GAUSSIAN NOISE REMOVAL FOR WET CHEMISTRY DATA FROM THE PHOENIX MISSION

Y. Mu[1], W. Ding[1], X. Ren[1], E. Oberlin[2], S. Kounaves[2], [1] Department of Computer Science, University of Massachusetts Boston, 100 Morrissey Blvd., Boston, MA02125 (yangmu, ding@cs.umb.edu, xiangren2000@gmail.com), [2]Department of Chemistry, Tufts University, 62 Talbot Avenue, Medford, MA 02155 (elizabeth.oberlin, fsamuel.kounaves@tufts.edu).

Introduction: The initial in-situ chemical analysis of the four soil samples delivered to the Wet Chemistry Laboratory (WCL) provided concentrations of the soluble ionic species in the soil along with the pH and conductivity. Even though at first the results appeared to present a simple picture of the soluble soil chemistry, it has become clearly evident that the analyses were complicated by an unplanned titration of the soil/water mixture with barium chloride, the presence of chemical species whose signal was convoluted into other sensor readings, and the effects of temperature and noise. The combined effect of these interferences has put into question the concentrations and identity of the ionic species that were directly measured. Use of such incorrect values in conjunction with chemical speciation modeling programs will result in misinterpretation of the parent salt composition of the soil analyzed by the Phoenix WCL.

The WCL on board the phoenix lander performed the first comprehensive wet chemical analysis of the soil on Mars, during the summer of 2008. Each WCL consisted of a lower cell whose walls were lined with an array of sensors and an upper assembly for adding water, reagents, soil, and stirring. The sensor array included ion selective electrodes (ISE) for K^+, Na^+, Mg^+, Ca^+, etc and electrodes for conductivity, redox potential, cyclic voltammetry, chronopotentiometry, and an PH electrode [1]. But as the effect of temperature, inclement weather and several types of interferences on Mars, the WCL sensor data presents a significantly “cloudy” and noisy picture, which brought trouble to the further research. Therefore, the basis of solving the above problem will require significantly improved signal resolution and accuracy.

To obtain correct sensor readings for conversion to concentrations and to identify other chemical species present, will require a combination of complex processing/analysis of the returned WCL data, parallel laboratory analyses using the remaining two WCL flight spare units, and equilibrium modeling to validate such results. In this paper, we report on our investigation of a promising strategy to address the aforementioned problem. Our strategy consists of developing and implementing of a new methodology using a Bayesian Least Square estimator on signal denoising, a fuzzy sequence pattern matching approach on white noise removal.

Methodology: We propose to perform denoising in the Wavelet domain instead of the original signal domain in order to take full advantage of statistical image modeling, as shown in Figure 1. The signal data will be first transformed into a 2D matrix, and then the matrix will be decomposed into pyramid subbands at difference scales in the Wavelet domain. Denoising will be performed at each subband. Finally, we will invert the pyramid transformation and obtain the denoised signal.

Our technical approach to Bayesian white noise denoising for signal processing relies on a set of techniques taken from computer vision and statistics. For any noise input signal vector $\vec{y} = [a_1, a_2, \cdots a_n]$, we propose a new method denoted as Bayesian Probability Mapping. Specifically, the n samples are quantized into d bins according to the input value range. i.e., a_1 is quantized into bin b_1, which means a_1 has a probability of 1 to fall into this bin. Therefore, the input signal could be converted to a probability matrix $P \in \mathbb{R}^{d \times n}$. The $d \times n$ matrix is then transformed into the wavelet domain using a set of multiscale bandpass orthogonal filters. In our previous work on crater detection from HRSC images [2], wavelets have proved to be a very powerful tool for pattern analysis because wavelet decomposition can efficiently enhance statistical features of images.

In this project, we will exploit dependencies between Wavelet coefficients based on Gaussian scale mixtures. It has been observed that most responses of the subband of the Wavelet filter have a near zero value because most of the matrix entries are either not data entries or simply random noise data. A portion of the signal response that corresponds to the real signal has comparatively large amplitude responses.

For the WCL data, we propose to model local clusters of wavelet coefficients of a pyramid subband using Gaussian Scale Mixtures. A random vector $\vec{x}$ is a Gaussian scale mixture if and only if it is expressed as $\vec{x} = \sqrt{\lambda}\vec{u}$, where $\vec{u}$ is a zero-mean Gaussian vector, λ is an independent positive scalar random variable, and here $=$ indicates equality in distribution. We then calculate the local neighborhood noise covariance C_w and the observed signal in Wavelet domain local neighborhood covariance C_y. C_w can be computed by averaging the products pairs of coefficients over all local neighborhoods of the subband. Bayesian Least Squares estimate can be written as [3]:

$$E\{x|y,\lambda\} = \int E\{x|y,\lambda\}p(\lambda|y)d\lambda$$
$$= \lambda(C_y - C_w)(\lambda(C_y - C_w) + C_w)^{-1})y,$$

