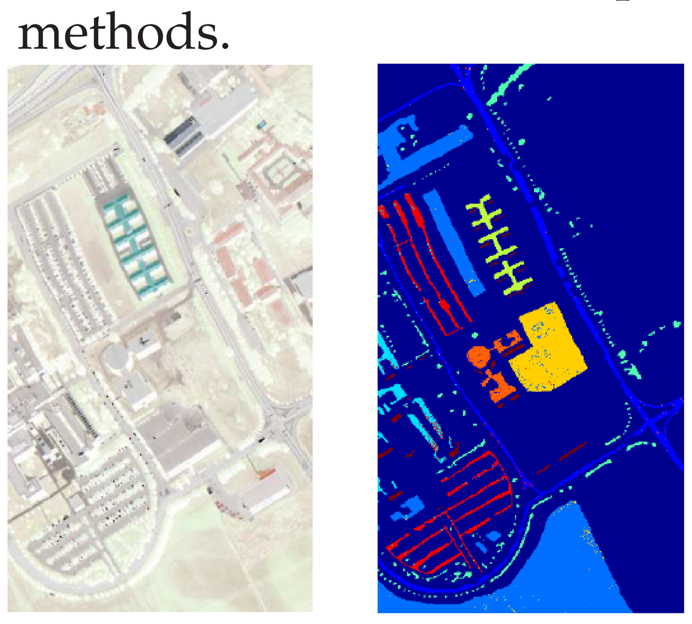
Predictive Lossless Compression of Regions of Interest in Hyperspectral Images via Maximum Correntropy Criterion Based Least Mean Square Learning

Hongda Shen, W. David Pan,
Department of Electrical and Computer Engineering
University of Alabama in Huntsville, USA

2016 Sept 25-28 Phoenix
The 23rd IEEE International Conference on Image Processing (ICIP)

Introduction

- There is a great need for massive hyperspectral data compression due to severely limited bandwidth and on-board storage space.
- We focused on predictive lossless compression of regions of interest (ROIs) with no-data regions.
- We proposed a MCC-LMS (Maximum Correntropy Criteria based Least Mean Square) filtering method to learn the prediction residuals with Laplacian or geometrically distributions.
- We aimed to improve several state-of-the-art ROI compression methods.



(a) Spectral band 50. (b) ROI Map.

MCC-LMS Predictor

Correntropy was developed as a local similarity measure between two random variables X and Y, defined by:

$$V_\sigma(X,Y) = E\left[\kappa_\sigma(X-Y)\right],$$

where κ_σ is a positive definite kernel with kernel width σ.
In this work, we use d_i and y_i to represent the actual pixel value and its estimate for i^{th} pixel, respectively. Furthermore, the estimate y_i can be computed as $y_i = \mathbf{W}_i^T\mathbf{X}_i$, a linear weighted average of input context vector $\mathbf{X}_i$. where $\mathbf{W}_n$ is the filter weight at n^{th} time instant. Iterative gradient descent method is chosen with a small learning rate μ. The correntropy-based cost function, at n^{th} time instant, can be written as:

$$J_n = \frac{1}{N\sqrt{2\pi}\sigma}\sum_{i=n-N+1}^{n}\exp\left[\frac{-(d_i - \mathbf{W}_n^T\mathbf{X}_n)^2}{2\sigma^2}\right],$$

After computing the gradient of J_n w.r.t $\mathbf{W}_n$, we obtain:

