DEMPESTER-SHAFER AND MULTI-FOCUS IMAGE FUSION USING LOCAL DISTANCE

Ias Sri Wahyuni¹ and Rachid Sabre²

¹University of Burgundy, Gunadarma University
²Laboratory Biogéosciences CNRS, University of Burgundy/Agrosup Dijon, France

ABSTRACT

In this article, we give a new method of multi-focus fusion images based on Dempster-Shafer theory using local variability (DST-LV). Indeed, the method takes into account the variability of observations of neighbouring pixels at the point studied. At each pixel, the method exploits the quadratic distance between the value of the pixel I(x, y) of the point studied and the value of all pixels which belong to its neighbourhood. Local variability is used to determine the mass function. In this work, two classes of Dempster-Shafer theory are considered: the fuzzy part and the focused part. We show that our method gives the significant and better result by comparing it to other methods.

KEYWORDS

Multi-focus-images, Dempster-Shafer Theory, local distance, fusion images.

1. INTRODUCTION

Image fusion involves combining images from the same scene to produce a single image containing more information and detail found on at least one of the input images. Thus, image fusion can reduce uncertainty and minimize redundancy on the output image as well as maximize particular relevant information. This paper deals with the fusion of multi focus images caused by the limited depth of field of optical lenses in cameras. Thus, it is not possible to obtain an image containing all the relevant objects in a scene but with different focusing, one can obtain several images where each one contains a clear object and the others blurry. The image fusion method is used to get all objects in focus on a single image.

In the literature, we find several methods of fusion of multifocal images. These methods can be two types, spatial methods and multiscale methods. Those of the spatial domain directly concern the pixels of the source images as well as their neighbourhoods. Moreover, the fusion methods such as the mean, the principal component analysis (PCA) [1], the maximum selection rule, the methods based on the two-sided gradient [2] and the method based on the filter and Guided images (GIF) [3] and the maximum selection rule are considered as spatial approaches. The flaw of spatial domain approaches is that they cause spatial distortion in the fused image. On the other hand, the fusion by the methods at several scales is carried out on the source images after having decomposed them into several scales. As examples of these methods we cite among others: discrete wavelet transform (DWT) [4] - [7], the fusion of Laplacian pyramidal images [8] - [14], the discrete cosine transform with calculation of the variance (DCT + var) [15], the method based on the detection of salience (SD) [16].
The paper [17] show that, from the evidence point of view where the fusion decreases the imprecision and the uncertainty by using the redundancy and the complementary information of the source image. This means that evidence of weak inputs gives the best estimate. Shafer first proposed the evidence theory in the 1970s, based on Dempster's research. The advantage of the Dempster-Shafer theory (DST) is without a priori and without preference, because of the unavailability of information, which implies an indeterminacy, as detailed in [18] and [19]. This theory is very successful in many applications, including image segmentation [20], [21], pattern classification [22], [23], object recognition [24], medical imaging. [25], fusion of sensors [26].

This paper proposes the fusion of multi-focus images using Dempster-Shafer theory based on the following information: the variability between each pixel and its neighbours. This variability is calculated from the quadratic distance between the value of the pixel \( I(x, y) \) and the value of all neighbouring pixels. It is called "local variability". Local variability as a measure can detect the sharp intensity of the image such as the edge. This method also takes into account the behaviour of neighbouring pixels and preserves the edge.

The organization of this paper is as follows: Section 2 gives details of Dempster-Shafer evidence theory. Section 3 gives the definition of local variability and its use. Section 4 presents the proposed method in detail. Section 5 provides the definition of the evaluation metrics used in this article. The experimental study using different images and the comparison between the proposed method and other methods are provided in section 6. The section 7 is reserved for the conclusions and perspective of this work.

