantly higher scores in terms of both Piella and Petrovic metrics. Figs. 3 and 5 depict examples of fused images for the standard and proposed ICA fusion algorithm, where different numbers of training patches are used. Visual (subjective) comparison between methods indicates that our method is far superior to the basic ICA method: for example, it is clear that the fence detail from the visual image is far better transferred into the fused image in the proposed method. In addition, the details of the tree in the visual image are visually more pleasing and the human figure is much more bright in the proposed method than in the fused image obtained by the standard ICA method. It is also noticeable that the performance of the proposed method is less dependent on the number of training patches then the standard ICA fusion method. In addition, the proposed method trained by only 200 training patches outperforms the standard ICA method trained by 40 000 training patches, measured by both fusion metrics. Therefore, the proposed algorithm needs a significantly shorter training process in order to obtain fusion performance comparable to, or above, the performance of the standard ICA method.

B. Comparison With the State-of-the-Art Image Fusion Methods

The proposed image fusion method was tested against several state-of-the-art image fusion methods in two modalities: IR and visible. The images used in experiments are surveillance images from TNO Human Factors and Octec Ltd., publicly available at the Image Fusion web site [17]. Image sequence "UN Camp" consists of 32 images (32 visual and 32 IR images), image sequence "Octec" has 25 images and image sequences "Trees" and "Dune" contain 19 images. In order to make a comparison between the proposed method and the standard ICA method, the
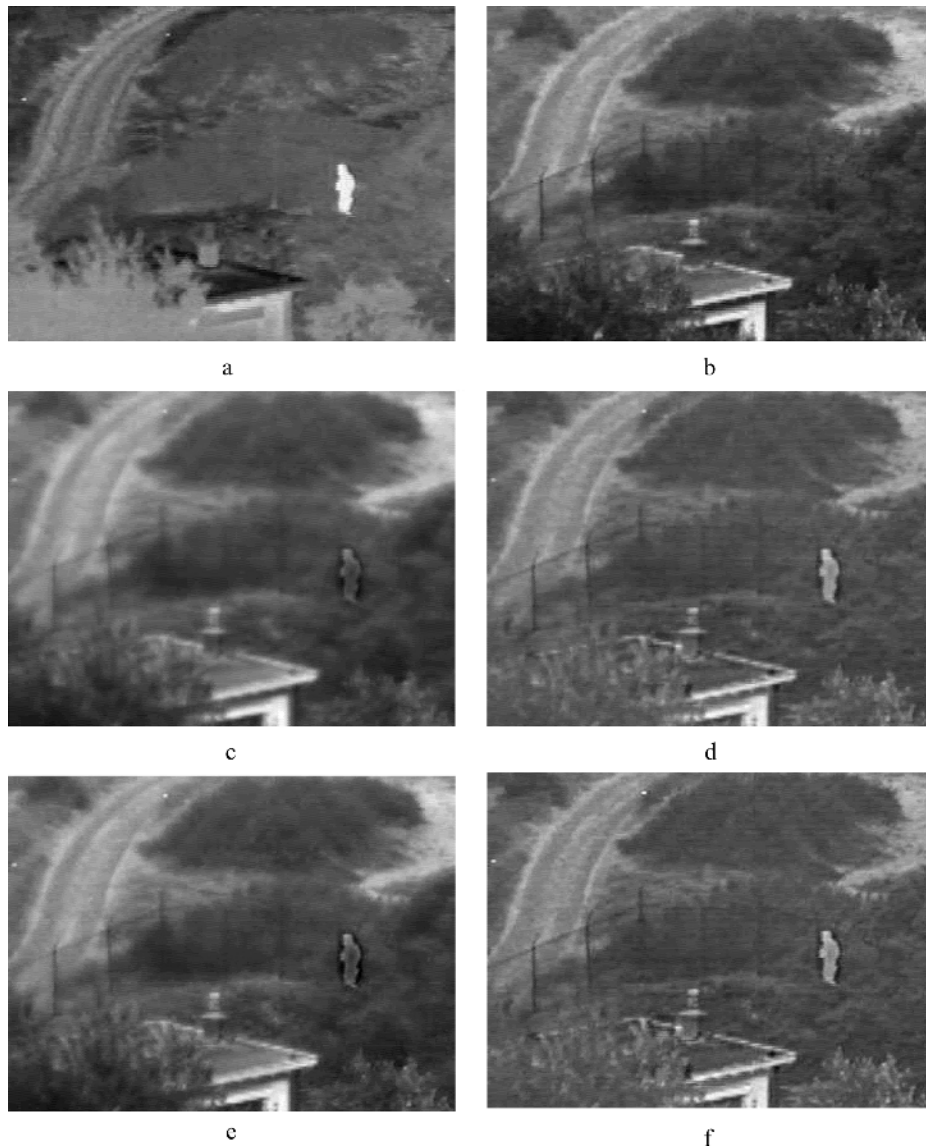
Fig. 3.   Subjective fusion results. (a) Input IR image. (b) Input visible image. (c) Fused image using standard ICA fusion and 100 training patches. (d) Fused image using proposed method and 100 training patches. (e) Standard ICA fusion and 10 000 training patches. (f) Proposed method and 10 000 training patches.
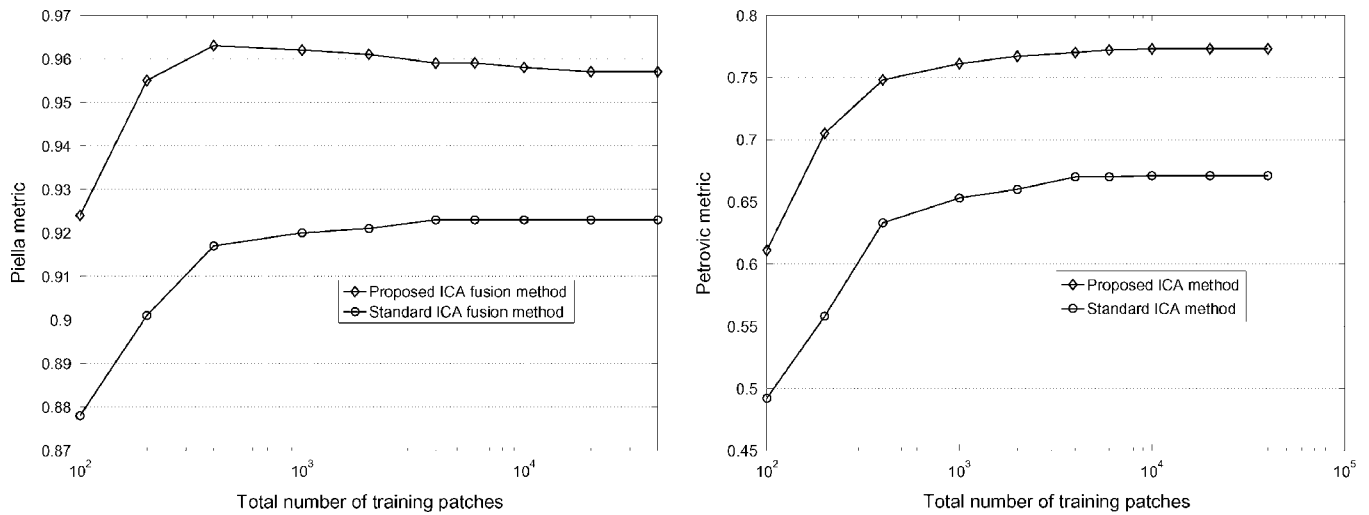

