

MULTIBAND IMAGE FUSION WITH CONTROLLABLE ERROR GUARANTEES

Unni V. S., Ruturaj G. Gavaskar and Kunal N. Chaudhury

Department of Electrical Engineering, Indian Institute of Science, Bengaluru, India

ABSTRACT

Multiband fusion involves combining an image having high spatial and low spectral resolution with another image having low spatial and high spectral resolution—resulting in a single multiband image with high spatial and spectral resolutions. In classical variational techniques, this problem is formulated as the minimization of an objective function consisting of two quadratic data-fidelity terms and an edge-preserving regularizer; the former account for blur, resolution mismatch and additive noise. In this work, we explore a constrained formulation of this problem where the regularization function is minimized subject to hard constraints on the data fidelity. Unlike the penalty approach, the advantage is that the user has direct control on the data fidelity of the reconstruction. We come up with an efficient ADMM solver for this constrained optimization problem. Moreover, for convex regularizers, we prove that the ADMM iterates converge to an optimal solution (this is somewhat standard but requires the verification of certain technical conditions). To our knowledge, the use of constrained optimization for image fusion is novel. The proposed framework is shown to model the observations well and its fusion quality is competitive with state-of-the-art methods.

Index Terms— Hyperspectral image, multiband fusion, Bayesian estimation, regularization, ADMM.

1. INTRODUCTION

Multiband imaging captures image data at specific wavelengths across the electromagnetic spectrum. Spectral imaging makes it possible to extract additional visual information that is not captured by grayscale and color images. A multiband image can be represented by using a 3D cube $\mathbf{Z} \in \mathbb{R}^{n_1 \times n_2 \times L}$, where the first two dimensions are the spatial dimensions ($n = n_1 \times n_2$ pixels) and the third dimension is the spectral dimension (L bands). Examples of multiband images are hyperspectral (HS) images [1], multispectral (MS) images, and images obtained via magnetic resonance spectroscopic imaging (MRSI) [2]. Multiband imaging has a wide range of applications in computer vision and remote sensing—object classification and recognition [1, 3], tracking [4], environmental monitoring [5], etc. With the rapid development of satellite sensors, remote sensing images are now widely used. However, due to the technical limitations of the sensors, existing sensors have to make a fundamental tradeoff between spatial and spectral resolutions [3, 6]. For example, HS images have hundreds of bands but are limited by their relatively low spatial resolution, whereas the opposite is true for MS images. An extreme case is a panchromatic (PAN) image, which has very high spatial resolution but only a single band. Therefore, reconstruction of multiband images with high spatial and spectral resolution is of

great interest since it admits various new applications that to date have only been possible locally via airborne imaging systems having high spatial resolution [1, 3].

Numerous image fusion methods have been developed in the past decade [3]. They can be broadly classified into three categories—component substitution methods (CS), multiresolution analysis (MRA) methods and Bayesian methods. Recently, Bayesian methods have received wider attention due to their superior performance. In the Bayesian framework, the image fusion problem is formulated as an optimization problem derived from a linear forward model and some suitable prior on the ground-truth image. For example, naive Gaussian and sparsity-promoting priors were considered in [7, 8] and in [9] the authors used high-dimensional total-variation as the regularizer. Recently, an approach based on solving a Sylvester equation was proposed, resulting in a closed-form solution with low computational complexity [10, 11].

In this work we focus on two types of multiband fusion—HS-MS fusion [3] and MS-PAN fusion [12]. Unlike existing Bayesian methods where the objective being minimized is the sum of data-fidelity terms and a regularizer (penalty method), we propose to explore the constrained formulation of this problem where we minimize just the regularizer and impose hard constraints on the data-fidelity terms (Section 2). This allows the user to control the fidelity (data consistency) error, which is not possible with existing penalty approaches. To solve the resulting constrained optimization problem, we come up with a provably convergent ADMM solver [13] (Section 3). We present visual and quantitative comparisons with state-of-the-art methods (Section 5) to demonstrate the superior performance of our algorithm.

2. BACKGROUND AND PROBLEM STATEMENT

In this section, we describe the forward model for multiband fusion. Let $\mathbf{Z}_g \in \mathbb{R}^{n_m \times L_h}$ be the target image with high spatial and spectral resolutions (L_h bands and n_m pixels). We need to estimate \mathbf{Z}_g from the observed low spectral resolution image $\mathbf{Y}_m \in \mathbb{R}^{n_m \times L_m}$ and low spatial resolution image $\mathbf{Y}_h \in \mathbb{R}^{n_h \times L_h}$ of the same scene, where $n_h \ll n_m$ and $L_m \ll L_h$. We use a standard forward model [14] where \mathbf{Y}_m and \mathbf{Y}_h are obtained via linear spectral and spatial degradations of \mathbf{Z}_g :

$$\mathbf{Y}_h = \mathbf{S}\mathbf{B}\mathbf{Z}_g + \mathbf{N}_h \quad \text{and} \quad \mathbf{Y}_m = \mathbf{Z}_g\mathbf{R} + \mathbf{N}_m; \quad (1)$$

