Volume 3, Issue 8, August 2013 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

Multifocus Image Fusion Using Sparse Representation by Adaptive Feature Matching

Madhu C K[1],
PG Student, Department of CS&E,
AIT, Chikmagalur,
Karnataka, India

Arjun B C[2],
Assistant Professor,
HOD, Department of IS&E, RIT, Hassan,
Karnataka, India

Prapulla Kumar M S[3],
Assistant Professor,
Department of CS&E,
RIT,Hassan, Karnataka, India

Abstract: Optical imaging cameras suffer from the problem of limited depth-of-field of optical lenses, so it is difficult to get an image with all objects in focus. One way to overcome this problem is by using multi-focus image fusion technique, in which several images with different focus points are combined to form a single image with all objects fully focused. So, it is crucial to effectively extract the image information of the original images and reasonably combine them into the final fusion image. The main concern regarding this approach is to represent source image in sparse coefficients from an over-complete dictionary and fuse by OMP algorithm. The proposed method is tested on several multifocus images and compared with Discrete Wavelet Transform based method. The existing method had many lack of disabilities to produce effective image which is of reduced blurring effects and shift variance, which may overcome by this method with better outcomes.

Keywords—Image Fusion, OMP algorithm, Multifocus Image, Sparse Representation, Over complete Dictionary

I. Introduction

Image fusion is widely recognized as an important aspect of information processing. It consists of combining information from several sources in order to improve the decision making process. In particular, multifocus image fusion combines images that depict the same scene but they are not in-focus everywhere. The task seeks to reconstruct an image which is more suitable for human visual perception and computer processing tasks, such as segmentation, feature extraction, and object recognition [2].Optics of lenses with a high degree of magnification suffers from the problem of limited depth of field. The larger the focal length and magnification of the lens, the smaller the depth of field becomes. As a result, it is often not possible to get an image that contains all relevant objects in focus. However would like to acquire images that have large depth of field, i.e., the images are in focus everywhere. A possible way to overcome this problem is by image fusion, in- which one can acquire a series of pictures with different focus settings and fuse them to produce an image with extended depth of field [1]-[4].

An Important preprocessing step in image fusion is image registration, i.e., corresponding pixel positions in the source images must refer to same location [5]. Basically, the fusion technique can be divided into two categories. One is the spatial domain-based methods, which directly select pixels or regions from clear parts in the spatial domain-based methods, which directly select pixels or regions from clear parts in the spatial domain to compose fused images [2]-[4]. Another category is combing the coefficients in multiscale transform domain under the assumption that image details are contained in the frequency sub-bands [5]-[8].

The simplest fusion method in spatial domain, which is to take the average of the source images pixel-by-pixel, would lead to several undesired side effects such as reduced contrast. To improve the quality of fused image, some more reliable methods are proposed to fuse source images with divided blocks or segmented regions instead of single pixels [2]-[4]. For the region based method, it not only considers the corresponding coefficients are in. First the source images are segmented, and the obtained regions are then fused using their properties, such as spatial frequency or SG. The segmentation algorithms usually complicated and time consuming [9]. Different from multiscale transformations, the sparse representation using over complete dictionary that contains prototype signal atoms describes signals by sparse linear combinations of these atoms [9]-[11]. Two main characteristics of sparse representation are it's over completeness and sparsity [12]. Over complete means that the number of basis atoms in the dictionary exceeds the number of image pixels or signal dimensions. The over complete dictionary that contains rich transform bases allows for more stable and meaningful representation of signals. Sparsity means that the coefficients corresponding to signal are sparse, that is to say, only "a few descriptions" can describe or capture the significant structure information about the object of interest. Benefiting from its sparsity and over completeness, sparse representation theory has successfully been applied in many practical applications, including compression, denoising, feature extraction, classification, and so on [11]-[13]. Earlier studies have shown that common image features can also be accurately described by only a few coefficients as the salient features of the images; we design a sparse representation (SR)-based image fusion scheme. In general, sparse representation is a global operation, in the sense that it is based on the gray-level contact of an entire image. However the image fusion quality depends on the accurate representation of the local salient features of source images. Therefore, a "sliding window" technique is adopted to achieve better performance in capturing local salient features and keeping shift invariance.

Matching pursuit (MP) [1] is a greedy strategy to decompose the signal into linear combination of atoms selected from redundant dictionary. Orthogonal matching pursuit (OMP) [14] is modified matching pursuit, which leads quicker convergences. In this paper, we attempt to use the orthogonal matching pursuit to implement image fusion. Using the OMP algorithm, different kinds of local frequency components from source images are literally extracted and fused together after comparison. Considering the proposed algorithm is redundant and overlapped, blocking or ring artifacts of the fused image can be efficiently avoided in certain degree. This feature can help us to preserve main features of source images in the fused image and ignore some unimportant details.

