Exposure fusion by Fast and Adaptive Bidimensional Empirical Mode Decomposition

I. Lousis, K. Ioannidis, I. Andreadis
Laboratory of Electronics, Department of Electrical and Computer Engineering
Democritus University of Thrace, GR-67100, Xanthi, Greece
Email: {ilialous, kioannid, iandread}@ee.duth.gr

ABSTRACT
A new method for fusing a scene of two or more different exposed images is proposed. The process of Fast and Adaptive Bidimensional Empirical mode Decomposition is adopted in order to decompose the illumination component of the input images in their Intrinsic Mode Functions (IMFs). The features of each image can then be detected by applying local energy operators within the IMFs and the final fused image is derived as a collection of features from the input images. Finally, the color information is selected from the "best exposed" pixels of the input sequence. Experimental results show that this method captures all the features from the input images while the fused images display uniform pixel values in the entire image region.

KEY WORDS
High Dynamic Range Imaging, Exposure Fusion, Fast and Adaptive Bidimensional Empirical Mode Decomposition (FABEMD).

1. Introduction

Current imaging technology needs to deal with the inefficiency of most digital image sensors to capture the full dynamic range of a scene. The most frequent approach to this problem involves capturing multiple exposures of the same scene [1], and combining them into a single image. Due to different exposure times for each image, this approach captures different details of the same scene. The resulted multi-exposure images are fused into a single image of a higher dynamic range (HDR) followed by a tone mapping process in order to be displayed in a conventional low dynamic range devices such as computer displays.

Several tone mapping operators have been suggested in the past which can be categorized into two main classes. Global tone mapping operators [1] - [4] take into consideration the whole set of image pixels while a linear or nonlinear curve is produced. All pixels are remapped using the resulted curve in order that the image can be displayed in a low dynamic range device. Local tone mapping operators [5] - [7] transform local areas within the image based on specific measures while the pixels are remapped to the available dynamic range. Thus, each pixel could be remapped by a different remapping curve. Compared with the local operators, global operators outperform in terms of speed since the transformation curve is calculated in one single step. On the other hand, local operators produce smoother results to the human perception.

In addition, Mertens et al. [8] introduced an exposure fusion method where the input images are processed in order to extract quality measures like contrast, saturation and well-exposedness. Subsequently, the images are fused directly to a single high quality low dynamic range (LDR) skipping the intermediate step of an HDR image creation which eventually simplifies the entire process. Moreover, Goshtasby in [9] proposed a block level method in order to increase the efficiency of the spatial information exploitation. The required weights for transforming each pixel of the input images are calculated according to the entropy of each image block. Vonikakis et al. [10] proposed a fusion method whose basis relies on illumination estimation. Once illumination is estimated, membership-like functions are applied to define the required weights for the best exposed pixels from each image. Finally, Ahmed et al. [11] applied a modified version of bidimensional empirical mode decomposition (BEMD), namely fast and adaptive bidimensional empirical mode decomposition (FABEMD) [12], in order to achieve multi-focused image fusion.

The FABEMD was adopted since it can detect features such as edges and texture within an image. Such features play an important role in the human visual system and ultimately can be used for the recognition of the "well-exposed" regions within the image. This ensures that the final image will include all the unique features of the input images. Experimental results of this approach show that the proposed method produces comparable or even better results than the existing techniques.

In the proposed technique, the FABEMD method is used for multi-exposure image fusion. Initially, the YCbCr counterpart of each image is computed and the luminance vector is decomposed into multiple Intrinsic Mode Functions (IMFs) by the application of FABEMD method. In addition, based on local energy measurements, the required weights for each IMF are extracted. The summation of the weighted IMFs represents the Y component of the final HDR image. Finally, the color information of the final image is obtained from the "best exposed" pixels of the input images. This yields that the final image will have a purely natural representation of