$\mathbf{B} \in \mathbb{R}^{n_m \times n_m}$ is the spatial blurring operator, $\mathbf{S} \in \mathbb{R}^{n_h \times n_m}$ is the subsampling operator, $\mathbf{R} \in \mathbb{R}^{L_h \times L_m}$ is the spectral degradation operator, \mathbf{N}_h and \mathbf{N}_m are white Gaussian noise. The objective is to estimate \mathbf{Z}_g from \mathbf{Y}_h , \mathbf{Y}_m and the forward model.

Note that \mathbf{Z}_g lies in a very high-dimensional space ($n_m L_h$). To reduce the number of variables, a widely used assumption [9–11] is that \mathbf{Z}_g lies in a subspace of lower dimension $n_m L_s$, where $L_s \ll L_h$. Mathematically, this can be expressed as $\mathbf{Z}_g = \mathbf{X}_g \mathbf{E}$, where $\mathbf{E} \in \mathbb{R}^{L_s \times L_h}$ is a matrix whose columns form a basis of the subspace in question, and $\mathbf{X}_g \in \mathbb{R}^{n_m \times L_s}$ is the image formed by projecting

Correspondence: unniv@iisc.ac.in, ruturajg@iisc.ac.in, kunal@iisc.ac.in. K. N. Chaudhury was supported by Core Research Grant CRG/2020/000527 and SERB-STAR Award STR/2021/000011 from Department of Science and Technology, Government of India.

the pixels of \mathbf{Z}_g onto this subspace. We can estimate \mathbf{E} using principal component analysis or vertex component analysis [15]. Once \mathbf{E} is determined, the problem reduces to one of estimating \mathbf{X}_g that has fewer variables than \mathbf{Z}_g .

Most existing Bayesian methods use maximum a-posteriori estimation to compute \mathbf{X}_g [3, 9, 10]. This reduces to solving an optimization problem of the form

$$\min_{\mathbf{X}} \frac{1}{2} \|\mathbf{SBXE} - \mathbf{Y}_h\|^2 + \frac{\lambda}{2} \|\mathbf{XER} - \mathbf{Y}_m\|^2 + \tau\phi(\mathbf{X}), \quad (2)$$

where $\lambda, \tau > 0$ are parameters, $\phi(\mathbf{X})$ is a convex regularizer, and $\|\cdot\|$ denotes the Frobenius norm. This can be solved using operator-splitting methods such as ADMM [13].

We propose to formulate the task as the following constrained optimization problem:

$$\begin{aligned} & \min_{\mathbf{X}} \phi(\mathbf{X}) \\ & \text{subject to } \|\mathbf{SBXE} - \mathbf{Y}_h\| \leq \delta, \quad \|\mathbf{XER} - \mathbf{Y}_m\| \leq \epsilon \end{aligned} \quad (3)$$

where δ and ϵ are tuneable parameters. This comes with the following advantages:

- Parameters δ and ϵ are readily interpretable: they are the maximum errors that we are willing to tolerate in ensuring consistency of the reconstructed image with the observed data \mathbf{Y}_h and \mathbf{Y}_m (data fidelity). It is thus easier to tune them—for example, by setting them proportional to the variances of \mathbf{N}_h and \mathbf{N}_m . In contrast, the parameters λ, τ in (2) are not easily interpretable, which makes them harder to tune.
- By changing δ and ϵ , the user can directly control the desired tolerance in (3). In contrast, it is difficult to bound the fidelity errors $\|\mathbf{SBXE} - \mathbf{Y}_h\|$ and $\|\mathbf{XER} - \mathbf{Y}_m\|$ by solving (2).

3. NUMERICAL SOLUTION

In this section, we derive an efficient ADMM solver for (3). We begin by defining the balls

$$\mathcal{H} = \{\mathbf{Y} : \|\mathbf{Y} - \mathbf{Y}_h\| \leq \delta\} \quad \text{and} \quad \mathcal{M} = \{\mathbf{Y} : \|\mathbf{Y} - \mathbf{Y}_m\| \leq \epsilon\}.$$

Note that we can write (3) as

$$\min_{\mathbf{X}} \phi(\mathbf{X}) + \iota_{\mathcal{H}}(\mathbf{SBXE}) + \iota_{\mathcal{M}}(\mathbf{XER}), \quad (4)$$

where ι_C denotes the indicator function of a set C : $\iota_C(\mathbf{X}) = 0$ if $\mathbf{X} \in C$ and $\iota_C(\mathbf{X}) = +\infty$ if $\mathbf{X} \notin C$. We apply variable-splitting to rewrite the above problem as

$$\begin{aligned} & \min_{\mathbf{X}, \mathbf{T}, \mathbf{U}, \mathbf{V}} \phi(\mathbf{T}) + \iota_{\mathcal{H}}(\mathbf{U}) + \iota_{\mathcal{M}}(\mathbf{V}), \\ & \text{subject to } \mathbf{T} = \mathbf{X}, \quad \mathbf{U} = \mathbf{SBXE}, \quad \mathbf{V} = \mathbf{XER}. \end{aligned} \quad (5)$$

