Multi-Scale Image Fusion Scheme based on Non-Sub Sampled Contourlet Transform and Four Neighborhood Shannon Entropy Scheme

Ankesh Raj
Dept. of CSE
IIT (ISM), Dhanbad
India-826004
Email: ankesh.1200@gmail.com

Jitesh Pradhan
Dept. of CSE
IIT (ISM), Dhanbad
India-826004
Email: jitpradhan02@gmail.com

Arup Kumar Pal
Dept. of CSE
IIT (ISM), Dhanbad
India-826004
Email: arupkrpal@gmail.com

Haider Banka
Dept. of CSE
IIT (ISM), Dhanbad
India-826004
Email: haider.banka@gmail.com

***Abstract*—Optical lenses have limited depth-of-focus, which makes it impossible to capture all significant objects in focus within single picture. Multi-focus image fusion techniques can be adopted to solve the above issue because this technique precisely selects every focused point from all parent images to create final fused image. So, the final fused image contains significantly more information regarding every salient objects. In this paper, we have proposed a novel image fusion technique for fusion of multi-focused images using non-sub sampled contourlet transform. Here, we have adopted contourlet transform due to its high edge discrimination property which enables us to capture all salient object edges from different parent images. In this approach first we have generated noise free gray scale parent images using non-linear anisotropic diffusion technique. Further, in all noise free images we have employed contourlet transform to discover all salient object edges. Later, we have used 4 neighborhood entropy calculation technique based winner-take-all approach to generate final fused image. We have also used different multi-focused image sets for experimental analysis. The outcomes of all the image fusion experiments show better performance as compared to the current-sate-of-arts.**

***Keywords*—Anisotropic diffusion, Contourlet transforms, Image registration, Multi-scale image fusion, Shannon entropy calculation.**

I. Introduction

A. Background

The visual discrimination power of the human visual unit is extremely high and this is the reason that we can see all significant objects in a single frame with high clarity. Human visual systems are capable of capturing every minute pictorial detail present in any image due to its high visual discriminative power. When it comes to the digital devices, it is nearly impossible to capture all the objects in focus within single picture frame. A typical single lens reflex camera can only capture single object region in focus and rest part of the image is treated as the background image. Those object regions which fall under the depth-of-field (DoF) of any camera will be in focus and treated as the foreground image. To capture all salient objects from a scene, generally more than one picture are taken under different DoF. Consequently, processing of large number of images take high transmission bandwidth as well as storage. All these problems can be solved using multi-focus image fusion techniques [1]. It drives all the visually appealing information from different set of images to a single fused image. So, its main goal is to combine two or more images in such a way that the resultant image retains most salient information of all source images. It has wide variety of applications in different fields like remote sensing, photography, military, machine vision, robotic and medical imaging. Therefore, an efficient and precise multi-focus image fusion technique is direly needed.

B. Related works

In the last decade, several contemporary researches have introduced different multi-focus image fusion techniques. Among all these techniques, the weighted averaging [2] scheme performed on every pixel of all input images to generate the fused images is the most common approach. In 1999, Zang et al. [3] have classified all the image fusion techniques which are based on multi-scale image decomposition principal. Multi-scale or multi-resolution technique based image fusion approaches use the transformed images to achieve the final fused image. In all these techniques, different focused input images have been first converted into multi-resolution images and later fused image has been generated by considering the salient visual information from decomposed images. But all these methods are high sensitive towards sensor noise and simultaneously, all these methods are time consuming and complicated. To overcome these limitations, researches have introduced multi-focus image fusion techniques based on multi-spectral domain information [4]. These multi-spectral image fusion schemes directly work on the pixels of the original image. Later, Nikolov et al. [5] have used discrete wavelet transform (DWT), discrete dyadic wavelet transform (DDWT) and dual-tree complex wavelet transform (DT-CWT) to perform multi-scale image fusion. In 2003, Piella et al. [6] have proposed a new framework for multi-resolution image fusion technique. They have employed a new pixel to region conversion scheme and performed image fusion based the reframed image regions. Later in 2008, Yang et al. [7] have used contourlet transform for multi-model medical image fusion. They have combined the weighted average and local energy of all low-pass and high-pass image sub-bands and used all these information to realize the image fusion operation. In 2009, Looney et al. [8] have used a new data driven approach for multi-scale image fusion. This new data driven scheme

