MDPI

Article

1. Full-Stokes imaging polarimetry for classification of polyethylene and non-polyethylene coated paperboard
2. Polarimetric detection and classification of polyethylene coating in paperboard

Javier Brugés Martelo [1,†,‡], Jan Lundgren [1,‡] and Mattias Andersson [2,*]

[1] Electronic Design department, Mid Sweden University, Holmagatan 10, Sweden; javier.bruges@miun.se, jan.lundgren@miun.se
[2] Department of Design, Mid Sweden University, Örnsköldsvik; mattias.anderson@miun.se
* Correspondence: javier.bruges@miun.se

Version October 13, 2020 submitted to Sensors

Abstract: Manufacturing of high quality extruded low-density polyethylene (PE) paperboard intended for the food packaging industry relies on manual and intrusive off-line inspection by the process operators to assess the overall quality and functionality of the product. Defects like pinholes, i.e. absence of the coating after extrusion or lamination, can occur at any location in the reel affecting the overall sealable property, which is important for customers intending to extend the life of the contained product. We investigate the use of novel full-Stokes imaging polarimetry for the material classification of PE coated paperboard, where the classification task is to discern between the presence or absence of the coating. From the different parameters obtained by polarimetry measurements, we propose the implementation of feature selection based on distance correlation statistical method and subsequently the implementation of a support vector machine algorithm that uses a nonlinear Gaussian kernel function. Our results open the possibility to adopt this optical system for in-process control of the manufacturing of coated paperboard, to increase its productivity and minimize waste of material.

Keywords: imaging polarimetry, Stokes parameters, extruded plastic coatings, supervised learning

1. Introduction

CHAGING STRUCTURE IN INTRO:1st. paperboard MM, 2nd. state of the art, 3rd. imaging polarimetry

Barrier functionality - contained product - extend life shelf

High quality products required constant monitoring of production parameters - in-process control vs off-line measurement.

Large area sensors are required - real-time monitoring, high spatial resolution with adequate form factor

Plastic coated paperboard for packaging purposes plays an important role in extending the life shelf of the contained product intended for the food, pharmaceutical, and cosmetic industry.

Extruded and laminated plastic coatings are equally used in one or both sides of the paperboard to increase its durability, barrier properties and it affects the printability and visual appearance. Barrier functionality is evaluated during production by human inspection, which is seen as destructive, intrusive, and time consuming, and in the meanwhile, production is running cannot avoid the

propagation of further defects. Pinholes are common defects found in extruded paperboard, i.e. missing areas of the coating. These defects decrease its barrier functionality and the quality of the end product is measured by the number of micrometer size holes, called pinholes, within the products area characterize during the off-line inspection. Low grammage on the plastic recipe, irregularities in the base paperboard (e.g. a non-desire texture, high surface roughness, or standing fibers), or unevenness of the coating profile can originate defects related to the top plastic coating in the paperboard [paperboard manufacturing]. Increasing the surface smoothness of the product affects directly the visual perception. If the extrusion of the plastic process is added, the requirements to maintain a higher smoothness property increases due to increased visibility of the base layer inhomogeneity. All these lead to a careful selection of the parameters for including the extrusion process. Compromises are made based on the desired level of smoothness and the functional barrier characteristics such as moisture, fat, and heat barrier. Defects on the surface structure can occur at any point in the manufacturing process, i.e. previous and after adding the extruded layer. While the product usually is inspected along and at the end of the production line, online monitoring of the product is of prime importance in early detection to avoid defect propagation resulting in big losses in the production. Optical systems are non-invasive techniques commonly used in metrology for high-quality products. [REFERENCE]TALK ABOUT NIR AND NUCLEAR SENSORS FOR PACKAGING INDUSTRY...blablabla... with the disadvantage that these sensors are expensive, point type sensor, and can harm humans along the production line (depend on the source). The need for a higher spatial resolution sensor.

