

Inverse-Model Space for Feature Extraction from Absorption Spectra

SAMUEL G. LAMBRAKOS

ANDREW R. SHABAEV

*Center for Materials Physics & Technology Branch
Materials Science & Technology Division*

April 21, 2022

DISTRIBUTION STATEMENT A: Approved for public release; distribution is unlimited.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

1. REPORT DATE (DD-MM-YYYY) 21-04-2022		2. REPORT TYPE NRL Memorandum Report		3. DATES COVERED (From - To) September 2021 – January 2022	
4. TITLE AND SUBTITLE Inverse-Model Space for Feature Extraction from Absorption Spectra				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S) Samuel G. Lambrakos and Andrew R. Shabaev				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER 1Y64	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Research Laboratory 4555 Overlook Avenue, SW Washington, DC 20375-5320				8. PERFORMING ORGANIZATION REPORT NUMBER NRL/6362/MR--2022/2	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) Office of Naval Research One Liberty Center 875 North Randolph Street, Suite 1425 Arlington, VA 22203-1995				10. SPONSOR / MONITOR'S ACRONYM(S) ONR	
				11. SPONSOR / MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION / AVAILABILITY STATEMENT DISTRIBUTION STATEMENT A: Approved for public release; distribution is unlimited.					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT This report describes a methodology for spectrum-feature extraction from absorption spectra, obtained using reflectance measurements, including those obtained by diffuse reflectance (DR) and attenuated total reflectance (ATR) spectroscopies, based on reflectance theory and phenomenological function decomposition of reflectance. Specifically, this methodology entails spectrum feature-extraction using iterative spectrum adjustment by phenomenological backgrounds. Formulation of the feature-extraction methodology, based on inverse-analysis theory, is that of an inverse-model space, whose concept and properties are described. The inverse-model space, for spectrum feature-extraction, consists of numerical and analytical functions for non-unique iterative spectrum-background adjustment, which formally combines existing methods. In this study, the methodology defined by these combined methods, is extended to include truncated basis-function representations of dominant spectral features. In addition, results of spectrum-feature extractions demonstrating application of the inverse-analysis methodology are described.					
15. SUBJECT TERMS Spectrum-feature extraction Absorption spectra Inverse analysis					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Andrew R. Shabaev
U	U	U	U	19	19b. TELEPHONE NUMBER (include area code) (202) 767-2619

This page intentionally left blank.

Table of Contents

1. Introduction.....	1
2. Inverse Model Space for Spectrum-Feature Extraction.....	2
3. Case Studies of Spectrum-Feature Extraction.....	6
4. Discussion.....	8
5. Conclusion.....	14
6. Acknowledgement.....	14
7. References.....	15

This page intentionally left blank.

1. INTRODUCTION

Extraction of spectral features associated with target materials within chemical and physical mixtures, including materials on surfaces having complex structure, for their identification and quantification, poses a specific problem for spectroscopic analysis. Parametric models of reflectance and absorption spectra can be applied for extraction of spectral features of target materials that are within a background of spectral features associated with mixtures and substrates, as well as scattering due to finite particle-size influences. Extraction of spectral features of target materials can be by means of signal-corrected spectra, where broad-background spectral features have been removed. These signal-corrected spectra can provide, in principle, extracted or residual spectra to be compared with reference or template spectra, for purposes of detection. This study describes a methodology for spectrum-feature extraction from absorption spectra, typically obtained by reflectance measurements, including those obtained by diffuse reflectance (DR) and attenuated total reflectance (ATR) spectroscopies, based on reflectance theory and phenomenological function decomposition of reflectance.

The model formulation is based on the inverse analysis approach [1,2], reflectance theory [3-7] and phenomenological scattering-matrix (S-matrix) representation of reflectance, consisting of multiplicative factors associated with different types of contributions to the reflectance. Among these factors are phenomenological reflectance functions representing adjustable parameters of the model formulation, whose adjustment is similar to that of spectrum normalization and spectrum-baseline correction with respect to phenomenological backgrounds, e.g., references [8-16], and signal-processing methods using signal templates [17].

There exists a large volume of work describing different spectrum-correction methods based on background removal. A proper review of this work poses a major task in itself, beyond the scope of this report. Reference [16], however, provides a reasonable review of such methods, and in addition, specifies the category of the methodology described here, which is based on non-unique iterative and segmented adjustment of spectra, which is in contrast to methods based on those of references 8-11. The inverse analysis methodology presented here, for spectral-feature extraction, follows from the observation that many spectrum-correction methods are formally equivalent, but defined in terms of different parameterizations. Specifically, these methods, which are for determination of signal-corrected spectra by adjustment of different contributions to absorbance spectra, are equivalent to phenomenological multiplicative-factor decomposition of the reflectance function. Accordingly, these methods, which can be applied in combination and iteratively for feature extraction, can be defined formally an “inverse-model space” as described by inverse problem theory [1,2], where model parameterizations are not unique.

Inverse-Model and Spectrum-Template Spaces. Following the inverse-problem approach [1,2], a system is represented by a parametric model. The particular choice of a parametric model, is termed a “parameterization” of the system [2]. The choice of a particular parameterization, however, is in general not unique. In order to address the property of non-uniqueness of system parameterization, inverse problem theory has adopted the concept of “model space,” where subregions of this space represent “conceivable” parametric models of the system [2]. Given a model space for diffuse reflectance, quantitative inverse analysis is further enhanced by isolating model-space regions associated with parameterizations that are physically consistent, conveniently adjusted with respect to given sets of measurements, and sufficiently general in term mathematical representation. A physically consistent, conveniently adjustable and sufficiently general parameterization of diffuse reflectance is significant for the following reasons. First, spectral-signature features calculated by

inverse methods represent a mapping from data space into parameter space. It is therefore preferable to adopt a parametric function representation whose form tends to minimize any bias resulting from its mathematical form. Second, a set of parameters associated with a physically consistent representation can in principle be used to extract relationships between parameters, which can provide further insight related to physical characteristics. Third, spectral-feature extraction for a specific application requires quantitative assessment of reflectance characteristics over a sufficient range of parameter values. Fourth, a sufficiently general parametric model formulation can be adjusted to include influences due to incomplete information concerning the system. And fifth, the “spectrum-data space” is that containing measurements of transmission and reflectance spectra, and any quantity related to these spectra via an assumed physical relationship. Finally, a general model formulation that can be applied, in principle, for inverse analysis of reflectance spectra, for a wide range of materials on substrates, can be a conveniently adjustable set of parametric functions.

An inverse-analysis framework for spectrum-feature extraction based on the concept of model space implies that feature extraction is via non-unique iterative operations. Feature extraction can be based on comparison to an ensemble of template spectra, which defines a “template space.” The concept of template space is similar to that of feature space, which is applied for image analysis. Shown is that previous feature-extraction methods, as a set, have a well defined interrelationship that defines a parametric model space.

In what follows, an inverse model space for spectrum feature-extraction is presented. Next, case studies concerning spectrum-feature extraction from IR absorbance spectra for a variety of coffees are described. Next, a discussion is given concerning aspects and utilization of the inverse model space for spectrum-feature extraction. Finally, a conclusion is presented.

