

Image fusion based on pixel significance using cross bilateral filter

B. K. Shreyamsha Kumar

Received: 19 April 2012 / Revised: 5 September 2013 / Accepted: 5 September 2013 / Published online: 5 October 2013
© Springer-Verlag London 2013

Abstract Like bilateral filter (BF), cross bilateral filter (CBF) considers both gray-level similarities and geometric closeness of the neighboring pixels without smoothing edges, but it uses one image for finding the kernel and other to filter, and vice versa. In this paper, it is proposed to fuse source images by weighted average using the weights computed from the detail images that are extracted from the source images using CBF. The performance of the proposed method has been verified on several pairs of multisensor and multifocus images and compared with the existing methods visually and quantitatively. It is found that, none of the methods have shown consistence performance for all the performance metrics. But as compared to them, the proposed method has shown good performance in most of the cases. Further, the visual quality of the fused image by the proposed method is superior to other methods.

Keywords Image fusion · Pixel significance · Multifocus · Multisensor · Bilateral filter · Cross bilateral filter · Detail image · Detail strength

Abbreviations

IR Infrared
WT Wavelet transform
DWT Discrete wavelet transform

BF Bilateral filter
CBF Cross bilateral filter
API Average pixel intensity
SD Standard deviation
AG Average gradient
MI Mutual information
FS Fusion symmetry
CC Correlation coefficient
SF Spatial frequency

1 Introduction

Image fusion has become a widely used tool to increase the visual interpretation of the image in various applications such as getting an ‘all-in-focus’ image from a given set of multifocus images, medical diagnosis, surveillance, military, machine vision, robotics, enhanced vision system, biometrics, and remote sensing. The main objective of any image fusion is that, to conglomerate all the significant visual information from multiple input images by retaining the more comprehensive, accurate and stable information than the individual source images without introducing any artifacts. This makes the human/machine perception or further processing easy [1].

Generally, one image of a complex scene does not contain enough information because of limitations in the system. Like, it is difficult to get all the objects in focus in a single image due to limited depth of focus by optical lens of a CCD camera. But, a series of images obtained by progressively shifting the focal plane through the scenery can be fused with a best fusion rule to produce an image with a quasi-infinite depth of field. This gives rise to the problem of multifocus image fusion. Similarly, the images obtained by CCD camera give information only in visible spectrum

B. K. Shreyamsha Kumar (✉)
Central Research Laboratory, Bharat Electronics,
Bangalore 560013, India
e-mail: shreyamsha@yahoo.com

B. K. Shreyamsha Kumar
Department of Electrical and Computer Engineering,
Concordia University, 1515 St. Catherine West, Montreal,
Quebec H3G 2W1, Canada

whereas Infrared (IR) camera in IR spectrum, and hence, the multispectral data from different sensors often present complementary information about the region surveyed, scene or object. In such scenarios, image fusion provides an effective method to enable comparison, interpretation, and analysis of such data, as the fused image facilitates improved detection and unambiguous localization of a target (represented in IR image) with respect to its background (represented in the visible image). Hence, the fusion of IR and visual images is gaining momentum in surveillance applications. A suitably fused representation of IR and visible images provides a human operator a more complete and accurate mental representation of the perceived scene, which results in a larger degree of situational awareness. Likewise in medical imaging, the MRI image shows brain tissue anatomy, whereas CT scan image provides details about bony structures. The integration of these medical images of different modalities into one image with the merits of both source images provides both anatomical and functional information, which is important for planning surgical procedure. The aim is to achieve better situation assessment and/or more rapid and accurate completion of a predefined task than would be possible using any of the sensors individually. In the literature, it has been defined as the synergistic combination of different sources of sensory information into a single representational format [1].

The objective of the paper is to improve the fusion performance by combining all the important visual information contained in the individual source images using the weights computed from the detail images of the cross bilateral filter (CBF). The paper is organized as follows: Sect. 2 deals with literature survey, Sect. 3 describes the proposed method, Sect. 4 discusses experimental results followed by conclusions and future work in Sect. 5.

2 Literature survey

Analogous to other forms of information fusion, image fusion is usually performed at any of the three processing levels: signal, feature and decision. Signal-level image fusion, also known as pixel-level image fusion, defines the process of fusing visual information associated with each pixel from a number of registered images into a single fused image, representing a fusion at lowest level. As the pixel-level fusion is part of the much broader subject of multifocus and multisensor information fusion, it has attracted many researchers in the last two decades [2–5]. Object-level image fusion, also called as feature-level image fusion, fuses feature, object labels and property descriptor information that have already been extracted from individual input images [6]. Finally, decision- or symbol-level image fusion, being the highest level of image fusion, represents fusion of probabilistic decision information obtained by local decision makers

using the results of feature-level processing on the individual images [7].

In the last two decades, a lot of research has been carried out in the area of multifocus and multispectral image fusion. Multispectral image fusion based on intensity-hue-saturation method is described in Carper et al. [8] and that based on Laplacian pyramid mergers in Toet [9]. The multifocus image fusion proposed in Haeberli [10] uses the fact that the focused area of the image will have highest intensity compared to that of unfocused areas. Further, the energy compaction and multiresolution properties of a wavelet transform (WT) were exploited by [11,12] for image fusion. In Qu et al. [11], the medical images are fused based on the WT modulus maxima of input images at different bandwidths and levels. Instead of convolution-based wavelet decomposition, Lifting-based wavelet decomposition is proposed in Ranjith and Ramesh [13] to reduce the computational complexity with less memory requirements. Further, induction of complex wavelets for image fusion has improved the fusion performance [14]. Fusion of satellite images using multiscale wavelet analysis is proposed in Du et al. [15] and various image fusion methods available in the literature are compared in Wang et al. [16]. Multisource image fusion using a sequence of support value images and a low frequency image has been described in Zheng et al. [17]. Also, pixel-level multifocus image fusion has been proposed based on image blocks and artificial neural networks [18]. Further, the comparative study of several pixel-level multispectral palm image fusion approaches for biometric applications is conducted [19]. Image fusion by averaging approximation subbands of the source images blurs the fused image and is reduced by combining approximation subbands based on the edge information present in the corresponding detail subbands [20]. Here, mean and standard deviation over 3×3 windows are used as activity measurement to find the edge information present in detail subbands. Further, a pixel-level image fusion has been proposed by decomposing the source images using wavelet, wavelet packet, curvelet [21] and contourlet transform [4,22]. Due to nonideal characteristics of the imaging systems, the source images will be noisy, and hence, fusion of these images requires a hybrid algorithm which addresses both image denoising and fusion. This type of scenario has been addressed in contrast-based fusion of noisy images using discrete wavelet transform (DWT) [23], which considers contrast of images using local variance of the denoised DWT coefficients as well as the noise strength. Recently, discrete cosine transform-based image fusion has been proposed for image fusion [24] instead of pyramids or wavelets, and its performance is comparable to both convolution and lifting-based wavelets. Further, the superior/similar fusion performance is achieved with discrete cosine harmonic wavelet [25] with reduced computational complexity than that in Naidu [24]. Also, a new multiresolution DCT

decomposition, by viewing DCT as a tree structure, has been proposed for multifocus image fusion to reduce the computational complexity without any degradation in the fusion performance [26].