978-1-5386-3039-6/18/$31.00

has been known as Empirical mode decomposition (EMD) and this scheme breaks the image signals into their natural scale components. Later in 2013, Ellmauthaler et al. [9] have used undecimated wavelet transform (UWT) to realize the multi-scale image fusion task. They have used UW T approach to break the process of image decomposition into two different filtering tasks. Further, they have used spectral factorization of the analysis filter and convolution operator to perform the final multi-scale image fusion. In 2017, Zhang et al. [10] have performed image fusion in remote sensing images. They have proposed a new framework for image fusion which first uses the saliency detection scheme to detect the salient image regions. Later, they have employed complex wavelet transform (CWT) to achieve the fused image.

C. Motivation and contribution

In general, most of the images contain well defined boundary regions and clear edge with in the image. But direct application of image transformation and image decomposition scheme on the entire image will gives distinct texture analysis. Multi-focus image fusion based on these distinct image features becomes more complex and time consuming in nature. To solve these problems it is essential to use such scheme which enhances the minor edges and boundaries present within image. All these local image region boundaries can be easily enhanced by using contourlet transform because contourlet transform is extremely suitable for human vision system. So in this paper, we have proposed a novel multi-scale image fusion scheme based on contourlet transform. The major contributions of this paper are as follows:

- Use of suitable image registration technique to generate identical frame images for multi-scale image fusion.
- Use of non-linear anisotropic diffusion scheme to generate the noise free images for multi-scale image fusion.
- Use of non-sub sampled contourlet transform technique to enhance the edges and boundaries of local regions of the image.
- Use of proposed four neighborhood entropy calculation technique based winner-take-all approach to perform multi-sale image fusion.

II. Preliminaries

A. Anisotropic diffusion

Every natural image has some homogeneous (i.e. similar pixels) and non-homogeneous (i.e. entirely different pixels) regions. Homogeneous region has been treated as a background scene whereas non-homogeneous region carries most vital image information. Anisotropic diffusion scheme [11] uses partial differential equation (PDE) preserves the non-homogeneous region of image and subsequently it smoothes the homogeneous region of image. It generates coarser resolution image by using intra-region smoothing techniques. Linear and non-linear schemes can be adopted in anisotropic diffusion to preserve non-homogeneous region. In this work, we have used non-liner anisotropic diffusion scheme because it considers pixels which are having dissimilar intensity values. It controls the diffusion of an image using a flux function which is defined as follows for an image I:

$$f(x, y, t) = df(\delta |I(x, y, t)|) \tag{1}$$

Here, f is the rate of diffusion or flux function; t is the time/scale/iteration and δ is the gradient operator. Simultaneously df is the diffusion function which can be defined using two different ways, which are as follows:

$$df(|A|) = e^{-\{|A|/K\}^2} \tag{2}$$

$$df(|A|) = \left(\frac{1}{1 + (|A|/K)^2} \right) \tag{3}$$

Both of these diffusion functions create a trade-off between edge preservation and smoothing of homogenous image regions. First diffusion function can be much useful if image possesses more high contrast edges. At the same time, second diffusion function can be more useful if image have wide regions. Here K is the free parameter used in both diffusion functions which is useful for deciding validity of region boundaries. Fig. 1 demonstrates the performance of non-linear anisotropic diffusion process.

