SCIENCE CHINA
Information Sciences

• LETTER •

April 2018, Vol. 61 049103:1–049103:3
https://doi.org/10.1007/s11432-017-9246-3

Visible and infrared image fusion using ℓ_0-generalized total variation model

Han PAN*, Zhongliang JING*, Lingfeng QIAO & Minzhe LI

School of Aeronautics and Astronautics, Shanghai Jiao Tong University, Shanghai 200240*, China*

Received 29 May 2017/Revised 23 August 2017/Accepted 19 September 2017/Published online 5 March 2018

Citation Pan H, Jing Z L, Qiao L F, et al. Visible and infrared image fusion using ℓ_0-generalized total variation model. Sci China Inf Sci, 2018, 61(4): 049103, https://doi.org/10.1007/s11432-017-9246-3

Dear editor,

Image fusion aims to generate a single composite image by combining information from two or more images. The resulting image is more informative and suitable for visual perception. This technology can improve the robustness of various vision tasks, e.g., visual tracking [1]. Different fusion methods, such as the ratio of low-pass pyramid (RLPP) [2], dual tree-complex wavelet transform (DTCWT), nonsubsampled contourlet transform (NSCT) [3] and curvelet [4], have been developed for integrating low-level features of multi-source images. Despite the great progress of image fusion, it remains a challenging task to enforce spatial or spectral coherence in the fusion process. Recently, optimization based fusion method [5] has been attracted much attention.

A common approach to preserve the spatial coherence in the process of image fusion is total variation (TV) minimization method. This scheme has proved to be robust to noise. Ma et al. [5] proposed an image fusion model involving TV minimization problem with ℓ_1-norm data fidelity. However, this model is far from optimal when the residual data does not obey the Laplace distribution. It is well-known that ℓ_1 norm induces sparsity. But, ℓ_1 norm suffers from the estimation bias because the sparsity profile of natural images is spatially varying. Moreover, ℓ_1-TV based fusion model does not take into account for high correlations in images of natural scenes, which may degrade the fusion performance.

To alleviate these problems, we propose a novel fusion model by considering ℓ_0-norm data fidelity with generalized total variation regularization term. Then, the equivalent half complementarity formulation of the ℓ_0 minimization problem is considered for obtaining a more sparse solution. An optimization framework with proximal alternating direction method of multipliers (PADMM) is developed for solving the resulting problem. Finally, experimental results with public available datasets show that the proposed method outperforms other state-of-the-art methods.

Principle of the proposed method. The main idea of the proposed method is to reformulate the fusion problem as a ℓ_0-generalized total variation minimization problem. This fusion model can combine the thermal radiation of infrared image and local image structures of visible image simultaneously. The generalized total variation [6] is adopted to reduce the staircase effect. However, ℓ_0 minimization problem lacks of convexity. To obtain an accurate solution, we reformulate ℓ_0 norm as an equivalent mathematical program with equilibrium constraints [7]. An efficient minimization scheme is proposed using the method of proximal alternating direction methods of multipliers.

For an image of size $m \times n$, the input images are assumed to be spatially registered. The visible, infrared and fused images are denoted by $v, u, x \in \mathbb{R}^{l\times 1}$, respectively, where $l = mn$ stands for the column-vector form of pixel intensities. The fusion model we focus is provided as follows:

* Corresponding author (email: hanpan@sjtu.edu.cn, zljing@sjtu.edu.cn)

$$f(x) = \min_x \|x - v\|_0 + \lambda \|\nabla x - \nabla u\|_p, \quad (1)$$

where ∇ denotes the discrete gradient operator. λ is a positive regularization parameter. $\|\nabla x - \nabla u\|_p$ denotes generalized total variation regularization term, i.e., ℓ_p based anisotropic total variation. Its formulation can be expressed as follows:

$$\sum_{i=1}^{m} \sum_{j=1}^{n} |x_{i+1,j} - x_{i,j}|_p + |x_{i,j+1} - x_{i,j}|_p, \quad (2)$$

where $0 < p \leqslant 1$. When $p = 1$, Eq. (2) reduces to standard anisotropic total variation.

However, minimizing generalized TV fusion problem with ℓ_0-norm data fidelity is non-convex and non-continuous. To handle this, we consider a complementarity formulation with equilibrium constraints [7]. Based on the half-complementarity formulation in [7], the objective function (1) can be rewritten as follows:

$$\begin{aligned} &\min_{0 \leqslant x, z \leqslant 1} \langle 1, 1 - z \rangle + \lambda \|\nabla x - \nabla u\|_p, \\ &\text{s.t.} \quad z \odot |x - v| = 0, \end{aligned} \quad (3)$$

where $\odot$ denotes the element-wise product. It can be seen that (3) is non-convex because of half complementarity constraint $z \odot |x - v| = 0$.