2. INVERSE MODEL SPACE FOR SPECTRUM-FEATURE EXTRACTION

Following the inverse analysis approach, a physically consistent parametric-model space provides a means for investigating the dielectric response characteristics of a given material, or combination of materials, and geometry. The model space is constructed using combinations of physical and phenomenological representations. For reflection from sparse distributions of particles (or in general, segments), whose material composition is heterogeneous, it is plausible that a general framework for modeling of reflectivity spectra can be based on background-subtraction procedures, which are for removal of spectral features due to base substrates, materials known to be present (other than those associated with target spectra), and unknown contributions to spectral features. An inverse-analysis methodology for extraction of target spectral features, which combines trend features of reflection measurements, e.g., diffuse and ATR, with background-normalization procedures, and is based on a phenomenological multiplicative-factor (or S-matrix) representation of reflectance [7], is given by

$$R_M(\omega) = R_r(\omega)R_p(\omega)R_b(\omega)R_{LC}(\omega) \quad (1)$$

where

$$R_{LC}(\omega) = \sum_{l=1}^N w_l \exp \left[- \sum_{n=1}^{M_l} c_{ln} A_{ln}(\omega) \right] \quad (2)$$

and

$$A_r(\omega) = \ln \left(\frac{1}{R_r(\omega)} \right). \quad (3)$$

The component decomposition defined by Eq.(1) provides for inverse analysis of spectra. The reflectance function $R_b(\omega)$ represents measured or modeled background contributions to the measured spectrum $R_M(\omega)$, e.g., spectrum of a substrate. The quantity $R_r(\omega)$ represents contributions to $R_M(\omega)$ that do not have broad-band variation, and assumed to contain spectral information of interest. Characteristically, this quantity is representative of spectral features not to be filtered, but more importantly, extracted. Equation (1) can be extended for inverse analysis by inclusion of components that are physical models based on theory and absorption coefficients for known materials. The function R_{LC} represents estimated known contributions to the measured spectrum associated with mixtures and surface segmentation. The normalized absorbance functions $A_{ln}(\omega)$ are determined in principle by inverse spectral analysis of reflectance or transmission spectra using the Beer-Lambert model, or of diffuse reflectance using the Kubelka-Munk model [6], where scaling with respect to path length and segmentation are represented parametrically by weight coefficients c_{ln} and w_l , respectively.

Following the general formulations of methods for spectrum normalization and spectrum-baseline correction with respect to phenomenological backgrounds [8-16], the function $R_p(\omega)$ can be decomposed phenomenologically in terms of absorbance basis functions and characteristic reflectance components, representing different types of spectral-feature characteristics, which are in terms of either analytical parametric models or numerical procedures. Accordingly,

$$R_p(\omega) = \exp(-A_f(\omega)) \langle R_{BN}(\omega) \rangle_S \quad (4)$$

where

$$A_f(\omega) = A_c + B_c \langle A_m(\omega) \rangle + \sum_{l=1}^N c_l \omega^l + F_{an}(\omega) \quad (5)$$

and

$$R_{BN}(\omega) = \frac{R_M(\omega)}{R_b(\omega)R_{LC}(\omega)}. \quad (6)$$

Equations (1)-(6) represent a phenomenological decomposition in terms of a product of reflectance functions, characterized by broadband variation, which are for iterative normalization. The quantity $\langle A_m(\omega) \rangle$ represents an average spectrum that is adopted as a reference spectrum. The quantity A_c represents additive baseline contributions to the spectrum. The coefficients B_c and c_l ($l = 1, \dots, N$) are a scaling factor of the reference spectrum, and those associated with phenomenological polynomial-function representation of broad baseline variations within spectra, respectively.

The quantity F_{an} is for analytical-function representation of spectral features that are a function of wavelength, broadband and have trend features with line-shapes. This quantity, which is included for completeness, includes within the model space phenomenological spectrum-fitting models that do not have simple functional forms, such as constant, linear and polynomial. For example, in principle, F_{an} could be defined in terms of Gaussian, Lorentzian and pseudo-Voigt functions, which are higher-order extension of simpler functional forms.

The multiplicative factor $\langle R_{BN}(\omega) \rangle_S$ represents the numerical operation of spectrum smoothing. The scaled average quantity $\langle A_m(\omega) \rangle$ can also be the result of spectrum smoothing, depending on the chosen sequence of model-space parameter adjustment. Reference [16] provides some discussion of spectrum-smoothing methods. The adjustable quantities of the model-space formulation, which are defined by the decomposition of $R_p(\omega)$ (Eqs. (4) and (5)) are not unique, and can be defined in terms of either analytical parametric models or numerical procedures.

The formal procedure underlying inverse analysis entails general concepts based successive normalization of spectra for determination of a residual spectrum $A_r(\omega)$ defined below. The inverse model space adopts as input information, the measured reflectance spectrum R_M of particles (or segments) distributed on a substrate, the reflectance spectrum R_b of substrate, and reflectance contributions of materials assumed known to be present R_{LC} , and an average reference spectrum $\langle A_m(\omega) \rangle$. As is characteristic of inverse models, the decompositions defined by Eqs.(1)-(6) are not unique. Different basis-function decompositions can be applied to the same measured spectrum R_M , which would be based on assumptions concerning the physical nature of spectral features. Interpretation of the general quantities comprising Eqs.(1)-(6) can be obtained by elucidating their relationship to spectrum-correction methods e.g., [8-16]. These methods, which are for determination of a signal-corrected absorbance spectrum by adjustment of different component contributions, are formally equivalent to a phenomenological multiplicative-factor decomposition of the reflectance function, i.e., a phenomenological S-matrix representation defined by Eq.(1), which entails adjustment of different “multiplicative-factor” contributions to the reflectance spectrum.

The general procedure for inverse analysis entails construction of a residual spectrum $A_r(\omega)$ according to the general relation

$$A_r(\omega) = A_{BN}(\omega) - A_p(\omega), \quad (7)$$

The absorbances A_{BN} and A_p , respectively, are given by

$$A_{BN}(\omega) = \ln\left(\frac{1}{R_{BN}(\omega)}\right) \quad (8)$$

and

$$A_p(\omega) = \ln\left(\frac{1}{R_p(\omega)}\right), \quad (9)$$

and components of $R_p(\omega)$, defined by Eqs.(4)-(6) are adjusted with respect to a template spectrum. The residual spectrum $A_r(\omega)$, which contains dominant spectral features, is calculated by successive adjustment of the component contributions of $A_f(\omega)$ and the smoothed spectrum $\langle R_{BN}(\omega) \rangle_S$, with given $R_M(\omega)$, $R_b(\omega)$ and $R_{LC}(\omega)$, as needed by characteristic trend features of measured spectra. The general formulation of the inverse analysis procedure defined by Eqs. (7)-(8) does not interpret A_r as a residual term associated with least-squares determination of parameters defined by Eqs.(4)-(6). Rather, the residual spectrum is calculated by successive, independent and non-unique adjustment of the component contributions, as needed by the characteristic trend features of the measured spectrum, as compared to those of a template spectrum. This interpretation of A_r provides flexibility with respect to spectrum-feature extraction associated with target materials. Specifically, Eqs. (4)-(6) define a set of independent operations for background normalization, which may be applied partially or in combination for sufficient feature extraction.