The bilateral filter (BF) introduced by Tomasi and Manduchi [27] has many applications in image denoising [28–30], flash photography enhancement [31,32], image/video fusion [33,34], etc. A variant of BF, Joint/Cross BF, which uses a second image to shape the filter kernel and operate on the first image, and vice versa was proposed in [31,32]. Both of these papers address the problem of combining the details of images captured with and without flash under ambient illumination. In Fattal et al. [35], BF has been used for multiscale decomposition of multilight image collections for shape and detail enhancement. Temporal joint BF and dual BF was proposed in Bennett et al. [34] for multispectral video fusion, where former uses IR video to find filter coefficients and filters the visible video, and latter uses both IR and visible video to compute the filter coefficients and filters the visible video. Further, the application of BF for hyperspectral image fusion has been proposed in Kotwal and Chaudhuri [36], which separates the weak edges and fine textures by subtracting the BF output from the original image. The magnitude of this difference image is used to find the weights directly. One more variant of BF, which uses center pixel from IR image and the neighboring pixels from visual image to find the filter kernel and operates on IR image, has been proposed for human detection by fusing IR and visible images [37]. Further, multiscale directional BF, a combination of BF and directional filter bank, has been proposed for multisensor image fusion to exploit the edge preserving capability of BF and directional information capturing capability of directional filter bank [33].

3 Proposed method

The proposed image fusion algorithm directly fuses two source images of a same scene using weighted average. The proposed method differs from other weighted average methods in terms of weight computation and the domain of weighted average. Here, the weights are computed by measuring the strength of details in a detail image obtained by subtracting CBF output from original image. The weights thus computed are multiplied directly with the original source images followed by weight normalization. The block diagram of the proposed scheme is show in Fig. 1 for two source images *A* and *B*.

3.1 Cross bilateral filter (CBF)

Bilateral filtering is a local, nonlinear and noniterative technique which combines a classical low-pass filter with an

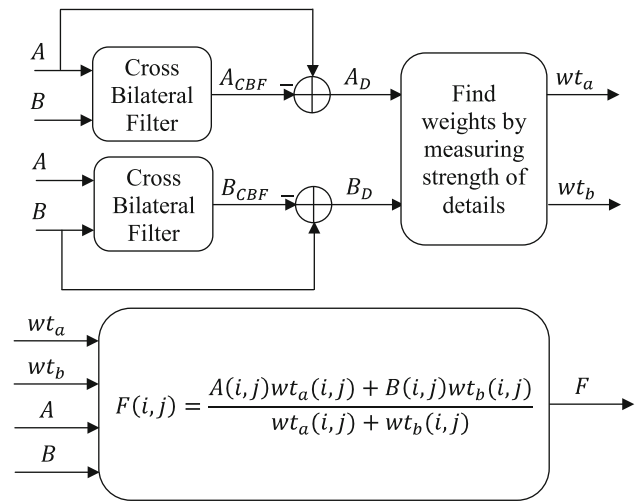


Fig. 1 Proposed image fusion framework

edge-stopping function that attenuates the filter kernel when the intensity difference between pixels is large. As both gray-level similarities and geometric closeness of the neighboring pixels are considered, the weights of the filter depend not only on Euclidian distance but also on the distance in gray/color space. The advantage of the filter is that it smooths the image while preserving edges using neighboring pixels. Mathematically, for an image *A*, the BF output at a pixel location *p* is calculated as follows [27]:

$$A_F(p) = \frac{1}{W} \sum_{q \in S} G_{\sigma_s}(\|p - q\|) \times G_{\sigma_r}(|A(p) - A(q)|) A(q) \tag{1}$$

where $G_{\sigma_s}(\|p - q\|) = e^{-\frac{\|p-q\|^2}{2\sigma_s^2}}$ is a geometric closeness function,

$G_{\sigma_r}(|A(p) - A(q)|) = e^{-\frac{|A(p)-A(q)|^2}{2\sigma_r^2}}$ is a gray-level similarity/ edge-stopping function,

$W = \sum_{q \in S} G_{\sigma_s}(\|p - q\|) G_{\sigma_r}(|A(p) - A(q)|)$ is a normalization constant,

$\|p - q\|$ is the Euclidean distance between *p* and *q*,

and *S* is a spatial neighborhood of *p*.

Since σ_s and σ_r control the behavior of BF, the dependency of σ_r/σ_s values and derivative of the input signal on the behaviors of the BF are analyzed in Zhang and Gunturk [28]. The optimal σ_s value is chosen based on the desired amount of low-pass filtering and blurs more for larger σ_s , as it combines values from more distant image locations [28]. Also, if an image is scaled up or down, σ_s must be adjusted accordingly in order to obtain equivalent results. It appears that a good range for the σ_s value is roughly (1.5–2.1); on the other hand, the optimal σ_r value will depend on the amount

of edge to be preserved. If the image is amplified or attenuated, σ_r must be adjusted accordingly in order to retain the same result.

CBF considers both gray-level similarities and geometric closeness of neighboring pixels in image A to shape the filter kernel and filters the image B . CBF output of image B at a pixel location p is calculated as [31]

$$B_{\text{CBF}}(p) = \frac{1}{W} \sum_{q \in S} G_{\sigma_s}(\|p - q\|) \times G_{\sigma_r}(|A(p) - A(q)|) B(q) \quad (2)$$

where $G_{\sigma_r}(|A(p) - A(q)|) = e^{-\frac{|A(p) - A(q)|^2}{2\sigma_r^2}}$ is a gray-level similarity/ edge-stopping function, and $W = \sum_{q \in S} G_{\sigma_s}(\|p - q\|) G_{\sigma_r}(|A(p) - A(q)|)$ is a normalization constant.

The detail image, obtained by subtracting CBF output from the respective original image, for image A and B is given by $A_D = A - A_{\text{CBF}}$ and $B_D = B - B_{\text{CBF}}$, respectively. In multifocus images, unfocused area in image A will be focused in image B and the application of CBF on image B will blur the focused area more compared to that of unfocused area in image B . This is because the unfocused area in image A anyway looks blurred with almost similar gray values in that area thereby making the filter kernel close to Gaussian. Now, the idea is to capture most of the focused area details in detail image B_D such that these details can be used to find the weights for image fusion using weighted average. Similarly, in multisensor images, the information in image B is absent in image A and the application of CBF on image B will blur the information in image B . This is because, as the information in A is absent, the gray levels in that region have similar values thereby making the kernel as Gaussian. Figure 2 shows the simulated multifocus lady source images along with the respective CBF output and detail images. From Fig. 2c and d, it is observed that CBF has blurred the focused area keeping the unfocused area as it is and the details in the focused area has been captured in the detail images (Fig. 2e and f). Now, these detail images are used to find the weights by measuring the strength of details.

