



## Spectral Signatures of Surface Materials in Pig Buildings

G. Zhang<sup>1</sup>; J.S. Strøm<sup>1</sup>; M. Blanke<sup>2</sup>; I. Braithwaite<sup>2</sup>

<sup>1</sup>Danish Institute of Agricultural Sciences, Research Centre Bygholm, Department of Agricultural Engineering, Schüttesvej 17, DK-8700 Horsens, Denmark; e-mail of corresponding author: [guoqiang.zhang@agrsci.dk](mailto:guoqiang.zhang@agrsci.dk)

<sup>2</sup>Section of Automation, Ørsted DTU, Technical University of Denmark, Elektrovej B 326, DK-2800 Lyngby, Denmark; e-mail: [mb@oersted.dtu.dk](mailto:mb@oersted.dtu.dk)

(Received 15 July 2005; accepted in revised form 3 May 2006; published online 3 July 2006)

Manual cleaning of pig production buildings based on high-pressure water cleaners is unappealing to workers, because it is tedious and health threatening. To replace manual cleaning, a few cleaning robots have been commercialised. With no cleanliness sensor available, the operation of these robots is to follow a cleaning procedure initially defined by the operator. Experience shows that the performance of such robots is poor regarding effectiveness of cleaning and utilisation of water. The development of an intelligent cleanliness sensor for robotic cleaning is thus crucial in order to optimise the cleaning process and to minimise the amount of water and electricity consumed. This research is aimed at utilising a spectral imaging method for cleanliness detection. Consequently, information on the reflectance of building materials and contamination in different spectral ranges is important.

In this study, the optical properties of different types of surfaces to be cleaned and the dirt found in finishing pig units were investigated in the visual and the near infrared (VIS–NIR) optical range. Four types of commonly used materials in pig buildings, *i.e.* concrete, plastic, wood and steel were applied in the investigation. Reflectance data were sampled under controlled lighting conditions using a spectrometer communicating with a portable computer. The measurements were performed in a laboratory with materials used in a pig house for 4–5 weeks. The spectral data were collected for the surfaces before, during and after high-pressure water cleaning.

The spectral signatures of the surface materials and dirt attached to the surfaces showed that it is possible to make discrimination and hence to classify areas that are visually clean. When spectral bands 450, 600, 700 and 800 nm are chosen, there are at least two spectral bands for each type of the materials, in which the spectral signals can be used for discrimination of dirty and clean condition of the surfaces.

© 2006 IAgrE. All rights reserved  
Published by Elsevier Ltd

### 1. Introduction

Within agriculture, manual cleaning of pig buildings is widely based on high-pressure water cleaners. The main purpose of the cleaning process is to reduce the risk of infection between batches of pigs. The cleaning process is unappealing to workers, however, because it is tedious and health threatening. The cleaning process contributes to deterioration of the working environment due to the stirring up of dirt and microorganisms, which is inhaled by the operator (Strøm *et al.*, 2003).

To replace manual cleaning, a few cleaning robots have been commercialised. Robotic cleaning often entails

subsequently manual cleaning because the robot could not detect the cleanliness of the surfaces since there is not a sensor available to do so. Further investigations have shown that the robot performance is poor regarding effectiveness and utilisation of water (Pedersen & Kai, 1998). The water consumption for robotic cleaning is up to 40% higher than what is used for manual cleaning. Therefore, it is crucial to develop an intelligent sensor for cleaning robots to detect the cleanliness of surfaces and subsequently minimise time for cleaning and the amount of water and electricity consumed by the cleaning robot.

A major issue is to find methods to discriminate effectively between dirt to be removed and background

Notation			
$D$	diagonal matrix	$\Pi; p$	positions, dimensionless
$P$	loading matrix	$\Theta; \theta$	cleanliness condition (clean; dirty)
$p$	loading vector	$\sigma$	singular value
$S$	spectral data matrix		
$s$	spectral data vector		
$T$	score matrix		
$U$	orthogonal matrix $k \times k$	<i>Subscript</i>	
$u$	column of $U$	$i; l$	index for wavelength
$V$	orthogonal matrix $l \times l$	$k$	index for measurement
$v$	column of $V$	$N; M$	index for measurement positions in clean and dirty conditions
$A; \lambda$	wavelengths, nm		

materials. Different approaches such as image analysis based on colour, texture or spectral properties of the surfaces and the residue on the surfaces could be applied for the sensor development. Our research is aiming at a possibility to utilise a spectral imaging method for cleanliness detection. Consequently, information on the reflectance of building materials and contamination in different spectral range is essential. The objective of this study was to investigate the optical properties in the visual and the near infrared (VIS–NIR) optical range of clean and dirty surfaces in a finishing pig unit.

## 2. Materials and methods

Samples of building materials in a finishing pig production building were used for studies on their spectral signatures in the VIS–NIR optical ranges. In practice, a number of different materials are used in pig housing. In this study, four commonly used materials were selected, namely concrete, plastic, wood and stainless steel. Concrete is widely used for solid and slatted floors and in some cases also for partition walls. Plastic is gaining popularity for partition walls at the expense of wooden partitions. Stainless steel is increasingly used for housing equipment such as drinkers and feeders.

### 2.1. Spectrometer

The spectrometer used was a diffraction grating spectrometer type EPP2000-VIS (StellarNet, USA, 2003) with a 2048 element charge-coupled device (CCD) detector that covered the spectral range from 400 to 1100 nm. It used a 10  $\mu$ m slit, giving a spectral resolution of 1.4 nm. The detector integration time was controlled by software in the range from 4 to 4000 ms. This range is equivalent to analogue to digital (A/D)

data acquisition rates of 500 down to 1 kHz. In this study, the integration time was set at 2000 ms. The 12-bit A/D converter gave a dynamic range from 1 to 4096 counts for any given detector integration setting. The signal-to-noise ratio was 1000:1. The spectrometer was connected to a portable computer *via* a USB-2 cable.

