Convolutional neural networks for the reconstruction of spectra in compressive sensing spectrometers

Cheolsun Kim, Dongju Park, Heung-No Lee

Cheolsun Kim, Dongju Park, Heung-No Lee, "Convolutional neural networks for the reconstruction of spectra in compressive sensing spectrometers," Proc. SPIE 10937, Optical Data Science II, 109370L (1 March 2019); doi: 10.1117/12.2509548

SPIE.

Event: SPIE OPTO, 2019, San Francisco, California, United States

Convolutional neural networks for the reconstruction of spectra in compressive sensing spectrometers

Cheolsun Kim, Dongju Park, and Heung-No Lee*

School of Electrical Engineering and Computer Science, Gwangju Institute of Science and Technology, Gwangju, Republic of Korea, 61005

ABSTRACT

In optical filter based compressive sensing (CS) spectrometers, an input spectrum is multiplexed and modulated by a small number of optical filters which have different sensing patterns. Then, detectors read out the modulated signals called measurements. By exploiting the CS reconstruction algorithms that utilize the measurements and the sensing patterns of optical filters, the spectrum is recovered. However, there exists a drawback on CS reconstruction algorithms. The input spectrum should be a sparse signal or be sparsely represented by a pre-determined sparsifying basis. In practice, however, the input spectrum could not be sparse or be sparsely represented by the pre-determined sparsifying basis. Therefore, the performance of spectral recovery using the CS reconstruction algorithms is varying according to the sparsity of the input spectrum and the sparsifying basis. In this paper, we implement a convolutional neural networks (CNNs) structure to reconstruct the input spectrum from the measurements of the CS spectrometers. The CNNs structure learns the way of solving the inverse problem of the underdetermined linear system. As an input of the CNNs structure, a spectrum calculated by multiplying a fixed transform matrix and the measurements is used. We investigate the reconstruction performance of the CNNs structure comparing with the CS reconstruction algorithm with different sparsifying basis. The experiment results indicate the reconstruction performance of the CNNs structure is compatible with the CS reconstruction algorithm.

Keywords: spectrometers, inverse problems, compressive sensing, deep learning, convolutional neural network

1. INTRODUCTION

Spectrometers with high resolution, wide operating range, small size, and single exposure acquisition have attracted much interest in research and development. Such spectrometers can provide on-site detection and quantification of various fields such as mobile applications[1], and medical services[2]. Optical filter based compressive sensing spectrometers can be one candidate for this kind of spectrometers. Thanks to the structural advantage that fewer filters are required and filters are attached to detectors, the size of the spectrometers can be reduced. In addition, exploiting compressive sensing technique, the measurements from fewer filters can reconstruct the input spectrum, which means that the resolution improvement is achieved.

Compressive sensing (CS)[3] is the framework for sampling and reconstruction of an input signal. In CS spectrometers, the different sensing patterns of filters spectrally modulate the input spectrum. Detectors read out these modulated inputs called measurements. Reconstruction algorithms recover the input spectrum using the sensing patterns and the measurements. Because the length of the input spectrum set to be far larger than the number of the measurements, the reconstruction algorithms deal with the underdetermined linear system. Numerous studies have proposed reconstruction algorithms such as orthogonal matching pursuit[4], and basis pursuit algorithm[5]. The sparsity of the input spectrum affects the reconstruction algorithms that solving the problem of recovering the input spectrum from the measurements (inverse problem). The input spectrum should be a sparse signal or could be sparsely represented by sparsifying basis such as wavelet transform matrix, Fourier transform matrix, and discrete cosine transform (DCT) matrix. In practice, however, not all kinds of spectra are sparse in a fixed sparsifying basis. Therefore, the reconstruction performance changes according to the sparsifying basis and spectra.

Previously, the deep learning based approach to solve inverse problems in CS have been proposed in various studies[6–8]. In this paper, we propose a simple framework based on convolutional neural networks (CNNs) to solve inverse problems in CS spectrometers. As a deep learning structure, we have trained the CNNs structure so that the output of the networks

*heungno@gist.ac.kr; phone +82-62-715-2237

Optical Data Science II, edited by Bahram Jalali, Ken-ichi Kitayama, Proc. of SPIE Vol. 10937, 109370L · © 2019 SPIE · CCC code: 0277-786X/19/$18 · doi: 10.1117/12.2509548

becomes the input spectrum of CS spectrometers. The input of the CNNs structure is the vector calculated by multiplying a fixed transformation matrix and the measurements. We investigate the reconstruction performance of the proposed framework comparing with the CS reconstruction algorithm with different sparsifying basis. The experiment results indicate the reconstruction performance of the proposed framework is compatible with the CS reconstruction algorithm.