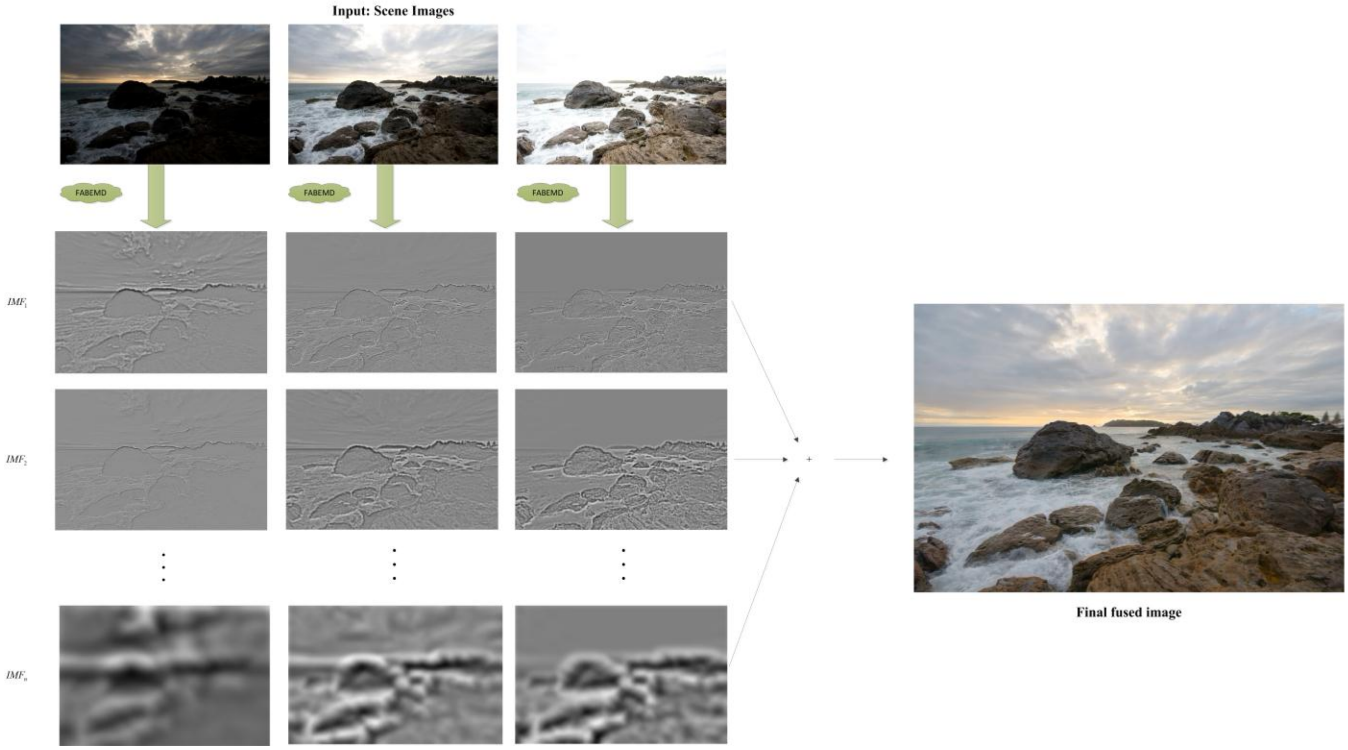

Figure 1. Overview of the proposed method.

colors such as those of the original images, since there is no further processing of the color components. A brief demonstration of the method is presented in Fig. 1. Furthermore, the process is performed only in the luminance component of the images instead of all components as applied in fusion in the RGB domain. Indeed, this is essential in reducing the computationally cost and could be a useful approach in the design of future real time applications.

The rest of the paper is organized as follows. The EMD and the FABEMD methods are briefly described in Section 2. In Section 3, the proposed method is presented in details while in Section 4, experimental results are provided. Finally, conclusions are provided in Section 5.

2. EMD & FABED

Empirical Mode decomposition (EMD) was originally proposed by Huang, in order to overcome the drawbacks of the Hilbert Transform in non-linear and non-stationary signals. The combination of the EMD process with the Hilbert Transform forms the Hilbert-Huang Transform (HHT) [13], a powerful time-frequency data analysis tool. The EMD process decomposes a signal to its basis functions, called IMFs. An IMF represents a function that satisfies the following two conditions:

a. The number of extrema and the number of zero crossings must be equal or differ by at most one in the whole data.
b. The mean value of the envelope defined by the local maxima and the envelope of the local minima must be zero at every point.

If the above two conditions are not fulfilled, then the candidate IMF is processed by an iterative process, called sifting process. The sifting process is repeated until the candidate IMF fulfills the demanded conditions. The processing steps of the EMD process are:

i. Input the original data.
ii. Detect the local extrema (maxima and minima).
iii. Create the upper and lower envelopes by interpolating the maxima and the minima points respectively, and calculate the mean envelope.
iv. Subtract the mean envelope from the original data.
v. Check if the residue of the subtraction fulfills the requirements of an IMF. If so, store the IMF, subtract it from the original data and repeat the above steps for the residue, until the residue either appears one extrema or is a monotonic function. If not, treat the candidate IMF as the original data, and repeat the process until it fulfills the required conditions.

