PROCEEDINGS OF SPIE

SPIEDigitalLibrary.org/conference-proceedings-of-spie

Signal recovery for compressive spectrometers

Lu, Chih-Cheng, Chen, Kevin, Huang, Li-Ren, Kung, H.

Chih-Cheng Lu, Kevin Chen, Li-Ren Huang, H. T. Kung, "Signal recovery for compressive spectrometers," Proc. SPIE 10665, Sensing for Agriculture and Food Quality and Safety X, 106650U (15 May 2018); doi: 10.1117/12.2304367

Event: SPIE Commercial + Scientific Sensing and Imaging, 2018, Orlando, Florida, United States

Signal Recovery for Compressive Spectrometers

Chih-Cheng Lu[a], Kevin Chen[b], Li-Ren Huang[a], and H.T. Kung[c]

[a]Industrial Technology Research Institute, Hsinchu, Taiwan
[b]Google Inc., Cambridge, USA
[c]Harvard University, Cambridge, USA

ABSTRACT

Spectrometers are widely used for characterizing materials. Recently, filter-based spectrometers have been proposed to lower the manufacturing cost by replacing optical components with low-cost wavelength-selective filters, but at the expense of possibly lowered signal quality. We present compressive spectrometers which, based on the compressive sensing principle, are able to recover signal with improved quality from measurements acquired by a relatively small number of low-cost filters. We achieve high quality recovery by leveraging the fact that spectrometer measurements typically follow the shape of a smooth curve with a few spikes. We validate our method with real-world measurements, and release our dataset to facilitate future research.

Keywords: Filter-based spectrometer, compressive sensing, sparse coding, SWIR

1. INTRODUCTION

Spectral analysis is a well-established technique used in physics, chemistry, and biology. It provides detailed information related to the chemical bonds of the molecule, and thus can identify the compositions of the sample and their concentrations [1][2]. Conventional optics-based spectrometers are expensive due to high-cost optics components and their large physical footprints. Recently, miniature filter-based spectrometers [3][4] have emerged to provide cost and size advantages over optics-based spectrometers. Instead of using dispersers, the new approach employs a bank of wavelength-selective filters to detect the corresponding spectrum. However, these miniature spectrometers usually cannot resolve the spectrum at a fine-gain level due to the difficulty of manufacturing filters with small leaks, resulting in lower signal quality. Additionally, many filters are needed in order to capture a large set of target wavelengths, and the non-ideal filtering mechanism makes reconstruction necessary.

In [5], the linear equation problem is converted to be overdetermined with Gaussian kernels in which the number of spectral components is smaller than the number of equations. They use ℓ_2-norm minimization to solve the problem, however the reconstructed accuracy is limited by the number of filters. According to the compressive sensing theory, sparse nature of signal can be capture and represented at a rate significantly below the Nyquist rate [6]. The method employs non-adaptive linear projections that preserve the structure, and then reconstructed from these projections using an optimization process. [7] utilizes the sparse nature of the spectrum, and convert the linear equation problem with more atoms in the Gaussian kernels as those in [5]. By using ℓ_1-norm minimization, they demonstrate the accuracy can be increased multiple times compared to the results in [5].

To overcome the drawbacks of filter-based methods [8], our proposed compressive spectrometers use a small number of filters to capture information from multiple wavelengths at the same time. Based on sparse signal recovery principles in compressive sensing [9], we present a high-quality signal reconstruction method that exploits the fact that spectrum signal normally exhibits itself as a smooth curve with a few spikes.

Further author information: (Send correspondence to Chih-Cheng Lu)
Chih-Cheng Lu: E-mail: chris.lu@itri.org.tw, Telephone: 886 3 591 5763
Li-Ren Huang: E-mail: lrhuang@itri.org.tw, Telephone: 886 3 591 3755

Sensing for Agriculture and Food Quality and Safety X, edited by Moon S. Kim, Kuanglin Chao, Bryan A. Chin, Byoung-Kwan Cho, Proc. of SPIE Vol. 10665, 106650U · © 2018 SPIE
CCC code: 0277-786X/18/$18 · doi: 10.1117/12.2304367

2. BACKGROUND

2.1 Compressive Sensing

Compressive sensing is a framework in which signals are sampled through a linear projection Φ [6]. In the case of compressive spectrometers, this linear projection is implemented as the filters in analog domain. In standard compressive sensing, it is assumed that all signals $x \in X$ can be sparsely expressed as $x = \Psi z$, where Ψ is a (overcomplete) basis and z is a sparse vector. The measurement y is defined as follows:

$$y = \Phi x + \epsilon = \Phi \Psi z + \epsilon \tag{1}$$

where ϵ denotes the noise in measurements.