II. Sparse Signal Representation

Sparse signal representations are useful in different applications. In commonly used block-oriented transforms, e.g. the DCT, and the more recent wavelet based decomposition methods, the sparseness is introduced through thresholding of the transform or wavelet coefficients. Thus only a limited number of non-zero coefficients represent the signal. This introduces errors in the reconstructed signal. The goal is to minimize these errors while fulfilling a sparsity constraint making the number of non-zero coefficients small compared to the number of samples in the original signal. The non-zero coefficients as well as their position information constitute a sparse representation of the signal, useful in many applications like compression, feature extraction, modeling and classification.

Sparse coding is the process of computing the representation coefficients, x, based on the given signal y and the dictionary D. This process commonly referred to as "atom decomposition", requires solving:

$$(P_{0,e}) \min ||x||_0 \text{ subject to } ||y - Dx||_2 \leq \varepsilon, \qquad (1)$$

Where $||x||_0$ denotes the number of non zero components in x.

Algorithms for finding approximating solutions have been extensively investigated and indeed, several effective decomposition algorithms are available. Thorough theoretical work studied the quality of these algorithms solution, in order to evaluate their similarity to the exact solutions. In all those methods, there is a preliminary assumption that the dictionary is known and fixed. In this research we address the issue of designing the proper dictionary, in order to better fit the sparsity model imposed.

This is typically done by a pursuit algorithm that finds an approximate solution, as exact determination of sparsest representation proves to be a NP-hard problem [12]. The simplest pursuit algorithms are the Matching pursuit and Orthogonal matching pursuit (OMP). These are the greedy algorithms that select the dictionary atoms sequentially These methods are very simple involving computation of inner products between signal and the dictionary atoms, and possibly deploying some least square solvers. The approximation derived from a matching pursuit can be refined by orthogonalizing the directions of projection. The resulting orthogonal pursuit converges with a finite number of iterations in finite dimensional space, which is not the case for non-orthogonal pursuit.

III. Orthogonal Matching Pursuit Algorithm

Greedy strategy starts with an empty set and iteratively adds one column at a time, which reduces the residual the most. Once the column is added, a sub-problem is solved in order to calculate a locally optimal solution with the expectation of eventually reaching the global optimum. The Matching Pursuit (MP) algorithm is an example of greedy algorithm which is very popular in signal approximation [1]. The algorithm has a simple interpretation: at each iteration, seek the direction along which s has a longest projection. The word "greedy" derive from the idea that the residual should be maximally reduced at every step. For each iteration, the computational complexity is determined by the matrix vector multiplication.

Orthogonal Matching Pursuit (OMP) was developed as an improvement to MP. Compared to MP, after the column is chosen, instead of only updating the new added co-efficient, a sub-problem is solved involving al selected entries as variables in order to minimizing the current residual. After this process, the new residual is orthogonal to all the selected columns [13]. It is formally described in the table 3.1. Similar to MP, after the K^{th} iteration, the algorithm provides an approximation solution with K non-zero terms. The complexity $O(K_nN)$ if the algorithm steps after K iteration which is much better than a through search OMP improves MP in the way it make the most of the selected columns. The improvement in accuracy comes with a sacrifice in the algorithm efficiency. OMP is more time consuming than MP since it requires the solution to a minimization problem with smaller size.

$$C_{K+1} = \arg_V ||s - \phi_{K+V}|| \qquad (2)$$

There are several alternative options for stopping the algorithm. One may wait until the norm of the residual R_S equals zero, i.e., $\epsilon_0=0$ in our proposed algorithm. This criterion is appropriate to recover a sparse input signal exactly or one may halt the procedure when the residual norm decreases below a specified threshold. The criterion is an error constrained problem. It is also appropriate to halt algorithm after K atoms have been selected to get an approximate for the sparsity constrained problems. Though greedy algorithms are considered to be fast, their efficiency is impacted when the dictionary ϕ is "implicit". An "explicit" matrix refers to a matrix that is explicitly stores, i.e., we have access to any arbitrary entry or column directly. In other words, it is easy to extract a column from an implicit matrix directly. They are usually expressed through some operation.

Table I: OMP Algorithm for approximating (P_O)

Orthogonal Matching Pursuit (OMP):

Goal: Approximate the solution of (P_O): $\min_c \|c\|_o$ s.t $\phi_c = s$.