Fig. 4.   Fusion performance for "Octec" image sequence. Comparison of fusion performance versus number of training patches for the proposed and standard ICA fusion method. (left) Piella metric. (right) Petrovic metric.
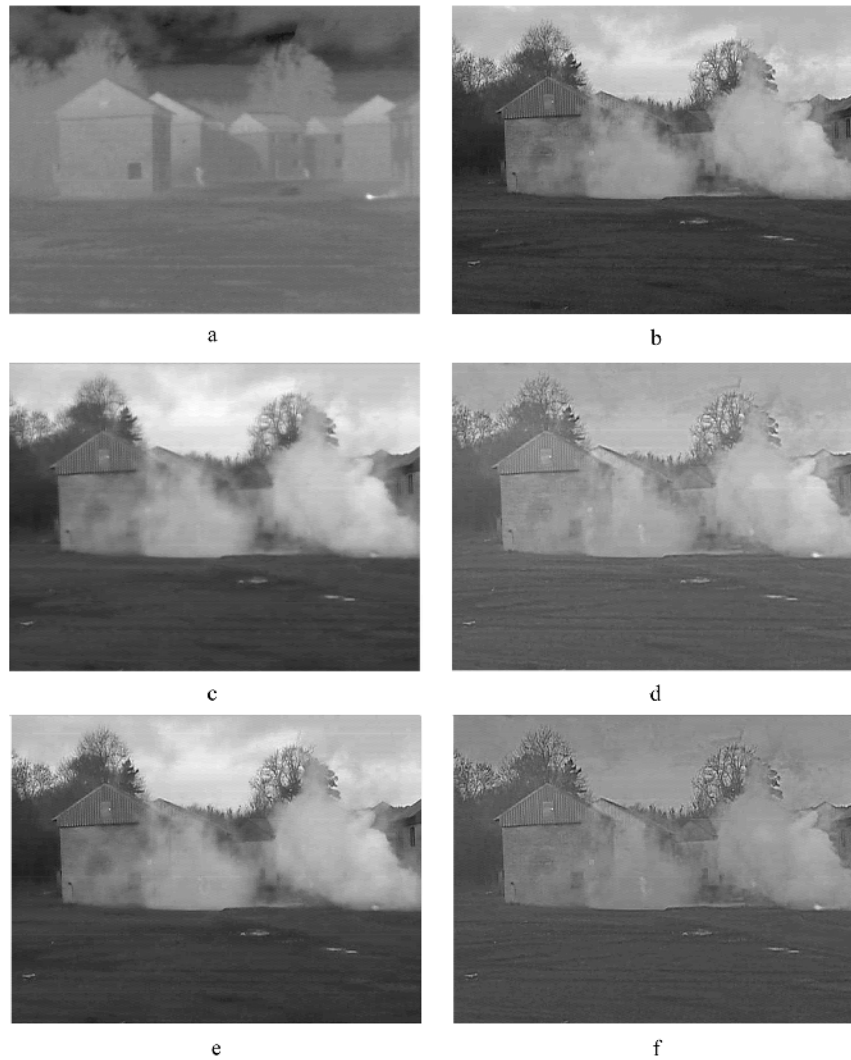
Fig. 5.   Subjective fusion results. (a) Input IR image. (b) Input visible image. (c) Fused image using standard ICA fusion and 100 training patches. (d) Fused image using proposed method and 100 training patches. (e) Standard ICA fusion and 10 000 training patches. (f) Proposed method and 10 000 training patches.

