Color Multifocus Image Fusion Using Empirical Mode Decomposition

Slavica Savić, Student member, IEEE and Zdenka Babić, Member, IEEE

Abstract — In this paper, a recently proposed grayscale multifocus image fusion method based on the first level of Empirical Mode Decomposition (EMD) has been extended to color images. In addition, this paper deals with low contrast multifocus image fusion. The major advantages of the proposed methods are simplicity, absence of artifacts and control of contrast, while this isn’t the case with other pyramidal multifocus fusion methods. The efficiency of the proposed method is tested subjectively and with a vector gradient based objective measure, that is proposed in this paper for multifocus color image fusion. Subjective analysis performed on a multifocus image dataset has shown its superiority to the existing EMD and DWT based methods. The objective measures of grayscale and color image fusion show significantly better scores for this method than for the classic complex EMD fusion method.

Keywords — Empirical Mode Decomposition (EMD), Discrete Wavelet Transform (DWT), image fusion, multifocus, low contrast.

I. INTRODUCTION

Usuall y, a scene to be photographed contains objects at different distances from camera, so there isn’t a possibility to acquire the image of that scene in which all the areas are well focused because of the limited camera depth of focus. The camera depth of focus is the range of distance within which objects appear to be sharp in the image. In acquisition process, just objects in the depth of focus are well focused, while objects in front of and behind the depth of focus are blurred. To solve this problem, during the process of image acquisition, camera can be focused on objects one after the other, creating a set of multifocus images.

Multifocus image fusion is a process of combining information of two or more multifocus images of the same scene. It results in a highly informative image that contains more information than any of the orginal images. Combining regions of multifocus images during fusion process, the "all-in-focus" image that contains better focus across all image can be obtained. Because of its ability to integrate information from multiple images into a fused image, those methods have particular importance in many digital image processing applications such as medical imaging, robotics, computer vision, etc.

The fusion process can be performed at different levels of information representation [1]. A common categorization of fusion methods is to distinguish between a pixel, feature or symbol level of fusion. Fusion at the lowest processing level is known as image fusion at pixel level and refers to the merging of measured physical parameters. Fusion at feature level requires extraction of features such as size, shape, contrast and texture. Effectively combining information at the highest level of abstraction means fusion on a symbol level. The choice of the appropriate level depends on different factors such as data sources, application and available tools. In this paper the fusion process at feature level is analyzed.

The remainder of this paper is organized as follows. Section 2 describes previous research work related to EMD and DWT methods for image fusion. In Section 3, grayscale multifocus image fusion based on the first level of EMD decomposition for high and low contrast images is described. Extension of one-level EMD fusion method to color multifocus images is proposed also. Section 4 gives a comparative analysis of the proposed, DWT and EMD multifocus image fusion methods using subjective and objective measures. For multifocus color image fusion a new objective measure is proposed. Conclusions are drawn in Section 5.

II. MULTIFOCUS IMAGE FUSION

DWT is a very popular technique for signal decomposition, so there are many image fusion methods based on DWT [2]-[5]. On the other hand, EMD is an intuitive and direct method for signal decomposition [6], [7], because a signal is represented by its Intrinsic Mode Functions (IMFs).

A. DWT Multifocus Image Fusion

DWT based image fusion selectively combines the decomposed coefficients of the original images in the transform domain using some fusion rule. The basic fusion rule is pixel-by-pixel maximum selection rule [2] which, for fusion process, uses two distinct modes: selection and averaging. Considering that the image useful features usually are determined by area larger than one pixel, in [3] an area-based selection rule is proposed for a pixel chosen process. The variance of each image patch over N x N (3x3 or 5x5) window is computed as an activity measure associated with the centered window pixel. If these measures at the same pixel locations in source images are close to each other, the average of the two is considered as the new value, otherwise the larger value is chosen. Also, the feature selection algorithm based on a maximum absolute value within the window as an activity measure associated with the center pixel is proposed in [4]. After a
pixel chosen process, the inverse DWT applied to the fused coefficient representation gives the resulting “all in focus” image.