The light source was a Tungsten-Krypton lamp, type SL1 from StellarNet Inc., with colour temperature of 2800 K, suitable for applications in the range VIS–NIR from 350 to 1700 nm.

A Y-type armoured fibre optic reflectance probe was used to connect the light source and the spectrometer to the probe head. The probe contained six illuminating fibres around a single read fibre each being 400  $\mu$ m in diameter and specified for VIS/NIR applications.

The probe head was placed in the cavity of a sensor block as shown in *Fig. 1*. The sensor block was made of black plastic with dimensions 100 mm long, 50 mm wide and 37 mm high. The block had a 13 mm deep circular cavity with diameter 42 mm for the sensor head to avoid undesired background illumination effects. In the experiments, the sensor head was kept at 45° and at a distance of 6 mm to measured surfaces. A vertical sensor hole was covered with black tape to avoid external light effects.

### 2.2. Experimental design

Small test plates of the different materials were placed on the floor and on the partition walls in a pig pen. The dimensions of the plates were 250 by 400 mm<sup>2</sup>. After 4–5 weeks the plates were brought into the laboratory for investigation. *Figure 2* shows the photos of the four materials in wet condition before and after cleaning.

Spectral data were sampled for both dirty and cleaned plates in both dry and wet conditions, giving a total of 16 measurement setups, as listed in *Table 1*. The dry condition was defined as the plates were placed in the

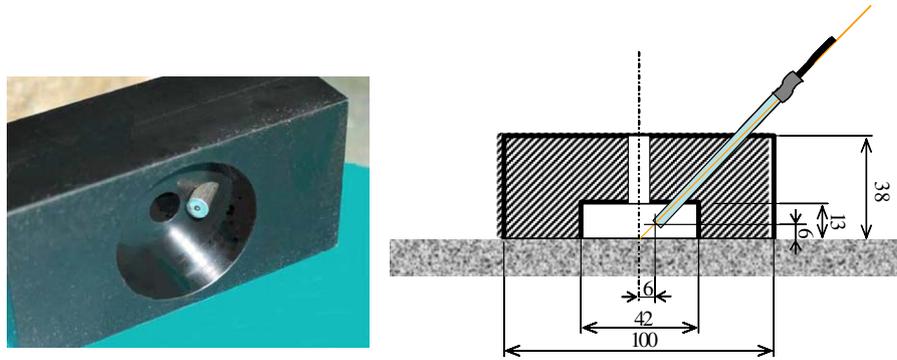


Fig. 1. View and dimensions of the sensor block; all dimensions in millimetres

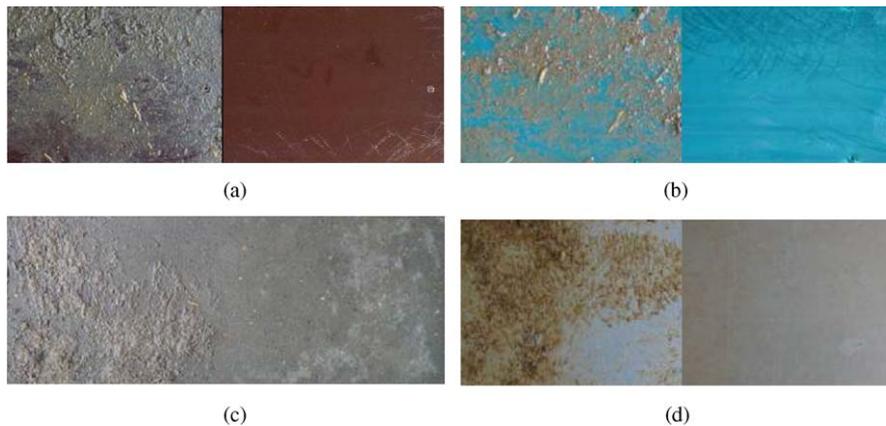


Fig. 2. Photos of the materials in wet condition before and after cleaning: (a), wood plate; (b), plastic plate; (c), concrete slab; (d), steel plate

laboratory room for 24 h after being brought back from the pig house or after cleaning. The wet condition was defined as the plate surfaces after water sprinkling before cleaning and immediate after final cleaning. During the measurement periods, the room air temperatures were 18–21 °C and the relative humidity were 40–55%.

For each measurement setup spectral data were sampled at 20 randomly selected positions on each test plate in order to include the effect caused by the non-homogeneous property of the surfaces. At each position, the reflectance data was recorded as an average of five scans with 2 s integration time to minimise signal noise. The values of the reflectance in the measured spectral ranges were used as the spectral signatures of the materials under the defined conditions.

### 2.3. Data analysis

The primary data of the reflectance measurements were explored *via* visual interpretation of the means and standard deviations for each experimental setup. In

Table 1  
Measurement set-ups in laboratory

Surface type	Clean dry	Clean wet	Dirty dry	Dirty wet
Concrete	x	x	x	x
Plastic	x	x	x	x
Wood	x	x	x	x
Steel	x	x	x	x

addition, principal component analysis (PCA) and analysis of variance (ANOVA) were used for spectral data analysis.