2. COMPRESSIVE SENSING SPECTROMETER

2.1 System Description

For CS spectrometers, numerous optical structures have been proposed such as multilayer thin-film filters[9] and Fabry-peort filters[10,11]. In this paper, we use the multilayer thin-film filters for CS spectrometers. Let $\mathbf{x} \in \mathbb{R}^{N\times 1}$ denote the input spectrum of light source and $\mathbf{y} \in \mathbb{R}^{M\times 1}$ denote the measurements, The relation between $\mathbf{x}$ and $\mathbf{y}$ can be expressed by:

$$\mathbf{y} = \mathbf{T}\mathbf{x}, \tag{1}$$

where $\mathbf{T} \in \mathbb{R}^{M\times N}$ is the sensing matrix. Each sensing patterns of the filters represents the row of the sensing matrix. To take advantage of the CS framework, the number of measurements is set to be smaller than the number of spectral component of light source, i.e., $M < N$. If the input spectrum $\mathbf{x}$ is a sparse signal or sparsely represented in some basis, reconstruction algorithms can recover the $\mathbf{x}$ in the underdetermined system of Eq. (1). An input spectrum $\mathbf{x}$ can be sparsely represented by the sparsifying basis $\mathbf{\Phi} \in \mathbb{R}^{N\times N}$:

$$\mathbf{x} = \mathbf{\Phi}\mathbf{s}, \tag{2}$$

where $\mathbf{s} \in \mathbb{R}^{N\times 1}$ is the sparse vector. Then, Eq. (1) becomes:

$$\mathbf{y} = \mathbf{T}\mathbf{\Phi}\mathbf{s}. \tag{3}$$

In order to recover the $\mathbf{s}$ in Eq. (3), numerous algorithms have been proposed which are based on greedy iterative algorithms and convex optimization algorithms. From the recovered $\hat{\mathbf{s}}$, the reconstructed input spectrum of light source $\hat{\mathbf{x}}$ is $\mathbf{\Phi}\hat{\mathbf{s}}$.

In this paper, unlike the aforementioned algorithms, we use a simple framework based on CNNs to reconstruct the input spectrum of light.

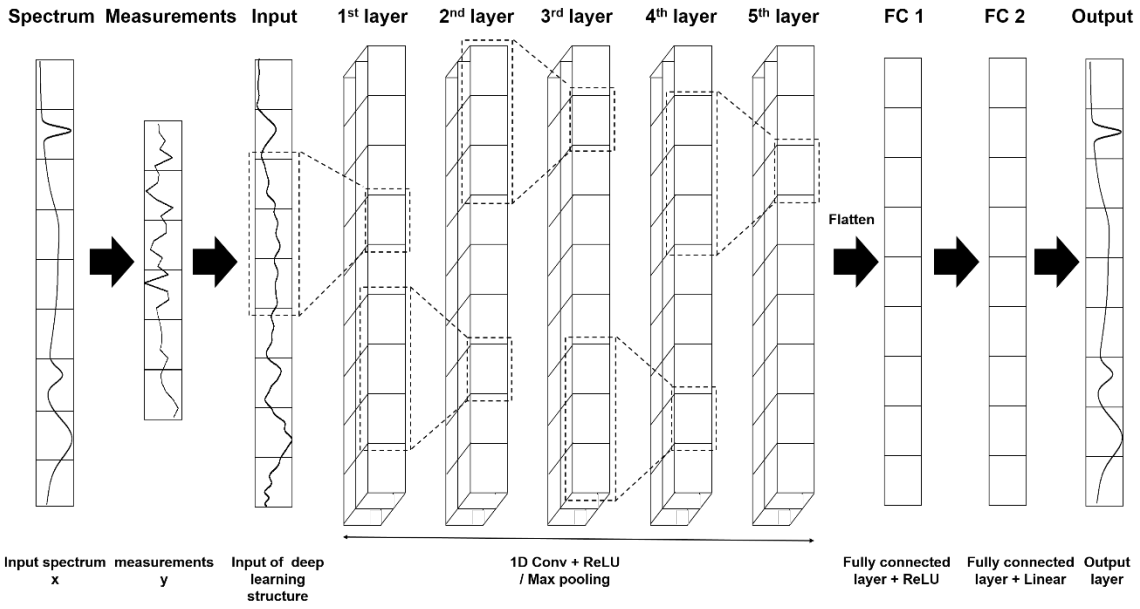

Figure 1. The schematic of the proposed framework to solve inverse problems in CS spectrometers.

