Diffuse Reflectance Spectroscopy of Hidden Objects. Part II: Recovery of a Target Spectrum

Alexey L. Pomerantsev^{1,2}, Oxana Ye. Rodionova¹, and Alexej N. Skvortsov³



Applied Spectroscopy 0(0) 1–12 © The Author(s) 2017 Reprints and permissions: sagepub.co.uk/journalsPermissions.nav DOI: 10.1177/0003702817695288 journals.sagepub.com/home/asp



Abstract

In this study, we consider the reconstruction of a diffuse reflectance near-infrared spectrum of an object (target spectrum) in case the object is covered by an interfering absorbing and scattering layer. Recovery is performed using a new empirical method, which was developed in our previous study. We focus on a system, which consists of several layers of polyethylene (PE) film and underlayer objects with different spectral features. The spectral contribution of the interfering layer is modeled by a three-component two-parameter multivariate curve resolution (MCR) model, which was built and calibrated using spectrally flat objects. We show that this model is applicable to real objects with non-uniform spectra. Ultimately, the target spectrum can be reconstructed from a single spectrum of the covered target. With calculation methods, we are able to recover quite accurately the spectrum of a target even when the object is covered by 0.7 mm of PE.

Keywords

Near-infrared diffuse reflectance spectroscopy, multilayer systems, reflectance, scattering, polyethylene films, spectrum recovery, multivariate curve resolution, MCR

Date received: 7 October 2016; accepted: 30 January 2017

Introduction

Diffuse reflectance (DR) spectroscopy of multilayer systems is a highly demanded but poorly studied area. The literature survey provides us with an overview of a few applications in various fields. Pi et al.¹ studied cheese properties through polyethylene (PE) film. The work by Saleem et al.² describes the usage of near-infrared (NIR) spectroscopy for the detection of specific chemicals when concealed under a layer of clothing. It is found that concealment modifies the spectrum of a particular chemical in a DR experiment. The development of optical methods in modern medicine and the investigation of optical properties of various biological tissues are considered in this study.³ It presents an overview of absorption and scattering properties of skin and subcutaneous tissues measured in a wide wavelength range. Basic principles of measurement of the optical properties of tissues and techniques used for processing the collected data are also outlined in literature. In particular, Yang et al.⁴ describe mathematical techniques to correct for analyte-irrelevant optical variability in spectral properties of a tissue overlying and within the muscle. The application of a principal component analysis (PCA) is discussed for the removal of inter-subject, analyte-irrelevant variations in

muscle scattering from continuous wave DR NIR spectra. In Zieba-Palus and Borusiewicz,⁵ IR microspectrometry and Raman spectroscopy have been applied to the examination of multilayer fragments of paints. A spatially offset Raman spectroscopy is applied to a nondestructive analysis of thin painted layers.⁶

The current study investigates the spectra of various samples covered by the layers of material of different thickness. Diffuse reflectance NIR spectroscopy has been chosen as the instrumental method. To begin with, and to simplify the problem, we have chosen layers of PE as covering material. Previously, we have succeeded in decomposition of the collected DR NIR spectra of PE layers into three components and interpretation of the physicochemical

Corresponding author:

¹Semenov Institute of Chemical Physics RAS, Moscow, Russia ²Branch of Institute of Natural and Technical Systems RAS, Sochi, Russia ³Peter the Great St Petersburg State Polytechnic University, St. Petersburg, Russia

Alexey L. Pomerantsev, Semenov Institute of Chemical Physics RAS, Kosygin str. 4, 119991, Moscow, Russia. Email: forecast@chph.ras.ru

meaning of these components. The approach is described in detail in the first part of the study.' A natural continuation is the application of the described procedure for the recovery of spectra of samples covered by PE. The goal of the present study is reconstruction of the DR spectra of the object of interest, further referred to as a target object, from the spectra acquired through PE layers of unknown thickness. At the first stage, we have investigated a system of PE layers of different thickness disposed on the standard underlayers, which are spectrally flat over a wide wavelength range and have known reflectance factors. In this paper, we apply the obtained results⁷ to systems that consist of varying number of PE layers and real objects as an underlayer. We deal with a case in which a measurement is performed through several layers of PE and we cannot change (or evaluate) the thickness of the PE layer.

For ease of understanding, the 'Methods' section begins with an overview of the results obtained in the first part of this research.⁷

Methods

Summary of the Results Obtained in the First Part of the Research

During the first stage, as covered in the 'Introduction,'⁷ we have investigated the DR spectra of a system: "the multilayer PE cover of depth h disposed on a solid underlayer, which is spectrally flat over a wide wavelength range with the known reflectance factor r." Five data sets with underlayers of various reflectance factors, 0%, 10%, 40%, 80%, and 100% have been measured. For each underlayer, a set of spectra with a different number of PE layers has been acquired. The number of layers was in the range of 1-48, and the overall corresponding thickness was in the range of 0.087-4.18 mm.

We refer to the spectral readings presented in reflectance units as the R-domain. These values are converted using logarithmic transformation,* $A = -\ln(R)$, into "absorbance units," which are referred to as the Adomain. It has been shown that any spectrum acquired from data sets $x(\lambda)$, presented in the A-domain, can be decomposed as a sum of three components

$$x(\lambda) = c_t(h, r)s_t(\lambda) + c_a(h, r)s_a(\lambda) + c_s(h, r)s_s(\lambda) + e$$
(1)

The error term e is rather small, namely $\|\mathbf{e}\| / \|\mathbf{x}\| \approx 0.4\%$. Hereinafter, $\|\mathbf{x}\| = \sqrt{\sum x_i^2}$.

The model in Eq. I includes factors c, which depend on the PE thickness, h, and the reflectance factor of an underlayer, r, as well as spectra s, which depend on the

0.0 8500 7500 6500 5500 4500 wavenumbers **Figure 1.** Polyethylene pure component spectra: \mathbf{s}_{t} , transmission (1, green), \mathbf{s}_{a} , absorption (2, red), \mathbf{s}_{s} , scattering (3, blue).

