Computer Science & Engineering Department

Project on

“Multi-focus Color Image Fusion based on Stationary Wavelet Transform (SWT)”

Nov 28, 2018,

Rajnish Kumar Ranjan
15JE001619

Dual Degree. Computer Science & Engineering

Indian Institute of Technology(ISM), Dhanbad - 826004

Session 2018-19

Acknowledgement

We would like to express the deepest appreciation to my project mentor Dr. Rajendra Pamula, who has the attribute and the substance of a genius. He continually and convincingly conveyed a spirit of adventure in regard to the project, and an excitement in regard to the teaching. Without his guidance and persistent help this project would not have been possible. Finally, we must say that no height is ever achieved without some sacrifices made at some end and it is here where we owe our special dept to our parents and our friends for showing their generous love and care throughout the entire period of time.

Nov 28, 2018

Rajnish Kumar Ranjan

Place – Indian Institute of Technology (ISM), Dhanbad

ABSTRACT

This project is intended for the fusion of two or more multi-focus color images using Stationary wavelet transform. The feature which has been widely used for it is related to image quality is focus. Sharp images provide better information than blurry images. We have implemented the code in MATLAB. The degree of fusion depends on the decomposition level and number of combinations

CONTENTS

1. Introduction

Generally two types of vision are classified. They are human vision and computer vision. Human vision is sophisticated system that senses and acts on visual stimuli. It has evolved for millions of years, primarily for defense or survival. Basic computer vision system requires a camera, a camera interface and a computer.

A feature closely related to image quality is focus. Sharp images provide better information than blurry images. However, in some situations it is not possible to obtain totally focused images in just one single camera shot, since some regions appear to be blurred due to variations in the depth of the scene and of the camera lenses focus. This means that if the camera is focused at one specific object, another region of the scene can be out of focus. An interesting solution is to take more pictures of the desired landscape in the same position, but with focus centered in different elements of the scenery. Then, using the image fusion concept, all source images are combined, creating a single image that contains all the best focused regions. Image fusion is becoming very popular in digital image processing.

The main aim of any image fusion algorithm is to coalesce all the important visual information from multiple input images such that the resultant image contains more accurate and complete information than the individual source images, without introducing any artifacts.

The good image fusion method has the following properties. First, it can preserve most of the useful information of different images. Second, it does not produce artifacts which can distract or mislead a human observer or any subsequent image processing steps. Third it must be reliable and robust. Finally it should not discard any salient information contained in any of the input images.

Image fusion can take place at the signal, pixel, feature and symbol level, and many image fusion approaches have been developed and presented in literatures, including IHS, PCA, and HPF and so on. These approaches are mostly focus on grayscale images but seldom color images, so this paper presents a method to fuse multi-focus color images, and it is based on SWT and IHS at pixel level which is described below.

1.1 Evolution of Image Fusion

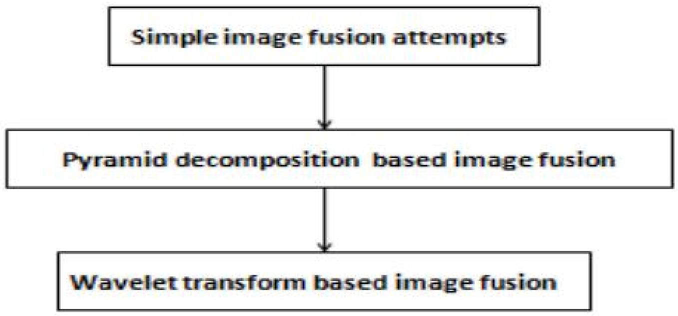


Figure 2: evolution of image fusion

The first evolution of image fusion research is simple image fusion, which perform the basic pixel by pixel related operations like addition, subtraction, average and division. Normally fusion techniques which rely on simple pixel operations on the input image values. Now the following operations are described.

Addition: Addition is the simplest fusion operation. It works by estimating the average intensity value of the input images on a pixel-by-pixel basis. The technique assumes a semantic alignment and requires very accurate spatial and radiometric alignment. The technique has the advantage of suppressing any noise which is present in the input images.

Average: The pixel average technique has the disadvantage that it tends to suppress salient image futures producing a low contrast image with a “washed-out” appearance.

Subtraction: Subtraction is the complement to addition and is used as a simple fusion operator in change detection algorithms.

Multiplication: Multiplication is not widely used as image fusion operators. However, one important image fusion application where multiplication is used in Brovey pan sharpening.

The second evolution of image fusion research is pyramid decomposition based image fusion. The primitive fusion schemes perform the fusion right on the source images, which often have serious side effects such as reducing contrast. With the introduction of pyramid transform in mid80’s, some sophisticated approaches began to emerge. People found that it would be better to perform the fusion in the transform domain.

