Survey on multifocus image fusion techniques

Gurpreet Kaur
Department of Computer engineering and technology
Guru Nanak Dev University
Amritsar, India
gurpreetkaur2091@gmail.com

Prabhpreet Kaur
Department of Computer engineering and technology
Guru Nanak Dev University
Amritsar, India
prabhsince1985@yahoo.co.in

***Abstract*— Multi-focus image fusion is considered to be a vast research topic. Image fusion is the process in which source images are combined to get a single focused image. This focused image obtained contains relatively more information with all objects in focus and better description of scene. It is applied in various applications like medical imaging, remote sensing etc. Various multi-focus image fusion techniques are discussed in this paper, using focus measures such as energy of gradient of image, spatial frequency etc. The performance of these techniques depends on the methods used to determine the focused regions to get a fused image.**

***Keywords*— image fusion; multi-focus; focus measure; fused image; performance evaluation.**

I. INTRODUCTION

Multi-focus image fusion technique integrates two or more source images to get a single focused image. The main purpose of image fusion is to obtain a fused image with greater quality and relatively more information than any of the source images. Since optical lenses have limited depth of field, it is not possible to obtain a sharp image containing all objects with sharp focus. So, image fusion is used in which a set of pictures are captured with camera by different focus adjustments. These pictures are then fused to obtain an image whose depth of field is extended. It is used in various applications such as image processing, remote sensing, computer vision, object recognition, medical imaging etc. There are three levels for process of image fusion which are pixel, decision and feature levels. The lowest level of fusion is pixel level which takes pixels of images into consideration. This level aims for visual enhancement. At this level, image fusion technique processes intensity values of pixels of input images. The advantages of fusion based on pixel level include detection of unwanted noise, less complexity and ease of implementation. However, these methods cause blurring artifacts and do not handle mis-registration. To overcome the problems related to registration and noise sensitivity, methods based on region can be used. In feature level fusion process, feature extraction is done from input images, segmentation of image is done in continuous regions and then combined using fusion method. Features of source images such as contrast, shape, size are combined in feature level. Decision level fusion process deals with image descriptors. Compared to decision and region level image fusion, image fusion based on pixel level provides information with relatively more details. Image fusion process is divided into two main domains based on spatial and frequency. Spatial domain works with pixels of image and combines significant information to get the desired result. Methods such as Principal Component Analysis (PCA), Brovey method, Intensity Hue Saturation[16], high pass filtering fall under category of spatial domain. In frequency domain technique, image is first converted to frequency components and the resultant image fusion method is obtained by combining the frequency coefficients. Pyramid based methods such as ratio of low pass pyramid[17], laplacian pyramid[18], wavelet based methods such as the Discrete Wavelet Transform[19], Haar wavelet[20], Stationary Wavelet Transform[21] and Daubechies wavelets[22] fall under the category of frequency domain methods. The basic steps of fusion are image registration[23] followed by combination of images using fusion method. Image registration is the process of transformation of different images into the same coordinate system. Fusion methods are problem dependent. Different focus measures for analysis of performance of fusion of multi-focus images include variance, energy of gradient, energy of laplacian of image, sum-modified-laplacian. Various multi-focus image fusion techniques are described in this paper and their performance evaluation is done based on focus measures.

II. IMAGE FUSION TECHNIQUES

Y.Phamila and R.Amutha [1] proposed a method based on Discrete Cosine Transform for multi-focus image fusion in visual sensor networks. The source images are divided into blocks of size 8*8. The DCT coefficients of each block is computed. The transformed block with higher valued AC coefficient is chosen for image fusion. Consistency verification is done on the fused DCT blocks. The inverse DCT is applied to the fused DCT coefficients for the reconstruction of original fused image. Higher value of AC coefficients implies higher variance and fine details of the image. This fusion method is more efficient if the fused image is saved in JPEG format. The performance analysis is done based on metrics of Peak Signal to Noise Ratio (PSNR), mean square error, Structural Similarity Index and Petrovic's metric. No complex arithmetic floating point operations such as mean and variance are allowed in this fusion method. This proposed method is more fast and energy efficient as compared to other DCT [25] based fusion methods. This energy efficient method is suitable for resource constrained devices.