The flowchart of the EMD process is provided in Fig. 2.

Bidimensional empirical mode decomposition (BEMD) is the extension of the EMD process in two dimensional signals, such as images. In both EMD and BEMD, the resulted IMFs strictly depend on the interpolation method that is used in order to construct both the envelopes. Especially in the case of BEMD, traditional ways of interpolation include 2-d spline interpolation, radial basis functions, finite elements methods etc. All these methods suffer in terms of accuracy in regions near the boundary of the image, since few points are determined to support the interpolation process. The effects of these drawbacks are unwanted oscillations in these regions such as overshoots or undershoots, which produce incorrect IMFs. Furthermore, most of the classic interpolation methods are computationally intense. Thus, their continuous application during the sifting process renders the entire algorithm time consuming.

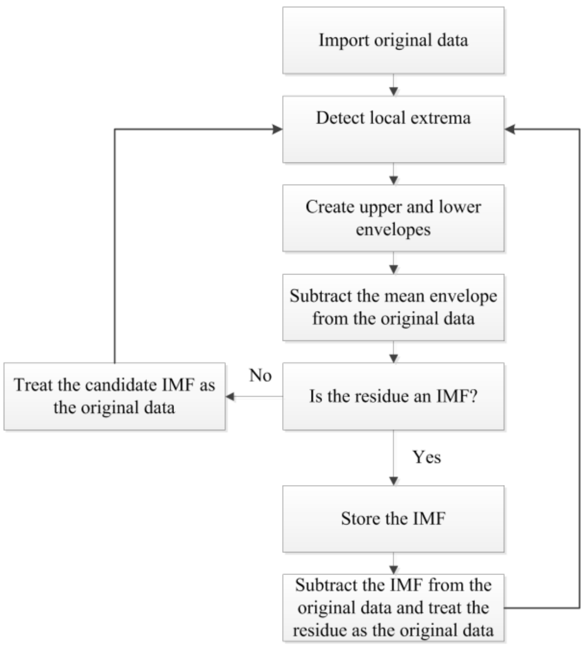

Figure 2. Flowchart of the EMD process.

An efficient algorithm to decompose an image to its IMFs was proposed by Bhuiyan [12] and was called Fast and Adaptive Bidimensional Empirical Mode Decomposition (FABEMD). Instead of applying the classic interpolation methods, the FABEMD algorithm calculates the required envelopes by using order statistic filters. A MAX filter is applied for the upper envelope construction, whereas for the lower envelope a MIN filter is used. Once these filters are applied to the image, smoothing operators are used in order to obtain the desired envelope characteristics.

The most crucial processing step of the FABEMD algorithm is to determine correctly the window size for both the MAX and the MIN filter. The window size is calculated as follows:

i. Detect the maxima and the minima within the image, by comparing each element with its 3×3 neighbors. If its value is higher than the value of the others then consider it as a local maximum point and if its value is higher than the value of the others then consider it as a local minimum point.

ii. For each point determined as extrema, calculate the Euclidean distance between itself and its nearest neighbor. By the end of this step, there will have been determined the maxima and minima maps $M(x,y)$ and $N(x,y)$, respectively.

iii. The window size is calculated as

$$\begin{aligned} d_1 &= \min\{\min\{M\}, \min\{N\}\} \\ d_2 &= \max\{\min\{M\}, \min\{N\}\} \\ d_3 &= \min\{\max\{M\}, \max\{N\}\} \\ d_4 &= \max\{\max\{M\}, \max\{N\}\} \end{aligned} \tag{1}$$

where, min and max denotes the minimum and maximum operators, respectively and the order statistics filters are considered as Type-1, Type-2, Type-3 and Type-4, respectively.

The selection of the filter type relies on the application of the algorithm. The method will produce a bigger number of IMFs when using a filter of Type-1 while, on the contrary, the application of a Type-4 filter will produce the minimum set of IMFs. Once the filtering process is completed, the following operators are applied in order to create smooth and continuous surfaces which represent the required envelopes:

$$\begin{aligned} U_E &= \frac{1}{w \times w} \sum_{(k,l) \in Z_{xy}} U_{Ei}(k,l) \\ L_E &= \frac{1}{w \times w} \sum_{(k,l) \in Z_{xy}} L_{Ei}(k,l) \end{aligned} \tag{2}$$

where, U_{Ei}, L_{Ei} are the upper and lower envelopes respectively, Z_{xy} denotes a square region $w \times w$ centered at any point (x,y) of U_{Ei} and L_{Ei}, and U_E, L_E are the smoothed upper and lower envelopes.