3.2 Pixel-based fusion rule [38]

Fusion rule proposed in Shah et al. [38] is discussed here for completeness to compare the performance of proposed method. Here, the weights are computed using statistical properties of a neighborhood of detail coefficient instead of wavelet coefficient as in Shah et al. [38]. A window of size $w \times w$ around a detail coefficient $A_D(i, j)$ or $B_D(i, j)$ is considered as a neighborhood to compute its weight. This neighborhood is denoted as matrix X . Each row of X is treated as an observation and column as a variable to compute

unbiased estimate $C_h^{i,j}$ of its covariance matrix [39], where i and j are the spatial coordinates of the detail coefficient $A_D(i, j)$ or $B_D(i, j)$.

$$\text{covariance}(X) = E[(X - E[X])(X - E[X])^T] \quad (3)$$

$$C_h^{i,j} = \frac{\sum_{k=1}^w (x_k - \bar{x})(x_k - \bar{x})^T}{(w - 1)} \quad (4)$$

where x_k is the k th observation of the w -dimensional variable and \bar{x} is the mean of observations. It is observed that diagonal of matrix $C_h^{i,j}$ gives a vector of variances for each column of matrix X . Now, the eigenvalues of matrix $C_h^{i,j}$ is computed and the number of eigenvalues depends on size of $C_h^{i,j}$. Sum of these eigenvalues are directly proportional to horizontal detail strength of the neighborhood and are denoted as $HdetailStrength$ [38]. Similarly, an unbiased covariance estimate $C_v^{i,j}$ is computed by treating each column of X as an observation and row as a variable (opposite to that of $C_h^{i,j}$), and the sum of eigenvalues of $C_v^{i,j}$ gives vertical detail strength $VdetailStrength$. That is,

$$HdetailStrength(i, j) = \sum_{k=1}^w \text{eigen}_k \text{ of } C_h^{i,j}$$

$$VdetailStrength(i, j) = \sum_{k=1}^w \text{eigen}_k \text{ of } C_v^{i,j}$$

where eigen_k is the k th eigenvalue of the unbiased estimate of covariance matrix. Now, the weight given to a particular detail coefficient is computed by adding these two respective detail strengths. Therefore, the weight depends only on the strength of the details and not on actual intensity values.

$$wt(i, j) = HdetailStrength(i, j) + VdetailStrength(i, j)$$

After computing the weights for all detail coefficients corresponding to both the registered source images, the weighted average of the source images will result in a fused image.

If wt_a and wt_b are the weights for the detail coefficients A_D and B_D belonging to the respective source images A and B , then the weighted average of both is computed as the fused image using Eq. 5.

$$F(i, j) = \frac{A(i, j) wt_a(i, j) + B(i, j) wt_b(i, j)}{wt_a(i, j) + wt_b(i, j)} \quad (5)$$

3.3 Parameters to evaluate the fusion performance

Evaluation of fusion performance is a challenging task as the ground truth is not available in most of the applications. In literature, various parameters have been proposed and used

Fig. 2 Simulated multifocus lady source images in **a** and **b**, CBF output images in **c** and **d**, and detail images in **e** and **f**, respectively



to evaluate the performance of image fusion. Among them, several classical evaluation parameters reported in literature are considered for exhaustive study. They are [21,38]

1. Average pixel intensity (API) or mean (\bar{F}) measures an index of contrast and is given by $API = \bar{F} = \frac{\sum_{i=1}^m \sum_{j=1}^n f(i,j)}{mn}$, where $f(i, j)$ is pixel intensity at (i, j) and $m \times n$ is the size of the image

2. Standard deviation (SD) is the square root of the variance, which reflects the spread in data and is given by $SD = \sqrt{\frac{\sum_{i=1}^m \sum_{j=1}^n (f(i,j) - \bar{F})^2}{mn}}$

3. Average gradient (AG) measures a degree of clarity and sharpness and is given by

$$AG = \frac{\sum_i \sum_j ((f(i,j) - f(i+1,j))^2 + (f(i,j) - f(i,j+1))^2)^{1/2}}{mn}$$

4. Entropy (H) estimates the amount of information present in the image and is given by $H = - \sum_{k=0}^{255} p_k \log_2(p_k)$,

where p_k is the probability of intensity value k in an 8-bit image

5. Mutual information (MI) quantifies the overall mutual information between source images and fused image, which is given by $MI = MI_{AF} + MI_{BF}$, where $MI_{AF} = \sum_k \sum_l p_{A,F}(k,l) \log_2 \left(\frac{p_{A,F}(k,l)}{p_A(k)p_F(l)} \right)$ is the mutual information between source image A and fused image F , and $MI_{BF} = \sum_k \sum_l p_{B,F}(k,l) \log_2 \left(\frac{p_{B,F}(k,l)}{p_B(k)p_F(l)} \right)$ is the mutual information between source image B and fused image F
6. Information symmetry or fusion symmetry (FS) indicates how much symmetrical the fused image is with respect to source images and is given by $FS = 2 - \left| \frac{MI_{AF}}{MI} - 0.5 \right|$
7. Correlation coefficient (CC) measures a relevance of fused image to source images and is given by $CC = (r_{AF} + r_{BF})/2$, where

$$r_{AF} = \frac{\sum_i \sum_j (a(i,j) - \bar{A})(f(i,j) - \bar{F})}{\sqrt{(\sum_i \sum_j (a(i,j) - \bar{A})^2)(\sum_i \sum_j (f(i,j) - \bar{F})^2)}}$$
, and

$$r_{BF} = \frac{\sum_i \sum_j (b(i,j) - \bar{B})(f(i,j) - \bar{F})}{\sqrt{(\sum_i \sum_j (b(i,j) - \bar{B})^2)(\sum_i \sum_j (f(i,j) - \bar{F})^2)}}$$
8. Spatial frequency (SF) measures the overall information level in the regions (activity level) of an image and is computed as $SF = \sqrt{RF^2 + CF^2}$, where $RF = \sqrt{\frac{\sum_i \sum_j (f(i,j) - f(i,j-1))^2}{mn}}$ and $CF = \sqrt{\frac{\sum_i \sum_j (f(i,j) - f(i-1,j))^2}{mn}}$

In addition to these, objective image fusion performance characterization [5,40] based on gradient information is considered. This provides an in-depth analysis of fusion performance by quantifying: total fusion performance, fusion loss and fusion artifacts (artificial information created). The procedure for computing these parameters are given in [5,40] and their symbolic representation is given below:

$Q^{AB/F}$ = Total information transferred from source images to fused image,
 $L^{AB/F}$ = Total loss of information, and
 $N^{AB/F}$ = Noise or artifacts added in fused image due to fusion process.

