DEMPSTER-SHAFER AND ALPHA STABLE DISTANCE FOR MULTI-FOCUS IMAGE FUSION

Rachid Sabre¹ and Ias Sri Wahyuni²

¹Laboratory Biogéosciences CNRS, University of Burgundy/Agrosup Dijon, France
²Faculty of Computer Science and Information Technology, Universitas Gunadarma, Indonesia

ABSTRACT

The aim of multi-focus image fusion is to integrate images with different objects in focus so that obtained a single image with all objects in focus. In this paper, we present a novel multi-focus image fusion method based using Dempster-Shafer Theory and alpha stable distance. This method takes into consideration the information in the surrounding region of pixels. Indeed, at each pixel, the method exploits the local variability that is calculated from quadratic difference between the value of pixel I(x,y) and the value of all pixels that belong to its neighbourhood. Local variability is used to determine the mass function. In this work, two classes in Dempster-Shafer Theory are considered: blurred part and focus part. We show that our method give the significant result.

KEYWORDS

Multi-focus-images, Dempster-Shafer, distance, Alpha-Stable.

1. INTRODUCTION

Image fusion is the technique of combining relevant information from multiple images to produce a single image that contains more information than the input images. The goal of image fusion is to reduce uncertainty and minimize redundancy on the output as well as maximize relevant information specific to an application or task. In this article, we deal with merging multifocal images. Due to the limited depth of field of optical senses in cameras, it is often not possible to obtain an image containing all relevant objects "in focus". So that a scene image can be taken from a set of images with different focus. Image fusion method is used to get all focus objects.

There are different approaches of multifocal image fusion techniques that have been performed in the literature. These approaches can be divided into two types, the spatial domain method and the multi-scale fusion method. The spatial domain fusion method is performed directly on the source images. In spatial domain techniques, we work directly on the pixels of the image. Fusion methods such as averaging, principal component analysis (PCA) [1], maximum selection rule, two-sided gradient based methods [2], guided image filter based method (GIF) [3] and the maximum selection rule fall under spatial domain approaches. The disadvantage of spatial domain approaches is that they produce spatial distortion in the merged image. Spatial distortion can be very well managed by multi-scale approaches on image fusion. In multi-scale blending methods, the blending process is performed on the source images after decomposing them into multiple scales. Discrete wavelet transform (DWT) [4]-[7], Fusion of Laplacian pyramidal
images [8]-[14]. Discrete cosine transform with variance calculation (DCT+var) [15], method based on the Salience detection (SD) [16] are examples of domain transformed image fusion techniques.

As stated in [17], from the point of view of proof, fusion degrades imprecision and uncertainty by using redundancy and complementary information from the source image. This means that the weak evidence of the inputs is used to give the best estimate. The proof theory was first proposed by Shafer in the 1970s, based on Dempster’s research. The advantage of the Dempster-Shafer theory (DST) is that it allows to deal with the lack of preference, due to the limitations of available information, which leads to indeterminacy, as in [18] and [19]. This theory has been successful in many applications, including image segmentation [20], [12], pattern classification [22]-[24], object recognition [25], imaging technology [26], sensor fusion [27], [28].

In this article, we use Dempster Shafer's theory, which has been successful in various image-processing methods. It is based on the plausibility and the weight of dependence of each pixel from a well-chosen distance. We propose the fusion of multi-focal images using the Dempster-Shafer theory, which derives from information: the variability of each pixel with its neighbourhood. This variability is calculated from the stable distance alpha between the value of the pixel I(x,y) and the value of all the pixels belonging to its neighbourhood. The stable alpha distance (alpha is between 0 and 2) generalizes the quadratic distance. This distance was introduced at the time of the discovery of stable alpha stochastic variables and processes [29], [30]. Several works have shown that working in a stable alpha space can improve the estimation and visibility of certain phenomena whose variability increases significantly [31]-[33]. The alpha-stable distribution is widely used in the processing of impulsive or spiky signals. It also has been applied in image processing field. [34] Models the sea clutter in SAR images using alpha stable distribution for ship detection while [35] removes speckle noise using alpha stable based Bayesian algorithm in the wavelet domain. Furthermore, alpha stable distribution is also used in image segmentation [36] and compressive image fusion [37] and alpha stable filter in fusion image [38]. Both [35], [36], and [36] and Wan employ alpha stable in wavelet domain. This section provides a brief of the alpha-stable distribution.