2. **DEMPSTER-SHAFER EVIDENCE THEORY**

Define \( \Theta \) as the set of hypotheses for a problem domain, called frame of discernment. Let \( m \) the function defined from \( 2^\Theta \) to \([0,1]\) where \( 2^\Theta \) be the set of all subsets \( \Theta \):

\[
2^\Theta = \{ A | A \subseteq \Theta \}. \tag{1}
\]

The function \( m \) is called a basic probability assignment whenever

\[
m(\emptyset) = 0 \text{ and } \sum_{A \subseteq \Theta} m(A) = 1. \tag{2}
\]

Where \( m(A) \) is the measure of belief that is attributed to \( A \). According to [27], \( m(A) \) is the degree of evidence supporting the claim that a specific element of \( \Theta \) belongs to the set \( A \), but not to any special subset of complementary of \( A \). If \( A \) is a element of \( \Theta \) such that \( m(A) > 0 \) then \( A \) is called the focal element of \( m \). The belief measure is given by using \( m \) as follows:

\[
Bel: 2^\Theta \rightarrow [0,1]: \quad Bel(A) = \sum_{B \subseteq A} m(B). \tag{3}
\]

The paper [28] define the plausibility measure \( Pl: 2^\Theta \rightarrow [0,1]: \)

\[
Pl(A) = \sum_{A \cap B = \emptyset} m(B) = 1 - Bel(A). \tag{4}
\]

\( Bel(A) \) measures the degree of evidence that the element in question belongs to the set \( A \) as well as to the various special subsets of \( A \). An important aspect of Dempster Shafer Theory (DST) is the aggregation of evidence given by different sources, see [17]. If two mass function \( m_1 \) and \( m_2 \)
applied at 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 $\Theta$, then they are combinable by means of Dempster’s rule. [29], [30]. The combination (joint mass) of two sets of masses $m_1$ and $m_2$ is defined as follows

$$m_1 \oplus m_2(\emptyset) = 0$$  \hspace{1cm} (5)

$$m_1 \oplus m_2(A) = \frac{\sum_{B \cap C = A} m_1(B)m_2(C)}{1-\sum_{B \cap C = A} m_1(B)m_2(C)}$$  \hspace{1cm} (6)

Equation (6) becomes

$$m_1 \oplus m_2(A) = \frac{\sum_{B \cap C = A} m_1(B)m_2(C)}{\sum_{B \cap C = A} m_1(B)m_2(C)}$$  \hspace{1cm} (7)

If the mass on a subset $A$ is zero does not mean that the set is impossible, simply that we are not capable of assigning a level precisely to $A$ (see [31]), since we could have non-zero masses on subsets of $A$, which would lead us to $Bel(A) \neq 0$.

### 3. Local Distance

This paper considerate the information of the neighbouring pixels of the fixed pixel. Indeed, at each pixel $I(x, y)$, the method uses the quadratic difference called local variability between the value of this pixel $I(x, y)$ and the value of its neighbours. The idea comes from the fact that the variability of the value in blurred region is smaller than the variability of the value in the focused region, the proof of this assertion is given in [32]. The neighbour of a pixel $(x, y)$ used in this paper, with the size $"a"$ is:

$$(x + i, y + j)\text{ where } i = -a, -a + 1, \ldots, a - 1, \text{ and } j = -a, -a + 1, \ldots, a - 1, a$$

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.

![Figure 1. Pixel at (x, y) within its neighborhood, a = 1](image-url)

Consider $p$ source images $(I_1, I_2, \ldots, I_p)$ with same size $(R \times C)$. Local variability of every source image at pixel $(x, y)$:
where \( k \) is the index of \( k^{th} \) source image \((k = 1, 2, \cdots, p)\).

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

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

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

The following proposition that the local variability is small enough where the location is on the blurred area \((B_1 \text{ or } B_2)\). Indeed, we consider, without loss the generality, that we have a focus pixel \((x, y)\) in image \(I_1\) and blurred in image \(I_2\), \(((x, y) \in B_2)\).

![Figure 2](image_url)

Figure 2. Two multi focus images, the yellow part is blurred area and the white part is clear(focused) area.