DWT based image fusion methods are an integral part of MATLAB’s Wavelet Toolbox. DWT_maxmax and DWT_minmax methods use specific wavelet ‘db2’, with 3 decomposition levels. The method denoted as DWT_maxmax takes a maximum of approximations and details, while the method DWT_minmax takes a minimum of approximations and maximum details from input images. Other, more advanced, DWT based fusion methods can be found in [5].

B. EMD Multifocus Image Fusion

Empirical Mode Decomposition is an appropriate method for decomposing nonlinear and nonstationary signals. EMD adaptively decomposes a signal into a finite set of AM/FM modulated components called Intrinsic Mode Functions [6]. Therefore, EMD is suitable for image fusion process based on a feature level of information representation.

Only a few multifocus image fusion algorithms are based on EMD, because EMD is usually used for multispectral image fusion [8]-[10]. But, some of EMD multifocus image fusion methods can be found in [11],[12]. A combination of EMD and support vector machines (SVMs) is proposed in [11] to produce a better EMD representation of the fused image than any other multifocus images. A solution that uses complex EMD to make multifocus image fusion is proposed in [12]. These authors use all complex EMD decomposition for the image fusion process as follows.

For each A and B image matrices transformation into 1D vectors \( v_1 \) and \( v_2 \) is made by concatenating rows of matrices. Using the complex EMD, the complex vector \( v = v_1 + jv_2 \) is decomposed into \( M \) complex IMFs. Separating real and imaginary components from IMFs and converting each of them into their 2D form, sets of \( M \) scale images \( \{A_i\} \) and \( \{B_i\} \), \( i = 1,2,...,M \), can be obtained, see Fig. 1.

The fused image \( F \) is then given by:

\[
F(x,y) = \alpha_i A(x,y) + \beta B(x,y),
\]

where \( \alpha(x,y) \) and \( \beta(x,y) \) are weighted coefficients satisfying condition \( \alpha(x,y) + \beta(x,y) = 1 \). The values for the coefficients \( \alpha(x,y) \) and \( \beta(x,y) \) at each scale and each location are determined by comparing the local variance:

\[
a) \quad \alpha(x,y) = 0, \quad \text{if} \quad \text{var}(A(x,y)) - \text{var}(B(x,y)) < \varepsilon,
b) \quad \alpha(x,y) = 0.5, \quad \text{if} \quad \text{var}(A(x,y)) - \text{var}(B(x,y)) < \varepsilon,
c) \quad \alpha(x,y) = 1, \quad \text{if} \quad \text{var}(A(x,y)) - \text{var}(B(x,y)) > \varepsilon,
\]

where \( \text{var}() \) denotes the local variance in the \( N \times N \) window centered at \( (x,y) \), and where \( \varepsilon > 0 \).

Because the fusion is performed in a complex domain, we denote this method as Complex_EMD.

III. MULTIFOCUS IMAGE FUSION BASED ON THE FIRST EMD LEVEL

In order to avoid the whole process of decomposition of input multifocus images using complex EMD, the results obtained at all decomposition levels are analyzed in [13]. It is concluded that the first image EMD decomposition level carries the most important information about well focused image regions. This is so because the first IMF contains information about the highest frequencies within the signal that is decomposed. Therefore, in [13] multifocus image fusion based on the first level of the EMD of input images is proposed and denoted as L1_EMD. Unlike the complex EMD fusion method proposed in [11], L1_EMD uses the real domain for fusion process.

As earlier, grayscale images A and B are represented by 1D vectors \( v_1 \) and \( v_2 \), respectively. Using real EMD, the first IMFs of each vector \( v_1 \) and \( v_2 \) are calculated. For each of those components their 2D forms are reconstructed and denoted by \( A_i \) and \( B_i \). Afterwards, mask image \( \alpha \) is created according to (3), where \( i = 1 \), i.e. \( \alpha = \alpha_1 \). The fused image \( F \) is given by:

\[
a) \quad F(x,y) = A(x,y), \quad \text{if} \quad \alpha(x,y) = 0,
b) \quad F(x,y) = B(x,y), \quad \text{if} \quad \alpha(x,y) = 1,
c) \quad F(x,y) = (A(x,y) + B(x,y))/2, \quad \text{if} \quad \alpha(x,y) = 0.5.
\]

A block diagram of this method is given in Fig. 2, while examples of two defocused input images, created image mask \( \alpha \) and fused image using L1_EMD are shown in Fig. 3.

