Peer Reviewed Paper openaccess

Impact of water vapour on polymer classification using *in situ* short-wave infrared hyperspectral imaging

Muhammad Saad Shaikh and Benny Thörnberg

Department of Electronics Design, Mid Sweden University, Sundsvall, Sweden

Contacts
M. Shaikh: saad.muhamadshaikh@miun.se
https://orcid.org/0000-0002-2538-5205

B. Thörnberg: benny.thornberg@miun.se
https://orcid.org/0000-0001-5521-7491

Hyperspectral remote sensing is known to suffer from wavelength bands blocked by atmospheric gases. Short-wave infrared hyperspectral imaging at *in situ* installations is shown to be affected by water vapour even if the pathlength of light through air is only hundreds of centimetres. This impact is especially noticeable with large variations of relative humidity, the coefficient of variation reaching 5 % in our test case. Using repeated calibrations of imaging system at the same relative humidity as in the measurement, we were able to reduce the coefficient of variation to 1 %. The measurement variations are also shown to induce significant error in material classification. Polymer type identification was selected as the test case for material classification. The measurement variations due to the change in relative humidity are shown to result in 20 % classification error at its minimum. With repeated calibrations or by eliminating the most affected wavelength bands from measurements, we were able to reduce the classification error to less than 1 %. Such improvement of measurement and classification precision may be important for industrial applications such as waste sorting, polymer classification etc.

Keywords: hyperspectral imaging, humidity, infrared, calibration, InGaAs, waste sorting, plastic detection, material classification

Introduction

Hyperspectral imaging is an optical contactless sensing method that can be used to acquire spectral signatures of various materials at each pixel. Unique spectral signatures are the result of molecules having quantified energy levels of vibrations. When incident photons cause molecules to enter higher energy levels, absorption of light occurs.[1] Most of these transitions happen in the mid-range infrared region (2.5–25 µm) of light but can be observed as overtones in the short-wave infrared (SWIR) region (0.9–2.5 µm).

Remote sensing is an important application area for hyperspectral imaging. Chemical composition of ground surfaces can be analysed and classified from their spectral signatures using instrumentation onboard satellites or aircraft.[2] Waste sorting and recycling is another example of industrial use of hyperspectral imaging for material

Correspondence
M. Shaikh: saad.muhamadshaikh@miun.se

Received: 17 May 2022
Revised: 28 May 2022
Accepted: 29 May 2022
Publication: 1 June 2022
doi: 10.1255/jsi.2022.a5
ISSN: 2040-4565

Citation
M.S. Shaikh and B. Thörnberg, "Impact of water vapour on polymer classification using *in situ* short-wave infrared hyperspectral imaging", *J. Spectral Imaging* **11,** a5 (2022). https://doi.org/10.1255/jsi.2022.a5

classification.[3–5] Push-broom is a commonly used technique for hyperspectral cameras. This type of camera captures a line of pixels (scan line), while hundreds of spectral channels provide a dense sampling of the wavelength spectrum at each pixel.[6,7] Push-broom cameras are well suited for industrial use, where typically the motion from a conveyer belt provides the second spatial image dimension.[5]

There is at least one major difference in using hyperspectral imaging for remote sensing compared to industrial applications. That is the availability of windows and blocked regions in the transmittance of the atmosphere.[2,8] Water vapour is one of the major reasons for the blocked wavelength regions that cannot be used for remote sensing. However, the relatively very short path length of light for industrial applications suggests to us that the whole SWIR region should be available for measurements. But two questions still remain: (1) How much will water vapour have an impact on the practical use of hyperspectral imaging for industrial applications? (2) How much can calibration procedures reduce the impact from water vapour? We intend to provide answers to those questions in this paper.

Radiometric calibration,[9] calibration of true reflectance[10,11] or relative reflectance[12] of hyperspectral cameras, or hyperspectral imaging systems are known computational methods used to suppress impact from variations in light intensity, pixel sensitivities and environmental parameters. Despite the well-known impact of water vapour in the SWIR region of light,[2,8] we could not find any published studies of the impact of water vapour on *in situ* hyperspectral imaging systems and material classification. The scientific contribution of this paper is an analysis of this impact.

Figure 1. Materials used in the research: (a) Specim FX17e hyperspectral camera, (b) Rallye 3003 compact halogen lamp, (c) PTFE tile, (d) DEM500 humidity sensor, (e) Beurer Gmbh's LB 88 humidifier.

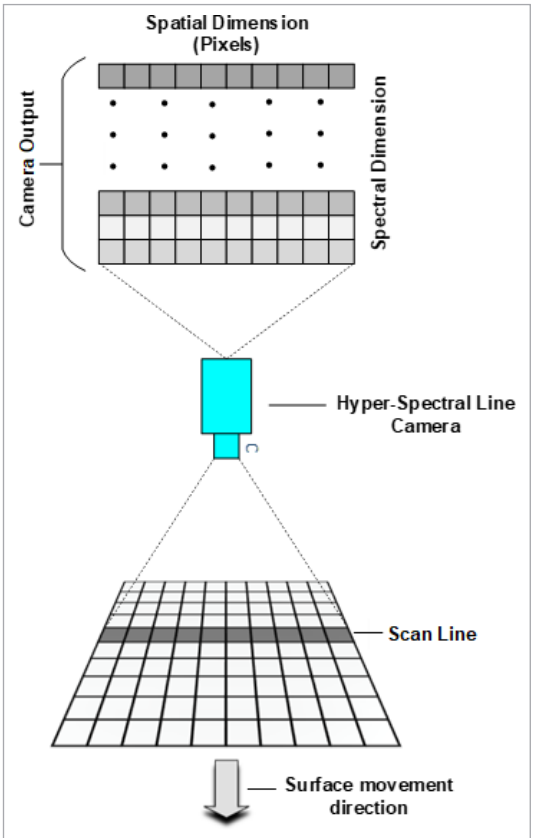

Figure 2. A general scheme for a push-broom camera.

Materials and methods

In this section we provide the brief details of hardware used in this research. The description of methods used to acquire and process the spectral data are also part of this section.

Hyperspectral camera

In this research, the FX17e hyperspectral camera from Specim was used, shown in Figure 1a. This camera captures 2D hyperspectral images in the SWIR region, with an exact range of 900–1700 nm. Each image contains information on 224 spectral bands with a full width half maximum (FWHM) of 8 nm. This camera uses

the push-broom[6] method to acquire spectral information of the surface along the scan line, as depicted in Figure 2. For each scanned line, the camera outputs a 2D image where each row represents the spectral reflection in one of the 224 spectral bands.

Calibration target

To measure the relative reflectance of the test samples, a 10 mm thick, 295 × 295 mm polytetrafluoroethylene (PTFE) tile manufactured by Amsler & Frey was used as the calibration target. This board is shown in Figure 1c. This PTFE tile is known to diffusely reflect light with as even a distribution as Spectralon™ over the range of all wavelengths in the SWIR region. It is also known to be insensitive to humidity, a property of particular interest when experimenting in humid environments.