#### 2.3.1. Principal component analysis

The PCA analysis technique is commonly used to reduce dimensionality of data-sets with a large number of interdependent variables (Jackson, 1981). The objective is to provide the most compact representation of all the variation in a data matrix  $S$ . The technique summarises the original  $S$  variables into much fewer, more informative variables called scores  $T$ . These new variables (or scores) are weighted averages of the

original  $S$  variables. The weighting profiles are called loadings  $\mathbf{P}$ . For each score variable  $\mathbf{t} \in \mathbf{T}$ , the influence (weight) of the original variables is found in its corresponding loading profile,  $\mathbf{p} \in \mathbf{P}$ .

The measured data based on the experimental design described in Section 2.2 are arranged as data matrices for PCA. The measurements show reflectance as functions of wavelength and cleanliness conditions. A formal representation of a reflectance measurement is

$$s(i, k, \theta) = S(\lambda_i, p_k, \theta_k), \quad \lambda_i \in A, p_k \in \Pi, \theta_k \in \Theta \quad (1)$$

where  $S(\lambda_i, p_k, \theta_k)$  is the reflectance at wavelength  $\lambda_i$  in nm, measured at position  $p_k$  by the spectrometer. The parameter  $\theta_k$  describes the *a priori* condition as clean or dirty. The main hypothesis to be investigated is whether the spectral components have a unique relation with the parameter  $\theta$ . The sets  $A, \Pi, \Theta$  are wavelengths, positions and condition, respectively.

The reflectance vector obtained from a single position is defined by

$$\mathbf{s}_k(\theta) = \begin{cases} \mathbf{s}_k^c & \theta_k = \text{clean} \\ \mathbf{s}_k^d & \theta_k = \text{dirty} \end{cases} \quad (2)$$

and the sets of measurements taken at  $N$  different clean positions and  $M$  different dirty positions are arranged into a matrix

$$\mathbf{S} = [\mathbf{s}_1^c, \mathbf{s}_2^c, \mathbf{s}_3^c, \dots, \mathbf{s}_N^c, \mathbf{s}_1^d, \mathbf{s}_2^d, \mathbf{s}_3^d, \dots, \mathbf{s}_M^d]^T \quad (3)$$

Each of the  $l$  columns in this matrix represent a spectral line obtained from the spectrometer, the  $k$  rows ( $k = N+M$ ) represent the measurements. Applying singular value decomposition (SVD) to  $\mathbf{S}$  from Eqn (3),

$$\mathbf{S} = \mathbf{U} \cdot \mathbf{D} \cdot \mathbf{V}^T = \sum_{j=1}^{nm} \mathbf{u}_j \sigma_j \mathbf{v}_j^T \quad nm = \min(k, l) \quad (4)$$

where  $\mathbf{D}$  is the diagonal matrix that has the singular values of  $\sigma_j \geq 0$  on its diagonal, sorted in decreasing order, with  $\mathbf{u}_j$  the columns of singular vectors  $\mathbf{U}$  (the object space), and  $\mathbf{v}_j$  the columns of singular vectors  $\mathbf{V}$  (the variable space) and  $T$  is the transpose operator. The diagonal matrix immediately shows the structure of the matrix: a few large singular values in the upper part represent the main phenomena, while the (near) zero diagonal elements below are the noise dimensions.

In PCA,  $\mathbf{t}_j = \mathbf{u}_j \sigma_j$  are denoted the scores,  $\mathbf{p}_j = \mathbf{v}_j$  the loadings. Hence the SVD of Eqn (4) reads

$$\mathbf{S} = \mathbf{T} \mathbf{P}^T = \sum_{j=1}^{nm} \mathbf{t}_j \mathbf{p}_j^T \quad (5)$$

The scores,  $\mathbf{T}$  describe the basis vectors along which the original  $S$  variables show their largest variance. The columns in  $\mathbf{T}$  are orthogonal. When different rows in  $\mathbf{S}$  (measurements) are related to each other, the PCA will

project these along the basis vectors in  $\mathbf{T}$ . Loading  $\mathbf{P}$  is the weighting (influence) of the  $S$  variables on the scores  $\mathbf{T}$ . The  $\mathbf{t}_i, \mathbf{p}_i$  pairs are arranged in descending order according to the associated  $\sigma_i$ . These  $\sigma_i$  are a measure of the amount of *variance* described by the  $\mathbf{t}_i, \mathbf{p}_i$  pair. In this context, variance can be considered as *information*. As the  $\mathbf{t}_i, \mathbf{p}_i$  pairs are in descending order of  $\sigma_i$ , the first pair captures the largest amount of information of any pair in the decomposition.

The detailed techniques in PCA are well established (Jackson, 1981; Wold *et al.*, 1987; Wu *et al.*, 1997; Andersen *et al.*, 1999).

In the PCA method, the components are embedded in an abstract data space and each component is a linear combination of all original variables. Variables with a high degree of systematic variation typically have large absolute variances, and consequently large loadings. The loading is a measure of how much each of the original data contributed to each of the components. The principal component (PC) will account for a high percentage of the variation in the data. Generally, large loading values imply important variables, i.e. spectral bands. When assessing importance, it is mandatory also to consider the proportions of the total explained variance along each component (Esbensen, 2000). If a PC explains a significantly larger portion than others, then variables with large loading in this PC are also more important than those with large loading in other PCs.

### 2.3.2. Analysis of variance

In addition to characterising the spectral reflectance of the surface in dirty and clean conditions in the selected wavelength ranges by PCA, ANOVA was used to check in which wavelength range the reflectance data of dirty and clean surfaces were statistically significant different.

A normality test was performed using the Lillefors method (Conover, 1980) to check if the measured spectral data were normally distributed. Both the classical ANOVA method and a non-parametric ANOVA–Kruskal–Wallis method (ANOVA—analysis of variance) (Hollander & Wolfe, 1973) were applied for data analysis. The Kruskal–Wallis method is non-parametric and does not assume a normal distribution of data whereas this is the case for the classical ANOVA algorithm.