A. Fusion of Low Contrast Images

It is noticed that L1_EMD, as well as other methods, have unsatisfactory results when input images have a narrow histogram, i.e. low image contrast. In that case, the mask image could not be calculated properly, and as a consequence there is a fused image with incorrectly detected regions of good focus. An example of incorrectly generated image mask in fusion process of low contrast multifocus images using the L1_EMD method is shown in Fig. 4.
Fig. 2. Block diagram of L1_EMD fusion method.

(a) Multifocus images.

(b) Mask image.

(c) Fused image.

Fig. 3. Fusion of two multifocus images using L1_EMD.

As a solution of this problem, a modified L1_EMD fusion method is proposed in [14]. Using linear histogram stretching, input images can be preprocessed and good contrast images created. The mask image can be calculated based on these good contrast images according to (3), after which it can be applied on input multifocus images with low or high contrast using (4).

Fig. 5. shows an example of fusion process of low contrast multifocus images using modified L1_EMD fusion method. As can be noticed in Fig 5., mask image is created based on preprocessed images but it can be applied on input multifocus images with low or high contrast. In this way, contrast on all-in-focus image is controlled by end-user or machine. All-in-focus image, that is a result of the L1_EMD fusion based on preprocessed input images, has a better subjective and objective image quality than the all-in-focus image that is a result of the L1_EMD fusion method based on low contrast input images. A result of fusion process with preprocessed images gives better visual representation for end-user.

As a consequence of fusion process at lower decomposition levels, artifacts in fused images obtained with Complex_EMD, DWTmaxmax, and DWTminmax have been observed. One more deficiency of these methods is contrast distortion. In the L1_EMD fusion method, only the first EMD decomposition level is used for creating the mask image. Image fusion is done by combining original multifocus images regions, thus the fused image artifacts and contrast distortion are avoided [13] (see Fig. 6.).

B. Fusion of Color Images

The proposed fusion process can be easily adapted to multifocus color images. Grayscale representation of color images is used for mask image α calculation according to (3), where i=1, i.e., α=α_i. Created mask image in the same way should be applied to all color channels. So, the fused color image F is given by:

a) \( F(x,y,k) = A(x,y,k) \) if \( α(x,y) = 0 \),
b) \( F(x,y,k) = B(x,y,k) \) if \( α(x,y) = 1 \),
c) \( F(x,y,k) = \frac{(A(x,y,k) + B(x,y,k))}{2} \) if \( α(x,y) = 0.5 \),

where \( k = 1,2,3 \) for R,G,B image components, respectively.

Fig. 7. shows an example of the fusion process of two color multifocus images using the L1_EMD method.

The recently published modified L1_EMD for low contrast grayscale images is applicable to color images, also. In contrast to the L1_EMD algorithm, Complex_EMD can't be easily adapted for fusion of color images. The possible solution for Complex_EMD can be to perform fusion on each color channel separately. Using the RGB color space, the pixel colors and textures that do not exist in original images can be obtained in the fused image. Also, there are artifacts as a consequence of image reconstruction from its IMFs. When fusion process is performed with mentioned DWT fusion methods, there are possibilities for artifacts on fused image, also. All of the above mentioned defects are avoided in the proposed algorithm, i.e. the fused color image has no artifacts as a consequence of image reconstruction from its IMFs, there isn't a possibility for additional unnatural pixel colors in
the fused image as a consequence of separated fusion process in different color contrast to the L1_EMD algorithm, Complex_EMD can’t be easily adapted for fusion of color images. The possible solution for Complex_EMD can be to perform fusion on each color channel separately. Using the RGB color space, the pixel colors and textures that do not exist in original images can be obtained in the fused image. Also, there are artifacts as a consequence of image reconstruction from its IMFs. When fusion process is performed with mentioned DWT fusion methods, there are possibilities for artifacts on fused image, also. All of the above mentioned defects are avoided in the proposed algorithm, i.e. the fused color image has no artifacts as a consequence of image reconstruction from its IMFs, there isn’t a possibility for additional unnatural pixel colors in the fused image as a consequence of separated fusion process in different color channels, as well as there is no uncontrolled histogram stretching. Examples of fusion of color multifocus images using all of these methods are shown in Fig. 8.