The sifting process is excluded and thus, a single iteration is required. The entire process is repeated until the limit of six extrema or less is reached. The initial image can be expressed as:

$$I(x,y) = \sum_{k=1}^{n} IMF_k(x,y) + R(x,y) \tag{3}$$

where, n denotes the total number of the produced IMFs and $R(x,y)$ is the residue signal. The flowchart of the FABEMD process is demonstrated in Fig. 3.

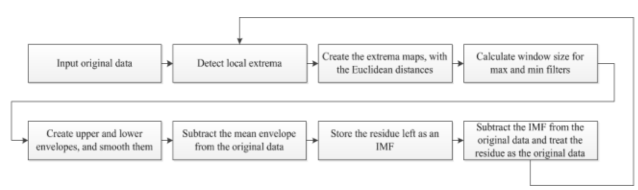

Figure 3. Flowchart of the FABEMD algorithm.

3. Exposure fusion via FABEMD

Exposure fusion comprises a set of techniques where images with different exposure times are fused avoiding the step of creating an HDR image [8]. The images are directly fused into a high quality LDR image which contains the best exposed details from the input image sequence. A set of image fusion approaches includes the decomposition of the original images into their fundamental components and the fused image is reconstructed by appropriate processing of the resulted components. A representative approach for such type of image fusion includes the application of wavelet transform [14], [15]. Hariharan in [16] used the BEMD algorithm in order to fuse images which were captured from different sensors (thermal and visual images). The method produces visually better results than the wavelet decomposition approach as well as than similar image fusion methods.

The main drawback that may occur by the application of the BEMD method is that the process may result a different number of IMFs for each image and thus, their scales cannot be matched correctly. Looney et al. in [17] used complex extensions of the EMD process in order to decompose two images into the same number of IMFs. The FABEMD process was proven to be efficient in decomposing two different images into the same number of IMFs by correctly choosing the window of the order statistic filters. The FABEMD is applied to each RGB channel of the input images and each set of IMFs is fused to create the IMFs of the RGB channels of the fused image. The nature of these images is suitable for applying the FABEMD method to the RGB domain since they contain the same color in the corresponding pixels, and thus the color of the fused image will be exactly the color of the initial images.

Nevertheless, in case of multiple exposure images, each image contains different chromatic information in multiple image regions, and thus fusion via FABEMD in the RGB domain fails to produce satisfactory results. To overcome this issue, the method initially transforms the input images from the RGB to the YCbCr color space in order to keep the computational burden at low levels and exploit the luminance vector at the next processing step.

After the color space conversion, the FABEMD process is applied to the Y-channel of each image. The size of the order statistics filter is calculated by the formula:

$$w = \max\{\max\{d_1\}, \max\{d_2\}, \ldots, \max\{d_i\}\} \quad (4)$$

where, i denotes the number of the input images and d denotes the type of the applied filter.

The proposed method utilizes a Type-4 filter for the decomposition process, as discussed Section 2. In addition, since the same filter size is used for all input images, the same number of IMFs, n, will be produced. This yields that the Y-component of the fused image can be reproduced by the weighted summation of the resulted n IMFs. The resulted residues of each image are processed as the corresponding IMFs. The IMFs of the fused image are extracted by weighting the local amount of information that each IMF carries:

$$IMF_{fused}(x,y) = \frac{\sum_{i=1}^{k}\left(\alpha_i \times IMF_i(x,y)\right)}{\sum_{i=1}^{k}\alpha_i} \quad (5)$$

where, (x,y) denotes the position of each pixel, k is the number of the input images and α_i are the coefficients that are equal with the energy included in the 3 × 3 neighborhood of the (x,y) pixel.

The energy metric was selected due to its feature in defining the contained information of a specific region. The above process is repeated n+1 times, because there are n-IMFs plus the residue of each image. In addition, in order to increase the visual quality of the final fused image, a local weighting approach was adopted meaning that the applied weighting factor is proportional to the calculated energy within the neighborhood of each pixel. The applied region dimensions were selected to be 3 × 3 pixels for best optical results and speed issues.