## 3. Results and discussion

The measured spectral signatures are shown as the average spectral data for the four selected materials before and after cleaning in both dry and

wet conditions. In this paper it is assumed that evaluation of cleanliness is performed after high-pressure water cleaning, where the surface is wet. Therefore, the data analysis is focused on the differences between clean and dirty surfaces in wet conditions only.

### 3.1. Spectral signatures

The primary results of the measured reflectance are shown in *Fig. 3*. The curves are average values of the measured data in the 20 random points for each measurement set-up.

#### 3.1.1. Dry clean versus dirty surfaces

The spectral ranges that may be used for classification of clean *versus* dirty surfaces in dry condition depend on background material. For concrete the reflectance was higher for clean surface within the whole wavelength region, but the difference tended to decrease with increasing wavelength, particularly above 700 nm [*Fig. 3(a)*]. For steel the reflectance was higher for the clean surface below 700 nm, above which the reflectance shifted to lower values for the clean surface [*Fig. 3(b)*]. For the brown wood there was no difference in reflectance up to 900 nm above which the reflectance from the clean surface increased more rapidly than from the dirty [*Fig. 3(c)*]. The reflectance from the green plastic plate varied in a peculiar way compared with the other three materials as the reflectance from the clean surface was higher at wavelengths below 600 nm, was similar to the dirty surface up

to 800 nm, above which the reflectance from the clean surface again became higher than from the dirty surface [*Fig. 3(d)*].

In dry conditions, wavelength regions can thus be identified where there is a noticeable difference between a clean and a dirty surface.

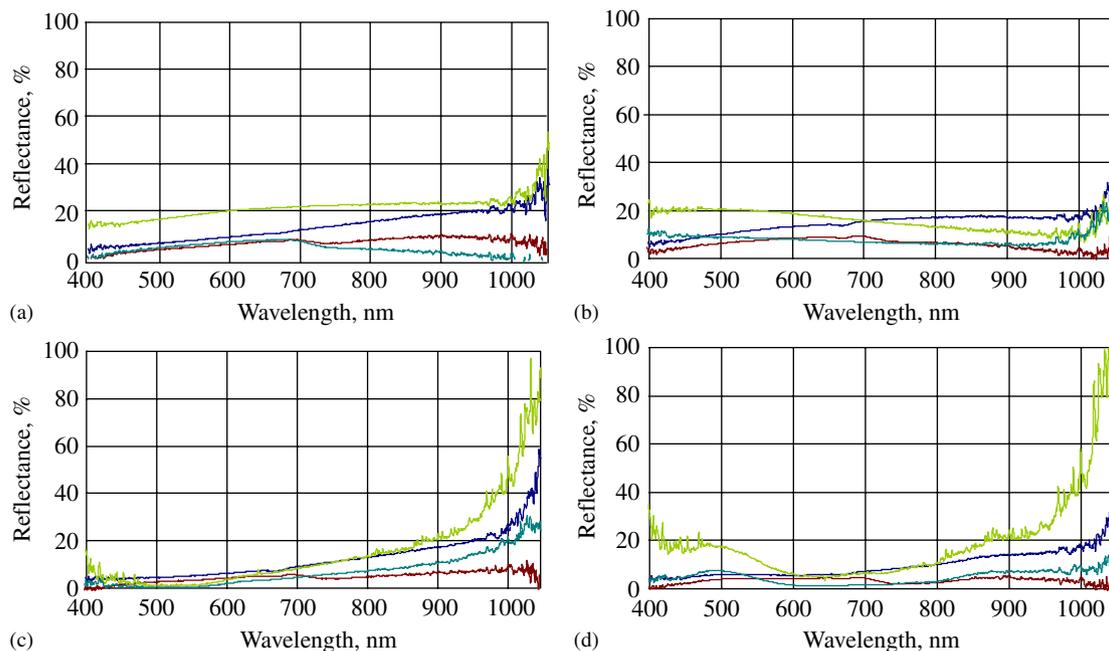
#### 3.1.2. Wet clean versus dirty surfaces

For wet concrete, a significant difference was seen in wavelength range above 750 nm [*Fig. 3(a)*]. For stainless steel, however, the differences were only seen in the wavelength ranges below 550 nm and above 950 nm with only slight differences in between [*Fig. 3(b)*]. For the brown wood plate [*Fig. 3(c)*], the reflectance for the clean–wet condition was slightly lower in the wavelength ranges 500–700 nm and higher above 750 nm compared with dirty–wet conditions. For the green plastic plate [*Fig. 3(d)*], the reflectance for clean–wet condition was higher for wavelengths below 550 nm and above 800 nm, but lower between 600 and 700 nm compared with dirty–wet condition.

Thus, within the investigated wavelength ranges there is also the potential for discrimination between clean and dirty surfaces in wet condition. The summarised spectral signatures for the surfaces in wet condition are listed in *Table 2*.

#### 3.1.3. Standard deviation

*Figure 4* shows the measured reflectance data in the 20 measurement points and the standard deviations for the



*Fig. 3.* Average spectral data for the four selected materials before and after cleaning in dry and wet condition: (a) concrete slab; (b) steel slab; (c) brown wood plate; (d) green plastic plate; — dirty dry; — dirty wet; — clean dry; — clean wet

concrete slab and the brown wood plate in wet conditions. It is seen that the standard deviation for the dirty condition was considerably larger than clean condition, especially in the NIR. A probable reason is the less homogeneous texture of the dirt attached on the surface due to different thickness and smoothness. A large standard deviation may increase the difficulty in discriminating between clean and dirty areas.