After introducing two auxiliary vectors A_1 and A_2, we reformulate (3) as follows:

$$\begin{aligned} &\min_{0 \leqslant x, z \leqslant 1} \langle 1, 1 - z \rangle + \lambda \|A_1\|_p, \\ &\text{s.t.} \ \nabla x - \nabla u = A_1, x - v = A_2, z \odot |A_2| = 0. \end{aligned} \quad (4)$$

The corresponding augmented Lagrangian function is presented as follows:

$$\begin{aligned} &\mathcal{L}(x, z, A_1, A_2, a_1, a_2, a_3) \\ &= \langle 1, 1 - z \rangle + \lambda \|A_1\|_p + \langle \nabla x - \nabla u - A_1, a_1 \rangle \\ &\quad + \frac{\beta}{2} \|\nabla x - \nabla u - A_1\|^2 + \langle x - v - A_2, a_2 \rangle \\ &\quad + \frac{\beta}{2} \|x - v - A_2\|^2 + \langle z \odot |A_2|, a_3 \rangle \\ &\quad + \frac{\beta}{2} \|z \odot |A_2|\|^2, \end{aligned}$$

where $\beta > 0$ is penalty parameter. a_1, a_2, a_3 are the Lagrange multipliers corresponding to the constraints $\nabla x - \nabla u$, $x - v$ and $z \odot |A_2|$, respectively.

Next, the resulting three subproblems and their solutions are summarized as follows:

(1) x-subproblem. After adding a proximal term $\frac{1}{2}\|x - x^k\|_W^2$ to $\mathcal{L}$, we have

$$\begin{aligned} x^{k+1} = \min_{0 \leqslant x \leqslant 1} &\frac{\beta}{2} \|\nabla x - \nabla u - A_1^k + a_1^k/\beta\|^2 \\ &+ \frac{\beta}{2} \|x - v - A_2^k + a_2^k/\beta\|^2 \\ &+ \frac{1}{2} \|x - x^k\|_W^2. \end{aligned} \quad (5)$$

This subproblem can be rearranged as follows:

$$x^{k+1} = \arg \min_{0 \leqslant x \leqslant 1} \frac{1}{2\kappa} \|x - h^k\|,$$

where $h^k = x^k - \kappa(\nabla^T a_1^k + a_2^k) + \kappa[\beta \nabla^T (A_1^k + \nabla u - \nabla x^k) + \beta(v + A_2^k - x^k)]$. At last, the solution to this subproblem is

$$x^{k+1} = \min(1, \max(0, h^k)).$$

(2) z-subproblem. The equation with respect to z can be rewritten as follows:

$$z^{k+1} = \min_{0 \leqslant z \leqslant 1} \frac{\beta}{2} \|z \odot |A_2|\|^2 - \langle z, c_k \rangle, \quad (6)$$

where $c_k = 1 - A_3^k \odot |A_2^k|$. Then, the solution can be provided as follows:

$$z^{k+1} = \min\left(1, \max\left(0, \frac{c^k}{\beta} A_2^k \odot A_2^k\right)\right).$$

(3) (A_1^{k+1}, A_2^{k+1})-subproblem. A_1^{k+1} is obtained by solving the next minimization problem:

$$A_1^{k+1} = \min_{A_1} \frac{\beta}{2} \|A_1 - t^k\| + \lambda \|A_1\|_{p,1},$$

where $t^k = \nabla u - \nabla x^{k+1} - a_1^k/\beta$. Based on the definition of p-shrinkage thresholding method, we have

$$A_1^{k+1} = \operatorname{sign}(t^k) \odot \max(|t^k| - \lambda^{2-p} |t^k|^{p-1}, 0).$$

For A_2^{k+1}, it is updated by minimizing the following problem:

$$A_2^{k+1} = \min_{A_2} \frac{\beta}{2} \|A_2 - q^k\|^2 + \frac{\beta}{2} \|z^{k+1} \odot |A_2| + a_3^k/\beta\|^2,$$

where $q^k = x^{k+1} - v + a_2^k/\beta$. Then, we have

$$A_2^{k+1} = \operatorname{sign}(q^k) \odot \max\left(0, \frac{|q^k| - a_3^k \odot z^{k+1}/\beta}{1 + z^k \odot z^{k+1}}\right).$$

At last, a brief description of the proposed method based on the optimization framework of PADMM [8] is provided in Algorithm C1 (see Appendix C for detail).