Illumination source

For illumination of the test samples and calibration target, two halogen lamps, Rallye 3003 compact, manufactured by Hella[13] were used. One of these lamps is shown in Figure 1b. The lamps were mounted on a tripod stand and connected to a general-purpose power outlet of 220 V via an AC/DC converter.

Humidity/temperature sensor

To measure the relative humidity (RH) and air temperature during experiments, a DEM500 sensor manufactured by Velleman[14] was used. The sensor is able to sense RH from 0 % to 100 % with tolerance of ±3 %. It also measures temperature in the range of −20–70 °C with tolerance of ±1 °C. This sensor device is shown in Figure 1d.

Humidifier

A Beurer Gmbh LB 88 humidifier[15] was deployed inside the lab to increase the RH to the desired values of 80–85 %. But the RH level did not go higher than 60 % with the LB 88 alone. An electric stove was additionally used, to boil water inside the lab to reach RH levels above 80 %. The LB 88 humidifier is shown in Figure 1e.

Experiment setup

Two experiments were conducted in this research. Experiment #1 was conducted to analyse the impact of RH change on the hyperspectral measurements of materials. It is also analysed how much the measurement error can be reduced by frequent calibrations of the imaging system. In Experiment #2 we analysed if the inaccuracies in the measurements induced by the change of RH have any impact on the classification of materials present in the scene. For the second experiment, polymer classification was taken as the test application case.

Both experiments were conducted using the experiment setup shown in Figure 3. The left side of Figure 3 shows a schematic depiction of the camera and light source. Distance L from the stand to the analysed object is approximately 85 cm. Height of stand is also 85 cm, which makes the travelled pathlength of light in air $2 \cdot \sqrt{2} \cdot L = 240$ cm. The right side of the same figure shows a front view photo of the experiment setup. Label A at the bottom indicates the position of a 3D printed test

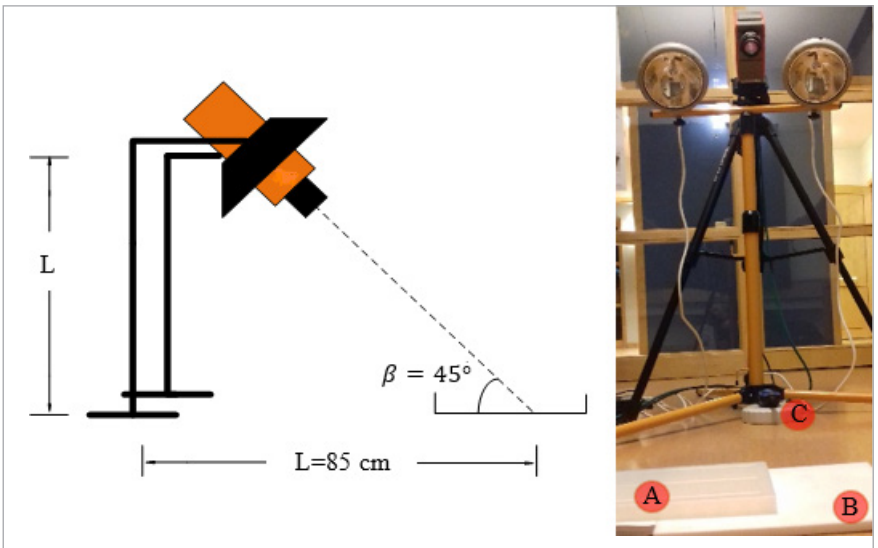

Figure 3. Left side: schematic drawing of camera and light source. Right side: front view photograph of experiment.

object made of polylactic acid (PLA). Label B indicates a 10 mm thick calibration object made of polytetrafluoroethylene (PTFE).

A sensor for RH and temperature was positioned close to the stand at floor level, label C. This experiment setup was arranged to mimic the optical situation at a typical industrial site for, e.g., sorting of waste material. All equipment was positioned inside a 10 m^2 office, including an ultrasonic humidifier and an electric stove for boiling water.

Sample preparation

For Experiment #1, a 3D printed test object made of PLA was placed on the right half of the camera line of scan. The calibration target was placed on the other half of the line of scan. All the spectral measurements for Experiment #1 were made without moving the test object or calibration target. The test object and calibration target were placed together for every measurement to avoid the impact of any potential slight change in the environment.

For Experiment #2, different types of recyclable plastic were placed on the camera scan line. The calibration target was also included in the sample as shown in Figure 4.

These recyclable plastics are generally known by their recycling codes from 1 to 7. The first six types are included in this sample. Table 1 presents the details of the plastics used with their polymer names, recycling numbers and abbreviations.

Table 1. Types of polymers used in test case 2 samples.

Recycling number	Abbreviation	Polymer name
1	PETE or PET	Polyethylene terephthalate
2	HDPE or PE-HD	High-density polyethylene
3	PVC or V	Polyvinyl chloride
4	LDPE or PE-LD	Low-density polyethylene
5	PP	Polypropylene
6	PS	Polystyrene

Experimentation procedure

Experiment #1

The first experiment was started in a long-term stable climate inside the office room, 20 % RH and 22.4 °C. Two hyperspectral images were captured using the push-broom camera such that the scan line covered both PLA and PTFE materials within the same capture. The humidifier was started to run at variable powers and times to reach higher RH. Hyperspectral images were captured for a series of increasing RH. It became necessary to also boil water using an electric stove to attain RH levels higher than 60 %.

Experiment #2

For the second experiment, the spectral measurements of all three polymer samples were made individually at 23 % RH and 83 % RH. The same experiment setup was used for both experiments.

Calibration model

The calibration model depicted in Figure 5 was used as a dataflow graph of spectral transformations.[12] $L(\lambda)$ corresponds to the spectral distribution of the light source, where λ is the wavelength of light. $A(\lambda)$ is the wavelength-dependent absorption in air along a path-length defined by the experiment setup. $M(\lambda)$ is the wavelength-dependent reflection of light in the analysed surface and at the angle β defined by the experiment setup. $C(\lambda)$ is the spectral distribution of sensitivity