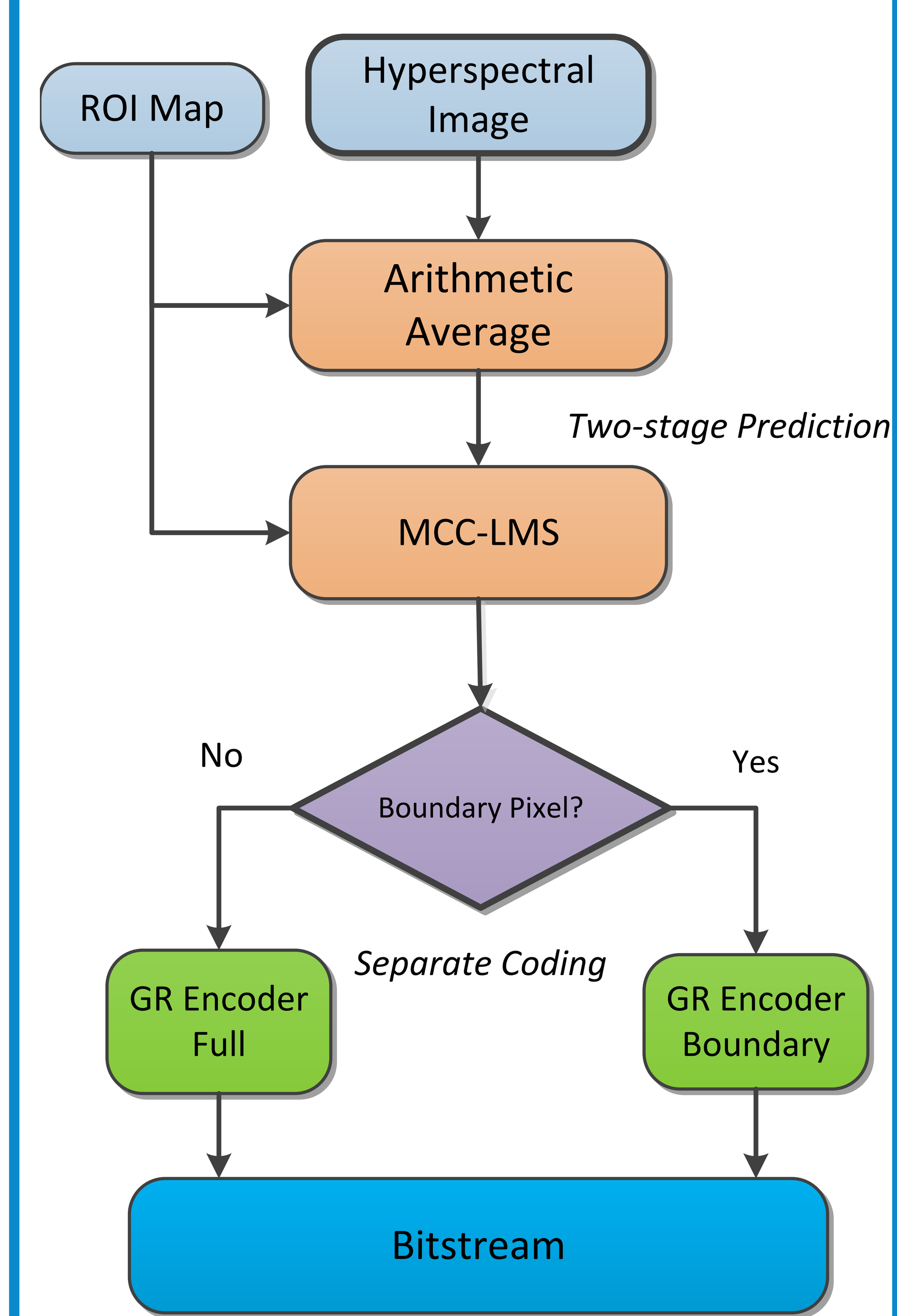
$$\mathbf{W}_{n+1} = \mathbf{W}_n + \frac{\mu}{N\sqrt{2\pi}\sigma^3}\sum_{i=n-N+1}^{n}\left[\exp\left(\frac{-{e_i}^2}{2\sigma^2}\right)e_i\mathbf{X}_i\right],$$

where $e_i = d_i - \mathbf{W}_n^T\mathbf{X}_n$. Inspired by the stochastic gradient, N is set to 1. Therefore,

$$\mathbf{W}_{n+1} = \mathbf{W}_n + \frac{\mu}{\sqrt{2\pi}\sigma^3}\exp\left(\frac{-{e_n}^2}{2\sigma^2}\right)e_n\mathbf{X}_n,$$

This correntropy-induced updating function can be viewed as LMS with a self-adjusting learning rate, which reflects the outlier rejection property of the correntropy.