978-1-4673-9939-5/16/$31.00

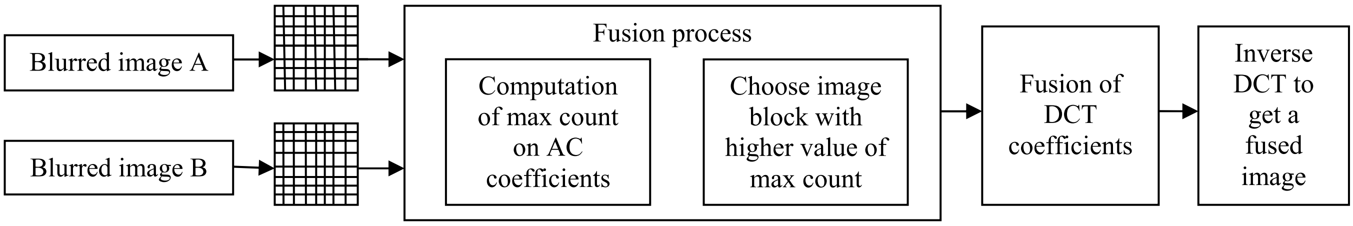

Fig.1. Block diagram of image fusion technique[1].

T.Wan, C. Zhu and Z. Qin [2] proposed a method based on robust principal component analysis (RPCA). The input matrix of data is split into a principal matrix of low rank and a sparse matrix. The principal components in sparse matrix represent dissimilar information. This method is used to form a robust fusion technique to distinguish focused and defocused areas. The extracted features from sparse matrix represent the salient information from source images. These local sparse features are combined to form the resultant image. For reducing blocking artifacts, technique of sliding window is also used for computation of the frequency matrix by searching whole image and select pixels according to sharper regions. This method outperforms wavelet based fusion methods and provides better visual perception. Cost of computation of this technique is high. The performance analysis of this method is done using three metrics, mutual information, petrovic's metric and Structural Similarity Index (SSIM).

S.Li and B.Yang [3] proposed a multi-focus image fusion method based on region. It is processed in the spatial domain. It comprises of three steps: segmentation of image, measuring region based clarity and formation of fused image. The input images are fused by averaging. Segmentation of fused image is done using normalization based cuts. Criterion of normalized criterion can measure total similarity as well as dissimilarity within different groups. Using this segmentation result, partitioning of source images is done. Finally, the fusion of corresponding regions of source images is done based on the measure of spatial frequency. Objective quality index is developed for evaluation of fusion based on spatial frequency. The performance evaluation is done based on mutual information and Petrovic's metric. Petrovic metric is based on the quantity of edge based information transferred to fused images from input images. The proposed method is robust and overcomes problems of pixel based fusion methods such as pixel mis-registration, blurring effects and sensitivity to noise. Representation of image formation using features makes the method more reliable and reduces complexity.

Y.Liu et al. [4] proposed a fusion technique for multi-focus images based on dense scale invariant feature transform (SIFT). The SIFT image is calculated from each source image. Accumulation of all elements of unnormalised SIFT descriptor is done to form an activity level map, containing focus information. Normalisation rule is applied to get a normalized SIFT image. Then, the initial decision map is formed by combining focus information in two activity level maps. Feature matching is done for the further refinement of decision map. This refined fusion map is used to form a fused image. High memory is required. This method outperforms other techniques in respect of objective based performance evaluation and visual perception.

Q.Ziang and B.Guo [5] proposed a fusion technique for multi-focus images using non-subsampled contourlet transform (NSCT). It is a contourlet transform based on shift invariance property. It is based on the principles of selecting low pass and bandpass directional subband coefficients. For the low subband coefficients, selecting scheme is combined with the averaging scheme based on the normal directional vector. For the bandpass directional coefficients, principle of selection is based on standard deviation of directional vector and directional bandlimited contrast. For the shift invariance property, elimination of downsamplers and upsamplers is done during decomposition and image reconstruction. NSCT depends on non-subsampled directional filter and pyramid filter banks. This method is better in terms of both objective based performance evaluation and visual quality compared to other wavelet based methods. It is highly redundant, has high cost and computational complexity