IV. COMPARATIVE ANALYSIS

The main goal of image fusion is collecting as much information as possible about a scene from more multifocused images. The all-in-focus image has to have the ability to represent the most of important information about the scene of interest. The quality of the resulting image usually is justified with subjective analysis. When subjective quality measurement is not possible or is too expensive, in automated processes for example, the objective analysis of fusion methods is required.

A. Subjective Analysis

To demonstrate the efficiency of the mentioned fusion methods, subjective assessment from 14 users with low expertise in this area is performed. They didn’t know which image is the result of which image fusion method. Fused image is evaluated with a feedback value from 1 to 4, where 1 indicates the best and 4 the worst rank, so an average rank for each compared methods is calculated as:

$$
\text{Average R}_{KP} = \frac{1}{K \cdot P} \sum_{i=1}^{K} \sum_{j=1}^{P} R_{i,j}
$$

where $i$ denotes users’ number, $j$ fused image number and $R_{i,j}$ denotes the $i$-th user’s feedback for the $j$-th fused image. In our experiment, the values of these parameters are: $K=14, P=26$.

The best feedback, as another measure of subjective analysis is calculated as:

$$
BF = \frac{N_{of1}}{N} \cdot 100\%,
$$

where $N_{of1}$ denotes the number of feedback values 1, while $N$ denotes the number of all feedbacks for each of the tested fusion methods, i.e. $N=K \cdot P$.

In the testing process, the same value of the parameter $\varepsilon=2$ and window size $N=3$ for all defocused images are used in the Complex_EMD and L1_EMD methods.

By analyzing the subjective assessment results given in Table 1, we have noticed that the method marked as L1_EMD has the best end-user feedback. The worst feedback has been shown by Complex_EMD fusion method, while the second and third place are occupied by DWT based image fusion methods. Also, from the results obtained in the testing process it is noticed that fused images using the DWT fusion methods have nonexisting textures which isn’t the case in the L1_EMD fusion. Hence, fused images using the L1_EMD fusion method have the best subjective image feedback.

In order to compare the efficiency of the proposed L1_EMD to the other mentioned fusion methods, they are applied to low contrast multifocus images. An example of low contrast multifocus images fusion using the analyzed methods is shown in Fig. 6. As in the case of high contrast multifocus images fusion, there are artifacts in the fused image using Complex_EMD. This is a consequence of fusion process at lower EMD decomposition levels, while the fused images using the DWT fusion methods have uncontrolled contrast distortion. Unlike these methods, the
TABLE 1: RESULTS OF OBJECTIVE ASSESSMENT.

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Dataset images - Average assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complex_EMD</td>
<td>0.77</td>
</tr>
<tr>
<td>L1_EMD</td>
<td>0.87</td>
</tr>
<tr>
<td>DWTmaxmax</td>
<td>0.87</td>
</tr>
<tr>
<td>DWTminmax</td>
<td>0.87</td>
</tr>
</tbody>
</table>

C. Objective Analysis of Color Image Fusion

With the previously described similarity measure proposed in [15] just grayscale image fusion can be evaluated. Therefore an extension of similarity measure for the evaluation of color images fusion method is proposed here. This measure is based on the vector images magnitude gradient proposed in [16] that can be calculated as follows. Vectors $u$ and $v$ are obtained by:

$$u = \frac{\partial R}{\partial x} + \frac{\partial G}{\partial x} + \frac{\partial B}{\partial b}, \quad v = \frac{\partial R}{\partial y} + \frac{\partial G}{\partial y} + \frac{\partial B}{\partial b},$$

where $R, G, B$ are color channels of color image, while $r, g, b$ are their unit vectors.

