HYPERSPECTRAL IMAGE SUPER-RESOLUTION VIA LOCAL LOW-RANK AND SPARSE REPRESENTATIONS

Renwei Dian[1,2]*, Shutao Li*[1]*, Leyuan Fang*[1]*, and José Bioucas-Dias*[2]

College of Electrical and Information Engineering, Hunan University, Changsha, China[1]
Instituto de Telecomunicacões, Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal[2]

ABSTRACT

Remotely sensed hyperspectral images (HSIs) usually have high spectral resolution but low spatial resolution. A way to increase the spatial resolution of HSIs is to solve a fusion inverse problem, which *fuses* a low spatial resolution HSI (LR-HSI) with a high spatial resolution multispectral image (HR-MSI) of the same scene. In this paper, we propose a novel HSI super-resolution approach (called LRSR), which formulates the fusion problem as the estimation of a spectral dictionary from the LR-HSI and the respective regression coefficients from both images. The regression coefficients are estimated by formulating a variational regularization problem which promotes local (in the spatial sense) low-rank and sparse regression coefficients. The local regions, where the spectral vectors are low-rank, are estimated by segmenting the HR-MSI. The formulated convex optimization is solved with SALSA. Experiments provide evidence that LRSR is competitive with respect to the state-of-the-art methods.

***Index Terms*—** Hyperspectral image super-resolution, low rank, superpixels

1. INTRODUCTION

HSIs have been used in many Earth remote sensing applications [1], due to its high spectral resolution. However, owing, namely, to the finite sun irradiance, there is the trade-off between the spectral resolution and spatial resolution for existing cameras. Hence, HSIs with high spectral resolution are often acquired with low spatial resolution, which limits the applications of HSIs. In contrast, MSIs usually have higher spatial resolution (HR-MSI) and much lower spectral resolution. The availability of pairs HR-MSI and LR-HSI have fostered several research efforts to obtain fused high spatial resolution HSI (HR-HSI) images.

This work was supported by the National Natural Science Fund of China for International Cooperation and Exchanges under Grant 61520106001, by the National Natural Science Fund of China for Distinguished Young Scholars under Grant 61325007, by the Fund of Hunan Province for Science and Technology Plan Project under Grant 2017RS3024, and by the Fundaçáo para a Ciência e Tecnologia, Portuguese Ministry of Science and Higher Education, projects UID/EEA/50008/2013 and ERANETMED/0001/2014.

Recently, HSI super-resolution based on matrix factorization has been actively investigated. These methods, introduced firstly in [2, 3], decompose the HR-HSI as a spectral dictionary, estimated form the LR-HSI, and the coefficients from the HR-MSI. The methods proposed in [4, 5] use spectral unmixing concepts to regularize the fusion problem; instead of estimating spectral basis in advance and keeping it fixed, these methods alternately update the dictionary atoms and the coefficients with non-negative and sum-to-one constraints. Furthermore, Simoẽs *et al.* [6] and Wei *et al.* [7] assume the HSIs live on a low-dimensional subspace, and use, respectively, total variation regularization to promote spatial piecewise smoothness solutions and spatial dictionaries learned from the HR-MSI to promote similarity, respectively.

In this paper, we propose a new formulation to the fusion problem. Pixels in properly selected neighborhoods are likely to represent the same materials, and therefore HSIs are often locally low-rank. To use the locally low rank prior, Veganzones *et al.* [8] solve the fusion problem for each local patch independently. Different from reference [8], we first oversegment the HR-MSI using the superpixel segmentation method [9], and then design a variational regularization problem which promotes superpixels with low rank. Besides, the spectral vectors in the HR-HSI are represented as regressions over a spectral dictionary learned from the LR-HSI with the constraint to produce sparse representations. The thus obtained variational regularization problem is convex and solved with SALSA [10], which is an instance of ADMM.

2. PROBLEM FORMULATION

In this paper, HR-HSIs, LR-HSIs, and HR-MSIs are represented by matrices. The target HR-HSI is denoted by $\mathbf{X} \in \mathbb{R}^{S\times N}$, where N and S represent the number of pixels and spectral bands, respectively. $\mathbf{Y} \in \mathbb{R}^{S\times n}$ represents the acquired LR-HSI, where $n < N$ and S denote the number of pixels and bands in the LR-HSI, respectively. $\mathbf{Z} \in \mathbb{R}^{s\times N}$ represents the HR-MSI of the same scenario, where $s < S$ represent the number of bands of the HR-MSI.

