International Journal of Scientific Research and Innovation ISSN 2455-0140
Volume 1, Issue 1, July 2015 IJSRI 1001

A STUDY ON THE FUSION OF REGISTERED INFRARED AND VISUAL IMAGES

K.R.Siva Bharathi

Assistant Professor
Sri Krishna College of Engineering and Technology, Coimbatore
Email id: sivabharathi@skcet.ac.in

R.Nivedha

Assistant Professor
Sri Krishna College of Engineering and Technology, Coimbatore
Email id: nivedhar@skcet.ac.in

***Abstract*—Fusion of registered pair of infrared (IR) and visual images of the same scene provides greater information of the scene. Two challenges of the process, addressed in this paper are– matching the resolution of both the image acquisition modes, identification of the optimal algorithm in the sense of providing greater composite information from the input images. Interpolation operators and fusion methods existing in literature were studied to address the challenges. Based on this study the optimal interpolation operator and fusion method for low contrast images has also been identified in this paper.**

***Index Terms*—IR, interpolation, fusion, low contrast**

I. INTRODUCTION

Images from multiple sensors provide better understanding of the scene. The some of the information in a composite image is always higher than that obtained from a single image of the scene. In particular, we discuss about the topic of fusing visual and IR images of the scene. IR cameras, whether standalone or integrated, that are used in defense and commercial sectors have resolutions lesser than their visual counterparts. This is due to limitations in the packing density of IR sensor. Therefore bringing images from both the modes to the same resolution is crucial for further image processing tasks. Four interpolation operators, namely Bilinear, Bicubic, lanczos2 and lanczos3 are tested for resizing the IR image. This resizing corresponds to increasing the resolution to bring it to the same resolution as the visual image. The next part of the paper is a study of various fusion techniques. Direct fusion results in reduced contrast of the images and hence the images become unsuitable for further processing. Therefore few existing algorithms starting from pyramid methods to discrete wavelet transform methods are studied and their performances are compared.

The rest of the paper is arranged as follows. Section-II briefly reviews about the interpolation operators used in this study and the metrics used for comparison. Section-III briefly discusses about the fusion methods and fusion rules used. Section-IV. discusses the results and a possible way of automating the identification of fusion method for low contrast images.

II. INTERPOLATION OPERATORS

A. Bilinear:

The simplest and computationally efficient operator is the bilinear operator. This operator linearly interpolates in both the directions. Bilinear interpolation of a point G surrounded by 4 points A, B, C and D is shown in Fig.1. It is implemented in two steps first row-wise and then column-wise.

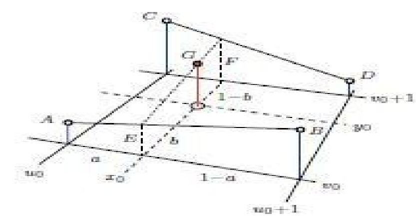

Fig.1.Bilinear Interpolation.

B. Bicubic:

The next operator in the family is the Bicubic operator. It extrapolates the concept of cubic interpolation in 1-D, wherein acubic kernel is used, to 2-D. It is also implemented in two steps.A rough sketch of the interpolation process is shown in Fig.2.

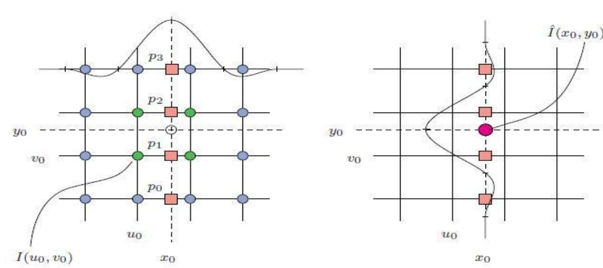

Fig .2.Bicubic Interpolation

where the horizontal and vertical axis is marked as y_0 and x0 respectively and the desired point is marked as $\hat{I}(x_0,y_0)$. Interpolation is performed in two steps: first along the row and then along the column as shown in the figure.

