

Comparison study of image fusion based on multi-scale transforms and sparse representation

Bin Sun^{1, a}, Qiao Deng¹, Jiajun Rui¹, Kai Hu¹ and Qi Yang¹

¹School of Aeronautics and Astronautics, University of Electronic Science and Technology of China, Chengdu 611731, China.

^asunbinhust@uestc.edu.cn

Keywords: multi-scale transform, sparse representation, registration, shift-invariance, robust image fusion.

Abstract. Nowadays image fusion becomes very intensive research. Multi-scale transform and sparse representation are typical tools in fusion processing. However, most of efforts are focus on the accurate registration images. A comparative study of MST and SR method over mis-registration are proposed to investigate the robust performance. Five metrics are adopted to evaluate the quality of different methods. The experimental results suggest that the kinds of transform play a crucial role for MST methods, and SR generally shows better performance.

Introduction

Recently many image fusion methods [1] are developed to integrate more comprehensive information from different sensors which provide complementary and redundant information. During the past two decades, multi-scale transform (MST) are the most popular tools [1] used in various image fusion scenarios. These methods assume that the underlying salient information of the source images can be extracted from the decomposed coefficients. Apparently, the selection of transform domain plays a crucial role. But those transform have fixed bases and cannot be used to form an arbitrary family of signals. Recently Yang and Li [2] introduced sparse representation (SR) into image fusion, which utilize the learning-based approach to obtain adaptive bases for source images.

Although both the MST- based and SR-based methods have achieved great success in image fusion [3], it is worthwhile to notice that the quality of fused results will be deteriorated rapidly because of the tiny difference between the pixels of the source images. Furthermore, it is practically hard to make an accurate registration between multimodal sensors. In this paper, we concern comparison of MST and SR methods on mis-registration source images to further investigated about the robust performance.

Image fusion algorithms

Multi-scale transforms. As shown in Fig.1, the detailed fusion scheme using multi-scale transforms contains the following steps: (1) Applying a specific multi-scale transform decomposed the source images into one low-pass and several high-pass sub-bands. (2) Merge L_A and L_B to obtain L_F with the 'average' rule. (3) Merge H_A and H_B to obtain H_F with different rules such as Burts method [4]. (4) Perform the corresponding inverse MST over L_F and H_F to reconstruct the final fused image I_F .

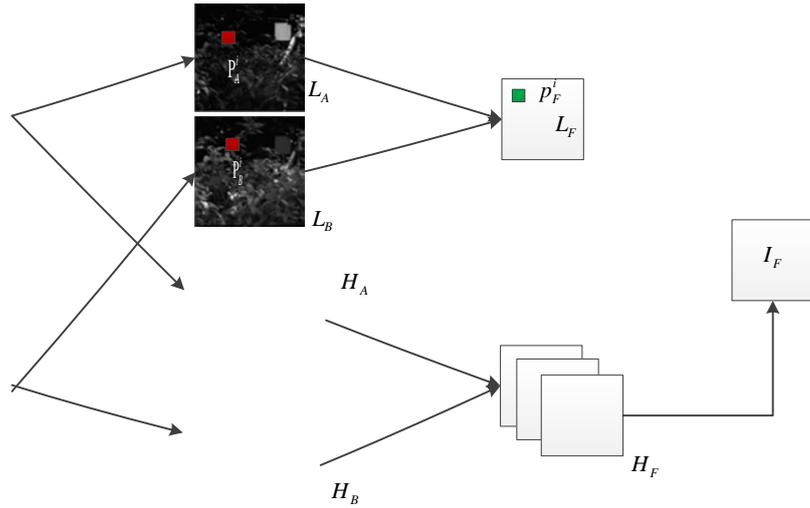


Fig 1. Image fusion using multi-scale transform (MST)

Sparse Representation. The ability of sparse representation is mostly determined by the overcomplete dictionary. Applying K-SVD [5] to obtain the dictionary from training image patches which are randomly sampled from source images.

As illustrated in Fig. 2, the detailed fusion scheme using SR contains the following steps: (1) the source images firstly are sampled using a fixed size slide window, each source images patches are transformed into vectors and arranged to matrix. (2) Then using the OMP method [6] to calculate the sparse vectors. (3) The fused sparse vector is obtained using ‘max-L1’ rule. (4) The vector representation of the fused image v_F can be calculated by $v_F = D\alpha_F$. (5) Finally, the fused image L_F is reconstructed using v_F . Reshape each vector v_F^i in v_F into a block and then arrange the block to I_F at its responding position.

