A Novel Image Fusion Metric for Intelligent Manufacturing Information System

Yao YAN*, Mei-hong ZHENG, Jian-hua LI, Qi DONG and Bin HUANG
Lianyungang JARI Deepsoft Technology Co., LTD Hailian Road 42, Jiangsu, P.R. China
*Corresponding author

Keywords: Image fusion, Fusion quality, Manufacturing information system, Image fusion metric.

Abstract. An image fusion metric is commonly used to evaluate a fusion scheme because subjective evaluation cannot work in an intelligent manufacturing information system. In this study, an objective image fusion metric based on a log-Gabor filter (LGIMF) is presented. This metric can be calculated in five steps: (1) filtering the source and fused images into sub-bands, (2) constructing an ideal synthesis image by applying the maximization principle from the sub-band of the source images, (3) capturing the variation information between the real fused image and the ideal synthesis image in each sub-band, (4) measuring the sub-band fusion visual information by using the signal-to-noise ratio model, and (5) weighting the sub-band fusion visual information to determine the overall quality. In our experiment, the proposed fusion metric is compared with other well-known metrics by using a subjective test dataset. We found that the LGIMF was more consistent in subjective perception compared with the other metrics.

Introduction

An image sensor often cannot offer complete information for a complex scene; thus, multisensor image fusion was proposed to combine information from two or more images of a scene into a single composite image. This is widely used in intelligent manufacturing information system. After decades of development, many image fusion methods have been introduced [1]. However, a new problem has emerged: how to assess the fusion quality. As subjective judgment cannot serve as feedback information in an automatic system, the use of objective fusion quality metrics or fusion assessment algorithms in fusion performance assessment has been receiving more attention.

Many researchers have published their fusion quality metrics. Qu et al. [2] and Hossny [3] proposed objective image fusion metrics based on mutual information (MI) and normalized mutual information (NMI), respectively, which have been widely accepted. Xydeas and Petrovic [4] used edge information as a guide in setting up their image fusion metric, which has shown high performance. In the works of Piella [5], Cvejic [6], and Yang [7], the structural similarity index measure (SSIM) [8] was applied in fusion assessment due to its excellent performance in image quality prediction. Applying a similar idea, Han [9] used visual information fidelity (VIF) to construct an image fusion metric, which was rigorously compared with other metrics. Human vision system (HVS) based metrics have also emerged in image fusion assessment research. Chen [10] proposed a new fusion metric based on contrast sensitivity function (CSF) and local spatial information. Blum [11] set up a new fusion metric that used local contrast in an empirical CSF filtered image. Toet [12] was the first one to discuss the assessment of feature fusion. In his work [12], the subject contour was depicted for both the source and the fused image. Thus, the result of fusion assessment is apparent. Petrovic [13] presented a method that pooled the localized quality scores of an optimal spatial space to evaluate the image fusion. The method seemed to be regarded as an improvement of the block-based fusion metric, and could be used to enhance the performance of existing image fusion metrics.

Because subjective assessment is the final evaluation criterion of a fused image, the HVS still affects fusion quality research. In the present study, we present a new objective image fusion quality assessment algorithm based on the log-Gabor filter, which has been widely used to model the HVS. The remainder of this paper is organized as follows. Section 2 briefly introduces log-
Gabor filters in image processing. Section 3 describes the principle of the fusion metric based on log-Gabor filters. Section 4 compares the proposed metric with several other fusion metrics in Petrovic’s database [14]. Finally, Section 5 presents the conclusions drawn.

Log-Gabor Filter for Image Analysis

Gabor filters [15] are widely accepted as analysis tools for modeling the spatial behavior of vision. However, the classic Gabor filters present certain difficulties in multi-resolution analysis. First, they overlap more in low frequencies than in higher ones, which results in nonuniform coverage of the spectrum domain. Second, Gabor filters are easily affected by DC components due to their non-zero mean.