images were fused using the approach described in [9]. We compared these results with a simple averaging method, the ratio method [15], the LT [16] and the DT-CWT [12]. In the multiresolution methods (LT, DT-CWT), a five-level decomposition is used and fusion is performed by selecting the coefficient with a maximum absolute value, except for the case of the lowest resolution subband where the mean value is used.

The images fused using these algorithms are given in Figs. 5 and 6, together with IR and visible input images. The proposed and standard ICA method were trained using 10 000 training patches taken from a set of images with similar content. Size of training patches was $8 \times 8 (N = 8)$ and 32 of the most significant bases obtained by training are selected using the PCA algorithm. It should be noted that the adaptive fused image reconstruction adds 1%–2% of computational overhead to the standard, nonadaptive ICA fusion algorithm.

Visual (subjective) comparison between methods indicates that our method is far superior to the basic ICA method, but also that the proposed weighted ICA method outperforms the LT and DT-CWT methods: for example, in Fig. 6, it is clear that the fence detail from the visual image is far better transferred into the fused image in the proposed method than in the stan-

dard ICA method. In addition, the details of the tree in the visual image are visually more pleasing in the proposed method than in the DT-CWT approach, although the person is brighter in the DT-CWT fused image. In Fig. 7, it is obvious that the proposed method outperforms standard ICA as the landscape structure is better represented in the fused image and the terrain information is more clear in the proposed ICA method compared with the DT-CWT and LT methods.

