

## Plastic identification by remote sensing spectroscopic NIR imaging using kernel partial least squares (KPLS)

W.H.A.M. van den Broek<sup>a,\*</sup>, E.P.P.A. Derks<sup>a</sup>, E.W. van de Ven<sup>a</sup>, D. Wienke<sup>a</sup>,  
P. Geladi<sup>b</sup>, L.M.C. Buydens<sup>a</sup>

<sup>a</sup> *Laboratory for Analytical Chemistry, Faculty of Science, Catholic University of Nijmegen, Toernooiveld 1, 6525 ED Nijmegen, Netherlands*

<sup>b</sup> *Department of Organic Chemistry, Research group of Chemometrics, Umeå University, S-90187 Umeå, Sweden*

Received 1 March 1996; revised 3 July 1996; accepted 29 July 1996

---

### Abstract

This work describes the application of partial least squares (PLS) modeling in data reduction purposes for the classification of spectroscopic near infrared (NIR) images. Given multi-dimensional images (i.e.  $p$  images taken at  $p$  different wavelengths regions in the NIR-range), PLS projects the (nearly void) high dimensional space into a low dimensional latent space using the coded class information of the sample objects. Hence, PLS can be considered as a supervised latent variable analysis. In addition, data reduction by PLS increases the speed of on-line classification which is attractive in, e.g., process control. In order to apply these conditions on imaging problems a rapid PLS version, kernel PLS, is investigated. Emphasis is put on the performance of PLS as a supervised data decomposition technique for the classification of collinear image data, applied on a real world application. This application entails the discrimination between the materials plastics, non-plastics and image backgrounds.

*Keywords:* Kernel PLS; NIR imaging; Multivariate image analysis

---

### 1. Introduction

#### 1.1. Introduction to PLS

Partial least squares (PLS) has become an accepted and well understood method for modeling linear relationships in multivariate data. However, the problems attained so far have generally focused to multivariate calibration [1–3]. PLS provides an excellent means to enhance the predictability of the

model by finding the optimal compromise between bias and variance of the PLS estimates. Although the underlying PLS-algorithms show a resemblance to the methods used for multivariate classification (e.g. canonical correlation (CC), linear discriminant analysis (LDA)), so far little theory for classification analysis has been derived for PLS. Since PLS decomposes the input-data using the output-data in a supervised way (i.e. the coded classification membership directs the decomposition of the input space), the method provides some interesting properties suitable for classification purposes [4].

---

\* Corresponding author. Fax: +31-24-3652653.

Classification of image objects based on Mahalanobis distances is performed in the latent (score) space where the most relevant physical or chemical information is contained. Especially in a high dimensional space containing collinear data, this approach is highly efficient considering the practical unacceptable computational cost of calculating distance matrices. It is known that NIR-spectra are likely to contain collinear data since physical and chemical phenomena manifest themselves at various places in the infrared region [5]. Especially in this particular situation, PLS provides estimates with a low variance in contrast to the ordinary least squares approach where (in the collinear data case) the nearly singular covariance matrix has to be inverted. More advanced PLS-algorithms such as kernel PLS are also able to deal with large data in a very efficient, fast and elegant way [6,7].

The properties described are relevant in situations where models have to be created, validated and used in on-line or at-line situations. Since a straightforward classification approach like statistical pattern recognition techniques puts some practically unfeasible constraints on the data used (e.g. assumptions of normality) a more pragmatic classification method, based on a similar soft-modeling approach, is described.

### 1.2. Real world application

Since the last decade, legislation is forcing producers and distributors to become more and more responsible for their products, even to their final disposal [8]. Besides prevention and reduction of waste at their sources, the increased use of recycling and re-use of materials is being under investigation. An extra problem arises for recycling plastic fractions, because these are often not compatible due to different physical properties, which in particular affects the mixability and the thermal stability. In order to obtain a valuable recovery of plastics, a sorting system is needed which does not need any pretreatment of the waste entities, which is able to discriminate between the materials to be identified and which is rapid, accurate, robust, safe and easy for operators to work with.

Several studies [9–13] have shown that NIR spectroscopy is capable to meet the aforementioned sort-

ing requirements for plastics. One of these studies is performed by SIRIUS<sup>1</sup>, a cooperation project between the Laboratory for Analytical Chemistry of the University of Nijmegen, The Netherlands and the Institute for Chemical and Biochemical Sensor Research in Münster, Germany. Within this project, the present authors have developed an identification system to sort plastics from non-plastics which is based on NIR imaging spectroscopy [14,15].