Thus, the stable alpha distance measures local variability around a pixel. This distance taken as an activity measure can detect the abrupt intensity of the image such as the edge. This method also takes into consideration the information in the surrounding region of the pixels and preserves the edge.

The originality of this work lies in the fact of combining Dempster Shafer method with the alpha stable distance adapted to large variations. Thus, we propose a new method that we compare to other existing methods in the literature and we show that it gives better fusion results.

This article is organized as follows: in section 2, we detail the main elements of the Dempster-Shafer proof theory. The definition of the stable distance alpha and the proof are presented in section 3. Section 4 provides the details of the proposed method. Section 5 defines the evaluation measures used in this article. Experiments are performed on different types of images and the results are compared with other works are provided in section 6. Section 7 gives the conclusion of this work.

2. **Dempster-Shafer Evidence Theory**

Let Θ represent a finite set of hypotheses for a problem domain, called frame of discernment.
Define a function \( m \) from \( 2^\varnothing \) to \([0,1]\) where \( 2^\varnothing \) be the set of all subsets of \( \varnothing \)
\[
2^\varnothing = \{ A | A \subseteq \varnothing \}. \tag{1}
\]
The function \( m \) is called a basic probability assignment whenever
\[
m(\varnothing) = 0 \text{ and } \sum_{A \subseteq \varnothing} m(A) = 1. \tag{2}
\]
\( m(A) \) is the measure of the belief that is committed exactly to \( A \). According to [39], \( m(A) \) is the degree of evidence supporting the claim that a specific element of \( \varnothing \) belongs to the set \( A \), but not to any special subset of \( A \). Each \( A \) of \( \varnothing \) such that \( m(A) > 0 \) are called the focal element of \( m \). By applying the basic assignment function, several evidential functions can be created. A belief measure is given by the function \( \text{Bel}: 2^\varnothing \mapsto [0,1] \):
\[
\text{Bel}(A) = \sum_{B \subseteq A} m(B). \tag{3}
\]
The plausibility measure \( \text{Pl}: 2^\varnothing \mapsto [0,1] \) is defined by [28] as follows:
\[
\text{Pl}(A) = \sum_{A \cap B = \varnothing} m(B) = 1 - \text{Bel}(\bar{A}). \tag{4}
\]
\( \text{Bel}(A) \) measures the degree of evidence that the element in question belongs to the set \( A \) as well as to the various special subsets of \( A \). In stated in [17], an important aspect of DST concerns the aggregation of evidence given by different sources. If two mass function \( m_1 \) and \( m_2 \) induced by distinct items of evidence are such that \( m_1(B) > 0 \) and \( m_1(C) > 0 \) for some non-disjoint subsets \( B \) and \( C \) of \( \varnothing \), then they are combinable by means of Dempster’s rule. [41], [42] followed by [40] suggested a rule of combination which allows that the basic assignments are combined. The combination (joint mass) of two sets of masses \( m_1 \) and \( m_2 \) is defined as follows
\[
m_1 \oplus m_2(\varnothing) = 0 \tag{5}
\]
\[
m_1 \oplus m_2(A) = \frac{\sum_{B \cap C = A} m_1(B) m_2(C)}{1 - \sum_{A \cap B = \varnothing} m_1(B) m_2(C)}. \tag{6}
\]
The numerator represents the accumulated evidence for the sets \( B \) and \( C \), which supports the hypothesis \( A \) and the denominator sum quantifies the amount of conflict between the two sets. Equation (6) can be written as
\[
m_1 \oplus m_2(A) = \frac{\sum_{B \cap C = A} m_1(B) m_2(C)}{\sum_{B \cap C = \varnothing} m_1(B) m_2(C)}. \tag{7}
\]
As stated in [43], having a zero mass on a subset \( A \) does not mean that the set is impossible, simply that we are not capable of assigning a level precisely to \( A \), since we could have non-zero masses on subsets of \( A \), which would lead us to \( \text{Bel}(A) \neq 0 \).