2.2 CNN framework: Solving inverse problems in CS spectrometers

Convolutional neural networks (CNNs)[12] are one of the most popular deep learning structures in various imaging research areas such as image classification[13], image super-resolution[14], and image segmentation[15]. Typically, the CNNs consist of an input layer, multiple hidden layers, and an output layer. The hidden layers are consist of convolutional layers with activation functions, pooling layers, and fully connected layers.

We propose a simple framework based on CNNs that the output of the framework becomes the input spectrum of CS spectrometer. Figure 1 shows the schematic of the proposed framework. By the sensing matrix $\mathbf{T}$ of CS spectrometer, the input spectrum $\mathbf{x}$ is modulated as the measurements $\mathbf{y}$. As the input of the CNNs structure, we generate an N-dimensional vector from the measurements $\mathbf{y}$ by multiplying the transform matrix $\mathbf{\Theta} \in \mathbb{R}^{N\times M}$, i.e., $\mathbf{\Theta y}$. We use the pseudo-inverse matrix of the sensing matrix $\mathbf{T}$ as the transform matrix. For the convolutional layers, 1D CNN is used to cover one-dimensional input signals (spectra). After the convolutional layer, input signal passes through an activation function. We utilize the rectified linear unit (ReLU) as the activation function, i.e., $\mathrm{f}(x) = \max(0, \mathrm{x})$. Then, the output of the ReLU is fed into a pooling layer. We use a max pooling that takes a maximum value within a specific pooling window size of the output of the previous layer. We stack the convolutional layer and the pooling layer multiple times, alternatively. Finally, the output of the last layer is flattened and go through fully connected layers. The output layer is the vector $\hat{\mathbf{x}} \in \mathbb{R}^{N\times 1}$. We train the framework so that the vector $\hat{\mathbf{x}}$ becomes the input spectrum $\mathbf{x}$.

The specific information about the CNNs structure is as follows: it has five convolutional layers and two fully connected layers. The first and second convolutional layers have 128 filters with the stride of three. The third layer has 256 filters with the stride of three. The fourth and fifth layers have 128 filters with the stride of three. In each convolutional layers, the ReLU work as the activation function. We use max pooling with the pool size of two and the stride of two. In addition, the first fully connected layer gives an output of the vector size of 512. We use the ReLU for the first fully connected layer. The second layer gives a vector with the size of 500. As the activation function of the second layer, we use the linear function.

3. EXPERIMENT AND RESULTS

A multilayer thin-film CS spectrometer with 36 filters is considered (M = 36). We set the wavelength range from 500 to 1000 nm as the range of interest (N = 500). The input light is assumed to be in normal incidence. We numerically generate the multilayer thin-film filters by using multiple layers of high and low refractive index materials with different thicknesses. Each sensing patterns of filters are calculated based on thin-film theory[16].

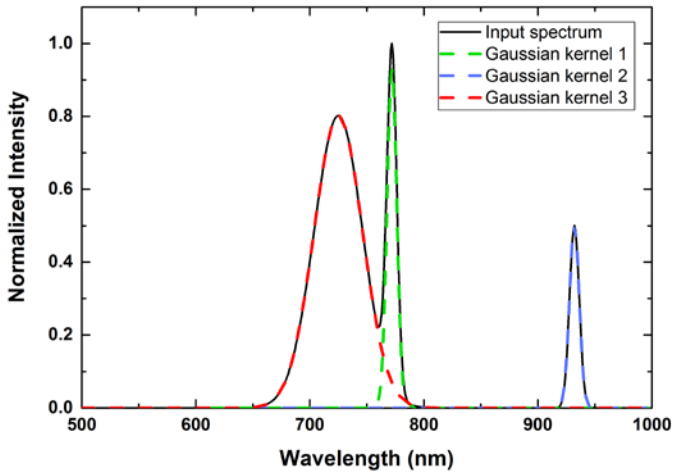

Figure 2. An example of the input spectrum composed of three Gaussian kernels.

