Efficient Compressive Multi-Focus Image Fusion

Chao Yang, Bin Yang*
College of Electric Engineering, University of South China, Hengyang, China
Email: ychao0906@163.com, yangbin01420@163.com

Received April 2014

Abstract

Two key points of pixel-level multi-focus image fusion are the clarity measure and the pixel coefficients fusion rule. Along with different improvements on these two points, various fusion schemes have been proposed in literatures. However, the traditional clarity measures are not designed for compressive imaging measurements which are maps of source sense with random or likely random measurements matrix. This paper presents a novel efficient multi-focus image framework for compressive imaging sensor network. Here the clarity measure of the raw compressive measurements is not obtained from the random sampling data itself but from the selected Hadamard coefficients which can also be acquired from compressive imaging system efficiently. Then, the compressive measurements with different images are fused by selecting fusion rule. Finally, the block-based CS which coupled with iterative projection-based reconstruction is used to recover the fused image. Experimental results on common used testing data demonstrate the effectiveness of the proposed method.

Keywords

Clarity Measures, Compressive Imaging, Multi-Focus Image Fusion

1. Introduction

Multi-focus image fusion, which is a major branch of multi-sensor data fusion, is aimed to produce an all-in-focused image from a sequence images with focus on different parts. The multi-focus image fusion can be performed in the transform domain or the spatial domain. In the transform domain, image fusion algorithms are fully implemented via a certain transform technique. In Refs [1] wavelet transform and curvelet transform are applied to image fusion. Refs [2] also exploited a multi-focus image fusion method based on Laplacian pyramid. Moreover, the nonsubsampled contourlet transform (NSCT) is presented in Refs [3]. In spatial domain, fusion rules are directly applied to image pixels or image regions [4]. The simplest fusion method in spatial domain is to take the average of the source images pixel by pixel [5] [6]. A multi-focus image fusion technique proposed by Li et al. [7] is that input images are divided into blocks and better focused ones are selected to compose a resultant image by majority filtering of the source image. Refs [8] proposed an algorithm that focused blocks can

*Corresponding author.
be detected via measuring blurriness. In Refs [9], focus measurements are obtained by sum of gradient values of all pixels in each block. Refs [10] presented a simple method using a morphology-based focus measure in a quad-tree structure. The schemes using divided blocks can make calculation simple and fast.

In recent years, the compressive sensing theory (CS) presented by D. L. Donoho, E. Candès, and T. Tao has attracted many attentions [11] [12], and [13]. Based on CS, various compressing imaging (CI) systems have been presented. Single-pixel camera architecture [14] implements random linear measurement process of the scene under view that combines sampling and compression. Wan et al. firstly introduced a CS-based image-fusion algorithm which uses “double-star” sampling pattern in a 2-D Fourier domain and shows the recovery via a total variation optimization [15]. However, the partial Fourier matrix is only incoherent with signals that are sparse in the time domain, restricting its practical applications. Ref [16] presented an image fusion scheme under SBHE fast CI that fuses the measurements through a linearly weighted average according to the entropy of measurements. However, as long as the entropies of the measurements did not carry enough information, it could not construct a good reconstructed image. Luo et al. revealed a method that directly performed fusion on the measurement vectors from multiple CI sensors according to the similarity classifications [17]. However, data similarity of measurement vectors needs to analyze much information before fusing. In traditional multi-focus image fusion, pixel clarities are calculated with neighboring pixels or regions. However, the traditional clarities calculations cannot be directly used for the compressive measurements since the random projections contain no geometry structure of source images. In order to solve this problem, a novel multi-focus image fusion method in CS domain is presented in this paper. In the new fusion scheme, the clarity measures bases are trained from natural multi-focus samples. Then, the clarity measures of the raw compressive measurements are not obtained from the random sampling data itself but from the coefficients of Hadamard clarity measure bases trained which can also be obtained from compressive imaging system efficiently. The clarity measures calculated from measurement values are used to guide the fusion rule to get fused image measurements. Finally, the full resolution fused image is constructed from the fused CS sampling measurements. In experimental simulations, the proposed method can get an attractive performance that usually exceeds the quality of other current methods.