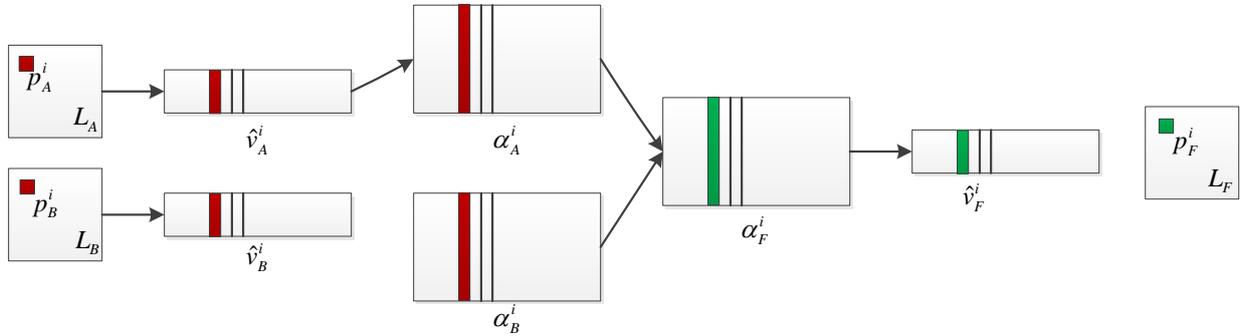
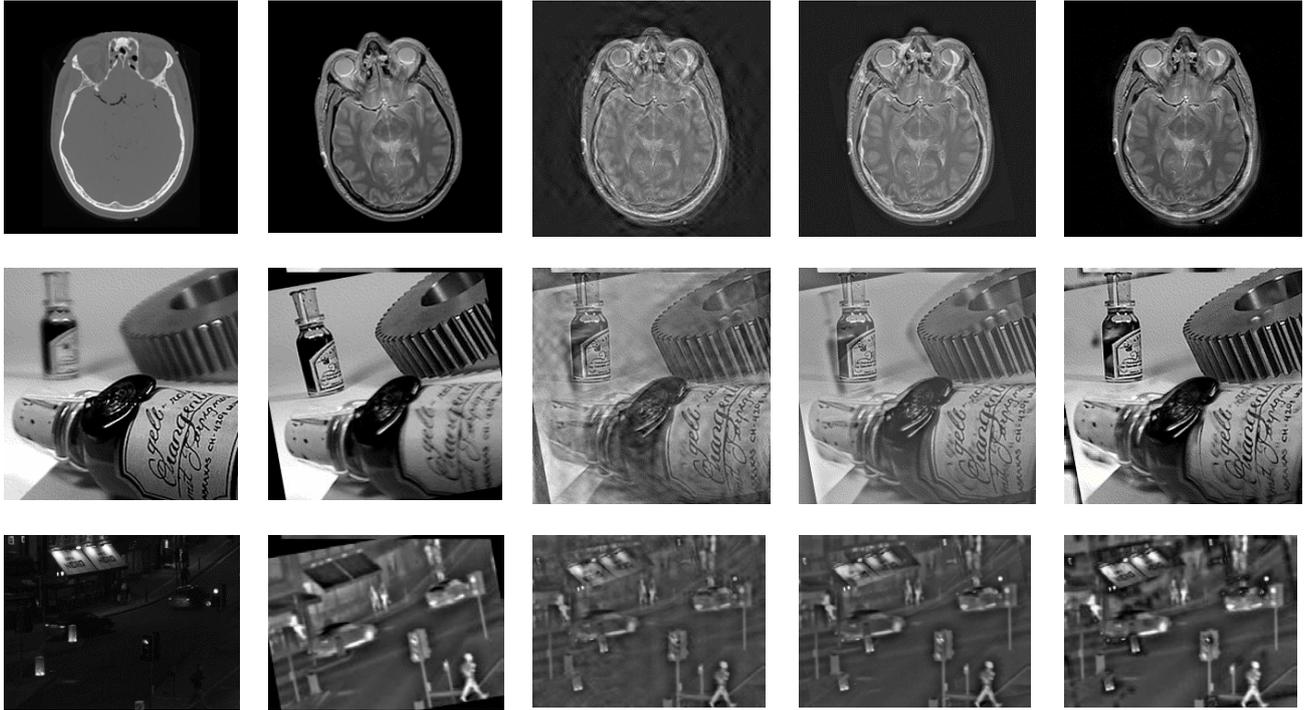


Fig 2. Image fusion using sparse representation (SR)

Experimental setup and results

In this paper, we perform the experiments to investigate the methods includes contourlet transform (CT), nonsubsampling contourlet transform (NSCT) and SR, over three types of images, such as medical, multi-focus and visible-infrared images. To verify the robustness to mis-registration those source images are shifted and rotated relatively.

The visual qualities of the fused images are shown in Fig.3. For medical image fusion, the SR has larger contrast and maintains the details in spatial. However, the CT result has disorder artifacts along the border. For multi-focus image fusion, the fused result of SR has the greatest sharpness, typically on the gear and bottles. For visible-infrared image pairs, the SR has much better performance especially at the words on the roof. Compared with the CT and NSCT, the SR results maintain important spatial details and much more robust to mis-registration image pairs.



