LOCAL DISTANCE AND DEMPSTER-DHAFER FOR MULTI-FOCUS IMAGE FUSION

Ias Sri Wahyuni¹ and Rachid Sabre²

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

ABSTRACT

This work proposes a new method of fusion image using Dempster-Shafer theory and local variability (DST-LV). This method takes into account the behaviour of each pixel with its neighbours. It consists in calculating the quadratic distance between the value of the pixel \(I(x, y)\) of each point and the value of all the neighbouring pixels. Local variability is used to determine the mass function defined in Dempster-Shafer theory. The two classes of Dempster-Shafer theory studied are: the fuzzy part and the focused part. The results of the proposed method are significantly better when comparing them to results of other methods.

KEYWORDS

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

1. INTRODUCTION

The role of image merging is to use images of the same scene with different focus to produce a single image containing the details found on at least one of the input images. Thus, image merging can reduce uncertainty and minimize redundancy on the output image as well as maximize particular relevant information. This work mainly concerns the fusion of multifocal images caused by the limited depth of field of the optical lenses of cameras. Thus, with a different focus, we can obtain several images where each contains a bright object and the others blurry. In this case, the method of fusion image is used to bring all the objects into focus on a single image.

Multifocal image fusion methods found in the literature are divided into two types, spatial methods and multiscale methods. The spatial methods use directly the pixels of the source images and the neighbouring pixels. Some fusion methods like: the mean, the principal component analysis (PCA) [1], the maximum selection rule, the two-sided gradient-based methods [2] and the filter-based method and guided images (GIF) [3] and the maximum selection rule are considered as spatial approaches. Among the disadvantages of spatial domain approaches is that they create spatial distortion in the fused image. On the other hand, the fusion using the methods at several scales is carried out on the source images after having decomposed them into several scales. We cite some examples of these methods: the discrete wavelet transform (DWT) [4] - [7], the fusion of Laplacian pyramidal images [8] - [14], the discrete cosine transform with variance calculation (DCT + var) [15], the method based on the detection of salience (SD) [16].

The article [17] shows that the fusion decreases the imprecision and the uncertainty by the redundancy as well as the complementary information of the source image. Shafer was the first to
introduce evidence theory in the 1970s, based on Dempster's research. The Dempster-Shafer theory (DST) has the advantage of being without a priori and without preference, due to the unavailability of information, this leads to indeterminacy, as detailed in [18] and [19]. Numerous applications of this theory, particularly in the field of: image segmentation [20], [21], pattern classification [22], [23], object recognition [24], medical imaging [25], sensor fusion [26].

This work gives a multi-focus image fusion method based on Dempster-Shafer theory using the variability between each pixel and its neighbors as distance. The calculation of this variability is made from the quadratic distance between the value of each pixel \( I(x, y) \) and the value of all the neighboring pixels called “local variability”. It allows as a measure to detect the sharp intensity of the image, the edges and preserves the edge.

This paper is organized as follows: Section 2 details the evidence Dempster-Shafer theory. Section 3 defines local variability and its use. Section 4 details the proposed method. Section 5 details the assessment metrics used in this article. The experimental study on different images with comparison to other methods are provided in section 6. Section 7 gives conclusions and perspective of this work.

### 2. Dempster-Shafer Evidence Theory

We start with Define \( \Theta \) as the set of assumptions for a problem area, called a framework for discernment. The function \( m \) defined from \( 2^\Theta \) to \([0,1]\) where \( 2^\Theta \) by where the set of all the subsets \( \Theta \):

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

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

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

where \( m(A) \) is the measure of belief that is attributed to \( A \). According to [27], \( m(A) \) measures the degree of evidence supporting the assertion that a specific element of \( \Theta \) belongs to the set \( A \), but not to a special complementary subset of. If \( A \) belongs to \( \Theta \) such that \( m(A) > 0 \) then \( A \) will be called the focal element of \( m \). The belief measure is calculated using the function \( m \) as follows:

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

The plausibility measure is defined in the paper [28] by: \( Pl: 2^\Theta \mapsto [0,1] \):

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

\( Bel(A) \) measures the degree of evidence that an element belongs to the set \( A \) as well as to the various special subsets of \( A \). The aggregation of evidence given by different sources, see [17], is an important aspect of Dempster Shafer Theory (DST).
Let $B$ and $C$ of non-disjoint subsets of $\emptyset$, such that $m_1(B) > 0$ and $m_1(C) > 0$. Then the functions $m_1$ and $m_2$ 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

$$m_1 \oplus m_2(\emptyset) = 0 \quad (5)$$

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

Equation (6) becomes

$$m_1 \oplus m_2(A) = \frac{\sum_{B \cap C = A} m_1(B)m_2(C)}{\sum_{B \cap C = \emptyset} m_1(B)m_2(C)} \quad (7)$$

when the mass of a subset $A$ is equal to zero it simply means that we are not able to assign a level precisely to $A$ (see [31]), since we could have non-zero masses on sub-sets of, which would lead us to $Bel(A) \neq 0$.