We solve this using ADMM which requires us to work with the augmented Lagrangian of (5). We call $\mathbf{X}, \mathbf{T}, \mathbf{U}$ and \mathbf{V} the primal variables. The primal update consists of minimizing the augmented Lagrangian of (5), first with respect to \mathbf{X} and then with respect to \mathbf{T}, \mathbf{U} and \mathbf{V} . We denote the dual variables corresponding to the three equality constraints in (5) by Θ, Λ, Γ . We skip the derivation details, which can be found in [13, 18].

Starting with $\mathbf{T}_0, \mathbf{U}_0, \mathbf{V}_0, \Theta_0, \Lambda_0$, and Γ_0 and $\rho > 0$, the updates for $k = 0, 1, 2, \dots$ are

$$\mathbf{X}_{k+1} = \argmin_{\mathbf{X}} \left\{ \|\mathbf{X} - \mathbf{T}_k + \Theta_k\|^2 + \right.$$

```

1 Input: Observations  $\mathbf{Y}_h, \mathbf{Y}_m$ , operators  $\mathbf{B}, \mathbf{S}, \mathbf{R}$ , basis  $\mathbf{E}$ .
2 Initialization:  $\mathbf{T}, \mathbf{U}, \mathbf{V}, \Theta, \Lambda, \Gamma$ .
3 Parameters:  $\delta, \epsilon, \rho$ .
4 while stopping criterion is not satisfied do
5    $\mathbf{X} \leftarrow$  Solve (7) using the method in [11].
6    $\mathbf{T} \leftarrow \text{prox}_{\phi/\rho}(\mathbf{X} + \Theta)$ .
7    $\mathbf{U} \leftarrow \Pi_{\mathcal{H}}(\mathbf{SBXE} + \Lambda)$ .
8    $\mathbf{V} \leftarrow \Pi_{\mathcal{M}}(\mathbf{XER} + \Gamma)$ .
9    $\Theta \leftarrow \Theta + \mathbf{X} - \mathbf{T}$ .
10   $\Lambda \leftarrow \Lambda + \mathbf{SBXE} - \mathbf{U}$ .
11   $\Gamma \leftarrow \Gamma + \mathbf{XER} - \mathbf{V}$ .
12 end

```

Algorithm 1: Estimation of \mathbf{X}_g .

$$\|\mathbf{SBXE} - \mathbf{U}_k + \Lambda_k\|^2 + \|\mathbf{XER} - \mathbf{V}_k + \Gamma_k\|^2 \Big\}, \quad (6a)$$

$$\mathbf{T}_{k+1} = \argmin_{\mathbf{T}} \phi(\mathbf{T}) + \frac{\rho}{2} \|\mathbf{T} - \mathbf{X}_{k+1} - \Theta_k\|^2, \quad (6b)$$

$$\mathbf{U}_{k+1} = \argmin_{\mathbf{U}} \iota_{\mathcal{H}}(\mathbf{U}) + \frac{\rho}{2} \|\mathbf{U} - \mathbf{SBX}_{k+1}\mathbf{E} - \Lambda_k\|^2, \quad (6c)$$

$$\mathbf{V}_{k+1} = \argmin_{\mathbf{V}} \iota_{\mathcal{M}}(\mathbf{V}) + \frac{\rho}{2} \|\mathbf{V} - \mathbf{X}_{k+1}\mathbf{ER} - \Gamma_k\|^2, \quad (6d)$$

$$\Theta_{k+1} = \Theta_k + \mathbf{X}_{k+1} - \mathbf{T}_{k+1}, \quad (6e)$$

$$\Lambda_{k+1} = \Lambda_k + \mathbf{SBX}_{k+1}\mathbf{E} - \mathbf{U}_{k+1}, \quad (6f)$$

$$\Gamma_{k+1} = \Gamma_k + \mathbf{X}_{k+1}\mathbf{ER} - \mathbf{V}_{k+1}, \quad (6g)$$

The \mathbf{X} -update consists of minimizing a convex quadratic function. Setting the gradient of this function to zero, we obtain the equation

$$\mathbf{XC}_1 + \mathbf{C}_2\mathbf{X} = \mathbf{C}_3, \quad (7)$$

where $\mathbf{C}_1 = [\mathbf{I} + \mathbf{ER}(\mathbf{ER})^\top](\mathbf{EE}^\top)^{-1}$, $\mathbf{C}_2 = (\mathbf{SB})^\top \mathbf{SB}$, and

$$\mathbf{C}_3 = [(\mathbf{T}_k - \Theta_k) + (\mathbf{SB})^\top (\mathbf{U}_k - \Lambda_k) + (\mathbf{V}_k - \Gamma_k)(\mathbf{ER}^\top)](\mathbf{EE}^\top)^{-1}.$$