The rest of this paper is organized as follows: the background of CS and CI are described in Section 2. In Section 3, the proposed method is introduced explicitly. The experiments and analysis are demonstrated in Section 4. The paper is concluded in Section 5.

2. Compressive Sensing and Compressive Imaging

The theory of compressive sensing becomes very popular due to its wide areas of applications. As states in CS theory, the signals can be well recovered from the few measurements by nonlinear optimization if the signal is sparse or compressible. When a \( n \)-pixels image is arrangement as a compressible vector \( x \in \mathbb{R}^n \), it can be represented as

\[
x = \Psi \theta,
\]

where \( \Psi \) indicates a certain basis and \( \theta \) is a sparse vector containing only \( k \) nonzero coefficients. A random matrix \( \Phi \in \mathbb{R}^{k \times n} \) projects the images from high-dimensional space \( x \in \mathbb{R}^n \) to a lower-dimensional space \( y \in \mathbb{R}^k \) as

\[
y = \Phi x,
\]

where \( k \ll n \). As proved in [18], the random matrix \( \Phi \in \mathbb{R}^{k \times n} \) provides a stable embedding that preserves the distance between all pairs of original signals with high probability. According to CS theory, the recovery of signal \( x \) from measurement vector \( y \) is possible.

Imaging techniques have a strong affinity with compressive sensing. Based on CS theory, various hardware implementations of compressive imaging system have been proposed. Single-pixel camera directly acquires random projections of a scene without initial collecting the pixels. The camera architecture employs a digital micro mirror array to optically calculate linear projections of the scene onto pseudorandom binary patterns. CI measurements are the projections of a continuous scene onto a discrete space instead of collecting the lighting intensity at distinct locations [19]. Each measurement is a linear combination of several pixels. Comparing with conventional imaging system, compressive imaging system requires less computation and storage because of compressing during sensing. Due to above some advantages of CI, the compressive imaging system is very fit
for applications in visual sensor network.

3. The Proposed Method

Figure 1 presents the fusion strategy for compressive imaging in visual sensor network. A common scene is observed by two node cameras with different optic setting. Compressive sensing encoding schemes within each sensor directly and independently sample the scene into a small number of randomized measurements denoted by \( y_1 \) and \( y_2 \). Then the measurements \( y_1 \) and \( y_2 \) are fused according to certain fusion rule. Finally, the resultant fused image is reconstructed from the combined compressive measurements \( y_F \). The key point of this fusion scheme is the focus-level evaluating for random measurements according to which \( y_1 \) or \( y_2 \) is focused. Clearly, the focus-level cannot be directly calculated by random sampling coefficients since the random projections don’t contain geometry structure of the source scene. In this work, the innovative point is to design or train image clarity measures bases to evaluate focus-level for compressive imaging system. Firstly, the absolute or norm-1 of the projection measurements with image clarity measures bases can be used as focus-level evaluating. Secondly, the projection measurements can be obtained directly from the compressive imaging system. Thirdly, there is no complex calculation for node cameras. So these characteristics are very suitable for visual sensor network application.

3.1. Clarity Measures Bases

Recently, to more effectively fuse the focused regions from multi-focus images, various fusion methods are proposed based on the clarity measures. This paper puts forward a novel image fusion scheme on clarity measures. In this work, we choose 14 pairs of multi-focus images. A fraction collection of the 14 pairs of multi-focus images are presented in Figure 2(a) where the up row gives the clear images and the bottom row gives the blurred images. The two training data sets are randomly taken from a database of 14 natural pairs of multi-focus images. Each set consists of 3584 16 × 16 patches and the Hadamard coefficients of those patches are represented as A and B, respectively. So each column of A and B denotes the Hadamard coefficient of each patch, respectively. Then the coefficient \( y_{ij} \), the value of the \( i \)th row and \( j \)th column of A, corresponds to a Hadamard base which is meaningful to detect and emphasize image salient features. In this work, the aim is to find clarity measures bases from Hadamard matrix and use them to evaluate image focus-level. we suppose each row of coefficient matrix A and B as vector M and vector N, respectively and the length of vector M is same as vector N’s. So \( x_i \) is the expression value of the sample \( i \) in the M and \( y_j \) is the value of the sample \( j \) in the N. s represents the clarity measures. We employ feature selection using Wilcoxon rank sum test [20] that is following