C. Lanczos2 and Lanczos3:

At this point, to give a quick reminder, it should be remembered that a good interpolation operator more closely resembles a sinc function. Since a sinc function is infinitely long, approximations to it consider few neighboring points alone with increasing complexity providing better approximations. The lanczos2 and lanczos3 operators come from a family of "windowed sinc" operators. Their 1-D representation is shown overlaid against the sinc functions in Fig.3

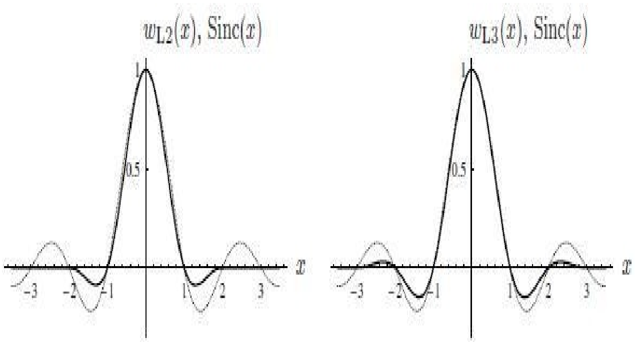

Fig.3:Lanczos2 and Lanczos3 operators in the left and right figures overlaid with the sinc function.

Metrics such as the Peak Signal to Noise Ratio (PSNR), Mean Squared Error (MSE) and correlation coefficient (ρ) are used for comparing these operators. All these operators are available under Image Processing toolbox of MATLAB V7.0. The reader is referred to [1, 2] for additional insights on concepts of interpolation.

III. FUSION METHODS AND FUSION RULES

Pyramid fusion methods, or also called multi-resolution analysis, are available since 1980"s and provide of plethora of implementation ways. [3 - 9] They are also accompanied by a specific fusion rule which establishes the rule for fusion starting from the lowest level. Other simple rules such as the weighted average, spatial frequency methods are also discussed in this paper.

A. Pyramid based fusion:

An image pyramid consists of a set of low pass or band pass copies of an image, each copy representing pattern information of a different scale [6]. The generic flow of a pyramid algorithm is as follows and is repeated for „L" number of desired levels:

Step.I: Preprocess the original image using a suitable approximation filter and down sample by a factor of 2. Use this as the input for next level of reduction. Up sample the image using a suitable approximation filter and then compare with the original to get the difference image. Repeat this step for „L" levels to get a coarse approximation at level „L" and difference images stacked on top of it to get the individual pyramid.

Step.II: Compare the two pyramids at each level using a suitable fusion rule and form a composite pyramid

Step.III: Reconstruct the output image from the composite pyramid by following the decomposition steps in reverse order.

1. LaplacianPyramid(LP):

Here a Gaussian filter is used at each level to form the approximation image. The downsampled image is up sampled using an interpolation filter and a set of bandpass copies is obtained from the difference between the interpolated image and original L^{th} level image at each such step giving rise to a difference pyramid. The entire image can be reconstructed using the last level approximation image and the difference pyramid. During reconstruction two fusion rules are followed in this paper. First rule creates an output pixel as the average (AVG) of the corresponding pixels in difference pyramid and coarse approximation to form the composite pyramid. The next rule selects the pixel whose 8-neighborhood has greater average intensity or local sum (LS).

2. Gradient Pyramid(GP):

At each step of decomposition two averaging filters are used and as the name indicates a gradient pyramid is formed, instead of a difference pyramid, at each level with the help of four directional gradient filters.

$$\text{Horizontal filter} \quad \begin{bmatrix} 0 & 0 & 0 \\ 1 & -2 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\text{Vertical filter} \quad \begin{bmatrix} 0 & 1 & 0 \\ 0 & -2 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

$$\text{Diagonal filter} \quad \begin{bmatrix} 0 & 0 & 0.5 \\ 0 & -1 & 0 \\ 0.5 & 0 & 0 \end{bmatrix}$$

$$\text{Diagonal filter} \quad \begin{bmatrix} 0.5 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0.5 \end{bmatrix}$$

A max" fusion rule was followed at each level. Since the gradients are used, a max operator here implies choosing a pixel from two pyramids that has a greater local contrast among each other.