Qi.Miao et al. [6] proposed a method for image fusion using shearlets. Directionality, multi-scalability, localization and anisotropy are the features possessed by shearlet transform. Using shearlet transform, image decomposition can be done in any direction and any scale. Shear matrix is used for multi-direction decomposition of image. Wavelet packets decomposition is used to decompose each direction based on multiscale. Average of coefficients of low frequency of two source images replaces low frequency coefficients of fused image. Then coefficients of high frequency are chosen selectively. The resulting decision map is subjected to the region based consistency check. Performance evaluation is based on metrics such as entropy, difference of entropy, overall cross entropy, standard deviation, mean square error, PSNR and SSIM. The proposed method provides detailed information and smaller distortions.

W.Wu et al. [7] proposed image fusion method based on Hidden Markov Model (HMM). The input images are split to form patches that overlap each other. The clarity is calculated of each patch. The compatibility is computed for each of two neighbouring patches. The fused image is modeled based on clarity and fidelity of each patch using HMM. It is also based on compatibility of neighbouring patches and the patches obtained from input images. The algorithm of belief propagation is applied to form the fused patches. The fused patches are then integrated to form a final fused image. The clarity of the patch is measured using spatial frequency matrix. This method provides better visual perception compared to multiscale transform methods and is robust to mis-registration.

R.Singh and A.Khare [8] proposed a method for image fusion using Daubechies Complex Wavelet Transform (DCWT).It has properties of shift invariance, phase information and multi-scale edge information. It is based on multi-resolution principle. DCWT is used to decompose source images at different levels. Then, wavelet coefficients are formed. Rule of maximum selection is applied to fuse these wavelet coefficients. Inverse DCWT is applied to form a fused image. This method has no redundancy and is symmetric. The performance evaluation is done based on entropy, edge strength, standard deviation, fusion symmetry and fusion factor. This method outperforms other wavelet based and spatial domain methods. This technique is robust to handle Gaussian, speckle, salt and pepper noise. It is shift invariant and preserves multi-scale edge based information. It gives phase based information immune to contrast distortions and noise.

X.Bai et al. [9] proposed an edge preserving image fusion method based on multi-scale toggle cast operator. The same shape and increasing size structuring elements are used by the multi-scale toggle cast operator. This method extracts the multi-scale dilation and erosion features. Edge information of source images is represented by these features. These extracted features are used for the construction of final dilation and erosion fusion features. These constructed features are finally combined to form a fused image. The number of scales and structuring elements affect the performance of this method. Performance evaluation is done based on mean gradient and spatial frequency. This method is effective for preserving edge information.

X.Bai et al. [10] proposed fusion technique for multi-focus images based on a quadtree using a weighted focus measure. Decomposition of input images is done into blocks of optimal sizes based on the structure of quadtree. Detection of focused blocks is done based on a weight based measure, known as sum of the weighted modified laplacian. Calculation of modified laplacian gradients of input images is done. The weighted based sum of these gradients in a local window represent weighted modified form of laplacian gradient at pixel location. Noise interference is suppressed using modified laplcaian. These detected focused regions are combined to form focused blocks. For consistency verification, focused regions are reconstructed using sequential filters. The morphological filter is used which eliminates small lines or blurrs and connect nearby regions. Then, a filter based on small region is applied which removes isolated blocks of small size. Finally, these regions are fused to get a resultant image. The performance evaluation is done using gradient similarity metric, edge based similarity metric and normalized mutual information. This method outperforms other transform domain methods based on computational speed.

H.Zin et al. [11] proposed image fusion technique based on compressive sensing. Decomposition of source images is done using non-subsampled contourlet transform. The subbands of low pass are combined using dual layer Pulse Coupled Neural Network model. The high pass subbands are combined using edge-retention based fusion rule. Gaussian matrix is used to combine the sparse coefficients. Finally, reconstruction of fused image is done using Compressive Sampling Matched Persuit algorithm. The most significant feature can be extracted from source images by directly fusing sparse coefficients. The reconstruction error is reduced by using related measurement matrix for measuring the fused coefficients. This method performs better than other traditional based methods in respect of visual quality. High details and saliency structure is obtained in the fused image by this technique.