(a)Source A (b)Source B (c)CT result (d)NSCT result (e)SR result

Fig 3. Three types of fusion images, the top row is medical , the middle one is multi-focus and the bottom is infrared-visible images. (a) and (b) show the pairs of source images, in which the Source B images are shifted over 5 pixels and rotated 10 degree anticlockwisely. Respectively, (c), (d), (e) show the fusion results obtained using the CT, NSCT and SR.

Table 1 Objective metrics of CT, NSCT and SR methods

Image	metrics	CT	NSCT	SR
Medical	VIFF	0.3499	0.3718	0.3568
	FMI	0.8515	0.8582	0.8799
	SSIM	0.5808	0.6839	0.7093
	MI	1.8915	2.2787	3.3556
	Qabf	0.4561	0.5612	0.6165
Multi-focus	VIFF	0.3663	0.4418	0.5475
	FMI	0.7591	0.7627	0.7846
	SSIM	0.3614	0.4501	0.5488
	MI	1.8258	2.0989	3.8013
	Qabf	0.3206	0.4067	0.4667
Infrared-visible	VIFF	0.2655	0.2996	0.3185
	FMI	0.8721	0.8801	0.8804
	SSIM	0.4002	0.4702	0.4256
	MI	1.0309	1.3599	2.6937
	Qabf	0.3812	0.5080	0.4619

Furthermore, we employ five metrics, i.e. visual information fidelity (VIFF) [7], Feature Mutual Information (FMI) [8], structural similarity index measure (SSIM) [9], Mutual Information (MI) [10], and Edge preservation values (Qabf) [11] to quantitatively evaluate the fusion performance. Table 1 illustrated the results obtained by CT, NSCT and SR for each metric. The best results are labeled in bold. It can be seen that the SR generally shows better performance than CT and NSCT.

Discussion and Conclusions

It should be a conclusion that NSCT performs better than CT. For CT, due to the upsampling operator, the error will be enlarged and has artifacts in fused results, especially for mis-registration condition. Moreover, the SR performs better than NSCT. This is because the sample patches used to learn the dictionary are overlapped, which contain redundant information to reconstruction. It is worthy to notice that image registration and fusion have much common processing. Both of them need similar measurements to determine what kinds of information should be matched and transferred. It will be a direction which might lead to more robust method in practical research.

Acknowledgments

The corresponding author Bin Sun thanks Yu Liu from University of Science and Technology of China, Oliver Rockinger from Technical University of Kaiserslautern, and Prof. Zheng Liu from Toyota Technological Institute (Japan) for sharing some images and codes used in the publications. The work is supported by the Fundamental Research Funds for the Central Universities of China (No. ZYGX2015J122 and No.ZYGX2015KYQD032), and New Characteristic Teaching Material Construction (No. Y03094023701019427).

References

- [1] Li, S., B. Yang and J. Hu, Performance comparison of different multi-resolution transforms for image fusion, *Information Fusion*. 12 (2011) 74-84.
- [2] Bin, Y., L. Shutao, Multifocus image fusion and restoration with sparse representation, *IEEE Transactions on Instrumentation and Measurement*. 59 (2010) 884-892.
- [3] Liu, Y., S. Liu and Z. Wang, A general framework for image fusion based on multi-scale transform and sparse representation, *Information Fusion*. 24 (2014) 147-164.
- [4] Rockinger O. Image Sequence Fusion Using a Shift-Invariant Wavelet Transform, *IEEE Computer Society. International Conference on Image Processing*. 3 (1997) 288-291.
- [5] M. Aharon, M. Elad, A. Bruckstein, K-svd: an algorithm for designing overcomplete dictionaries for sparse representation, *IEEE Trans. Signal Process*. 54 (11) (2006) 4311-4322.
- [6] S. Mallat, Z. Zhang, Matching pursuits with time-frequency dictionaries, *IEEE Trans. Signal Process*. 41 (12) (1993) 3397-3415.
- [7] Han Y, Cai Y Z, Cao Y and Xu X M. A new image fusion performance metric based on visual information fidelity, *Information Fusion*. 14 (2013) 127-135.
- [8] Haghghat, M.B.A., A. Aghagolzadeh and H. Seyedarabi, A non-reference image fusion metric based on mutual information of image features. *Computers & Electrical Engineering*, 37 (5) (2011) 744-756.
- [9] Zhou Wang, A.C.B.H., Image Quality Assessment: From Error Measurement to Structural Similarity. *IEEE Transactions on Image Processing*, 13 (4) (2004) 600-612.
- [10] G. Qu, D. Zhang, P. Yan, Information measure for performance of image fusion, *Electron. Lett*. 38 (7) (2002) 313-315.
- [11] Xydeas, C.S. and V. Petrovic, Objective image fusion performance measure. *Electronics Letters*, 36 (2000) 308-309.