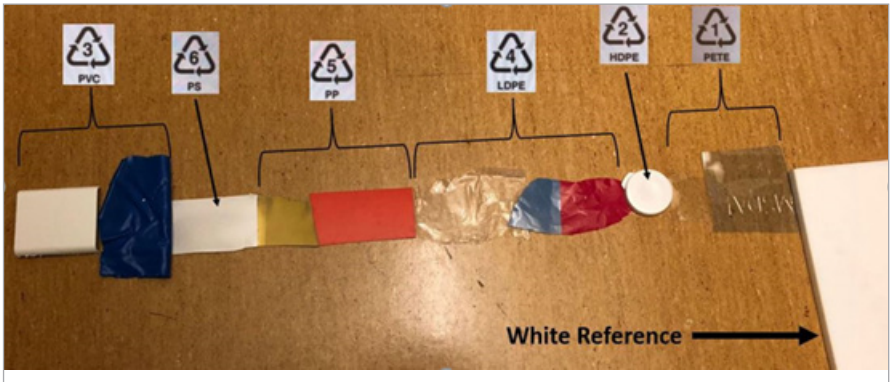

Figure 4. One of the three polymer samples for Experiment 2.

for the hyperspectral camera being used. $W(\lambda)$ is the wavelength-dependent reflection of light in the surface of the calibration object and at the angle β. The lower path of the dataflow graph shown in Figure 5, labelled as reference measurement "*ref*" corresponds to a calibration at air absorption $A_r(\lambda)$. The upper path is labelled as material measurement "*m*" and corresponds to the measurement of light reflected in an analysed material at air absorption $A_m(\lambda)$. As per Equation 1, the recorded spectral signatures of analysed materials are made independent of light source, air absorption and camera sensitivity by the computation of relative reflectance $RR(\lambda)$ when $A_m(\lambda) = A_r(\lambda) = A(\lambda)$.

$$RR(\lambda) = \frac{m}{ref} = \frac{A(\lambda) \bullet M(\lambda) \bullet C(\lambda) \bullet L(\lambda)}{A(\lambda) \bullet W(\lambda) \bullet C(\lambda) \bullet L(\lambda)} = \frac{M(\lambda)}{W(\lambda)} \qquad (1)$$

Hence, this is no absolute radiometric calibration since $RR(\lambda)$ is still dependent on the calibration target. However, if the calibration object has a spectral distribution of reflected light $W(\lambda)$ close to constant over the spectral range of measurements, this makes the relative reflectance $RR(\lambda)$ dependent only on the analysed material.

Data preparation

Experiment #1

For Experiment #1, six hyperspectral images of the first sample were captured at 20, 31, 42, 56, 78 and 89 % of RH, respectively. From each image, mean spectra of PLA test object and calibration target were calculated.

Experiment #2

For Experiment #2, three polymer samples were prepared. For each sample, two measurements were made at 23 and 83 % RH respectively. From these six measurements, three matrices of $RR(\lambda)$ were calculated using Equation 1. First, the $RR_T(\lambda)$ matrix was computed using both $M(\lambda)$ (polymers) and $W(\lambda)$ measured at 23 % RH such that: $A_m(\lambda) = A_r(\lambda)$. Second, $RR_D(\lambda)$ was computed using $A_m(\lambda)$ at 83 % RH and $A_r(\lambda)$ at 23 % RH. The third matrix $RR_H(\lambda)$ was computed using $A_m(\lambda) = A_r(\lambda)$ at 83 % RH. The matrix $RR_D(\lambda)$ represents a scenario when RH in material measurement has increased considerably with respect to the reference measurement at lower RH. The third matrix $RR_H(\lambda)$ represents a scenario when the reference and material measurements were made at the same higher RH (83 %). Table 2 presents the number of spectra per data class, from all three samples, for training and test datasets.

Principal component analysis

Principal component analysis (PCA) is a mathematical method to reduce the number of dimensions of large data sets. *N* axes called principal components are defined within a *K*-dimensional input data space such that a new orthonormal coordinate system is formed. The orientation of the principal components is defined by a 2-dimensional matrix of variable loadings

$$\boldsymbol{P} = \left\{ P_{i,j} | i \in 1..K \wedge j \in 1..N \right\}.$$

An input data vector $\vec{X}$ is a set of *K* number of input scalar variables. Linear projection of $\vec{X}$ on the set of *N* principal components will result in an *N*-dimensional score vector $\vec{S}$,

$$\vec{S} = \vec{X} \cdot \boldsymbol{P} \qquad (2)$$

The graphical plotting of a set of vectors $\vec{S}$ is usually referred to as a "score plot".[16] The *K*-dimensional loading vector

$$\vec{P}_i = \left\{ P_{i,j} | j \in 1..N \right\}$$

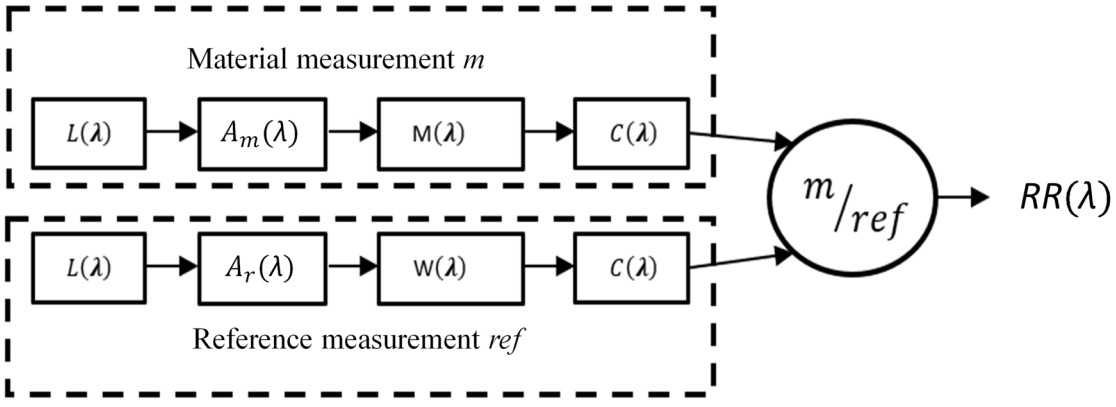

Figure 5. Spectral calibration of measurement setup.

Table 2. Number of spectra per data class for training and test datasets.

Class	Number of spectra		
	$RR_T(\lambda)$ (training data)	$RR_D(\lambda)$ (Test data 1)	$RR_H(\lambda)$ (Test data 2)
PETE	183	183	183
HDPE	60	60	60
PVC	243	243	243
LDPE	333	333	333
PP	270	270	270
PS	138	138	138
Background	93	93	93

explains how much an input variable indexed by *i* contributes to the *N* dimensions of a score vector $\vec{S}$. The graphical plotting of a set of loading vectors $\vec{P_i}$ is usually referred to as a "loading plot". If a loading vector $\vec{P_i}$ has large magnitude, it means that the input variable indexed by *i* has a significant impact on the PCA model, less magnitude means less impact.

$$\vec{M} = \left\{ M_i | M_i = \sqrt{\sum_{j=1}^{N} P_{i,j}^2} \; \forall i = 1..K \right\} \tag{3}$$

The vector of loading magnitudes $\vec{M}$ can thus reveal information about which subset of input variables contribute the most to the PCA model. For applications using hyperspectral imaging, this vector of loading magnitudes can be used for wavelength band selection. Other researchers have previously used PCA and the loading coefficients for band selection.[17,18]

In this study, PCA was performed on $RR_T(\lambda)$ to create a matrix of variable loadings and a low-dimensional training dataset of score vectors. Test datasets of score vectors were computed from $RR_D(\lambda)$ and $RR_H(\lambda)$ using the same matrix of variable loadings. The centring and scaling of $RR_T(\lambda)$, $RR_D(\lambda)$ and $RR_H(\lambda)$ were conducted using mean and variance vectors of $RR_T(\lambda)$ as shown in Figure 6.