Thus, instead of Gabor filters, log-Gabor filters [16] were chosen for modeling the receptive field of simple cortical cells. Log-Gabor filters avoid the problem of non-zero DC components and can produce a fairly uniform coverage of the frequency domain in an octave scale scheme [17]. Log-Gabor filter banks have been proven to be effective tools for describing the vision response in the frequency domain. The log-Gabor filter is defined by the following equation:

$$G_{s,t}(\rho, \theta) = \exp\left(-\frac{1}{2}\left(\frac{\log(\rho) - \log(\rho_s)}{\log(\sigma_s / \rho_s)}\right)^2\right) \exp\left(-\frac{1}{2}\left(\frac{\theta - \theta_t}{\sigma_\theta}\right)^2\right)$$

where $G_{s,t}(\cdot)$ is the filter denoted by the spatial scale index $s$ and the orientation index $t$. Here, $(\rho, \theta)$ refers to the log-polar coordinates in the frequency domain. The parameter $\rho$ is the normalized radial frequency and $\theta$ is the orientation.

In Eq.1, spatial scale index $s$ influences the filter through parameter $\rho_s$, which denotes the normalized center frequency of the scale. The parameter $\theta_t$ is the center orientation of the filter; $\sigma_s / \rho_s$ and $\sigma_\theta$ determine the bandwidth of the spatial scale and the angular orientation, respectively.

According to [18], log-Gabor filters could approximate the cortical responses in the primary visual cortex. By selecting appropriate $\rho_s$, $\sigma_s / \rho_s$, $\sigma_\theta$ and $\theta_t$, the log-Gabor filter could simulate the corresponding estimates obtained from the mammalian visual system.

Previous researchers [16] have typically used five scales ($s=1, \ldots, 5$), and eight orientations ($t=1, \ldots, 8$) (resulting in 40 sub-bands per image) to analyze vision.

Fusion Quality Metric Based on a Log-Gabor Filter

Log-Gabor filters have been shown to be capable of analyzing images; however, the process of transforming the analysis information into a fusion quality problem still needs to be discussed. According to [9], the signal-to-noise ratio (SNR) could reflect how much visual information is contained in the test image compared with the reference image.

Here, we construct our fusion metric based on the similar idea; Fig.1 shows a schematic of the proposed metric. By using a log-Gabor filter, an ideal synthesis image is constructed from the source images. Then, the visual information is defined as the SNR of the contrast image between the fused image and the ideal synthesis image in each sub-band. Finally, the fusion metric is determined by weighting the visual information in each sub-band.

For general fusion issues, the source image for fusion is indexed by $I_1, \ldots, I_n$, and the fused image is indexed by $I_F$. Here, we denote the log-Gabor filtered image as LGI. $LGIL_{s,t}$ represents a sub-band image of $I_L$ generated by the log-Gabor filter in spatial space; the center scale and center orientation are denoted as $s$ and $t$, respectively. $LGIL_{s,t}(i, j)$ refers to the pixel of the sub-band image $LGIL_{s,t}$ in row $i$ and column $j$. We apply the following five steps in describing the metric:

Filtering the Source and Fused Images into Sub Bands

In scale $s$ and orientation $t$, the filtered source image and the fused image are $LGIL_{s,t}$, $LGIF_{s,t}$.
Constructing an Ideal Synthesis Image

The ideal synthesis image (ISI) is a virtual image that could be recovered by all its sub-band images (although it is unnecessary to recover it in the present work). The sub-band image of ISI in scale $s$ and orientation $t$ ($ISI_{s,t}$) is determined by Eq.2.

$$ISI_{s,t}(i,j) = LGI_{s,t}^I(i,j)$$

$$L = \arg\max_p (LGI_{s,t}^1(i,j), \ldots, LGI_{s,t}^p(i,j), \ldots, LGI_{s,t}^n(i,j))$$

Figure 1. Schematic of the LGIMF.

In Eq.2, $ISI_{s,t}(i,j)$ denotes the pixel of the $i$-th row and $j$-th column in the sub-band image $ISI_{s,t}$.