### 1.3. Spectroscopic NIR imaging

The near infrared (NIR) region of the electromagnetic spectrum spans the region between 750–2500 nm. In NIR spectroscopy, the absorption of light from this part of the electromagnetic spectrum is measured. This is done by illuminating the material under investigation with NIR radiation. Waste objects are then allowed to absorb their fraction of material specific NIR radiation. The remaining reflected photons are collected and passed through a wavelength selection device which is transparent for preselected wavelength regions only. The selection of these regions is based on prior information about the material being studied and is described in van den Broek et al. [16]. These preselected reflected photons are eventually measured by a detector. As there are many detector materials available which can measure in different modes [17], the newest sort of detectors in NIR spectroscopic imaging are the focal plane array (FPA) detectors [18]. In contrast to conventional detectors, these FPA detectors individually consist of thousands of sensing elements called pixels. Each pixel can integrate light that has been reflected by a different part of the material surface. In this way, geometric material information can be obtained which gives information about the composition of the material surface. The described experimental setup will be used to classify the measured municipal waste materials into the classes plastic and non-plastic by using the material information per detector pixel. A main problem in spectroscopic imaging, also mentioned by other authors [19,20], is to deal with the large amount of data that is generated by these detection systems. Therefore, research is necessary in mathematical pro-

<sup>1</sup> SIRIUS stands for Sensors and artificial Intelligence for Recognition and Identification of Used plasticS.

cessing of these data to extract only the most relevant data with a minimum of effort.

#### 1.4. Summarized

The performance of PLS as a data reduction technique for the classification of collinear image data is investigated. PLS performs a data decomposition where pre-knowledge about the material composition is taken into account. Once the optimal PLS model is calculated, it has to be tested by offering unknown images to the model. Classification is then performed by using a distance measure (Mahalanobis and Euclidian distance measures). The strategy of this investigation includes the preparation of a representative calibration and test set from images, measured from the waste objects. These two datasets are used to optimize the PLS model. The classification results are compared with the classification performance of principal component analysis (no pre-knowledge used in the data decomposition) and with classification of the original data (without data reduction). Finally, the optimized model is used to classify complete three-dimensional stacks of raw images.

## 2. Theory

### 2.1. PLS modeling

Partial least squares regression (PLS) has generally been accepted as a powerful statistical analysis tool. In the last decade, a considerable amount of papers emerged [1–3] providing the method with a solid theoretical foundation. The PLS algorithms have originally been developed as a reduced rank regression method in order to stabilize its estimates. The original formulation of standard PLS can be recognized by the decomposition of the data matrices  $\mathbf{X}$  and  $\mathbf{Y}$  in order to produce successive orthogonal scores  $\mathbf{t}$  of the  $\mathbf{X}$ -space, commonly performed by the NIPALS algorithm which equations are given by

$$\mathbf{w}_i = \mathbf{X}^T * \mathbf{u}_i \quad (1)$$

$$\mathbf{t}_i = \mathbf{X} * \mathbf{w}_i \quad (2)$$

$$\mathbf{q}_i = \mathbf{Y}^T * \mathbf{t}_i \quad (3)$$

$$\mathbf{u}_i = \mathbf{Y} * \mathbf{q}_i \quad (4)$$

It should be noted that  $\mathbf{w}_i$  and  $\mathbf{q}_i$  are normalized in each step prior to further processing. In cases when the number of objects is much larger than the number of variables, the criss-cross projections of NIPALS become very inefficient. Lindgren [6] proposed a very efficient and elegant alternative for the deflation of the  $\mathbf{X}$  and  $\mathbf{Y}$  matrices into latent variables by means of the condensed covariance matrices  $\mathbf{X}^T \mathbf{X}$ ,  $\mathbf{X}^T \mathbf{Y}$  and  $\mathbf{Y}^T \mathbf{Y}$ . Since only small covariance matrices are used to update successive principal components, a significant gain in computation speed is obtained. For obvious reasons the method is referred to as Kernel PLS. The PLS loadings can directly be calculated from the covariance matrices, as is extensively described in [6,7]. A clear regression

$$\hat{\mathbf{Y}} = \mathbf{X} * \mathbf{B} \quad (5)$$

$$\mathbf{Y} = \mathbf{X} * \mathbf{B} + \mathbf{F} \quad (6)$$

can be derived by expressing the loadings of PLS as

$$\mathbf{B} = \mathbf{W}(\mathbf{P}^T * \mathbf{W})^{-1} * \mathbf{Q}^T \quad (7)$$

A more detailed description of the composition of data matrices  $\mathbf{X}$  and  $\mathbf{Y}$  is given in Section 4.

### 2.2. PLS in multivariate NIR-imaging

Multivariate NIR-images yield data which match the situation described above. To illustrate matters, a real world example from SIRIUS is given:

A typical dataset for multivariate NIR-image analysis is an unfolded three-dimensional stack of images ( $m, n, p$ ) into a two-dimensional ( $m * n, p$ ) data matrix<sup>14</sup>, where  $m$  (64) and  $n$  (64) represent detector dimensions and  $p$  (6) the number of interference filters used in the experimental setup. Clearly, the update of the  $6 * 6$  kernel or covariance matrices requires substantially less computation time than the original NIPALS computations. Even in the case of the 6 wavelengths in the NIR region an expensive time price has to be paid since each wavelength has to be selected by means of a rotating filter wheel causing delays in the (on-line) classification process. Obviously, the classification speed is proportional to the number of filters. The need for reduction of the number of collinear filters from a non-mathematical point of view is evident.

The PLS-model is obtained by means of cross-validation, i.e. internal calibration and test data are generated in order to estimate the optimal number of latent variables. For extensive theory on cross-validation, the reader is referred to [21,22].