Attenuation spectrum (4, magenta).

wavelength λ . Three pure spectra $s(\lambda)$ involved in this formula are presented in Figure 1. Spectrum $\mathbf{s}_t \equiv \mathbf{I}$ does not depend on the wavelength λ and it stands for the direct transmission/reflection effects. Spectrum $s_a(\lambda)$ can be interpreted as the PE absorbance coefficient. Spectrum $s_s(\lambda)$ predominantly characterizes the PE scattering.

Coefficients \mathbf{c}_t , \mathbf{c}_a , and \mathbf{c}_s are the weighting factors, which describe the contributions of each of the pure spectra into spectrum x collected by a detector. These factors depend on h and r, and they can be modeled by the following explicit functions

$$c_t = F_t(h, r) \quad c_a = F_a(h, r) \quad c_s = F_s(h, r)$$
(2)

Functions for c_a , and c_s are rather complicated. The reader can find them in the online Supplemental Material.

In Part I of this research,⁷ we have demonstrated that the weighting profiles obtained for the first component can be perfectly described by a simple relaxation model

$$c_t(h, r_{\rm UL}) = -\ln[r_{\rm PE} + (r_{\rm UL} - r_{\rm PE})e^{-kh}]$$
(3)

where r_{UL} is the reflectance factor of an underlayer, $r_{\rm PE} = 0.27$ is the reflectance factor of PE, and k is an empirical decay constant, which was found to be close to 1 mm⁻¹ for the used PE samples. We have also shown that the c_r values for all data sets agree well with the spectral baseline shifts, i.e.,

$$c_t(h, r) = \min x(\lambda, h) \tag{4}$$

In case of a thin PE layer (h < 0.6 mm), it has been shown that the general Eq. I can be simplified in the



 $^{^*}$ In practice, a decimal logarithm is routinely used, but we prefer a Napierian logarithm as it is more suitable for a theoretical discussion.

following way. Consider conventionally pre-preprocessed spectra

$$\tilde{x}(\lambda, h) = x(\lambda, h) - \min x(\lambda, h)$$
(5)

Finally, we obtain the equation,

$$\tilde{x}(\lambda, h) = h \cdot c_u(r)s_u(\lambda) + e \tag{6}$$

In this equation, spectrum $s_u(\lambda)$ (curve 4, Figure 1) represents the attenuation coefficient found in Pomerantsev et al.⁷ Weighting coefficient $c_u(r)$ depends on the underlayer reflectance factor r, but not on the PE depth, h.

Thus, in the first part of the research we obtained the following outcomes. We have found three pure spectra $s_t(\lambda)$, $s_a(\lambda)$, $s_s(\lambda)$, and three corresponding weighting profiles $c_t(h, r)$, $c_a(h, r)$, $c_s(h, r)$. For the thin PE layer, we can use a linear dependence on h and the attenuation coefficient $s_{\mu}(\lambda)$. These results were obtained using a series of experiments with the multilayer PE cover disposed on the exemplar underlayers with a known uniform reflection spectrum. In the current research, we set two goals. The first one is to test our "calibration" model and demonstrate its applicability to a situation when an underlayer is a real object with a complex NIR absorbance/reflectance spectrum. For this purpose, we assume the spectrum of a target object to be known. The second goal is to apply the 'calibration' model for the reconstruction of the unknown spectrum of the real target object in case the PE thickness is unknown.

Multivariate Curve Resolution (MCR)

Our approach to the determination of the target spectrum is based on a fundamental bi-linear relation

$$\mathbf{X} = \mathbf{C}\mathbf{S}^{\mathsf{t}} + \mathbf{E} \tag{7}$$

Here, **X** is the $(I \times J)$ matrix that contains the spectra in the A-domain of I samples recorded for J wavelengths, **C** is the $(I \times K)$ matrix of the weighing factors (pseudo-concentrations) related to K components, **S** is the $(J \times K)$ matrix of the components' spectra, and **E** is the $(I \times J)$ matrix of errors.

In general MCR theory, matrices C and S in Eq. 7 are unknown. They can be estimated using the multivariate curve resolution alternating least squares (MCR-ALS) method, which comes to the following optimization problem

$$\begin{array}{l} \underset{C,S}{\text{minimize}} \|\mathbf{X} - \mathbf{CS}^{t}\|^{2} \\ \text{subject to constraints} \end{array} \tag{8}$$

Multivariate curve resolution alternating least squares is a well-developed technique that is described elsewhere.^{9,10} This procedure consists of two steps, the C-type step and the S-type step, which are repeated until convergence. At the C-type step, the spectral profiles are fixed, $S \equiv S_{hat}$, and the C matrix is estimated using the unconstrained least squares (LS) method. Afterwards, the estimated matrix C_{in} is transformed into matrix C_{hat} to incorporate the known constraints. For the S-type step, the concentration profiles are fixed, $C \equiv C_{hat}$, and matrix S is found applying a similar LS estimator. Subsequently, matrix S_{in} is transformed into matrix S_{hat} to account for the spectral constraints. The manner of application of MCR-ALS for our system is minutely presented in the first part of this research⁷ and briefly described in the next subsection.

Multivariate Curve Resolution for Multi-Layer System and Problem Ambiguity

Coming back to our problem, we suppose that the three pure spectra presented in Eq. I may be used in a general case of an unknown target. Our hypothesis is based on the assumption that the sought-for spectrum of the underlayer can be empirically represented as a sum of the reflectance factor r, and an additional term, which models the specific spectral features of the underlayer. The former term does not depend on λ , it goes into the weighting factors in Eqs. I-3 and does not add any new components. The latter term depends on λ and comprises specific spectral properties of the underlayer. Thus, the second term adds one new component to the model. Therefore, we select the number of components in Eq. 7 to be K = 3 + 1 = 4. The first three spectra $\mathbf{s}_1 - \mathbf{s}_3$ in matrix **S** are known from the analysis of PE samples with standard underlayers and the last component of **S** is the unknown spectrum s_4 of the target. Using this approach, we arrive at the following optimization problem