It is noted that total fusion performance $Q^{AB/F}$, fusion loss $L^{AB/F}$ and fusion artifacts $N^{AB/F}$ are complimentary indicating that the sum of all these should result in unity [25,40], i.e.,

$$Q^{AB/F} + L^{AB/F} + N^{AB/F} = 1. \quad (6)$$

In most of the cases, this may not lead to unity, and hence, these parameters are reviewed and a modification of the parameter which measures the fusion artifacts has been proposed in Shreyamsha Kumar [25], and its equation is given here for completeness.

$$N_m^{AB/F} = \frac{\sum_{vi} \sum_{vj} AM_{i,j} \left[(1 - Q_{i,j}^{AF}) w_{i,j}^A + (1 - Q_{i,j}^{BF}) w_{i,j}^B \right]}{\sum_{vi} \sum_{vj} (w_{i,j}^A + w_{i,j}^B)} \quad (7)$$

where $AM_{i,j} = \begin{cases} 1, & g_{i,j}^F > g_{i,j}^A \text{ and } g_{i,j}^F > g_{i,j}^B \\ 0, & \text{otherwise} \end{cases}$, indicates

locations of fusion artifacts where fused gradients are stronger than input.

$g_{i,j}^A$, $g_{i,j}^B$ and $g_{i,j}^F$ are the edge strength of A , B and F , respectively,

$Q_{i,j}^{AF}$ and $Q_{i,j}^{BF}$ are the gradient information preservation estimates of source images A and B , respectively,

$w_{i,j}^A$ and $w_{i,j}^B$ are the perceptual weights of source images A and B , respectively.

The procedure for computing these parameters $g_{i,j}^A$, $g_{i,j}^B$, $g_{i,j}^F$, $Q_{i,j}^{AF}$, $Q_{i,j}^{BF}$, $w_{i,j}^A$ and $w_{i,j}^B$ is given in [5,40]. With this newly modified fusion artifact measure $N_m^{AB/F}$, the Eq. 6 can be rewritten as

$$Q^{AB/F} + L^{AB/F} + N_m^{AB/F} = 1. \quad (8)$$

4 Results and discussion

Experiments were carried out on various standard test pairs of multifocus, medical and IR–visible images provided by online resource for research in image fusion (<http://www.imagefusion.org>). Due to lack of space, fusion performance comparison is given only for three standard test pairs, namely medical (MRI), multisensor (gun), multifocus (office). Fused image by the proposed method is compared with different methods discussed in [4,20,22,36,38] with the same simulation parameters indicated in the respective methods with db8 wavelet decomposition. The parameters used for the proposed method are $\sigma_s = 1.8$, $\sigma_r = 25$, neighborhood window = 11×11 (for CBF) and neighborhood window = 5×5 (to find detail strength). The output of CBF is subtracted from the respective source image to get the detail image. For two source images, two detail images are obtained, which are used to find the weights w_{t_a} and w_{t_b} by measuring the detail strengths. These weights are used to find the fused image by weighted average.

Conventional performance measures, API, SD, AG, H , MI, FS, CC and SF are tabulated in Table 1, and the objective performance measures $Q^{AB/F}$, $L^{AB/F}$, $N^{AB/F}$ and $N_m^{AB/F}$ along with their respective sums are tabulated in Table 2. The quality of the fused image is better when these parameters have higher value, excluding $L^{AB/F}$, $N^{AB/F}$ and $N_m^{AB/F}$, which should have lower values. In these Tables, higher values are bolded except for $L^{AB/F}$, $N^{AB/F}$ and $N_m^{AB/F}$, where lower values are bolded. As the goal of image fusion is to enhance comprehensive, accurate and stable information such that the fused image is more suitable for human

Table 1 Conventional image fusion performance measure

Measure	API	SD	AG	<i>H</i>	MI	FS	CC	SF
Input image	Multisensor MRI							
Proposed	54.7351	57.6902	11.0334	6.7311	5.3292	1.6142	0.6565	19.8160
[38]	46.3165	52.3071	8.8572	6.6926	2.3097	1.6899	0.6374	15.6624
[20]	36.4330	51.3242	9.7398	6.1728	1.5464	1.7651	0.5563	17.4809
[4]	40.1711	46.8869	8.9507	6.6467	1.9804	1.7126	0.6185	15.4078
[22]	44.1301	51.3010	9.3093	6.6820	2.0679	1.6880	0.6011	15.8166
[36]	33.8052	36.1403	6.8810	5.9619	3.6685	1.6455	0.6922	12.2309
Input image	Multisensor gun							
Proposed	6.3412	5.5260	3.9668	3.9412	2.2981	1.8756	0.6774	4.8476
[38]	5.7801	6.1473	3.6183	4.0012	1.6561	1.9054	0.6630	4.4755
[20]	4.7303	6.4447	3.5464	3.5656	1.3278	1.7196	0.5956	4.6016
[4]	4.2140	4.0842	3.6094	3.5773	0.9895	1.8138	0.6736	4.3948
[22]	4.0907	4.1445	3.4600	3.5328	0.8361	1.6651	0.6126	4.2559
[36]	4.3483	3.8317	2.7239	3.5857	2.1754	1.8888	0.7247	3.3778
Input image	Simulated multifocus office							
Proposed	81.3517	66.8859	13.2674	7.4217	6.1486	1.9963	0.9643	24.6667
[38]	80.6225	64.0681	12.8277	7.4846	4.8645	1.9802	0.9656	23.7224
[20]	81.0247	65.9277	13.8664	7.3366	5.0888	1.9420	0.9586	24.9933
[4]	80.6444	64.3795	12.7710	7.4750	4.9359	1.9507	0.9654	23.5254
[22]	80.8079	64.3038	12.8823	7.4497	4.8638	1.9476	0.9636	23.5419
[36]	80.4006	59.5628	8.7694	7.5245	5.1310	1.9865	0.9747	14.9741

Table 2 Objective image fusion performance measure

Measure	$Q^{AB/F}$	$L^{AB/F}$	$N^{AB/F}$	Sum (Eq. 6)	$N_m^{AB/F}$	Sum (Eq. 8)
Input image	Multisensor MRI					
Proposed	0.8932	0.0961	0.0950	1.0844	0.0107	1
[38]	0.8065	0.1856	0.0735	1.0657	0.0078	1
[20]	0.6900	0.2776	0.2172	1.1849	0.0323	1
[4]	0.7760	0.2137	0.0924	1.0820	0.0103	1
[22]	0.7300	0.2531	0.1310	1.1142	0.0168	1
[36]	0.6724	0.3260	0.0133	1.0116	0.0017	1
Input image	Multisensor gun					
Proposed	0.9024	0.0953	0.0363	1.0340	0.0022	1
[38]	0.9167	0.0811	0.0524	1.0502	0.0022	1
[20]	0.8744	0.1174	0.1870	1.1787	0.0082	1
[4]	0.7899	0.2094	0.0256	1.0250	0.0006	1
[22]	0.6799	0.3184	0.0337	1.0321	0.0017	1
[36]	0.7619	0.2380	0.0031	1.0029	0.0002	1
Input image	Simulated multifocus office					
Proposed	0.8848	0.1130	0.0116	1.0093	0.0023	1
[38]	0.8766	0.1157	0.0802	1.0724	0.0078	1
[20]	0.8712	0.1102	0.1627	1.1441	0.0186	1
[4]	0.8635	0.1287	0.0523	1.0445	0.0078	1
[22]	0.8571	0.1328	0.0586	1.0485	0.0102	1
[36]	0.8021	0.1978	0.0005	1.0004	0.0001	1