- Input: ϕ, s_1 an error threshold ϵ_o; Output: c, Γ, where the entries off the indices Γ equal to 0 and on the indices Γ equal to c_Γ.
- Initialize: Input s, error threshold ϵ_o, threshold ϵ_o, residual $R^o s = s$, $c^o = 0$, initial solution support Γ^o =support (c^o)= ϕ.
- Main iteration, at step k.

  -Update Support: choose ϕ_i, from $\phi_{(\Gamma\ )}$, where (Γ) is the complementary of Γ, such that $|(R\ s,\phi_i)|$ and=Γ $\cup${i}.

  -Update Current Solution: Compute c_Γ^{k+1} which minimizes $\|s- \phi_\Gamma^{k+1} c_\Gamma^{k+1}\|_2^2$.

  -Update Residual: $R^{k+1}s = s- \phi_\Gamma^{k+1} c_\Gamma^{k+1}$.

  -Stopping Rule: If $\|R^{k+1}s\|_2 < \epsilon_0$, stop. Otherwise, apply another iteration.

The greedy OMP algorithm selects the atom with the highest correlation to the current residual at each iteration. Once the atom is selected, the signal is orthogonally projected to the span space of the selected atom and the residual will be recomputed. The procedure reports until the stopping criterion is satisfied.

IV. Multifocus Image Fusion

In the proposed method, source images A and B are pair of registered images of same size and F is the fused image. Source images are treated as overlapped images patches of size $\sqrt{n} \times \sqrt{n}$ pixels, ordered lexicographically as columns vectors $x \epsilon R^n$. Then define a dictionary matrix $\epsilon R^{n \times k}$ (with k>n, implying that dictionary is redundant). For natural images, dictionaries such as curvlets, contourlets and steerable wavelets are all preferable. Here, we choose separable discrete gabor dictionary [15]. Assorting to the OMP algorithm, the image patch can be represented by the linear combination of the few vectors from the redundant dictionary D.

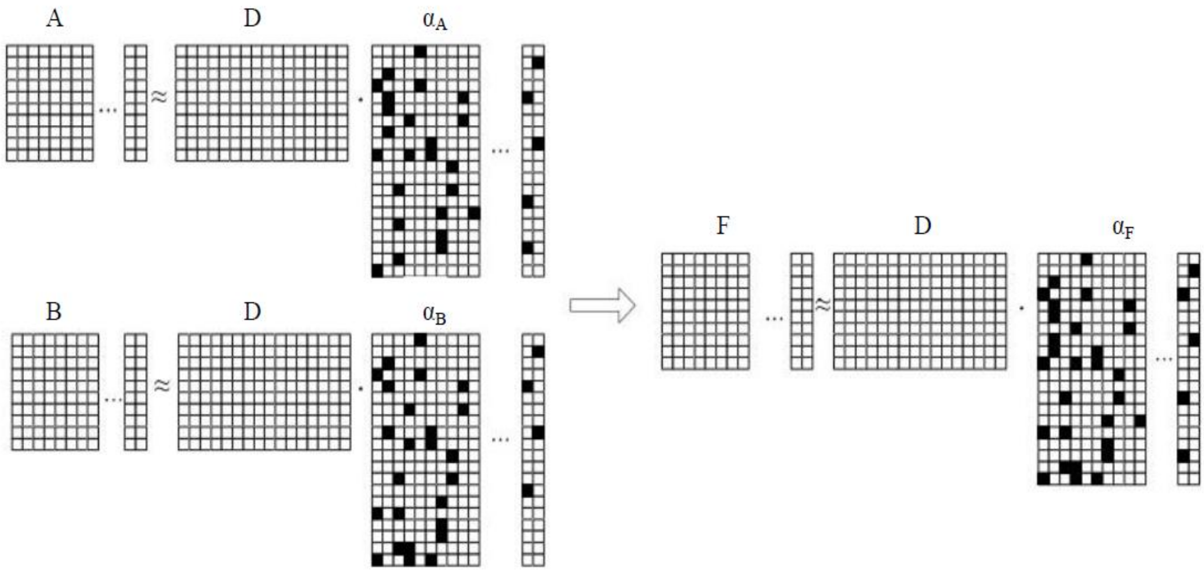


Fig.1: Fusion Framework using OMP

The solution of the equation (1) is very sparse indeed, $\|\alpha\|_o << n$. And these few vectors reflects the main features, such as edges or contours in the image patches, which is most sensitive information to the human visual system. Therefore we could get the approximation of image patch by

$$X_A \approx D\alpha_A \quad (3)$$

$$X_B \approx D\alpha_B \quad (4)$$

The fusion framework using OMP is demonstrated in the fig 1. Firstly the source images A and B are both decomposed by the OMP algorithm and the coefficients matrix α_A and α_B are obtained. The black marks on the columns of the coefficients matrix show which are selected for the sparse representation. Then the coefficients matrixes of the source images are fused together by taking the maximum value of each atom position. It is referred to as α_F, which is the coefficients matrix of the fused image. At last the fused image can be reconstructed by following equation:

$$X_F = D\alpha_F \tag{5}$$

V. Experiments

In order to evaluate the performance of the proposed method, a series of experiments are conducted on three sets of multifocus images: “air balloon”, “wheel” and “plane” as shown in fig.2.