3. **ALPHA STABLE DISTANCE**

Our method takes into consideration the information in the surrounding region of pixels. Indeed, at each pixel \( I(x,y) \), the method exploits the local variability calculated from alpha stable distance between the value of pixel \( I(x,y) \) and the value of all pixels that belong to its neighborhood. The idea comes from the fact that the variability value in blurred region is smaller
than the variability value in focused region, the proof is provided in this section. We use in this work the neighbor, with the size "a", of a pixel \((x, y)\) defined as follows:

\[(x + i, y + j)\] where \(i = -a, -a + 1, \ldots, a - 1, a\) and \(j = -a, -a + 1, \ldots, a - 1, a\). For example the neighbor with the small size \((a = 1)\) contains: \((x - 1, y - 1)\), \((x - 1, y)\), \((x - 1, y + 1)\), \((x, y - 1)\), \((x, y + 1)\), \((x + 1, y - 1)\), \((x + 1, y)\), \((x + 1, y + 1)\) as we can see in Fig. 1.

We consider \(p\) source images \((I_1, I_2, \cdots, I_p)\) where each image has size \((R \times C)\). Alpha stable distance of every source image at pixel \((x, y)\):

\[
v_{a,k}(x, y) = \left(\frac{1}{T} \sum_{m=-a}^{a} \sum_{n=-a}^{a} |I_k(x, y) - I_k(x + m, y + n)|^a\right)^{1/\alpha} \tag{8}
\]

where \(k\) is the index of \(k^{th}\) source image \((k = 1, 2, \cdots, p)\) is the number of source images.

\[
I_k(x + m, y + n) = \begin{cases} 
I_k(x + m, y + n), & \text{if } 1 \leq x + m \leq R \text{ and } 1 \leq y + n \leq C, \\
I_k(x, y), & \text{otherwise}
\end{cases}
\]

\[
T = (2a + 1)^2 - \text{card}(S)
\]

\[
S = \{(m, n) \in \{[-a, a]^2 \setminus \{0,0\}\} \text{ such that } I_k(x + m, y + n) = I_k(x, y)\}
\]

We show in the following that this local variability is small enough where the location is on the blurred area \((B_1 \cup \partial B_2)\). Indeed, we consider, without loss the generality, that we have a focus pixel \((x, y)\) in image \(I_1\) and blurred in image \(I_2\), \(((x, y) \in B_2)\)
The local variability of image $I_1$ and image $I_2$ are respectively: $(\frac{1}{r_1(x,y)})^{1/\alpha}$ and $(\frac{1}{r_2(x,y)})^{1/\alpha}$, where $r_1(x,y)$ and $r_2(x,y)$ can be written as follow:

$$r_1(x,y) = \sum_{m=0}^{2\alpha} \sum_{n=0}^{2\alpha} |I_1(x,y) - I_1(x+(m-a),y+(n-a))|^{\alpha}$$

$$r_2(x,y) = \sum_{m=0}^{2\alpha} \sum_{n=0}^{2\alpha} |I_2(x,y) - I_2(x+(m-a),y+(n-a))|^{\alpha}$$

**Proposition**

Let $(x,y)$ a pixel belongs to blurred area of the image $I_2$ ($(x,y) \in B_2$), then the local variability of $(x,y)$ in image $I_2$, is smaller that the local variability of $(x,y)$ in image $I_1$, $r_2(x,y) < r_1(x,y)$.

**Proof**

The proof of this is given by using same arguments in [46].