H.Li et al. [12] proposed image fusion technique for multi-focus images using decomposition based on sparse matrix and morphological filtering. The source images are decomposed after the extraction of sparse feature matrices. These sparse matrices contain salient features of source images. A temporary matrix is formed by weighting these sparse matrices. Morphological filtering is applied to this temporary matrix for the extraction of bright and dark regions. These extracted features are built into the base image and combined to form a fused image. Decomposition approach used in this method extracts more salient information than robust principal component analysis (RPCA). Pixel wise fusion used in this method takes comparatively less running time and form fused images with higher contrast. Performance evaluation is done based on spatial frequency, mean gradiant, information entropy and mutual information. This method do not preserve original pixel value of source images.

R.Vijayarajan and S.Muttan [13] proposed image fusion method using discrete wavelet transform based on averaging of principal components. The input images are decomposed using discrete wavelet transform with db3 wavelets and one or two decomposition levels. Principal component analysis is done for detailed and approximate coefficients of input images. Evaluation of principal components is done for multi-scale coefficients. Averaging of these principal coefficients is done, which constitutes weights for the fusion rule to be applied. Using these weights, PCA fusion is applied to source images. Performance evaluation is done using non-reference quantitative and qualitative metrics such as average mutual information, average quality index, average peak signal to noise ratio and Hossny, Nahavandi-Creighton metric. This method outperforms other existing state-of-art methods. It has wide application in area of medical imaging for fusion of CT and MRI images.

Y.Liu and F.Yu [14] proposed an automatic algorithm for fusion of multi-focus images which are unregistered. The input images are aligned using image registration. Based on entropy theory, one image is chosen as a reference image for image registration process. For feature matching, Binary Robust Invariant Scalable Keypoints (BRISK) with Speeded up Robust Features (SURF) feature descriptor is used. To reject incorrect matches, improved technique of Random Sample Consensus (RANSAC) is applied followed by transformation and resampling of image. Then stationary wavelet transform (SWT) is used to decompose the registered and reference images. The focused regions are obtained in each reference image. Inverse stationary wavelet transform is used to reconstruct the focused image. The fusion result is determined by the wavelet basis and coefficient selection rules. This method outperforms traditional method of discrete wavelet transform[24]. It is scale invariant and robust to rotation and translation.

III. Comparative analysis

TABLE I. Comparison of different image fusion techniques

No.	Techniques	Domain	Benefits	Limitations
1.	Discrete Cosine Transform	Spatial	Fast, energy efficient, less computational cost and complexity, overcomes computation of energy and low power devices.	Output with less clarity and low peak signal to noise ratio.
2.	Principal Component Analysis	Spatial	Better visual perception, outperforms Discrete Wavelet Transform methods, robust to noise interference.	High computational cost.
3.	Region based method	Spatial	Reduces complexity, overcomes problems related to registration, blurring effects and noise sensitivity.	Larger computational time than wavelet based methods.
4.	Scale Invariant Feature Transform (SIFT)	Transform	Independent of scale, changes in 3D view, orientation of camera and level of illumination.	High memory required.
5.	Non-subsampled Contourlet Transform (NSCT)	Transform	Better visual perception compared to other wavelet based methods and avoid introduction of artificial information.	High redundancy, cost, computational complexity and high memory required.
6.	Shearlets	Transform	Output with detailed information and smaller distortions, possess features of directionality, multi-scalability, localization and anisotropy.	Less sharpness and standard deviation compared to contourlets.
7.	Hidden markovModel	Spatial	Better visual perception, less artifacts compared to multi-scale transform methods and robust to mis-registration.	Not sensitive to patch size, no improvement when patch size is larger than 10*10.
8.	Daubechies complex wavelet transform	Transform	Robust against Gaussian, speckle, salt and pepper noise, preserves multiscale edge and phase information, outperforms other wavelet based methods.	High computational complexity.
9.	Multiscale toggle cast operator	Spatial	Preserves edge information.	Higher computational time.
10.	Quadtree based fusion	Spatial	Outperforms other transform domain methods based on computational speed.	Higher computational cost than transform based methods.
11.	Compressive sensing	Transform	Preserves high details and saliency structure, avoid blocking artifacts and poor fidelity problems, performs better than other traditional methods in respect of visual quality.	High redundancy, running time and computational complexity.
12.	Sparse feature matrix decomposition and morphological filtering	Transform	Output images with higher contrast, takes comparatively less running time.	Do not preserve original pixel values of source images.

