In-line industrial contaminants discrimination for the packaging sorting based on near-infrared reflectance spectroscopy

A proof of concept

Manuele Bonaccorsi, Giovanni Rateni, Filippo Cavallo and Paolo Dario
The BioRobotics Institute
Scuola Superiore Sant'Anna
Pisa, Italy
manuele.bonaccorsi@santannapisa.it; giovanni.rateni@santannapisa.it; filippo.cavallo@santannapisa.it; paolo.dario@santannapisa.it

***Abstract*—The Industry 4.0 paradigm requires new technologies and methods not only to improve the profitability and the quality of the industrial production and products, but also new strategies to reduce the social and environmental impact of the production process. Many line manufacturing chains unbox and assembly components to create products, but create a large amount of waste that sometimes can't be recycled because of the exposure to contaminants. When it comes to the automotive industry, mineral oils may contaminate plastic packaging and cardboard boxes during manufacturing, making hard to recycle them. In this paper we propose a proof of concept of a packaging sorting system based on NIR spectroscopy, to automate sorting and get high quality outputs for the recycling of cardboard package boxes. Spectral datasets have been pre-processed and dimensionally reduced using PCA. A SVM algorithm has been trained to distinguish between oil contaminated and non contaminated materials. Two NIR spectrometers with sensing range 640-1050 nm and 950-1650 nm have been used and evaluated, to select the proper sensor configuration. Eventually, the system classification accuracy was respectively up to the 98,68% and 98,64% using the 950-1650 nm and the 640-1050 nm spectrometers, demonstrating the opportunity to detect mineral oil contamination on boxes.**

Keywords—Industry 4.0; contaminant discrimination; packaging sorting; recycling; NIR spectroscopy; chemometrics; multivariate analysis.

I. Introduction

Traditional waste sorting methods in the recycling industry mainly depend on the processed material since metals, glass, plastics, paper, and woods have very different physical properties. Most advanced sorting techniques, like indirect sorting, introduce advanced sensors and techniques including Laser Spectroscopy, X-rays, Optical Sorting, Hyper-Spectral Imaging and Reflectance Spectroscopy [1]. In the indirect sorting, sensors are used for detecting recyclable materials in the bulk input waste followed by segregation using various actuators or robotic systems as in [2]. The indirect sorting using Near-Infrared Reflectance (NIR) spectroscopy allows faster and safer identification of different materials respect to other methods such as X-Ray and Laser spectroscopy, while is cheaper than using expensive hyperspectral cameras. It is successfully used for chemical analysis, materials classification and quality increase of recycled products. Indeed, high purity recycled materials ensure optimum technical and environmental performances, while impurities or contaminations can cause a reduction in quality. For this reason it is today more and more investigated for advanced sorting. Recently, the use of NIR spectrometers has been investigated in the industrial field to meet specific standards determined by industrial applications, such as characterizing exhausted lubricating oils [3,4], in the separation of plastic packaging materials [5], to separate recycled building and demolition inerts [6], and in the separation of composite wood-plastic materials [7].

Taking in account for the state of the art, the identification of mineral oil contamination on cardboard and plastic packaging is still a research challenge, with significant implication in recycling and industry 4.0. The proposed solution could positively contribute to the definition of new automated methods for advanced and high quality sorting.

II. Materials and Methods

A. Samples

Traditional packaging from automotive industry were selected including four different type of cardboard boxes. Boxes had three different background colours (light brown, dark brown and white) and black and blue inks printed inscriptions. The different backgrounds and the coloured inks of prints have been introduced as confounding factors both for the NIR sensors and for the classification algorithm. Then the

978-1-5090-1012-7/17/$31.00 ©2017 IEEE

cardboard boxes were greased using engine lubricant oil and used as contaminated samples.

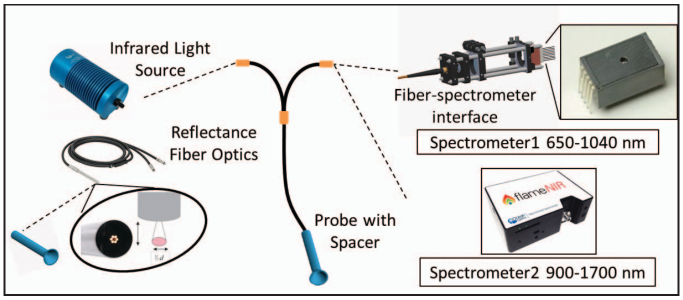

Fig. 1. Experimental setup schematics.