Figure 1. The fusion strategy for compressive imaging.
Figure 2. The training data and the selected bases. (a) Fourteen pairs of multi-focus images; (b) Fifteen selected Hadamard bases.

$$S = \sum_{i\in M} \sum_{j\in N} P\left( (x_i - y_j) \leq 0 \right),$$ (3)

where $P$ is the logic judge function, if the logic expression in the bracket is true, the value of $P$ is 1, or else it is 0. When $s$ is closer to zero or closer to the value of $m \times n$ ($m = |M|$, $n = |N|$), the corresponding row of coefficient matrix is more important to the feature classification. The importance degree of each row of coefficient matrix be calculated.

$$w = \max\left( (s, mn - s) \right).$$ (4)

According to $w$ of each row of coefficient matrix, the row of coefficient matrix can be ranked. The bigger the value of $w$ is, the more Hadamard basis contributes. So the Hadamard bases are selected from Hadamard matrix responding to the top $k$ rows of coefficient matrix. Figure 2(b) presents fifteen Hadamard bases derived from above method.

3.2. Fusion Scheme

Given two multi-focus images $I_k$, $k = \{A, B\}$ with size of $m \times n$, they are transformed into vectors $\{X_i\}_{i=1}^{mn}$ using CI hardware. The pixel values in the source blocks can present clear structural information of source image. However, by applying a randomize measurement matrix, the obtained measurements do not preserve the structural information of source blocks any more. In compressive imaging fusion, the focus-level evaluating cannot be obtained by calculating randomized sampling coefficients. Furthermore, it is a key for multi-focus images to evaluate focus-level. In the above section, clarity measures bases are obtained by proposed method, which are integrated a clarity measures matrix $\Phi' \in \mathbb{R}_+^{P \times N}$ ($P$ the number of Hadamard bases) for compressive imaging system. Then, measurement vectors $H_A^1, H_A^2, \ldots, H_A^i$ and $H_B^1, H_B^2, \ldots, H_B$ can be produced by $H_k^i = \Phi'X_i$, $k = \{A, B\}$. In addition, $H'$ can preserve the structural information of source images and have correlation between each image block comparing with randomize measurements. It is observed that the measurements vectors $H'$ are greatly sensitive to sampling rate and block size varies. So the focus-level evaluating exist some difference under various conditions. The coefficients $c'$ are absolute value sum of measurements $H'$ respectively, which characterize evaluating focus-level. In theory, compared with signal-by-signal recovery, joint recovery should provide reliable reconstruction quality from a given collect of measurement vectors, which is equivalent to reduce the measurement burden needed to get better reconstructed quality. The sampling coefficients $y_A$ and $y_B$ obtained by CS are fused using the maximum rule.

$$y_A^i = \begin{cases} y_A^i & \text{if } c'_A \geq c'_B \\ y_B^i & \text{Otherwise} \end{cases},$$ (5)

$$y_B^i = \begin{cases} y_B^i & \text{if } c'_B \geq c'_A \\ y_A^i & \text{Otherwise} \end{cases},$$ (6)

$$y_F^i = \frac{(y_A^i + y_B^i)}{2},$$ (7)
where \( c^i_a \) and \( c^i_b \) are the \( i \)th coefficients of corresponding \( X^i_a \) and \( X^i_b \), respectively. The \( y_f \), the average of measurement vector \( y_1 \) and \( y_2 \), also includes enough information to reconstruct resultant fused image \( I_f \). Furthermore, it is very important to use the median filter in fusion scheme, which can make fused images smooth and reduce blocking artifacts. Combined with all the vectors, the fused image \( I_f \) is reconstructed via a single stage of SPL iterations. In [21], SPL used hard threshold with Wiener filter. Using directional transforms can preserve interesting features of traditional transforms in [18].