### 2.3. Classification of images

The first scores vectors in the (smaller) latent space calculated by projecting the  $\mathbf{X}$ -matrix to the transposed loading matrix as defined in Eq. (2) contain the most relevant physical or chemical information and can be used to classify the image objects by means of the Mahalanobis distance ( $\mathbf{D}^2$ ) measure defined by equation  $(\mathbf{X}_i - \mathbf{X}_j)' \mathbf{S}^{-1} (\mathbf{X}_i - \mathbf{X}_j)$ , whereas  $\mathbf{S}$  represents the pooled within the covariance matrix and  $i$  and  $j$  denote the indices of the objects [23]. More details about the use of Mahalanobis distances can be found in the Section 4. The  $\mathbf{D}^2$  matrix is an extension of the Euclidian distance measure accounting for correlations that might exist between the variables. These correlations are largely present in the presented application. The final output of a classification of a complete three dimensional stack of images is a two dimensional matrix, where each element contains a coded representation of the material class.

The PLS-estimates ( $\hat{Y}$ , representing binary code) can also be used as an indication to identify classes. However, since PLS is a biased estimator, systematic errors are to be expected using the code space ( $\hat{Y}$ ). Moreover, the number of coding bits relative to the number of latent variables are also examined on their classification performance. In this work most emphasis is given to the classification of image objects in the latent score space ( $\mathbf{T}$ , see Eq. (2)), computed in a supervised way, by means of the Mahalanobis distance measure.

## 3. Experimental

### 3.1. Experimental setup

The experimental macro setup for plastic identification in municipal waste is shown schematically in Fig. 1. An infrared light source is used to illuminate the samples positioned on e.g. a conveyor belt. The

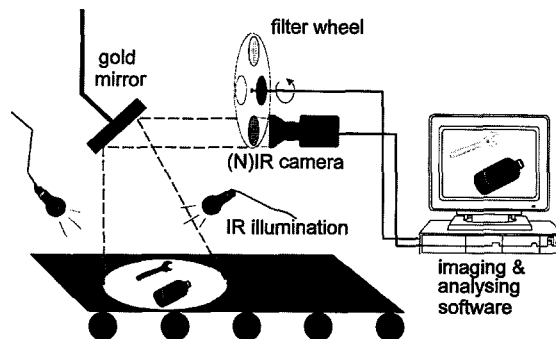


Fig. 1. The experimental setup for material identification.

total distance between sample and detector is about 2 m (remote sensing mode), and therefore a powerful light source is required to get sufficient photons on the detector for a safe identification of the samples. The reflected radiation from the samples is passed through 6 interference filters which are transparent in the following wavelength regions: 1548–1578, 1545–1655, 1655–1745, 1700–2150, 2207–2321 and 2115–2550 nm, respectively. The remaining photons pass a standard 50 mm EFL  $f/1.8$  multiple element lens (Pentacon). This objective lens is transparent from 300 to 2700 nm<sup>2</sup>. As a result, the objective lens projects the photons on an InSb focal plane array (FPA) detector (Cincinnati Electronics, OH). The FPA detector contains 64 \* 64 pixels, and each of these is capable of registering incoming photons. Moreover, each image pixel exhibits a sensitivity range of 10 bits resulting in 1024 gray levels. The output of the camera is a 12 bit digital data port which is able to send 51.44 frames (complete images) per second. No framegrabber is used since the digital output of the camera is sent indirectly, via an external interface SC-01 (electronic buffer) and a high speed 16 bit S16D I/O interface (Engineering Design Team, OR), to the internal memory of the computer. The computer is a SUN SPARC 10 workstation containing 32 Mbyte RAM memory. Several materials have been investigated for background ma-

<sup>2</sup> Measurements taken from different objective lenses, performed by H.W. Siesler and students, NIR research group, University of Essen, Germany.

terial in the experimental setup. Sanded aluminum gave the highest contrast for macroscopic plastic samples.

### 3.2. Software and computations

The images are recorded and processed using the Khoros 1.0 software environment (The Khoros Group, University of New Mexico, Albuquerque, USA), with its accompanying graphical user interface CANTATA. Khoros can be used as a visual programming tool for software development in scientific image visualization. This package entails a library of over 260 routines to facilitate research in image processing, pattern recognition, remote sensing and machine vision.

An alternative way to register images is by direct data access using the S16D interface combined with a library of C procedures. The library enables real time image acquisition by self written software. The necessary acquisition software and external interface SC-01 were developed and installed by Starling Consultancy, Hengelo, The Netherlands. Matlab (The MathWorks, MA) was used for the PLS calculations and image representations.

### 3.3. Dataset

Objects or samples with known material composition are needed in order to calibrate (learn) the material properties by PLS. As a consequence, a number of waste samples are gathered from the household waste of the city of Nijmegen.

In total, 40 waste objects are collected composing 17 plastics (1 PVC, 2 PET, 4 PS, 4 PE, and 6 PP) and 23 non-plastics (4 ceramics, 5 metals, 3 glass, 5 paper, 5 wood and 1 fiber). The objects differed in size, color, composition, mechanical flexibility and thickness. Images are measured using 6 interference filters per object and corrected for the dark current and reference images as described in literature [14]. After preprocessing, 40 stacks of corrected images ( $64 \times 64 \times 6$ ) are obtained which are unfolded to 40 matrices of size (4096, 6). The  $i$ th row in the matrix is referred to as a mini-spectrum measured by the  $i$ th detector pixel for six different wavelength ranges.