Next, noted is that further background removal can be applied to residual spectra $A_r(\omega)$, which contain dominant spectral features. This removal is achieved by constructing truncated basis-function representations of dominant spectral features, i.e.,

$$A_r(\omega) = A_{rt}(\omega) + E_r(\omega) \quad (10)$$

where

$$A_{rt}(\omega) = \sum_{k=1}^{N_k} w_k f_k(\omega) \quad (11)$$

and is a linear combination of Gaussian or Lorentzian functions. Inherent incompleteness of the basis-function set f_k ($k=1-N_k$) is assumed, but sufficiently complete for representation of dominant spectral features. The expansion Eq.(11) is constructed using discrete spectra resulting from sampling at critical points of $A_r(\omega)$, which are then transformed to a continuous-spectrum estimation of absorbance spectra. To be noted is that discrete sampling of spectra at critical points is *not* related to discrete sampling of a continuous spectrum according to a Nyquist criterion, and truncated basis-function representation is *not* related to phenomenological basis-function decomposition of spectra, because they are an encoding of physical resonant structure. To be noted additionally, is the the non-unique sequence of operations defined by Eqs. (1)-(9), is for feature extraction related not only to “peaks” or local maxima, but all critical points associated with function variation,

Equations (1)-(9), not including the general function F_{an} within Eq.(5), define a model subspace that formally combine existing methods (references [8-16]). In this study, the methodology defined by this subspace, is extended to include truncated basis-function representation of dominant spectral features, $A_{rt}(\omega)$ defined by Eqs. (10)-(11).

As described above, the formulation defined by Eqs.(1)-(11) represents an inverse model space for spectrum-feature extraction. The problem of “spectrum-feature extraction” for comparison with spectral features of template spectra (or a spectrum feature-space) should be considered different from that of determining absorbance spectra of materials, by inverse analysis, for calculation of dielectric-response functions. Specifically, the problem of spectrum-feature extraction entails comparison of calculated residual spectra with template spectra associated with target materials. Development of procedures for such comparison defines a separate problem. In principle, for the methodology presented here, template spectra or feature-space spectra can be approximate absorbance spectra that are sufficient in terms of characteristic features for comparison with extracted spectral features.

The inverse model space, Eqs. (1)-(11), defines a set of independent and general feature-extraction operations, which are in terms of reflectance and absorbance, are equivalent representations of spectral features, and can be applied in terms of both. These operations are guided by spectra within spectrum templates or feature spaces. This follows in that spectrum feature-extraction operations are applied to spectra that are assumed to contain spectral features dominated on average by those of a target material, but also contain additional features resulting from various physical process of both known and unknown origin. Again, the general formulation defined by Eqs.(1)-(11) represents an “inverse-model space” for spectrum-feature extraction, as described above. Conceptually, application of the inverse-model space Eqs.(1)-(11) for spectrum-feature extraction represents an iterative process, where sequences of spectrum adjustments are not unique, consistent with non-unique inverse-model parameterization. In addition, it is expected that a given set of spectrum adjustments will be relatively optimal for certain localized regions of frequencies, but not all. Thus, feature extraction over different wavelength regions, using the above methodology, is in general, by means of spectrum adjustments over different segmented wavelength-regions. The concept of comparing extracted features to those of feature space, which includes spectrum templates, is motivated in part by algorithms applied in signal processing [17].

References [18,19] apply the above methodology to surface-distributed caffeine particles on substrates, i.e., aluminum, glass, polyethylene (PE), and gold. Studies described in these references demonstrate adjustment of both reflectance and absorbance spectra for feature extraction. Described

in the next section are only adjustments of absorbance spectra for feature extraction, and extension of the model space for inclusion of truncated basis-function representation.

3. CASE-STUDIES OF SPECTRUM-FEATURE EXTRACTION

Presented in this section are case studies of spectrum-feature extractions from absorbance spectra, demonstrating application of the inverse-model space related to adjustments of absorbance spectra. For these case studies, spectrum-feature extraction is applied to absorbance spectra of different coffee types using the model space defined by Eqs.(1)-(11). This analysis entails calculation of residual absorbance functions, discrete spectra of critical points, and truncated basis-function representations. In principle, operations for calculation of these different types of spectra, for feature extraction from a given spectrum, can be done in combination (sequentially). The present case studies apply these operations individually, for the purpose of demonstration. In addition, for these studies, quantitative correlation or comparison of extracted features with template spectra is not considered, but discussed qualitatively.

Our first set of spectrum-feature extractions considers construction of an ensemble of template spectra for comparison with extracted spectral features. In that our prototype analyses concern types of coffee, template spectra of caffeine are considered as prototypical. Shown in Figs. (1)-(3) is a measured absorbance spectrum for caffeine, its discrete-function approximation, and its truncated basis-function representation, respectively. The significance of other representations, e.g., Figs. (2) and (3) is that of “feature enhancement,” which should be an important characteristic of spectrum templates.

Our next set of spectrum-feature extractions consider absorbance spectra for three different types of coffee [20], i.e., Arabian, Instant and Roasted. Shown in Figs. (4) and (5) is a measured absorbance spectrum for Arabian coffee and a scaled bin-average of this spectrum, respectively. Shown in Fig.(6) is a residual absorbance $A_r(\omega)$ spectrum obtained by subtraction of background spectrum in Fig. 5. Shown in Figs. (7) and (8) are a discrete-function approximation and truncated basis-function representation, respectively, of the absorbance function of Fig.(4). For this spectrum-feature analysis, the significance of different feature extraction procedures, e.g., Figs. (6)-(8), is that of “enhancing” different characteristics of individual features. The same spectrum-feature analysis procedure applied for Arabian coffee, Figs. (4)-(8), are applied for Instant coffee, Figs. (9)-(13), and for Roasted coffee, Figs. (14)-(18).

Our last spectrum-feature extraction demonstrates the use of different phenomenological spectrum backgrounds, e.g., Fig. (19), for calculation of residual absorbance spectra $A_r(\omega)$, e.g., Fig. (20). Shown in Fig.(20) is a residual absorbance $A_r(\omega)$ spectrum for Arabian coffee, obtained by subtraction of background spectrum in Fig. (19), which has been constructed using Gaussian, Lorentzian and linear functions and segmentation of the wavelength range. Again, the significance of different feature extraction procedures, e.g., comparison of Figs. (6) and (20) , is that of enhancing different characteristics of individual features. With respect to this goal, a comparison of residual spectra shown in Figs. (6) and (20) would provide a sensitivity analysis of spectral features with respect to their extraction.

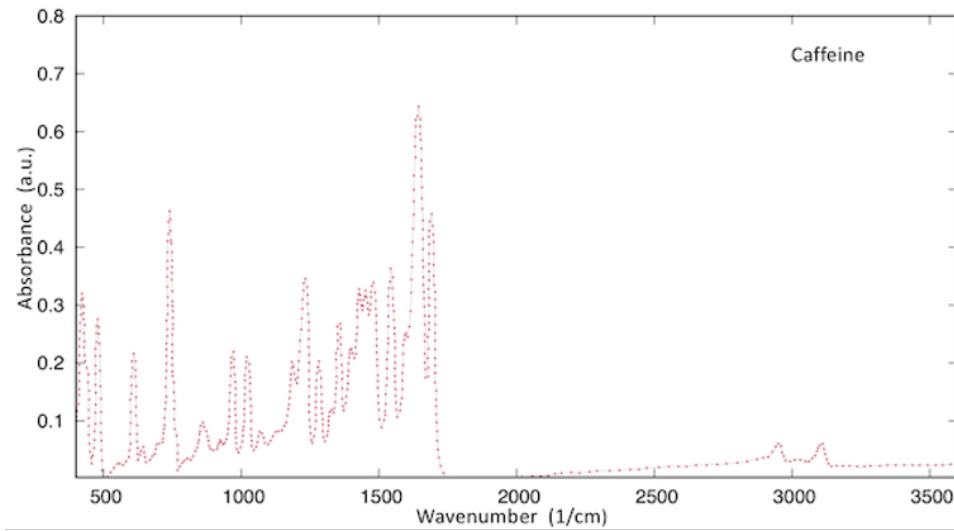


Figure 1. Measured absorbance function of caffeine [20].