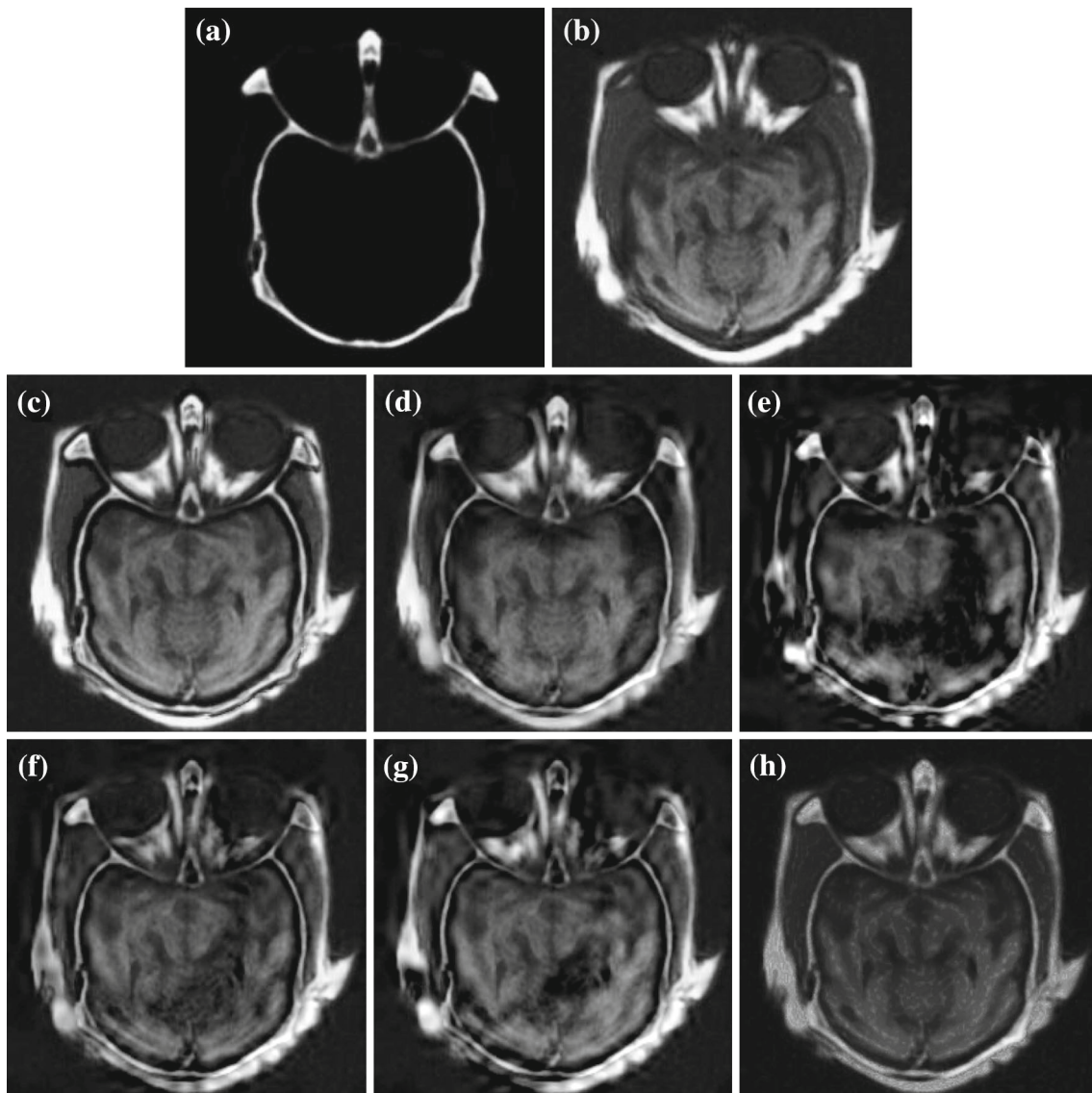


Fig. 3 Multisensor MRI source images in **a** and **b**, fused images by **c** proposed, **d** [38], **e** [20], **f** [4], **g** [23], **h** [36]

perception, visual analysis is also very important in addition to quantitative analysis. There are three criteria that are widely used in the literature for visual analysis: (1) information transferred from each individual image to fused image, (2) information lost from the source images and (3) artifacts introduced because of fusion. To compare the performance visually, the fused images of MRI, gun, office and zoomed version of office are shown in Figs. 3, 4, 5 and 6, respectively. In all the figures, (a) and (b) show the source images, fused images by proposed method in (c), [4, 20, 22, 36, 38] in (d), (e), (f), (g), (h), respectively.

It is observed from Fig. 3 that image quality of the fused MRI image by the proposed method (Fig. 3c) is better than the other methods considered and has all details from both the source images Fig. 3a and b with less information loss and

artifacts. That is, it could be able to fuse much of the information from both the source images when compared with other fused images. The other methods considered could be able to transfer all the details from the source image shown in Fig. 3a, but not from the other source image shown in Fig. 3b. Among the other methods, the method discussed in Shah et al. [4] has better performance than the remaining methods. Now, to compare the performance of fused MRI image quantitatively, Tables 1 and 2 are considered. For multisensor MRI image, it is observed that the proposed method has shown good performance for all the parameters except FS, CC, $N^{AB/F}$ and $N_m^{AB/F}$. The methods in Kotwal and Chaudhuri [36] and Arivazhagan et al. [20] have shown good performance in terms of CC, $N^{AB/F}$, $N_m^{AB/F}$ and FS, respectively. The proposed method has high $Q^{AB/F}$ (more