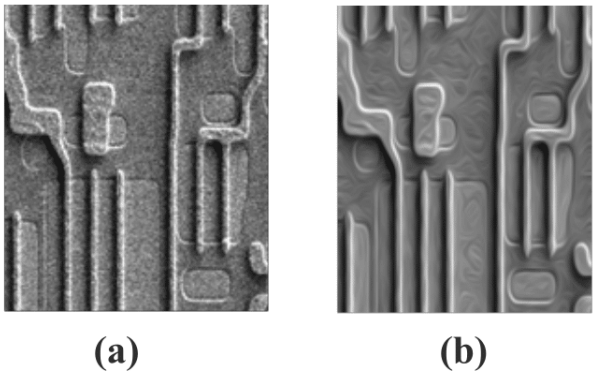

Fig. 1. (a) Normal image and (b) Output image generated after use of non-linear anisotropic diffusion scheme on normal image.

B. Non-sub sampled contourlet transform (NSCT)

Contourlet transform is basically used for representing most salient visual features of an image such as contours, curves and edges of the image because it consist a rich set of basis function. It gives the multi-resolution image representation with enhanced anisotropy and directionality. It uses directional filter banks (DFB) and Laplacian pyramid (LP) to follow the double filtration scheme for construction of contourlet transform. This double filtration scheme is known as pyramidal directional filter banks (PDFB). The contourlet transform works in two different stages where in first stage image has been decomposed into sun-bands. Later, in second stage directional transformed has been applied on the sub-band images to realize the contourlet transform. Here, point discontinuities can be easily captured using Laplacian pyramid and point of discontinuities can be linked into a linear structure using the directional filter bank. So, use of different filter banks

978-1-5386-3039-6/18/$31.00©2018 IEEE

in contourlet transform, not only captures all the geometrical structures but also generates the enhanced high directional representation of input image. These all features of contourlet transform make it suitable for multi-scale image fusion. The framework of the contourlet transform has been shown in fig. 2.

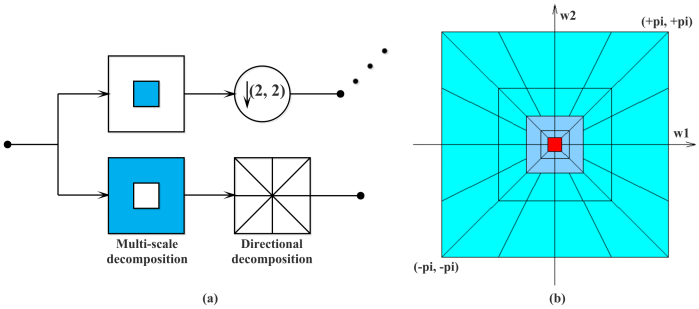

Fig. 2. (a) Contourlet decomposition framework and (b) Frequency partitioning structure.

III. Proposed Multi-focus Image Fusion

In this paper we have proposed a new scheme for multi-focus image fusion. In this proposed image fusion scheme first we have selected n number of images having different depth-of-field (DOF) of a same scene. Let us say these n different DOF images are $I_1, I_2, I_3, ..., I_n$ and in first step we have performed image registration to transform all images into a single coordinate system. After image registration, we have converted all RGB images into gray-scale images i.e. $g_1, g_2, g_3, ..., g_n$. Later, we have employed non-linear anisotropic diffusion (NLAD) technique in all gray-scale images to preserve their salient visual image features. After use of NLAD techniques, all gray-scale images become noise free and all these images can be written as $gf_1, gf_2, gf_3, ..., gf_n$. Further, we have applied non-sub sampled contourlet transform on all noise free images to generate enhanced transformed images i.e. $T_1, T_2, T_3, ..., T_n$. At last we have used 4 neighborhood entropy calculation scheme followed by winner-take-all approach to generate final fused image i.e. I_f. The detailed steps of our proposed image fusion framework have been explained in following sub section.