The spectral vectors in the HR-HSI may be represented as

a linear combination of spectral signatures. That is, we may factor $\mathbf{X}$ as

$$\mathbf{X} = \mathbf{EA} = \mathbf{E}[\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_N], \tag{1}$$

where $\mathbf{E} \in \mathbb{R}^{S\times L}$ is a spectral dictionary, and $\mathbf{A} \in \mathbb{R}^{L\times N}$ is a matrix with coefficients.

We adopt the following observation models for the LR-HSI $\mathbf{Y}$ and the HR-MSI $\mathbf{Z}$:

$$\mathbf{Y} = \mathbf{XBS} + \mathbf{N}_h \tag{2}$$

$$\mathbf{Z} = \mathbf{RX} + \mathbf{N}_m, \tag{3}$$

where $\mathbf{B} \in \mathbb{R}^{N\times N}$ is a blur matrix, which models a convolution between the HR-HSI bands (represented by the rows of $\mathbf{X}$) and point spread function of the imaging sensor, $\mathbf{S} \in \mathbb{R}^{N\times n}$ is a spatial downsampling matrix, $\mathbf{R} \in \mathbb{R}^{s\times S}$ is the spectral response matrix of the mulspectral sensor, and $\mathbf{N}_h$ and $\mathbf{N}_m$ represent Gaussian additive noise affecting the LR-HSI and the HR-MSI, respectively.

Combining the linear mixture model (1) with imaging models (2) and (3) leads to

$$\mathbf{Y} = \mathbf{EABS} + \mathbf{N}_h, \;\; \mathbf{Z} = \mathbf{REA} + \mathbf{N}_m. \tag{4}$$

In this formulation, the fusion amounts to estimate the dictionary $\mathbf{E}$ and the coefficients $\mathbf{A}$ from $\mathbf{Y}$ and $\mathbf{Z}$.

3. THE PROPOSED LRSR METHOD

The proposed method mainly contains two steps: learning spectral dictionary $\mathbf{E}$ and estimating the coefficients $\mathbf{A}$. Since a spectral dictionary with sparsity representation ability may be learned from LR-HSI [3], we learn the spectral dictionary $\mathbf{E}$ from the LR-HSI via dictionary learning method proposed in [11]. With spectral dictionary known, the estimation of coefficients $\mathbf{A}$ is formulated as

$$\min_{\mathbf{A}} ||\mathbf{Y} - \mathbf{EABS}||_F^2 + \lambda||\mathbf{Z} - \mathbf{REA}||_F^2, \tag{5}$$

where $\lambda > 0$. Formulation (5) assumes implicitly that the noise in i.i.d in both images. Parameter λ accounts for different noise variances in both images. Generic noise covariance matrices are easily introduced in (5).

Due to the downsampling matrices $\mathbf{R}$ and $\mathbf{S}$, the problem (5) is usually ill-posed and, hence, its solution is not unique. Therefore, we need some prior information to regularize the problem. Since the columns of $\mathbf{X}$ may be sparsely represented over the dictionary $\mathbf{E}$, the regularizer $\|\mathbf{A}\|_{1,1} = \sum_{n=1}^{N} \|\mathbf{a}_n\|_1$ is added to (5) to promote spase coefficient vectors $\mathbf{a}_n$.

Spectral vectors in properly selected neighborhoods often represent the same materials, and therefore HSIs are locally low-rank. Here, the nuclear norm is used to locally promote low-rank. Since HR-HSIs preserves most of the spatial information of HR-HSI counterparts, we firstly oversegment the HR-MSI into K superpixels $\{\mathbf{Z}^i =\}_{i=1}^K$ using the entropy rate based superpixel segmentation method [9]. In this way, the HR-HSI is also segmented into K superpixels $\{\mathbf{X}^i = \mathbf{EA}^i\}_{i=1}^K$ according to the corresponding spatial locations. Then, we impose the low-rank prior to the coefficients of spectral pixels in the same superpixel. The obtained variational regularization problem is therefore

$$\min_{\mathbf{A}} ||\mathbf{Y} - \mathbf{EABS}||_F^2 + \lambda||\mathbf{Z} - \mathbf{REA}||_F^2 + \eta_1||\mathbf{A}||_1 + \eta_2 \sum_{i=1}^{K} ||\mathbf{A}^i||_*, \tag{6}$$

where $||\mathbf{A}^i||_*$ denotes the nuclear norm of the coefficients in i^{th} superpixel, and $\eta_1.\eta_2 \geq 0$ is regularization parameters.