The standard deviations for the concrete slab and the wood plate were different at different wavelengths. The standard deviations for the painted brown wood plate were much lower in entire wavelength range 400–1000 nm compared with the concrete slab. A

surface of painted wood thus appeared more homogeneous than the concrete slab.

### 3.2. Statistical analysis of spectral data

Since the spectral data has lower signal/noise (S/N) ratio in the ranges below 450 nm and above 800 nm, the analysis was focused on the signal within this range for discrimination of clean and dirty surfaces.

#### 3.2.1. Principal component analysis of spectral data

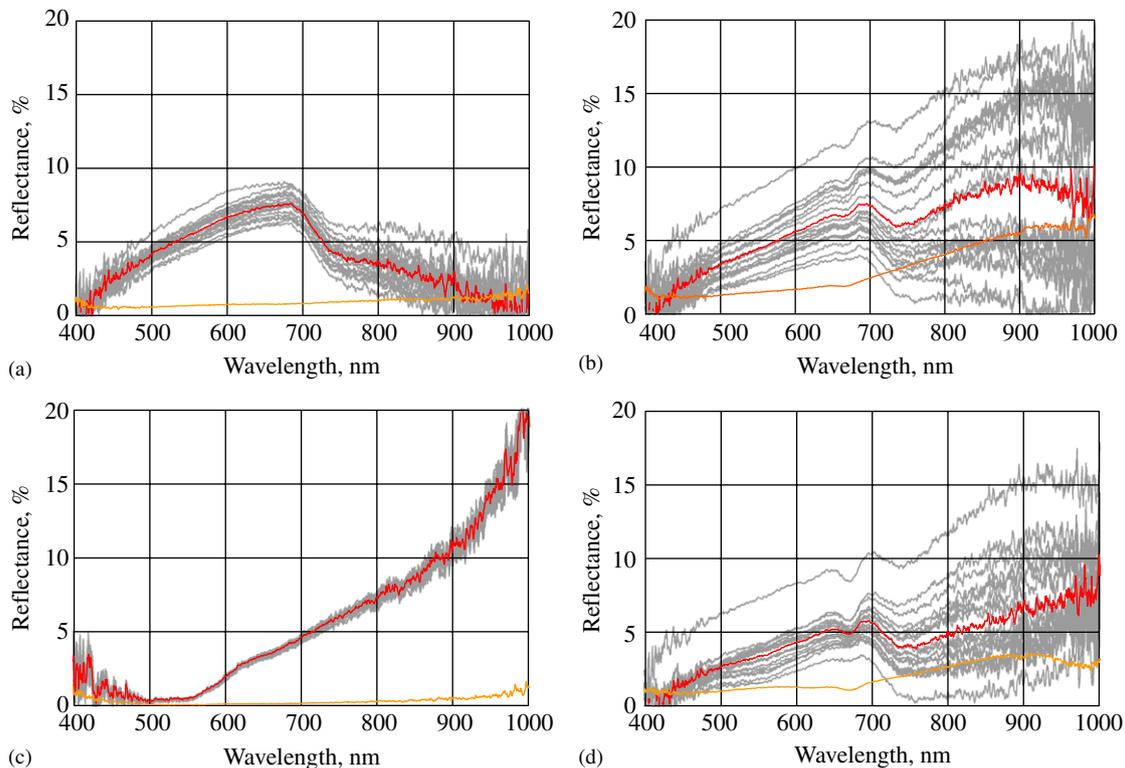
For determination of important spectral bands for discrimination of dirty and clean surface condition, the PCA method was applied for spectral data analysis.

The potential spectral bands were dependent on the surface materials investigated. For wet concrete, the loadings for PC1 increased linearly from 700 to 900 nm, *Fig. 5*. This implies that the spectral data in the range 800 to 900 nm would provide valuable information. Recognising that the loading for PC3 has maximum negative value at 800 nm and to avoid the lower S/N level at longer wavelength bands, the signal at 800 nm would be significantly important for wet concrete surfaces. Another important band for wet concrete

**Table 2**  
Summarised spectral signatures for clean surface compared to dirty in wet conditions

Surface type	Significant spectral bands, nm		
Concrete		750–1000 low*	
Plastic	400–550 high	600–700 low	800–1000 high
Wood	500–700 low	750–1000 high	
Steel	400–550 high	600–700 low	

\*The 'low' and 'high' in the table means that the reflectances measured were lower or higher for clean surfaces compared to dirty.



*Fig. 4. Spectral data and the standard deviations of grey concrete (a) clean; (b) dirty; and brown painted wood slabs (c) clean; (d) dirty in wet conditions: —, measurements; —, average and —, standard deviation*

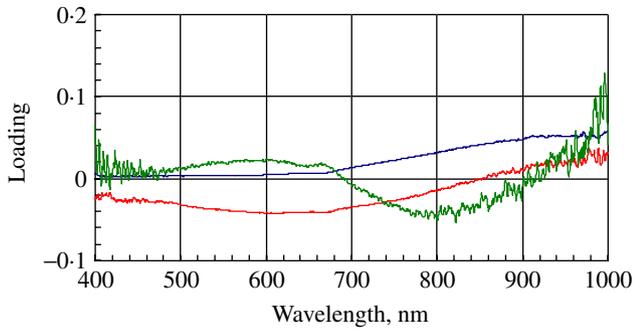


Fig. 5. Loading plot for a principal component analysis of the spectral signatures for wet concrete, with the total explained variance along each component: —, PC1 of 91%; —, PC2 of 7% and —, PC3 less than 1%

would be 630 nm, where, the loading for PC2 has a maximum negative value and that for PC3 has also a peak value. Therefore, for wet concrete, the important spectral bands are 800 and 630 nm.