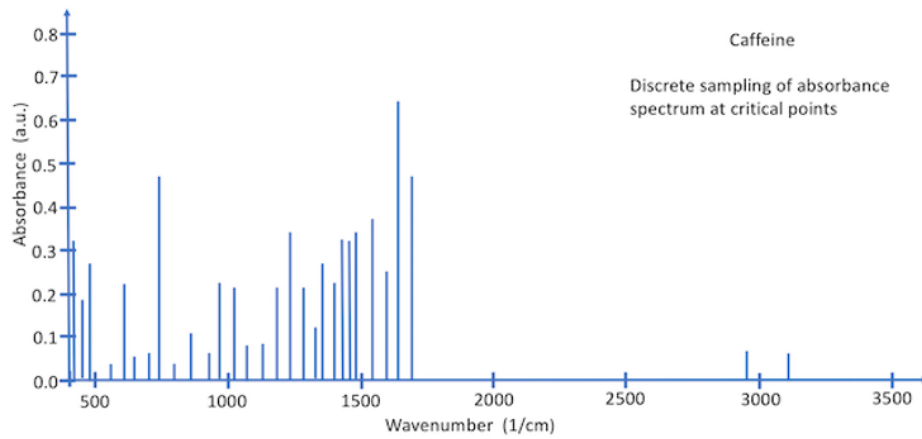


Figure 2. Discrete-function approximation of absorption function (Figure 1) for caffeine.

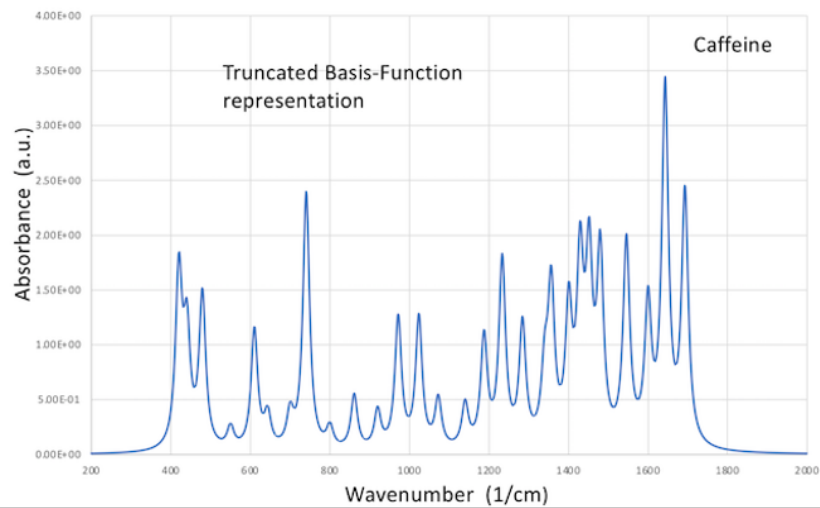


Figure 3. Truncated basis-function representation of caffeine absorbance spectrum constructed using sampled critical points (Figure 2).

Figures (1)-(20) demonstrate a prototype analysis of absorption spectra for different types of coffee, which is motivated by that described in reference [20]. For a realistic analysis, e.g., reference [20], spectral features would be correlated with respect to some type of “quality metric.” The quality metric could be based, in principle, on comparing the spectral features of any given coffee to either caffeine, or those of another coffee, designated as a reference standard. Accordingly, spectrum features of a specific coffee, adopted as a reference standard, would represent a spectrum template

4. DISCUSSION

The inverse-model space Eqs. (1)-(11) defines a set of independent feature-extraction operations, which are in terms of both reflectance and absorbance. These operations are guided by template spectra, which impose constraints on extraction of spectral features. This follows in that these operations are applied to spectra that are assumed to contain spectral features dominated on average by those of adopted templates, but also contain additional features resulting from various physical process of both known and unknown origin. For this study, comparison to a spectrum template was qualitative, and not based on a quantitative procedure. Quantitative procedures for comparison of extracted features to template spectra defines a separate problem.

The operation of spectrum smoothing, represented by the multiplicative factor $\langle R_{BN}(\omega) \rangle_S$, can be defined by various procedures. Application of this factor was demonstrated in other studies [18,19]. For this study, smoothing based on bin-averaging of spectra was applied for determining the scaled average quantity $\langle A_m(\omega) \rangle$. It is significant to note that spectrum-feature extraction using smoothing can be achieved, equivalently, by adjustment of parameters defined by Eq.(5).

Referring to the prototype inverse analyses above, it is seen that calculated residual absorbance spectra show good qualitative agreement with template spectra. Quantitative agreement is not expected in general, in that residual spectra, although containing spectral features dominated by those of adopted spectrum templates, also contain additional features resulting from other processes.

Calculation of residual spectra with respect to template spectra is effected non-uniquely by adjusting components of Eq.(4). For this study, calculation of residual spectra included background subtraction with respect to $\langle A_m(\omega) \rangle$ and phenomenological-baseline functions, including scaling, i.e., A_c defined by Eq.(5). These adjustable quantities are a subset of the parametric model space Eqs.(1)-(11). Procedures for spectrum-feature extraction by adjustment of other components of this model space, in combination with spectrum smoothing, remains for more investigation. In particular, construction of functions $F_{an}(\omega)$ for representation of spectral features that are not conveniently represented by polynomial functions.

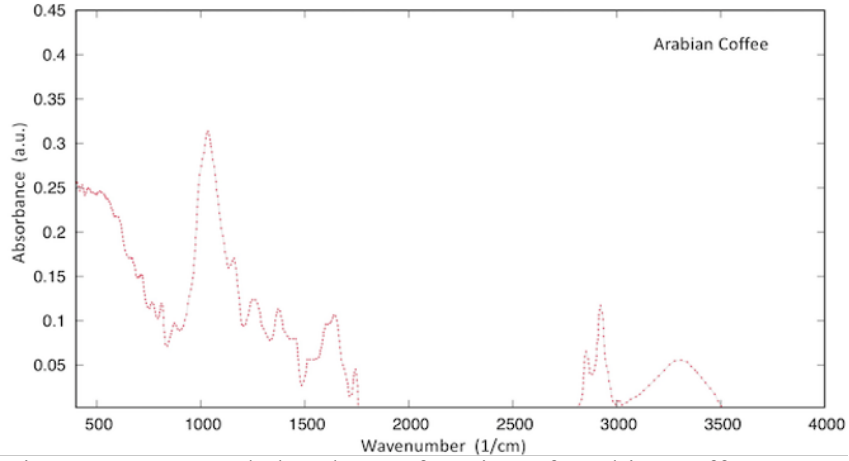


Figure 4. Measured absorbance function of Arabian coffee [20].

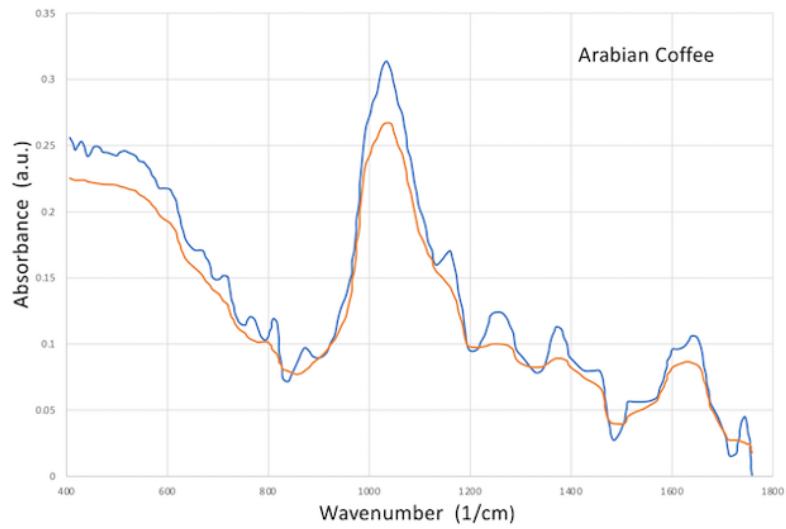


Figure 5. Measured absorbance function of Arabian coffee (blue) and scaled bin-averaged spectrum (red), background spectrum for feacher extraction.