A. Four neighborhoods Shannon entropy calculation

In this technique, we need to select the i^{th} pixel of the transformed image T_i. Subsequently, based on this i^{th} pixel we have extracted a pixel window w_i of size 3×3 where, center pixel represents the i^{th} pixel of the transformed image Ti and other pixels are adjacent neighbors of the i^{th} pixel. Fig. 3 shows the i^{th} pixel representation in the transformed image Ti with its corresponding 3×3 neighboring pixel window selected for entropy calculation.

After extraction of window w_i, we have selected the actual adjacent neighbors of the center pixel P_i i.e. top pixel (P_T), bottom pixel (P_B), right pixel (P_R) and left pixel (P_L) for entropy calculation. Further, transformed image T_i has been converted into entropy image E_i, where i^{th} pixel of E_i has been defined as follows:

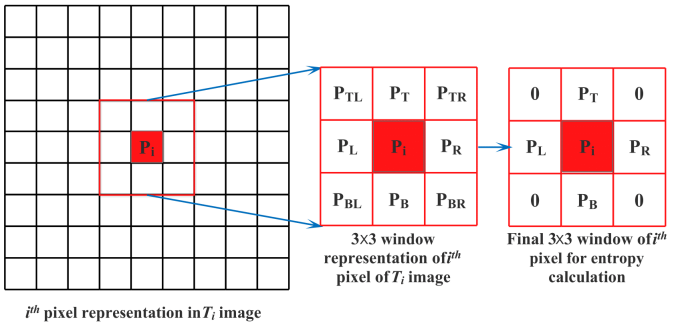

Fig. 3. 3×3 neighboring pixel window selections for entropy calculation.

$$E_i(P_i) = -\sum_{k}^{k \in T_i(P_i), T_i(P_T), T_i(P_B), T_i(P_R), T_i(P_L)} P_k \times log_2(P_k) \tag{4}$$

Further all entropy images $E_1, E_2, E_3, ..., E_n$ have been considered for final fused image I_f generation.

B. Winner-take-all approach

In this step, all entropy images $E_1, E_2, E_3, ..., E_n$ have been used to generate the final fused image I_f. First, we need to select the i^{th} pixel of all entropy images $E_1, E_2, E_3, ..., E_n$ and find the winner entropy pixel among them. The original RGB image pixels of the winner entropy pixel will get a place in the final fused image I_f. Final fused image I_f has been calculated as follows:

$$I_f^C(P_i) = \begin{cases} I_f^C(P_i) = I_1^C(P_i), & \text{If } Wf(E(P_i)) = E_1(P_i) \\ I_f^C(P_i) = I_2^C(P_i), & \text{If } Wf(E(P_i)) = E_2(P_i) \\ I_f^C(P_i) = I_3^C(P_i), & \text{If } Wf(E(P_i)) = E_3(P_i) \\ ... & \\ I_f^C(P_i) = I_i^C(P_i), & \text{If } Wf(E(P_i)) = E_i(P_i) \\ ... & \\ I_f^C(P_i) = I_n^C(P_i), & \text{If } Wf(E(P_i)) = E_n(P_i) \end{cases} \tag{5}$$

Here, $C \in \{Red, Green, Blue\}$ color components hence, $I_f^{Red}(P_i)$, $I_f^{Green}(P_i)$ and $I_f^{Blue}(P_i)$ represents the i^{th} pixel of red, green and blue component of final fused image I_f. Similarly, $I_i^{Red}(P_i)$, $I_i^{Green}(P_i)$ and $I_i^{Blue}(P_i)$ represents the i^{th} pixel of red, green and blue component of original RGB image I_i . Simultaneously, $E_i(P_i)$ represents the i^{th} pixel of E_i entropy image and $Wf(.)$ is the winner function which is defined as follows:

$$Wf(E(P_i)) = maxof\{E_1(P_i), E_2(P_i), E_3(P_i), ..., E_n(P_i)\} \tag{6}$$

C. Proposed image fusion framework

Fig. 4 demonstrates the schematic flow diagram of the proposed multi-scale image fusion framework and algorithm of this proposed work is given below:

978-1-5386-3039-6/18/$31.00©2018 IEEE

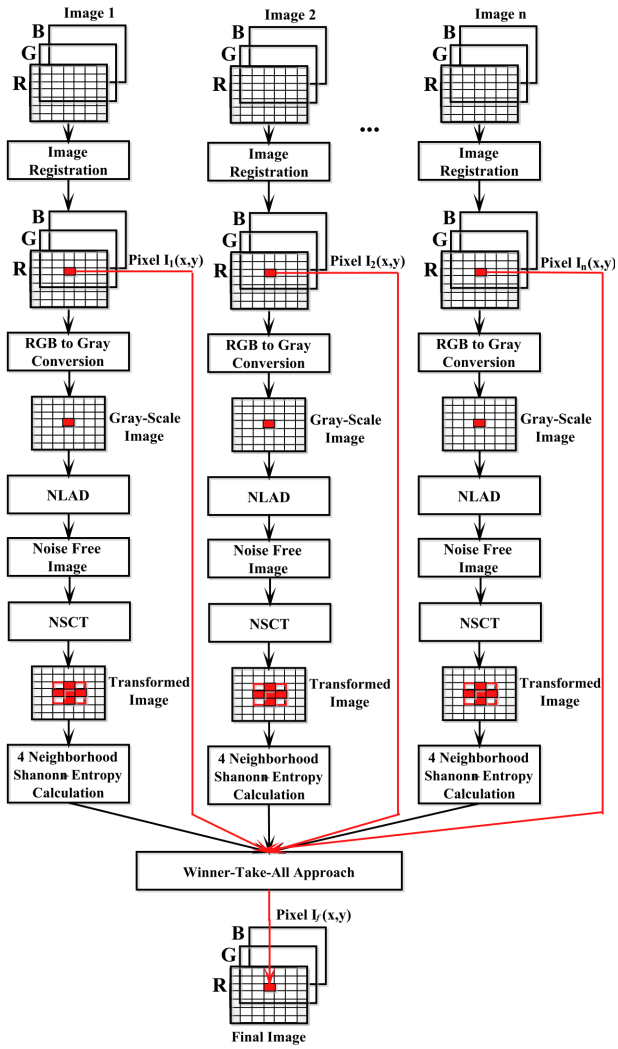

Fig. 4. Basic block diagram of the proposed multi-focus image fusion technique.

Algorithm 1 Multi-scale Image Fusion.

Input: n input images i.e. $I_1, I_2, I_3, ..., I_n$.
Output: Final multi-scale fused image I_f.

1: Select n input images of same scene with different DOF i.e. $I_1, I_2, I_3, ..., I_n$.
2: Perform image registration to transform all images into a single coordinate system.
3: Convert all n registered images into gray-scale image i.e. $g_1, g_2, g_3, ..., g_n$.
4: Apply NLAD scheme in all gray-scale image to generate noise free and more sharp edge images i.e. $gf_1, gf_2, gf_3, ..., gf_n$.
5: Apply NSCT technique in all noise free image to generate transformed images i.e. $T_1, T_2, T_3, ..., T_n$.
6: Apply 4 neighborhoods Shannon entropy calculation technique to generate n entropy images i.e. $E_1, E_2, E_3, ..., E_n$.
7: Consider all n entropy images and apply winner-take-all scheme to generate final multi-scale fused image I_f.

IV. Experimental Results and Discussion

We have used Mat-Lab R2013a software for image fusion experiments in a personal computer which is having windows 10 platform with 4 GB RAM. In this work we performed different multi-scale image fusion experiments on different standard image sets. Subsequently, we have used different performance evolution parameters such as root mean square error (RMSE), percentage fit error (PFE), peak signal to noise ratio (PSNR), etc. In first experiment we have considered set of colored book image. This set of image contains two different DOF images along with the reference image of dimension 1024×768. Fig. 5 shows the output of the proposed multi-scale image fusion scheme in this image set.