$$\underset{\mathbf{c}_{1}, \mathbf{c}_{2}, \mathbf{c}_{3}, \mathbf{c}_{4}, \mathbf{s}_{4}}{\text{minimize}} \left\| \mathbf{X} - [\mathbf{c}_{1}, \mathbf{c}_{2}, \mathbf{c}_{3}, \mathbf{c}_{4}] [\mathbf{s}_{1}, \mathbf{s}_{2}, \mathbf{s}_{3}, \mathbf{s}_{4}]^{\mathsf{t}} \right\|^{2}$$

$$\text{subject to } \mathbf{s}_{1} \equiv \mathbf{1}; \mathbf{s}_{2} \equiv \mathbf{s}_{a}; \mathbf{s}_{3} \equiv \mathbf{s}_{s}$$

$$(9)$$

It is evident that the solution of the task given in Eq. 9 is not unique. As an example, we consider an important case in which matrix **X** consists of a single spectrum **x**, i.e., I = I. If we select arbitrary non-negative values for c_1 , c_2 , and c_3 , and set $c_4 = I$, the spectrum

$$\mathbf{s}_4 = \mathbf{x} - c_1 \mathbf{1} - c_2 \mathbf{s}_a - c_3 \mathbf{s}_s \tag{10}$$

satisfies Eq. 9 exactly, **E**=0. In particular, we can obtain the following trivial solution: $c_1 = c_2 = c_3 = 0$, $c_4 = 1$, $s_4 = x$. Needless to say that in a general case, when l > 1, we have a similar problem with ambiguity. To resolve this issue, we suggest using the relationships given in Eq. 2 for calculation of the *c*-profile values.

Let us presume that both parameters, the reflectance factor of the target, r, and the PE depth, h, are known, or can be estimated (see the section, Assessment of r and h).

In case l = 1, the weighting factors c in Eq. 10 are selected with respect to Eq. 2,

$$c_1 = F_t(h, r), \quad c_2 = F_a(h, r), \quad c_3 = F_s(h, r), \quad c_4 = 1$$
(11)

It is clear that the proposed estimation of the target spectrum solves the Eq. 7 with the zero error E. Therefore, no optimization in Eq. 9 is needed.

When l > l, we use a similar approach. Vectors \mathbf{c}_1 , \mathbf{c}_2 , and \mathbf{c}_3 are selected as $\mathbf{c}_1 = \mathbf{F}_t$, $\mathbf{c}_2 = \mathbf{F}_a$, $\mathbf{c}_3 = \mathbf{F}_s$, where vectors \mathbf{F} are calculated using the corresponding functions given in Eq. 2, e.g., $\{\mathbf{F}_a\}_i = F_a(h_i, r)$, where h_i is the PE depth in the *i*th sample. The unknown target components, \mathbf{c}_4 and \mathbf{s}_4 , are found in the course of the following optimization problem

$$\begin{array}{l} \underset{\mathbf{c}_{4}, \mathbf{s}_{4}}{\text{minimize}} \left\| \mathbf{X} - \mathbf{C}\mathbf{S}^{\mathsf{t}} \right\|^{2} \\ \text{subject to} \\ \mathbf{s}_{1} \equiv \mathbf{1}; \quad \mathbf{s}_{2} \equiv \mathbf{s}_{a}; \quad \mathbf{s}_{3} \equiv \mathbf{s}_{s}; \\ \mathbf{c}_{1} = \mathbf{F}_{t}; \quad \mathbf{c}_{2} = \mathbf{F}_{a}; \quad \mathbf{c}_{3} = \mathbf{F}_{s} \end{array}$$

$$(12)$$

Here, matrices **C** and **S** consist of two parts: the first three components in each matrix, $C_{13} = \{c_1, c_2, c_3\}$ and $S_{13} = \{s_1, s_2, s_3\}$, are known and only the last component, represented by vectors c_4 and s_4 , is unknown. In that case, the LS estimators for matrices **S** and **C** are given by the equations

$$C_{in} = \{ C_{13}, \quad (\mathbf{X} - C_{13} \mathbf{S}_{13}^{t}) \mathbf{s}_{4}^{+} \}$$

$$S_{in} = \{ S_{13}, \quad (\mathbf{X} - C_{13} \mathbf{S}_{13}^{t})^{t} \mathbf{c}_{4}^{+} \}$$
(13)

where $^+$ means the matrix pseudo-inverse, i.e., $\bm{A}^+ = \bm{A}(\bm{A}^t\bm{A})^{-1}$

The thin-layer approach presented in Eq. 6 can also be used to estimate the target spectrum. Here we should solve the following optimization problem

$$\underset{\mathbf{c}_{23}, \mathbf{c}_{4}, \mathbf{s}_{4}}{\text{minimize}} \| \mathbf{\tilde{X}} - \mathbf{c}_{23} \mathbf{s}_{23}^{t} - \mathbf{c}_{4} \mathbf{s}_{4}^{t} \|^{2}$$
subject to $\mathbf{s}_{23} \equiv \mathbf{s}_{u}$

$$(14)$$

Spectral matrix $\tilde{\mathbf{X}}$ is obtained from the raw matrix \mathbf{X} using the offset pre-processing given in Eq. 5. The first component spectrum \mathbf{s}_{23} is set to be equal to the attenuation coefficient \mathbf{s}_u . When matrix $\tilde{\mathbf{X}}$ contains a single spectrum $\tilde{\mathbf{x}}$, the weighting vectors \mathbf{c}_1 and \mathbf{c}_4 turn into the numbers c_1 and c_4 . To avoid scaling ambiguity, we should select $c_4 = 1$.

Spectra Comparison and Accuracy

After reconstruction of the target spectrum, we should evaluate the quality of the result. The accuracy of the

MCR modeling is conventionally evaluated by the root mean square error (RMSE)

$$RMSE = \frac{1}{\sqrt{IJ}} \left\| \mathbf{X} - \mathbf{C}_{hat} \mathbf{S}_{hat}^{t} \right\|$$
(15)

which estimates the error term **E** in Eq. 7. This measure can be used in case a true target spectrum is unknown, but it cannot be applied to the recovery of the target from the single spectrum (l = l), because in the latter case **E** = **0**.