3. Filter, Subtract and Decimate Pyramid(FSD):

This is a computationally efficient version of the Laplacian pyramid. Here the difference pyramid is generated immediately after approximation filtering and then down sampled. This mitigates the need of interpolation filter during the decomposition process whereas it is used in the reconstruction process.

A. max fusion rule:

Max fusion rule was used during reconstruction.

B. Wavelet based fusion:

The Discrete Wavelet Transform(DWT) captures notonly a notion of the frequency content of the input, by examining it at different scales, but also temporal content, i.e. the times at which these frequencies occur [6]. This is also an iterative way of implementation such as the pyramids but their underlying mathematics is well proved and established. Fusion methods using DWT have been well examined in [10, 11]. Further it has been proved in [12] that in the fusion of visual and IR images one level of decomposition provided the optimal tradeoff between computational load and output efficiency irrespective of the wavelet function chosen. Here we choose a Daubechies 4-tap wavelet function. Using the „max‟ operator corresponds to selecting image pixels with sharper brightness changes which may represent salient changes in the image such as the line, edges and region boundaries.

C. Spatial Frequency based Image fusion:

Spatial Frequency (SF) measures the overall activity in an image. This method can be implemented efficiently in real time. Exhaustive treatment of the subject is dealt in [13-14]. For an M x N image F, with the gray value at pixel position (m, n) denoted by F (m, n), its spatial frequency is defined as

$$RF = \sqrt{\frac{1}{MN}\sum_{m=1}^{M}\sum_{n=2}^{N}\left(F(m,n) - F(m,n-1)\right)^2}$$

$$CF = \sqrt{\frac{1}{MN}\sum_{n=1}^{M}\sum_{m=2}^{N}\left(F(m,n) - F(m-1,n)\right)^2}$$

Where RF is called the row frequency and CF is called as the column frequency and are defined as:

The images are to be decomposed into blocks and the corresponding blocks from each image are compared against each other. The block with higher SF is chosen to be a part of the final composite image.

D. Adaptive Weight Averaging based fusion (AWA):

In this method, each picture element is assigned a weight proportional to the interest associated with it [15]. The thermal weights are taken from the divergence of the pixels intensity from the mean pixel intensity and the visual weights are determined by the local variance in time. The pixel (x, y) in the final fused image is obtained as accordingly:

$$I^{F}_{(x,y)} = \frac{1}{Wt_{(x,y)} + Wv_{(x,y)}}\left[Wt_{(x,y)} * I^{T}_{(x,y)} + Wv_{(x,y)} * I^{V}_{(x,y)}\right]$$

where: $I^F_{(x,y)}$ denotes the intensity in (x, y) location of the final image. $W^V_{(x,y)}$ denotes the thermal and visual weights used for the input images. $I^T_{(x,y)}$ and $I^V_{(x,y)}$ denote the intensity levels at location (x, y) in the input thermal and visual images respectively.

E. Principal Component Analysis (PCA)

The optimal weighting coefficients, with respect to information content and redundancy removal, can be determined by a principal component analysis (PCA) of all input intensities. By performing a PCA of the covariance matrix of input intensities, the weightings for each input image are obtained from the eigenvector corresponding to the largest eigenvalue. A simple diagrammatic representation of the PCA is shown in Fig.4 where I1 and I2 represent the input visual and IR images respectively

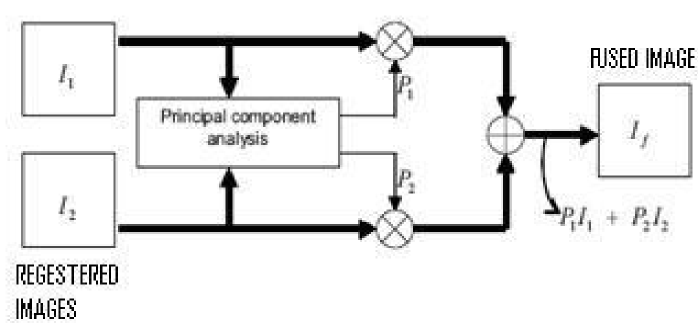

Fig .4.Algorithm of PCA used in our study

IV. RESULTS AND DISCUSSION

The methods were tested using a registered set of IR and visual images obtained from the database of [16] which were supported by TNO, Netherlands. It consisted of the image of a surveillance scene captured both in visual and IR modes. In this case the bimodal image resolutions were 270*360 pixels. The images are low-contrast and are shown in Fig. 5. The results are presented two-fold.