### 3.3. Experimental Results and Analysis

In order to evaluate performance of the proposed fusion scheme, two sets of multi-focus source images which contain two natural multi-focus image pairs and ten artificial multi-focus image pairs are used to perform the experiments. The natural multi-focus image pairs contain the common used Clock and Lab images with size of 256 × 256 and 256 × 384, respectively. In addition ten standard images are used as ground truth. For each image, two blurred artificial images are created by Gaussian blurring at the left part and right part respectively. Then, the blurred images with different focus points are taken as the source images.

In this work, we select two objective evaluation metrics \( Q^{ab/f} \) [22] and \( Q_w \) [23] which have been proved to be validated in large degree to perform the quantitative performance evaluation. To reflect a better fusion result, the values of metrics should be as close to 1 as possible.

To demonstrate the performance of the proposed method, some well-known multi-resolution based methods including the discrete wavelet (DWT) and gradient pyramid as well as Luo’s method are used to compare with the proposed method. For DWT method, at first, the source images are reconstructed by using sensing coefficients. Then the reconstructed images are fused using based on wavelet method and a final image is gotten. In the gradient pyramid method, first of all, source multi-focus images are reconstructed via using sensing coefficients. Then under these conditions including decomposition levels 2, highpass component adopted absolute maximum and approximate images taken averaging method, the final fused images is obtained by applying an inverse transformation. In method of Luo et al., block Walsh-Hadamard transform matrix is chosen as measurement matrix. Here, the dual-tree discrete wavelet transform (DDWT) [24] is selected as reconstructed method. Contrasting to traditional DWT and discrete cosine transform (DCT), the DDWT has a better preference with shift invariance and significant directional selectivity.

Without loss of universality, we provide results at three sampling rates 0.3, 0.5, and 0.7, respectively. Corresponding to each sampling rates, three kinds of block sizes 8, 16, and 32 are individually elected. Then, for all the tested sampling rates and block sizes, the measurements of Clock and Lab would be changed, which is resulted form the increased details of images. The objective evaluation results of two experiments are shown in Table 1, Table 2, respectively. The results in Table 1 and Table 2 indicate that the proposed method achieves superior fusion results compared with other methods. Though the measurements with absolute maximum cannot directly present salient features of images, it can preserve the contrast information of the source images. Figure 3 presents fusion examples about Clock source images with different focuses.

The fused images Figure 3(a), Figure 3(b) and Figure 3(e) have different subjective visual qualities. When sampling rate is set to 0.3 or 0.5, some blocking artifacts are observed on both large and small clock, but the resultant images display better visual effect with sampling rate 0.7. Supposing that sampling rate 0.7 is constant, the quality of fused images are changed along with block-size varies. The analysis from fused results indicates that Figure 3(c) and Figure 3(d) both emerge several significant breakages, while the resultant images display better performance with block-size 16 than other results.

In Clock experiment, the maximum values of \( Q^{ab/f} \) and \( Q_w \) are corresponding to block-size16, but corresponding to block-size 8 in Lab experiment. However, combining with subjective visual sense, the resultant fused images display better performance at block-size16, because fused images become clearer and the values of \( Q^{ab/f} \) and \( Q_w \) also become larger at the same time along with increasing sampling rates. Even though the fused images of the proposed method become smoother, it also improves the quality of the fused images. The fused results of the DWT are obtained using the same reconstructed algorithm, which introduced some artifacts from edges. Figure 4 shows the fused images of Lab which are fused by the proposed method, gradient pyramid, DWT and method of Luo et al. with sampling rate 0.7 and block-size 16, respectively. When two multi-focus images are individually reconstructed via method of CS, two preferable images are obtained. Then a resultant image is fused using above two images by the different methods, such as gradient pyramid or DWT. Each original
Table 1. Quantitative assessments of various fusion methods for clock images.