The matrix of variable loadings computed from $RR_T(\lambda)$ were used to transform the $RR_D(\lambda)$ and $RR_H(\lambda)$ into the same lower-dimensional space as with the training data. This was done to ensure that the test vectors had no impact on the training of the classifier.

Support vector machines

Support vector machines (SVMs) are supervised learning models with associated learning algorithms that are used for data classification and regression analysis. In this study, we trained a multiclass SVM classifier to analyse the impact of water vapour on the material classification for a selected test case. To implement the multiclass SVM, MATLAB multiclass error-correcting codes (ECOC)[19] were used with radial base function (RBF) kernels. A previous study[20] motivated us to use the combination of PCA for dimension reduction and SVM for classification to get better classification of hyperspectral data.

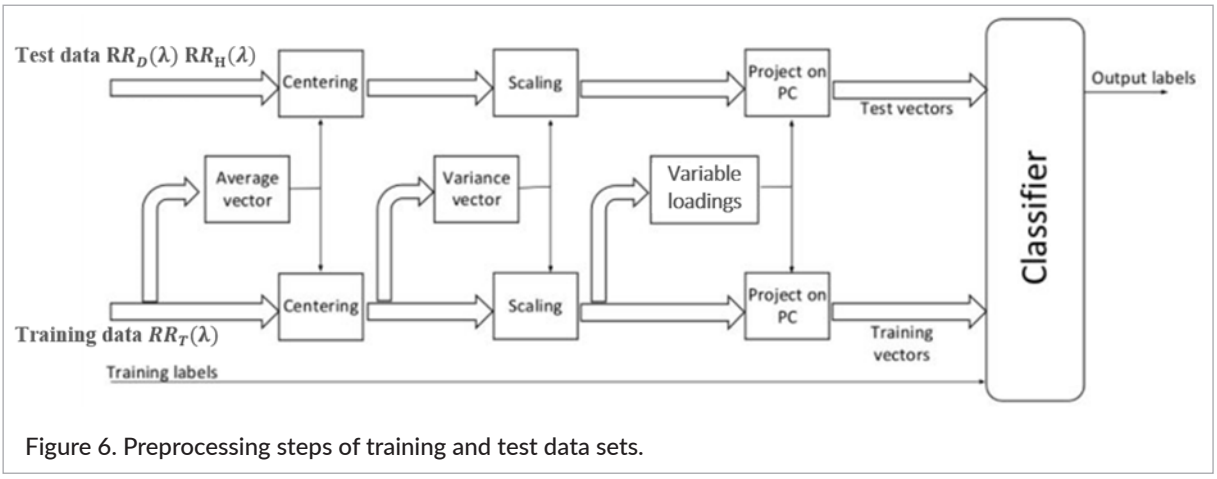

Figure 6. Preprocessing steps of training and test data sets.

Result and analysis

Figure 7 shows the measurements of relative reflectance $RR(\lambda)$ for the calibration target repeated for various RH. Hence, the material $M(\lambda)$ is the same as the calibration target $W(\lambda)$ and ideally $RR(\lambda) = 1$. The computation of $RR(\lambda)$ in this figure is done with a calibration measurement at 20 % RH. As expected, we see an almost perfect straight line such that $RR(\lambda) = 1$ at 20 % RH. But for the other levels of RH, we see increased absorption of light in two spectral regions: approximately 1.1 µm and 1.38 µm, as also reported as blocked wavelength bands by the remote sensing community.[2,8] To understand why Figure 7 shows deviations from a straight line, we need to consider the different compositions of air, causing different absorption spectra at the time of reference measurement $A_r(\lambda)$, and at the time of material measurement, $A_m(\lambda)$.

If the analysed material $M(\lambda)$ is the same as the calibration target, it means that Figure 7 shows the impact from variations of RH, where $RR(\lambda) = A_m(\lambda) / A_r(\lambda)$.

Figure 8 shows the relative reflectance $RR(\lambda)$ for a 3D-printed object made of PLA and at various RH using a single reference measurement at 20 % RH. We can see that the alignment of the six curves is far from perfect. Misalignments are in particular visible close to 1.1 µm and 1.38 µm.

The same measurements of PLA are used to generate graphs in Figure 9, but using reference measurements at corresponding RH. Consequently, $A_m(\lambda) = A_r(\lambda) = A(\lambda)$ ideally corresponds to the same composition of air for both measurement of analysed PLA as well as for the reference measurement of the calibration target. The alignment of the curves seems visually close to perfect.

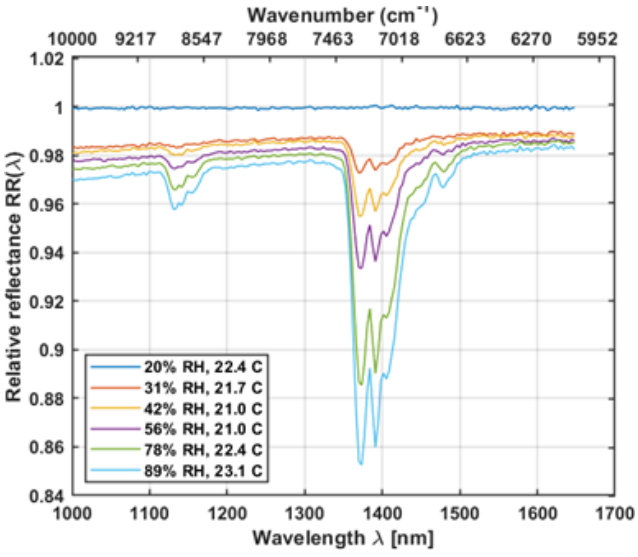

Figure 7. Measured relative reflectance of PTFE at different humidity. Calibration at 20 % RH.