More recently, with the development of wavelet theory, people began to apply wavelet decomposition to take the place of pyramid decomposition for image fusion. Actually wavelet transform can be taken as one special type of pyramid decompositions

1.2 Image Fusion Techniques

In general, fusion techniques can be classified into different levels. They are signal level, pixel/data level, feature level and decision level.

Signal level fusion in signal based fusion, signals from different sensors are combined to create a new signal with a better signal to noise ratio than the originals signals.

Pixel/ Data level fusion is the combination of raw data from multiple sources into single resolution data, which are expected to be more informative and synthetic than either of the input data or reveal the changes between data sets acquired at different times.

Feature level fusion extracts various features, e.g. edges, corners, lines, texture parameters etc., from different data sources and then combines them into one or more feature maps that may be used instead of the original data for further processing. It used as input to preprocessing for image segmentation or change detection.

Decision level fusion combines the result from multiple algorithms to yield a final fused decision. When the results from different algorithms are expressed as confidences rather than decisions, it is called soft fusion. Otherwise it is called hard fusion. Methods of decision fusion include voting methods, statistical methods and fuzzy logic based methods.

2. Image Fusion Methods

Now a day's many fusion methods are available in research, but every new method based on the common characteristics on basics method. Here above methods described in theoretically.

2.1. Spatial Average Image Fusion Method

This is a simple image fusion method that is to take the average of the source image pixel by pixel. The block diagram is shown in Figure 2.

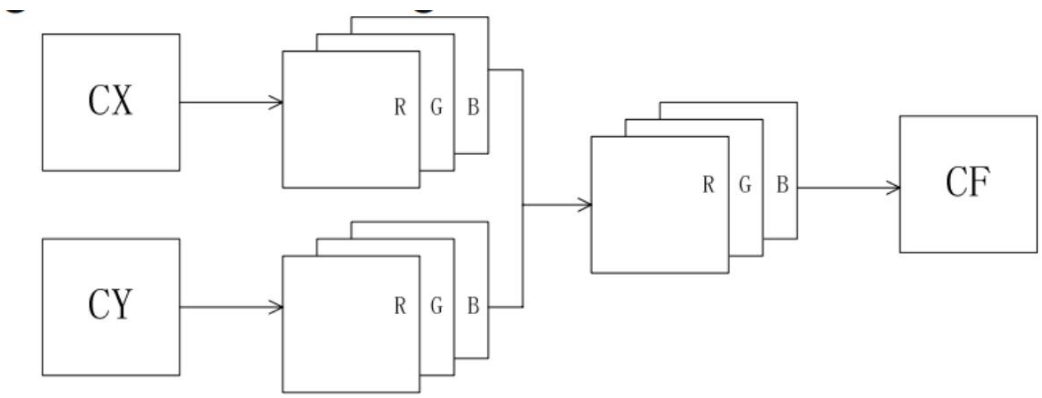


Figure 2. Block diagram of spatial average method

We know that each image have three multispectral bands of R, G, B. This method is to take the average of the corresponding two bands of CX and CY as one band of CF, respectively.

$$CF(m,n,i) = \frac{CX(m,n,i) + CY(m,n,i)}{2} \quad i = R,G,B$$

(m,n) denotes the pixel in spatial domain in an image. However, along with simplicity this method comes several undesired side effects including reduced contrast.

2.2. Intensity-Hue-Saturation (IHS) Image Fusion Method

IHS is a common way of fusing high spatial resolution, single band, pan and low spatial resolution, multispectral remote sensing image. The R, G and B bands of the multispectral image are transformed into HIS components, replacing the intensity component by the pan image, and performing the inverse transformation to obtain a high spatial resolution multispectral image (see fig 3). HIS can enhance spatial details of the multispectral image and improve the textural characteristics of the fused, but the fusion image exist serious spectral distortion. The HIS transform is used for geologic mapping because the IHS transform could allow diverse forms of spectral and spatial landscape information to be combined into a single data set for analysis.

Transformation of RGB to HIS is given as

$$\begin{bmatrix} I \\ V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{-1}{\sqrt{6}} & \frac{-1}{\sqrt{6}} & \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} & \frac{-1}{\sqrt{6}} & 0 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

$$H = \tan^{-1}\left(\frac{V_1}{V_2}\right)$$

$$S = \sqrt{(V_1)^2 + (V_2)^2}$$

Limitations of IHS Fusion Method

Although the HIS method has been widely used, the method cannot decompose an image into different frequencies in frequency space such as higher or lower frequency. Hence the IHS method cannot be used to enhance certain image characteristics. The color distortion of HIS technique is often significant.