The summation of the resulted weighted IMFs corresponds to the Y component of the final image. The maximum available dynamic range of the Y channel requires that its values will be in the interval [16,236]. In order to guarantee that the fused image will take advantage of the full dynamic range, the following linear operator is applied:

$$Y_{out} = 16 + \frac{Y - Y_{\min}}{Y_{\max} - Y_{\min}}(235 - 16) \quad (6)$$

where, Y_{out} and Y are the stretched values and the initial values of the Y channel, respectively. Furthermore, Y_{max} and Y_{min} are the maximum and minimum values of the initial Y channel, respectively. The purpose of this operator is to place the resulted values in the interval [16,236]. Since the Y channel processing step is completed, the appropriate color values are calculated. A pixel is defined as well-exposed only if its luminance value is approximately equal to 128, meaning that its luminance is close to the median value of the values in the interval [0,255]. The color information of each fused pixel (Cb and Cr channels) is selected as the color information (Cb and Cr channels) from the input pixel whose luminance is closest to the 128 value. The chromatic information for each fused pixel is calculated through the following formula:

$$\begin{gathered} C_i^{fused}(x,y) = C_i(x,y) \\ i = \arg\min(abs(128 - Y_1), abs(128 - Y_2), \ldots, abs(128 - Y_k)) \end{gathered} \quad (7)$$

where, C_i corresponds to either the Cb or the Cr chromatic vector, x and y are the pixel coordinates, i is the index where the difference becomes minimum and *argmin* is the argument of minimum function. The same process is applied in both Cb and Cr chromatic vectors.

Finally, a YCbCr to RGB color space transformation is performed in order to display or save the final fused image to a compatible electronic equipment or an appropriate image format, respectively.

4. Experimental Results

In order to study the results of the proposed method, the produced images were quantitatively compared with other related approaches, namely with Mertens et al. method [8], Vonikakis et al. method [10], the Essential HDR [18] and the Photomatix [19]. Mertens et al. method is a classic exposure fusion algorithm. Vonikakis et al. algorithm is a very recent algorithm in the specific field and the Essential HDR and Photomatix are commercial software products. The latter provide the user the option to enhance further the produced image, by controlling the local contrast, color saturation and other parameters. In order to have an as objective comparison as possible, the images compared in this section are the default fused images, without further enhancement. The Root Mean Square Error (RMSE) of a well-exposed region (defined by observing the images) in the initial image sequence is calculated in each case, in order to have an objective comparison among them all. For example, Fig. 4 depicts the “Church” scene.

Results reveal that in RMSE terms, the proposed method produces comparable numerical results to the Vonikakis et al. approach, and outperforms the rest of the other methods. Moreover, it can be observed that the produced images lack of overexposed or underexposed regions. It is also clear that the proposed method perceives a pure natural representation since it does not involve any color processing. The advantage of the proposed method in preserving the original colors is better illustrated by observing Figs. 5 and 6.

More specifically, Fig. 5 depicts the “Venice” scene, where the proposed method produces similar results in terms of the RMSE metric with the corresponding results of Vonikakis et al. approach. However, it can be observed that the proposed method preserves the color information of the fused images better than all the other methods. This is due to the fact that the proposed method uses the color information of the initial images. No further processing is carried out and thus, the colors of the fused image are the same as the colors of the original images. This can be observed in the right column of Fig. 5. The proposed method captures the most vivid red color, which is closer to the original red color.

Table 1. Summary of the RMSE metric results.

Algorithm	**Image scenes**		
	Church	**Venice**	**Coast**
Essential HDR	49.27	76.27	145.07
Photomatix	47.18	95.05	114.89
Vonikakis et al.	16.87	65.05	113.75
Mertens et al.	82.13	96.82	96.08
Proposed method	19.08	75.44	91.19

Figure 4. Comparison for the “Church” scene.

Figure 5. Comparison for the "Venice" scene.

In addition, the advantage of the proposed method in perceiving the chromatic information is clearly demonstrated in Fig. 6. In this case, the proposed method outperforms all the other methods in RMSE metrics. Moreover, by observing the right column of the figure (zoomed areas), it is clear that the proposed method produces the best result as long as the colors contained are concerned. These areas contain information both from Original (1) and Original (3) images. This makes it really difficult to compare these regions in terms of RMSE, and thus no RMSE metrics are included.