13.	Discrete wavelet Transform based Principal Component Averaging	Transform	Output with detailed information, background enhancement, better visual perception compared to other transform based methods.	Output with spatial distortions.
14.	Image registration using SURF, BRISK and fusion using Stationary Wavelet Transform.	Transform	Shift invariant and robust to translation and rotation.	Dependent on dynamic conditions such as change in 3D view, orientation of camera and level of illumination.

IV. PERFORMANCE MEASURES

Quantitative and qualitative metrics are used for performance evaluation of image fusion process. Energy of image gradient, variance, energy of laplacian of the image, spatial frequency and sum-modified-laplacian are some of the focus measures for evaluation[15]. Sum-modified-laplacian gives better performance compared to other measures, when evaluation does not take execution time into consideration. Quantitatively, statistical analysis is used. Reference and non-reference metrics are used for qualitative assessment. Peak signal to noise ratio, mean square error are the reference based metrics. Spatial frequency, Mutual information, Quality index and structural similarity index are non-reference based metrics.

$$RSME=\sqrt{1/(M*N)) \sum_{i=1toM} \sum_{j=1toN} ((S(I,j)-T(I,j))^2} \quad (1)$$

$$PSNR=10\log_{10} G^2/ (RSME)^2 \ (db) \quad (2)$$

$$MI_{ST}= \sum_{i=1 \text{ to } N} \sum_{j=1 \text{ to } N} p_{ST}(i,j)\ 10 \log_2 (p_{ST}(i,j) / p_T(j) * p_S(i)) \quad (3)$$

where RSME is Root Mean Square Error, PSNR is Peak Signal to Noise Ratio, MI_{ST} is Mutual Information, p_{ST} is the normalized histogram of joint gray levels of images S and T having size M*N, p_S, p_T are the normalized marginal histograms of images, G is number of gray levels, T(i,j) and S(i,j) are the values of pixels of reference and fused image.

.

V. CONCLUSION

Image fusion is being used on a vast scale in various applications such as remote sensing, computer vision, medical imaging, object recognition, etc. Some techniques of multi-focus image fusion and metrics based performance evaluation are introduced in this paper. Various issues have been found based on the detailed study of related papers. Image fusion is dependent on problem context. Mis-registration, blurring and noise sensitivity are the limitations of pixel based techniques. Region based methods overcome these limitations, provide good performance but computational complexity is higher than pixel based methods. Spatial domain techniques are proved to be better than those of frequency domain as they are shift invariant and retain more detailed information. Better and improved results can be achieved by using image fusion techniques with varying image block size. Further improvement can be done by hybrid techniques using both spatial and transform based techniques.

ACKNOWLEDGMENT

We sincerely thank to all those who helped us to accomplish this work.

REFERENCES

[1] Y. Phamila and R. Amutha, 'Discrete Cosine Transform based fusion of multi-focus images for visual sensor networks', Signal Processing, vol. 95, pp. 161-170, 2014.

[2] T.Wan , C. Zhu and Z. Qin, 'Multifocus image fusion based on robust principal component analysis', Pattern Recognition Letters, vol. 34, no. 9, pp. 1001-1008, 2013.

[3] S. Li and B. Yang, 'Multifocus image fusion using region segmentation and spatial frequency', Image and Vision Computing, vol. 26, no. 7, pp. 971-979, 2008.

[4] Y. Liu, S. Liu and Z. Wang, 'Multi-focus image fusion with dense SIFT', Information Fusion, vol. 23, pp. 139-155, 2015.

[5] Q. Zhang and B. Guo, 'Multifocus image fusion using the nonsubsampled contourlet transform', Signal Processing, vol. 89, no. 7, pp. 1334-1346, 2009.

[6] Q. Miao, C. Shi, P. Xu, M. Yang and Y. Shi, 'A novel algorithm of image fusion using shearlets',Optics Communications, vol. 284, no. 6, pp. 1540-1547, 2011.