Polarization is one of the main constituents of light, along with intensity and wavelength. As humans we can perceive optical material properties by the way light is reflected from the material and the colors are perceived. However, polarization is only available to us with the use of external hardware and optical components. Similarly, in many industrial applications, machine vision systems are being used by employing camera based sensor, which often uses color or intensity from the scene to inspect the product. Wavelength is uncorrelated to polarization and the relation of intensity and polarization in an machine vision application is often used to eliminate the haze created by high specularities in the scene. Machine vision is implemented in industry to aid in-process monitoring of product quality, defect inspection and material classification.

Imaging polarimeters are not new in the used for material classification. The technique has been employed in remote sensing [ref - Tyo, Goldstein], astronomy [ref - Bueno], aerospace and defence (target detection) [ref - Walter], biomedical [ref -Valery], and industrial manufacturing [ref - Meriaudeau]. Polarimeters can provide partial or full polarization information of the captured incoming light and this depends on the optical elements employed and the imaging system configuration. The metrological properties of polarization has being restricted industrially to ellipsometry measurement systems. These systems are bulky and require apriori knowledge of the samples to obtain its optical properties. Instruments are are often single-pixel detector and their use are limited by surface roughness and light source coherence length. Imaging polarimetry offers an alternative to measure polarization cues that are described by the Stokes parameters, i.e. $\vec{S} = (S_0, S_1, S_2, S_3)^T$. The mathematical formulation of the Stokes parameters is found convenient since it displays several degrees of freedom and it can be further computed to obtain complementary parameters influenced directly from the light - material interaction. Stokes parameters are measured as radiance quantities, which also make them convenient since read out from a camera system is performed by the scene radiance. Different configurations of imaging polarimeters are found in literature serving for specific metrological purposes [ref - ??]. The advent of new camera systems with embedded micro-polarizer arrays in a division-of-focal-plane configuration has increased the possibilities to have more compact polarimeters. These polarimeter camera systems measures the scene in terms of the first three Stokes parameters, i.e. linear polarization, and to complement the third component, i.e. the circular polarization relation, under the same captured scene an instantaneous full-Stokes imaging polarimetry has been devised [ref - Tu].

CTo provide a solution for the online monitoring of coated paperboard, we implemented the use of a full stokes imaging polarimetry camera system for the detection of defects presented in different samples of coated paperboard, which evaluates the polarization cues and detects defects in the material associated to the coating. This novel optical system can measure all stokes parameters on a single shot [ref - Tu] providing direct information of the material optical properties. Material classification is one part of industrial inspection.

In this article, we present a polarization reflectometer based on full stokes polarimetry system capable of pixel-wise classification for the presence or absence of extruded PE coating on high quality paperboard. The generalization of the classification technique proves that the systems are capable to discern between the presence or absence of the extruded coating on the paperboard. In our understanding, this serves as the base for the detection of defects like pinholes, where the absence of coating can be realized from this experiments and a modification of some optical elements in the system is possible to meet the requirements for this kind of problem.

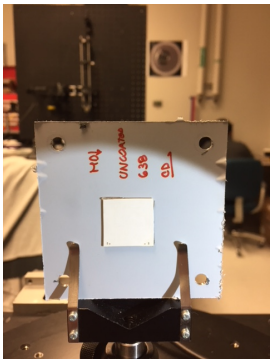
2. Material and methods

We summarize our strategy on three main parts as follows.

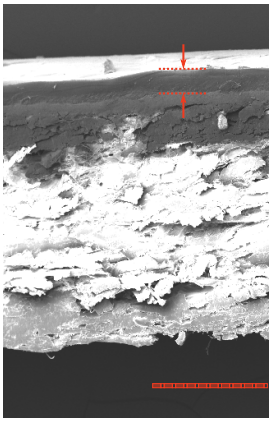
1. Polarimetric measurement of PE and non-PE coated paperboard samples by the full-Stokes imaging polarimeter
2. Implementation of distance correlation function for feature selection of the polarimetric measurements
3. Model training and validation for material classification using supervised learning algorithm, i.e. support vector machine (SVM).