Constructing the Synthesis Fusion Contrast Image

The synthesis fusion contrast image (SFCI) is also a virtual image. After the above computation, the synthesis fusion contrast image (SFCI) in scale $s$ and orientation $t$ ($SFCI_{s,t}$) is defined by Eq.3.

$$SFCI_{s,t}(i,j) = \frac{LGI_{s,t}^I(i,j)}{ISI_{s,t}(i,j)}$$

Computing the SNR of the Synthesis Fusion Contrast Image in Each Sub-Band

The signal-to-noise ratio (SNR) of an image $I$ is defined by Eq.4.

$$SNR = \sum_{x} \sum_{y} SNR(SFCI_{x,y})$$

$$LGIMF = \sum_{x} \sum_{y} LGIMF$$
\[ SNR(I) = \frac{1}{M \cdot N} \sum_{i,j} I(i, j) I(i, j) \left( \frac{1}{M \cdot N - 1} \sum_{i,j} (I(i, j) - \frac{1}{M \cdot N} \sum_{i,j} I(i, j))^2 \right) \]

where \( I(i, j) \) is the gray value of the \( i \)-th row and \( j \)-th column in image \( I \); \( M \) and \( N \) indicate the size of image \( I \). The equation follows the traditional definition of SNR, and could be easily accepted.

**Weighting the SNR in Each Sub-Band**

Because the orientation is assumed to be uniformly distributed in nature, we use \( LSNRO_s \) to synthesize the SNR of \( SFCI_{x,y} \) in each direction. The process could be expressed by the following equation.

\[ LSNRO_s = \sum_r SNR(SFCI_{x,y}) \]  

Finally, the log-Gabor–based image fusion quality metric (LGIMF) could be calculated by weighting \( LSNRO_s \).

\[ LGIMF = \sum_s w_s LSNRO_s \]  

4. Validation Experiment and Discussion

To ensure the reliability of the quality metrics (in terms of statistical significance), a comparison based on Video Quality Experts Group [19] was carried out in present work. The subjective test database released by Petrovic [14] was selected for the metric validation. The database contained 120 pairs of registered gray image, all of which were captured in real or realistic conditions by using different sensors. For each pair of registered source images, two different fusion methods were used, and then two fused images were generated. All the fused images were evaluated based on the ITU recommendation [20] and widely used. In the work of Petrovic[14], the two indicators, namely, subjective relevance (SR) and correct ranking (CR), was suggested for an objective comparison of the fusion metrics. The higher the SR and CR, the better the performance of the image fusion metric.

4.1 Parameter Selection

The parameters of the log-Gabor filter were selected and validated in [16, 21]; in the present work, we use the recommended parameters directly. Thus, only the weighting parameter \( w_s \) needs to be determined. To obtain \( w_s \) in (8), we optimized the CR of the \( LGIMF \) by using the classic optimization algorithms in the database. We found that when \( w_s \) is \([1, 1, -1.4, -1.16, 0.7]\) for 5 scales, the \( LGIMF \) performs well in prediction.

4.2 Main Results

We computed the SR and CR of all the above-mentioned metrics in the database; Table 1 shows the results.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SR</strong></td>
<td>0.563</td>
<td>0.563</td>
<td>0.713</td>
<td>0.649</td>
<td>0.573</td>
<td>0.700</td>
<td>0.446</td>
<td>0.708</td>
<td>0.745</td>
<td>0.781</td>
</tr>
<tr>
<td><strong>CR</strong></td>
<td>0.650</td>
<td>0.650</td>
<td>0.725</td>
<td>0.717</td>
<td>0.616</td>
<td>0.717</td>
<td>0.575</td>
<td>0.750</td>
<td>0.792</td>
<td>0.825</td>
</tr>
</tbody>
</table>

According to our experiment, the LGIMF correctly predicted 99 of 120 pairs. Thus, the LGIMF clearly has the highest predictive capability among the different fusion metrics. If we consider the influence of the empirical parameters, the LGIMF correctly predicted 94 pairs; that is, the prediction was correct for nearly 80% of the samples. Therefore, we believe that the LGIMF has better generalization performance.
Conclusion

A new image fusion performance metric, the LGIMF, is proposed in this report. The LGIMF first uses a log-Gabor filter to decompose the source and fused images into sub-bands. Then, the metric applies the maximization principle to build an ideal synthesis image from the source images in each sub-band. Thereafter, a ratio model is introduced to capture the contrast image between the real fused image and the ideal synthesis image in each sub-band. By calculating the information index SNR, the LGIMF collects the effective visual information of the sub-band fused image. Finally, the metric is calculated by weighting all the SNRs of the sub-band image. To compare the LGIMF with other fusion metrics, a parallel experiment is carried out by using Petrovic’s database; the experimental results show that the LGIMF has better predictive performance.

References