Let us suppose that the spectrum of the target object, \mathbf{s}_{tar} is known. For example, it can be obtained in an independent experiment or found in literature. Obviously, the target spectrum is not unique because it can be arbitrary scaled and shifted. To account for such an issue, we should select coefficients k and m, which minimize the distance between the estimate, \mathbf{s}_{hat} , and the target, \mathbf{s}_{tar} i.e.,

$$\underset{k,m}{\text{minimize}} \|\mathbf{s}_{\text{tar}} - k\mathbf{s}_{\text{hat}} - m\mathbf{1}\|^2$$
(16)

After that, the linearly corrected spectrum $\mathbf{s}_{cor} = k\mathbf{s}_{hat} + m\mathbf{I}$ can be compared with the target spectrum \mathbf{s}_{tar} using conventional measures of similarity

$$d = d_0^{-1} \| \mathbf{s}_{\text{tar}} - \mathbf{s}_{\text{cor}} \| / \sqrt{J}; \quad C = \operatorname{cor}(\mathbf{s}_{\text{tar}}, \, \mathbf{s}_{\text{cor}})$$
(17)

Here, *d* is the relative distance between the spectra, and *C* is the correlation coefficient. The scaling factor d_0 can be selected arbitrarily in order to simplify the interpretation of the matching spectra. For example, RMSE can be used as a scaling parameter in Eq. 17, i.e., $d_0 = \text{RMSE}$. In this case, the quality of the target spectrum recovery can be assessed in such terms as "better (or worse) than the modeling of the whole spectral set **X**."

Note that both quality criteria are not sensible to the difference in the reflectance factor and primarily compare the shapes (spectral features) of the two spectra. It is also important to emphasize that the true spectrum \mathbf{s}_{tar} is employed only for the assessment of the recovering quality and is not used in the target reconstruction algorithm.

Assessment of **r** and **h**

We cannot expect that the reflectance factor of the target is given in advance. The PE thickness is also a questionable parameter, which is often unknown. This case has been considered Rodionova et al.¹¹ where the NIR spectrum of a packed pharmaceutical substance was acquired through the PE film. Despite the fact that the PE film thickness has been carefully measured, the actual PE depth under the sensor probe was unknown because of the folds of the film in different parts of the package. In this subsection, we show how to get rid of the burdensome assumptions that both the target reflectance factor, *r*, and the PE depth, *h*, are known.

Factor r is found using the following approach. Let us select a single spectrum $x(\lambda)$ from the data matrix **X**. In a general case (l > 1), we should select the spectrum, which is acquired at the minimal PE depth, h_0 . Let $x_0 = \min x(\lambda)$. Applying Eqs. 3 and 4, we obtain the estimate of the target reflectance factor

$$r = r_{\rm PE} + (e^{-x_0} - r_{\rm PE})e^{kh_0}$$
(18)

Note that r_{PE} and k are the specific PE properties, which have been obtained in Pomerantsev et al.⁷ at the development stage of model (Eq. 1).

The method used for the assessment of the PE depth is more complicated. First, we should pay attention to the strong peaks around 5600–5900 cm⁻¹ (Figure 1). The reasons for the special attention to the abovementioned spectral region are high values in the absorption \mathbf{s}_a and scattering \mathbf{s}_s spectra of PE, and simultaneous rapid changes in both spectra. As a result, we obtain a high absolute error in the course of spectrum modeling in this range. Exactly in this area, a wrong selection of PE depth *h* causes artifacts, which are manifested as additional peaks (PE peaks or their negative images) folded into the target spectrum.

For example, Figure 2a demonstrates how the recovered target spectra \mathbf{s}_{hat} in the range 5900–5600 cm⁻¹ depend on the chosen PE thickness *h*. The true target spectrum (black, 1) has no peaks in this area, but the recovered spectrum with h = 0.07 (blue, 2) has a distinctive peak near 5778 cm⁻¹ which corresponds to the PE absorbance band. As the chosen PE depth increases (h = 0.078, green curve 3), the solution approaches the target, but when *h* increases further (h = 0.1, red curve 4), the result deviates farther from the true spectrum. This example illustrates the idea that lies at the heart of the proposed method, namely the assessment of the value *h* of the PE depth that minimizes variations of the sought spectrum \mathbf{s}_{hat} .

To be more specific, we define an index of spectrum variations, V, by the following formula

$$V(h) = \frac{1}{J} \sum_{j=1}^{J-1} \left| s(\lambda_j) - s(\lambda_{j+1}) \right|$$
(19)

The optimal depth h corresponds to the minimum of the function V(h):

$$h_{\rm opt} = \arg\min_{h} V(h) \tag{20}$$

In other words, we search for h, which minimizes peaks and smoothens out the predicted spectrum, as index V is directly related to the first derivative of $s(\lambda)$. If we ignore the unlikely case when all peaks of the underlayer overlap PE peaks, the spectrum with the least number of spectral features is most likely the spectrum with the least number of artifacts.

Certainly, it is necessary to verify whether the target spectrum really fits well at the optimal depth h. The measure of closeness d, defined in Eq. 17, can be employed for this assessment. In Figure 2b we demonstrate a plot in which the quality measure d (the left axis, red curve, 1) is shown in parallel with the variation index V (the right axis, blue curve, 2); both values are plotted in dependence on the selected depth h. An amazing match between the locations of the minima can be seen. This confirms our idea that the PE thickness can be calculated by finding the minimum of V(h).

Materials and Measurement

Materials

Cover Material. The cover material is the same PE, which was used in Pomerantsev et $al.,^7$ i.e., a big sheet of low



Figure 2. (a) Spectra \mathbf{s}_{hat} reconstructed for the selected PE depths *h*: (1, black) target, (2, blue) h = 0.07, (3, green) h = 0.078, (4, red) h = 0.1. (b) Recovering quality *d* (curve, 1, the left axis, in red) and the variation index *V* (curve, 2, the right axis, in blue) in dependence on the selected PE depth, *h*.