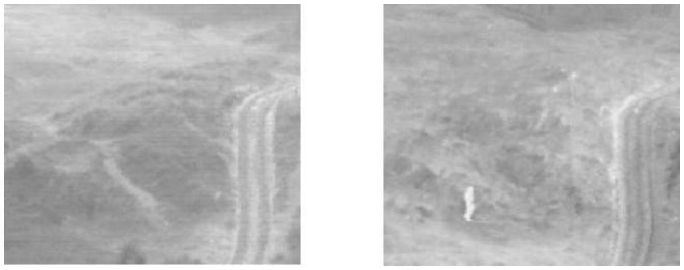

Fig .5. Input Visual image (left) and Input IR image (Right)

A. *Interpolation Operators:*

The original IR image was down sampled by a factor of 2 and 4. The interpolation operators were used against these two images and compared with the original.

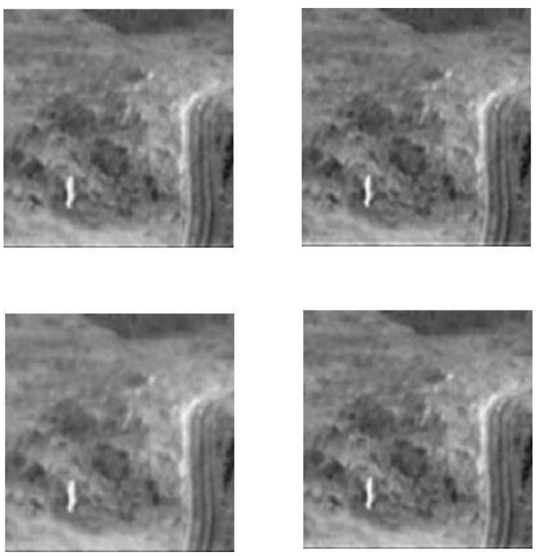

Fig. 6 Output of using Bilinear, Bicubic, Lanczos3, Lanczos2 (clockwise direction) operator

The output figures are shown in Fig. 7 for a resizing factor of 4. The results are presented in Table I and Table II for aresizing factor of 2 and 4 respectively. Owing to their computational sophistications, it could be observed that both Bicubic and lanczos3 are better in terms of PSNR and MSE than compared to bilinear and lanczos2. Both Bicubic and lanczos3 were found to perform similar for a resize factor of 2 but for a factor of 4 Bicubic performed better than lanczos3

B. *Fusion methods:*

The fusion operators might cause the pixel value to go beyond the dynamic range of 8-bit gray scale level [0,255]. Therefore proper scaling was ensured for a uniform comparison. For PCA, since occluded information always lays in the IR band the eigenvector corresponding to the largest eigenvalue was used as the weighting function for the IR image. Since the SF is a block based method partitioning the input images into block sizes of 9*9 gave the optimum result.

Choosing a block size lesser than this leads to a patchy effect in the output whereas a block size below this dint effectively choose the prime features. The images obtained are shown in Fig. 7. To test the efficiency of the fusion operators' three metrics namely Entropy, Relative variances (RV) and mutual information (MI) were used [7, 11]. The variances were scaled within the different methods to form RV which could give a better inference. In addition subjective criterion was also used to arrive at the results. A good fusion operator needs to have greater entropy and variance which corresponds to more information and greater mutual information. The results are shown in Table III.

On inspection of the results we come to the conclusion that gradient pyramid turns out to be the best in meeting 2 of the 3 quantitative metrics and proving to be superior subjectively also. Spatial Frequency method is block based and hence it tends to show high values for MI. The adaptive weight averaging methods also performed closely well. In fact gradient pyramid performs superior on most of the low contrast input images. This is because the gradient "amplifies" the differences and "attenuates" the similarity between adjacent pixels. For an image to be classified as low contrast its histogram must have a constrained spread about its mean. After empirical study on a test set of 10 images it was decided that a naturally occurring image with a normalized variance of approximately 0.28 or less may be considered as low contrast. Based on this fact, it can be concluded that if the variance of the input images were less than 0.28 a gradient pyramid based fusion method provides the desired optimal fusion.