This is a special form of Sylvester's equation which has a closed-form solution [11]. The \mathbf{T} -update can be written as $\text{prox}_{\phi/\rho}(\mathbf{X}_{k+1} + \Theta_k)$, where

$$\text{prox}_{\phi}(\mathbf{Q}) := \argmin_{\mathbf{P}} \frac{1}{2} \|\mathbf{P} - \mathbf{Q}\|^2 + \phi(\mathbf{P}) \quad (8)$$

is the proximal map of ϕ [13]. The proximal map can be evaluated in closed form for a variety of ϕ used in fusion [9, 11]. The \mathbf{U} - and \mathbf{V} -updates do not depend on ρ (due to the presence of the indicator functions) and can be written as

$$\mathbf{U}_{k+1} = \Pi_{\mathcal{H}}(\mathbf{SBX}_{k+1}\mathbf{E} + \Lambda_k), \quad \mathbf{V}_{k+1} = \Pi_{\mathcal{M}}(\mathbf{X}_{k+1}\mathbf{ER} + \Gamma_k),$$

where Π_C denotes orthogonal projection onto the set C . Note that the projections are well-defined since \mathcal{H} and \mathcal{M} are closed and convex. In fact, the projections can be computed in closed form [19]. More precisely, if we let $C = \{\mathbf{Q} : \|\mathbf{Q} - \mathbf{C}\| \leq r\} \subset \mathbb{R}^{m \times n}$, then

$$\Pi_C(\mathbf{Q}) = \begin{cases} \mathbf{Q}, & \text{if } \|\mathbf{Q} - \mathbf{C}\| \leq r, \\ \mathbf{C} + r \|\mathbf{Q} - \mathbf{C}\|^{-1} (\mathbf{Q} - \mathbf{C}), & \text{if } \|\mathbf{Q} - \mathbf{C}\| > r. \end{cases}$$

The updates are summarized in Algorithm 1. The final reconstruction is given by $\mathbf{Z}_g = \mathbf{X}^*\mathbf{E}$, where \mathbf{X}^* is the output of Algorithm 1.

The choice of regularizer ϕ is discussed in the next section. As far as theoretical convergence is concerned, convexity of ϕ is sufficient to ensure convergence under some mild assumptions.

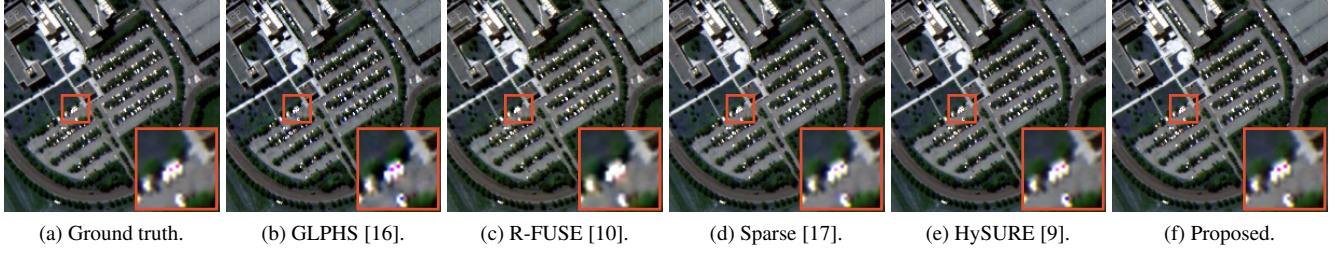


Fig. 1. HS-MS fusion results for the Pavia dataset using various methods. Note that the proposed method is able to preserve the red dots.

Theorem 1. *Let ϕ be closed, proper and convex. Suppose that problem (3) is feasible and has a solution. Then the sequence $(\mathbf{X}_k)_{k \geq 0}$ generated by Algorithm 1 converges to a minimizer of (3) as $k \rightarrow \infty$ for any $\rho > 0$.*

However, the choice of ρ is crucial as far as empirical convergence is concerned [13].

We note that Algorithm 1 is similar to C-SALSA [19], which is an instance of ADMM for a special class of constrained problems. The difference is that the variables in our algorithm are matrices (as opposed to vectors), and our problem contains two constraints as opposed to one in C-SALSA. The dominant cost of Algorithm 1 are the \mathbf{X} - and \mathbf{T} -updates. The \mathbf{T} -update depends on ϕ and is discussed in the next section. The \mathbf{X} -update requires us to solve a Sylvester equation and this can be done using the fast solver proposed in [11]. In this regard, we note that the spatial blur \mathbf{B} can be applied using FFT, and multiplications by \mathbf{E} and \mathbf{R} are not too expensive due to their small sizes (typically, L_s has a value between 10 and 30).