The results in Table I show the proposed algorithm significantly outperforms the standard ICA fusion algorithm for the tested surveillance image sequences, with constantly higher scores in terms of fusion metrics. The proposed method also exhibits higher performance than the multiresolution methods in most cases, except for the case of the Dune sequence, where the best results are obtained by the DT-CWT fusion method. The advantage in terms of metric values is accentuated for the Petrovic metric, because of its higher dynamics (smaller differences in the fused image are discriminated with larger difference in the metric grade). The metrics' values confirm the subjective assessment that the images obtained using the proposed algorithm generally incorporate more information from the visible image together with the important details from the IR image.
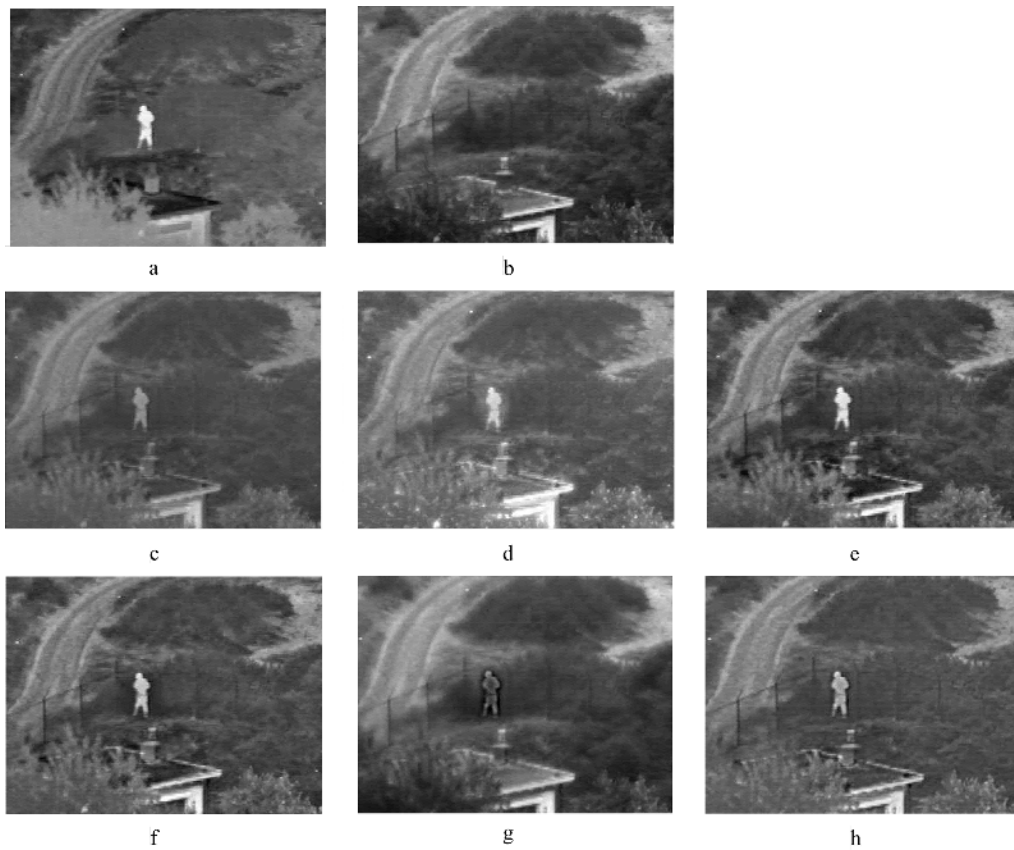
Fig. 6.   Subjective fusion results. (a) Input IR image. (b) Input visible image. (c) Fused image using averaging. (d) Fused image using ratio pyramid. (e) Fused image using Laplace pyramid. (f) Fused image using DT-CWT. (g) Fused image using standard ICA method. (h) Fused image using the proposed ICA method.
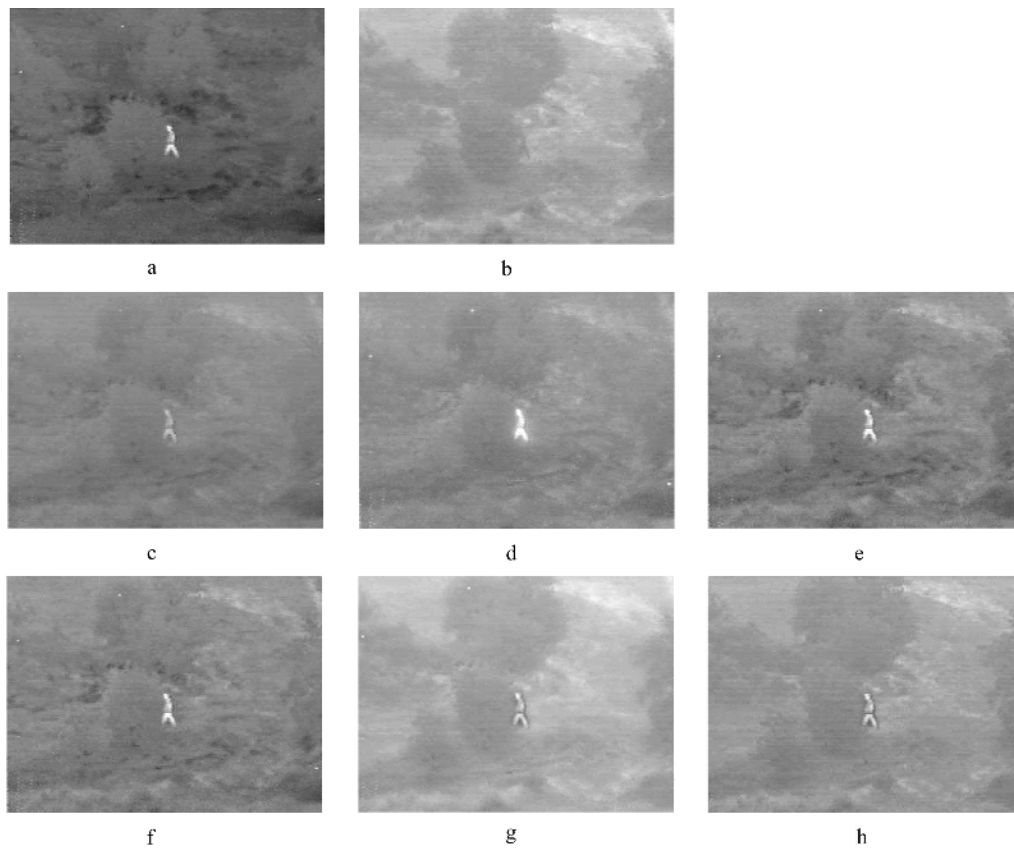