Target Objects. The first target object is an artificial solid object. It is the NIST Traceable Extended Range Near-Infrared Wavelength Standard MRC-910-1920x (MiddletonResearch).⁸ This commercial standard consists of four components, three rare earth oxides (Dy_2O_3 , Er_2O_3 , Ho_2O_3) and talc. All the components are sintered in a spectrally neutral matrix. The standard has a rich spectrum that covers a wide NIR range of 10 000–4000 cm⁻¹.

The second object is a medicine, a derma-protection powder consisting of zinc oxide 10 wt%, starch 10 wt%, and talc 80 wt%. This example is used to illustrate the applicability of the method to target objects such as a highly reflective powder. The sample is packed in a small PE bag 0.046 mm thick in such a way that the powder is about 2 cm deep in the illuminated area. Thus, we can consider that the target object has an optically infinite depth.

Measurements

The measurements technique is similar to that used in the first part of the study. Fourier transform NIR Spectrometer Nicolet 6700 (ThermoFisher, USA) with a InGaAs detector is used for the spectra acquisition in the range of 4000– 12400 cm^{-1} , with resolution 8 cm⁻¹. The number of scans is equal to 64. The 0.99 DR disk, SRS-99, is used for the background measurements. The DR spectra are collected with the help of a fiber optic probe fixed in a mounting. Several layers of the PE are placed on the mounting table with the target sample put on the top. The light source within the fiber probe accessory illuminates the samples

from below. A working range of $9000-4450 \text{ cm}^{-1}$ is used for the analysis of all spectral data in this study. The acquired reflectance spectra are converted to the A-domain. They constitute spectral matrix **X**.

Pure spectra for both target objects have been acquired but used only for the assessment of the quality of the results. Pure spectra were not used in modeling.

Case Study I: The NIST Traceable Standard

Figure 3a demonstrates the acquired spectra \mathbf{X} . The arrow follows the increase of the PE depth. Figure 3b is presented here as a reference and shows the target spectrum (1, brown), the pure PE absorption component (2, red), and the pure PE scattering component (3, blue). All data are used in their original form; no pre-processing is applied.

All Spectra Together

We begin our analysis with a case where all spectra acquired at an increasing PE depth are modeled together using the entire spectral matrix X. As mentioned above, such a case does not have much to do with real practice. However, this case is theoretically important, as it serves to confirm that the developed approach-the modeling of the optical properties of multilayer cover by three predefined pure spectra plus their corresponding weighting factors-is valid not only for spectrally flat underlayers, but also for the real-world targets. In other words, we test the practical usefulness of the hypothesis, described in the Methods section. To do so, we check whether spectra \mathbf{s}_a and \mathbf{s}_s , which have been obtained in the experiments with dedicated underlayers (Figure 3b, curves 2, 3) and their conjugated weighting coefficients \mathbf{c}_a and \mathbf{c}_s (Eq. 2) are able to model the data in the presence of spectral variations of the



Figure 3. Example I. (a) Collected data set; (b) reference spectra: target, measured; (1, brown), PE absorption term, multilayer model (2, red); PE scattering term, multilayer model (3, blue).

underlayer. Additionally, we replace the sought target spectrum \mathbf{s}_4 with the measured target spectrum. After that, we check, whether Eq. 9 is still able to handle the contribution from the PE layers and whether the empirical models given in Eq. 2 are applicable. This validation is converted to the following task

$$\underset{\mathbf{C}}{\operatorname{minimize}} \| \mathbf{X} - \mathbf{C}\mathbf{S}^{\mathsf{t}} \|^{2}$$
subject to
$$\mathbf{s}_{1} \equiv \mathbf{1}; \, \mathbf{s}_{2} \equiv \mathbf{s}_{a}; \, \mathbf{s}_{3} \equiv \mathbf{s}_{s}; \, \mathbf{s}_{4} \equiv \mathbf{s}_{\operatorname{tar}}$$

$$(21)$$

In fact, the problem given in Eq. 21 does not require a recurrent ALS optimization, because matrix \mathbf{S} is completely known, thus a solution can be obtained in single S-step

$$\mathbf{C}_{\text{hat}} = \mathbf{X}(\mathbf{S})^+ \tag{22}$$

The estimated weighting factors, $c_1 - c_4$, can be represented in dependence of the PE depth, *h*.

These results are shown in Figure 4. The markers correspond to the weighting factors c obtained by Eq. 22. The curves represent the theoretical profiles, F(h,r), calculated by Eq. 2. Parameter r is the target reflectance factor evaluated by Eq. 18. In this calculation, we employ the first spectrum acquired through a single PE layer of depth, h = 0.087 mm, to obtain the minimal spectral reading x_0 . The calculated value is r = 0.396, which expectedly matches the baseline of the known target spectrum $-\ln(r) = 0.93$ (Figure 3b, curve I).

Figure 4 shows a good agreement between values **c** and corresponding curves **F**. To demonstrate this numerically, we calculated both the mean, μ , and the standard deviation, σ , for each of the three error vectors, **c–F**. The results are as follows: $\mu_1 = 0.06$, $\sigma_1 = 0.12$; $\mu_2 = 0.01$, $\sigma_2 = 0.03$;



Figure 4. The *c*-profiles (markers) and their models (curves): c_1 , transmission (1, green), c_2 , absorption (2, red), c_3 , scattering (3, blue).

 $\mu_3 = -0.01$, $\sigma_3 = 0.03$. The Student's t-test confirms the conclusion of a good compliance. The transmission profile c_1 , which demonstrates a large, but non-regular dispersion around the curve, does not significantly differ from its model values according to the t-test. Therefore, we can conclude that the theory developed earlier⁷ is confirmed and may be employed for the analysis of practical cases.

Additionally, using these results, we calculate the RMSE defined in Eq. 15. This value, $d_0 = 0.02$, represents the overall accuracy obtained in the MCR modeling of the entire data set **X**. We will use this value as a scaling factor in Eq. 17 that assesses the quality of the recovery of the target. In case we obtain d < 1, this means the target spectrum is recovered better than the experimental spectra were explained.