## 4. Materials and methods

### 4.1. Training and test set

A complete three-dimensional stack of images (size:  $n = 64$ ,  $m = 64$ ,  $p = 6$ ) contains only information from a single (waste) object. Since each stack of images contains  $64 \times 64$  pixels, representing 4096 mini-spectra each of length  $p = 6$ , a huge amount of data is obtained for only one material sample. To include sufficient variation in material type for both the calibration and test set, many material samples need to be characterized by their image stacks. Therefore, a compromise is made between a representative variation in material type and the total size of training and test dataset. This is done by extracting a pre-defined number of material pixels (plastic or non-plastic) and background pixels from a series of image stacks. The pre-defined numbers are chosen such, that the calibration dataset contains about the same number for each material class. The pixels are selected by means of principal component analysis (PCA) [14]. Pixel selection by principal component analysis guarantees the selected homogeneity of the mini-spectra and is very easy to perform. The PCA selection is based on the spectral correlation of pixels in a score plot. A cluster in a score plot can be selected by mouse-assisted handling on a computer screen. Since there is little control about the number of pixels in a selected cluster of a PCA score plot, it is necessary to reduce the amount of selected pixels to a pre-defined number. This selection is done randomly. The final number of pixels in the calibration set (1250, 6) are: plastics: 400, non-plastics: 450, and background: 400, whilst those for the test set (1750, 6) are: plastics: 450, non-plastics: 700, and background: 600.

### 4.2. PLS modeling

Before each measured image pixel can be classified, a PLS model has to be constructed. The PLS model transforms the image pixel to the score and code space. It should be noted that in both cases classification can be performed. Classification in the score space is further called PLS-score, whereas classification in the code space is called PLS-code.

Table 1

Calibration dataset: X-block containing reflectance values in au. Pixels were extracted from 22 stacks of images

# pixels	Materials	Filter 1	Filter 2	Filter 3	Filter 4	Filter 5	Filter 6
1	plastic	...	...	...	...	...	...
...		...	...	...	...	...	...
400		...	...	...	...	...	...
401	non-plastic	...	...	...	...	...	...
...		...	...	...	...	...	...
850		...	...	...	...	...	...
851	background	...	...	...	...	...	...
...		...	...	...	...	...	...
1250		...	...	...	...	...	...

The model is constructed using a calibration set, Tables 1 and 2. The calibration dataset consisted of NIR data of known material type (plastic, non-plastic and background), called X-block, and a coded representation of the material type, coding block Y. Each row of the X-block contains six NIR gray values (mini-spectrum) of a specific geometric position inside the camera view, measured for six different wavelength regions. The X-block is a collection of material pixels, extracted from a large number of measured images. The rows of the coded Y-block represent the three corresponding material types of the data block. Both blocks can significantly influence the model. The X-block must be without falsely measured data and must be a good representation of the three classes. The model was finally tested with an independent test set of known composition.

The choice of the coding block influences the PLS model through the choice of the type of coding, e.g.

Table 2

Calibration dataset: Y-block coding using three bits. Pixels were extracted from 22 stacks of images

# pixels	Materials	Bit 1	Bit 2	Bit 3
1	plastic	1	0	0
...		1	0	0
400		1	0	0
401	non-plastic	0	1	0
...		0	1	0
850		0	1	0
851	background	0	0	1
...		0	0	1
1250		0	0	1

1, 2, 3, ... bits coding, especially if the classification is performed in the coding space. The dimension of this space will be defined by the coding block. Also scaling the data can improve the classification. Scaling the data is often applied for creating a better classification model. For our problem, scaling techniques such as unit length scaling, range scaling, auto scaling and mean centering did not improve the model, in contrary, the classification results were even worse in comparison with the original data. One effect that caused this result was observed when the offset in the mini-spectra was eliminated by scaling. The reason why the offset in the mini-spectra contains useful classification information can be explained by the surface properties of the materials. Non-plastics such as glass, stone and metals have a low infrared absorptivity and therefore can reflect most illuminated infrared radiation. In addition, the surface of aforementioned materials was rather smooth which reduces the diffuse reflectance proportions.

#### 4.3. Classification with Mahalanobis distances

When the PLS model has been calculated, it can be used to classify unknown image pixels. An unknown image pixel is classified to the class membership  $c$  (plastic, non-plastic or background material), if the Mahalanobis distance between a PLS projected image pixel and the centroid in the projected space of class  $c$  is the smallest. In the classification of complete images, the same procedure can be applied, although now, the individual pixels in the image are classified in one step using matrix algebra. From the resulting classification matrix the total number of pixels from each material class needs to be calculated. The classification results can also be visualized by plotting them after range scaling, see Figs. 6 and 7.