Can you explain more about the sample? What are the good sample and pin holes sample that you used for the training of your SVM? How you determine the how bad is it with the number of pin holes and what are the threshold for the pin holes density to be classified as bad? Is that possible to correlate between the pin hole density and the measurement results? The paperboard samples in this study were industrially manufactured by a local company. A4 sheet sizes were extracted from the paperboard reel in pairs, before and after the extruded low density polyethylene (PE) coating was applied. The paperboard is of a multi-ply structure,which is composed of a cellulose bottom layer coated with a mixture of binders and fillers on one side of the cellulose base. The coating mixture is added before the extruded PE coating to increase the smoothness of the multi-ply structure which promotes good adhesion and homogeneity of the thickness of the PE coating. It also attempt to eliminate the effects of standing fibers, i.e. protruded fibrous from the cellulose material, that can brake the coating layer and cause the appearance of pinholes in the product. Pinholes are a main problem related to the decreased overall quality in packaging materials and reducing its barrier properties. For the polarimetric measurement, 20x20mm area from the samples were extracted and attached to a sample holder as shown in figure 1(a). An SEM image of the structure of the multi-ply PE coated paperboard is also shown in figure 1(b), were the PE coating is in average 13 thick. The low average roughness from the multi-ply base, the close similarity of the base and PE coating refractive indexes and the micrometer thickness dimension of the PE coating, makes challenging the classification for traditional machine vision systems based only on intensity and spectral radiance values.

What is the other hand?On the other hand, light reflection by dielectrics tend to polarized the light and materials were refraction and scattering occur polarization is modified at the different optical paths the beam encounters. In this sense, the measure of polarization states conveys more information if possible to measure by a detector. As shown in figure 2, each sample is measured independently by the imaging polarimetry, capturing the complete polarization radiance by the sample. A white LED illuminates the sample with vertically polarized light. If the incident light is fully polarized, as it is in this experiment, the beam at the the sample will be partially polarized or completely depolarized. This is quantified by the imaging polarimetry in the form of the Stokes vector **??**. For a complete description



(a)



(b)

Figure 1. *(a)* adapted 20x20mm sample illuminated by the optical system. *(b)* SEM image of the PE coated paperboard sample. The base layer composition is similar for both samples, i.e. with and without the extruded PE coating

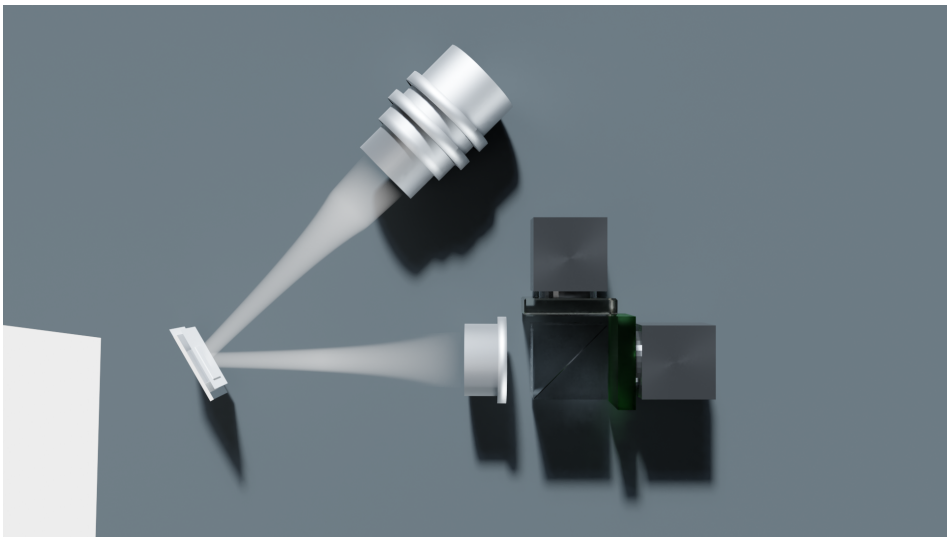


Figure 2. Schematic of the optical system composed of a polarized broadband LED for illuminating the samples and a full-Stokes imaging polarimetry for recording the polarization and spectral information