2.3. Stationary Wavelet Transform (SWT) fusion method

2.3.1 Stationary Wavelet Transform

The wavelet transform has become a very useful tool for image fusion. The wavelet based approach is appropriate for performing fusion takes for the following reasons

1. It is a multi-scale (multi-resolution) approach well suited to manage the different image resolutions.

2. The discrete wavelet transform allows the image decomposition in different kinds of coefficients preserving the image decomposition

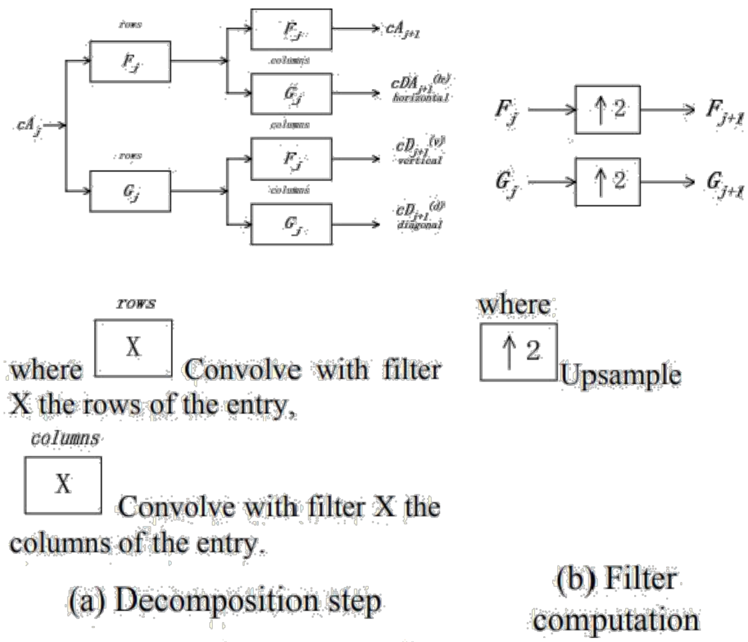
3. Such coefficients coming from different images can be appropriately combined to obtain new coefficients so that the information in the original images is collected appropriately.

4. Once the coefficients are merged, the final fused image is achieved through the inverse discrete wavelets transform (IDWT), where the information in the merged coefficients is also preserved. The key step in image fusion based on wavelets is that of coefficients combination in order to obtain the best quality in the fused image. This can be achieved by set of strategies.

First, decompose the two source images into smaller blocks. Second, compute the focus measure for each block. Then, compare the focus measure of two corresponding blocks, and select the blocks with bigger focus measure as the corresponding blocks of the fused image. Select the energy of image gradient (EOG), energy of Laplacian (EOL), spatial frequency (SF) or sum-modified Laplacian (SML) as focus measure.

The fusion rule which is commonly used is that the low frequency coefficients are obtained with average method and the high frequency coefficients by setting each coefficient equal to the corresponding input image wavelet coefficient that has the greatest absolute value.

However, the DWT yields a shift variant data representation by the down sampling process and is not available for the image fusion [3]. The SWT algorithm is very simple and is close to the DWT one. SWT was presented in [4] which is time invariant. A clear description and analysis of the method is presented in [5]. In summary, the SWT method can be described as at each level, when the high and low pass filters are applied to the data, the two new sequences have the same length as the original sequences. To do this, the original data is not decimated, however, the filters at each level are modified by padding them out with zeros. The algorithm for SWT on image is visualized in Figure 1.



(a) Decomposition step (b) Filter computation

Figure 1. SWT

2.3.2 SWT+IHS fusion method

The block diagram of this method is shown in Figure 4.

1. Perform IHS transform on CX and CY so as to the color space from RGB to IHS and get the I-components of CX and CY: X and Y. Since this method only uses the I-component but H component and S component, we can ignore the transformation of H and S to cost low computation.

2. Get wavelet coefficients of X and Y by taking SWT: Dx(m,n,k,l) and Dy(m,n,k,l) , where m,n indicate the spatial position in a given frequency band, l the decomposition level and k the frequency band of the

MSD representation, k=a,h,v,d , a indicates the approximation band, h the horizontal band, v the vertical band and d the diagonal band.