3.1 Data set

We numerically generate spectra using Gaussian kernels. As shown in Fig. 2, three Gaussian kernels with the different full width at half maximums (FWHMs), peak locations, and peak heights can make an input spectrum. We make a data set composed of a training set, a validation set, and a test set. For generating spectra, we define the spectral types according to the number of Gaussian kernels used to make the spectrum. From one to 18 kernels, pairs of 18 types-spectra and their corresponding inputs of the CNNs structure are used as the data set. For the given number of Gaussian kernels, we can randomly generate spectra by changing FWHMs, peak locations, and heights. We use 800, 200, and 200 spectra for each type in the training set, validation set, and test set, respectively. Therefore, the number of spectra for the training set, the validation set, and the test set are 14400, 3600, and 3600, respectively.

3.2 Experiment and results

Let $\boldsymbol{D}_{\text{train}} = \{(\boldsymbol{\Theta}\mathbf{y}_1, \mathbf{x}_1), (\boldsymbol{\Theta}\mathbf{y}_2, \mathbf{x}_2), \cdots, (\boldsymbol{\Theta}\mathbf{y}_l, \mathbf{x}_l)\}$ denotes the training set where l is the number of pairs of input spectra and their corresponding inputs of the CNNs structure for the training. $\boldsymbol{D}_{\text{validation}} = \{(\boldsymbol{\Theta}\mathbf{y}_1, \mathbf{x}_1), (\boldsymbol{\Theta}\mathbf{y}_2, \mathbf{x}_2), \cdots, (\boldsymbol{\Theta}\mathbf{y}_v, \mathbf{x}_v)\}$ is the validation set where v is the number of pairs for the validation. $\boldsymbol{D}_{\text{test}} = \{(\boldsymbol{\Theta}\mathbf{y}_1, \mathbf{x}_1), (\boldsymbol{\Theta}\mathbf{y}_2, \mathbf{x}_2), \cdots, (\boldsymbol{\Theta}\mathbf{y}_t, \mathbf{x}_t)\}$ is the test set and t is the number of pairs for the test. Using the 14400 pairs of training set, we train the CNNs structure. As a loss function, we exploit the mean squared error (MSE) over the training set $\boldsymbol{D}_{\text{train}}$. The MSE is defined as:

$$MSE = \frac{1}{l}\sum_{k=1}^{l}\left\|\mathbf{x}_l - \hat{\mathbf{x}}_l\right\|_2^2. \tag{4}$$

Applying the backpropagation, the CNNs structure is trained in the direction of reducing the MSE. To predict the performance of the CNNs structure and determine the stopping point of the backpropagation, we exploit the validation set. We assess the performance of our fully trained framework using the test set.

Figure 3 shows the training and validation loss of our CNNs structure compared with the result using CS reconstruction algorithm[17]. We use the DCT matrix and the Haar wavelet matrix as sparsifying basis. As the epoch increases, training loss decreases and the validation loss converges to the MSE of 0.00078. We evaluate the MSE of the reconstruction using CS reconstruction algorithm over the test set. The MSE of CS reconstruction algorithm using the DCT matrix and the Haar wavelet matrix is 0.00276 and 1.30318, respectively. In the case of the Haar wavelet matrix, the reconstructions are failed in most of the test spectra. We achieve the optimized structure through the training. The MSE of the trained structure over the test set is 0.00078.

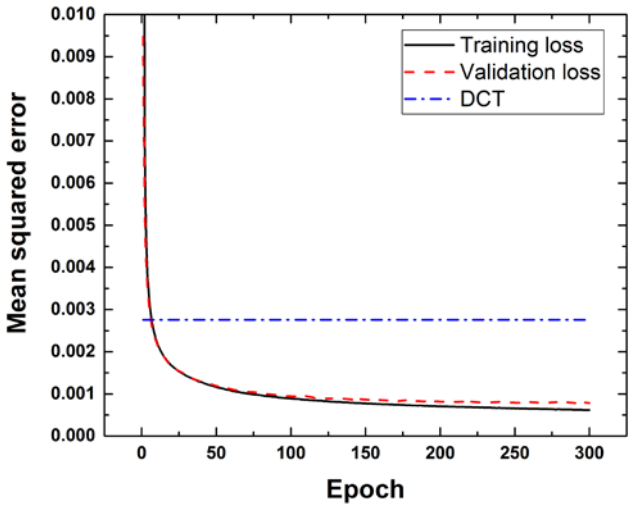

Figure 3. The training and validation loss of our CNNs structure.