## 5. Results and discussion

### 5.1. Cross-validation

The PLS model is dependent on the number of latent variables. This number can be estimated by a

cross-validation experiment. Cross-validation is performed by splitting the training dataset randomly into five groups of the same size. A PLS model is calculated from a reduced dataset where only one group is deleted. This model is used to predict the class membership of the objects in the deleted group. The differences between the actual class membership and the predicted ones are squared and summed to yield the PRESS value (predicted residual error sum of squares). This is repeated for each deletion group successively. All PRESS values are eventually summed to one overall PRESS. The whole cross-validation procedure is repeated 10 times (each time, the dataset is split into different groups) whereupon a mean PRESS is calculated. For each latent variable a mean PRESS value is calculated, see Fig. 2. All experiments are repeated with four different types of coding: one, two, three and four bit coding. A significant decrease in PRESS is observed when three latent variables are used. On the other hand, only a small decrease in PRESS is observed when four or more latent variables are used. Therefore, the optimal number of latent variables is chosen to be three.

### 5.2. Class coding

Now the number of latent variables is known, a PLS model is constructed where the influence of the coded Y-block is investigated. To illustrate this effect with an example, the material type of the Y-block

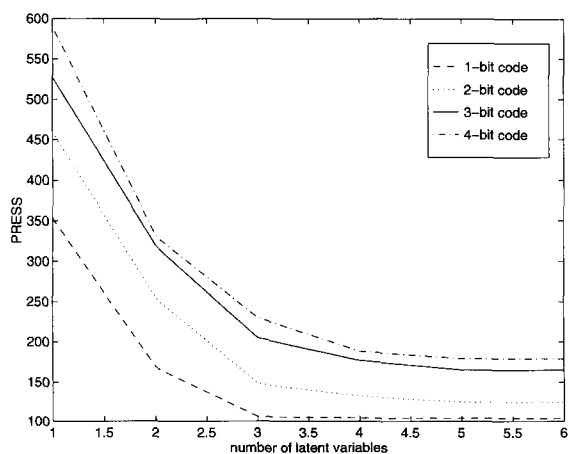


Fig. 2. Plot of the cross-validation experiments. The number of latent variables is depicted horizontally, the PRESS vertically.

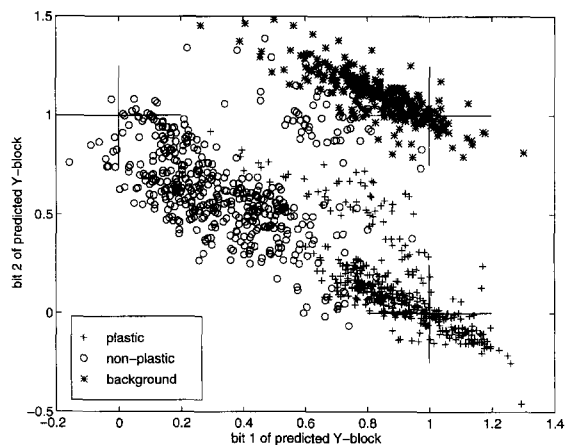


Fig. 3. Plot of the reproduced class memberships of the calibration set in the code space.

of the calibration set is coded with two bits: *10* to code for a plastic material, *01* for non-plastic and *11* for background. The prediction of the mini-spectra in the test set by the PLS model can be seen in Fig. 3. This figure shows that PLS does not exactly reproduce the original codes. Although the background pixels are clearly separated from the object pixels, the latter have little overlap with respect to each other. Furthermore, the clusters in Fig. 3 show a bias, especially the non-plastic class, which is the intrinsic property of PLS [2]. The bias is the deviation between the predicted mini-spectra in the code space and the original code. The non-plastic class was originally coded as *01*, but the PLS model was not able to reproduce this. Further investigation showed that this effect could not be caused by the bit order of the Y-block. Exactly the same predictions were obtained when plastic was coded as *11*, *01* or *00*, taken into account that the corresponding other materials were changed also.

Fig. 4 shows the optimization of the number of coding bits in the Y-block belonging to the training set against the classification error in the test set. This is done for a classification in the code space as well as in the score space. When using a one bit code, as in PLS-1, the total classification error in the code space is more than 12%, whereas a total error of about 4% is obtained when the classification is performed in the score space. The difference in classification error can be caused by the fact that the corresponding

coding space is only one-dimensional whereas the score space is three-dimensional. Probably a projection from a three-dimensional score space to a one-dimensional code space results in a higher classification error. When using a two bit code, the difference between the total errors of the score and code space is decreased as the difference in dimension has decreased. Here, a projection from a three-dimensional score space to a two-dimensional code space results in a lower classification error. If the number of coding bits is equal to the number of latent variables, the error of classification is of the same order. The similarity in the error using a three bits code is obvious because the dimensionality of both the score and the code space are the same. The estimates in the code space are then rotated scores. In the four bits code, the dimension of code space is one larger than that of the score space. The code space is obtained by the projection of the score space in this code space, which leaves the discrimination ability the same. Fig. 4 shows that the lowest total error is obtained using the three bits code in the score and the coding space.

To summarize the optimal PLS conditions for the identification of plastic waste, the model should contain three latent variables and a three bit code block. The best classification is performed either in the score and code space of the PLS model. It should be noted that this comparison is only valid when the correct number of coding bits is used.