Problem (6) is convex and thus it can be solved efficiently via the the split augmented Lagrangian shrinkage algorithm (SALSA) [10], which is an instance of ADMM. By introducing the constrains $\mathbf{V}_1 = \mathbf{AB}$, $\mathbf{V}_2 = \mathbf{A}$, $\mathbf{V}_3 = \mathbf{A}$, and $\mathbf{V}_4 = \mathbf{A}$, the optimization (6) is converted into an equivalent one, whose augmented Lagrangian function is

$$\begin{aligned} L(\mathbf{A}, \mathbf{V}_1, \mathbf{V}_2, \mathbf{V}_3, \mathbf{V}_4, \mathbf{G}_1, \mathbf{G}_2, \mathbf{G}_3, \mathbf{G}_4) = \\ ||\mathbf{Y} - \mathbf{EV}_1\mathbf{S}||_F^2 + \mu||\mathbf{V}_1 - \mathbf{AB} + \mathbf{G}_1||_F^2 + \\ \lambda||\mathbf{Z} - \mathbf{REV}_2||_F^2 + \mu||\mathbf{V}_2 - \mathbf{A} + \mathbf{G}_2||_F^2 + \\ \eta_1||\mathbf{V_3}||_1 + \mu||\mathbf{V}_3 - \mathbf{A} + \mathbf{G}_3||_F^2 + \\ \eta_2 \sum_{i=1}^{K} ||\mathbf{V}_4^i||_* + \mu||\mathbf{V}_4 - \mathbf{A} + \mathbf{G}_4||_F^2 \end{aligned} \tag{7}$$

The SALSA algorithms results in the following updates:
(1) Update $\mathbf{A}$:

$$\begin{aligned} \mathbf{A} &\in \operatorname*{argmin}_{\mathbf{A}} L(\mathbf{A}, \mathbf{V}_1, \mathbf{V}_2, \mathbf{V}_3, \mathbf{V}_4, \mathbf{G}_1, \mathbf{G}_2, \mathbf{G}_3, \mathbf{G}_4) \\ &= [(\mathbf{V}_1 + \mathbf{G}_1)\mathbf{B}^T + \mathbf{V}_2 + \mathbf{G}_2 + \mathbf{V}_3 + \mathbf{G}_3 + \mathbf{V}_4 + \mathbf{G}_4]\times \\ &\quad (\mathbf{BB}^T + 3\mathbf{I})^{-1} \end{aligned} \tag{8}$$

(2) Update $\mathbf{V}_1$:

$$\mathbf{V}_1 \in \operatorname*{argmin}_{\mathbf{V}_1} L(\mathbf{A}, \mathbf{V}_1, \mathbf{V}_2, \mathbf{V}_3, \mathbf{V}_4, \mathbf{G}_1, \mathbf{G}_2, \mathbf{G}_3, \mathbf{G}_4) \tag{9}$$

$\mathbf{V}_1$ can be solved by separating $\mathbf{V}_1$ as $\mathbf{V}_1\mathbf{S}$ and $\mathbf{V}_1\bar{\mathbf{S}}$. $\bar{\mathbf{S}}$ is the downsamping matrix which selects the pixels not selected by $\mathbf{S}$. We have then

$$\begin{aligned} \mathbf{V}_1\mathbf{S} &= (\mathbf{E}^T\mathbf{E} + \mu\mathbf{I})^{-1}[(\mathbf{E}^T\mathbf{Y} + \mu(\mathbf{AB} - \mathbf{G}_1)\mathbf{S}] \\ \mathbf{V}_1\bar{\mathbf{S}} &= (\mathbf{AB} - \mathbf{G}_1)\bar{\mathbf{S}} \end{aligned} \tag{10}$$