TABLE I

COMPARISON OF INTERPOLATION OPERATORS
(RESIZE FACTOR=2)

	PSNR(dB)	MSE	ρ
Bilinear	98.57	3.4054	0.9882
Bicubic	102.72	2.2472	0.9910
Lanczos3	102.56	2.2844	0.9907
Lanczos2	102.77	2.2393	0.9909

TABLE II
COMPARISON OF INTERPOLATION OPERATORS
(RESIZE FACTOR=4)

	PSNR(d)	MSE	ρ
Bilinear	89.90	8.1056	0.9429
Bicubic	90.29	7.7995	0.9519
Lanczos3	90.17	7.8866	0.9505
Lanczos2	89.98	8.0448	0.9524

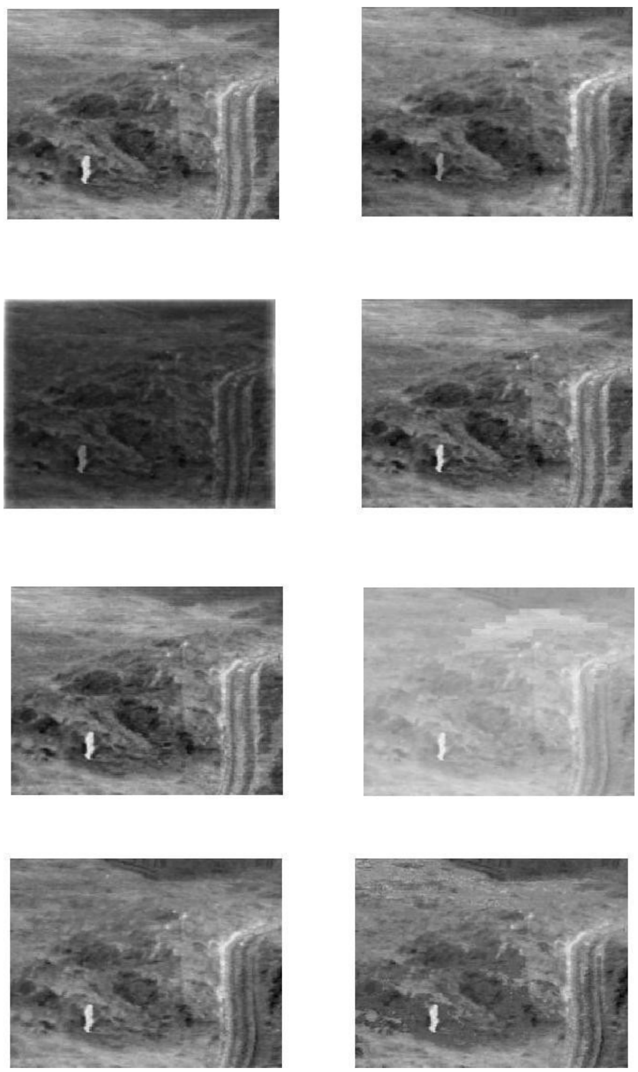

Fig.7 : (In clockwise order) Output of LP(LS), LP(AVG), FSD(Max), SF, AWA, PCA, DWT(Max), GP(Max), fusion methods using specified fusion rules in brackets

V. CONCLUSION

This study has resulted in the following results. It has identified Bicubic kernel among the three other kernels as the best for the interpolation of infrared images. Also naturally occurring images (visual/IR) having a variance of approximately 0.28 can be classified as low-contrast images. Finally it has been concluded that gradient pyramid provides a superior way of fusion of such bi-modal images effectively.