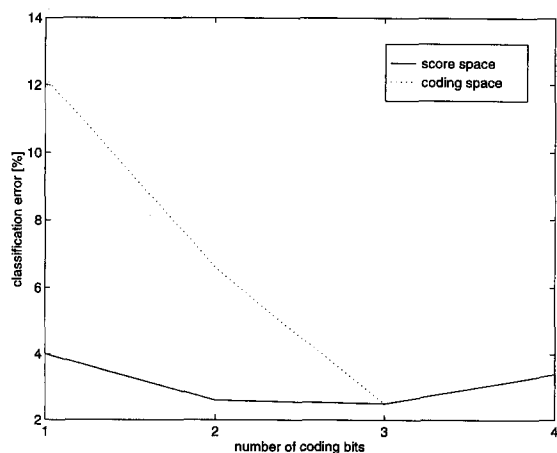


Fig. 4. Optimization of the number of coding bits. This figure shows the influence of the number of coding bits versus the classification error. This is done for a classification in the code space as well as in the score space.

Table 3

$\alpha$ ,  $\beta$  and total error in percentages and number of pixels respectively, calculated for the test set

	Comparison of classification errors			
	PLS-score	PLS-code	PCA	original data
$\alpha_{\text{plastics}}$	0.4 (2)	0.4 (2)	0.7 (3)	0.4 (2)
$\alpha_{\text{non-plastics}}$	4.3 (30)	4.3 (30)	5.3 (37)	1.3 (9)
$\alpha_{\text{background}}$	2.0 (12)	2.0 (12)	2.2 (13)	4.5 (27)
$\beta_{\text{plastics}}$	2.5 (33)	2.5 (33)	1.4 (38)	2.8 (9)
$\beta_{\text{non-plastics}}$	1.0 (11)	1.0 (11)	1.4 (15)	2.8 (29)
$\beta_{\text{background}}$	0.0 (—)	0.0 (—)	0.0 (—)	0.0 (—)
Total percentage	2.5 (44)	2.5 (44)	3.0 (53)	2.2 (38)

### 5.3. Classification of the test set

Table 3 shows the classification results of the test set performed in the PLS space (score and coding) with classification performed in the PCA and original space. For a comparison of the classification techniques a zero-hypothesis ( $H_0$ ) was formulated for each material type:

$H_0$ : The classification result is material  $c$  if the Mahalanobis distance of this object pixel in the score or code space to the centroid of class  $c$  is the smallest, where  $c$  can be plastic, non-plastic or background.

The  $\alpha$  and  $\beta$  errors were calculated for each class, the  $\alpha$  error was defined as the rejection of  $H_0$  while it is true ('false negative'). The  $\beta$  error was defined as the acceptance of  $H_0$  while it is false ('false positive'). The  $\alpha$  values of the plastic class show that the amount of misclassified plastic pixels is very low (between 0.4 and 0.7%). The misclassification of the background and the non-plastic image pixels are somewhat larger but still acceptable. The  $\beta$  values show that no image pixel non-plastic or plastic are classified as background. There are image pixels plastic, non-plastic and background which are misclassified as plastic or non-plastic but the percentages are also acceptable low. A good insight in the goodness of the applied model (PLS, PCA or no data reduction model) can be given by the total errors of Table 3. The percentage total error of both classifications in the score space and the code space appeared to be the same when the optimal coding was used (three bits): in both cases 2.5%. When a PCA model is used, a larger error was expected in contrary to PLS (3.0% error), because the model was made unsuper-



vised and therefore gave less good results. Performing a classification in the original space gives the best result (2.2% error). The data reductive methods (PLS and PCA) lose information with this reduction and this leads to a (small) increase in the classification error of 0.3 and 0.7%, respectively. The gain of data reduction can be found in the noise reduction and computation time; although time is needed to reduce the image data, the overall computation time will be significantly decreased without much error increase.

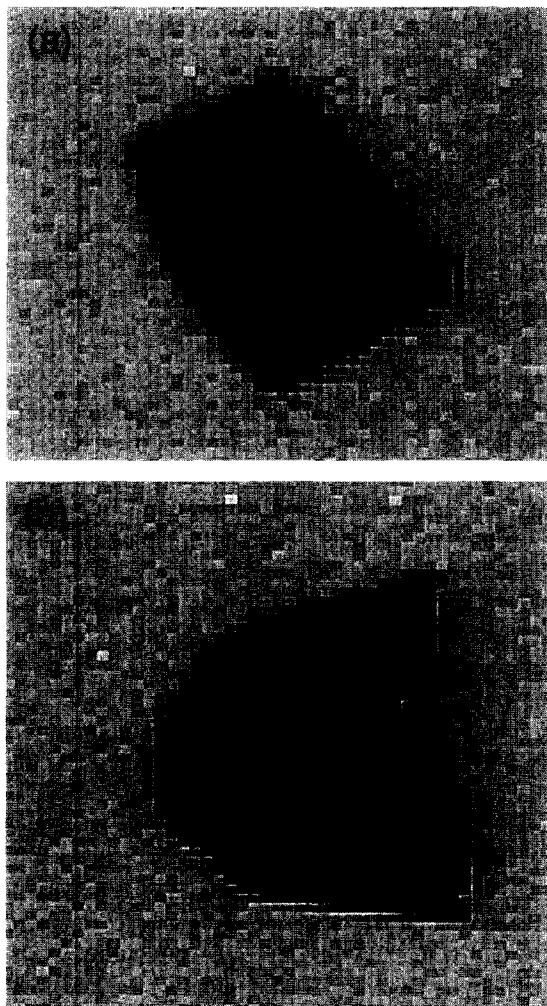


Fig. 5. Example of two waste materials, measured in the wavelength region between 1548–1578 nm. The image a is measured from a plastic object, whereas b is measured from a non-plastic object.