(3) Update $\mathbf{V}_2$:

$$\begin{aligned} \mathbf{V}_2 &\in \operatorname*{argmin}_{\mathbf{V}_2} L(\mathbf{A}, \mathbf{V}_1, \mathbf{V}_2, \mathbf{V}_3, \mathbf{V}_4, \mathbf{G}_1, \mathbf{G}_2, \mathbf{G}_3, \mathbf{G}_4) \\ &= [\lambda(\mathbf{RE})^T\mathbf{RE} + \mu\mathbf{I}]^{-1}[\lambda(\mathbf{RE})^T\mathbf{Z} + \mu(\mathbf{A} - \mathbf{G}_2)] \end{aligned} \tag{11}$$

(4) Update $\mathbf{V}_3$:

$$\mathbf{V}_3 \in \underset{\mathbf{V}_3}{\operatorname{argmin}} L(\mathbf{A}, \mathbf{V}_1, \mathbf{V}_2, \mathbf{V}_3, \mathbf{V}_4, \mathbf{G}_1, \mathbf{G}_2, \mathbf{G}_3, \mathbf{G}_4) = \text{soft}(\mathbf{A} - \mathbf{G}_3, \frac{\eta_1}{2\mu}). \tag{12}$$

where $\text{soft}(a, b) = \text{sign}(a) * \max(|a| - b, 0)$

(5) Update $\mathbf{V}_4$:

$$\mathbf{V}_4 \in \underset{\mathbf{V}_4}{\operatorname{argmin}} L(\mathbf{A}, \mathbf{V}_1, \mathbf{V}_2, \mathbf{V}_3, \mathbf{V}_4, \mathbf{G}_1, \mathbf{G}_2, \mathbf{G}_3, \mathbf{G}_4). \tag{13}$$

Problem (13) can be solved separately for each superpixel. For coefficients of i^{th} superpixel $\mathbf{V}_4^i$, we have

$$\mathbf{V}_4^i = \mathbf{U} \left(\mathbf{\Sigma} - \frac{\eta_2}{2\mu} \right)_+ \mathbf{V}^T \tag{14}$$

where $\mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$ is the SVD of $\mathbf{A}^i - \mathbf{G}_4^i$ and $(\cdot)^+$ is the positive part.

(6) Update Lagrangian multipliers $\mathbf{G}_1$, $\mathbf{G}_2$, $\mathbf{G}_3$, and $\mathbf{G}_4$:

$$\begin{aligned} \mathbf{G}_1 &= \mathbf{G}_1 + \mathbf{V}_1 - \mathbf{A}\mathbf{B} \\ \mathbf{G}_2 &= \mathbf{G}_2 + \mathbf{V}_2 - \mathbf{A} \\ \mathbf{G}_3 &= \mathbf{G}_3 + \mathbf{V}_3 - \mathbf{A} \\ \mathbf{G}_4 &= \mathbf{G}_4 + \mathbf{V}_4 - \mathbf{A} \end{aligned} \tag{15}$$

The above six steps are iterated until convergence.

4. EXPERIMENTS

In this Section, we present results resulting from fusing LR-HSI with HR-MSI (HSI-MSI fusion) and LR-HSI with a panchromatic image (HSI-PAN fusion).

4.1. Data Set

We use the Pavia University [12], with size $N = 128 \times 128$ and $S = 115$, to test the proposed method. Since some bands have very low signal-to-noise-ratio (SNR) due to water vapor absorption, they were removed, reducing the HSI to $S = 93$ bands. To generate an LR-HSI of size $32 \times 32 \times 93$ (downsampling factor of 4), we firstly convolve the HR-HSI bands with a 7×7 symmetric Gaussian filter (standard deviation 2) and then downsample the lowpass filtered image by a factor of 4 in both spatial dimensions. The IKONOS-like reflectance spectral response filter is used to generate the HR-MSI of size $128 \times 128 \times 4$ and the HR panchromatic (PAN) of size $128 \times 128 \times 1$. Noise with i.i.d Gaussian distribution is added to LR-HSI, HR-MSI, and PAN defining SNRs of, respectively, 30dB, 40dB, and 40dB.

Table 1. Quantitative results for HSI-MSI fusion on the Pavia University [12].