TABLE III

COMPARISON OF FUSION SCHEMES

Method (Fusion rule)	Entropy	RV	MI
LP (LS)	5.3242	0.5556	0.9429
LP (AVG)	5.0892	0.1045	0.9519
GP (Max)	5.5955	1.0000	2.1358
FSD (Max)	5.3217	0.5501	0.9524
DWT(Max)	5.3333	0.5683	2.1460
SF	5.3307	0.6739	5.1485
PCA	5.2004	0.4197	4.0626
AWA	5.2840	0.3102	3.8246

REFERENCES

[1] Principles of Digital Image Processing: Core Algorithms (Undergraduate Topics in Computer Science) – Wilhelm Burger; Mark.J.burge, pp. 210-238.

[2] Unser, M.; , "Splines: a perfect fit for signal and image processing," *Signal Processing Magazine, IEEE* , vol.16, no.6, pp.22-38, Nov 1999doi: 10.1109/79.799930

[3] Burt, P.J.; Kolczynski, R.J.; , "Enhanced image capture through fusion," *Computer Vision, 1993. Proceedings., Fourth International Conference on* , vol., no., pp.173-182, 11-14 May 1993doi:10.1109/ICCV.1993.378222 Burt, P.J.; Kolczynski, R.J.; , "Enhanced image capture through fusion," *Computer Vision, 1993. Proceedings., Fourth International Conference on* , vol., no.,pp.173-182, 11-14 May 1993 doi: 10.1109/ICCV.1993.378222

[4] Filter subtract decimate hierachial pyramid signal analyzing

and synthesizing techniques - http://www.google.com/patents?id=h2svAAAAEBAJ&zoom=4&dq=Filter%20subtract%20decimate%20hierachial%20pyramid%20signal%20analyzing%20and%20synthesizing%20technique&pg=PA1#v=onepage&q&f=false

[5] ShivsubramaniKrishnamoorthy and K P Soman. Article:Implementation and Comparative Study of Image Fusion Algorithms. *International Journal of Computer Applications* 9(2):25– 35, November 2010. Published By Foundation of Computer Science.

[6] ManjushaDeshmukh, UdhavBhosale**,** Image Fusion and Image Quality Assessment of Fused Images, International Journal of Image Processing (IJIP), Volume (4): Issue (5)

[7] FiroozSadjadi, "Comparative Image Fusion Analysais," cvprw, pp.8, 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPRW'05) - Workshops, 2005

[8] Schnelle, S.R.; Chan, A.L.; , "Enhanced target tracking through infrared-visible image fusion," *Information Fusion (FUSION), 2011Proceedings of the 14th International Conference on* , vol., no., pp.1-8,5-8 July 2011

[9] Gonzalo Pajares, Jesús Manuel de la Cruz, A wavelet-based image fusion tutorial, Pattern Recognition, Volume 37, Issue 9, September 2004, Pages 1855-1872, ISSN 0031-3203, 10.1016/j.patcog.2004.03.010.

[10] Naidu, VPS and Raol, JR (2008) *Pixel-level image fusion usingwavelets and principal component analysis.* Defence Science Journal,58 (3). 338 -352. ISSN 0011-748X

[12] Su-xia Xing, Pei-yuanGuo, Tian-hua Chen, "Study on Optimal Wavelet Decomposition Level in Infrared and Visual Light Image Fusion," icmtma, vol. 3, pp.616-619, 2010 International Conference on Measuring Technology and Mechatronics Automation, 2010

[13] Project Report on "Image Fusion" for the Degree of M.Eng in Electrical Engineering by Eduardo FernándezCanga, University of Bath.

[14] Shutao Li, James T. Kwok, Yaonan Wang, Combination of images with diverse focuses using the spatial frequency, Information Fusion, Volume 2, Issue 3, September 2001, Pages 169-176, ISSN 1566-2535, 10.1016/S1566-2535(01)00038-0.

[15] Lallier, E.; Farooq, M.; , "A real time pixel-level based image fusion via adaptive weight averaging," *Information Fusion, 2000. FUSION 2000.Proceedings of the Third International Conference on* , vol.2, no., pp.*WEC3/3-WEC313 vol.2, 10-13 July 2000 doi: 10.1109/IFIC.2000.859841*

[16] www.imagefusion.org -The Online Resource for Research in ImageFusion.