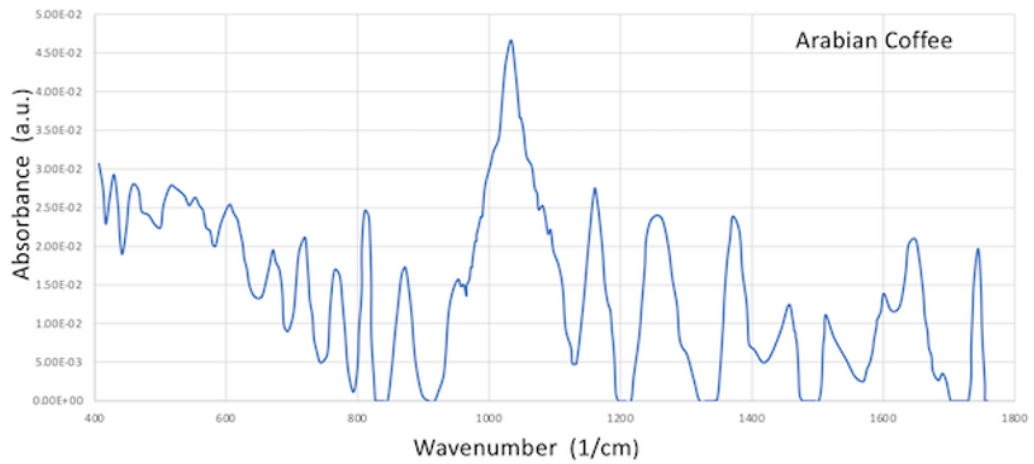


Figure 6. Residual absorbance $A_r(\omega)$ spectrum of Arabian coffee obtained by subtraction of background spectrum (Figure 5).

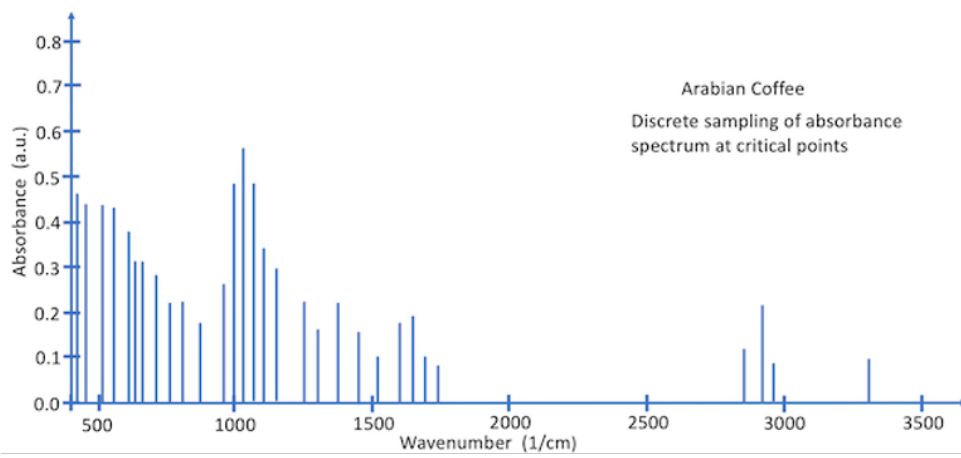


Figure 7. Discrete-function approximation of absorbtion function (Figure 4) for Arabian coffee.

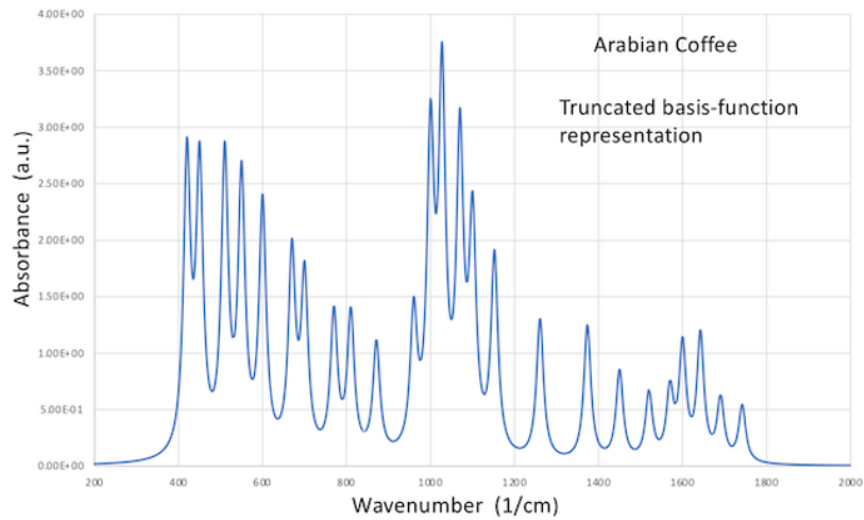


Figure 8. Truncated basis-function representation of caffeine absorbance spectrum constructed using sampled critical points (Figure 7).

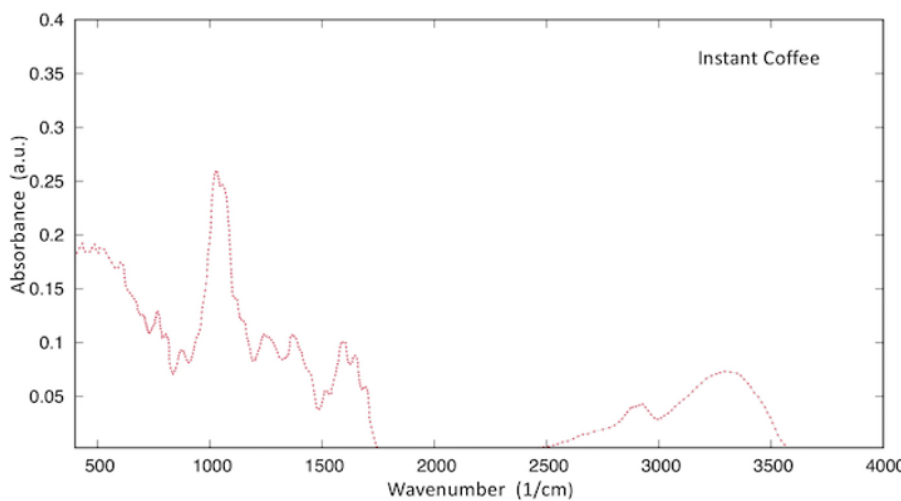


Figure 9. Measured absorbance function of Instant coffee [20].