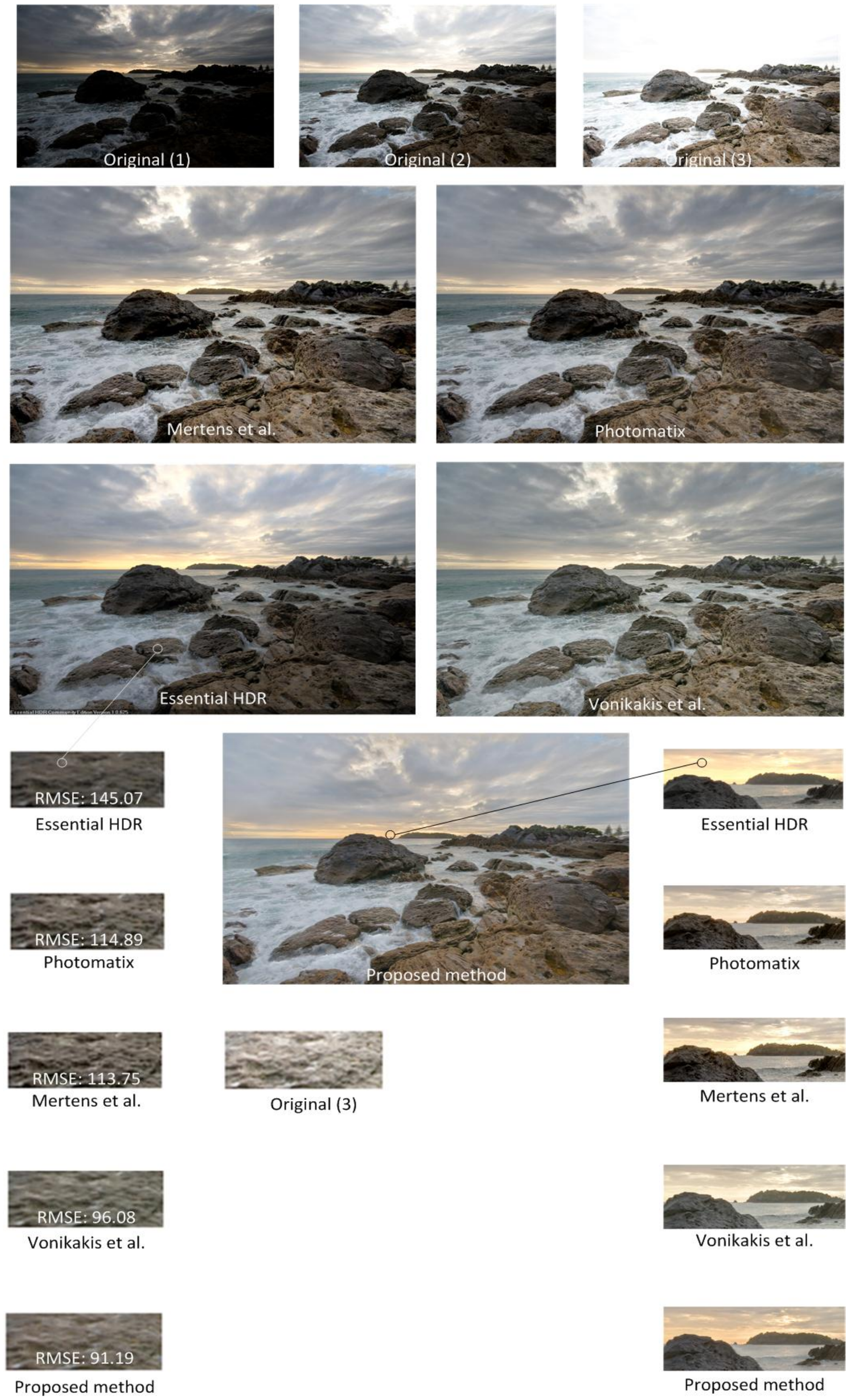

Figure 6. Comparison for the "Coast" scene.

The performance of all the compared algorithms, as long as the RMSE metric is concerned, is summarized in Table 1. From the experimental results it is clear that the proposed method produces comparable or even better results than the rest of the methods. The resulted fused images display the best exposed image regions, since the spectral content of each input image is used. The main advantage of the proposed method is that it processes regions which display significant difference in edge or texture content, instead of fussing an entire image sequence. Furthermore, compared to other similar methods, it has a clear advantage in preserving the color information, due to the fact that the colors are simply selected from the initial images, without further processing.