4. CHOICE OF REGULARIZER

The quality of fusion depends crucially on the regularizer $\phi(\mathbf{X})$. Prior works have explored the use of regularizers such as Tikhonov regularizer [11], sparsity-promoting functions based on learnt dictionaries [17], vector total variation (VTV) [9], and plug-and-play (PnP) regularization [20, 21]. To demonstrate the versatility of the proposed framework with different regularizers, we show results using two convex regularizers for $\phi(\mathbf{X})$ in (3)—VTV and PnP, both of which are considered state-of-the-art [9, 20].

We use VTV regularization for the MS-PAN fusion experiments in Section 5. In this case, the \mathbf{T} -update can be computed numerically using the technique proposed in [22]. For the HS-MS fusion experiments, we use PnP regularization where a powerful denoiser is directly substituted for the proximal map in ADMM. We give the main idea in PnP and refer the interested reader to [20, 21, 23] for further details. It is known that certain linear and symmetric image denoisers (smoothing operators) can be expressed as the proximal map of a closed, proper and convex function. Specifically, suppose $\mathcal{D} : \mathbb{R}^{n_m \times L_s} \rightarrow \mathbb{R}^{n_m \times L_s}$ is a high-dimensional image denoiser of

the form $\mathcal{D}(\mathbf{Q}) = \mathbf{W}\mathbf{Q}$, where $\mathbf{W} \in \mathbb{R}^{n_m \times n_m}$ is a symmetric positive semidefinite matrix with eigenvalues in $[0, 1]$. Then it was shown in [20, 21] that there exists a closed, proper and convex function ψ such that

$$\forall \mathbf{Q} \in \mathbb{R}^{n_m \times L_s} : \quad \mathcal{D}(\mathbf{Q}) = \text{prox}_{\psi}(\mathbf{Q}).$$

Moreover, linear denoisers with this property have been constructed in [20, 21]. The function ψ can be derived in closed-form in terms of \mathbf{W} (see [20]); however, we skip this since it is not directly relevant.

For HS-MS fusion using PnP regularization, we replace the \mathbf{T} -update (6b) by

$$\mathbf{T}_{k+1} = \mathbf{W}(\mathbf{X}_{k+1} - \Theta_k), \quad (9)$$

where \mathbf{W} is the high-dimensional linear denoiser in [21]. This is a variant of the popular nonlocal means (NLM) denoiser [24] which comes with a fast implementation. The other updates in (6) are kept intact. Since \mathbf{W} can be associated with a regularizer ψ (as discussed above), the denoising in (9) in effect amounts to the proximal update (6b) with $\phi = \psi$. Moreover, since ψ is closed, proper and convex, convergence of the ADMM iterations is guaranteed by Theorem 1.

5. EXPERIMENTS

We present experimental results for two fusion applications—HS-MS fusion and MS-PAN fusion (also called pansharpening).

5.1. HS-MS Fusion

In HS-MS fusion [25], the HS image serves the role of \mathbf{Y}_h and the MS image serves the role of \mathbf{Y}_m (both represent a common geographical region). We show results for the Pavia dataset [9] acquired over the city of Pavia, Italy. We remark that the ground-truth is not available for this dataset; we implement the standard procedure in [9] to generate the ground-truth and subsequently synthesize the HS and MS images. The HS image has a spatial resolution of 50×50 and contains 93 bands, whereas the MS image has a spatial resolution of 200×200 and contains 4 bands. Thus the subsampling operator \mathbf{S} performs $4 \times$ -downsampling along each spatial dimension. We add Gaussian noise \mathbf{N}_h and \mathbf{N}_m corresponding to a signal-to-noise ratio (SNR) of 30 dB to the HS and MS images [11].

We compare the performance of our method with the following methods which are considered state-of-the-art: CNMF [26], GLPHS [16], Sparse [17], R-FUSE [10], and HySURE [9]. A visual comparison of the results is shown in Fig. 1, where we display just three bands. Notice that the highlighted region in the image is more clear and smooth for the proposed method. For a quantitative comparison, we use the following evaluation metrics [9]: Erreur Relative Globale Adimensionnelle de Synthèse (ERGAS), Spectral Angle Mapper (SAM), Universal Image Quality Index (UIQI), Root Mean Squared

Metrics	RMSE	ERGAS	SAM	UIQI	PSNR (dB)	Time (s)
CNMF	0.0234	3.6843	3.3789	0.9684	32.6303	4.0231
GLPHS	0.0266	4.0379	4.2094	0.9606	31.5102	2.3261
R-FUSE	0.0215	3.4095	4.2165	0.9717	33.3453	0.3317
Sparse	0.0077	1.1527	1.9520	0.9954	42.2590	161.2587
HySURE	0.0081	1.2046	1.9588	0.9952	41.7837	5.4281
Proposed	0.0077	1.1432	1.9480	0.9954	42.2810	8.8712

Table 1. HS-MS fusion quality metrics for the Pavia dataset.

