

OPTICAL COMPUTATION OF CHEMOMETRICS PROJECTIONS USING A DIGITAL MICROMIRROR DEVICE

Purvesh Karkamkar, Ricardo Gutierrez-Osuna

Texas A&M University, College Station, Texas 77840, USA

{kpurvesh, rgutier}@tamu.edu

ABSTRACT

We present an approach to perform chemometrics computations optically using a digital micro-mirror device (DMD) coupled with a near infrared source and diffraction grating. The DMD can be configured to project specific wavelengths onto (or away from) a detector. We show how to program the DMD to perform linear operations (i.e., inner products) on chemical spectra, thus avoiding the need for additional processing. We validate the approach on two chemical discrimination problems using linear discriminant analysis (LDA). Our results show that computing the LDA projections optically yields higher signal-to-noise ratio, and therefore improved discrimination, than performing it digitally (e.g., after scanning the full spectra). The approach is general and can be applied to other chemometrics problems beyond discrimination.

Index Terms— multivariate optical computing (MOC), NIR spectroscopy, digital micromirror device (DMD), integrated sensing and processing (ISP)

1. INTRODUCTION

A number of chemometrics techniques involve computing the linear projection of chemical spectra onto a handful of eigenvectors. Examples of such techniques include partial least squares, principal components regression, and linear discriminant analysis. Conventionally, one measures the full spectrum and then computes the projections in software. An alternative, first suggested by Bialkowski [1], is to compute these projections optically at acquisition time. Doing so avoids having to record the full spectrum in the first place, and can also increase the signal-to-noise ratio (SNR), a phenomenon known as Fellgett's advantage [2].

One approach to performing the computation optically is to construct an optical filter, also known as a multivariate optical element (MOE), to mimic the projection vectors. MOEs were originally proposed by Dobrowolski et al. [3] for the monitoring of chemical reactions, and were later demonstrated by Nelson et al. [4] and Soyemi et al. [5] for predictive spectroscopy. Depending on the complexity of the filter function, however, fabricating an MOE may be

difficult. In addition, the resulting filter is static and cannot be adapted to compute a different filter function. To overcome these limitations, approaches based on spatial light modulators (SLMs) as variable MOEs have been suggested. For example, Uzunbajakava et al. [6] and Davis et al. [7] used liquid crystal panels as SLMs to generate variable optical filters. A major drawback of SLMs is that they require linearly polarized light. An alternative design, proposed by Smith et al. [8], is to image the output of a spectrograph onto a digital micro-mirror device (DMD), which then can be used to project (optically) a combination of the wavelengths onto a single-point detector. A low-cost spectrometer based on DMDs has recently become available [9]. The device, known as the DLP NIRscan Nano, illuminates a chemical sample with a near-infrared lamp, and then uses a diffraction grating to disperse light reflected from the sample into different wavelengths across the DMD plane.

The NIRscan software allows the experimenter to select individual wavelengths or combination of wavelengths, but not their relative strength. This paper describes how the NIRscan can be modified to perform multivariate optical projections. Our focus will be on linear discriminant analysis (LDA) but can be trivially extended to other linear projection models (e.g., PCA, PLS). We validate the system experimentally and compare it against performing the projection in software (after measuring the full spectrum). In agreement with Fellgett's advantage, our results show that performing the computation optically improves the SNR of the projections.

The rest of this manuscript is organized as follows. Section 2 describes the NIRscan, the proposed software modifications, and how they enable computation of multivariate projections. Section 3 describes the experimental methods and results on two discrimination problems (alcohol dilutions, oil adulteration). The article concludes with a discussion and directions for future work.

2. METHODS

Our solution is based on the DLP NIRscan Nano developed by Texas Instruments. As illustrated in Fig 1, the device illuminates a chemical sample with broadband near-infrared light (900-1700nm) and passes the reflection through a

slit/collimating lens onto a diffraction grating, which then disperses it onto the DMD. The DMD is an array of micro-mirrors arranged in two dimensions (854×480), each of which can be individually controlled to be in an ON state or an OFF state. When the mirrors are in the ON state, they reflect light onto a single point InGaAs detector; when they are in the OFF state, they reflect light away from it.

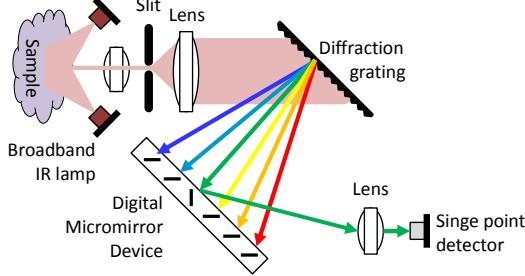


Fig 1. Optical configuration of the NIRscan Nano (adapted from [9]).