Further, table I shows the comparison between our proposed image fusion scheme and other state-of-arts i.e. discrete cosine transform (DCT), curvelet transform, block-level DCT, TV-l_1 model and Morphological component analysis (MCA) decomposition based image fusion methods. Table I shows the comparisons in terms of execution time and PSNR value between reference image and output fused image. All values of the table I have been obtained by

performing image fusion experiments on the fig. 5 image set. Further, table II shows the comparison between our proposed method and some others methods in terms of different evolution criterias which are Figure definition (FD), Cross-Entropy (CE), Average Gradient (AG), Edge Retention $Q_{AB/F}$, Structural Similarity (SSIM), Relatively Warp (RW), Edge-Intensity (EI) and Mutual Information (MI) [12]. Table II shows the comparison between our proposed method and other method which are Laplacian Pyramid method (LP), Ratio Pyramid method (RP), Morphological Pyramid method (MP), Gradient Pyramid method (GP), FSD Pyramid method (FSD), DWT-based method (DWT), Contrast Pyramid method (CP), PCA-based method (PCA), Select maximum based method (Max) and average-based method (AV). Form table I and II, it is clear that our proposed image fusion scheme is giving acceptable results. At the same time it is also giving high PSNR values for the book image set data as compared to other methods. Later, fig. 6 shows the graphs of PSNR values obtained with respect to different window size taken in the image fusion process. Simultaneously, fig. 7 and 8 shows the graphs of RMSE values and Correlation coefficients obtained with respect to different window size taken in the image fusion process. Subsequently, fig. 9 shows the graphs of Absolute mean values obtained with respect to different window size taken in the image fusion process.

Fig. 5. (a) Input book image 1 (b) Input book image 2 (c) Reference book image for fusion and (d) output fused book image.

TABLE I. COMPARISON BETWEEN PROPOSED SCHEME AND OTHER-STATE-OF-ART METHOD IN TERMS OF EXECUTION TIME AND PSNR VALUE

S. No.	Methods	Execution Time	PSNR
1.	DCT	1.9588 sec.	22.4146
2.	Curvelet	1.6655 sec.	25.2725
3.	Block DCT	5.3015 sec.	25.5751
4.	TV	6.5781 sec.	21.7943
5.	MCA	6.8715 sec.	27.7943
6.	Proposed	5.7570 sec.	74.4209

TABLE II. COMPARISON BETWEEN PROPOSED SCHEME AND OTHER-STATE-OF-ART METHODS WITH RESPECT TO DIFFERENT EVOLUTION CRITERIA

S. No.	Methods	$Q_{AB/F}$	RW	MI	SSIM	FD
1.	AVMCA	0.7978	0.027	5.3814	0.9626	8.8379
2.	MAX	0.7666	0.0205	5.8521	0.9547	9.2181
3.	PCA	0.799	0.0266	5.4402	0.9587	8.8619
4.	CP	0.7712	0.0087	5.3669	0.9168	11.9958
5.	DWT	0.7621	0.0033	4.8076	0.9136	12.3349
6.	FSD	0.7310	0.0861	4.0011	0.8927	10.493
7.	GP	0.7416	0.0862	4.0176	0.8954	10.4233
8.	MP	0.7296	0.0177	4.4438	0.09048	12.1859
9.	RP	0.7625	0.0097	5.6569	0.9436	10.0131
10.	LP	0.7746	0.0062	5.2433	0.9211	11.6845
11.	TV-l1	0.6386	0.1928	3.8879	0.7931	12.5483
12.	MCA	0.6329	0.0408	4.3807	0.8977	7.1332
13.	Proposed	0.8026	0.0913	7.0136	0.8510	11.9198

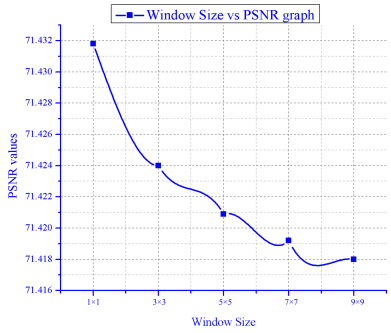

Fig. 6. Graph of PSNR values obtained for different window size.