Given y, one can recover x indirectly by first finding a sparse vector z that explains the measurements. There are several formulations for this with different performance guarantees, but the general solution is

$$z^* = \underset{z}{arg\,min}\, g(z) \text{ subject to } \|y - \Phi\Psi z\|_2 \leq \|\epsilon\|_2 \tag{2}$$

where g is an objective function that promotes sparsity (e.g., $g(z) = \|z\|_1$). Once z^* is recovered, one can obtain an estimation of x via $x^* = \Psi z^*$. In this paper, we extend this idea and consider a slightly more complex signal model that encourages both sparsity and smoothness.

2.2 Spectrometer

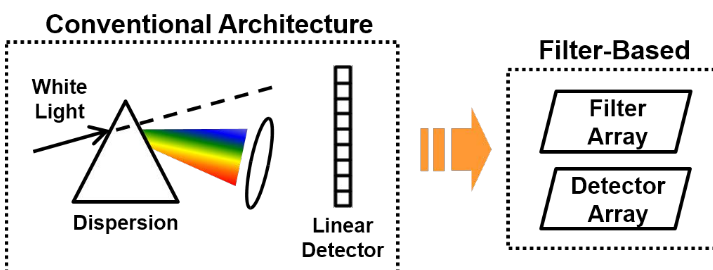


Figure 1. Block diagram of the conventional and filter-based architecture.

The left-hand side of Fig. 1 shows conventional spectrometers which typically composed of a collimating lens, a dispersion element, a focusing lens, and a linear detector. To eliminate the optical path, filter-based architecture shown in the right-hand side of Fig. 1 is introduced. Fig. 2 depicts the model of the filter-based architecture. Let the input spectrum $x(\lambda)$ which represents the reflection or transmission light power at a given wavelength λ passes through each filter, and output y_i by the corresponding sensor. The transmittance of each filter $f_i(\lambda)$ define the wavelength that can pass through and be detected by the sensor; the output y_i of the i^{th} sensor is given by $y_i = \int f_i(\lambda)\, x(\lambda)\, d\lambda$. By measuring the output y_i, the input spectrum $x(\lambda)$ can be theoretically recovered as $\widehat{x}(\lambda)$. Ideally, we can make a filter that only one wavelength can pass through, and thus no reconstruction algorithm is needed. The reconstruction effort will be depending on the complexity of filter design.

2.3 Previous Signal Recovery Methods

Several signal recovery methods for spectrometers have been proposed in the literature. In essence, these methods can be viewed as solving generalized least square problem with varying regularizations that correspond to different signal model. Specifically, the recovery methods are typically solving problems of the following form:

$$z^* = \underset{z}{arg\,min} \|y - \Phi f(z)\|_2 + \lambda g(z) \tag{3}$$

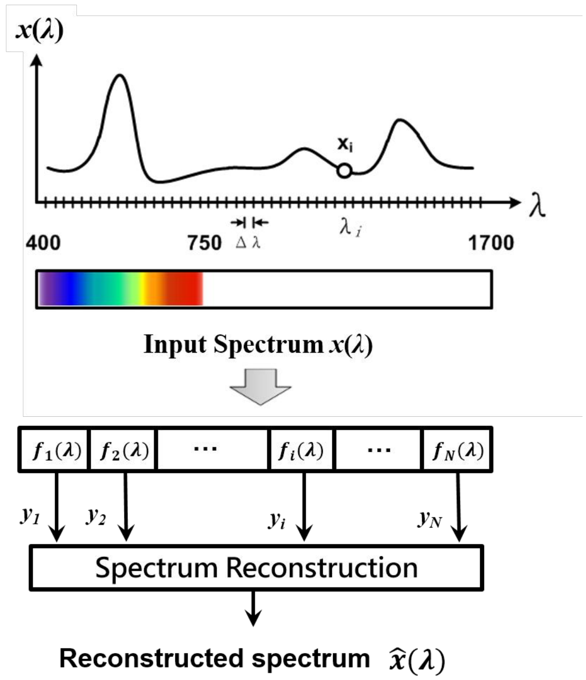


Figure 2. Model of the filter-based spectrometer.