Scalar components can be calculated as follows:

$$g_{xx} = u \cdot u = u^2 = \left| \frac{\partial R}{\partial x} \right|^2 + \left| \frac{\partial G}{\partial x} \right|^2 + \left| \frac{\partial B}{\partial b} \right|^2,$$

$$g_{yy} = v \cdot v = v^2 = \left| \frac{\partial R}{\partial y} \right|^2 + \left| \frac{\partial G}{\partial y} \right|^2 + \left| \frac{\partial B}{\partial b} \right|^2,$$

$$g_{xy} = u \cdot v = u^T v = \frac{\partial R}{\partial x} + \frac{\partial R}{\partial y} + \frac{\partial G}{\partial x} \frac{\partial G}{\partial y} + \frac{\partial B}{\partial b} \frac{\partial B}{\partial b}.$$
\[
\theta_i = \frac{1}{2} \tan^{-1}\left( \frac{2g_{xx}'}{g_{xx} - g_{yy}'}, \right), \quad i = 1,2,...,n.
\] 

G is created by taking the maximum gradient value from each \( G_i \) at each position \((x,y)\) using (9), and the similarity is calculated using (10).

Table 3 shows the similarity measure between a referent gradient image and the gradient of fused color images using the mentioned fusion methods. By analyzing the results, we have recognized the superiority of the proposed fusion algorithm over the Complex EMD method. As with grayscale images, the fused images obtained with the L1 EMD method have equal objective scores to those of the DWT based fusion methods.

The most significant advantage of the proposed algorithm for color images based on the first level of EMD is that the resulting image has no artifacts which is not the case with other fusion methods. Conversely, the fused image obtained by the DWT based methods can contain textures as a consequence of no ideal input images acquisitions, i.e. multifocus images aren’t aligned pixel by pixel [13]. This problem has no influence on the L1 EMD fusion process.

In some cases the DWT based methods have high objective scores because of the fact that the objective measure recognizes image artifacts as regions of good focus. Hence, there is a need for envisaging a new objective measure for the quality of fusion that will be robust to image artifacts, will have similar feedback as a subjective measure. The fused image using L1 EMD is a result of the combination of original, input multifocus images regions, thus artifacts and contrast distortion are avoided, so it has the best subjective feedback from end-user.

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Color Image Fusion Measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>ComplexEMD</td>
<td>0.74</td>
</tr>
<tr>
<td>L1_EMD</td>
<td>0.89</td>
</tr>
<tr>
<td>DWTmaxmax</td>
<td>0.89</td>
</tr>
<tr>
<td>DWTminmax</td>
<td>0.89</td>
</tr>
</tbody>
</table>

V. CONCLUSION

In this paper, the image fusion methods that use pyramidal decomposition algorithms such as DWT and EMD in the fusion process are analyzed. Applying these methods to low contrast multifocus images is considered. A comparative analysis is performed on an in-house multifocus images dataset. Subjective and objective analysis has shown superiority of the fusion method based on the first level of EMD. Especially, when input images have a narrow histogram, the L1 EMD method shows its superiority and gives better results than other EMD and DWT fusion methods.

A modification of the L1 EMD fusion method for the fusion of color multifocus images, as well as a novel objective fusion measure for color all-in-focus images are proposed in this paper. Objective and subjective analysis of fused color images shows that the proposed color L1 EMD fusion has better assessments than the complex EMD based fusion method.

Fusion method based on the first level of EMD has many advantages such as simplicity and straightforward applicability to color multifocus images unlike the other EMD based fusion method. Fusion is performed in real domain using just the first IMFs which speeds up the image fusion process. Artifacts in the resulting images are avoided. Unlike previous fusion methods, the proposed fusion method controls histogram stretching, i.e. contrast distortion is approved by end-user or machine or not.

REFERENCES