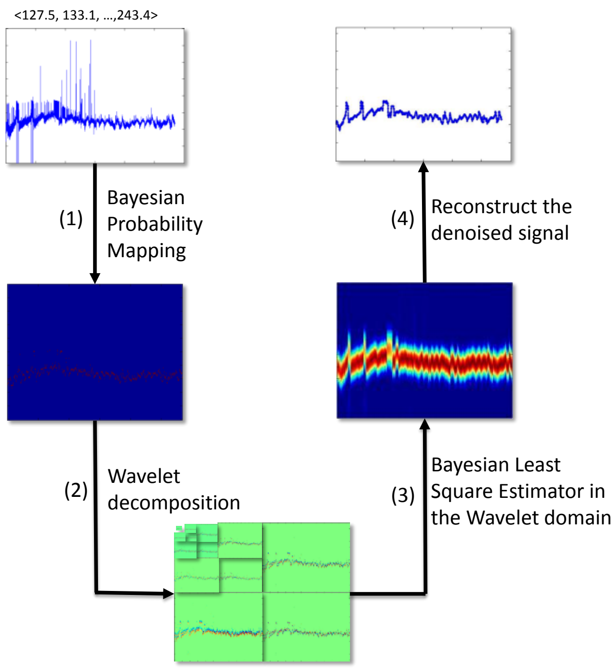

Figure 1. Bayesian Least Squares estimator denoising on Cl^- in Sol 30.

where $E\{x|y,\lambda\}$ is the estimated Wavelet coefficients of the denoised signal. The last step is the standard routine in image denoising by inverting the pyramid subbands to produce the denoised sensor data. A major branch on existing signal denoising methods relies on user-defined parameters, i.e., the thresholds of noise filters. However in the WCL data, we cannot validate the parameter settings with ground truth data. Thus the proposed nonparametric Bayesian probability approach is a natural choice to be examined.

Results: First, we have performed simulations with white Gaussian noises in a standard deviation of 0.5 added to the standard Sine function (see Figure 2). Then, we take our algorithm on this noise signal, the result shows that denoised Sine function using our proof-of-concept algorithm matches well with ground truth, which confirms our assumption for the proposed approach on the white noise denoising.

Second, we use this algorithm on Li^+, phA, Irph, and Na^+ of the WCL data on sol 30 (see Figure 3). Compareing the WCL data before denoising, denoised signals show a clear trend and also preserve intervals which represent different phases of the chemical experiment. In Figure 3, we can see that phA, Irph and Na^+ produce a dramatic fluctuation at about 2000 and 4000 Mars log time, which is caused by the addition of cali-

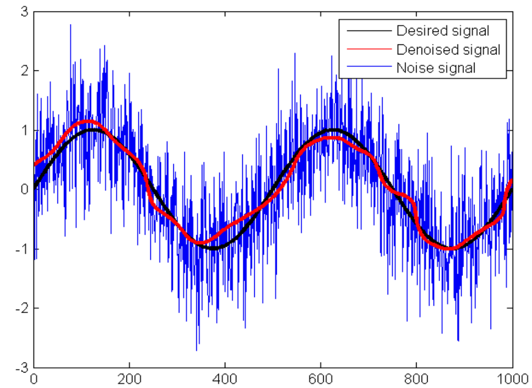

Figure 2. An illustrating denoising example for Sine function.

brant crucible and the soil sample, whereas Li^+ keep stable during the whole process. According to this property, we take Li^+ as a reference, i.e., the ISE of Li^+ is subtracted before analysis.

Conclusion: In this project, we have proposed an approach using Bayesian least squares estimation to remove white noise influence from the WCL data. We have conducted experiments on synthetic added Gausian noise data and WCL ions data from to evaluate our algorithm. The results demonstrate that white noise has been removed by using this approach. This idea is promising and requires further research on studying the denoised data.

Acknowledgements: This work was supported by NASA under Grant NNX13AJ69G.

References: [1] Kounaves, S. P.,(2009) J. Geophys. Res.,114, E00A19. [2] Ding, W. et al. (2011) *Transactions on Intel-ligent Systems and Technology*, Vol. 2, Issue 4, July. [3] Portilla, J. et al, (2003) *IEEE Transactions on Image Processing*, 12(11), 1338-1351.

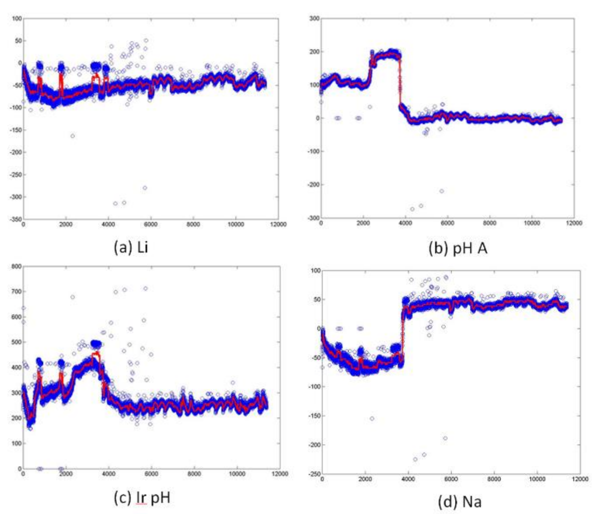

Figure 3. Sample denoising results.