where f is a predefined function that maps the solution to the signal space (e.g., in compressive sensing $f(z) = Dz$), and g is the regularization function with respect to z (e.g., in compressive sensing $g(z) = \|z\|_1$). An estimation of the signal is obtained by $x^* = f(z^*)$.

Tikhonov regularization is one of the most common method used in the literature [5]. Using the previous general form, this method can be expressed as having $f(z) = \Psi z$ for some basis Ψ, and $g(z) = \|Az\|_2^2$ for some transform matrix A. The basis Ψ can be a collection of Gaussian kernels as suggested in [5]. There are several choices of A, notably the identity matrix (which treats the coefficients as multivariate Gaussian variables, and leads to coefficients with smaller ℓ_2 norms) and the 1st order gradient matrix (which treats the gradient of the coefficients as multivariate Gaussian variables, and leads to smooth signal).

Another method used in the literature is based on sparse coding [7]. In this case, the regularization function is something that promotes sparsity, such as $g(z) = \|z\|_1$. It has been shown that in very simple cases such as LED spectrum reconstruction, the sparse coding method tends to do better than Tikhonov regularization [5]. However, none of the methods in the literature can recover complex signals such as the spectrum of real plastic materials. The difficulty that arises in these practical applications is due to the fact that real signals are neither entirely sparse or smooth.

3. SIGNAL ACQUISITION

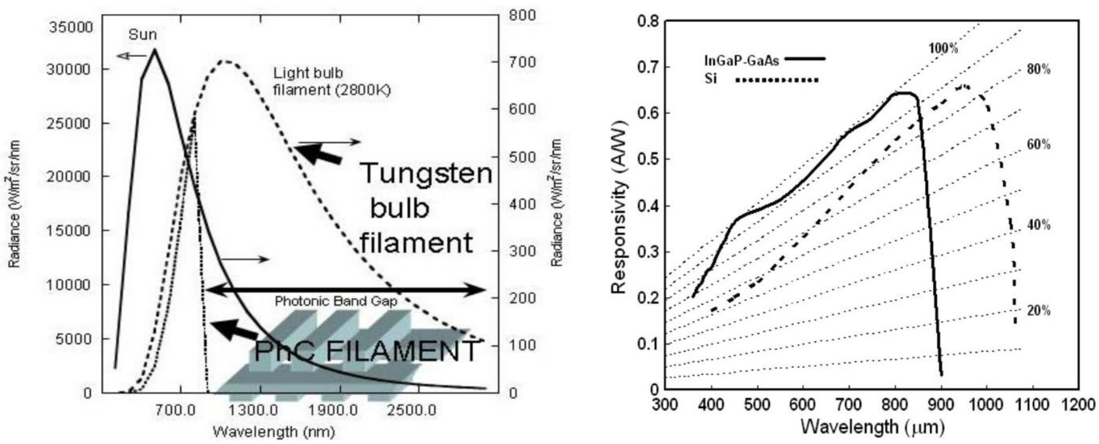
Signals are sensed through a filter-based spectrometer. The measurements y obtained by the spectrometer is modeled by

$$y = \Phi x \tag{4}$$

where x is target signal (spectrum) and Φ is the sensing matrix that is composed of 3 components: the filter response, sensor gain in a given bandwidth, and the individual sensor gain. For simplicity, we assume that the effect Φ is already known. In practice, there are many limitations to the structure of Φ due to the manufacture process, and often has high coherence between columns in Φ. Examples of actual filters can be found in Fig. 6.