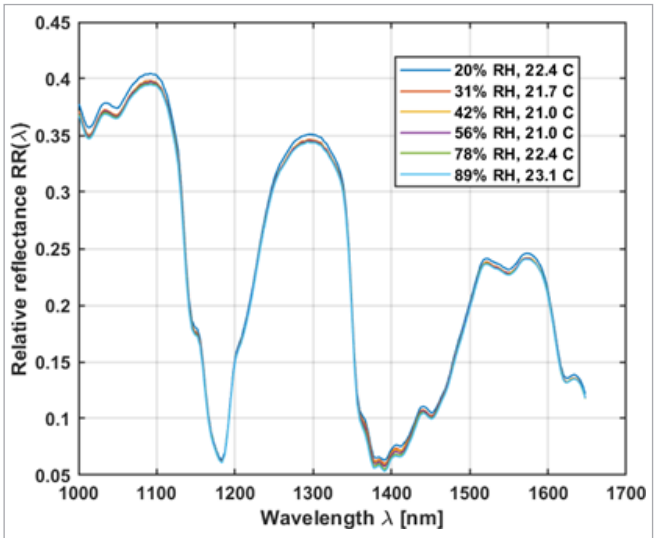

Figure 8. Relative reflectance of PLA at different humidity using calibration at 20 % RH.

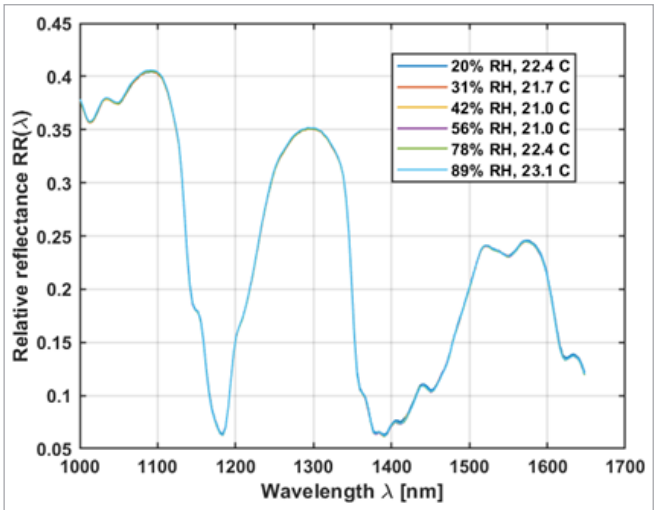

Figure 9. Relative reflectance of PLA using calibration at each RH.

The alignment of curves can further be quantified as the coefficient of variation versus the wavelength as shown in Figure 10. Coefficient of variation is commonly used to evaluate the relative precision of any measurement. In this particular case, it is expected to ideally obtain the same spectral signature of PLA independently of RH. In the 1.38 µm region, there is a reduction from about 5 % to 1 % misalignment when using repeated calibrations at each RH.

Figure 11 shows the mean relative reflectance of PTFE versus RH and for three spectral bands. Calibration was done at 20 % RH. The higher sensitivity to RH for the spectral bands close to 1.1 µm and 1.38 µm is clearly visible.

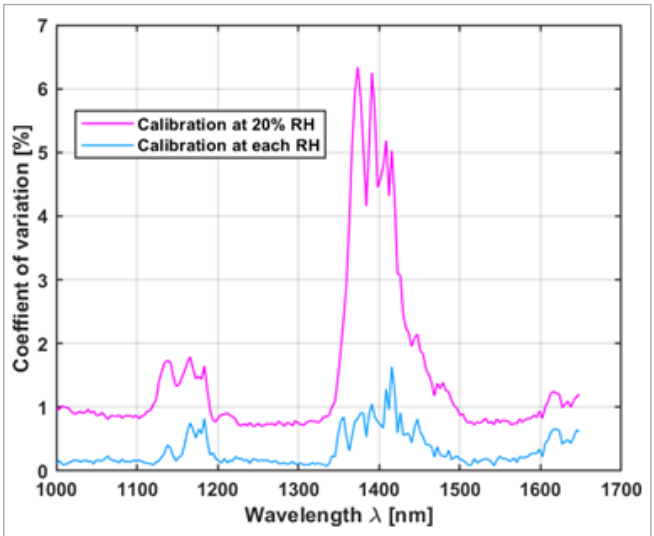

Figure 10. Precision of measured PLA signature.

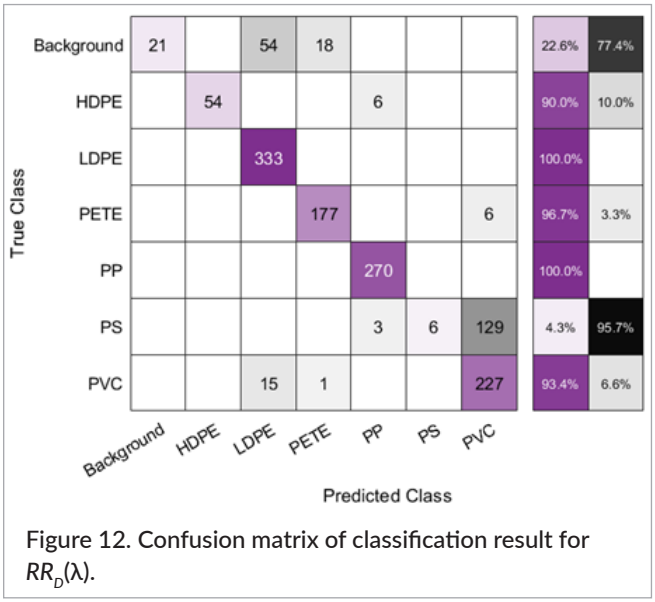

Figure 12. Confusion matrix of classification result for $RR_D(\lambda)$.

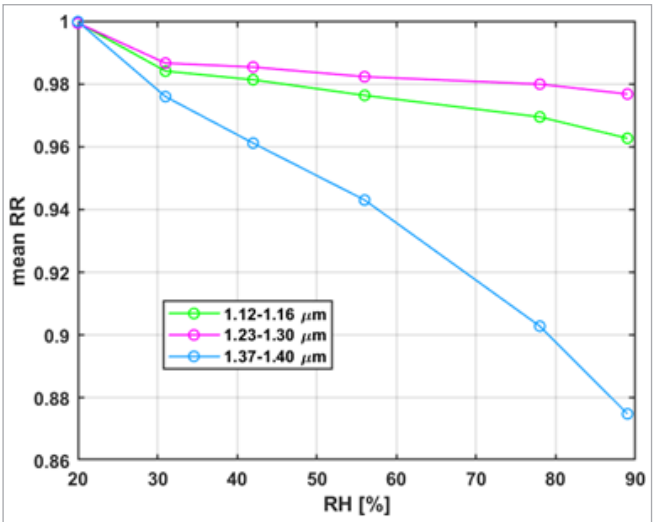

Figure 11. Measured mean relative reflectance of PTFE versus RH and for three spectral bands. Calibration at 20% RH.