Fig. 7.   Subjective fusion results. (a) Input IR image. (b) Input visible image. (c) Fused image using averaging. (d) Fused image using ratio pyramid. (e) Fused image using Laplace pyramid. (f) Fused image using DT-CWT. (g) Fused image using standard ICA method. (h) Fused image using the proposed ICA method.

TABLE I
PERFORMANCE OF IMAGE FUSION METHODS MEASURED BY THE
STANDARD FUSION METRICS

Metric	Method	UnCamp	Dune	Octec	Trees
Piella	Average	0.85	0.95	0.87	0.91
	Ratio	0.86	0.96	0.87	0.91
	DT-CWT	0.91	**0.97**	0.94	0.92
	Laplace	0.91	0.96	0.93	0.92
	Standard ICA	0.86	0.95	0.92	0.90
	Proposed ICA	**0.93**	0.96	**0.95**	**0.93**
Petrovic	Average	0.33	0.51	0.43	0.43
	Ratio	0.41	0.53	0.50	0.47
	DT-CWT	0.46	0.60	0.76	0.55
	Laplace	0.50	0.60	0.76	0.55
	Standard ICA	0.40	0.52	0.67	0.44
	Proposed ICA	**0.60**	**0.66**	**0.79**	**0.61**

V. CONCLUSION

In this paper, we have described an improved image fusion algorithm based on ICA. In the proposed method, images used for training ICA bases are separated in two groups prior to the training process, one consisting of IR images and the second consisting of visible images. Region-based fusion of ICA coefficients is implemented, where segmentation is performed in the spatial domain and ICA coefficients from separate regions fused separately. Weighting of the ICA bases during reconstruction of the fused image by using the fusion metrics is used to maximize the performance of the proposed method.

Experimental results confirm that the proposed method exhibits significantly better fusion than basic ICA method, achieving higher scores using both Piella and Petrovic metrics. The proposed method also outperforms the state-of-the-art algorithms, both in terms of subjective quality and fusion metric values. Finally, it is important to note that the proposed ICA-based algorithm clearly offers improved performance in scenarios where contextual information is available. However, the increased performance comes at a cost of increased computational complexity of the fusion process.

REFERENCES

[1] J. A. Ratches, R. H. Vollmerhausen, and R. G. Driggers, "Target acquisition performance modeling of infrared imaging systems: Past, present, and future," *IEEE Sensors J.*, vol. 1, no. 1, pp. 31–40, Jun. 2001.

[2] D. Bank, "A novel ultrasonic sensing system for autonomous mobile systems," *IEEE Sensors J.*, vol. 2, no. 6, pp. 597–606, Dec. 2002.