The NIRscan is configured such that all the elements in each column (i.e., 480 mirrors) receive light from the same range of wavelengths. If all the elements in a column are configured to the ON state, energy from the corresponding wavelengths is projected onto the detector in full; if m elements are set to ON, the detector receives a fraction (i.e., $m/480$) of the energy. This allows the NIRscan to compute arbitrary linear projections, as long as they are non-negative.

Denote by x the full spectrum of a sample (854 spectral lines in the case of the NIRscan). Assume that we wish to project spectrum x onto eigenvector v ; v may be a principal component of the data, a PLS latent variable or any other multivariate projection. As such, v will likely contain both positive and negative elements. To perform the inner product $x^T v$ optically, we must then decompose v into two non-negative vectors: one containing the positive elements (v^+), the other containing the negative elements (v^-) such that $v = v^+ - v^-$. Thus, computing each multivariate projection requires two separate optical projections.

2.1. Enabling multivariate optical computing

The NIRscan system stores a maximum of 624 patterns in a 32MB SDRAM, each pattern being an 864×480 matrix containing the state of the mirrors in the DMD. Patterns are grouped into 26 frames (24 binary images each) and streamed frame-wise to a DLPC150 controller over a 24-bit RGB bus. The DLPC150 is responsible for setting the mirrors to their desired state based on pattern data. The system architecture is shown in Fig 2(a).

Two types of patterns can be loaded onto the DMD using the default firmware on the NIRscan. The first type of pattern, illustrated in Fig 2(b), turns ON all the mirrors for a particular column. This allows the device to scan the spectrum sequentially, one wavelength at a time. The second type of pattern, known as a Hadamard scan (not shown), directs multiple wavelengths to the detector at any one time; a decoding procedure allows individual wavelengths to be

recovered afterwards. The Hadamard scan collects more light on the detector than a column scan, and therefore increases SNR (Fellgett's advantage).

Neither type of pattern, however, can be used to compute multivariate projections optically, as was described in the previous section. For that to happen, one must be able to select an arbitrary number of mirrors on each column, as illustrated in Fig 2(c). Accordingly, we added a function to the original firmware that accepts an array of column identifiers along with the number of ON state mirrors in each column, as parameters. This function returns an 864×480 binary matrix with the corresponding bits set to 1, and is stored as a pattern in the SDRAM. This enables us to compute optical projections. The modified firmware (including documentation on how to install it onto the NIRscan) will be made available to the research community on github once we are able to de-anonymize the submission.

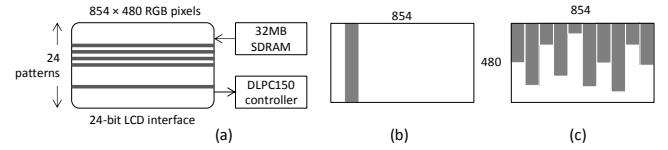


Fig 2. (a) System architecture; (b) scan pattern in the original system; (c) pattern in the modified system.

3. EXPERIMENTAL

We evaluated our optical computing approach on two different classification problems. The first problem consisted of discriminating water from isopropyl alcohol (IPA) at five aqueous dilutions: 100% (i.e., undiluted IPA), 50%, 25%, 10%, and 1%. For each dilution level, we created 5 replicate vials, and then measured the NIR spectrum of each vial 10 times, for a total of 50 sampled spectra. An additional 50 NIR spectra were measured from water. Next, 5-fold cross validation was performed on the resulting dataset. All the replicates from each vial were assigned to either the training set or to the test set, i.e., a different vial was selected for the test set in each fold. Eigenvectors obtained from LDA on training data were used to project the test data both digitally and optically, giving two sets of measurements. A k-nearest-neighbors ($k=1$) classifier was used on the projected data for obtaining classification rates for the binary classification problem of water vs IPA at each dilution level. The results are reported in Table 1. For 1% aqueous dilution of IPA, we can see a remarkable difference in classification rates between the two methods, indicating that our approach performs better.

Table 1. Classification rate for different dilutions of isopropyl alcohol. Number in parenthesis represents standard deviation.

Concentration level	Classification rate (%)	
	Digitally	Optically
100%	100 (± 0)	100 (± 0)

50%	100 (± 0)	100 (± 0)
25%	85 (± 13)	100 (± 0)
10%	83 (± 9)	96 (± 9)
1%	62 (± 36)	92 (± 18)

Fig 3 shows the LDA projections computed digitally (top row), and optically (bottom row) for the five different dilution levels of IPA. As we lower the concentration levels (from right to left), the overlap between the two classes increases at a faster rate for digitally computed projections as compared to optical projections. Especially, at 1% concentration of IPA we can see a significant difference in overlap between the two distributions, with the optical projection showing significantly higher class separability. This result is consistent with the classification rates reported in Table 1.