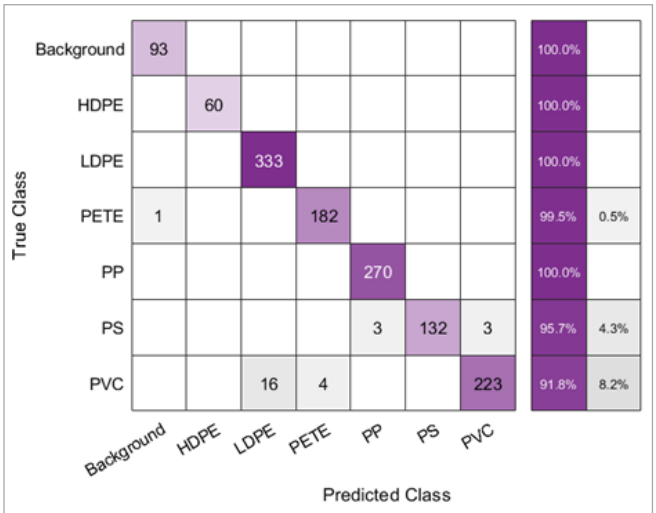

Figure 13. Confusion matrix of classification result for $RR_H(\lambda)$.

The next two figures, Figure 12 and Figure 13 show the confusion matrices of the classification results, when three PCs were used for training, for $RR_D(\lambda)$ and $RR_H(\lambda)$ matrices, respectively.

The PCA of $RR_T(\lambda)$ was further analysed to find the most contributing variables (wavelength bands) by calculating the vector of loading magnitudes using Equation 3, in the space spanned by the first five PCs ($N = 5$). The vector of loading magnitudes $\vec{M}$ are shown in Figure 14a, where the peaks represent the most contributing wavelengths. We included only the first five PCs in our analysis as they explain 99.64% of the total variance, as shown in Figure 14b.

Figure 14a shows that some of the most contributing wavelengths lie in the regions (1.1–1.2 µm and 1.3–1.5 µm) that are susceptible to the humidity change (see Figure 10), but the majority of important wavelengths are outside these regions. In order to suppress the effect of wavelengths from susceptible regions on the classification, these wavelengths in $RR(\lambda)$ of polymers were multiplied by zero, as shown in Figure 15.

Figure 15a shows the mean spectra of each data class, multiplied by the function presented in Figure 15b to prune the spectra such that the wavelengths susceptible to water vapour weigh zero, as shown in Figure 15c. This procedure was applied to all the data vectors of $RR_T(\lambda)$,

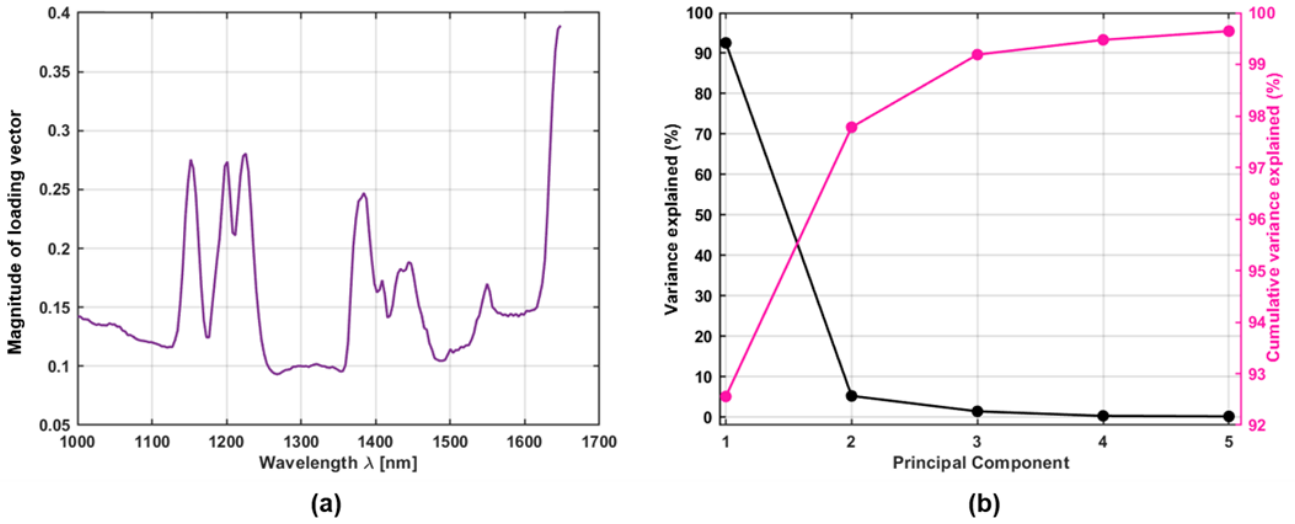

Figure 14. Analysis of $RR_T(\lambda)$ PCA, (a) the magnitude of loading vectors for first five PCs, (b) percent variance explained by the first five PCs.