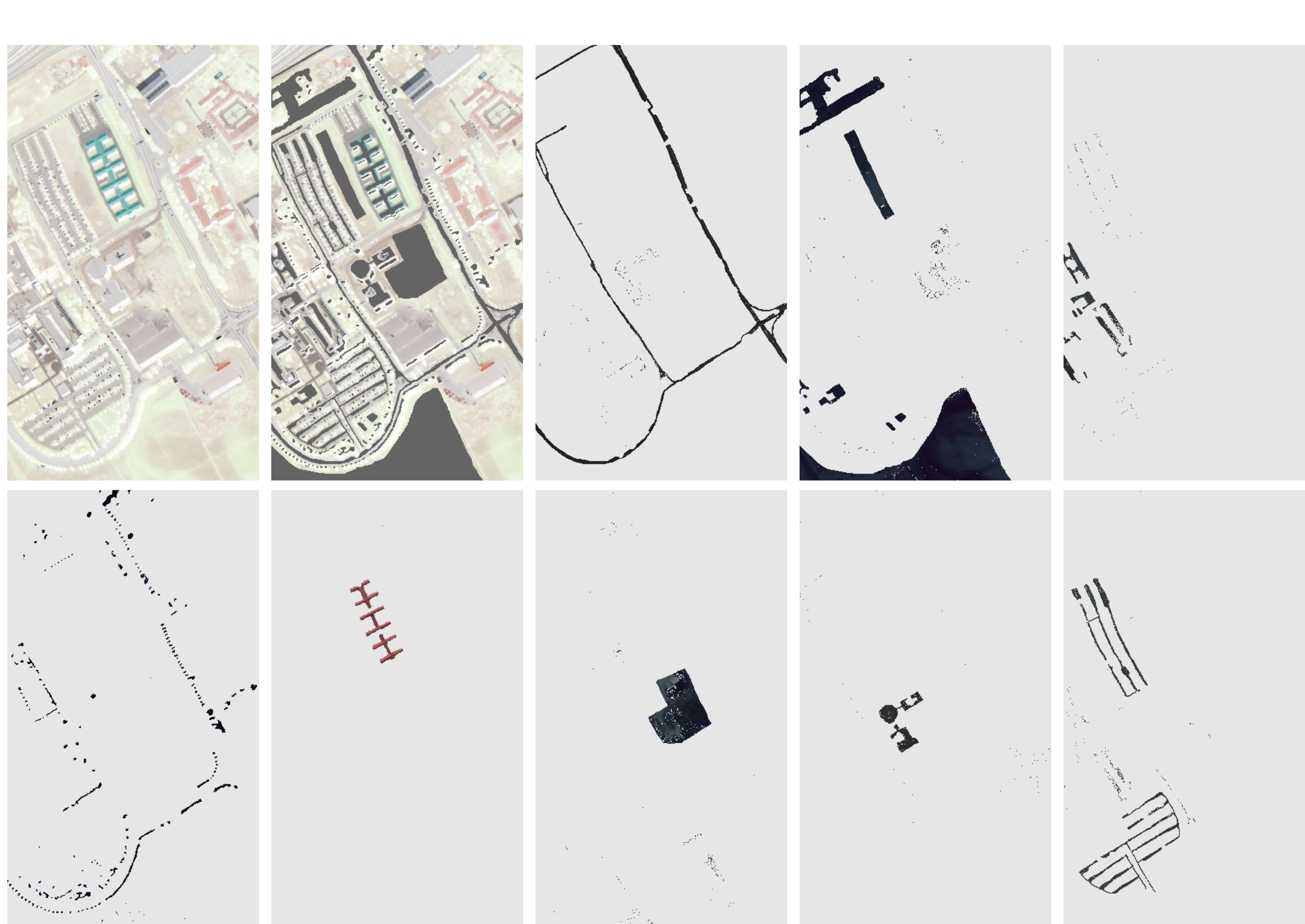
Algorithm



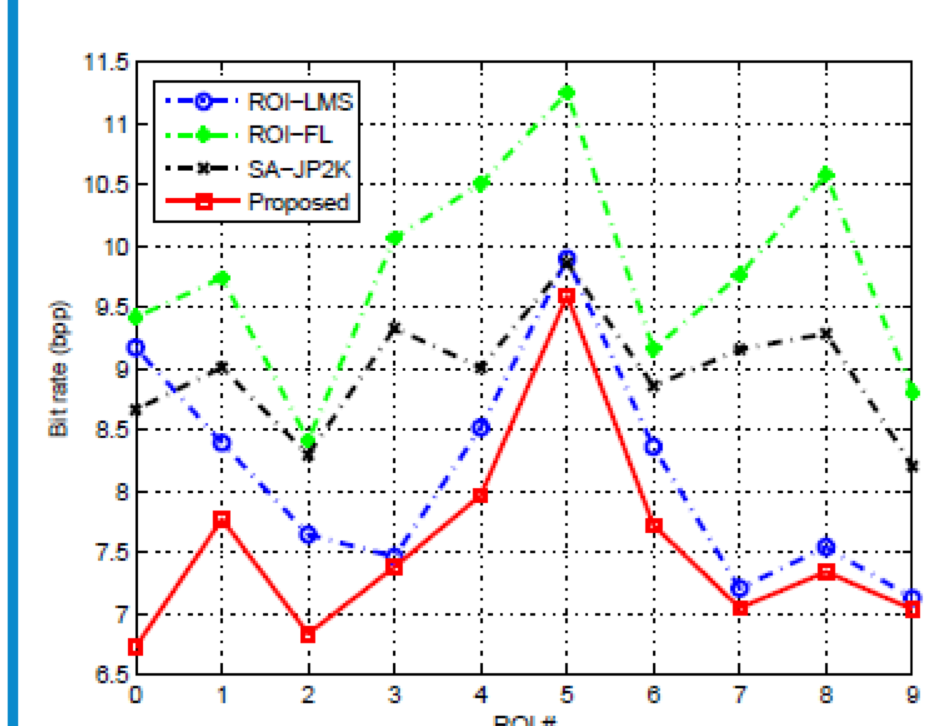
Compression Results for Individual ROIs

The table below shows the bit rates of the four methods as a result of compressing each of the ten ROIs in the dataset, "Pavia University".