Method	**Pavia University** [12]			
	PSNR (dB)	SAM	ERGAS	UIQI
Best Values	$+\infty$	0	0	1
CNNMF [4]	43.30	1.87	1.042	0.9937
CSU [5]	41.87	1.87	1.220	0.9918
HySure [6]	42.89	1.94	1.137	0.9931
SR [7]	43.84	**1.77**	1.003	0.9942
LRSR	**43.89**	1.81	**0.981**	**0.9943**

Table 2. Quantitative results for HSI-PAN fusion on the Pavia University [12].

Method	**Pavia University** [12]			
	PSNR (dB)	SAM	ERGAS	UIQI
Best Values	$+\infty$	0	0	1
CNNMF [4]	31.75	5.36	4.045	0.9259
CSU [5]	29.64	7.88	5.841	0.8900
HySure [6]	32.71	5.19	3.612	0.9434
SR [7]	31.89	5.85	4.025	0.9272
LRSR	**33.69**	**4.56**	**3.251**	**0.9516**

4.2. Quantitative Metrics

In our study, we use four quantitative metrics to evaluate the quality of the fusion results: the peak signal to noise ratio (PSNR), the spectral angle mapper (SAM) [13], the relative dimensionless global error in synthesis (ERGAS) [14], and the universal image quality index (UIQI) [15].

4.3. Compared Methods

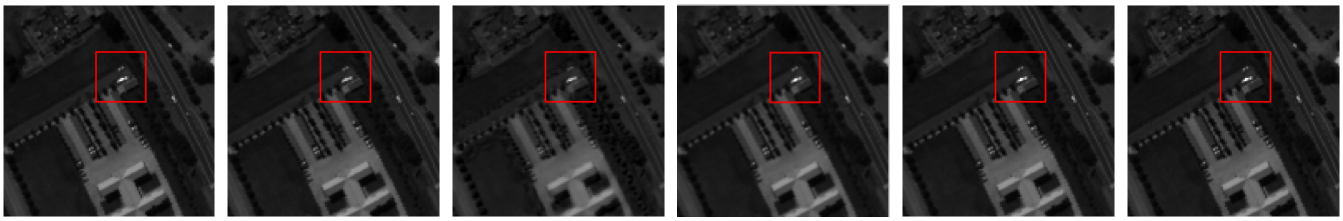
We compare the proposed algorithm with three state-of-the-art HSI super-resolution methods: the coupled non-negative matrix factorization (CNNMF) [4], the coupled spectral unmixing (CSU) [5], HySure [6], and sparse representation (SR) based method [7].

4.4. Parameters Selection

There are five parameters of the proposed method are set as follows: relative ration between the two data fidelity terms, $\lambda = 1$; number of atoms in the dictionary, $L = 24$; number of superpixels. $K = 200$; sparsity regularization parameter, $\eta_1 = 1 \times 10^{-4}$; and low rank regularization parameter $\eta_2 = 1 \times 10^{-3}$.

4.5. Experimental Results

Tables 1 and 2 show the quantitative results of the test methods for HSI-MSI fusion and HSI-PAN fusion on Pavia U-



(a) CNNMF [4] (b) CSU [5] (c) HySure [6] (d) SR [7] (e) LRSR (f) Ground truth

Fig. 1. The reconstructed Pavia University at 30^{th} band of HSI-MSI fusion for the test methods: (a) CNNMF [4], (b) CSU [5] (c) HySure [6], (d) SR [7], (e) LRSR, and (f) Ground truth.

niversity, respectively. All of the CNNMF, CSU, Hysure, and SR are existing state-of-the-art methods. As can be seen from Tables 1 and 2, the LRSR outperforms the test methods on both of the HSI-MSI fusion and HSI-PAN fusion, and HySure, CNNMF, and SR perform better than the CSU. The HSI-PAN fusion is a much more challenging problem than the HSI-MSI fusion since the panchromatic image has only one band. Fig. 1 shows the reconstructed Pavia University at 30^{th} band by the test methods. Besides, the LRSR and SR perform better than the other methods in reconstructing the HR details for the marked region, as can be seen from Fig. 1.