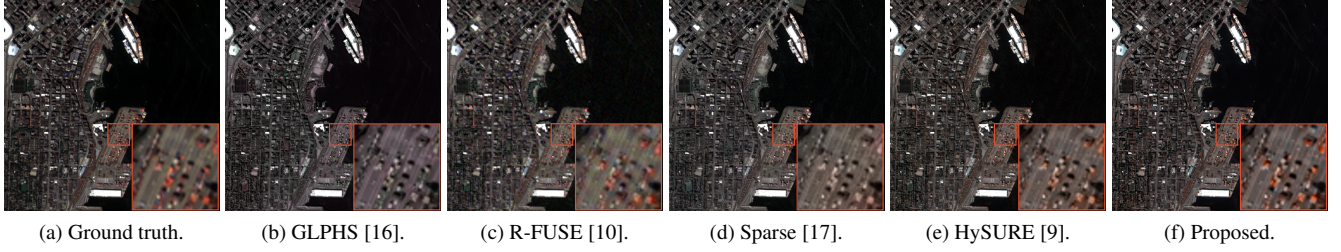


Fig. 2. MS-PAN fusion results for the Vancouver dataset using various methods. The proposed method is able to preserve the spectral quality.

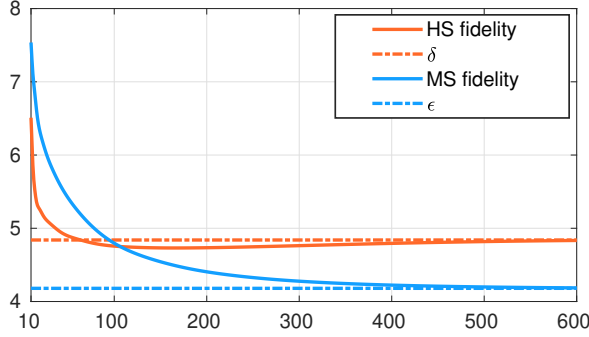


Fig. 3. Evolution of data-fidelity with iterations for HS-MS fusion. HS fidelity: $\|\mathbf{SBX}_k\mathbf{E} - \mathbf{Y}_h\|$, MS fidelity: $\|\mathbf{X}_k\mathbf{ER} - \mathbf{Y}_m\|$.

Metrics	RMSE	ERGAS	SAM	UIQI	PSNR (dB)	Time (s)
CNMF	0.0219	6.5809	5.1804	0.6980	33.1879	0.2498
GLPHS	0.0374	10.9172	15.0623	0.5695	28.5300	0.0209
R-FUSE	0.0230	6.8811	9.7413	0.6564	33.6094	0.3458
Sparse	0.0209	6.2454	7.4298	0.6589	32.7654	133.4253
HySURE	0.0201	5.8996	6.6710	0.7671	33.9015	31.3416
Proposed	0.0201	5.9497	6.4368	0.7688	33.9093	35.3011

Table 2. MS-PAN fusion metrics for the Vancouver dataset.

Error (RMSE) and Peak Signal to Noise Ratio (PSNR) [3]. Smaller values of ERGAS, SAM and RMSE signifies better fusion quality, whereas the converse is true for UIQI and PSNR. From the values reported in Table 1, we see that the proposed method is competitive with state-of-the-art methods. In terms of runtime (see Table 1), our method is comparable to HySURE and much faster than Sparse; we are slow compared to the other methods but our fusion quality is significantly better.

Finally, in Fig. 3, we verify that our algorithm indeed converges to the desired values of δ and ϵ in (3). The plot shows the values of $\|\mathbf{SBX}_k\mathbf{E} - \mathbf{Y}_h\|$ and $\|\mathbf{X}_k\mathbf{ER} - \mathbf{Y}_m\|$ as a function of the iteration index k . Note that with the progress in iterations, the above quantities converge to a value that is within the prescribed δ and ϵ bounds.

5.2. MS-PAN Fusion

In MS-PAN fusion [27], the MS image serves the role of \mathbf{Y}_h and the PAN image the role of \mathbf{Y}_m . We show results for a dataset representing a geographical region over Vancouver, Canada [28]. As before, the MS and PAN images are generated synthetically, with additive Gaussian noise corresponding to a SNR of 25 dB. The MS images contains 4 bands and PAN image contains a single band. The spatial resolutions of the MS and PAN images are 125×125 and 500×500 .