4. HYBRID MODEL OF SMOOTHNESS AND SPARSITY

There are several factors that contribute to the final spectrum detected by a spectrometer. First, the light source is smooth and relatively broadband as shown in Fig. 3(a) [10]. This source signal then reflects off some surface, which selectively absorbs specific narrow band depending on the material properties of the surface. Then, the reflected signal is captured by the sensor with characteristics as shown in Fig. 3(b) [11]. As a results, the signal acquired at the sensor would be roughly smooth except around the narrowband wavelength segments that got absorbed by the reflective surface.



(a) (b)

Figure 3. (a) Signal strength of light source [10]. (b) Selectivity at the sensor [11].

Spectrum signal tend to be a smooth curve with a few spikes of varying magnitudes. This is the result of several contributing factors throughout the sensing process, as illustrated in Fig. 3. We propose a signal model that treats the signal x as the composition of a sparse component and a smooth one: $x = v + \Psi z$ where v is smooth, Ψ is a sparsifying basis and z is sparse. The measurements y is defined as $y = \Phi x$ where Φ is the sensing matrix.

5. ALGORITHM

The sparsity and smoothness assumption manifests as separate regularization terms in the optimization problem for signal reconstruction:

$$\underset{v,z}{arg\,min} \|y - \Phi(v + \Psi z)\|_2^2 + \lambda_1 \|z\|_1 + \lambda_2 \|Av\|_2^2 \tag{5}$$

where A is a bidiagonal (1, -1) matrix such that Av captures gradients in adjacent components of v. The choice of using ℓ_2 norm rather than ℓ_1 norm for Av reflects the fact that v is more likely to have many small changes instead of few large ones. Note that (5) is convex and can be solved efficiently with gradient descent methods [12].

It can be shown that the optimization in (5) is convex, which means that it can be solved very efficiently. At larger scale, gradient descent methods tends to have performance. Since the ℓ_1 sparsity regularization term is not differentiable, we recommend using proximal gradient descent as the solver. Note that all solvers would eventually reach the same solution due to convexity, so the choice of solver is mostly for better efficiency.

6. EVALUATION

In this section, we present details of the collected dataset [13], and performance comparisons between our hybrid model versus previous approaches.

6.1 Experiment Setup and Dataset

Spectrum signals are collected using a RED-Wave-NIRX-SR spectrometer with SL1 tungsten lamp. This spectrometer acquires the reflection spectra off plastics from 1000nm to 1656nm at a very high resolution (1nm). We use these measurements as the ground-truth signals x.

Table 1 tabulates the two settings in our plastic spectrum dataset. We measure seven different types of plastics based on ASTM international standards. In setting (A), we measure several spectrum from different items within the same plastic type. This captures the inter and intra-class variations of different plastic types. In setting (B), for each plastic type we measure several spectrum of the same item with varying distance, location, angle, etc. This captures the variations in measuring the same material.

In order to validate our signal reconstruction approach with filter-based spectrometers, we measure filter characteristics matrix Φ (i.e., the sensing matrix) using the Oriel Cornerstone 130 monochromator with 64 filters at the same wavelength interval between 1000nm and 1656nm. The signal resolution of the monochromator is 12nm, so the measured matrix Φ was upsampled to match the resolution of the ground-truth signals.

Table 1. Number of samples from each plastic type in the dataset.

	1 PETE	2 HDPE	3 V	4 LDPE	5 PP	6 PS	7 OTHER	
	No.1	No.2	No.3	No.4	No.5	No.6	No.7	Total
(A)	10	13	7	11	11	11	2	65
(B)	20	20	20	20	40	20	20	160

6.2 Comparison between Methods

In this Section, we compare our hybrid method with state-of-the-art methods: conventional sparsity-based recovery method using a sparse model in ℓ_1 regularization[7] and the Tikhonov regularization method in ℓ_2 (based on smoothness assumption) [5]. To our best knowledge, these models have produced the best signal recovery quality in the literature. However, we note that the comparisons done in these papers are only considering simple LED signals, and may not reflect the signal quality under practical sensing scenarios. In our experiments, we evaluate all methods using the real datasets described in Section 6.1. Specifically, we train and validate our models on dataset (B), and report the signal reconstruction error of dataset (A).