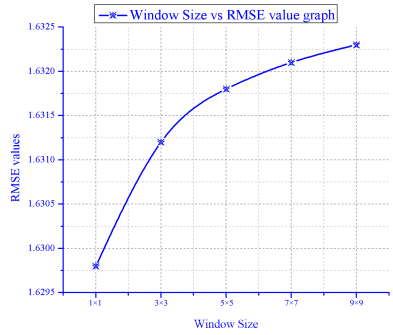

Fig. 7. Graph of RMSE values obtained for different window size.

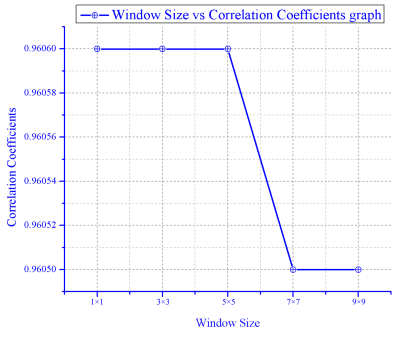

Fig. 8. Graph of Correlation Coefficients values obtained for different window size.

We have also considered different image sets for experi-

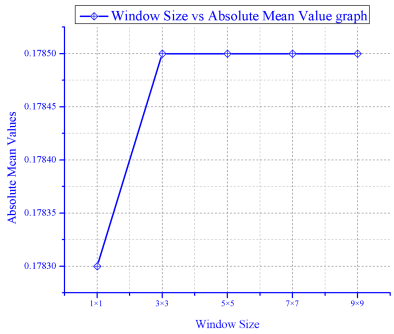

Fig. 9. Graph of Absolute mean values obtained for different window size.

mental analysis. The results of image fusion for all these image sets have been shown bellow in fig 10. These image sets are Clock, Disk, Doll and Jug image datasets. In fig. 10 the first and second column represents the different DOF input image; third column contains the fused reference image whereas last column presents the output fused image. In fig. 10 we can see that, our proposed multi-scale image fusion technique is generating satisfactory results for different types of image sets.

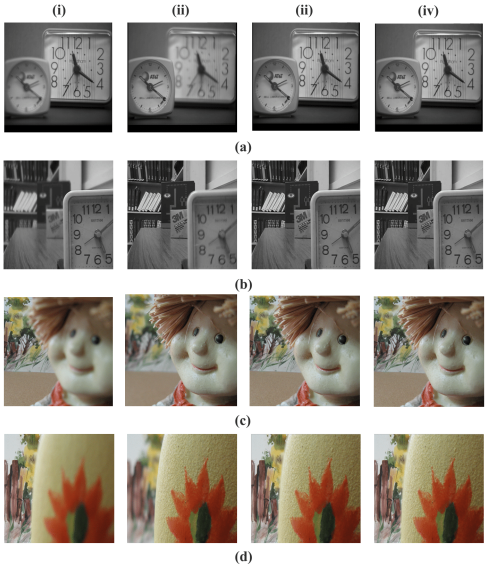

Fig. 10. Experimental image fusion outcomes for different image sets i.e. (a) Clock image set (b) Disk image set (c) Doll image set and (D) Jug image set.