3. Calculate the activity level of each pixel for X and Y, Ax(m,n) and Ay(m,n)

$$A_I(m,n) = \sum_{k\in K, l\in L} w(k,l)\left|D_I(m,n,k,l)\right|$$
$$K = \{h,v,d\} \quad L = \{1,2,\cdots,N\}$$

Where N indicates the maximum decomposition level and w (k, l) the weight coefficient of the Corresponding pixel in k band at level l such that

$$\sum_{k\in K, l\in L} w(k,l) = 1$$

N.G. Kharatishvili et al argued that for the detail information in different directions at same decomposition scale, one's sensibility to the information's loss and error in the diagonal direction is lower than that in the horizontal and vertical direction. And in fact the contributions of high frequency coefficients at different decomposition scale are different: in general the contribution of coefficients at low decomposition scale is greater than that at high decomposition scale. Based on this point, there are four combinations rules of alternatives:

(1) Combination rule 1: same weight of coefficients in different directions, same weight of coefficients at different decomposition scales, for example w1 in SWT with 3 decomposition level.

(2) Combination rule 2: different weight of coefficients in different directions, same weight of coefficients at different decomposition scales, for example w2 in SWT with 3 decomposition level.

(3) Combination rule 3: same weight of coefficients in different directions, different weight of coefficients at different decomposition scales, for example w3 in SWT with 3 decomposition level.

(4) Combination rule 4: different weight of coefficients in different directions, different weight of coefficients at different decomposition scales, for example w4 in SWT with 3 decomposition level.

$$w_1 = \begin{bmatrix} \frac{1}{9} & \frac{1}{9} & \frac{1}{9} \\ \frac{1}{9} & \frac{1}{9} & \frac{1}{9} \\ \frac{1}{9} & \frac{1}{9} & \frac{1}{9} \end{bmatrix} \quad w_2 = \begin{bmatrix} \frac{2}{15} & \frac{2}{15} & \frac{2}{15} \\ \frac{2}{15} & \frac{2}{15} & \frac{2}{15} \\ \frac{1}{15} & \frac{1}{15} & \frac{1}{15} \end{bmatrix} \quad w_3 = \begin{bmatrix} \frac{4}{21} & \frac{2}{21} & \frac{1}{21} \\ \frac{4}{21} & \frac{2}{21} & \frac{1}{21} \\ \frac{4}{21} & \frac{2}{21} & \frac{1}{21} \end{bmatrix} \quad w_4 = \begin{bmatrix} \frac{8}{35} & \frac{4}{35} & \frac{2}{35} \\ \frac{8}{35} & \frac{4}{35} & \frac{2}{35} \\ \frac{4}{35} & \frac{2}{35} & \frac{1}{35} \end{bmatrix}$$

4. M(m,n) is calculated using the coefficients Ax(m,n) and Ay(m,n) as

$$M(m,n) = \begin{cases} 1, if\ A_X(m,n) \geq A_Y(m,n) \\ 0, if\ A_X(m,n) < A_Y(m,n) \end{cases}$$

This decision map decides which image the pixel of the fused image to come from.

5. Get a fused image F from CX, CY and the decision map M (m, n)

$$F(m,n,i) = M(m,n)CX(m,n,i) + \overline{M(m,n)}CY(m,n,i)$$
$$i = R,G,B$$

6. Consistency verification is performed on each pixel of F. If a center pixel comes from image CX while the majority of surrounding pixels come from image CY, the center pixel is then changed to come from image CY and vice versa.