For wet stainless-steel surfaces, the loading of PC1 increased also linearly in the spectral range between 700 and 900 nm, with a zero value at 700 nm, Fig. 6. However, a large negative value was found in the bands about 500 nm. For PC2 and PC3, the largest loadings were found in the bands of 650 and 700 nm, with negative values. The results indicated that for the wet steel surfaces, important spectral bands are 500, 650 and 750 nm.

For the brown painted wood plate, the loadings of PC1 had a similar pattern as the wet steel Fig. 7. However, the loadings increased quickly from a zero value in the bands of 700–730 nm. A large negative value was found in the bands of 500–600 nm, in which PC2 had also a maximum negative value. For PC2, another large negative loading was found near 700 nm, where, loadings for PC1 and PC3 were near to zero. Anyhow, the loadings for PC2 and PC3 had reasonable large negative values at 800–850 nm, where the loadings for PC1 were close to the maximum. These results mean that the valuable bands for brown painted wood are 500–600, and 800 nm.

For the wet green plastic plate in Fig. 8, the loadings of PC1 had a small peak value at 480 nm, where the loadings for PC2 and PC3 also had their peak values. At ranges of 650 and 700 nm loadings for PC2 had two negative peak values, where the loadings for PC3 were also reasonable large. At 800 nm, the loadings for PC1 reached a value that was close to the maximum, while the loadings for PC2 and PC3 had very small negative values. The results indicate that useful bands for the wet green plastic surface are 480 and 650, and possibly 700 nm.

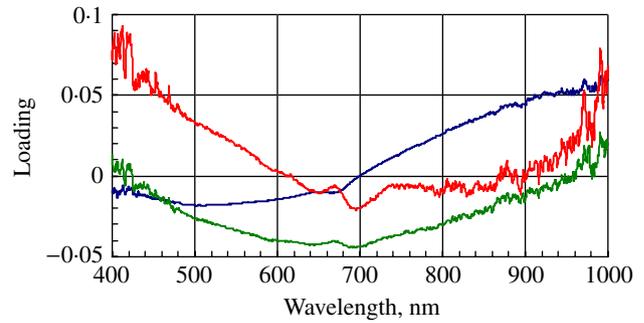


Fig. 6. Loading plot of a principal component analysis for the spectral signatures of wet stainless steel, with the total explained variance along each component: —, PC1 of 61%; —, PC2 of 27% and —, PC3 of 10%

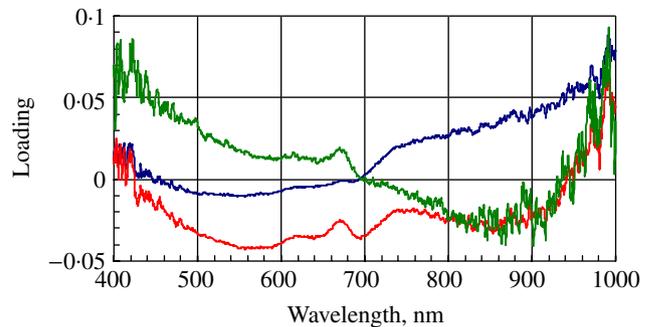


Fig. 7. Loading plot of a principal component analysis for the spectral signatures of wet brown painted wood, with the total explained variance along each component: —, PC1 of 80%; —, PC2 of 18% and —, PC3 of 1%

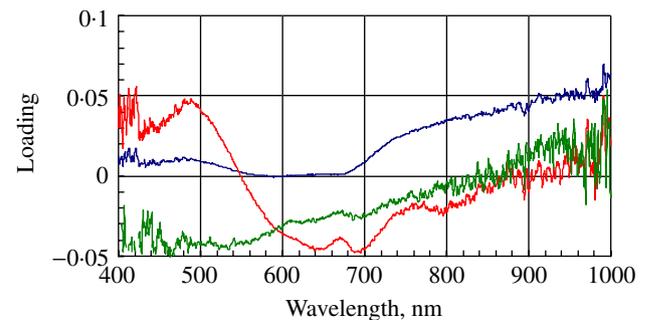


Fig. 8. Loading plot of a principal component analysis for the spectral signatures of wet green plastics, with the total explained variance along each component: —, PC1 of 69%; —, PC2 of 28% and —, PC3 of 1%

### 3.2.2. Analysis of variances of spectral data

A set of measurements on a surface will have a distribution of reflectivity due to differences in the clean surface itself and due to the uneven distribution of residues to be cleaned. It is therefore necessary to make detailed statistic analysis to determine the spectral bands

that are important for discrimination of clean and dirty surfaces. To evaluate the potential bands determined with PCA, further analysis was performed by applying the ANOVA. In the analysis, the spectral data from the dirty and clean surfaces of each type of the materials were compared in the selected spectral bands.

Since the classical ANOVA test requires normal distribution of data, test for normality were performed using the Lilliefors method (Conover, 1980). With a 5% level of confidence, results show that a normal distribution was found at most wavelengths for clean conditions, while normal distributions were not found for many spectral ranges for dirty conditions. Therefore, the analyses of the spectral data were performed in both the classical one-way ANOVA method and Kruskal–Wallis method—a non-parametric one-way ANOVA (Hollander & Wolfe, 1973), noting that the classical ANOVA is only valid for normally distributed data. A pre-set confidence parameter of 5% was applied for the classical ANOVA analysis.

A probability  $P$  that represents the statistical significance of the differences between the data in the clean and dirty conditions for each spectral band was determined. Specifically, the  $P$  value represents the probability of error that is involved in accepting the measured result as valid, that is, as representative of the population. In the bands, where the significant differences were found between the spectral data of clean and dirty surfaces, the  $P$  values should be close to zero. Consequently, the spectral signal in these bands could be used for discrimination of cleanliness of the surfaces. However, with the confidence level of 5%, a result may be considered as significant if the value of  $P$  is less than 0.05.