We consider the signal reconstruction error as a function of the number of filters (i.e., number of measurements) used, where the sensing matrix Φ is drawn from Gaussian distribution. As shown in Fig. 4, our hybrid method consistently outperform the other methods significantly. Note that all methods are capable of achieving the target error margin (marked as the black dotted line) at 64 random measurements (the dimension of original signal is 1150), which shows that the spectrum signals are indeed highly compressible. Plot examples of reconstructed signals (plastic type II) under different methods are shown in Fig. 5. The hybrid model is able to capture the valley more accurately (see the inset).

In practice, the filters used in spectrometers are far from random. Due to the manufacturing process, the actual filters tend to be very smooth and Φ is in fact quite coherent. Figure 6 shows some examples of real filter responses. We performed the recovery error comparison using Φ measured from an actual filter-based spectrometer, and report the reconstruction error in Table 2. This specific spectrometer has 64 filters, which correspond to 64 measurements. Our method achieves significant better performance in minimizing recovery error than the state-of-the-art methods for our dataset.

Table 2. Comparison of Reconstruction Error using real filters

	SC	TV	Hybrid	Target
Error	.021	.038	**.013**	.013

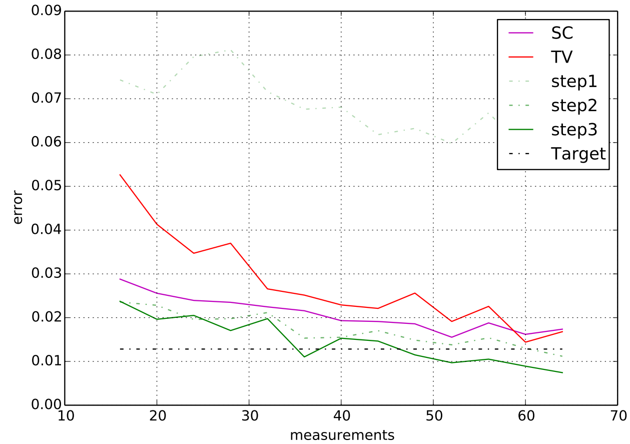


Figure 4. Performance of different methods over number of measurements. The parameters for each method is screen under $m = 64$ and kept fixed for other number of measurements. Note that our proposed hybrid method outperforms methods that only relies on either sparsity or smoothness.

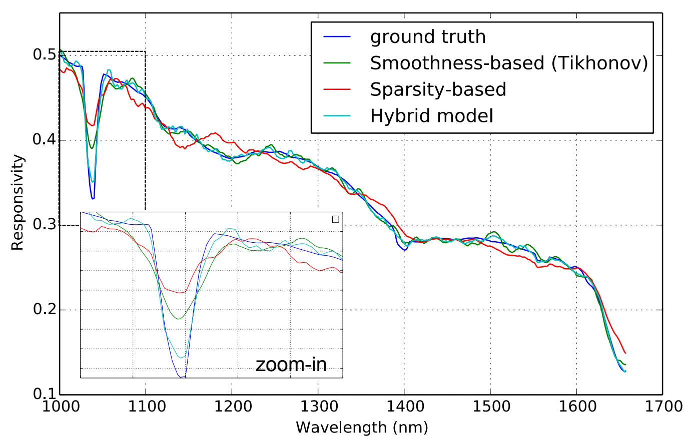


Figure 5. Examples of recovered signals(plastic type II) using different methods.

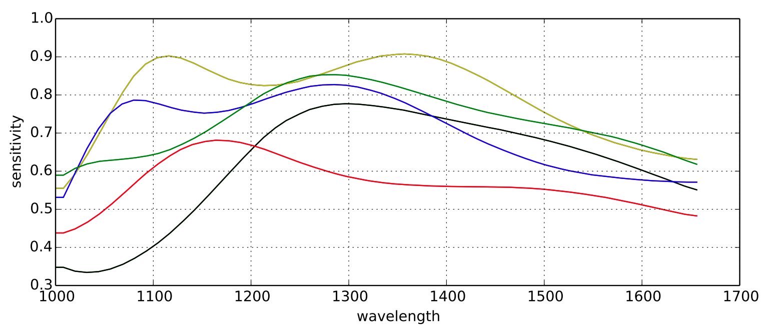


Figure 6. Examples of filter responses from a filter-based spectrometer with 64 filters.