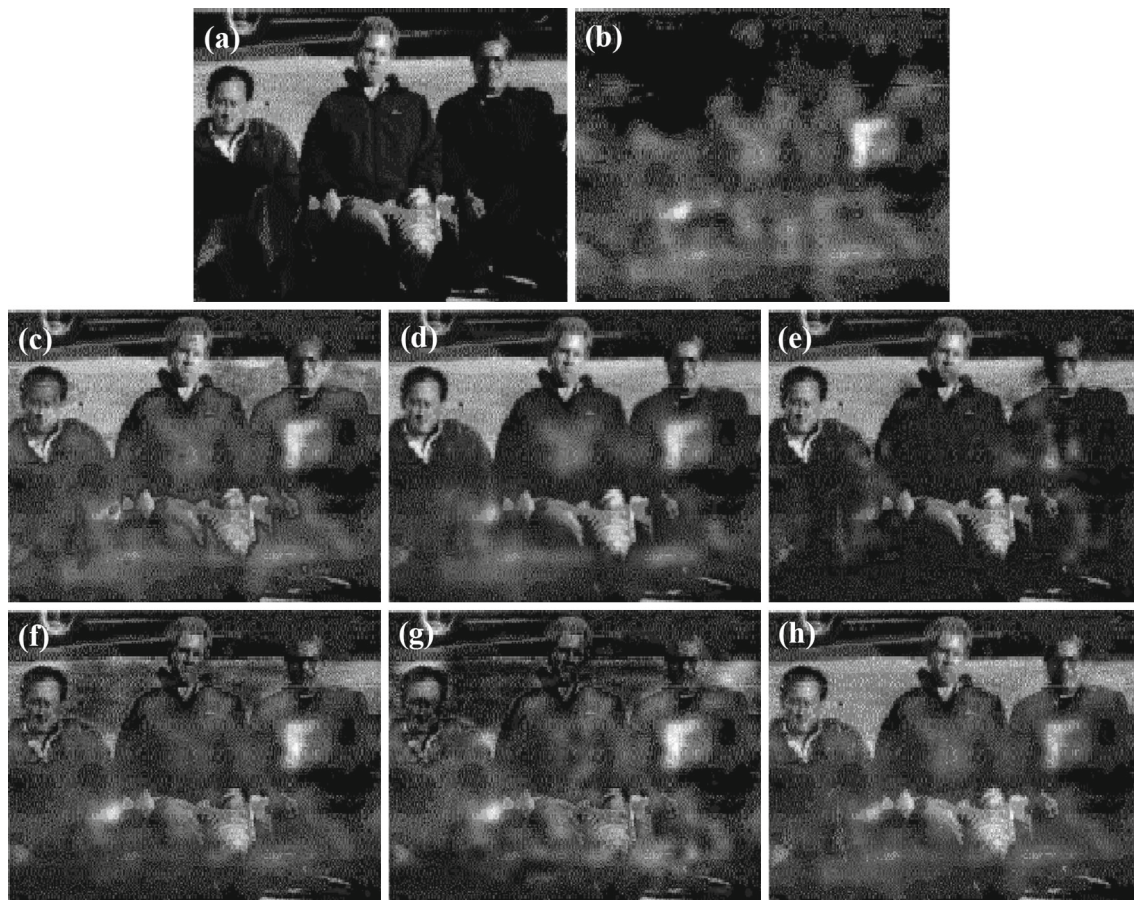


Fig. 4 Multisensor gun source images in **a** and **b**, fused images by **c** proposed, **d** [38], **e** [20], **f** [4], **g** [23], **h** [36]

information has been transferred from source images) and low $L^{AB/F}$ (loss of information is less), and this has been found true by visual inspection of Fig. 3c.

All the fused images shown in Fig. 4 could be able to show the presence of gun except in Fig. 4e [20]. Further, the visibility of the gun in Fig. 4h [36] is feeble compared to other fused images. Eventhough Fig. 4g [22] is showing the presence of gun, it is very difficult to identify the face of the person. Visual performance of the proposed method (Fig. 4c) is almost akin to the methods in Shah et al. [4] (Fig. 4f) and [38] (Fig. 4d). From the Tables, it is observed that the performance of the proposed method is superior to other methods in terms of API, AG, MI, SF with exception of SD, FS, CC, $Q^{AB/F}$, $L^{AB/F}$, $N^{AB/F}$ and $N_m^{AB/F}$. The methods in Kotwal and Chaudhuri [36] and Shah et al. [38] have shown good performance in terms of CC, $N^{AB/F}$, $N_m^{AB/F}$ and H , FS, $Q^{AB/F}$, $L^{AB/F}$ respectively. As in MRI, Kotwal and Chaudhuri [36] has less artifacts in the fused image, i.e., both $N^{AB/F}$ and $N_m^{AB/F}$ have lower values.

Figures 5 and 6 show the visual performance comparison of the fused images of simulated multifocus office and its

zoomed version by different methods. It is observed that the image quality of fused images by Shah et al. [38] (Fig. 5d) and the proposed method (Fig. 5c) looks similar and have almost all the information from both the source images. From Fig. 5h, it is observed that the fused image by Kotwal and Chaudhuri [36] looks blurred with loss of information from both the source images. At the same time, it has fewer artifacts compared to fused images by [4,20,22] (Fig. 5e, f, g) and this is confirmed by almost zero values in $N^{AB/F}$ and $N_m^{AB/F}$. The artifacts can be observed in right top corner (glass of the window above monitor) of the fused images in Fig. 5e, f, g. A horizontal line between focused and unfocused portion of the source image is visible near top portion of the fused images by the proposed method (Fig. (c)) and [38] (Fig. (d)) and is an artifact introduced because of fusion, which has increased the value of $N^{AB/F}$ and $N_m^{AB/F}$. Hence, the performance of the proposed method is better than [4,20,22,38] in terms of $N^{AB/F}$ and $N_m^{AB/F}$ and worse than [36]. As in MRI, here also Kotwal and Chaudhuri [36] has shown good performance in terms of H , CC, $N^{AB/F}$ and $N_m^{AB/F}$. Also, Arivazhagan et al. [20] has shown good performance in terms of AG, SF and $L^{AB/F}$, and in terms of other parameters, the



Fig. 5 Simulated multifocus office source images in **a** and **b**, fused images by **c** proposed, **d** [38], **e** [20], **f** [4], **g** [23], **h** [36]

proposed method supersedes the other methods. The fused image of Kotwal and Chaudhuri [36] has blurred edges (edges of window strip and monitor) (Fig. 6h) as compared to other fused images in Fig. 6.

From Table 1, it is observed that the proposed method has scored well in most of the parameters with very good visual performance. In some cases, [20, 36, 38] have scored only in some of the parameters. From the simulation results, it is clear that the fused image having highest values in any of API, SD, AG, H , MI, FS, CC and SF can be visually degraded compared to other fused images with low value of API, SD, AG, H , MI, FS, CC and SF. This shows that API, SD, AG, H , MI, FS, CC and SF, which may be a good criteria to evaluate fusion performance, may work well for some images and may not for other images. Therefore, performance com-

parison using more appropriate criteria based on the idea of measuring localized preservation of input gradient information in the fused image is used, which is consistent with visual quality of the fused image. From Table 2, it is observed that the proposed method has performed well in terms of $Q^{AB/F}$ for all the images except for multisensor gun and $L^{AB/F}$ for MRI image. But for simulated multifocus office image, Arivazhagan et al. [20] have scored over the proposed method in terms of $Q^{AB/F}$ and $L^{AB/F}$. Whereas, in terms of $N^{AB/F}$ and $N_m^{AB/F}$, Shah et al. [38] has performed well over the other methods. It is clear from the simulation results that the objective image fusion performance metric [2, 40]-based evaluation is in agreement with visual quality. Further, from these experiments, it is found that the visual quality of the proposed method is superior/similar to the methods considered.