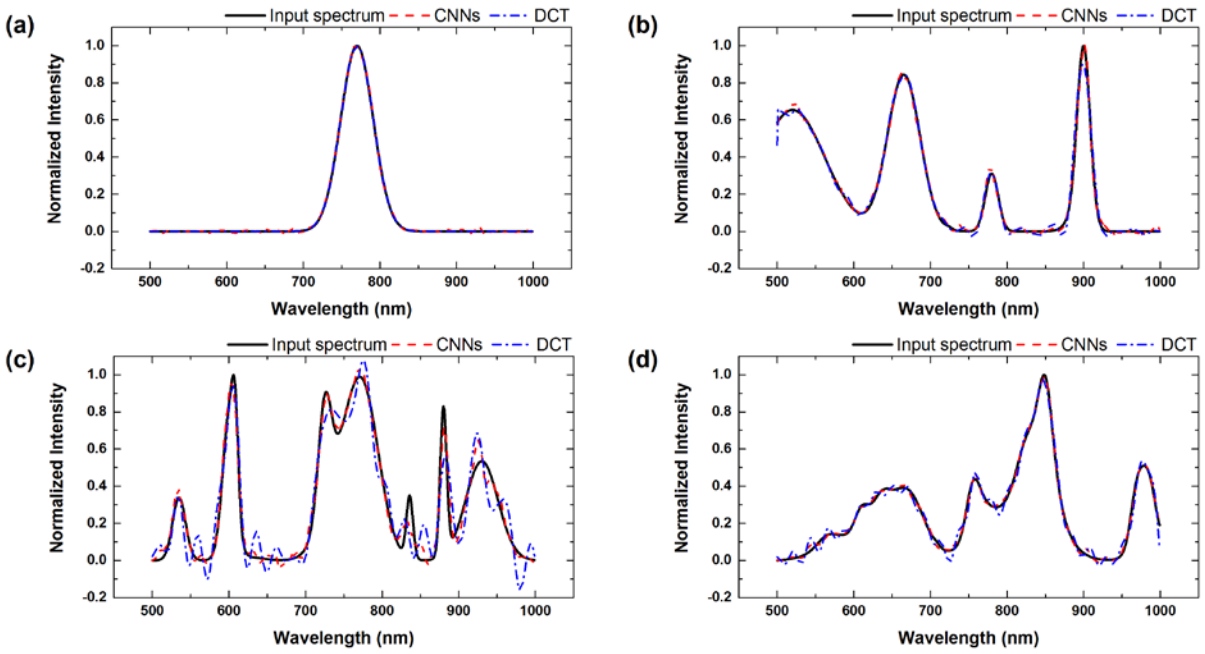

Figure 4. Reconstruction results of the test spectra: (a) The test spectrum generated with one Gaussian kernel, (b) The test spectrum generated with seven Gaussian kernels, (c) The test spectrum generated with nine Gaussian kernels, (d) The test spectrum generated with 18 Gaussian kernels.

Figure 4 shows the reconstruction results of four types of the test spectra. To generate the input spectrum, we use one, seven, nine, and 18 Gaussian kernels in Figs. 4(a-d), respectively. The reconstructions of spectra using the proposed CNNs based framework (red dash) are compared with original input spectra (black solid) and reconstructions using CS reconstruction algorithm with DCT sparsifying basis (blue dash-dot). The reconstruction peak signal-to-noise ratio of the proposed framework is 36.2 dB, 32.3 dB, 25.9 dB and 35.6 dB for Figs. 4(a-d), respectively. The reconstruction peak signal-to-noise ratio of CS reconstruction algorithm using the DCT sparsifying basis is 63.3 dB, 24.4 dB, 17.7 dB and 28.2 dB for Figs. 4(a-d), respectively. The proposed framework shows compatible performance with CS reconstruction algorithm. In addition, the proposed framework shows stable performance over the test data. However, CS reconstruction algorithm using the DCT sparsifying basis has significant a performance change over the type of spectrum.