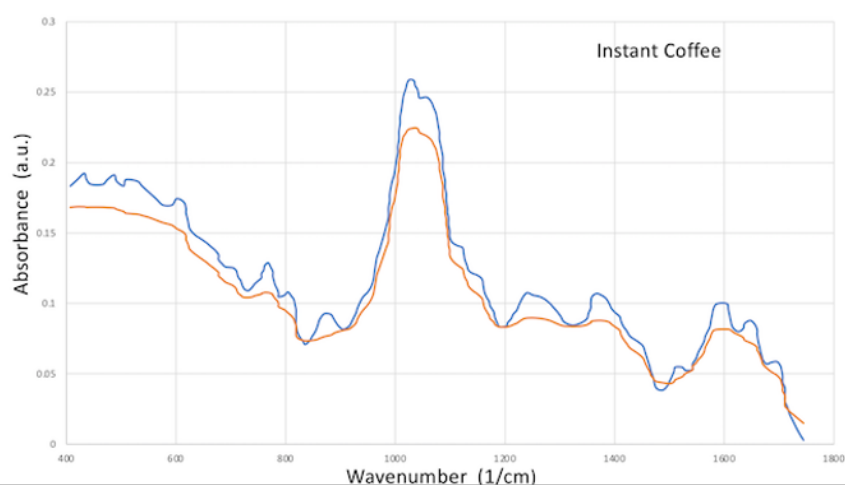


Figure 10. Measured absorbance function of Instant coffee (blue) and scaled bin-averaged spectrum (red), background spectrum for feacher extraction.

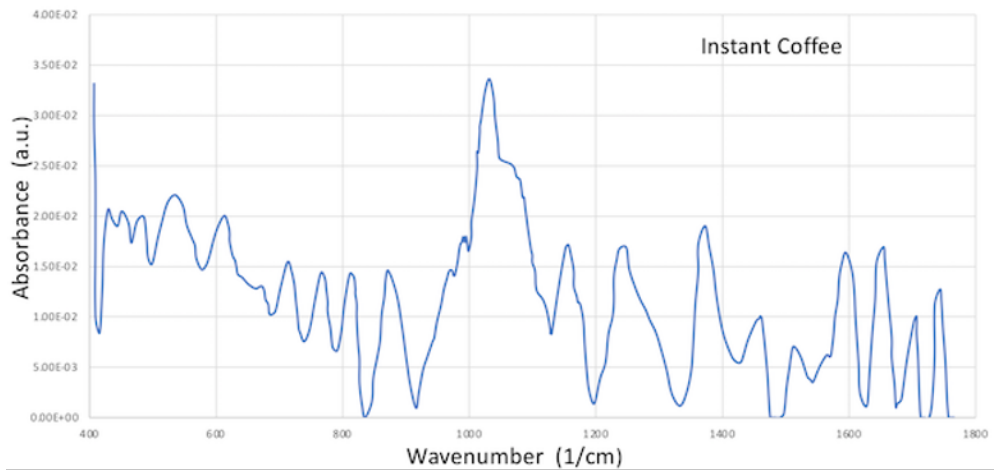


Figure 11. Residual absorbance $A_r(\omega)$ spectrum of Instant coffee obtained by subtraction of background spectrum (Figure 10).

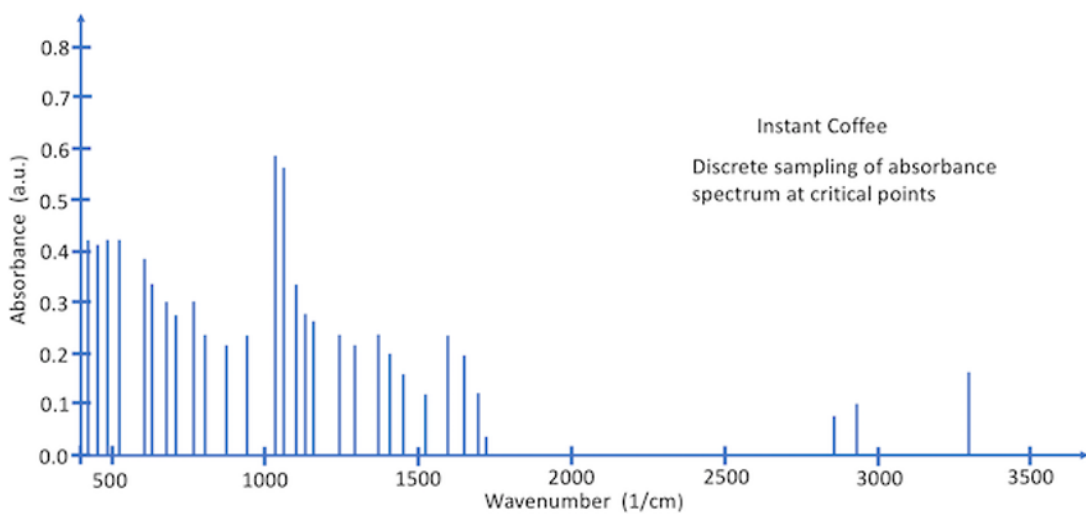


Figure 12. Discrete-function approximation of absorption function (Figure 9) for Instant coffee.

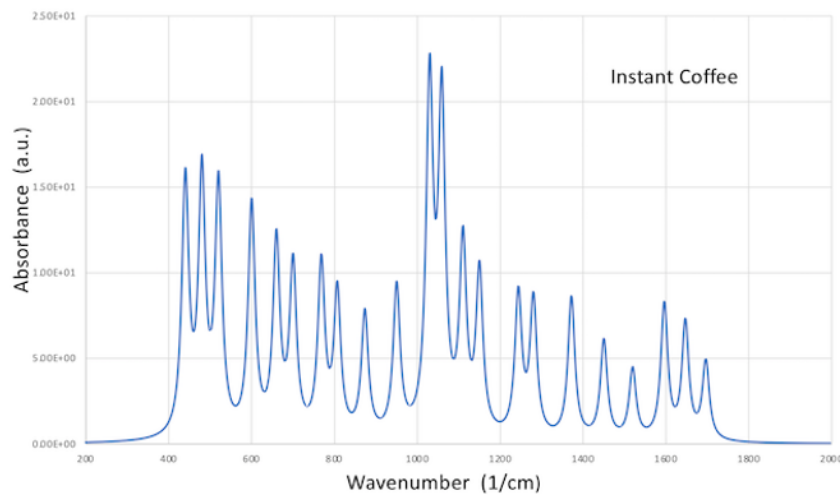


Figure 13. Truncated basis-function representation of caffeine absorbance spectrum constructed using sampled critical points (Figure 12).

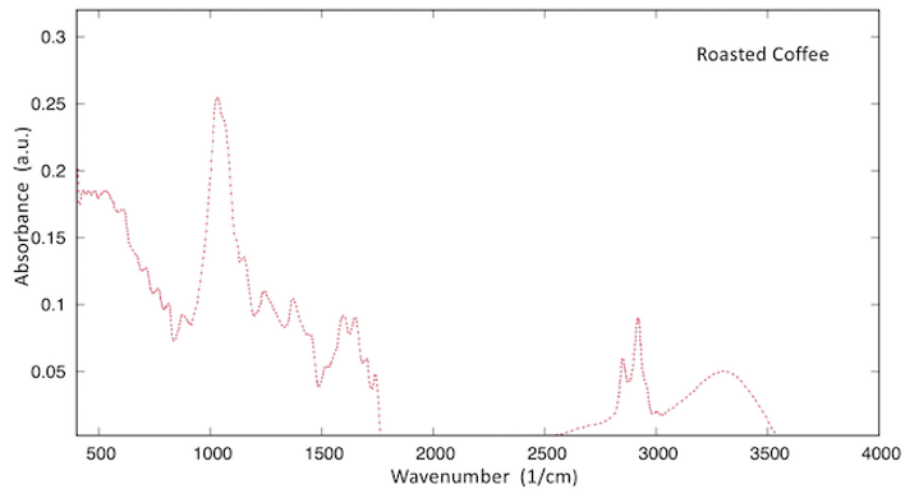


Figure 14. Measured absorbance function of Roasted coffee [20].