The variability of image expresses the behavior of pixel relative to all pixels belong to their neighborhood. The precision of this fusion is depending on the size of the neighborhood, "$\alpha$". For each image we try with different values of "$\alpha$" in the set {1,2,…,10} and we get the value of "$\alpha$" that corresponds to the minimum of root mean square error (RMSE). This operation is repeated for set of 150 multi-focus images [35].

**4. THE PROPOSED METHOD**

One of the essential problems of image fusion using Dempster-Shafer Theory is to construct the evidential representation of images. In this paper, we use one information as the evidential representation images: local variability. We consider two classes in the Dempster-Shafer theory. Either a pixel belongs to blurred part $\omega$ or it belongs to the focus part $\overline{\omega}$. There is also uncertainty $\Theta$ inherent in the theory of evidence. All this constitute the frame of discernment in $\Theta$ our case [20].

$$\Theta = \{\omega, \overline{\omega}, \Theta\}$$

For each pixel one value of evidence for information will be obtained, $m$.

$$\{m(\omega), m(\overline{\omega}), m(\Theta)\}$$

with the condition $m(\omega) + m(\overline{\omega}) + m(\Theta) = 1$.

The steps of image fusion in this work as follows. Suppose there are $p$ original source images, $I_1, I_2, \ldots, I_p$, where each image has size $(R \times C)$ with different focus to be fused.

**Step 1:**

1. To calculate mass function:

   For each image where we use different values of size of neighborhood, $\alpha \in \{1,2,\ldots,10\}$, we define: $d'_{\alpha,k}(x,y)$
where \( k \) is the \( k \text{th} \) source image, \( k \in \{1, 2, \cdots, p\} \) and \( \alpha \) is size of neighborhood of local variability. We set the standard deviation of \( d'_{a,k}(x, y) = \sigma_{a,k}(x, y) \),

for \((x, y)\) belongs to \( \omega \), we calculate:

\[
m_{a,k}(\omega) = (1 - \sigma_{a,k}(x, y))d'_{a,k}(x, y)
\]  

(27)

for \((x, y)\) belongs to \( \theta \), we calculate:

\[
m_{a,k}(\theta) = \sigma_{a,k}(x, y)
\]  

(28)

for \((x, y)\) belongs to \( \overline{\omega} \), we calculate:

\[
m_{a,k}(\overline{\omega}) = 1 - \left(1 - d'_{a,k}(x, y)\right)\sigma_{a,k}(x, y) - \sigma_{a,k}(x, y)
\]

\[
= (1 - d'_{a,k}(x, y))(1 - \sigma_{a,k}(x, y))
\]  

(29)

The final result of this method is obtained by showing which pixels belong to focus area or which do not, we use concept plausibility. In our case the plausibility of \( \omega \) is the sum of the masses of the evidence for \( \omega \) and the uncertainty \( \theta \):

\[
PL_{a,k}(\omega) = m_{a,k}(\omega) + m_{a,k}(\theta)
\]

And for fusion image of the pixel \((x, y)\) due to \( \omega \) is a set of pixel on blurred area, we take pixel \((x, y)\) from image \( k_0 \) that assigned to minimum \( PL_{i,k}(\omega), k = 1, 2, \cdots, p \).

**Step 2.**

For \((x, y)\), we take \( F_\alpha \) as fused image with size of neighborhood = \( \alpha \)

\[
F_\alpha(x, y) = I_{k_0}(x, y), \text{where } k_0 \in \{1, 2, \cdots, p\} \text{ and } PL_{a,k_0}(\omega)(x, y) = \min_{k \in \{1, 2, \cdots, p\}} (PL_{a,k}(\omega)(x, y)).
\]

**Step 3.**

For the proposed method, we use different values of size of neighborhood, \( \alpha \in \{1, 2, \cdots, 10\} \), and choose the value of \( \alpha \) that corresponds to the minimum value of RMSE, such that our final fused image

\[
F = F_{a_3} \text{ where } a_3 \in \{1, 2, \cdots, 10\} \text{ and } \text{RMSE}(F_{a_3}) = \min_{\alpha \in \{1, 2, \cdots, 10\}} (\text{RMSE}(F_\alpha))
\]