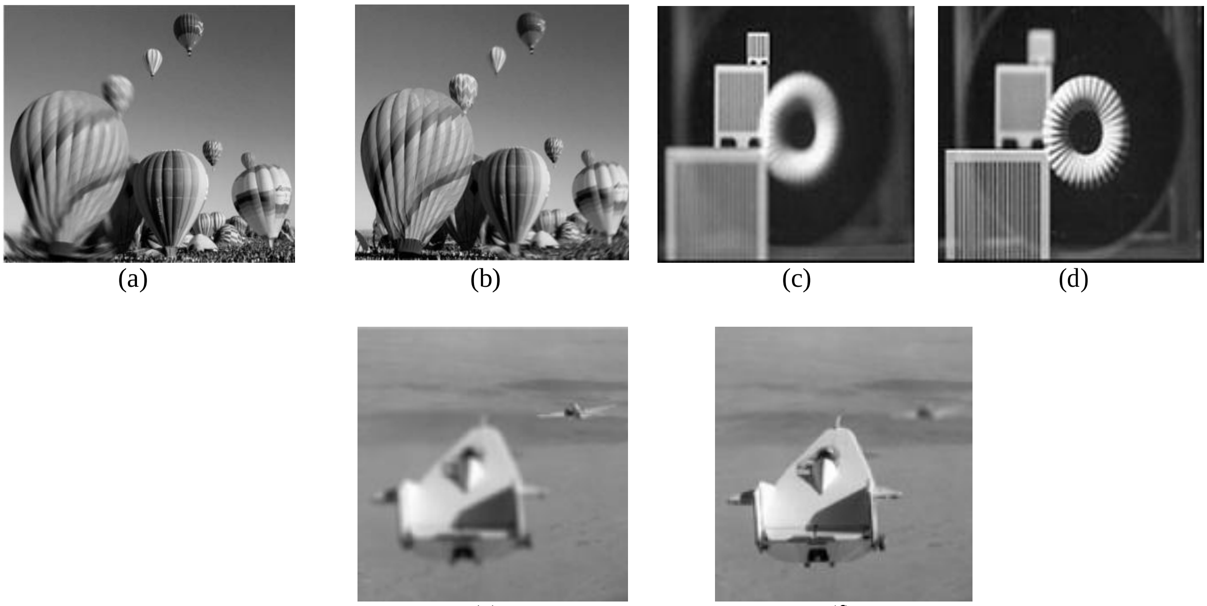
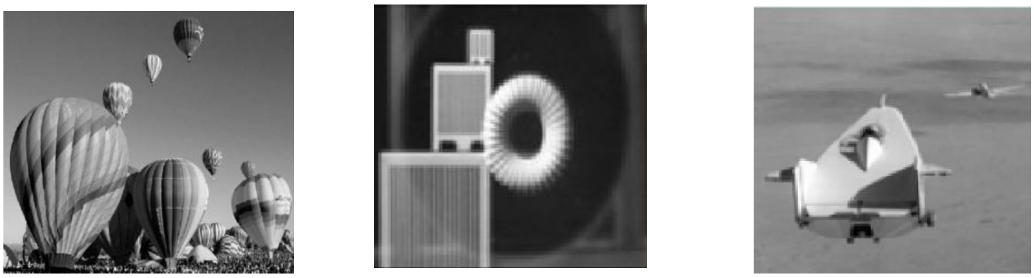


Fig. 2: Multifocus Source Images

Each set contains two sets of source images with different depth of focus. In fig. 2 (a) balloon appear bigger is “in focus” and which is smaller is out of focus. In fig.5 (b) balloon appear smaller is “in focus” and which is bigger is out of focus. In the other images, the left column focus on the near objects and the right column on the far objects.



(a) (b) (c)

Fig 3: Fused Result of Proposed Method

The proposed method is compared with discrete wavelet transform based methods. For the proposed method, the over complete DCT dictionary is used which is obtained by sampling the cosine wave in different frequencies [14]. The over complete DCT dictionary can be implemented fast and has effective performance.