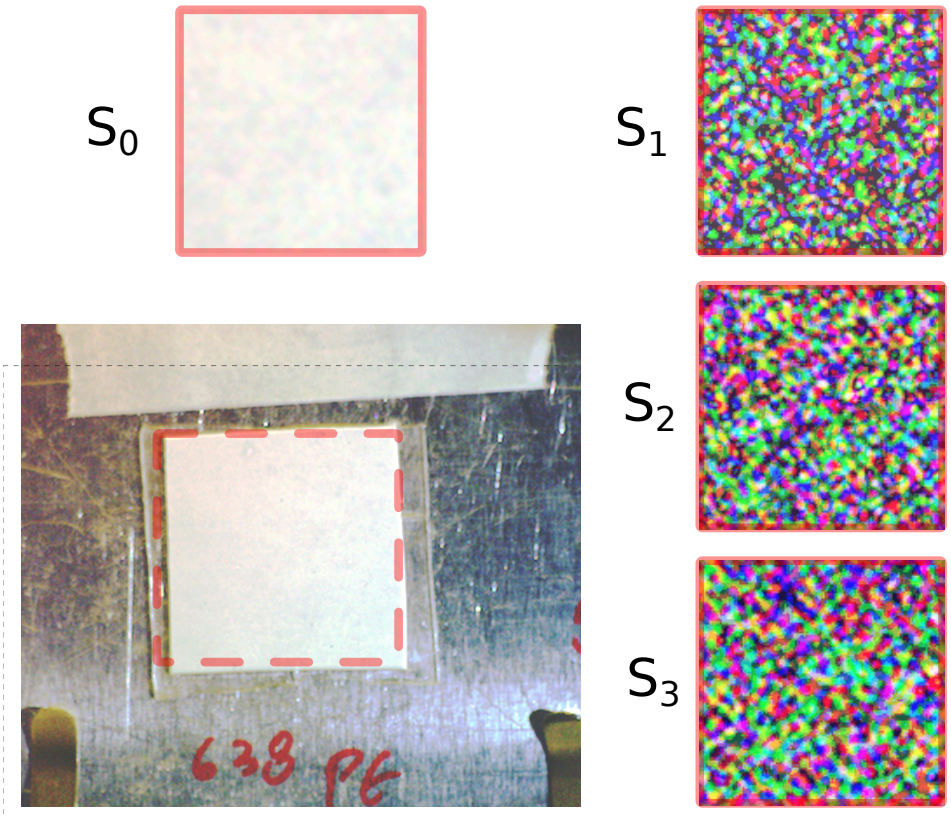


Figure 3. Stokes vectors from image add some colormap 0 to 1... check napari app to choose appropiate colorcode

of the full-Stokes imaging polarimeter, the authors suggest to the reader referring to *Tu - , et.al.*[ref - Tu], where the design, calibration and demosaicing algorithm is detailed. A brief description of the main components in the optical system is described here. The optical system uses a division-of-amplitude configuration where two balanced optical paths with complementary polarization information are divided. This allows the instantaneous measurement of the polarization. The first element in the system is an imaging lens for focusing the system on the surface of the imaged object. A 50/50 beamsplitter (BS) divides the incident beam in two balanced paths. As mentioned, two complementary polarization measurements are required to obtain the Stokes vector. In one arm, an achromatic quarter wave plate (AQWP) will modify the beam into its circularly polarized components. On the second arm, to balance the path difference created by the AQWP, a glass of similar refractive is placed. Finally, in both arms linear Stokes polarimetric cameras captured the scene. With the appropriate demosaicing and calibration algorithms, calibration and synchronization of the cameras makes possible to obtain pixel-wise polarization and spectral radiance of the scene in one single capture.

Figure 3 shows the complete Stokes image representation of the samples. The $20x20mm$ sample is resolved with $37\mu m$ pixel information. The on-axis set-up, shown in figure 2 was built to rotate the camera system around the samples and illumination was fixed at 45 degrees respect to the sample's surface normal. Several in-plane measurements were taken at different angles around the sample with 5 degrees steps. Previous to the parameter selection for the experiment, we analyzed all different positions in order to determine the best angle configuration. It was found that the best angle between the camera and the normal to the surface of the sample was between 35 and 55 degrees, i.e. around the specular direction. In dielectrics, when illuminating with vertical polarization, roughness increases affect to a great extent the angles of polarization due to the retardation in the light wave phase component [diffusereflectionImagingAtkinson].