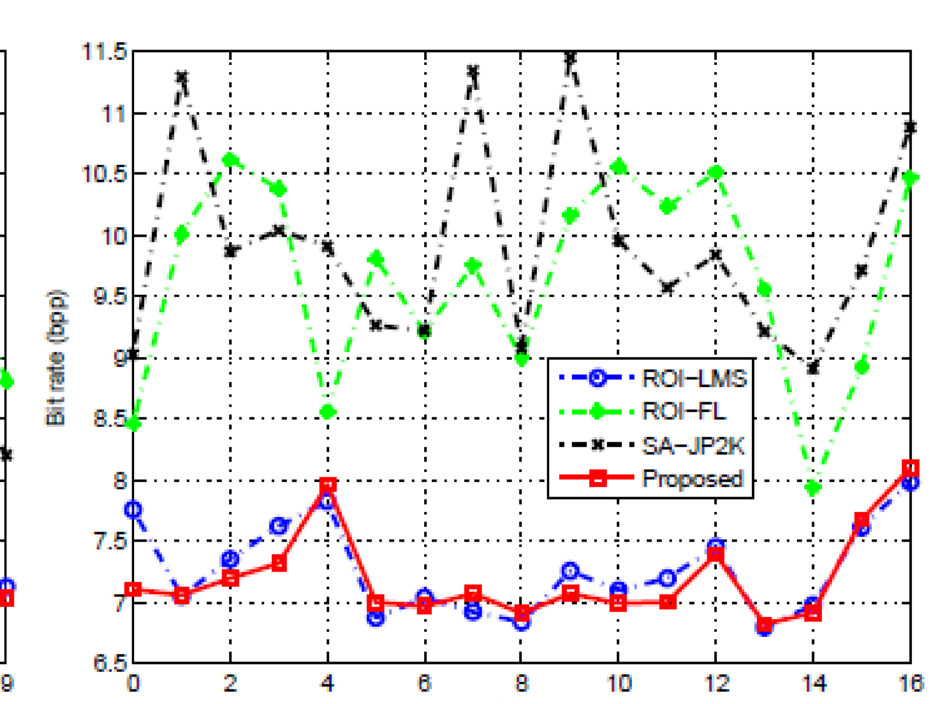
ROI#	Proposed	ROI-LMS	ROI-FL	SA-JPEG 2000
0	**6.73**	9.17	9.42	8.67
1	**7.77**	8.40	9.74	9.01
2	**6.84**	7.65	8.42	8.30
3	**7.38**	7.47	10.07	9.33
4	**7.97**	8.52	10.51	9.01
5	**9.60**	9.90	11.25	9.86
6	**7.72**	8.36	9.16	8.86
7	**7.05**	7.21	9.76	9.16
8	**7.34**	7.55	10.58	9.29
9	**7.04**	7.13	8.81	8.21