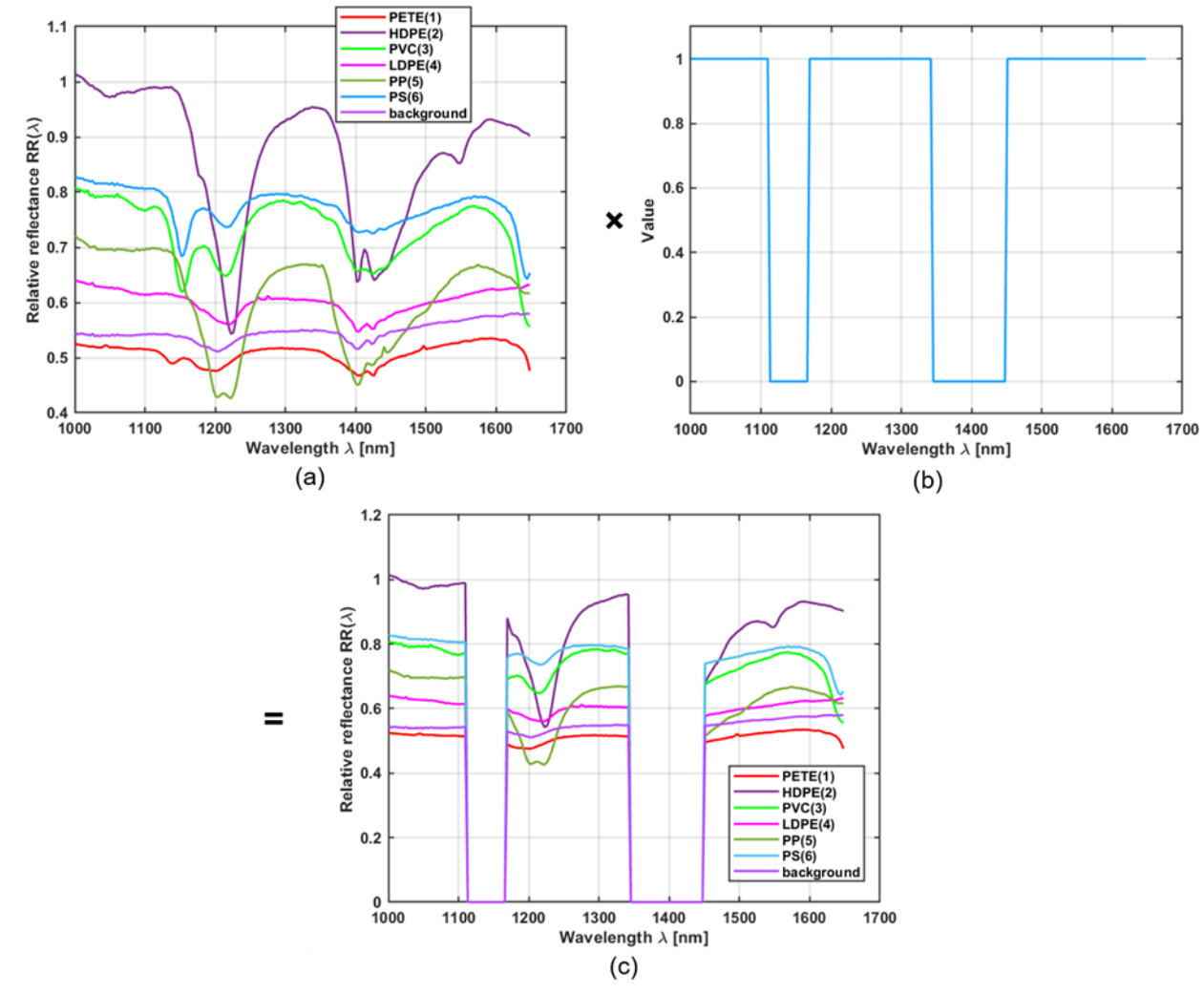

Figure 15. Pruning of wavelengths (λ) susceptible to water vapour.

$RR_D(\lambda)$ and $RR_H(\lambda)$. After the pruning of susceptible wavelength bands, PCA was applied to create the low-dimensional training and testing datasets as detailed in above. The classification, also detailed above, was performed. Table 3 shows the classification results for each method when full and pruned spectra were used, respectively. The classification results are shown with respect to the number of PCs used for the training and testing of the classifier.

Discussion

There is a major difference between using hyperspectral imaging for remote sensing and using it *in situ* at an industrial site for, e.g., classification of waste material. While the travelled path of light through the atmosphere for remote sensing could be hundreds of kilometres, the light path through air at an industrial site is hundreds of centimetres. The composition of atmospheric gases constitutes a complex optical filter causing wavelength bands to be blocked for remote sensing. Water molecules are, for example, known to create blocked bands within the SWIR region of light. It is, nevertheless, customary to assume that the SWIR region is completely available at *in situ* installations for hyperspectral measurements.

In this paper, it is shown that vapourised water has an impact even at pathlengths of hundreds of centimetres, especially in the case of large variations of RH. The method proposed to overcome this source of uncertainty in measurements is to apply frequent calibrations of the hyperspectral measurement setup such that the impact of varying composition of air is suppressed. The spectral signature of a 3D printed object made of PLA was used as a test case. We found out that the uncertainty of measurement, computed as coefficient of variation, could be reduced from 5 % to 1 % in the most error prone wavelength region.

It is also shown that the RH-induced variations in measurement can drastically affect the classification of materials using an *in situ* setup. Such an effect can be seen in confusion matrices presented in Figure 12 and Figure 13. For the given example of recyclable plastic types, PS and background were misclassified at the rate of 95.7 % and 77.4 %, respectively when $A_m(\lambda) \neq A_r(\lambda)$, whereas, for $A_m(\lambda) = A_r(\lambda)$ the misclassification rate was dropped to 4.3 % and 0 % for the same classes. As per Table 3, for the given test case, it was possible to reduce the overall misclassification rate to less than 0.4 % when $A_m(\lambda) = A_r(\lambda)$ and full spectra of polymers were used. On the other hand, it was not possible to further reduce the misclassification to less than 20 % when $A_m(\lambda) \neq A_r(\lambda)$.

Frequent reference measurements could be used as a method to make sure that $A_m(\lambda) = A_r(\lambda)$. However, it might be difficult to frequently switch analysed material with a calibration target at an industrial installation. Another solution is to include the reference surface as part of the background.[10] A second alternative method is also proposed to prune the wavelengths that are susceptible to water vapour. The classification results presented in Table 3 show that even a significant change in humidity has a minor impact on the classification if the susceptible bands are pruned. This method is shown to work well for the selected test case of polymer classification. However, for classification problems where most of the contributing variables (wavelengths) lie in the susceptible regions, this method will not work well, theoretically. Separate instrumentation for on-line recording of absorption in air could possibly be another feasible alternative. Figure 11 shows non-linear dependencies for the response of spectral bands versus RH. Alternatively, we think it might be possible to create a model for prediction of $A_m(\lambda) = A_r(\lambda)$ at measured temperatures and RH. Possible methods could be multivariate regression of empirical data or using data from publicly available spectroscopic databases.[21] However, such more advanced models of calibration are beyond the scope of this publication.

Table 3. Misclassification rate for $RR_D(\lambda)$ and $RR_H(\lambda)$.

Wavelength band	RH(%) $A_m(\lambda)$	RH(%) $A_r(\lambda)$	Matrix	Misclassification rate (%)						
				1 PC	2 PCs	3 PCs	4 PCs	5 PCs	6 PCs	7 PCs
Full	83	23	$RR_D(\lambda)$	51.5	28.4	17.5	18.6	39.2	37.3	19.9
Full	*83*	*83*	$RR_H(\lambda)$	51.9	12.5	2.04	1.3	0.8	0.9	0.37
Pruned	83	23	$RR_D(\lambda)$	51.5	14.24	3.48	2.27	1.74	1.36	0.6
Pruned	*83*	*83*	$RR_H(\lambda)$	51.9	14.09	3.48	1.66	0.9	0.53	0.15

Instead of using the raw spectra for training the classifier, PCA was applied to preprocess the spectra. The combination of PCA for preprocessing and SVM for classification was used in this research as the previous study[20] reported, the combination results in better classification of hyperspectral data. To validate the claim, we trained and tested an SVM classifier with raw spectra. The results of this classification are presented in Table 4. Comparing the results presented in Table 3 and Table 4, the classifier performance is clearly superior when PCA was used.

Furthermore, it is very possible that any other classification algorithm would perform better for the polymer classification test case. However, it is beyond the scope of this research to compare different machine learning algorithms and find the best one for any particular classification problem. We performed the classification of polymers to demonstrate that the change in humidity can affect the classifier performance significantly, irrespective of which classifier is used.

The rather small volume of acquired experimental data presented in this paper limits the empirical evidence of this study. Still, in combination with knowledge developed by the remote sensing community,[2,8] the evidence for water vapour having a significant impact on *in situ* hyperspectral imaging becomes more convincing. The analysed coefficient of variation and the results from polymer classification give a clear indication, for example to practitioners of *in situ* hyperspectral imaging, that large variations of RH require special attention to be paid to the calibration procedure.