The objective evolution criterion i.e., 1) root mean square error (RMSE) [1] and 2) $Q^{AB/F}$ metric [15] is also used. For better fusion result, the value of the RMSE must be small. The $Q^{AB/F}$ measure should be a close to 1 as possible. The moving step for the sliding window technique is one pixel.

Table II Objective performances of the different image fusion techniques

Measures	**Methods**	**Fig 5(a) (b)**	**Fig 5(c) (d)**	**Fig 5(e)(f)**
RMSE	DWT	0.0922	0.0634	0.0848
RMSE	SR	0.0453	0.0394	0.492
$Q^{AB/F}$	DWT	0.4765	0.4650	0.388
$Q^{AB/F}$	SR	0.7943	0.6874	0.6104

VI. Conclusion

In this paper, we presented a multi focus image fusion method based on sparse representation theory. The image sparse representation theory to resolve the image fusion problem can simultaneously conduct image fusion and restoration is explored. Also the sliding window technique is introduced to overcome the shift variance problem. Under the restrain of certain approximation error, exploiting the orthogonal matching pursuit algorithm, the image features such as geometry structures can be expressed effectively and fused together. Then the residual images are continued to be fused through the Max Fusion method. Numerous experiments on evaluating the fusion performance have been made and the results show that the proposed method has practical meaning. It outperforms the traditional image fusion methods both in visual effect and objective evaluation criteria even in the case of noise. The performance comparison with different block sizes and global error have been made. Also with the DWT method and demonstrate the feasibility and effectiveness of the proposed method.

References

[1] S.G. Mallat and Z.Zhang, "Matching Pursuits and time frequency dictionaries," IEEE Trans. Sig. Proc., Vol.41, No. 12,p.33973415, Dec.1993.

[2] G.Pajares and J. Cruz: "A Wavelet –based image fusion tutorial," Pattern Recongt., vol.37, no. 9, pp. 1855-1872, Sep.2004.

[3] Z.Zhang and R. S. Blum, "A categorization of multiscale decomposition based image fusion schemes with a performance study for a digital camera application," Proc. IEEE, vol. 87, no. 8, pp. 1315-1326, Aug. 1999.

[4] A.A.Goshtasby and S.Nikolov: „Image Fusion: Advances in the state of the art,‟ Inf. Fusion, vol.8, no.2, pp.114-118, Apr.2007.

[5] S.T.Li, I.T.Kwrk and Y.N.Wang: "Combination of images with diverse focus using the spatial frequency," Inf. Fusion, vol.2, no.3, pp.169-176, Sep.2001.

[6] G. Piella, "A general framework for multiresolution image fusion: From pixels to regions," Inf. Fusion, vol. 4, no. 4, pp.259-280, Dec. 2003.

[7] Ishita Dea and Bhabatosh Chanda, "A simple and efficient algorithm for multifocus image fusion using morphological wavelets" Signal Processing, India, August 2005.

[8] A. Kubota, K. Aizawa, and T. Chen, "Reconstructing dense light field from array of multi focus images for novel view synthesis," IEEE Trans. image Process.vol.16, no.1, pp 269-279, Jan. 2008.

[9] P.L.Lin and P.Y.Huang, "Fusion methods based on dynamic segmented morphogical wavelet or cut and paste for multifocus images," Signal Process, vol.88, no.6, pp.1511-1527, Jun 2008.

[10] P.T.Burt and E.H.Andelson, "The Laplacian Pyramid as a compact image code," IEEE Trans, commun., vol.31, no.4, pp532-540, Apr.1983.

[11] G.Davis, S.Mallat and M.Avellaneda, "Adaptive greedy approximation," constr, Approx, vol.13, no.1, pp.57-98, Mar.1997

[12] I. F. Gorodnitsky and B. D. Rao, "Sparse signal reconstruction from limited data using FOCUSS: A re-weighted norm minimization algorithm," IEEE Trans. Signal Process., vol. 45, no. 3, pp. 600-616, Mar. 1997.

[13] Y.C.Pati, R.Rezaifar and P.S.Krishnaprasad "Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition", Conf. of Signals, Systems and Computation, Nov.1993.

[14] Bin Liao, Gang Xo and Yuguo Wang, "Multilayered image representation and coding based on mixed transform," Journal of china institute of communication, 2004, 25(6):120-125.

[15] C.S.Xydeas and V.Petrovic: "Objective image fusion performance measure," Electron. Lett, vol.36, no.4, pp: 308 -309, Feb.2000.