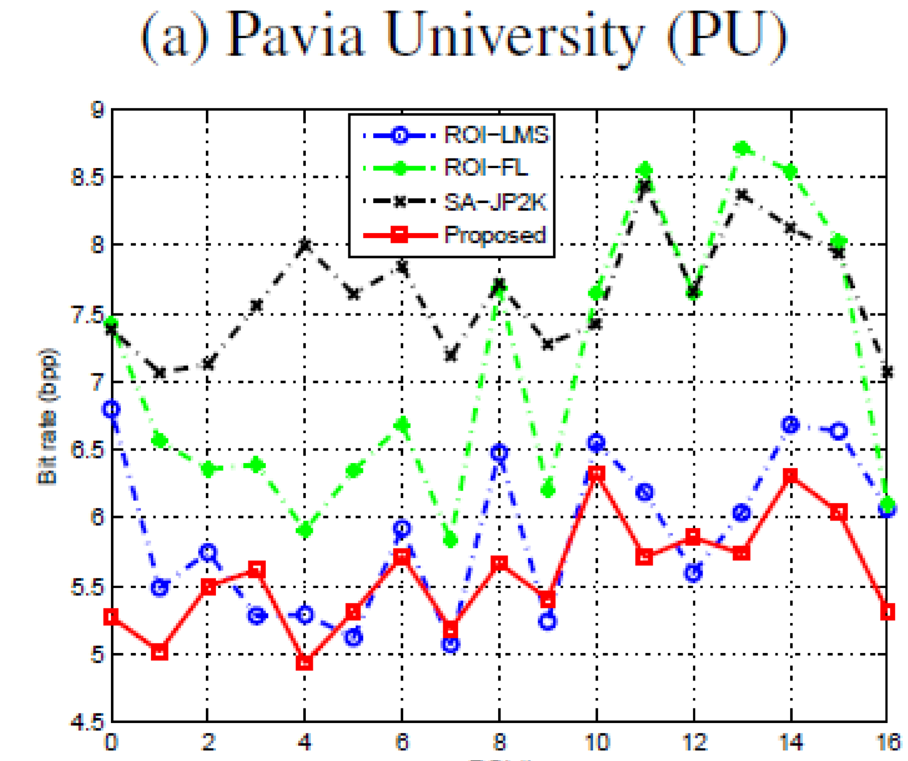
Compression Results for Several Datasets



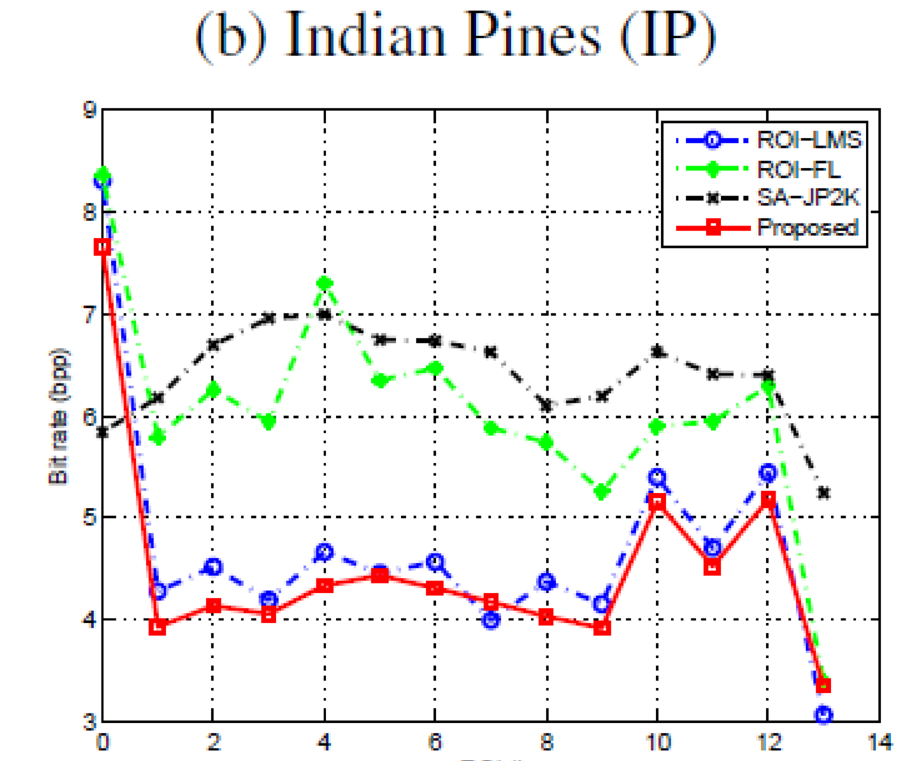
(a) Pavia University (PU)



(b) Indian Pines (IP)



(c) Salinas



(d) Kennedy Space Center (KSC)

The table below shows the average compression bit rates of the four datasets: the proposed method achieves the largest compression on PU, IP and Salinas, as well as on KSC (except for ROI0, which accounts for over 97% of the pixels of KSC and it is corrupted with significant impulse noise).

DATASETS	Proposed	ROI-LMS	ROI-FL	SA-JPEG 2000
PU	**6.85**	8.92	9.39	8.68
IP	**7.11**	7.49	9.12	9.30
Salinas	**5.45**	6.38	7.25	7.49
KSC	7.57	8.21	8.29	**5.85**

Conclusion

Simulation results demonstrated the proposed algorithm offered higher compression on ROIs with no-data regions than ROI-LMS, ROI-FL, and Shape Adaptive JPEG 2000.

Major Reference

[1] A. Singh and J. C. Principe, "Using correntropy as a cost function in linear adaptive filters," in *International Joint Conference on Neural Networks*, June 2009, pp. 2950–2955.