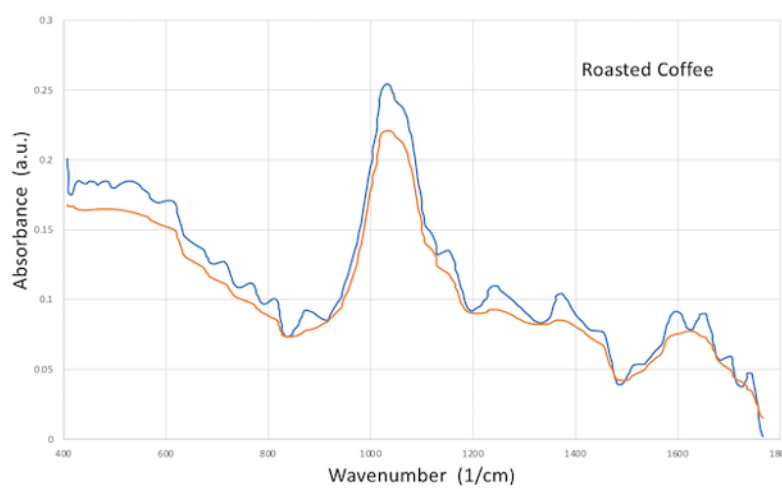


Figure 15. Measured absorbance function of Roasted coffee (blue) and scaled bin-averaged spectrum (red), background spectrum for feacher extraction.

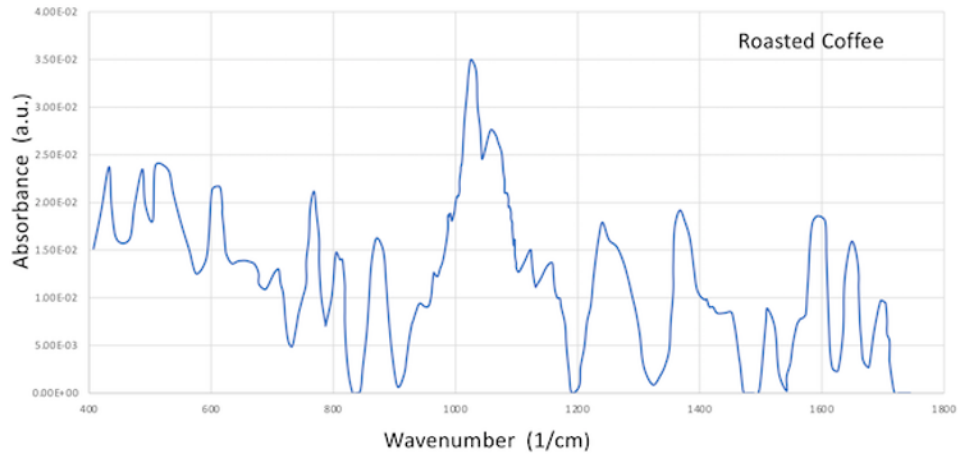


Figure 16. Residual absorbance $A_r(\omega)$ spectrum of Arabian coffee obtained by subtraction of background spectrum (Figure 15).

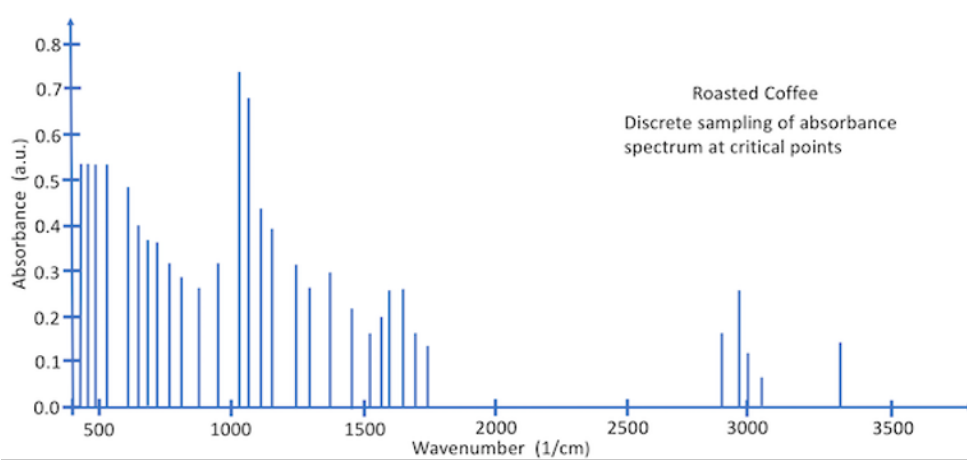


Figure 17. Discrete-function approximation of absorption function (Figure 14) for Roasted coffee.

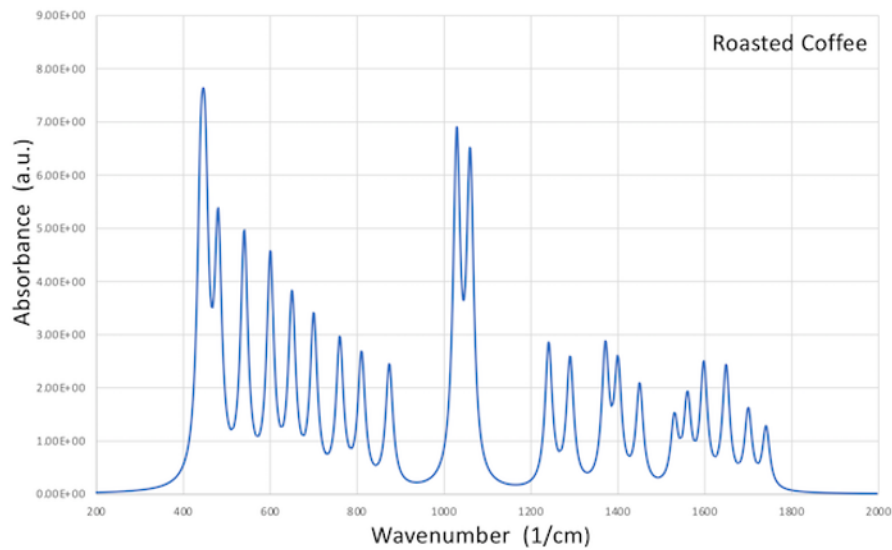


Figure 18. Truncated basis-function representation of caffeine absorbance spectrum constructed using sampled critical points (Figure 17).

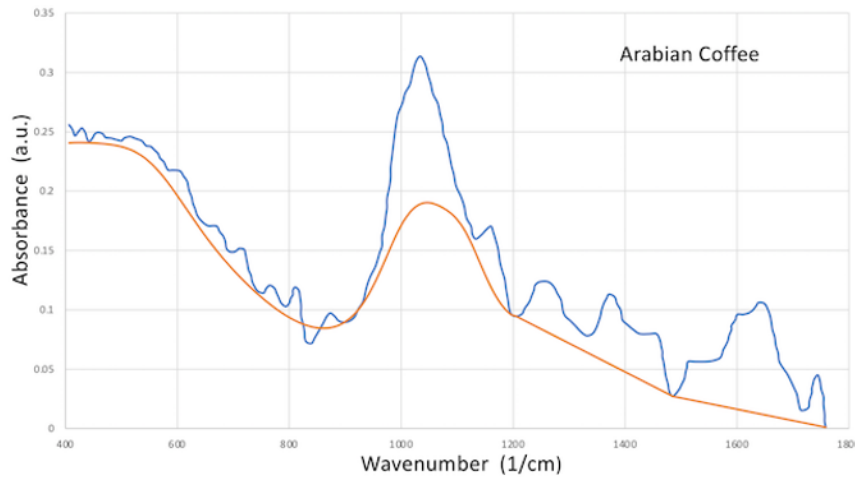


Figure 19. Measured absorbance function of Arabian coffee (blue) and background spectrum for feacher extraction (red), constructed using Gaussian, Lorentzian and linear functions.