V. Conclusions

In digital imaging process, multi-focus image fusion plays an important role such as remote sensing, robotics, military surveillance, clinical medicine and computer vision. In multi-scale image fusion different DOF images have been fused in a single image. Thus, the final image must contain all the most salient visual information of every input image. These salient visual information of all different DOF image can be easily capture by applying contourlet transform because, it has rich set of basis function which preserves the contours, curves and edges of the image. So in this paper, we have employed contourlet transform to achieve better quality of the fused image. Initially, we have registered all input images so that they can be represented using single coordinate systems. Further, we have used NLAD scheme to eliminate noise from all registered image. This approach will remove noise and also enhances the edges and boundaries of the image. Subsequently, we have used NSCT scheme to generate transformed images to preserve the salient image information from all noise free images. Finally, we have applied 4 neighborhoods Shannon entropy calculation followed by winner-take-all approach to generate final fused image. This scheme selects the most salient visual information from different DOF images and places them in the final fused image which ultimately improves the quality of the fused image. We have also performed experimental analysis on different image set and the results are acceptable in all cases. Furthermore, the experimental results have shown significant improvements in terms of PSNR values.

References

[1] A. Ardeshir Goshtasby and S. Nikolov, "Guest editorial: Image fusion: Advances in the state of the art," *Inf. Fusion*, vol. 8, no. 2, pp. 114–118, Apr. 2007. [Online]. Available: http://dx.doi.org/10.1016/j.inffus.2006.04.001

[2] S. K. Helmy A. Eltoukhy, "Computationally efficient algorithm for multifocus image reconstruction," pp. 5017 – 5017 – 10, 2003. [Online]. Available: http://dx.doi.org/10.1117/12.476754

[3] Z. Zhang and R. S. Blum, "A categorization of multiscale-decomposition-based image fusion schemes with a performance study for a digital camera application," *Proceedings of the IEEE*, vol. 87, no. 8, pp. 1315–1326, Aug 1999.

[4] A. P. James and B. V. Dasarathy, "Medical image fusion: A survey of the state of the art," *Information Fusion*, vol. 19, no. Supplement C, pp. 4 – 19, 2014, special Issue on Information Fusion in Medical Image Computing and Systems. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1566253513001450

[5] S. Nikolov, P. Hill, D. Bull, and N. Canagarajah, *Wavelets for Image Fusion*. Dordrecht: Springer Netherlands, 2001, pp. 213–241. [Online]. Available: https://doi.org/10.1007/978-94-015-9715-9_8

[6] G. Piella, "A general framework for multiresolution image fusion: from pixels to regions," *Information Fusion*, vol. 4, no. 4, pp. 259 – 280, 2003. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1566253503000460

[7] L. Yang, B. Guo, and W. Ni, "Multimodality medical image fusion based on multiscale geometric analysis of contourlet transform," *Neurocomputing*, vol. 72, no. 1, pp. 203 – 211, 2008, machine Learning for Signal Processing (MLSP 2006) / Life System Modelling, Simulation, and Bio-inspired Computing (LSMS 2007). [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0925231208003706

[8] D. Looney and D. P. Mandic, "Multiscale image fusion using complex extensions of emd," *IEEE Transactions on Signal Processing*, vol. 57, no. 4, pp. 1626–1630, April 2009.

[9] A. Ellmauthaler, C. L. Pagliari, and E. A. B. da Silva, "Multiscale image fusion using the undecimated wavelet transform with spectral factorization and nonorthogonal filter banks," *IEEE Transactions on Image Processing*, vol. 22, no. 3, pp. 1005–1017, March 2013.

[10] L. Zhang and J. Zhang, "A new saliency-driven fusion method based on complex wavelet transform for remote sensing images," *IEEE Geoscience and Remote Sensing Letters*, vol. PP, no. 99, pp. 1–5, 2017.

[11] J. Weickert, "Anisotropic diffusion in image processing," 1996.

[12] Z. Liu, Y. Chai, H. Yin, J. Zhou, and Z. Zhu, "A novel multi-focus image fusion approach based on image decomposition," *Information Fusion*, vol. 35, no. Supplement C, pp. 102 – 116, 2017. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1566253516300781