The  $P$  values of the ANOVA for the wet surface conditions are listed in Table 3. There are significant

differences at certain frequency bands where data are not normally distributed. In other spectral ranges, the two different ANOVA methods give similar results. According to the analysis, the optimal spectral bands for wet concrete surfaces are 500–600 and 750–800 nm, while the optimal spectral bands for wet stainless-steel surfaces are 450–500 nm and 650–700 nm. For the brown painted wood surface in wet condition, significant difference was found for clean and dirty surface in all the spectral bands analysed, while the bands of 500–650 was superior. For the green plastic plate, the spectral data of clean and dirty wet condition were significantly different in the 450–700 nm bands. The most significant bands were 450 and 500 nm.

The analysis indicates that, when spectral bands of 450, 600, 700 and 800 nm are chosen, there are at least two spectral bands for each type of the materials, in which the spectral signals can be used for discrimination of dirty and clean condition of the surfaces.

### 3.3. Possibility of utilising a camera and multi-spectral image analysis method

The spectral signatures and statistical analysis indicate that it is feasible to apply a CCD camera with appropriate filter or light source as a sensor for detection of the surface cleanliness. Since both CCD camera and VIS–NIR spectrometers are based on CCD array technology, their sensitivities/relative responses in the visible and NIR spectrums are similar. A spectrometer applies a line sensor array (one dimension) while a camera applies an area array (two dimension). Figure 9 shows a typical sensitivity/response curve of a CCD sensor array. The maximum sensitivity range varies from a blue range around 450 nm to a red around

**Table 3**  
Probability from the classical analysis of variance (ANOVA) and the Kruskal–Wallis non-parametric ANOVA (K–W) of the spectral data in the selected spectral bands for wet surfaces

Spectral bands, nm	Probability ( $P$ )							
	Gray concrete		Stainless steel		Brown-painted wood		Green plastics	
	Classical	K–W	Classical	K–W	Classical	K–W	Classical	K–W
450	0.054	<b>0.015</b> <sup>†</sup>	*	*	*	*	*	*
500	<b>0.017</b>	<b>0.005</b>	*	<b>0.007</b>	*	*	*	*
550	<b>0.006</b>	<b>0.004</b>	0.113	0.345	*	*	0.884	0.829
600	<b>0.012</b>	<b>0.012</b>	0.361	0.433	*	*	*	*
650	0.163	0.088	*	*	*	*	*	*
700	0.442	0.829	*	*	<b>0.010</b>	0.030	*	*
750	<b>0.011</b>	0.074	0.432	0.291	*	*	0.896	<b>0.035</b>
800	*	*	0.574	0.607	*	*	0.492	<b>0.027</b>

\*  $P < 0.001$ .

<sup>†</sup>The values in bold are  $P < 0.05$ .

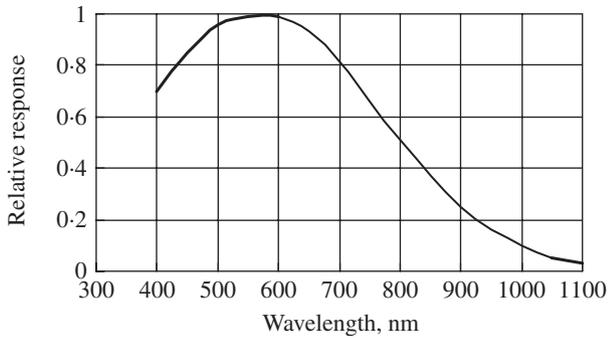


Fig. 9. A typical sensitive curve for a charged-couple device sensor array for spectral meters and cameras

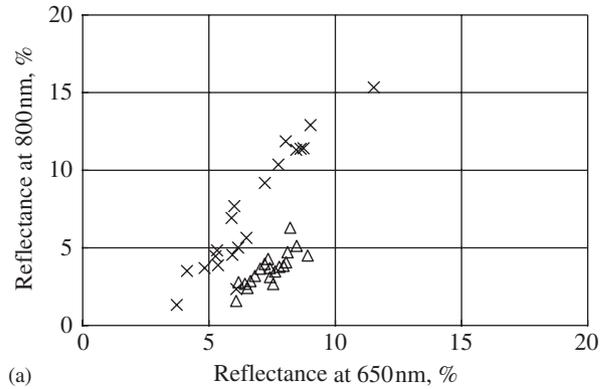
600 nm, depending on the types of CCD sensors. Many of them are able to catch information in the NIR range up to 1000 nm.

A spectrometer is designed to use each sensor in a line array for a specific wavelength range to characterise a measurement point. An individual sensor provides the information related to a well-defined wavelength in the range that the spectrometer covered.

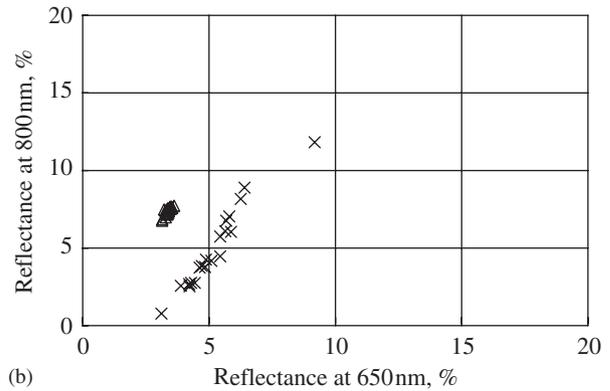
A CCD camera focuses a sensor array to a measurement area. An individual sensor provides the integrated information related to entire wavelength range on a part of the measured area. The information achieved by a CCD camera therefore is dependent on the filter or the illumination used. That means a pre-defined filter or illumination is necessary to provide information in a defined spectral range.