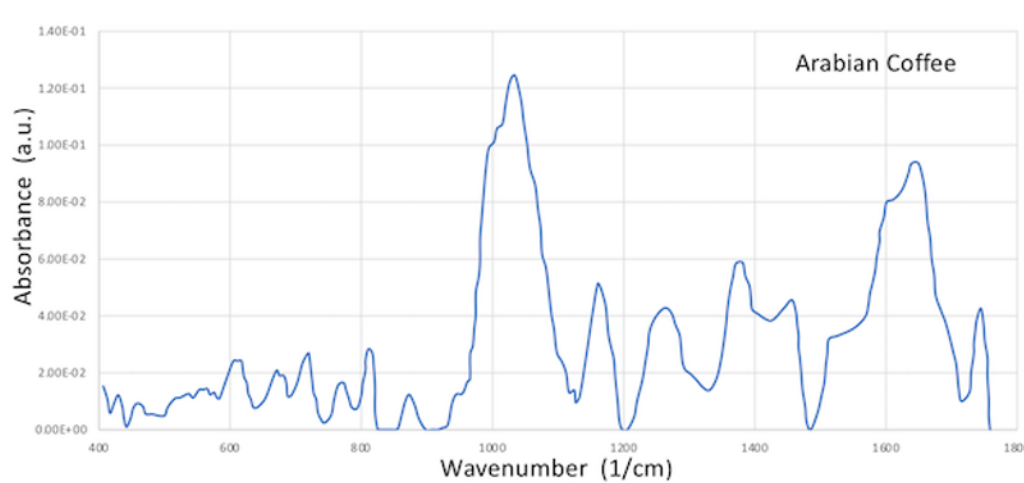


Figure 20. Residual absorbance $A_r(\omega)$ spectrum of Arabian coffee obtained by subtraction of background spectrum (Figure 19).

5. CONCLUSION

This report presents a feature-extraction methodology in terms of its general formulation, which is an inverse-model space. In addition, described here are case studies demonstrating spectrum-feature extraction from absorbance spectra for a prototype substance (coffee), using the methodology. The mathematical analysis and case studies described in this report are considered preliminary. In particular, futher application of the model space for spectrum-feature extraction is necessary to examine sensitivity of features with respect to parameter adjustment, for different types of IR-spectra, associated with different materials and spectroscopic measurements.

ACKNOWLEDGEMENT

Funding for this project was provided by the Office of Naval Research (ONR) through the Naval Research Laboratory's Basic Research Program.

REFERENCES

- [1] A. Kirsch, *An Introduction to the Mathematical Theory of Inverse Problems*, Springer-Verlag, New York, 1996.
- [2] A. Tarantola, "Inverse Problem Theory and Methods for Model Parameter Estimation," SIAM, Philadelphia, PA, 2005.
- [3] W.W. Wendlandt and H.G. Hecht, *Reflectance Spectroscopy*, John Wiley & Sons, New York (1966).
- [4] G. Kortum, *Reflectance Spectroscopy*, Springer-Verlag, Berlin (1969).
- [5] H.G. Hecht, "The Interpretation of Diffuse Reflectance Spectra," *Journal of Research of the National Bureau of Standards-A, Physics and Chemistry*, Vol. 80A, No.4, July-August (1976), pp. 567-583.
- [6] J. Torrent and V. Barron, *Diffuse Reflectance Spectroscopy*, Chapter 13, *Methods of Soil Analysis, Part 5, Mineralogical Methods*, SSSA Book Series, no. 5, Soil Science Society of America, Madison, WI, 2008, pp. 367-385.
- [7] C.F. Bohren and D.R. Huffman, *Absorption and Scattering of Light by Small Particles*, Wiley-VCH, 2004.
- [8] N.K. Afseth, A. Kohler, "Extended Multiplicative Signal Correction in Vibrational Spectroscopy, A Tutorial," *Chemometrics and Intelligent Laboratory Systems*, 117 (2012), pp. 92-99.
- [9] A. Kohler, J. Sule-Suso, G.D. Sockalingum, M. Tobin, F. Bahrami, Y. Yang, J. Pijanka, P. Dumas, M. Coote, D.G. van Pittius, G. Parkes, H. Martens, "Estimating and Correcting Mie Scattering in Synchrotron-Based Microscopic Fourier Transform Infrared Spectra by Extended Multiplicative Signal Correction," *Applied Spectroscopy*, 52 (3) (2009), pp. 259-266.
- [10] P. Geladi, D. MacDougall, H. Martens, "Linearization and Scatter-Correction for Near-Infrared Reflectance Spectra of Meat," *Appl. Spectrosc.*, 39 (1985) 377-562.
- [11] H. Martens, J.P. Nielsen, S.B. Engelsen, "Light scattering and light absorbance separated by extended multiplicative signal correction. application to near-infrared transmission analysis of powder mixtures," *Anal. Chem.* 75 (2003) 394-404.
- [12] S. Ebel, J.S. Kang, W. Windmann, "Normalization of in Situ-Spectra in Thin Layer Chromatography, chapter in *Software Developments in Chemistry 4*, pp. 207-219, 1990, Springer, Berlin. DOI: 10.1007/978-3-642-75430-2_23
- [13] B.B. H. Zhao, L. Zhang, X. Zhang, J. Liu, T. Wu, S. Wang, "Hyperspectral Feature Extraction Based on the Reference Spectral Background Removal Method," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, Vol. 8, Issue 6 (2015), pp. 1-27.
- [14] V. Mazet, C. Carteret, D. Brie, J. Idier, B. Humbert. Background removal from spectra by designing and minimizing a non-quadratic cost function, *Chemom. Intell. Lab. Syst.* 76 (2005) pp. 121-133.
- [15] G. Schulze, A. Jirasek, M.M. Yu, A. Lim, R.F. Turner, M.W. Blades, Investigation of selected baseline removal techniques as candidates for automated implementation, *Appl. Spectrosc.* 59 (2005) pp. 545-574.
- [16] P. Lasch, "Spectral Pre-Processing for Biomedical Vibrational Spectroscopy and Microspectroscopic Imaging," *Chemometrics and Intelligent Laboratory systems*, Vol. 117 (2012), pp. 100-114.
- [17] S.W. Smith, *The Scientist and Engineer's Guide to Digital Signal Processing*, California Technical Pub., 1997, ISBN: 0966017633.

- [18] S.G. Lambrakos, R. Furstenberg, C. Breshike, C.A. Kendziora, T. Huffman, R.A. McGill, A. Shabaev, "Inverse-Spectral-Analysis of Diffuse Reflectance for Surface-Distributed Particles," SPIE paper 12116-60, SPIE Defense and Commercial Sensing, 2022.
- [19] S. G. Lambrakos, R. Furstenberg, C. Breshike, C.A. Kendziora, T. Huffman, R.A. McGill, A. Shabaev, "Spectral-Feature Extraction from Diffuse Reflectance Using Multiplicative-Factor Decomposition," SPIE paper 11727-35, SPIE Defense and Commercial Sensing, 2021.
- [20] M.A. Abdalla, "Determination of Caffeine, the Active Ingredient in Different Coffee Drinks and Its Characterization by FTIR/ATR and TGA/DTA," International Journal of Engineering and Applied Sciences, Vol. 2, Issue 12, (2015), pp. 85-89.