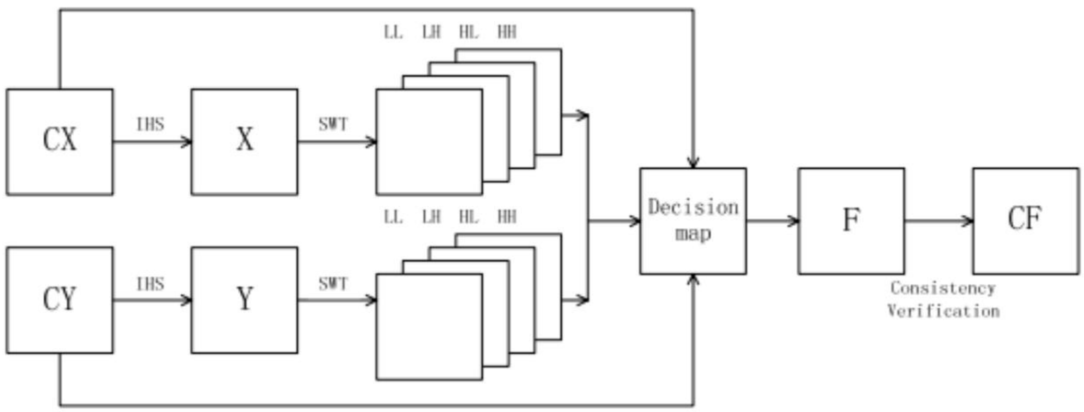


Figure 4. Block diagram of SWT+IHS method

3. Code in Matlab

```
%%Read Images
%The size of images must be equal and 2^N must divide size of image where N is level
of decomposition
[file, pathname] = uigetfile('*.jpg','Select the First Image ');cd(pathname);
a=imread(file);
a = imresize(a, [800 800]);
[file, pathname] = uigetfile('*.jpg','Select the Second Image ');cd(pathname);
b=imread(file);
b = imresize(b, [800 800]);

%%     RGB Component of First
Image
R1=a(:,:,1);G1=a(:,:,2);B1=a(:,:,3);

%%RGB Component of Second Image
R2=b(:,:,1);G2=b(:,:,2);B2=b(:,:,3);

%%  IHS
Transformation
I1=(R1+G1+B1)/3;
I2=(R2+G2+B2)/3;

%%Input from console
prompt = 'Enter the decomposition level : ';
x = input(prompt);
prompt = 'Enter the combination rule : ';
y = input(prompt);

%%     Taking swt of first
image switch x
   case 1
      [a1,h1,v1,d1] =swt2(I1,1,'db4');
   case 2
      [a1,h1,v1,d1] =swt2(I1,2,'db6');
    case 3
      [a1,h1,v1,d1] =swt2(I1,3,'db1');
    case 4
      [a1,h1,v1,d1] = swt2(I1,4,'db4');
end
H1 = wcodemat(h1(:,:,1),255);
V1 = wcodemat(v1(:,:,1),255);
D1 = wcodemat(d1(:,:,1),255);

%%     Taking swt of second
image switch x
```

```
   case 1
      [a2,h2,v2,d2] = swt2(I2,1,'db4');
   case 2
      [a2,h2,v2,d2] = swt2(I2,2,'db6');
   case 3
      [a2,h2,v2,d2] = swt2(I2,3,'db1');
   case 4
      [a2,h2,v2,d2] = swt2(I2,4,'db4');
end
H2 = wcodemat(h2(:,:,1),255);
V2 = wcodemat(v2(:,:,1),255);
D2 = wcodemat(d2(:,:,1),255);
```

```
%%     Combination
Rule switch y
   case 1
      w = [1/9 1/9 1/9;1/9 1/9 1/9;1/9 1/9 1/9];
   case 2
      w = [2/15 2/15 2/15;2/15 2/15 2/15;1/15 1/15 1/15];
   case 3
      w = [4/21 2/21 1/21;4/21 2/21 1/21;4/21 2/21 1/21];
   case 4
      w = [8/35 4/35 2/35;8/35 4/35 2/35;4/35 2/35 1/35];
end
```

```
%%     Calculating activity level and creating decision
map [r,c] = size(R1);
for i=1:r
   for
   j=1:c
      A1(i,j)=sum(sum(w*[H1(i,j);V1(i,j);D1(i,j)]));
      A2(i,j)=sum(sum(w*[H2(i,j);V2(i,j);D2(i,j)]));
      if(A1(i,j)>A2(i,j))
         M(i,j)=1;
      else
      M(i,j)=0;
      end
   end
end
```

```
%%Consistency verification
for i=1:r
   for j=1:c
      s=0; t=0;
      if(i>1)
         s=s+M(i-1,j); t=t+1;
      end
```

```
        if(j>1)
          s=s+M(i,j-1); t=t+1;
        end
        if(i<r)
          s=s+M(i+1,j); t=t+1;
        end
        if(j<c)
          s=s+M(i,j+1); t=t+1;
        end
        if(i>1&&j>1)
          s=s+M(i-1,j-1); t=t+1;
        end
        if(i>1&&j<c)
          s=s+M(i-1,j+1); t=t+1;
        end
        if(i<r&&j>1)
          s=s+M(i+1,j-1); t=t+1;
        end
        if(i<r&&j<c)
          s=s+M(i+1,j+1); t=t+1;
        end
        if(s>t/2)
           M(i,j)=1;
        elseif(s<t/2)
              M(i,j)=0;
        end
    end
end
Mb=ones(r,c)-M;

%%     Fusion of image using decision
map for i=1:r
    for                                          j=1:c
        RN(i,j)=M(i,j)*R1(i,j)+Mb(i,j)*R2(i,j
        );
        GN(i,j)=M(i,j)*G1(i,j)+Mb(i,j)*G2(i,
        j);
        BN(i,j)=M(i,j)*B1(i,j)+Mb(i,j)*B2(i,j
        );
    end
end
%%     Combining RGB component of fused
image F(:,:,1)=RN;F(:,:,2)=GN;F(:,:,3)=BN;
%%
%%Output in window
subplot(2,2,1), imshow(a,[]);
title('Focus on left')
```

```
subplot(2,2,2), imshow(b,[]);
title('Focus on right')
subplot(2,2,3), imshow(F,[]);
title('Fused Image Based on IHS+SWT Method ')
```