A visual comparison is shown in Fig. 2 using three bands of the

fused image. We see that our method is better able to preserve the spectral quality as compared to the other methods. The corresponding quality metrics are reported in Table 2, from which we see that the proposed method is competitive with other methods in terms of fusion quality. In terms of runtime, our method is competitive with best-performing methods such as Sparse and HySURE. We note that HySURE [9] uses VTV regularization but uses the unconstrained formulation in (2). The constrained formulation (proposed method) yields comparable results both visually and quantitatively, but is better able to retain the spectral characteristics as compared to HySURE.

6. CONCLUSION

We explored a novel constrained optimization formulation for multi-band image fusion, in which we minimize a convex regularizer subject to data-dependent spatial and spectral constraints. As compared to existing approaches, our formulation has the advantage of being interpretable and easily tunable and can achieve a prescribed level of data fidelity. The proposed ADMM solver was shown to be efficient in practice. We note that the proposed approach can be used with any convex regularizers provided its proximal map can be efficiently computed. Through experiments on standard datasets, we showed that our method is able to enhance the spatial resolution while retaining good spectral quality.

7. APPENDIX: PROOF OF THEOREM 1

Note that (3) can be expressed as the minimization of $F(\mathcal{G}(\mathbf{X}))$, where

$$F(\mathbf{X}^1, \mathbf{X}^2, \mathbf{X}^3) = \phi(\mathbf{X}^1) + \iota_{\mathcal{H}}(\mathbf{X}^2) + \iota_{\mathcal{M}}(\mathbf{X}^3), \quad (10)$$

and the map $\mathbf{X} \mapsto \mathcal{G}(\mathbf{X}) := (\mathbf{X}, \mathbf{SBX}\mathbf{E}, \mathbf{XER})$ is linear. Since \mathcal{H} and \mathcal{M} are closed sets, F is a closed function. Moreover, F is proper since the problem is assumed to be feasible. Thus, F is closed, proper and convex. By representing the matrix variable \mathbf{X} as a vector \mathbf{x} , the problem can be written in the form

$$\min_{\mathbf{x}} f(\mathbf{G}\mathbf{x}),$$

where \mathbf{G} is a matrix representation of \mathcal{G} and f is the function corresponding to (10), but for vectorized arguments. Algorithm 1 is then an instance of ADMM applied to

$$\min_{\mathbf{x}, \mathbf{w}} f(\mathbf{w}) \quad \text{subject to} \quad \mathbf{G}\mathbf{x} = \mathbf{w}. \quad (11)$$

It is known that the ADMM converges to a minimizer of (11) if f is closed, proper and convex, and \mathbf{G} has full column rank [18]. Thus, we just need to show that \mathbf{G} has full column rank. But this is immediate since \mathcal{G} is injective, i.e., $\mathcal{G}(\mathbf{X}) = (\mathbf{0}, \mathbf{0}, \mathbf{0})$ only if $\mathbf{X} = \mathbf{0}$.