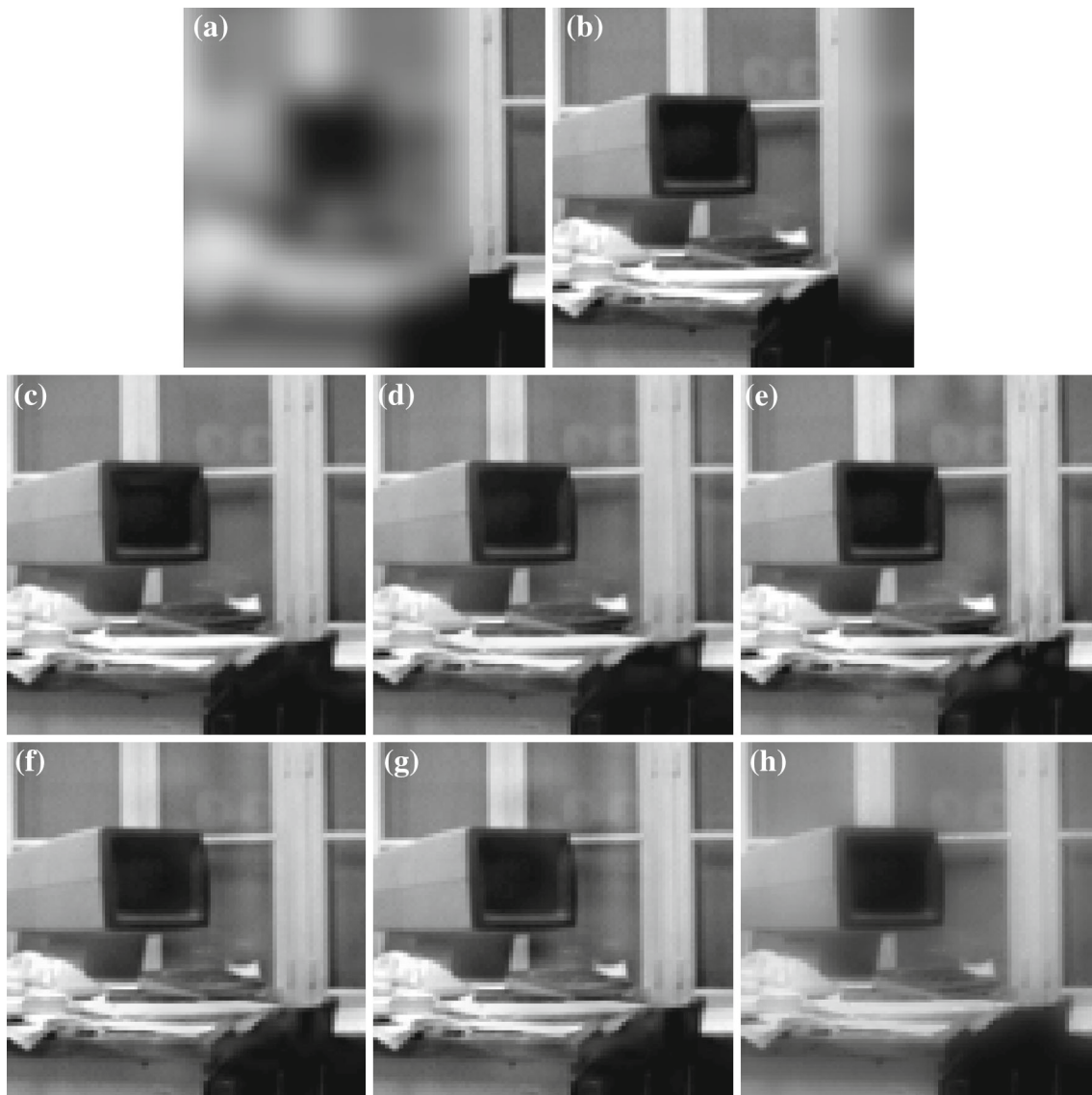


Fig. 6 Simulated multifocus office source images (zoomed) in **a** and **b**, fused images by **c** proposed, **d** [38], **e** [20], **f** [4], **g** [23], **h** [36]

Even in terms of the quantitative parameters, the proposed method is performed well when compared to other methods.

5 Conclusions and future work

In this paper, it was proposed to use detail images extracted from the source images by CBF for the computation of weights. These weights, thus computed by measuring the strength of horizontal and vertical details, are used to fuse the source images directly. Several pairs of multisensor and multifocus images are used to assess the performance of the proposed method. Through the experiments conducted on standard test pairs of multifocus, medical and IR–visible images, it was found that the proposed method has shown

superior/similar performance in most of the cases as compared to other methods in terms of quantitative parameters and in terms of visual quality, it has shown superior performance to that of other methods.

The application of other nonlinear filters instead of CBF for detail image extraction is left as future work and will inspire further research toward image fusion. Also, the performance of the proposed method could be improved by exploring the other methods of weight computation and the domain of weighted average to reduce the fusion artifacts. Further, it can be extended to fuse multiple source images.

Acknowledgments The author would like to express his gratitude to Mr. C. R. Patil, Member (Senior Research Staff), CRL-BEL, India, for his kind support and every possible help. Also, the author thanks the anonymous reviewers for their helpful and constructive comments.