4. CONCLUSIONS

In this paper, we propose a simple framework based on the CNNs structure that solving inverse problems in CS spectrometers. We train the framework to reconstruct the input spectrum of CS spectrometer. We investigate the reconstruction performance of the CNNs structure comparing with the CS reconstruction algorithm with different sparsifying basis. The MSE of CS reconstruction algorithm using the DCT matrix and the Haar wavelet matrix is 0.00276 and 1.30318, respectively. The MSE of the proposed framework is 0.00078. The MSE results indicate the reconstruction performance of the CNN structure is compatible with the CS reconstruction algorithm. In addition, the proposed framework shows stable performance over the various spectra.

5. ACKNOWLEDGEMENTS

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIP) (NRF- 2018R1A2A1A19018665).

REFERENCES

[1] Kim, S., Cho, D., Kim, J., Kim, M., Youn, S., Jang, J. E., Je, M., Lee, D. H., Lee, B., Farkas, D. L. and others., "Smartphone-based multispectral imaging: system development and potential for mobile skin diagnosis," Biomedical optics express **7**(12), 5294–5307 (2016).

[2] Bacon, C. P., Mattley, Y. and DeFrece, R., "Miniature spectroscopic instrumentation: applications to biology and chemistry," Review of Scientific instruments **75**(1), 1–16 (2004).

[3] Donoho, D. L., "Compressed sensing," IEEE Transactions on information theory **52**(4), 1289–1306 (2006).

[4] Tropp, J. A. and Gilbert, A. C., "Signal recovery from random measurements via orthogonal matching pursuit," IEEE Transactions on information theory **53**(12), 4655–4666 (2007).

[5] Chen, S. S., Donoho, D. L. and Saunders, M. A., "Atomic decomposition by basis pursuit," SIAM review **43**(1), 129–159 (2001).

[6] Palangi, H., Ward, R. K. and Deng, L., "Distributed Compressive Sensing: A Deep Learning Approach.," IEEE Trans. Signal Processing **64**(17), 4504–4518 (2016).

[7] Kulkarni, K., Lohit, S., Turaga, P., Kerviche, R. and Ashok, A., "Reconnet: Non-iterative reconstruction of images from compressively sensed measurements," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 449–458 (2016).

[8] Mousavi, A. and Baraniuk, R. G., "Learning to invert: Signal recovery via deep convolutional networks," Acoustics, Speech and Signal Processing (ICASSP), 2017 IEEE International Conference on, 2272–2276, IEEE (2017).

[9] Kim, C., Lee, W.-B., Lee, S. K., Lee, Y. T. and Lee, H.-N., "Fabrication of 2D thin-film filter-array for compressive sensing spectroscopy," Optics and Lasers in Engineering **115**, 53–58 (2019).

[10] Huang, E., Ma, Q. and Liu, Z., "Etalon Array Reconstructive Spectrometry," Scientific reports **7** (2017).

[11] Oiknine, Y., August, I., Blumberg, D. G. and Stern, A., "Compressive sensing resonator spectroscopy," Optics letters **42**(1), 25–28 (2017).

[12] LeCun, Y., Bottou, L., Bengio, Y. and Haffner, P., "Gradient-based learning applied to document recognition," Proceedings of the IEEE **86**(11), 2278–2324 (1998).

[13] Krizhevsky, A., Sutskever, I. and Hinton, G. E., "Imagenet classification with deep convolutional neural networks," Advances in neural information processing systems, 1097–1105 (2012).

[14] Dong, C., Loy, C. C., He, K. and Tang, X., "Image super-resolution using deep convolutional networks," IEEE transactions on pattern analysis and machine intelligence **38**(2), 295–307 (2016).

[15] Chen, L.-C., Papandreou, G., Kokkinos, I., Murphy, K. and Yuille, A. L., "Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs," IEEE transactions on pattern analysis and machine intelligence **40**(4), 834–848 (2018).

[16] Macleod, H. A., [Thin-film optical filters], CRC press (2001).

[17] Koh, K., Kim, S.-J. and Boyd, S., "An interior-point method for large-scale l1-regularized logistic regression," Journal of Machine learning research **8**(Jul), 1519–1555 (2007).