8. REFERENCES

- [1] C. Chang, *Hyperspectral Data Exploitation: Theory and Applications*, John Wiley & Sons, 2007.
- [2] A. Panda, S. Jones, H. Stark, R. Raghavan, K. Sandrasegaran, N. Bansal, and U. Dydak, "Phosphorus liver mrsi at 3 t using a novel dual-tuned eight-channel 31p/1h coil," *Magnetic Resonance in Medicine*, vol. 68, no. 5, pp. 1346–1356, 2012.
- [3] N. Yokoya, C. Grohnfeldt, and J. Chanussot, "Hyperspectral and multispectral data fusion: A comparative review of the recent literature," *IEEE Geoscience and Remote Sensing Magazine*, vol. 5, no. 2, pp. 29–56, 2017.
- [4] B. Uzkent, M. Hoffman, and A. Vodacek, "Real-time vehicle tracking in aerial video using hyperspectral features," *IEEE Conference on Computer Vision and Pattern Recognition Workshop*, pp. 36–44, 2016.
- [5] A. Plaza, Q. Du, J. Bioucas-Dias, X. Jia, and F. Kruse, "Foreword to the special issue on spectral unmixing of remotely sensed data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 11, pp. 4103–4110, 2011.
- [6] J. Bioucas-Dias, A. Plaza, G. Camps-Valls, P. Scheunders, N. Nasrabadi, and J. Chanussot, "A convex formulation for hyperspectral image superresolution via subspace-based regularization," *IEEE Geoscience and Remote Sensing Magazine*, vol. 1, no. 2, pp. 6–36, 2013.
- [7] Q. Wei, N. Dobigeon, and J. Y. Tourneret, "Bayesian fusion of multi-band images," *IEEE Journal of Selected Topics in Signal Processing*, vol. 9, no. 6, pp. 1117–1127, 2015.
- [8] N. Akhtar, F. Shafait, and A. Mian, "Bayesian sparse representation for hyperspectral image super resolution," *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3631–3640, 2015.
- [9] M. Simoes, J. Bioucas-Dias, L. B. Almeida, and J. Chanussot, "A convex formulation for hyperspectral image superresolution via subspace-based regularization," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 6, pp. 3373–3388, 2014.
- [10] Q. Wei, N. Dobigeon, J. Y. Tourneret, J. M. Bioucas-Dias, and S. Godsill, "R-fuse: Robust fast fusion of multiband images based on solving a Sylvester equation," *IEEE Signal Processing Letters*, vol. 23, no. 11, pp. 1632–1636, 2016.
- [11] Q. Wei, N. Dobigeon, and J. Y. Tourneret, "Fast fusion of multi-band images based on solving a Sylvester equation," *IEEE Transactions on Image Processing*, vol. 24, no. 11, pp. 4109–4121, 2015.
- [12] I. Amro, J. Mateos, M. Vega, R. Molina, and A. K. Katsaggelos, "A survey of classical methods and new trends in pansharpening of multispectral images," *EURASIP Journal on Advances in Signal Processing*, vol. 2011, no. 1, pp. 1–22, 2011.
- [13] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2011.
- [14] R. C. Hardie, M. T. Eismann, and G. L. Wilson, "MAP estimation for hyperspectral image resolution enhancement using an auxiliary sensor," *IEEE Transactions on Image Processing*, vol. 13, no. 9, pp. 1174–1184, 2004.
- [15] J. M. Nascimento and J. M. Dias, "Vertex component analysis: A fast algorithm to unmix hyperspectral data," *IEEE Transactions on Image Processing*, vol. 43, no. 4, pp. 898–910, 2005.
- [16] M. Selva, B. Aiazzi, F. Butera, L. Chiarantini, and S. Baronti, "Hyper-sharpening: A first approach on SIM-GA data," *IEEE Journal of Selected Topics in Applied Earth Observation and Remote Sensing*, vol. 8, no. 6, pp. 3008–3024, 2015.
- [17] Q. Wei, J. M. Bioucas-Dias, N. Dobigeon, and J. Y. Tourneret, "Hyperspectral and multispectral image fusion based on a sparse representation," *IEEE Transactions on Image Processing*, vol. 53, no. 7, pp. 3658–3668, 2015.
- [18] M. V. Afonso, J. Bioucas-Dias, and M. A. T. Figueiredo, "Fast image recovery using variable splitting and constrained optimization," *IEEE Transactions on Image Processing*, vol. 19, no. 9, pp. 2345–2356, 2010.
- [19] M. V. Afonso, J. M. Bioucas-Dias, and M. A. T. Figueiredo, "An augmented Lagrangian approach to the constrained optimization formulation of imaging inverse problems," *IEEE Transactions on Image Processing*, vol. 20, no. 3, pp. 681–695, 2010.
- [20] A. M. Teodoro, J. M. Bioucas-Dias, and M. A. T. Figueiredo, "A convergent image fusion algorithm using scene-adapted Gaussian-mixture-based denoising," *IEEE Transactions on Image Processing*, vol. 28, no. 1, pp. 451–463, 2019.
- [21] P. Nair, V. S. Unni, and K. N. Chaudhury, "Hyperspectral image fusion using fast high-dimensional denoising," *Proc. IEEE International Conference on Image Processing*, pp. 3123–3127, 2019.
- [22] X. Bresson and T. Chan, "Fast dual minimization of the vectorial total variation norm and applications to color image processing," *Inverse Problems & Imaging*, vol. 2, no. 4, pp. 455, 2008.
- [23] S. Sreehari, S. V. Venkatakrishnan, B. Wohlberg, G. T. Buzzard, L. F. Drummy, J. P. Simmons, and C. A. Bouman, "Plug-and-play priors for bright field electron tomography and sparse interpolation," *IEEE Transactions on Computational Imaging*, vol. 2, no. 4, pp. 408–423, 2016.
- [24] A. Buades, B. Coll, and J. M. Morel, "A non-local algorithm for image denoising," *Proc. IEEE Computer Vision and Pattern Recognition*, vol. 2, pp. 60–65, 2005.
- [25] R. Dian, S. Li, L. Fang, T. Lu, and J. M. Bioucas-Dias, "Non-local sparse tensor factorization for semiblind hyperspectral and multispectral image fusion," *IEEE Transactions on Cybernetics*, vol. 50, no. 10, pp. 4469–4480, 2020.
- [26] N. Yokoya, T. Yairi, and A. Iwasaki, "Coupled nonnegative matrix factorization unmixing for hyperspectral and multispectral data fusion," *IEEE Transactions on Image Processing*, vol. 50, no. 2, pp. 528–537, 2012.
- [27] X. Fu, Z. Lin, Y. Huang, and X. Ding, "A variational pansharpening with local gradient constraints," *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pp. 10265–10274, 2019.
- [28] Q. Wei, J. Bioucas-Dias, N. Dobigeon, J. Tourneret, and S. Godsill, "Blind model-based fusion of multi-band and panchromatic images," *Proc. IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems*, vol. 2, pp. 21–25, 2016.