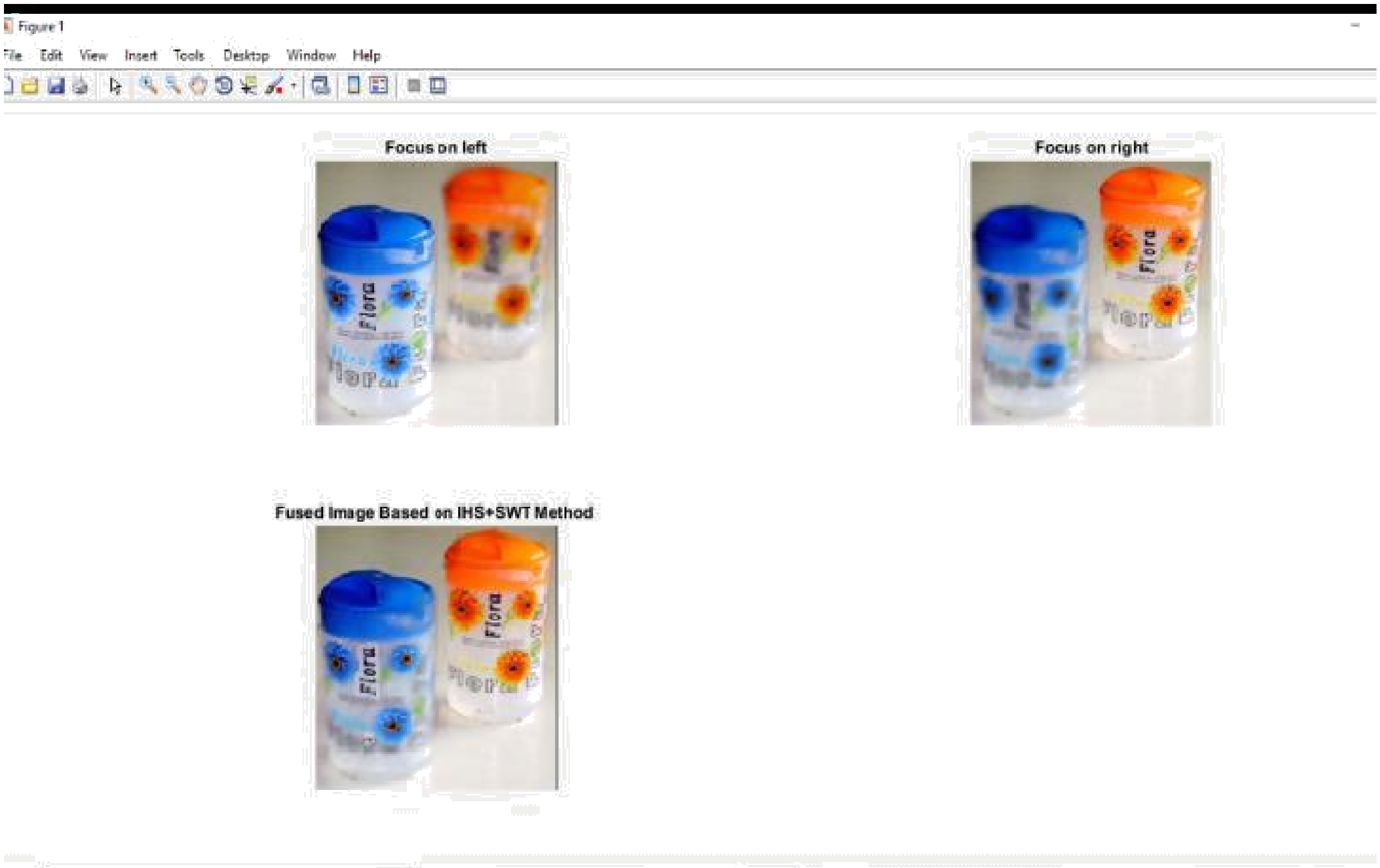
```
%%Experimental Criteria
IN=(RN+GN+BN)/3;
SF = sqrt(sum(sum((IN(1:r-1,2:c)-IN(1:r-1,1:c-1)).^2+(IN(2:r,1:c-1)-IN(1:r-1,1:c-1)).^2,2))/(r-1)/(c-1))
```