Fig. 6. Example of the same waste materials as in Fig. 5, but now classified by the calculated PLS model.

#### 5.4. Classification of real images

The PLS model has been tested with a pre-selected calibration and test set. In the following, the calculated optimized model is used to classify complete three-dimensional stacks of images into classification matrices. The PLS model has been constructed using three latent variables and a three bits code block. Images were measured from two different materials, a plastic (polyethylene) and non-plastic (ceramic) fragment. Two examples of the raw images are shown in Fig. 5. Fig. 6 shows the classification matrices of the

aforementioned materials. These were obtained by a PLS classification in the score space. This means that 4096 ( $64 \times 64$ ) mini-spectra had to be classified. Note that both images show falsely classified image pixels, especially around the objects. This can be due to mixed image pixels, object-edge effects and shadow. Mixed pixels occur in particular when low spatial resolution detectors are used. Reflected radiation overlaps with neighboring pixels, which causes mixed mini-spectra. Shadow effects occur when objects exceed a certain height [14]. The intensity of the radiation around the edges of the object will become less than elsewhere in the image. Object-edge effects oc-

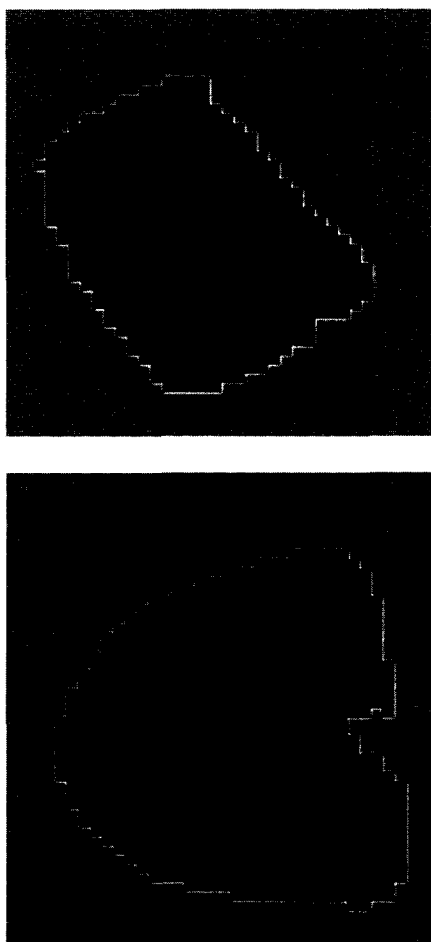


Fig. 7. Example of the same waste materials as in Fig. 5, but now classified by the calculated PLS model and filtered by a majority filter.

cur when images are not synchronized perfectly with respect to their spatial position in the image. Another measurement artifact that may occur are mirror reflections. These occur when the surfaces of the objects are so smooth that all illuminated radiation is reflected on the detector. This effect is expressed as spikes in the images. The problem of shadow can be solved by including shadow as an extra class and calibration of the model for this class, too. The other measurement artifacts can be reduced by applying a mathematical filter on the classification matrix. Information from neighboring pixels is used to reduce or eliminate these effects. When a majority filter [24] is applied on the new PLS model with the four classes (plastic, non-plastic, background and shadow) the classified images of Fig. 6 are shown in Fig. 7. The images in Fig. 7 show a low number of falsely classified pixels, which corroborates that the corresponding PLS model gave satisfactory results.

## 6. Conclusion

In the present work, we have shown that the kernel PLS is a viable tool for reducing collinear image data in order to classify large stacks of these data. The advantages of this technique with respect to imaging are short calibration stages during modeling and reduction of the number of latent variables which significantly increases the speed of image processing, especially, when new wavelength selection devices such as AOTF and interferometers are used, which can generate more data than the filter wheel used in this work. Moreover, spectral differences in the pixel mini-spectra contained sufficient information to identify 95% or more of the investigated materials: plastic, non-plastic and background material. The classification results for the wavelength reduced datasets are almost equal to those obtained with the original data. To emphasize this good result, the largest difference in total classification error was less than 1%, obtained with PCA. In this case, the dataset was reduced by a factor 0.5. Care should be taken in the modeling step where the number of latent variables and the number of coding bits have to be selected.

Spectroscopic NIR imaging seems to be a promising technique for material recognition. In combina-

tion with fast and robust chemometrical techniques, this imaging system may become mature for on-line process control. Quality control and process or reaction monitoring are then within reach.

## Acknowledgements

The authors gratefully acknowledge financial assistance from the Commission of European Communities for awarding Environment grant No. EVWA-CT-92-0001. Thanks to T. Huth-Fehre, R. Feldhoff, T. Kantimm (ICB, Münster, Germany) for their stimulating discussions.