**5. Experimental Result**
The images used in this section are taken from the database of the webpage [47]. We have blurred an area of each image using the convolution of Gaussian filter applied on the reference image. The choice of Gaussian is approved in the works [44]-[45]. Blurred areas are chosen to hide an object from the photographed scene when there are multiple objects. Thus, the size of blurred areas varies according to the size of the objects hidden in the images. We applied the method on 150 sets of multi focus images on a datasets of images [47]. In this paper, as the number of pages is limited, we present only 3 sets of multi focus images. Figures 4, 5, 7, 8, 10 and 11 show the multi focus images obtained by the convolution of Gaussian filter. Figures 6, 9 and 12 show the fused image by proposed method. Visually the image obtained by the proposed method gives a very satisfactory fusion.

Fig. 4 in focus on the right  Fig. 5 in focus on the left

Fig. 6 Fused image by proposed method
Fig. 7 in focus on the left

Fig. 8 in focus on the right

Fig. 9 Fused image by proposed method

Fig. 10 in focus on the left

Fig. 11 in focus on the right
For comparison purposes, we perform fusion using methods: PCA method [1], Discrete Wavelet Transform (DWT) method [6], Laplacian Pyramid LP_PCA [13], LP_DWT [14] and Bilateral gradient (BG) [2].

To objectively evaluate these fusion methods, quantitative measures of the fusion results are needed. According to the evaluation measure RMSE, the Table 1, gives the mean and standard deviation of RMSE for the given methods.

<table>
<thead>
<tr>
<th>Method</th>
<th>LP_AV</th>
<th>PCA</th>
<th>BG</th>
<th>LP_PCA</th>
<th>DWT</th>
<th>LP_DWT</th>
<th>Proposed_method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>6.351</td>
<td>6.245</td>
<td>7.7375</td>
<td>1.7456</td>
<td>3.0738</td>
<td>1.7841</td>
<td>0.3360</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>2.81099</td>
<td>2.76977</td>
<td>3.77837</td>
<td>0.62897</td>
<td>1.06387</td>
<td>0.638727</td>
<td>0.0338</td>
</tr>
</tbody>
</table>

The results show that the proposed method has a smaller mean of the RMSE. The histograms of RMSE for 150 images by different methods show for almost method that the values of RMSE are almost symmetrically centered around the mean value. In order not to clutter this paper, we present below only the histogram of the proposed method.
To compare analytically the proposed method to other methods we use the Analysis of variance (ANOVA) with dependent samples (dependence by image). The software R gives the following Anova table:

<table>
<thead>
<tr>
<th>Df</th>
<th>Sum Sq</th>
<th>Mean Sq</th>
<th>F value</th>
<th>Pr(&gt;F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>7310</td>
<td>1218.4</td>
<td>268</td>
<td>&lt;2e-16 ***</td>
</tr>
<tr>
<td>Residuals</td>
<td>1043</td>
<td>4741</td>
<td>4.5</td>
<td></td>
</tr>
</tbody>
</table>

As Pr(>F) is smaller than 1% the methods are significantly different. We use now the Newman Keuls test to compare the methods two-by-two and make groups having significantly the same mean. The software R gives the results below of the test.