7. CONCLUSION

We propose compressive spectrometers that can have lower manufacturing cost. This is because unlike conventional filter-based spectrometers, our method does not require filters with small leakage, and uses much fewer filters for signal reconstruction. The proposed method leverages the fact that spectrum signals tend to exhibit a few spikes over a smooth curve. By enforcing sparsity (for spikes) and smoothness in the signal recovery process, we achieve low reconstruction error even under significant compression (Figure 4). We validate our method with real measurements from spectrometers, and release our dataset to the community to facilitate research.

ACKNOWLEDGMENTS

This work is supported in part by the Industrial Technology Research Institute under the Hon-Hu Talent Development Program, in part by gifts from the Intel Corporation and in part by the Naval Supply Systems Command award under the Naval Postgrate School Agreements No. N00244-15-0050 and No. N00244-16-1-0018.

REFERENCES

[1] Zornoza, R., Guerrero, C., Mataix-Solera, J., Scow, K., Arcenegui, V., and Mataix-Beneyto, J., "Near infrared spectroscopy for determination of various physical, chemical and biochemical properties in mediterranean soils," *Soil Biology and Biochemistry* **40**(7), 1923–1930 (2008).

[2] Yu, H., Niu, X., Lin, H., Ying, Y., Li, B., and Pan, X., "A feasibility study on on-line determination of rice wine composition by vis–nir spectroscopy and least-squares support vector machines," *Food Chemistry* **113**(1), 291–296 (2009).

[3] Oliver, J., Lee, W.-B., and Lee, H.-N., "Filters with random transmittance for improving resolution in filter-array-based spectrometers," *Optics Express* **21**(4), 3969–3989 (2013).

[4] Lee, W.-B., Kim, C., Ju, G. W., Lee, Y. T., and Lee, H.-N., "Design of thin-film filters for resolution improvements in filter-array based spectrometers using dsp," in [*Next-Generation Spectroscopic Technologies IX*], **9855**, 98550Q, International Society for Optics and Photonics (2016).

[5] Kurokawa, U., Choi, B. I., and Chang, C.-C., "Filter-based miniature spectrometers: spectrum reconstruction using adaptive regularization," *IEEE Sensors Journal* **11**(7), 1556–1563 (2011).

[6] Donoho, D. L., "Compressed sensing," *Information Theory, IEEE Transactions on* **52**(4), 1289–1306 (2006).

[7] Oliver, J., Lee, W., Park, S., and Lee, H.-N., "Improving resolution of miniature spectrometers by exploiting sparse nature of signals," *Optics Express* **20**(3), 2613–2625 (2012).

[8] Neumann, N., Ebermann, M., Kurth, S., and Hiller, K., "Tunable infrared detector with integrated micromachined fabry-perot filter," *Journal of Micro/Nanolithography, MEMS, and MOEMS* **7**(2), 021004–021004 (2008).

[9] Candes, E. J., Romberg, J. K., and Tao, T., "Stable signal recovery from incomplete and inaccurate measurements," *Communications on pure and applied mathematics* **59**(8), 1207–1223 (2006).

[10] Zhou, Y., Sridhar, S., Mishra, K., Daubach, R., Garner, R., and Klotzkin, D., "Fabrication of tungsten photonic crystals for high efficacy incandescent light emitters," in [*2008 8th IEEE Conference on Nanotechnology*], 75–77, IEEE (2008).

[11] Wu, M.-C., Huang, Y.-H., and Ho, C.-L., "High-speed ingap/gaas pin photodiodes with wide spectral range," *IEEE Electron Device Letters* **28**(9), 797–799 (2007).

[12] Beck, A. and Teboulle, M., "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM journal on imaging sciences* **2**(1), 183–202 (2009).

[13] "Spectrometer dataset." https://github.com/kunglab. Accessed: 2017-04-05.