B. Instrumentation

Two different spectrometer sensors were tested:

- MS-series C11708MA from Hamamatsu Photonics (Shizuoka, JP). The mini-spectrometer integrates a CMOS sensor with wavelength sensing range 640-1050 nm and spectral resolution of max 20 nm. In order to provide SMA 905 connector to the sensor, it was coupled with a customized optical interface, realized by means of several optical components, including a SMA adapter, an aspheric collimator, and two cylindrical lens from Thorlabs (Newton, USA).
- FlameNIR from Ocean Optics (Dunedin, USA). The sensor, featuring an uncooled InGaAs detector with wavelength range 950-1650 nm, has a spectral resolution of max 10 nm;

A modular experimental setup was assembled as in Fig.1, consisting of a high-power infrared light source with range 360-2400 nm from Ocean Optics (Dunedin, USA), connected to the measuring system through a RP23 fiber optic reflection probe from Thorlabs (Newton, USA).

C. Experimental design

The two spectrometers were used to acquire NIR reflectances spectra directly on the sample surfaces. The probe was placed perpendicularly, 8 mm above 10 surfaces from different samples: (i) four different clean cardboard samples, (ii) black and blue prints on cardboards, and (iii) three contaminated cardboard surfaces. For each surface, 100 different acquisitions were saved, for a total of 1000 observations (TABLE I). The same integration time (465 μs) was set on the two spectrometers.

TABLE I. SAMPLES

Type	*Material type (repetitions)*	*Observations*
Clean	light brown cardboard (200) dark brown cardboard (200) white cardboard (100) black-ink prints (100) blue-ink prints (100)	700
Contaminated	Oil on light brown cardboard (200) Oil on dark brown cardboard (100)	300

D. Methodology

The observed 1000 spectra contained both chemical information and noise. Samples were initially pre-processed to filter out the noise (stray light, light scattering, detector non-linearities, etc), according to the following steps:

1. Outlier rejection: using the Hampel method [8] to improve the classification ability of multivariate methods;
2. Standardization: using the Standard Normal Variate (SNV) transformation;
3. Weighting: samples were weighted to compress the dynamics in the range [0,1].

The Leave-Multiple-Out (LeaveMOut) cross-validation method was implemented leaving out the 10%, 30% and 50% of the dataset to validate the proposed classification system. The dataset was randomly divided into training and test sets. Then the compositional information were extracted performing the Principal Component Analysis (PCA) on the training-set [9]. Eventually, the samples were projected on the first two principal components space. A Support Vector Machine (SVM) algorithm was trained on the projected training-set, to distinguish between contaminated and non contaminated cardboards. A fourth-order polynomial function was selected for the samples classification. After the training, the PCA was applied to the test-set, and the trained SVM classified the PCA outputs (first two components). The LeaveMOut method was repeated 1000 times on the dataset, to achieve statistical significance. The PCA, the training of the SVM and the LeaveMOut method were implemented in Matlab version R2012a from MathWorks (Natick, USA).

III. RESULTS

The proposed system successfully distinguished and classified contaminated cardboard samples from non contaminated ones, even in the presence of confounding factors, including black and blue ink prints. The classification accuracy was up to 99.68%, performing the spectral analysis in the range 950-1650 nm. Using the 640-1050 nm NIR bandwidth, the accuracy was up to 99.64%. Nevertheless, the latter can be considered a cost-effective solution for low-cost applications.

The worst classification performance was achieved leaving 50% of the samplings for the training of the SVM and using the remaining 50% of samples for testing. In this case the maximum number of classification mistakes were 79 and 59 respectively for the Hamamatsu and the Ocean Optics spectrometers. Using the 90% of the samples for the training and the 10% for the testing, the maximum classification error decreased up to 34 and 33 respectively for the Hamamatsu and the Ocean Optics sensors. The classification accuracy thus mainly depended on the training-set size, and in a residual manner on the spectrometer sensing range. The results of the LeaveMOut validation process are reported in TABLE II.