Single Spectrum Through Many Layers

In this subsection, we consider a more realistic case, namely, the recovery of the underlayer spectrum using a single spectrum acquired through several PE layers. To test our approach, we employ rows \mathbf{x}_i from matrix \mathbf{X} (i = 1, 2, ...) one by one and perform the analysis using this single spectrum only. Thus, we reconstruct the target through one, two, etc. PE layers. In each case, the target reflectance factor, r, is assessed individually using Eq. 18. The true target spectrum is not used in the modeling, except for the assessment of the quality of predictions.

The target spectrum can be evaluated using the methods explained above. The direct recovery, which requires no ALS optimization, is presented in Eqs. 10 and 11. This approach is used as the basis for two sets of calculations. The first one is an option, in which the PE depth, h, is assumed to be known. For sample \mathbf{x}_i , it is calculated as $h=ih_0$, where $h_0=0.087$ mm is the average thickness of a single layer. This method is further referred to as known depth (KD). In the second option, the PE depth, h, is considered to be unknown, and it is assessed by means of optimizing the variation index, presented in Eqs. 19 and 20. We denote this approach as OD (optimized depth). The third method utilizes the "thin layer" (TL) concept given by Eq. 14.

Figure 5 demonstrates the reconstructed spectra obtained by all three methods for the case of eight PE layers.

The most challenging problems are found in the range 5600–5900 cm⁻¹ that is shown in the insert of Figure 5. These artifacts happen due to following issues. As we have seen in Figure 4, the absorption profile c_2 is modeled by function F_a , presented in Eq. 2, with the error of about 0.1. Close to the band of 5778 cm⁻¹, we observe (see Figure 1) the highest absorbance peak that is ~1.6 AU. Thus, in this area, the target reconstruction error may be near 0.16 AU. The real deviations from the target, observed in the insert of Figure 5, are about 0.1–0.2 AU, which fit



Figure 5. Example 1. The target (0, red) and its estimates through eight PE layers: KD method (1, blue), OD (2, green), and TL (3, brown). The challenging range is shown in the insert.

 Table I. Example I. Quality characteristics of the target spectrum recovery.

Layers Sample no.	KD		OD		TL	
	d	С	d	С	d	С
I	0.140	1.000	0.201	0.999	1.332	0.976
2	0.357	0.998	0.357	0.998	1.348	0.976
3	0.512	0.997	0.425	0.998	1.376	0.975
4	0.673	0.994	0.673	0.994	1.501	0.97
5	1.015	0.986	0.618	0.995	1.478	0.971
6	0.757	0.992	0.578	0.996	1.447	0.972
7	1.786	0.957	0.809	0.991	1.833	0.955
8	1.769	0.958	0.792	0.992	2.286	0.929
9	4.143	0.741	2.285	0.929	2.911	0.882
10	I.857	0.954	1.271	0.979	3.852	0.781

Values given in bold represent the best solution among all methods.

into the estimate range of possible errors. At the same time, we see that the peak around 5690 cm^{-1} , which is attributed to the target object, is perfectly reconstructed in spite of it being located between two large peaks of PE. This proves that the variation optimization approach does not affect the original peaks of the target spectrum.

Numerical results are given in Table I, which summarizes the results of the three methods, KD, OD, and TL, for the respective number of PE layers listed in column I. Columns d and C contain the quality measures defined in Eq. 17.

Table I shows that the quality unsurprisingly deteriorates as the number of layers is growing. In general, the PE depth optimization (OD method) provides better results in comparison with the other methods. This is evident because this technique seeks out the optimal value among all possible depths, including the depth that is considered to be known. In this example, the TL method provides the results that are the worst for all depths. Applying the cutoff criterion, which is d < 1, we conclude that the target coated with eight layers of PE (h < 0.7 mm) is still accessible for reconstruction. This case is presented in Figure 5.

Both the KD and OD algorithms determine the reflectance factor r of a target object. The latter method also calculates the PE depth, h. It is interesting to compare the assessments of these parameters. Figure 6 demonstrates the reflectance factor, r (the left axis), and the PE depths, h (the right axis), calculated for a different number of the PE layers.

Figure 6 represents both r and h estimates shown in one plot. The reflectance values are shown in red; they relate to the left axis. The dots (Ia) stand for the KD method and the diamonds (1b) represent the OD method. The red line (1c) shows the reference value r = 0.396 found in the All Spectra Together section, where all spectral data are analyzed together. Obviously, the r value found for nine layers is an outlier. The abnormality of this observation may also be seen in Table I. If we exclude this point, the average reflectance factors become $r_{\rm KD} = 0.40 \pm 0.04$, $r_{OD} = 0.41 \pm 0.04$. This result is interesting, because in this experiment we obtain very unstable values of the baseline shifts, which can be clearly seen in Figure 4, where the c_1 profile has large deviations from the theoretical model. Evaluation of the reflectance factor is primarily based upon the baseline shift that is x_0 value in Eq. 18. Nevertheless, we can conclude that the reflectance factor is rather well assessed by both methods, and the different estimates coincide.

The PE depth values, h, are shown in blue; they relate to the right axis. The known values are represented by line (2a), described by the equation $h = 0.087 \cdot n$, where n is the number of layers. Blue triangles (2b) stand for the PE depth, which is found in the variation optimization procedure (Eqs. 19 and 20). For a small number of layers ($n \le 6$), the optimized depths are very close to their known counterparts. For a larger number of layers (n > 6), the estimated PE depth can essentially differ from the expected value. This effect cannot be fully explained by the variation in the PE sheet thickness, which is much smaller than deviations



Figure 6. Example 1. The target reflectance factor, r (left axis, in red) assessed by the KD method (dots, 1a), the OD method (diamonds, 1b), the reference value found in the All Spectra Together section (line, 1c). The PE depths, h (right axis, in blue): the known values (line, 2a), the OD method (triangles, 2b).

shown in Figure 6. In our opinion, there are two factors that may contribute to the error. The first one is the modeling uncertainty in calibration profiles obtained in the first part of this research, which generally increases with the number of layers. The second factor is the fact that h is a result of optimization, which heavily relies on minimization of the spectral features of the recovered spectrum. As the optical thickness becomes larger, the quality of the recovery of peaks of the underlayer spectrum decreases, making the optimization in Eq. 20 less stable.