5. Conclusions

In this paper, a novel method of fusing multiple exposed images is presented. The proposed approach is based on the FABEMD algorithm which is used to decompose the luminance vector of the original images into a number of sub-images, called IMFs. The resulted images provide a useful distribution of information within the images. In contrary with the majority of the reported methods that use the illumination values to define whether a region is overexposed or underexposed, the proposed approach detects such regions due to the definition of spectral information. The calculated IMFs are then fused to extract the corresponding IMFs of the final image by using proportional weights to the local energy of the neighborhood of each pixel. Subsequently, the color information is selected from the “best exposed” pixels of the image sequence in order to create the final HDR image.

Experimental results reveal that the proposed method produces comparable or even better results than other similar methods in terms of RMSE while the produced images display a more realistic appearance of the colors. Furthermore, the most computationally intense part of the processing is being inducted in only one component (Y-channel) of the initial images.

References

[1] E. Reinhard, G. Ward, S. Pattanaik, and P. Debevec. *High dynamic range imaging: Acquisition, display and image-based lighting* (Amsterdam: Elsevier/Morgan Kaufmann, 1st edition, 2006).

[2] F. Drago, K. Myszkowski, T. Annen, and N. Chiba. Adaptive logarithmic mapping for displaying high contrast scenes, *Computer Graphics Forum, 22*(3), 2003, 419-426.

[3] C. Schlick, Quantization techniques for the visualization of high dynamic range pictures, *P. Shirley, G. Sakas, and S. Muller (eds.), Photorealistic Rendering Techniques*, New York: Springer-Verlag, 1994, 7-20.

[4] G. W. Larson, H. E. Rushmeier, and C. D. Piatko. A visibility matching tone reproduction operator for high dynamic range scenes. *IEEE Trans. Vis. Comput. Graphics, 3*(4), 1997, 291–306.

[5] F. Durand and J. Dorsey. Fast bilateral filtering for the display of high-dynamic range images, *ACM Trans. Graphics, 21*(3), 2002, 257–266.

[6] R. Fattal, D. Lischinski, and M. Werman, Gradient domain high dynamic range compression, *ACM Trans. Graphics, 21*(3), 2002, 249–256.

[7] L. Meylan, and S. Süsstrunk, High dynamic range image rendering with a retinex-based adaptive filter, *IEEE Trans. Image Process., 15*(9), 2006, 2820–2830.

[8] T. Mertens, J. Kautz and F. Van Reeth, Exposure fusion, *Proc. 15th Pacific Conf. on Computer Graphics and Applications*, 2007, 382 – 390.

[9] A. A. Goshtasby, Fusion of multi-exposure images, *Image and Vision Computing, 23*(6), 2005, 611 – 618.

[10] V. Vonikakis, O. Bouzos and I. Andreadis, Multi-exposure image fusion based on illumination estimation, *Proc. IASTED Int. Conf. on Signal and Image Processing and Applications*, Crete, Greece, 2011, 135 – 142.

[11] M.U. Ahmed and D.P. Mandic, Image fusion based on fast and adaptive bidimensional empirical mode decomposition, *Proc. 13th Conf. on Information Fusion*, 2010, 1 – 6.

[12] S.M.A. Bhuiyan, R.R. Adhami, and J.F. Khan, Fast and adaptive bidimensional empirical mode decomposition using order-statistics filter based envelope estimation, *EURASIP Journal on Advances in Signal Processing 164*, 2008, 1-18.

[13] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu, The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis, *Proc. R. Soc. Lond. A, Math. Phys. Sci. 454*(1971), 1998, 903 – 995.

[14] M. H. Malik, S. A. M. Gilani and Anwaar-ul-Haq, Wavelet based exposure fusion, *Proc. of World Congress on Engineering*, 2008, 688 – 693.

[15] T. Stathaki, *Image fusion: Algorithms and applications* (Academic Press, 2008).

[16] H. Hariharan, A. Gribok, M. Abidi, and A. Koschan, Image fusion and enhancement via empirical mode decomposition, *Journal of Pattern Recognition Research, 1*(1), 2006, 16 – 32.

[17] D. Looney and D. P. Mandic, Multiscale image fusion using complex extensions of EMD, *IEEE Trans. Signal Process., 57*(4), 2009, 1626 – 1630.

[18] http://www.imagingluminary.com

[19] http://www.hdrsoft.com/