TABLE II. Classification accuracy of the proposed system using the Ocean Optics and the Hamamatsu sensors

Spectrometer sensing range	**950-1650 nm**			**640-1050 nm**		
Training-set [% of sample size]	**10%**	**30%**	**50%**	**10%**	**30%**	**50%**
Maximum classification errors	33	59	44	34	40	79
Mean classification errors	3.21	6.40	10.44	3.60	7.48	12.30
Mean classification accuracy	99.68 %	99.36 %	98.96 %	99.64 %	99.25 %	98.77 %

IV. Conclusions

The proposed work highlights the opportunity to perform an accurate classification of oil-contaminated cardboard from non contaminated one, to improve the purity of the sorting output and the quality of the next recycling process. The achieved results positively impact on to the research of new sorting methods, by analysing the combination of synthetic oil and cardboard in the NIR spectrum. Furthermore, the achieved classification accuracy fits the previous works on the classification/sorting of contaminated materials. Indeed the work proposed in [7], claimed an output purity of the 97.1% in the sorting of wood samples, respect to wood-plastic samples, and a purity of the 98.9 % in the sorting of wood respect to wood contaminated with preservatives. This preliminary work demonstrates that both the sensing ranges 950-1650 nm and 640-1050 nm can be profitably used. Future activities will concern the analysis of more cardboard samples and the introduction of more confounding factors, including water drops, different types of inks and prints, applied plastic tape, foils or sheets. Furthermore, different pre-processing techniques will be investigated, as well as the opportunity to implement different classification techniques, including K-Means and neural networks.

Acknowledgment

This work was supported by the Tuscany's regional research project CENTAURO: Colavoro, Efficienza, preveNzione nell'industria dei motoveicoli mediante Tecnologie di AUtomazione e RObotica. bando FAR-FAS 2014.

References

[1] Gundupalli, S. P., Hait, S., & Thakur, A. (2016). A review on automated sorting of source-separated municipal solid waste for recycling. Waste Management.

[2] Lukka, T. J., Tossavainen, T., Kujala, J. V., & Raiko, T. (2014). ZenRobotics Recycler–Robotic sorting using machine learning. In *Proceedings of the International Conference on Sensor-Based Sorting (SBS)*.

[3] Marinović, S., Jukić, A., Doležal, D., Špehar, B., & Krištović, M. (2012). Prediction of used lubricating oils properties by infrared spectroscopy using multivariate analysis. *goriva i maziva*, *51*(3), 205-215.

[4] Balabin, R. M., Safieva, R. Z., & Lomakina, E. I. (2011). Near-infrared (NIR) spectroscopy for motor oil classification: From discriminant analysis to support vector machines. Microchemical Journal, 98(1), 121-128.

[5] Kassouf, A., Maalouly, J., Rutledge, D. N., Chebib, H., & Ducruet, V. (2014). Rapid discrimination of plastic packaging materials using MIR spectroscopy coupled with independent components analysis (ICA). Waste management, 34(11), 2131-2138.

[6] Vegas, I., Broos, K., Nielsen, P., Lambertz, O., & Lisbona, A. (2015). Upgrading the quality of mixed recycled aggregates from construction and demolition waste by using near-infrared sorting technology. Construction and Building Materials, 75, 121-128.

[7] Mauruschat, D., Plinke, B., Aderhold, J., Gunschera, J., Meinlschmidt, P., & Salthammer, T. (2016). Application of near-infrared spectroscopy for the fast detection and sorting of wood–plastic composites and waste wood treated with wood preservatives. Wood Science and Technology, 50(2), 313-331.

[8] Zhao, Y., Lehman, B., Ball, R., Mosesian, J., & de Palma, J. F. (2013, March). Outlier detection rules for fault detection in solar photovoltaic arrays. In *Applied Power Electronics Conference and Exposition (APEC), 2013 Twenty-Eighth Annual IEEE* (pp. 2913-2920). IEEE.

[9] Berrueta, L. A., Alonso-Salces, R. M., & Héberger, K. (2007). Supervised pattern recognition in food analysis. *Journal of Chromatography A*, *1158*(1), 196-214.