The local variability of image \(I_1\) and image \(I_2\) are respectively:

\[
\sqrt{\frac{1}{T} \sum_{m=-a}^{a} \sum_{n=-a}^{a} |I_1(x, y) - I_1(x + (m - a), y + (n - a))|^2}
\]

\[
\sqrt{\frac{1}{T} \sum_{m=-a}^{a} \sum_{n=-a}^{a} |I_2(x, y) - I_2(x + (m - a), y + (n - a))|^2}
\]

Proposition

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

The proof of this proposition is given in [32].
4. **The Proposed Method**

The difficulty of merging images using Dempster-Shafer theory is to construct the evidential representation of images. In this paper, one information used as the evidential representation images is local variability and two classes are used 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 of this forms the framework for discernment in \( \Theta \):

\[
\Theta = \{ \omega, \overline{\omega}, \theta \}
\]  

(12)

For each pixel one value of evidence for information will be obtained, \( m \).

\[
\{m(\omega), m(\overline{\omega}), m(\theta)\}
\]

(13)

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

Suppose there are \( P \) original images, \( I_1, I_2, \ldots, I_p \), where each image has size \( (R \times C) \) with different focus to be fused. The fusion in this work follows 3 steps:

**Step 1:**

1. To calculate mass function:

For each image where we use different values of size of neighbourhood, \( \alpha \in \{1,2,\ldots,10\} \), we define:

\[
d'_{\alpha,k}(x,y) = 1 - \frac{v_{\alpha,k}(x,y) - \min_{(x',y')} \left( v_{\alpha,k}(x',y') \right)}{\max_{(x',y')} \left( v_{\alpha,k}(x',y') \right) - \min_{(x',y')} \left( v_{\alpha,k}(x',y') \right)}
\]

(14)

where \( k \) is the \( k \)th source image, \( k \in \{1,2,\ldots,p\} \) and \( \alpha \) is size of neighbourhood of local variability. We set the standard deviation of \( d'_{\alpha,k}(x,y) = \sigma_{\alpha,k}(x,y) \),

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

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

(15)

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

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

(16)

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

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

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

(17)

This method obtains the information whether or not a pixel belongs to the focus area, for this we
use the plausibility of $\omega$ which is the sum of the masses of the evidence for $\omega$ and the uncertainty $\theta$:

$$P_{\alpha,k}(\omega) = m_{\alpha,k}(\omega) + m_{\alpha,k}(\theta)$$

For fusion image at 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 $P_{\alpha,k}(\omega)$, $k = 1, 2, \cdots, p$.

Step 2.

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

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

Step 3.

The proposed method use different values of size of neighbourhood, $\alpha \in \{1, 2, \cdots, 10\}$, and choose the value of $\alpha$ that corresponds to the minimum value of RMSE, such that our final fused image

$$F = F_{\alpha_0} \text{ where } \alpha_0 \in \{1, 2, \cdots, 10\} \text{ and } \text{RMSE}(F_{\alpha_0}) = \min_{\alpha \in \{1, 2, \cdots, 10\}} \text{RMSE}(F_{\alpha}).$$

5. Experimental Result

In this section, we are using images taken from the web page database [35]. We blurred an area of each reference image by applying the convolution of the Gaussian filter. The works [33] - [34] justify the choice of the Gaussian filter. The reference images contain at least two objects of the photographed scene. We have chosen to hide an object from the reference image. Thus from each image we obtain multi-focus images whose number is equal to the number of objects belonging to the reference image. The size of the blurred areas depends on the size of the masked object. We applied the approach on 150 images of the web page [35]. To make this article easy to read, we have chosen to present only three reference images that we blurred by masking an object each time to extract multi-focus images from each image (figures 4, 5, 7, 8, 10 and 11). Figures 6, 9 and 12 show the images fused using the 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
To compare the proposed method to other methods, we perform the fusion using methods: PCA method [1], discrete wavelet transform (DWT) method [6], Laplacian pyramid LP_PCA [13], LP_DWT [14] and gradient bilateral (BG) [2].

To objectively evaluate these fusion methods, we will use quantitative measurements on the fused images. The RMSE evaluation measure was chosen for its efficiency in this comparison category. The table gives the mean and standard deviation of RMSE for the methods studied.