Experimental results. To evaluate the effectiveness of the proposed method, we conducted extensive experiments on three public available datasets, i.e., human factor dataset from Netherlands organization for applied scientific research

(TNO)[1], aerial images[2] and Ohio state university color-thermal dataset[3]. Some samples of these datasets are displayed in Figures B1, B3 and B5 (see Appendix B for detail), respectively. Three objective indexes were used to assess the fusion performance, which consist of mutual information (MI), edge information preservation index [9] $Q^{ab/f}$ and entropy (EN). It should be noted that a larger value means the better fusion performance.

The proposed method was compared to six fusion methods, including RLPP [2], Wavelet, DTCWT, NSCT [3], curvelet [4] and ℓ_1-TV based method [5]. In our experiments, the decomposition level is 4. The regularization parameter λ is 1. p for generalized total variation is 0.5. Some details about the choice of these parameters are provided in Appendix A.

The visual results of all competing methods are presented in Figures B2, B4 and B6 (see Appendix B for detail). Compared with the state-of-the-art methods, the visual results of our method in Figures B2(h), B4(4) and B6 (see Appendix B for detail) provide a more accurate fusion result, and have a significant improvement over the other fusion methods. Moreover, the visual comparisons indicate that the proposed method can preserve the fine image details without obvious artifacts. In contrast, the visual outcomes of wavelet, DTCWT and curvelet based methods may drop local image content, such as sharp edges. The main reason is that the smoothing effect of multi-scale analysis methods may lead to lose image brightness.

Tables B1–B9 (see Appendix B for detail) show the values of various fusion methods in terms of MI, $Q^{ab/f}$ and EN, respectively. The best values of scores are highlighted in bold font. For the metric of EN, our method can yield comparable results. For the metric of MI, the proposed method outperforms the other methods by a significant margin. The improvement on $Q^{ab/f}$ indicates that the proposed method can reconstruct more local details. Based on these numerical results, we can see that the proposed method is superior to the other ones.

Conclusion. We proposed an efficient visible and infrared image fusion method based on ℓ_0-generalized total variation model. To deal with the non-convexity of ℓ_0 norm, the proposed fusion model is reformulated as an equivalent mathematical program with equilibrium constraints. An efficient minimization scheme with proximal alternating direction methods of multipliers is proposed. Experimental results on a variety of datasets demonstrate that the proposed method achieves the state-of-the-art performance.

Acknowledgements This work was jointly supported by National Natural Science Foundation of China (Grant Nos. 61603249, 61673262, 61690210, 61690212), and Key Project of Science and Technology Commission of Shanghai Municipality (Grant No. 16JC1401100).

Supporting information Fusion results on three datasets (Figures B2, B4, and B6 in Appendix B; Tables B1–B9 in Appendix B), detailed algorithm (Algorithm C1 in Appendix C), and extensive discussions on parameters setting (Appendix A). The supporting information is available online at info.scichina.com and link.springer.com. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.

References

1 Yun X, Jing Z L, Xiao G, et al. A compressive tracking based on time-space Kalman fusion model. Sci China Inf Sci, 2016, 59: 012106

2 Toet A. Image fusion by a ratio of low-pass pyramid. Pattern Recogn Lett, 1989, 9: 245–253

3 Da Cunha A L, Zhou J P, Do M N. The nonsubsampled contourlet transform: theory, design, and applications. IEEE Trans Image Process, 2006, 15: 3089–3101

4 Nencini F, Garzelli A, Baronti S, et al. Remote sensing image fusion using the curvelet transform. Inf Fusion, 2007, 8: 143–156

5 Ma J Y, Chen C, Li C, et al. Infrared and visible image fusion via gradient transfer and total variation minimization. Inf Fusion, 2016, 31: 100–109

6 Jie Y, Lu W S. Compressive imaging by generalized total variation minimization. In: Proceedings of 2014 IEEE Asia Pacific Conference on Circuits and Systems (APCCAS), Ishigaki, 2014. 21–24

7 Feng M B, Mitchell J E, Pang J S, et al. Complementarity formulations of ℓ_0 norm optimization problems. 2013. http://www.optimization-online.org/DB_FILE/2013/09/4053.pdf

8 Yuan G Z, Ghanem B. A proximal alternating direction method for semi-definite rank minimization. In: Proceedings of the 30th AAAI Conference on Artificial Intelligence (AAAI-16), Phoenix, 2016. 2300–2308

9 Xydeas C S, Petrovic V. Objective image fusion performance measure. Electron Lett, 2000, 36: 308–309

1) https://doi.org/10.6084/M9.FIGSHARE.1008029.
2) http://www.metapix.de/.
3) http://vcipl-okstate.org/pbvs/bench/.