<table>
<thead>
<tr>
<th>$r$</th>
<th>$S$</th>
<th>$Q_r$</th>
<th>$Q^{ws}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>0.818 0.791 0.569 0.357</td>
<td>0.709 0.697 0.575 0.289</td>
<td></td>
</tr>
<tr>
<td>0.3</td>
<td>16</td>
<td>0.824 0.773 0.535 0.377</td>
<td>0.705 0.687 0.557 0.301</td>
</tr>
<tr>
<td>32</td>
<td>0.812 0.774 0.546 0.404</td>
<td>0.690 0.683 0.553 0.319</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>0.820 0.773 0.539 0.476</td>
<td>0.713 0.698 0.552 0.391</td>
<td></td>
</tr>
<tr>
<td>0.5</td>
<td>16</td>
<td>0.832 0.766 0.531 0.511</td>
<td>0.715 0.693 0.553 0.425</td>
</tr>
<tr>
<td>32</td>
<td>0.815 0.769 0.533 0.546</td>
<td>0.700 0.689 0.547 0.450</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>0.825 0.769 0.529 0.555</td>
<td>0.723 0.703 0.549 0.477</td>
<td></td>
</tr>
<tr>
<td>0.7</td>
<td>16</td>
<td>0.8350.768 0.523 0.585</td>
<td>0.722 0.698 0.542 0.515</td>
</tr>
<tr>
<td>32</td>
<td>0.816 0.769 0.525 0.620</td>
<td>0.709 0.695 0.539 0.536</td>
<td></td>
</tr>
<tr>
<td>average</td>
<td>0.822 0.772 0.537 0.492</td>
<td>0.709 0.693 0.549 0.411</td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Quantitative assessments of various fusion methods for lab images.

<table>
<thead>
<tr>
<th>$r$</th>
<th>$S$</th>
<th>$Q_r$</th>
<th>$Q^{ws}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>0.889 0.847 0.597 0.470</td>
<td>0.688 0.661 0.504 0.241</td>
<td></td>
</tr>
<tr>
<td>0.3</td>
<td>16</td>
<td>0.885 0.844 0.569 0.504</td>
<td>0.674 0.654 0.492 0.266</td>
</tr>
<tr>
<td>32</td>
<td>0.862 0.834 0.575 0.526</td>
<td>0.661 0.649 0.486 0.291</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>0.891 0.835 0.570 0.588</td>
<td>0.688 0.658 0.483 0.347</td>
<td></td>
</tr>
<tr>
<td>0.5</td>
<td>16</td>
<td>0.888 0.831 0.557 0.616</td>
<td>0.682 0.655 0.483 0.383</td>
</tr>
<tr>
<td>32</td>
<td>0.867 0.829 0.551 0.642</td>
<td>0.671 0.651 0.477 0.411</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>0.893 0.830 0.545 0.652</td>
<td>0.695 0.661 0.475 0.435</td>
<td></td>
</tr>
<tr>
<td>0.7</td>
<td>16</td>
<td>0.889 0.828 0.544 0.685</td>
<td>0.694 0.657 0.476 0.477</td>
</tr>
<tr>
<td>32</td>
<td>0.870 0.827 0.538 0.695</td>
<td>0.680 0.656 0.475 0.490</td>
<td></td>
</tr>
<tr>
<td>average</td>
<td>0.882 0.834 0.560 0.598</td>
<td>0.681 0.656 0.483 0.371</td>
<td></td>
</tr>
</tbody>
</table>