<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>0.44059</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>0.223299</td>
</tr>
</tbody>
</table>

From the table, we deduce that the proposed method has a smaller mean of the RMSE. The histograms of RMSE for 150 images by different methods (Figure 14, 15, 16, 17, 18 and 19) are for almost method symmetric and centred around the mean value.
Figure 13. The histogram of LP. PCA method

Figure 14. The histogram of PCA method

Figure 15. The histogram of DWT method
Figure 16. The histogram of LP.DWT method

Figure 17. The histogram of BG method

Figure 18. The histogram of proposed method
An analytic comparison, Analysis of variance (ANOVA) with dependent samples (dependence by image), is used. The software R gives the following Anova table:

<table>
<thead>
<tr>
<th>Method</th>
<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>Residuals</td>
<td>1036</td>
<td>4743</td>
<td>4.58</td>
<td></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.

We obtain Four different groups: First Group “a” contains only method BG having the bigger mean of RMSE (7.737). The Group “b” contains two 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 two 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

This paper, presents a new method to fuse two multi-focus images based on Dempster-Shafer theory using local variability. The originality of this method lies in the fact that we use the Dempster Shaper theory and the local variability of each pixel according to the quadratic distance. The fusion decision is obtained by pixels, which correspond to the minimum of plausibility. The experimental study shows that the proposed method gives a significant improvement in the result both visually and quantitatively. We extend this method to fuse more than two blurred images. We can use it in many applications, such as

1. The drone is becoming an essential tool in digital imaging, it offers interesting possibilities to improve photography. The drone can capture images on the same scene, which zooms in on different objects and at different altitudes. Thus, it will give several images on the same scene but with different objects in focus. With the proposed method, we obtain an image with all the objects at the point very similar to the real images

2. The method can also be used in medical imaging. In fact, it can be used to detect an object or cell anomaly due to the local variability indicating the behaviour of each pixel with its neighbourhood.

3. The food industry uses cameras to control the quality of the manufactured product. Each camera targets one of several objects to detect an anomaly. We would use the proposed method to get a photo containing all the objects in focus with more detailed information.

This work has several perspectives:

1. Our proposed method is performed on images in gray levels. We plan to extend it to colour images that convey important information.

2. Encouraged by the fusion of two images we intend to extend the method to more than two images by taking into account the local variability in each image (intra variability) and the variability between the images (inter variability). This inter-variability can detect "abnormal pixels" among images.

3. We are motivated to use the proposed method to fuse images containing different objects from different sensors (multimodal).

7. ACKNOWLEDGEMENTS

The authors would like to thank the anonymous reviewers for their important suggestions and remarks.

REFERENCES


AUTHORS

Ias Sri Wahyuni was born in Jakarta, Indonesia, in 1986. She earned the B.Sc. and M.Sc. degrees in mathematics from the University of Indonesia, Depok, Indonesia, in 2008 and 2011, respectively. In 2009, she joined the Department of Informatics (COMPUTING) System, Gunadarma University, Depok, Indonesia, as a Lecturer. She is currently a PhD student at University of Burgundy, Dijon, France. Her current research interests include statistics and image processing.

Rachid Sabre received the PhD degree in statistics from the University of Rouen, Rouen, France, in 1993 and Habilitation (HdR) from the University of Burgundy, Dijon, France, in 2003. He joined Agrosup Dijon, Dijon, France, in 1995, where he is an Associate Professor. From 1998 through 2010, he served as a member of “Institut de Mathématiques de Bourgogne”, France. He was a member of the Scientific Council of AgroSup Dijon from 2009 to 2013. In 2012, he has been a member of “Laboratoire Electronique, Informatique, et Image” (Le2i), France. Since 2019 has been a member of Laboratory Biogeosciences UMR CNRS, University of Burgundy. He is author/co-author of numerous papers in scientific and technical journals and conference proceedings. His research interests lie in areas of statistical process and spectral analysis for signal and image processing.

© 2021 By AIRCC Publishing Corporation. This article is published under the Creative Commons Attribution (CC BY) license.