Conclusion

The technological development of hyperspectral cameras has initially been driven by remote sensing applications. *In situ* applications are now emerging from different industries, such as waste sorting and the circular economy. Every possible way of improving measurement accuracy can potentially lead to improved classification results and higher value of recycled material. It is shown in this paper that vapourised water can cause a coefficient of variation in reflectance measurements as high as 5 % at large variation of RH for a chosen test case. It is also shown that the variations in reflectance measurements can have a drastic impact on the classification of polymers. Frequent re-calibrations of the imaging system ensure that the reference and the polymer measurements are done at the same RH levels, such that the misclassification is reduced from 20 % to less than 1 %. As an alternative to frequent re-calibrations, pruning of the wavelength bands that are susceptible to water vapour are also shown to produce equally good results for polymer classification.

References

1. B.H. Stuart and D.J. Ando, *Biological Applications of Infrared Spectroscopy*. Wiley (1997).
2. G. Sicot, M. Lennon, V. Miegebielle and D. Dubucq, "Estimation of the thickness and emulsion rate of oil spilled at sea using hyperspectral remote sensing imagery in the SWIR domain", *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* **XL-3/W3,** 445–450 (2015). https://doi.org/10.5194/isprsarchives-XL-3-W3-445-2015
3. M.S. Shaikh, K. Jaferzadeh and B. Thörnberg, "Extending effective dynamic range of hyperspectral line cameras for short wave infrared imaging," *Sensors* **22(5),** 1817 (2022). https://doi.org/10.3390/s22051817
4. G. Bonifazi, R. Capobianco, R. Palmieri and S. Serranti, "Hyperspectral imaging applied to the waste recycling sector", *Spectrosc. Europe* **31(2),** 8–11 (2019). https://doi.org/10.1255/sew.2019.a3
5. A.C. Karaca, A. Erturk, M.K. Gullu, M. Elmas and S. Erturk, "Automatic waste sorting using short-wave infrared hyperspectral imaging system," in *5th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS),*

Table 4. Misclassification rate for $RR_D(\lambda)$ and $RR_H(\lambda)$ (raw spectra).

Wavelength band	RH(%) $A_m(\lambda)$	RH(%) $A_r(\lambda)$	Matrix	Missclassification rate (%)
Full	83	23	$RR_D(\lambda)$	9.16
Full	83	83	$RR_H(\lambda)$	3.86
Pruned	83	23	$RR_D(\lambda)$	6.28
Pruned	83	83	$RR_H(\lambda)$	5.75

pp. 1–4 (2013). https://doi.org/10.1109/WHISPERS.2013.8080744
6. R. Gupta and R.I. Hartley, "Linear push-broom cameras", *IEEE Trans. Pattern Anal. Mach. Intell.* **19(9),** 963–975 (1997). https://doi.org/10.1109/34.615446
7. T. Hyvärinen, E. Herrala, W. Procino and O. Weatherbee, "Compact high-resolution VIS/NIR hyperspectral sensor," *Proc. SPIE 8032, Next-Generation Spectroscopic Technologies IV* 80320W (2011). https://doi.org/10.1117/12.887003
8. K.P. Shine, A. Campargue, D. Mondelain, R.A. McPheat, I.V. Ptashnik and D. Weidmann, "The water vapour continuum in near-infrared windows – Current understanding and prospects for its inclusion in spectroscopic databases," *J. Mol. Spectrosc.* **327,** 193–208 (2016). https://doi.org/10.1016/j.jms.2016.04.011
9. G. Yang, C. Li, Y. Wang, H. Yuan, H. Feng, B. Xu and X. Yang, "The DOM generation and precise radiometric calibration of a UAV-mounted miniature snapshot hyperspectral imager", *Remote Sens.* **9(7),** 642 (2017). https://doi.org/10.3390/rs9070642
10. A. Noviyanto and W. H. Abdulla, "Segmentation and calibration of hyperspectral imaging for honey analysis", *Comput. Electron. Agric.* **159,** 129–139 (2019). https://doi.org/10.1016/j.compag.2019.02.006
11. R. Pillay, J.Y. Hardeberg and S. George, "Hyperspectral imaging of art: acquisition and calibration workflows", *J. Am. Inst. Conserv.* **58(1–2),** 3–15 (2019). https://doi.org/10.1080/01971360.2018.1549919
12. M.S. Shaikh, K. Jaferzadeh, B. Thörnberg and J. Casselgren, "Calibration of a hyper-spectral imaging system using a low-cost reference", *Sensors* **21(11),** 3738 (2021). https://doi.org/10.3390/s21113738
13. *Rallye 3003 Compact* [Online]. https://hella.se/produkter/extraljus/rallye-3003-compact/. [Accessed: 09-Mar-2022].
14. *DEM500 Humidity Sensor* [Online]. https://www.velleman.eu/products/view/?id=420578. [Accessed: 10-Mar-2022].
15. *Beurer LB88 Humidifier* [Online]. https://www.beurer.com/web/gb/products/wellbeing/air-and-aroma/air-purifiers_air-washers_air-humidifiers/lb-88-black.php. [Accessed: 08-Mar-2022].
16. L. Eriksson, E. Johansson, N. Kettaneh-Wold, J. Trygg, C. Wikstrom and S. Wold, *Multi- and Megavariate Data Analysis Part I Basic Principals and Applications*, Second rev. Umetrics AB (2006).
17. P. Jonsson, J. Casselgren and B. Thornberg, "Road surface status classification using spectral analysis of NIR camera images", *IEEE Sens. J.* **15(3),** 1641–1656 (2015). https://doi.org/10.1109/JSEN.2014.2364854
18. H. Jiang, C. Zhang, Y. He, X. Chen, F. Liu and Y. Liu, "Wavelength selection for detection of slight bruises on pears based on hyperspectral imaging", *Appl. Sci.* **6(12),** 450 (2016). https://doi.org/10.3390/app6120450
19. *ClassificationECOC* [Online]. https://www.mathworks.com/help/stats/classificationecoc.html. [Accessed: 09-Mar-2022].
20. Y. Huang, J. Li, R. Yang, F. Wang, Y. Li, S. Zhang, F. Wan, X. Qiao and W. Qian, "Hyperspectral imaging for identification of an invasive plant *Mikania micrantha* Kunth", *Front. Plant Sci.* **12,** (2021). https://doi.org/10.3389/fpls.2021.626516
21. L.S. Rothman *et al.*, "*The HITRAN2012 molecular spectroscopic database*", *J. Quant. Spectrosc. Radiat. Transf.* **130,** 4–50 (2013). https://doi.org/10.1016/j.jqsrt.2013.07.002