Figure 3. Fusion result of Clock: (a) and (b) Sampling rate 0.3, 0.5 with block-size 16; (c) (d) Block-size 8, 32 with sampling rate 0.7; (e)-(h) Four methods with sampling rate 0.7 and block-size 16.
image is reconstructed by using sensing coefficients, which maybe change its feature, but sometimes also produce explicit results. For the same sampling rate and block-size, the quality of fused images via gradient pyramid-based after reconstructed is superior to those of DWT. Comparing to the other methods, the values of two metrics of the proposed method is larger than that. We can see that the fused image obtained by other methods cannot get more salient feature from two source images. It is observed that the resulting fused images of DWT contain some artifacts from blocking and loss more complementary information. Besides, the proposed scheme achieves better performance and has superiority to the method of Luo et al. to some extent. Furthermore, it takes us short time to fuse source images by compressive sensing. So it can greatly improve efficiency in program run.

In addition to classical multi-focus images, some natural images blurred by Gaussian filter are formatted artificial multi-focus images. The proposed method is applied to these images with sampling rate 0.7 as well as block-size 16. Figure 5 shows ten standard images with same size of 256 × 256. Table 3 lists the quantitative results by using two quality measures, which demonstrate that the proposed approach is superior to other methods and confirm successfully that the proposed approach has lots of advantages in fusing the artificial multi-focus images.

4. Conclusion

In this work, we present a novel image fusion scheme about efficient compressive multi-focus image fusion. The fusion rule is guided by clarity measures and fused image is reconstructed based on blocked CS. The experiments demonstrate that the proposed method provides superior fused images in terms of subjective visual sense as well as objective estimated index. Not only can this method present very clear fused result, but also accelerate the time of multi-focus image fusion. Because of compressing during sensing, fusing less data and reconstructing only one image, we usually spent about 30 seconds in fusing a pair of multi-focus source images on PC. So it greatly improves the efficiency of processing for multi-focus images fusion.

Figure 4. Fusion result of Lab: (a) The proposed method; (b) Gradient pyramid; (c) DWT; (d) Method of Luo et al.

Figure 5. Ten standard images.
Table 3. Quantitative assessments of various fusion methods for artificial multi-focus images.

<table>
<thead>
<tr>
<th>Images</th>
<th>$Q_r$</th>
<th>$Q^{avg}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.8730.8400.6240.737</td>
<td>0.8890.7240.7080.583</td>
</tr>
<tr>
<td>2</td>
<td>0.9050.8700.6960.754</td>
<td>0.9120.7120.6960.551</td>
</tr>
<tr>
<td>3</td>
<td>0.8970.8700.6580.736</td>
<td>0.9130.7110.7000.513</td>
</tr>
<tr>
<td>4</td>
<td>0.8990.8660.6770.777</td>
<td>0.9080.7210.7100.570</td>
</tr>
<tr>
<td>5</td>
<td>0.8720.8450.4890.645</td>
<td>0.8970.6880.6890.458</td>
</tr>
<tr>
<td>6</td>
<td>0.8880.8640.4950.625</td>
<td>0.9060.6770.6740.416</td>
</tr>
<tr>
<td>7</td>
<td>0.8750.8310.5510.723</td>
<td>0.8890.6990.6900.520</td>
</tr>
<tr>
<td>8</td>
<td>0.8860.8700.5340.659</td>
<td>0.9110.6730.6750.421</td>
</tr>
<tr>
<td>9</td>
<td>0.9100.8780.7220.782</td>
<td>0.9130.7610.7490.663</td>
</tr>
<tr>
<td>10</td>
<td>0.90750.8780.6310.748</td>
<td>0.9180.6760.6640.440</td>
</tr>
<tr>
<td>average</td>
<td>0.8910.8610.6080.719</td>
<td>0.9060.7040.6950.513</td>
</tr>
</tbody>
</table>

Acknowledgements

This paper is supported by the National Natural Science Foundation of China (Nos. 61102108, 11247214 and 61172161), Scientific Research Fund of Hunan Provincial Education Department (Nos. 11C1101 and 12A115), and the construct program of key disciplines in USC (No.NHXK04).

References