[7] W. Wu, X. Yang, Y. Pang, J. Peng and G. Jeon, 'A multifocus image fusion method by using hidden Markov model', Optics Communications, vol. 287, pp. 63-72, 2013.

[8] R. Singh and A. Khare, 'Fusion of multimodal medical images using Daubechies complex wavelet transform – A multiresolution approach', *Information Fusion*, vol. 19, pp. 49-60, 2014.

[9] X. Bai, F. Zhou and B. Xue, 'Edge preserved image fusion based on multiscale toggle contrast operator', *Image and Vision Computing*, vol. 29, no. 12, pp. 829-839, 2011.

[10] X. Bai, Y. Zhang, F. Zhou and B. Xue, 'Quadtree-based multi-focus image fusion using a weighted focus-measure', *Information Fusion*, vol. 22, pp. 105-118, 2015.

[11] H. Yin, Z. Liu, B. Fang and Y. Li, 'A novel image fusion approach based on compressive sensing', *Optics Communications*, vol. 354, pp. 299-313, 2015.

[12] H. Li, L. Li and J. Zhang, 'Multi-focus image fusion based on sparse feature matrix decomposition and morphological filtering', *Optics Communications*, vol. 342, pp. 1-11, 2015.K. Elissa, “Title of paper if known,” unpublished.

[13] R. Vijayarajan and S. Muttan, 'Discrete wavelet transform based principal component averaging fusion for medical images', *AEU - International Journal of Electronics and Communications*, vol. 69, no. 6, pp. 896-902, 2015.

[14] Y. Liu and F. Yu, 'An automatic image fusion algorithm for unregistered multiply multi-focus images', Optics Communications, vol. 341, pp. 101-113, 201

.

[15] W.Huang and Z.jing, ‘Evaluation of focus measures in multifocus image fusion’, Pattern Recognition Letters, vol.28, no.4,pp. 493-500, 2007.

[16] Myungjin Choi, “A new intensity-hue-saturation fusion approach to image fusion with a tradeoff parameter”, IEEE Trans. Geosci. Remote Sensing, vol.44, no. 6,pp. 1672-1682, 2006.

[17] A. Toet, “Image fusion by a ratio of low-pass pyramid”, Pattern Recognition Letters, Vol. 9, no. 4, pp. 245-253, 1989.

[18] P. Burt and E. Adelson, “The Laplacian Pyramid as a Compact Image Code”, IEEE Transactions on Communications, vol. 31, no. 4, oo. 532-540, 1983.

[19] T. Pu, “Contrast-based image fusion using the discrete wavelet transform”, Optical Engineering, vol. 39, no. 8, p. 2075, 2000.

[20] G. Singh, G. Singh and G. Singh Aujla, “MHWT-A Modified Haar Wavelet Transformation for Image fusion”, International Journal of Computer Applications, vol. 79, no. 1, pp. 26-31, 2013.

[21] H. Kim, S. Kang, J. Park, H.Ha, J. Lim and D. Lim, “Robust image fusion using stationary wavelet transform”, *Korean Journal of Applied Statistics*, vol. 24, no. 6, pp. 1181-1196, 2011.

[22] D. Clonda, J. Lina and B. Goulard, “Complex Daubechies wavelets: properties and statistical image modelling” Signal Processing, vol. 84, no. 1, pp.1-23, 2004.

[23] B. Zitova and J. Flusser, “Image registration methods: a survey”, Image and Vision Computing, vol. 21, no. 11, pp. 977-1000, 2003.

[24] Zhijun Wang, D. Ziou, C. Armenakis, D. Li and Qingquan Li, “A comparative analysis of image fusion methods”, IEEE Trans. Geosci. Remote Sensing, vol. 43, no. 6, pp. 1391-1402, 2005.

[25] L. Cao, L. Jin, H. Tao, G. Li, Z. Zhuang and Y. Zhang,”Multi-Focus Image Fusion Based on Spatial Frequency in Discrete Cosine Transform Domain”, IEEE Signal Processing Letters, vol. 22, no. 2, pp. 220-224, 2015.