References

- Blum, R., Liu, Z.: *Multi-Sensor Image Fusion and Its Applications*. CRC Press, London (2005)
- Li, H., Manjunath, B.S., Mitra, S.K.: Multisensor image fusion using the wavelet transform. *Graph. Models Image Process* **57**(3), 235–245 (1995)
- Hamza, A.B., He, Y., Krim, H., Willisky, A.: A multiscale approach to pixel-level image fusion. *Integr. Comput. Aid. Eng.* **12**(2), 135–146 (2005)
- Shah, P., Merchant, S.N., Desai, U. B.: Multifocus and multispectral image fusion based on pixel significance using multiresolution decomposition. *J. SIViP*. (2011). doi:[10.1007/s11760-011-0219-7](https://doi.org/10.1007/s11760-011-0219-7)
- Petrovic, V.: *Multisensor pixel-level image fusion*. PhD Thesis, Department of Imaging Science and Biomedical Engineering, Manchester School of Engineering, United Kingdom (2001)
- Sasikala, M., Kumaravel, M.: A comparative analysis of feature-based image fusion method. *Inf. Tech. J.* **6**(8), 1224–1230 (2007)
- Tao, Q., Veldhuis, R.: Threshold-optimized decision-level fusion and its application to biometrics. *Pattern Recogn.* **42**, 823–836 (2009)
- Carper, J.W., Lilles, T.M., Kiefer, R.W.: The use of intensity-hue saturation transformations for merging SPOT panchromatic and multi-spectra image data. *Photogr. Eng. Remote Sens.* **56**, 459–467 (1990)
- Toet, A.: Hierarchical image fusion. *Mach. Vis. Appl.* **3**(1), 1–11 (1990)
- Haerberli, P.: *A multi-focus method for controlling depth of field*. *Grafica Obscura* (1994)
- Qu, G., Zhang, D., Yan, P.: Medical image fusion by wavelet transform modulus maxima. *Opt. Exp.* **9**(4), 184–190 (2001)
- Zhi-guo, J., Dong-bing, H., Jin, C., Xiao-kuan, Z.: A wavelet based algorithm for multi-focus micro-image fusion. In: *Proceedings of the International Conference on Image and Graphics (ICIG)*, pp. 176–179, Dec 2004
- Ranjith, T., Ramesh, C.: A lifting wavelet transform based algorithm for multi-sensor image fusion. *CRL Tech. J.* **3**(3), 19–22 (2001)
- Hill, P., Canagaraj, N., Bull, D.: Image fusion using complex wavelets. In: *Proceedings of the British Machine Vision Conference (BMVC)*, pp. 487–496, Sept 2002
- Du, Y., Vachon, P.W., Vander Sanden, J.J.: Satellite image fusion with multi-scale wavelet analysis for marine applications. *Can. J. Remote Sens.* **29**(1), 14–23 (2003)
- Wang, Z., Ziou, D., Armenakis, C., Li, D., Li, Q.: A comparative analysis of image fusion methods. *IEEE Trans. Geosci. Remote Sens.* **43**(6), 1392–1402 (2005)
- Zheng, S., Shi, W.-Z., Liu, J., Zhu, G.-X., Tian, J.-W.: Multisource image fusion method using support value transform. *IEEE Trans. Image Process.* **16**(7), 1831–1839 (2007)
- Li, S., Kwok, J.T., Wang, Y.: Multifocus image fusion using artificial neural networks. *Pattern Recogn. Lett.* **23**, 985–997 (2002)
- Hao, Y., Sun, Z., Tan, T.: Comparative studies on multispectral palm image fusion for biometrics. *ACCV, Part II, LNCS* **4844**, 12–21 (2007)
- Arivazhagan, S., Ganesan, L., Subash Kumar, T.G.: A modified statistical approach for image fusion using wavelet transform. *J. SIViP* **3**(2), 137–144 (2009)
- Shah, P., Merchant, S.N., Desai, U.B.: Fusion of surveillance images in infrared and visible band using curvelet, wavelet and wavelet packet transform. *Int. J. Wavel. Multiresol. Inf. Process.* **8**(2), 271–292 (2010)
- Shah, P., Srikanth, T.V., Merchant, S.N., Desai, U.B.: A novel multifocus image fusion scheme based on pixel significance using wavelet transform. In: *Proceedings of the Image, Video, and Multidimensional Signal Process. (IVMSP)*, pp. 54–59, Aug 2011
- Rahman, S.M.M., Omair Ahmad, M., Swamy, M.N.S.: Contrast-based fusion of noisy images using discrete wavelet transform. *IET Image Process.* **4**(5), 374–384 (2010). doi:[10.1049/iet-ipr.2009.0163](https://doi.org/10.1049/iet-ipr.2009.0163)
- Naidu, V.P.S.: Discrete cosine transform-based image fusion. *Def. Sci. J.* **60**(1), 48–54 (2010)
- Shreyamsha Kumar, B.K.: Multifocus and multispectral image fusion based on pixel significance using discrete cosine harmonic wavelet transform. *J. SIViP* (2012). doi:[10.1007/s11760-012-0361-x](https://doi.org/10.1007/s11760-012-0361-x)
- Shreyamsha Kumar, B.K., Swamy, M.N.S., Omair Ahmad, M.: Multiresolution DCT decomposition for multifocus image fusion. In: *Proceedings of the Canadian Conference on Electrical and Computer Engineering (CCECE)*, pp. 1–4. Regina, Canada, May 2013. doi:[10.1109/CCECE.2013.6567721](https://doi.org/10.1109/CCECE.2013.6567721)
- Tomasi, C., Manduchi, R.: Bilateral filtering for gray and color images. In: *Proceedings of the International Conference on Computer Vision*, pp. 839–846, Jan 1998
- Zhang, M., Gunturk, B.K.: Multiresolution bilateral filtering for image denoising. *IEEE Trans. Image Process.* **17**(12), 2324–2333 (2008)
- Mustafa, Z.A., Kadah, Y.M.: Multi resolution bilateral filter for MR image denoising. In: *Proceedings of the 1st Middle East Conference on Biomedical Engineering (MECBME)*, pp. 180–184. Sharjah, UAE, Feb 2011
- Shreyamsha Kumar, B.K.: Image Denoising based on Gaussian/Bilateral Filter and its Method Noise Thresholding. *J. SIViP* (2012). doi:[10.1007/s11760-012-0372-7](https://doi.org/10.1007/s11760-012-0372-7)
- Petschnigg, G., Agrawala, M., Hoppe, H., Szeliski, R., Cohen, M., Toyama, K.: Digital photography with flash and no-flash image pairs. *ACM Trans. Gr.* **23**(3), 664–672 (2004)
- Eisemann, E., Durand, F.: Flash photography enhancement via intrinsic relighting. *ACM Trans. Gr.* **23**(3), 673–678 (2004)
- Hu, J., Li, S.: The multiscale directional bilateral filter and its application to multisensor image fusion. *Inf. Fus.* (2011). doi:[10.1016/j.inffus.2011.01.002](https://doi.org/10.1016/j.inffus.2011.01.002)
- Bennett, E.P., Mason, J.L., McMillan, L.: Multispectral bilateral video fusion. *IEEE Trans. Image Process.* **16**(5), 1185–1194 (2007)
- Fattal, R., Agrawala, M., Rusinkiewicz, S.: Multiscale shape and detail enhancement from multi-light image collections. *ACM Trans. Gr.* **26**(3) (2007). doi:[10.1145/1275808.1276441](https://doi.org/10.1145/1275808.1276441)
- Kotwal, K., Chaudhuri, S.: Visualization of hyperspectral images using bilateral filtering. *IEEE Trans. Geosci. Remote Sens.* **48**(5), 2308–2316 (2010)
- Choi, E.-J., Park, D.-J.: Human detection using image fusion of thermal and visible image with new joint bilateral filter. In: *Proceedings of the 5th International Conference on Computer Sciences and Convergence Information Technology (ICCIT)*, pp. 882–885, Nov 2010
- Shah, P., Merchant, S.N., Desai, U.B.: An efficient adaptive fusion scheme for multifocus images in wavelet domain using statistical properties of neighborhood. In: *Proceedings of the 14th International Conference on Information Fusion*, pp. 1–7, July 2011
- Devlin, S.J., Gnanadesikan, R., Kettenring, J.R.: Robust estimation and outlier detection with correlation coefficients. *Biometrika* **62**(3), 531–545 (1975)
- Petrovic, V., Xydeas, C.: Objective image fusion performance characterization. In: *Proceedings of the International Conference on Computer Vision (ICCV)*, vol. 2, pp. 1866–1871 (2005)