<table>
<thead>
<tr>
<th>Method</th>
<th>A$\bar{R}$MSE</th>
<th>std</th>
<th>r</th>
<th>Min</th>
<th>Max</th>
<th>Q25</th>
<th>Q50</th>
<th>Q75</th>
</tr>
</thead>
<tbody>
<tr>
<td>BG</td>
<td>7.7379533</td>
<td>3.7783706</td>
<td>150</td>
<td>0.7996</td>
<td>22.7874</td>
<td>5.433375</td>
<td>6.92975</td>
<td>9.827225</td>
</tr>
<tr>
<td>DWT</td>
<td>3.0737927</td>
<td>1.0638730</td>
<td>150</td>
<td>0.5012</td>
<td>6.0300</td>
<td>2.320575</td>
<td>2.97360</td>
<td>3.851425</td>
</tr>
<tr>
<td>LP_AV</td>
<td>6.3513993</td>
<td>2.0109920</td>
<td>150</td>
<td>1.1543</td>
<td>17.3144</td>
<td>4.616225</td>
<td>5.96120</td>
<td>7.539725</td>
</tr>
<tr>
<td>LP_DWT</td>
<td>1.78411140</td>
<td>0.6387217</td>
<td>150</td>
<td>0.2923</td>
<td>3.6641</td>
<td>1.335000</td>
<td>1.75805</td>
<td>2.166900</td>
</tr>
<tr>
<td>LP_PCA</td>
<td>1.7456267</td>
<td>0.6289769</td>
<td>150</td>
<td>0.2872</td>
<td>3.6378</td>
<td>1.322750</td>
<td>1.72560</td>
<td>2.130525</td>
</tr>
<tr>
<td>PCA</td>
<td>6.2446747</td>
<td>2.7697604</td>
<td>150</td>
<td>1.1535</td>
<td>17.1652</td>
<td>4.656875</td>
<td>5.92430</td>
<td>7.323285</td>
</tr>
<tr>
<td>proposed_method</td>
<td>0.3360487</td>
<td>0.1039922</td>
<td>150</td>
<td>-0.0385</td>
<td>0.8720</td>
<td>0.196550</td>
<td>0.30555</td>
<td>0.449750</td>
</tr>
</tbody>
</table>

$\bar{A}$R$\bar{M}$S$\bar{E}$ groups:

- BG: a
- LP_AV: b
- PCA: c
- DWT: d
- LP_DWT: e
- LP_PCA: d
- proposed_method: e

Four different groups: Group “a” contains only method BG has the bigger mean of RMSE (7.737). Group “b” contains 2 methods LP_AV and PCA that have significantly the same average. Group “c” contains only the method DWT which better than group “a” and “b”. Group “d” contains 2 methods LP_DWT and LP_PCA which better than group “a”, “b” and “c”. The last group “e” containing the proposed method that the best method because his mean is the smallest by comparing with other means.

6. CONCLUSION

In this paper, we present the multi-focus image fusion method based using Dempster-Shafer Theory based on local variability. The method calculates the local variability for each pixel of each image and determines the mass function from local variability. The decision of fusion is obtained by pixels that correspond to minimum plausibility. The result of experiment shows that the proposed method gives significant improvement result in both visually and quantitatively. This method can be extended to image fusion for more than two blurred images. Our proposed method can be used in many applications, such as
Drone is a new technology in digital imaging, it has opened up unlimited possibilities for enhancing photography. Drone can capture images on the same scene that zooms in on different objects, and at various altitudes. It will produces several images on the same scene but with different objects in-focus. The proposed method is used to obtain an image with all objects in-focus.

In medical imaging, the proposed can be used to detect an anomaly object or cell using local variability where the behavior of each pixel with its neighborhood is given.

For quality control in of food industry, cameras are used that take pictures. Each camera targets one of several objects to detect an anomaly. The objects are on a conveyor belt. To have a photo containing all the objects in-focus, we can use our proposed method which gives more details information.

There are several perspectives of this work:

1. As many work on image fusion have implemented on grayscale images. In this paper, the proposed method is performed on the grayscale image. However, the proposed method can be extended to color images as color image conveys significant information.
2. We are also encouraged to fuse more than two images by taking into account the local variability in each image (intra variability) and variability between image (inter variability). This inter variability can detect the ‘abnormal pixels’ among the images.
3. We are motivated to extend the proposed method to fuse images with different objects from different sensors (multimodal).

ACKNOWLEDGEMENTS

The authors would like to thank the anonymous reviewers for their valuable comments.

REFERENCES

298

Computer Science & Information Technology (CS & IT)