[3] S. Suranthiran and S. Jayasuriya, "Optimal fusion of multiple nonlinear sensor data," *IEEE Sensors J.*, vol. 4, no. 5, pp. 651–663, Oct. 2004.

[4] M.-C. Lu, W.-Y. Wang, and C.-Y. Chu, "Image-based distance and area measuring systems," *IEEE Sensors J.*, vol. 6, no. 2, pp. 495–503, Apr. 2006.

[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] S. Nikolov, D. Bull, and N. Canagarajah, "Wavelets for image fusion," in *Wavelets in Signal and Image Analysis*. Norwell, MA: Kluwer, 2001.

[8] A. Hyvrinen, J. Karhunen, and E. Oja, *Independent Component Analysis*. London, U.K.: Wiley, 2001.

[9] N. Mitianoudis and T. Stathaki, "Pixel-based and region-based image fusion using ICA bases," *Information Fusion*, vol. 8, no. 2, pp. 131–142, Apr. 2007.

[10] N. Cvejic, D. Bull, and N. Canagarajah, "A novel ICA domain multimodal image fusion algorithm," in *Proc. SPIE Defense and Security Symp.*, Orlando, FL, 2007.

[11] R. O'Callaghan and D. Bull, "Combined unsupervised image segmentation," *IEEE Trans. Image Process.*, vol. 14, pp. 49–62, 2005.

[12] J. Lewis, R. O'Callaghan, S. Nikolov, D. Bull, and N. Canagarajah, "Pixel- and region-based image fusion with complex wavelets," *Information Fusion*, vol. 8, no. 2, pp. 119–130, Apr. 2007.

[13] G. Piella and H. Heijmans, "A new quality metric for image fusion," in *Proc. IEEE Int. Conf. Image Processing*, Barcelona, Spain, 2003, pp. 173–176.

[14] V. Petrovic and C. Xydeas, "Objective evaluation of signal-level image fusion performance," *Opt. Eng.*, vol. 44, no. 8, p. 087003, 2005.

[15] A. Toet, "Image fusion by a ratio of low-pass pyramid," *Pattern Recogn. Lett.*, vol. 9, pp. 245–253, 1996.

[16] P. Burt and E. Adelson, "Laplacian pyramid as a compact image code," *IEEE Trans. Commun.*, vol. 31, no. 4, pp. 115–123, 1983.

[17] [Online]. Available: http://imagefusion.org

Nedeljko Cvejic received the Dipl.-Ing. degree in electrical engineering from the University of Belgrade, Belgrade, Serbia, in 2000 and the Dr. Tech. degree from the University of Oulu, Oulu, Finland, in 2004.

From 2001 to 2004, he was a Research Scientist at the Department of Electrical and Information Engineering, University of Oulu. He is currently a Research Associate with the Department of Electrical and Electronic Engineering, University of Bristol, Bristol, U.K. He has published more than 40 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, University of Bristol, Bristol, U.K. 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 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., specializing in wireless multimedia communications. He has worked widely in the fields of 1-D and 2-D signal processing and has published over 250 papers, various articles, and 2 books. He is widely supported in these areas by both industry, Europe, MoD, and the Engineering and Physical Sciences Research Council (EPSRC). His current research is focused on the problems of image and video communications for low bit rate wireless, Internet and broadcast applications.

Dr. Bull has won two IEE Premium Awards for this work. From 2003 to 2005, he was also a member of the Science and Technology Board of the U.K. Defence Technology Centre in Data and Information Fusion. He has been a member of the U.K. Foresight Panel and the Steering Group for the DTI/EPSRC LINK program in Broadcast Technology.

Nishan Canagarajah received the B.A. (Hons) degree and the Ph.D. degree in DSP techniques for speech enhancement from the University of Cambridge, Cambridge, U.K..

He is currently a Professor of Multimedia Signal Processing at the University of Bristol, Bristol, U.K. Prior to this, he was an RA and Lecturer at the University of Bristol investigating DSP aspects of mobile radio receivers. His research interests include image and video coding, image segmentation, content-based video retrieval, 3-D video, and image fusion. He is widely supported in these areas by industry, EU, and the Engineering and Physical Sciences Research Council (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.

Dr. Canagarajah is a member of the Executive Team of the IEE PN on Multimedia Communications and a member of EPSRC Peer Review College.