4. Examples

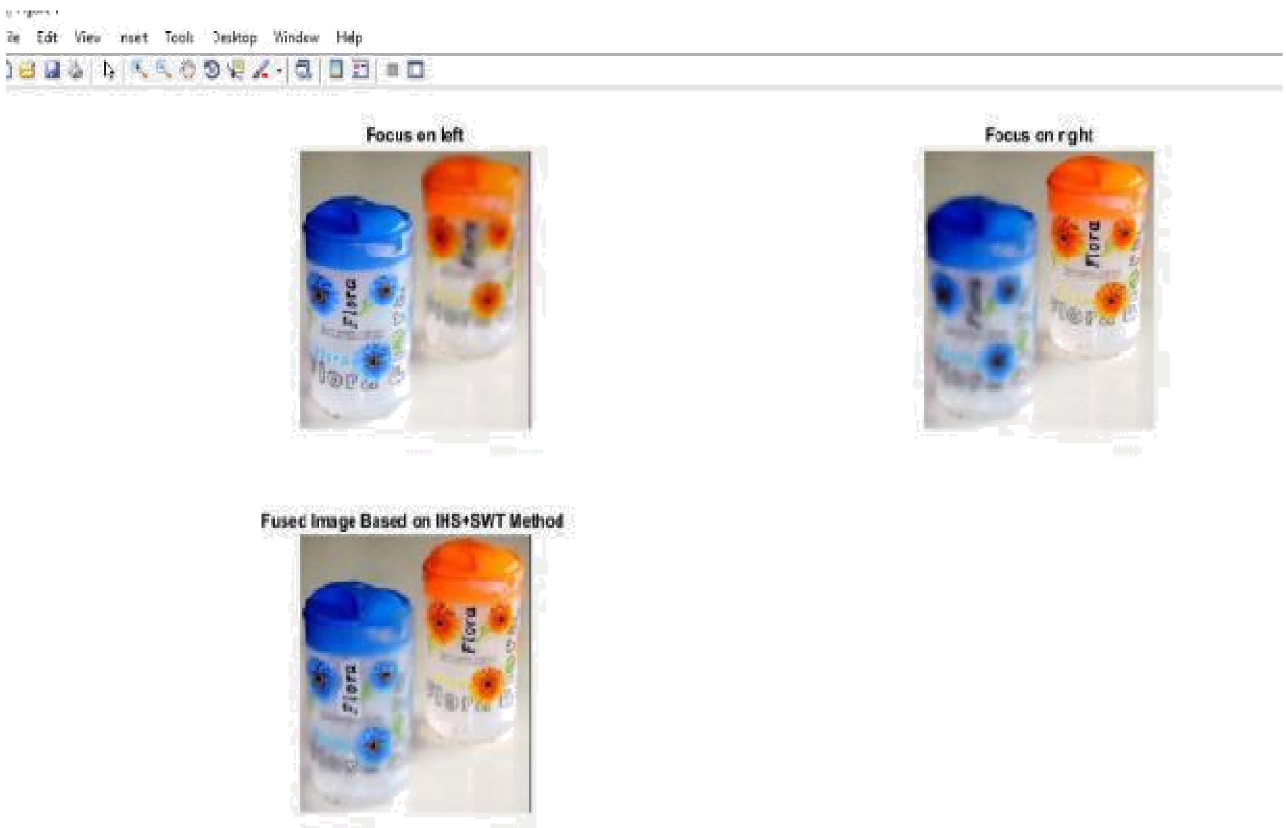
Enter the decomposition level : 1

Enter the combination rule : 2

SF = 2.8746



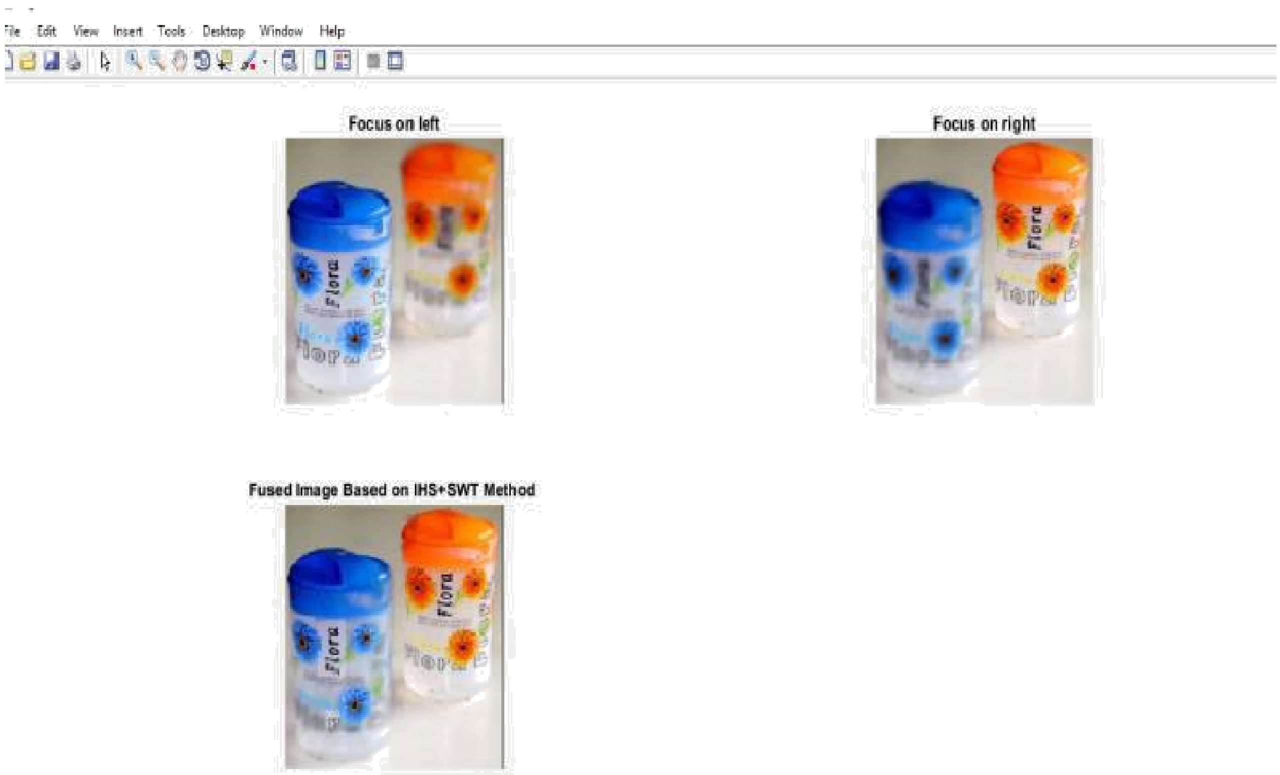
Decomposition level=1
Combination rule=1
SF = 2.1469



Decomposition level=1

Combination rule=4

SF =2.1450



decomposition level : 2

combination rule : 1

SF =2.1377

Enter the decomposition level : 2

Enter the combination rule : 2

SF =6.5254

Enter the decomposition level: 2

Enter the combination rule: 3

SF =6.5249

Enter the decomposition level : 2

Enter the combination rule : 4

SF =6.5272

Enter the decomposition level : 3

Enter the combination rule : 1

SF =6.5750

Enter the decomposition level : 3

Enter the combination rule : 2

SF = 6.5755

Enter the decomposition level : 3

Enter the combination rule : 3

SF =6.5668

Focus on left

Focus on right

Fused Image Based on IHS+SWT Method

Enter the decomposition level : 3

Enter the combination rule : 4

SF =6.5686

Decomposition Level	Combination Rule	Spatial Frequency
1	1	2.8746
1	4	2.8536
2	1	2.8348
2	2	2.8452
2	3	2.8503
2	4	2.8514
3	1	3.0024
3	2	3.0013
3	3	3.0287
3	4	3.0286

Above table displays Spatial Frequency for first example for all decomposition and combination rule

5. CONCLUSION

It is found that with the proposed methods of the fused images are much better than the other two HIS and SWT itself considering all the evaluation criteria.

From the spatial frequency results in previous Table, it can be concluded that in the proposed method, the one with 3 decomposition level and combination rule 3 provides the best performance, and the one with 3 decomposition level and combination rule 4 next. .

Just as the experimental results show that this method cannot overcome the image blur problem completely, methods based on quaternion curvelet transform (qct).

6. BIBLIOGRAPHY

Citing from books:

1. Image Fusion: Principles, Methods, and Applications by Jan Flusser, Filip Sroubek, and Barbara Zitov from Institute of Information Theory and Automation Academy of Sciences of the Czech Republic
2. Zhong Zhang and Rick S.Blum, “A categorization of multiscale-decomposition-based image fusion schemes with a performance study for a digital camera application”,
Proceedings of the IEEE

Citing from Encyclopedia:

1. Britannica Encyclopedia
2. Wikipedia Encyclopedia
3. Microsoft Encarta Encyclopedia

Citing from sits:

1. https://www.mathworks.com/support/learn-with-**matlab-tutorials**.htm
2. https://www.**tutorials**point.com/**matlab**/
3. www.cyclismo.org/**tutorial**/**matlab**