Case Study 2: Powder Sample

The pure target spectrum \mathbf{s}_{tar} is obtained directly from a high powder pile without any gap between the sample and the fiber probe. The acquired target spectrum manifests the high observed reflectance factor, which is greater than the 99% spectralon reference sample at most wavelengths. Such observations with r > 1 are possible because of the finite detection angle of the probe, when the detector captures only part of the reflected light. In some materials, light can be predominantly reflected toward the probe. In this case the fraction of light captured by the detector can be larger than that from an ideal diffuse reflector. The second target object contains 80% of talc powder consisting of the oriented tiny flat crystals, which produce predominant reflection in specific directions. In other terms, this effect can be explained by a non-Lambertian nature of scattering, i.e., a specific shape of the reflection indicatrix.

Figure 7a demonstrates the spectra acquired through the growing number of PE layers. The arrow represents the direction of the increasing PE depth. Figure 7b shows the reference spectra that include: the target (1, brown), the PE absorption (2, red), the PE scattering (3, blue). All data are used in their original form; there is no pre-processing applied.



Figure 7. Example 2. (a) Collected data set; (b) reference spectra: target, measured (1, brown); PE absorption term, multilayer model (2, red); PE scattering term, multilayer model (3, blue).

All Spectra Together

It should be mentioned, that the theory presented in Pomerantsev et al.⁷ for modeling the DR spectra of PE layers has been developed for the case in which the reflectance factor of the underlayer r varies from zero to one. At present, we come across a case of r > 1. In this example, we are also interested in verification of the applicability of the developed approach for a new target that is a highly reflective powder. We use the same approach, as described in the All Spectra Together section. The sought spectrum s_4 is temporarily replaced by the measured target spectrum, and the weighting factors are evaluated using the formulae presented in Eqs. 21 and 22. Afterwards the estimated weighting factors c_1 , c_2 , and c_3 are compared with the



Figure 8. Example 2. The c-profiles (marks) and their models (curves): reflection (I, green), absorption (2, red), scattering (3, blue).

theoretical models, F(h,r), calculated using Eq. 2. Parameter r, which is the target reflectance factor, is evaluated by Eq. 18, where x_0 is obtained from the first spectrum measured through the PE pack of depth h = 0.046 mm. The established value is r = 1.13.

The results are shown in Figure 8: the dots show the estimates and the curves represent the theoretical models, F(h).

Here, we have also calculated the mean, μ , and the standard deviation, σ , values for each of the three error vectors, $\mathbf{e} = \mathbf{c} - \mathbf{F}$. The results are as follows: $\mu_1 = 0.01$, $\sigma_1 = 0.06$; $\mu_2 = 0.003$, $\sigma_2 = 0.04$; $\mu_3 = -0.001$, $\sigma_3 = 0.04$. In this case, the Student's *t*-test also confirms a good compliance. Therefore, we can conclude that the theory developed earlier⁷ is also applicable for objects with a high reflectance factor. The RMSE, d_0 , defined in Eq. 15, is $d_0 = 0.005$. This value will be used as a scaling factor in Eq. 17 to assess the quality of the target spectrum reconstruction.

Single Spectrum Through Many Layers

In this subsection, we consider a reconstruction of the target spectrum using a single spectrum obtained through multiple PE layers. In every case, i.e., when the measurement is conducted through one, two, etc. layers, the target reflection factor, r, is assessed individually using Eq. 18.

The target spectrum is evaluated using the same three methods employed in the first example. The KD method utilizes the known depth of the PE layer. In the OD method, the PE depth, h, is considered unknown, and it is assessed using the variation optimization method presented in Eqs. 19 and 20. These two techniques use direct calculations given in Eqs. 10 and 11 for spectrum recovery. The third method utilizes the thin layer (TL) concept that is given in Eq. 14.



Figure 9. Example 2. The target (0, red) and its estimates through eight PE layers: KD method (1, blue), OD (2, green), and TL (3, brown). The challenging range is shown in the insert.

KD OD TL Layers С С С Sample no. d d d I 1.502 0.982 1.167 0.989 1.071 0.991 2 2.057 0.966 1.140 0.990 1.021 0.992 3 1.845 0.973 1.318 0.986 1.285 0.987 4 1.700 0.977 1.437 0.984 1.470 0.983 5 1.472 0.983 1.539 0.981 1.583 0.980 6 1.765 0.975 1.680 0.977 1.865 0.972 7 2.037 0.956 0.967 1.549 0.981 2.341 8 2.119 0.964 1.607 0.979 3.273 0.911 9 2.002 0.968 2.025 0.967 4.149 0.853 10 2.274 0.958 2.430 0.952 4.826 0.794

Table 2. Example 2. Quality characteristics of the target spectrum recovery.

Values given in bold represent the best solution among all methods.

Figure 9 presents the results obtained from all these methods for a case of eight PE layers. In this example, we see that the range $5600-5900 \text{ cm}^{-1}$, which is shown in the insert, is again the most challenging area. All the predicted spectra have an artificial peak near 5778 cm^{-1} . The peak is negative for the TL recovered spectrum. This effect is associated with imperfection in modeling of the sharp PE absorption/scattering peaks. The origins of this artifact have been explained in the Case Study I: The NIST Traceable Standard. section.

Numerical results are given in Table 2, structured in the same way as Table I. Here, it is apparent that the quality of reconstruction d yields worse results than in the first example. This happens due to a higher accuracy d_0 in modeling of the entire data set. In the first case the RMSEP was 0.02, whereas here it is four times smaller, namely, 0.005. Therefore, all samples demonstrate d > I. Another conclusion is that all the methods are rather competing. The TL approach is expectedly better for the shallow PE depth. The OD approach serves best at the middle range of depths, and the KD method is better when applied to thick PE layers.

The target reflectance factor, r, calculated by the KD and OD methods for different number of the PE layers, is shown in Figure 10 (left, red axis). The layout of the figure is similar to Figure 6.