3. LOCAL DISTANCE

This work is based on the use of the quadratic difference called local variability between the value of each pixel $I(x,y)$ and the value of its neighbours. The idea comes from the fact that in the fuzzy region, the local variability is smaller than that 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:

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

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)

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

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

The following proposition that the local variability is small enough where the location is on the blurred area ($B_1$ 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 $I_2$, $(x,y) \in B_2$

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

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

The following proposition that the local variability is small enough where the location is on the blurred area ($B_1$ 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. Two multi focus images, the yellow part is blurred area and the white part is clear(focused) area.](image)

The proof of this proposition is given in [32].
4. THE PROPOSED METHOD

The use of Dempster-Shafer theory in fusion images is based on the construction of the probative representation of images. In this article, local variability plays the role of information used as a convincing representation image by constructing two classes in Dempster-Shafer theory. A pixel belongs either to the blurred part \(\omega\) or it belongs to the focused part \(\overline{\omega}\). Obviously, there is an inherent uncertainty \(\Theta\) in the evidence theory. These elements form the framework of discernment in \(\Theta\):

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

(12)

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

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

(13)

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

Assume that there are \(P\) original images, \(I_1, I_2, \ldots, I_P\) where each image has size \((R \times C)\) with different focus area to be fused. The proposed fusion method follows 3 steps:

Step 1:

1. To calculate mass function:

For each image and for each size of neighbourhood, \(a \in \{1, 2, \ldots, 10\}\), we calculate:

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

(14)

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

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

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

(15)

for each pair \((x, y)\) belongs to \(\theta\), we give:

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

(16)

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

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

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

(17)

This method determines the information whether or not a pixel belongs to the focus area by using
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.

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

$$F = F_{a_0} \text{ where } a_0 \in \{1, 2, \cdots, 10\} \text{ and } \text{RMSE}(F_{a_0}) = \min_{\alpha\in\{1,2,\cdots,10\}}(\text{RMSE}(F_\alpha))$$

5. EXPERIMENTAL RESULT

The experimental part uses images extracted from the database of Web pages [35] containing 150 images with at least two objects of the photographed scene. We proceeded to blur an area of each reference image by the convolution of the Gaussian filter. The justification for the choice of the Gaussian filter is developed in works [33] - [34]. We have chosen to hide an object from the reference image. So we have a reference image and we get multifocus images. The size of the blurred areas depends of course on the size of the masked object. We applied the approach to 150 images from the web page [35]. In order not to clutter up this article, we have chosen to present only three reference images that we have blurred by masking each time an object to extract multifocus images from each image (figures 4, 5, 7, 8, 10 and 11). Figures 6, 9 and 12 show the images merged using the proposed method. The proposed method visually gives a very satisfactory fused image.
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
Now we compare the proposed method to other methods. For that, we apply the fusion methods: PCA method [1], discrete wavelet transform (DWT) method [6], Laplacian pyramid LP_PCA [13], LP_DWT [14] and gradient bilateral (BG) [2].

To evaluate the quality of these fusion methods, we will use metric tools such as the RMSE evaluation measure which has given satisfaction for this kind of comparison. The following table calculates the means and the standard deviations of the RMSEs between the reference images and the fused images by 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, the proposed method gives a smaller mean and smaller standard deviation of RMSE. We produce the histogram graphs of RMSE for 150 images by different methods (Figure 14, 15, 16, 17, 18 and 19). They 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
To complete our study and make homogeneous groups of methods, we perform an analysis of variance (ANOVA) with dependent samples (dependence by image). 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>

Since the probability Pr(>F) is smaller than 1% the methods are significantly different. To compare the methods two-by-two and determine the groups having significantly the same mean, we perform the Newman Keuls test. The software R gives the results below of the test.

We distinguish the different groups: First Group “a” containing only method BG that have the bigger mean of RMSE (7.737). The Group “b” containing two methods LP_AV and PCA that have significantly the same average. Group “c” containing only the method DWT which better
than group “a” and “b”. Group “d” containing two methods LP_DWT and LP_PCA which better than group “a”, “b” and “c”. The last group “e” containing the proposed method that have the smallest mean. Thus, consider the proposed method ca as the best method comparing to other methods.

6. CONCLUSION

The new method presented in this paper serves to merge two multifocal images using local variability and Dempster-Shafer theory. The originality of this work is to combine the theory of Dempster Shaper and the local variability based on the quadratic distance of each pixel and its neighbours. Thus, the merger decision is taken pixel by pixel, it is the one that is at least plausible. An experimental study is carried out in this work and highlights the fact that the proposed method presents a significant improvement in the result both visually and analytically. This work will be used to fuse more than two blurry images. This method can be used in many applications, such as

1. Using the images taken by drones that are essential tools 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 gives several images on the same scene but with different objects in focus. We will apply the proposed method to obtain one image with all the objects in focus very similar to the real image.

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

3. The method can be used in the food industry which uses cameras to control the quality of the produced product. Each camera targets one of several objects to detect an anomaly. The proposed method will give a photo with all the objects in focus and with more detailed information.

In terms of research, this work has several perspectives:

1. We plan to extend this work to color images that convey important information.

2. We intend to adapt 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 identify “abnormal pixels” among the images.

3. We will adapt the proposed method to fuse images containing different objects taken by different equipment: ultrasound, IREM, scanner ... (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 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 Laborotary 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.