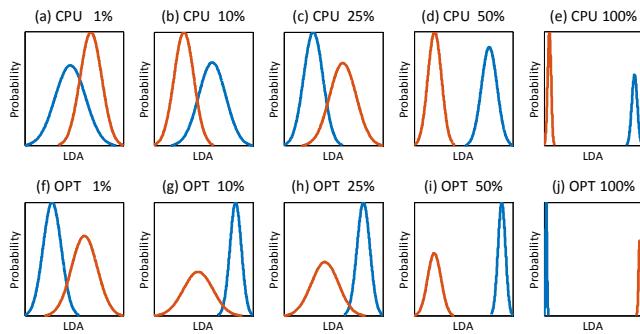


Fig 3. LDA projections for IPA dilutions, computed digitally (a-e) and optically (f-j). Red: IPA; blue: water.

For the second experiment, we considered the problem consisted of detecting adulteration of olive oil. Substitution of high grade extra-virgin olive oil with cheaper oils has commercial and health implications, and is an important problem plaguing the food industry. We simulate the scenario by mixing different proportions of inexpensive sunflower oil with extra-virgin olive oil. We repeated the same procedure as in the previous scenario for generating the datasets, and measuring classification rates for the binary classification problem of pure vs adulterated olive oil at different adulteration levels. Results are summarized in Table 2. As before, computing the LDA projections optically leads to higher classification rates with kNN than computing them digitally.

Table 2. Classification rates at different adulteration levels of extra virgin olive oil.

Adulteration level	Classification rate (%)	
	Digitally	Optically
50%	92 (± 13)	100 (± 0)
25%	87 (± 18)	96 (± 5)

Fig 4 shows the LDA projections computed digitally (top row), and optically (bottom row) for two adulteration levels of olive oil. Again, we can see a similar difference in overlap for the two methods, optical computation providing

better separability compared to computing projections digitally. This is again consistent with the classification rates in Table 2.

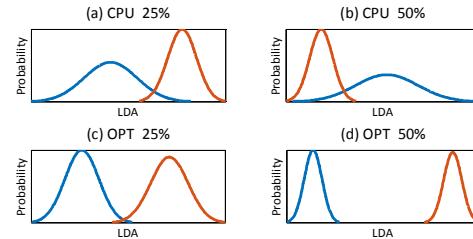


Fig 4. LDA projections for olive oil at two adulteration levels, computed digitally (a-b) and optically (c-d). Red: pure olive oil; blue: adulterated olive oil.

4. CONCLUSION

We have evaluated using a low-cost spectrometer device (NIRscan Nano) for multivariate optical computation. We have presented a modification to the NIRscan firmware code for computing arbitrary linear projections. We have experimentally validated our approach on two classification problems. Our results show that computing the multivariate projections optically leads to higher SNRs compared to computing the same projections digitally. Future work will explore using the modified NIRscan system to other applications and chemometrics problems.

5. REFERENCES

- [1] S. Bialkowski, "Species discrimination and quantitative estimation using incoherent linear optical signal processing of emission signals," *Anal. Chem.*, vol. 58, no. 12, pp. 2561-2563, 1986.
- [2] M. Farrow *et al.*, "Fourier transform photoacoustic spectroscopy," *Appl. Phys. Lett.*, vol. 33, no. 8, pp. 735-737, 1978.
- [3] J. A. Dobrowolski *et al.*, "Design of thin-film filters for the monitoring of chemical reactions," in *Optical Science, Engineering and Instrumentation'97*, 1997, pp. 38-45: SPIE.
- [4] M. Nelson *et al.*, "Multivariate optical computation for predictive spectroscopy," *Anal. Chem.*, vol. 70, no. 1, pp. 73-82, 1998.
- [5] O. Soyemi *et al.*, "Design and testing of a multivariate optical element: The first demonstration of multivariate optical computing for predictive spectroscopy," *Anal. Chem.*, vol. 73, no. 6, pp. 1069-1079, 2001.
- [6] N. Uzunbajakava *et al.*, "Low-cost spectroscopy with a variable multivariate optical element," *Anal. Chem.*, vol. 78, no. 20, pp. 7302-7308, 2006.
- [7] B. M. Davis *et al.*, "Multivariate hyperspectral Raman imaging using compressive detection," *Anal. Chem.*, vol. 83, no. 13, pp. 5086-5092, 2011.
- [8] Z. Smith *et al.*, "Multivariate optical computing using a digital micromirror device for fluorescence and Raman spectroscopy," *Opt. Express*, vol. 19, no. 18, pp. 16950-16962, 2011.
- [9] Texas Instruments DLP NIRscan Nano Evaluation Module. Available: <http://www.ti.com/tool/DLPNIRNANOEV>