Dots Ia in Figure 10 stand for the KD method and the diamonds (1b in Figure 10) represent the OD method. The red line (1c in Figure 10) shows the reference value r = 1.13 found in the All Spectra Together section. The assessed mean reflectance factors are $r_{\rm KD} = 1.08 \pm 0.04$, $r_{\rm OD} = 1.08 \pm 0.03$. The values of the PE depth, *h*, which are evaluated using the OD method, are shown in Figure 10 in blue; they relate to the right axis. In this case, the estimated depths are very close to the known values presented by line (2a in Figure 10), which has an equation



Figure 10. Example 2. The target reflectance factor, r (left axis, in red): KD method (dots, 1a), the OD method (diamonds, 1b); the reference value can be found in the All Spectra Together section (line, 1c). The PE depths, h (right axis, in blue): the known values (line, 2a), the OD method (triangles, 2b).

 $h = 0.046 + 0.087 \cdot n$, where *n* is the number of layers. In this case the spectrum of the target object has a very high apparent reflectance factor, but the spectrum is not as complex as in the first example. Thus, the additional artifacts from the PE spectrum are better detected by the V(h) criterion, resulting in a good estimation of *h*.

Conclusions

In the first part of this research we proposed a three-component model of the optical properties of a multilayer PE cover. This model was developed using the multivariate analysis of the reflectance spectra obtained for reference objects that have flat spectra. In this paper, we demonstrated that this model is applicable when the underlayer is a real-world object with a complex NIR spectrum. The proposed model has two nonlinear parameters. They are the underlayer reflectance factor, r, and the depth of the PE cover, h. We have tested that the values of r and h computed by the developed empirical functions are in good agreement with the experimental values.

The following procedure is developed for a practical application of recovery of DR spectra of an unknown target object. At the first step, the reflectance factor of the target, r, is determined using the function given in Eq. 18. At the second step, the PE depth h is calculated. This procedure is not very straightforward and requires the solution to a simple optimization problem given in Eq. 20. At the third step, the MCR procedure with three fixed pure spectra, which characterize PE, and one unknown spectra of the target, is used. In practice, the recovery of the DR spectra of two different target objects covered with PE

layer of various depths is performed. The results show that the target spectra can be reconstructed rather accurately when the PE depth is not greater than 0.7 mm. The simplified thin layer model works reliably when the PE depth is less than 0.4 mm. However, the methods were unable to remove the contribution of sharp and narrow PE peaks in the 5600–5900 cm⁻¹ range completely. Some artifacts remained in the predicted spectrum. This shortcoming is partly caused by local imperfection of the multilayer model and it can probably be corrected in future by introducing a fourth component into the model. However, even the three-component approach for modeling of the optical properties of the cover layer provides satisfactory results and can be used for reconstruction of DR spectra of objects, covered by absorbing and scattering layers.

Conflict of Interest

The authors report there are no conflicts of interest.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Supplemental Material

Supplementary material for this paper can be found at the journal website: http://journals.sagepub.com/home/asp.

References

 F. Pi, H. Shinzawa, Y. Ozaki, D. Han. "Non-Destructive Determination of Components in Processed Cheese Slice Wrapped with a Polyethylene Film Using Near-Infrared Spectroscopy and Chemometrics". Int. Dairy J. 2009. 19: 624–629.

- A. Saleem, C. Canal, D.A. Hutchins, L.A.J. Davis, R.J. Green. "NIR Spectroscopy with Multivariate Calibration and Lock-In Amplification to Detect Chemicals Concealed Behind Fabrics"Proc. SPIE 8018, Chemical, Biological, Radiological, Nuclear, and Explosives (CBRNE) Sensing XII, 801811 (June 03, 2011); doi:10.1117/12.883839.
- A.N. Bashkatov, E.A. Genina, V.V. Tuchin. "Optical Properties of Skin, Subcutaneous, and Muscle Tissues: a Review". J. Innovative Opt. Health Sc. 2011. 4(1): 9–38.
- Y.L. Yang, L. Shoer, O.O. Soyemi, M.R. Landry, B.R. Soller. "Removal of Analyte-Irrelevant Variations in Near-Infrared Tissue Spectra". Appl. Spectrosc. 2006. 60(9): 1070–1077.
- J. Zieba-Palus, R. Borusiewicz. "Examination of Multilayer Paint Coats by the Use of Infrared, Raman and XRF Spectroscopy for Forensic Purposes". J. Mol. Struct2006. 792–793: 286–292.
- C. Conti, C. Colombo, M. Realini, G. Zerbi, P. Matousek. "Subsurface Raman Analysis of Thin Painted Layers". Appl. Spectrosc. 2014. 68(6): 383–387.
- A.L. Pomerantsev, O.Y. Rodionova, A.N. Skvortsov. "Diffuse Reflectance Spectroscopy of Hidden Objects. Part I: Interpretation of the Reflection-Absorption-Scattering Fractions in NIR Spectra of Polyethylene Films". Appl. Spectrosc 2017.
- Middletonspectral.com. Hyperspectral Imaging Cameras and Systems | Middleton Spectral Vision. http://www.middletonspectral.com/ [accessed 29 May 2016].
- A. de Juan, R. Tauler. "Multivariate Curve Resolution (MCR) from 2000: Progress in Concepts and Applications". Crit. Rev. Anal. Chem. 2006. 36: 163–176.
- A. de Juan, S.C. Rutan, M. Maeder, R. Tauler. "Multivariate Curve Resolution". In: D. Brown, R. Tauler, B. Walczak (eds) Comprehensive Chemometrics. Vol. 2, Amsterdam: Elsevier, 2009325–344, 473–506.
- O.Y. Rodionova, Y.V. Sokovikov, A.L. Pomerantsev. "Quality Control of Packed Raw Materials in Pharmaceutical Industry". Anal. Chim. Acta. 2009. 642(1–2): 222–227.
- A.L. Pomerantsev, O.Y. Rodionova. "Concept and Role of Extreme Objects in PCA/SIMCA". J. Chemom. 2014. 28: 429–438.
- W.G. Haanstra, W. Hansen, M.J.G. Huys, B.J. Kip, P. Palmen, et al. "Information Depth of Vis-NIR Light in Polyethylene Films Using Transmission and Reflectance Measurements". Appl. Spectrosc. 1998. 52(6): 863–868.