5. CONCLUSIONS

In this paper, we present a locally spatial low-rank and spectrally sparse based HSI super-resolution method to fuse LR-HSIs and HR-MSIs. In the proposed method, we formulate the super-resolution problem as the estimation of a spectral dictionary and coefficients regularized by a prior promoting local low-rank and another prior promoting sparse spectral coefficients. We solve the variational regularization problem efficiently using the SALSA algorithm. Experiments using the remote sensing HSIs provide evidence of the effectiveness of the proposed method.

6. REFERENCES

[1] J. Bioucas-Dias, A. Plaza, G. Camps-Valls, P. Scheunders, N. Nasrabadi, and J. Chanussot, "Hyperspectral remote sensing data analysis and future challenges," *IEEE Geosci. Remote Sens. Mag.*, vol. 1, no. 2, pp. 6–36, Jun. 2013.

[2] R. Kawakami, J. Wright, Y. W. Tai, Y. Matsushita, M. Ben-Ezra, and K. Ikeuchi, "High-resolution hyperspectral imaging via matrix factorization," in *IEEE Conf. Comput. Vis. Pattern Recog.*, Jun. 2011, pp. 2329 –2336.

[3] A. Charles, B. Olshausen, and C. Rozell, "Learning sparse codes for hyperspectral imagery," *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 5, pp. 963–978, 2011.

[4] N. Yokoya, T. Yairi, and A. Iwasaki, "Coupled non-negative matrix factorization unmixing for hyperspectral and multispectral data fusion," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 2, pp. 528–537, Feb. 2012.

[5] C. Lanaras, E. Baltsavias, and K. Schindler, "Hyperspectral super-resolution by coupled spectral unmixing," in *IEEE Int. Conf. Comput. Vis.*, Dec. 2015, pp. 3586 –3594.

[6] M. Simões, J. Bioucas-Dias, L. Almeida, and J. Chanussot, "A convex formulation for hyperspectral image superresolution via subspace-based regularization," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 6, pp. 3373–3388, Jun. 2015.

[7] Q. Wei, J. Bioucas-Dias, N. Dobigeon, and J.-Y. Tourneret, "Hyperspectral and multispectral image fusion based on a sparse representation," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 7, pp. 3658–3668, 2015.

[8] M. Veganzones, M. Simões, G. Licciardi, N. Yokoya, J. Bioucas-Dias, and J. Chanussot, "Hyperspectral super-resolution of locally low rank images from complementary multisource data," *IEEE Trans. Image Process.*, vol. 25, no. 1, pp. 274–288, Jan. 2016.

[9] M. Liu, O. Tuzel, S. Ramalingam, and R. Chellappa, "Entropy rate superpixel segmentation," in *IEEE Conf. Comput. Vis. Pattern Recog.*, Jun. 2011, pp. 2097–2104.

[10] M. Afonso, J. Bioucas-Dias, and M. Figueiredo, "An augmented lagrangian approach to the constrained optimization formulation of imaging inverse problems," *IEEE Trans. Image Process.*, vol. 20, no. 3, pp. 681–695, Mar. 2011.

[11] W. Dong, F. Fu, G. Shi, X. Cao, J. Wu, G. Li, and X. Li, "Hyperspectral image super-resolution via non-negative structured sparse representation," *IEEE Trans. Image Process.*, vol. 25, no. 5, pp. 2337–2352, May 2016.

[12] F. Dell'Acqua, P. Gamba, A. Ferrari, J. Palmason, and J. Benediktsson, "Exploiting spectral and spatial information in hyperspectral urban data with high resolution," *IEEE Geosci. Remote Sens. Lett.*, vol. 1, no. 4, pp. 322–326, Oct. 2016.

[13] R. Yuhas, A. Goetz, and J. Boardman, "Discrimination among semi-arid landscape endmembers using the spectral angle mapper (SAM) algorithm," *JPL Airborne Geosci. Workshop*, vol. 1, pp. 147–149, 1992.

[14] L. Wald, "Quality of high resolution synthesised images: Is there a simple criterion," in *Int. Conf. Fusion Earth Data*, Jan. 2000, pp. 99–103.

[15] Z. Wang and A. C. Bovik, "A universal image quality index," *IEEE Signal Process. Lett.*, vol. 9, no. 3, pp. 81–84, Mar. 2002.