## References

- [1] P. Geladi and B.R. Kowalski, Partial Least Squares Regression: A Tutorial, *Analytica Chimica Acta* 185 (1986) 1–17.
- [2] I.E. Frank and J.H. Friedman, A statistical view of some chemometrics regression tools, *Technometrics* 35 (1993) 109–148.
- [3] I. Helland, Partial Least Squares Regression and Statistical Models, *Scand. J. Stat.* 17 (1990) 97–114.
- [4] R. Vong, P. Geladi, S. Wold and K. Esbensen, Source Contributions to ambient aerosol calculated by discriminant Partial least squares regression (PLS), *J. Chemom.* 2 (1988) 281–296.
- [5] H. Martens and T. Naes, Multivariate Calibration by Data Compression, in: *Near-Infrared Technology in the Agricultural and Food Industries*, P. Williams and K. Norris (Eds.) (American Association of Cereal Chemists, 1990).
- [6] F. Lindgren, Third Generation PLS: some Elements and Applications, Thesis Umeå University, Research Group for Chemometrics, Umeå, Sweden, 1994.
- [7] S. de Jong and C.J.F. ter Braak, Short Communication: Comments on the PLS Kernel Algorithm, *J. Chemom.* 8 (1994) 169–174.
- [8] G. Holzhey, Das Duale System und seine Auswirkungen, *Papier* 47 (1993) 337–344.
- [9] H. Ritzmann and D. Schudel, Kunststoffabfälle automatisch sortieren, *Kunststoffe* 84 (1994) 582–584.
- [10] N. Eisenreich, Schnelle Identifikation von Kunststoffen mit schneller nahinfrarotspektroskopie, in: 23rd Int. Annual Conf. of ICT 1992: Waste management of energetic materials and polymers, June 30–July 3, 1992, Karlsruhe, Germany, pp. 59.1–59.12.
- [11] Th. Huth-Fehre, R. Feldhoff, Th. Kantimm, L. Quick, F. Winter, K. Camman, W. van den Broek, D. Wienke, W. Melssen and L. Buydens, NIR-Remote Sensing and Artificial Neural Networks for Rapid Identification of Post Consumer Plastics, *J. Mol. Struct.* 348 (1995) 143–146.
- [12] M.K. Alam, S.L. Stanton and G.A. Hebner, Near-Infrared Spectroscopy and Neural Networks for Resin Identification, *Spectroscopy* 9 (1994) 30–35.
- [13] D.M. Scott and R.L. Waterland, Identification of Plastic Waste Using Spectroscopy and Neural Networks, *Polym. Eng. Sci.* 35 (1995) 1011–1015.
- [14] W.H.A.M. van den Broek, D. Wienke, W.J. Melssen, C.W.A. de Crom and L.M.C. Buydens, Identification of Plastics among non-plastics in Mixed Waste by Remote Sensing Near-Infrared Imaging Spectroscopy. 1. Image Improvement and Analysis by Singular Value Decomposition, *Anal. Chem.* 67 (1995) 3753–3759.
- [15] D. Wienke, W.H.A.M. van den Broek and L.M.C. Buydens, Identification of Plastics among non-plastics in Mixed Waste by Remote Sensing Near-Infrared Imaging Spectroscopy. 2. Multivariate Image Rank Analysis for Rapid Classification, *Anal. Chem.* 67 (1995) 3760–3766.
- [16] W.H.A.M. van den Broek, D. Wienke, J.C.A. Kraak, W.J. Melssen, and L.M.C. Buydens, Filter selection by Genetic Algorithms for spectroscopic NIR Imaging, to be submitted.
- [17] E.N. Lewis, P.J. Treado, R.C. Reeder, G.M. Story, A.E. Dowrey, C. Marcott and I.W. Levin, Fourier Transform Spectroscopic Imaging Using an Infrared Focal-Plane Array Detector, *Anal. Chem.* 67 (1995) 3377–3381.
- [18] P.T. Treado, I.W. Levin and E.N. Lewis, Indium Antimonide (InSb) Focal Plane Array (FPA) Detection for Near-Infrared Imaging Microscopy, *Appl. Spectrosc.* 48 (1994) 607–615.
- [19] B. Novales, D. Bertrand, M.F. Devaux and P. Robert, Assessment of the quality of wheat milling products by fluorescence video imaging, *Int. Conf. on Quality Control by Artificial Vision*, May 17–19th, 1995, Le Creusot, France.
- [20] E.N. Lewis and I.W. Levin, Vibrational Spectroscopic Microscopy: Raman, Near-Infrared and Mid-Infrared Imaging Techniques, *JMSA* 1 (1995) 35–46.
- [21] S. Wold, Cross-Validatory Estimation of the Number of Components in Factor and Principal Components Models, *Technometrics* 20 (1978) 397–405.
- [22] D.W. Osten, Selection of optimal regression models via cross-validation, *J. Chemom. Int. Lab. Syst.* 2 (1988) 39–48.
- [23] W.R. Dillon and M. Goldstein, in: *Multivariate Analysis, Methods and Applications* (John Wiley and Sons, New York, 1984).
- [24] W.H.A.M. van den Broek, D. Wienke, E.W. van de Ven, W.J. Melssen, R. Feldhoff, T. Huth-Fehre, T. Kantimm and L.M.C. Buydens, Application of a Spectroscopic Infrared Focal Plane Array (FPA) Sensor for On-line Identification of Plastic Waste, to be submitted.