By using the acousto-optic tunable filter (AOTF) method, a set of spectral images may be achieved in the defined spectral bands (Gupta *et al.*, 1999). However, this is an expensive solution. An alternative is to select a few spectral bands based on the spectral characteristics achieved in the signature measurements and to apply the filters or illumination in the defined spectral range to a CCD camera. For the most commonly used material, concrete, for example, the clean and dirty area may be identified by the reflectance information in the spectral range of 650 and 800 nm. That can be demonstrated in Fig. 10, where the clean and dirty areas for concrete and brown painted wood are clearly separated in the two spectral channels in wet condition.

The statistical analysis showed that it is possible to make a significant discrimination and hence classify areas that are visually clean. The spectral bands with potential for discrimination of clean and dirty condition may be identified and used for classification of spectral images of the building materials. The methods for classification, *e.g.* the Jeffrey–Matusita distance methods (Matusita, 1966) and Bayesian discrimination for multi-spectral image classification have been presented by Blanke *et al.* (2004) and by Braithwaite *et al.* (2005).



(a)



(b)

Fig. 10. Reflectance for concrete (a) and brown wood (b) slabs at wavelength 800 nm versus 650 nm, in wet condition, where,  $\Delta$ , clean and  $\times$ , dirty

#### 4. Conclusions

The results of an investigation of the optical properties in the visual and the near infrared (VIS–NIR) optical range of clean and dirty surfaces in a finishing pig unit are presented.

The reflectance of clean and contaminated building materials is different in the VIS–NIR wavelength range. For the spectral reflectance of a clean surface compared to a dirty in wet conditions, the values were lower in 750–1000 nm for grey concrete; higher in 400–550 nm, lower in 600–700 nm and higher in 800–1000 nm for green plastic; lower in 500–700 nm and higher in 750–1000 nm for brown printed wood; and higher in 400–550 nm and lower in 600–700 nm for stainless steel.

The principal component analysis (PCA) showed that for wet concrete, the spectral data in the ranges of 630 and 800–900 nm provide valuable information. The valuable bands for brown painted wood are 500–600, and 800 nm. For the wet green plastic surface the valuable bands are 480 and 650, and possibly 700 nm. For the wet steel surfaces, important spectral bands are 500, 650 and 750 nm.

According to the analysis of variance, when spectral bands 450, 600, 700 and 800 nm are chosen, there are at

least two spectral bands for each type of the materials, in which the spectral signals can be used for discrimination of dirty and clean condition on the surfaces.

The results of the spectral signatures studies of the surface materials with and without dirt showed that it is possible to make a significant discrimination and hence classify which area that is visually clean.

### Acknowledgements

This research was performed as part of the ISAC project supported by the Sustainable Technology in Agriculture programme funded jointly by the Danish Ministry of Science, Technology and Innovation and the Ministry of Food Agriculture and Fisheries under Grant number 2053-01-0021.

### References

- Andersen A H; Gash D M; Avison M J** (1999). Principal component analysis of the dynamic response measured by fMRI: a generalized linear systems framework. *Magnetic Resonance Imaging*, **17**(6), 795–815
- Blanke M; Braithwaite I; Zhang G** (2004). Classification of Clean and Dirty Pighouse Surfaces Based on Spectral Reflectance. Technical University of Denmark, Copenhagen 69 pp
- Braithwaite I; Blanke M; Zhang G; Carstensen J M** (2005). Design of a vision-based sensor for autonomous pighouse cleaning. *EURASIP Journal of Applied Signal Processing*, **2005**(13), 2005–2017
- Conover W J** (1980). *Practical Nonparametric Statistics*. Wiley, New York
- Esbensen K H** (2000). *Multivariate Data Analysis—In Practice* (4th Edn). CAMO ASA, Oslo
- Gupta N; Dahmani R; Gottlieb M; Denes L; Kaminsky B; Metes P** (1999). Hyperspectral imaging using acousto-optic tunable filters. In *Proceedings of SPIE*, Vol. 3718, 10 pp
- Hollander M; Wolfe D A** (1973). *Nonparametric Statistical Methods*. Wiley, New York
- Jackson J E** (1981). Principal components and factor analysis—part I: principal components. *Journal of Quality Technology*, **12**(4), 201–213
- Matusita K** (1966). A distance and related statistics in multivariate analysis. In: *Multivariate Analysis*. (Krishnaiah, ed), 187–200. Academic Press, New York
- Pedersen B K; Kai P** (1998). Rengøring af Svinestalde. [Cleaning of pig houses] (in Danish). Info Svin No. 376, 10pp. Landsudvalget for Svin. [National Committee for Pig Production.] Axelborg, Copenhagen
- StellarNet** (2003). *EPP2000 and ISA2000 Fiber Optic Spectrometer Manual*. StellarNet Inc., FL 34677, USA
- Strøm J S; Zhang G; Kai P; Lyngbye M; Braithwaite I; Arrøe S** (2003). Intelligent sensor for autonomous cleaning of pig houses—system in context. DIAS Internal Report 182, 24pp
- Wold S; Esbensen K H; Geladi P** (1987). Principal component analysis. *Chemometrics and Intelligent Laboratory Systems*, **2**(1–3), 37–52
- Wu W; Massart D s L; de Jong S** (1997). The kernel PCA algorithm for wide data—part I: theory and algorithms. *Chemometrics and Intelligent Laboratory Systems*, **36**(2), 165–172