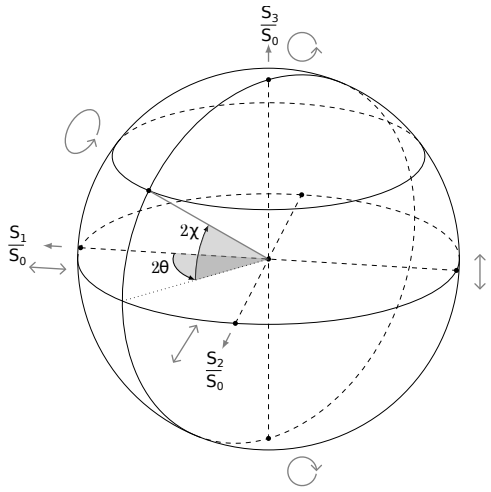


Figure 4. Observable Poincaré sphere of polarization states. To any point on the sphere is associated a unique polarization state either described by spherical angular coordinates $(2\theta, 2\chi)$ or normalized Stokes parameters S_1/S_0 , S_2/S_0 and S_3/S_0

In many imaging applications found within remote sensing and biomedic, the derivation of polarimetric parameters from the measured Stokes vector is rather desired than using directly the later for the classification task. These polarimetric parameters are associated with the material's optical properties and they can be easily mathematically formulated. We calculated the three polarimetric parameters from the initial Stokes vector, the degree of polarization (DoP) p, the azimuth θ and the ellipticity χ, and we first considered them as our initial features for the classification task. The degree of polarization (DoP) p is expressed as

$$p = \frac{\sqrt{{S_1}^2 + {S_2}^2 + {S_3}^2}}{S_0} \tag{1}$$

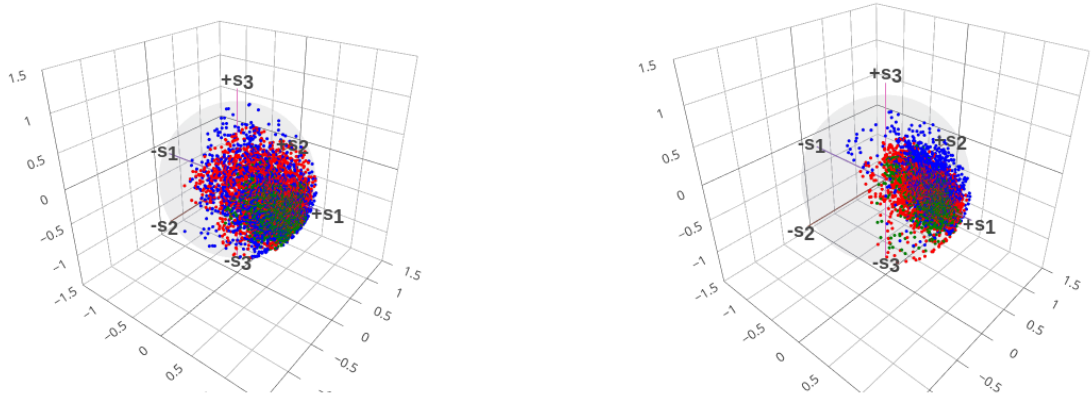
that is inversely related with the depolarization effect, i.e. the ratio at which polarized light is partially polarized or totally unpolarized by the material under polarized illumination. Using the three last components in the Stokes vector like axis in a Cartesian coordinates and representing the polarization measurements inside this coordinate space, angular relations can be derived. Two angles can be calculated, the azimuth θ and the ellipticity χ as,

$$\chi \ = \tan\left(\frac{1}{2}\arcsin\left(\frac{S_3^2}{\sqrt{S_1^2+S_2^2+S_3^2}}\right)\right) \ , \ -\pi/4\,\chi \ \ +\pi/4 \tag{2}$$

$$\varphi \ = \ \frac{1}{2}\arctan\left(\frac{S_2}{S_1}\right) \ , \ 0\,\varphi \ < \ \pi \tag{3}$$

where in the ellipticity angle χ, the extremes of the inequality represent a fully circularly polarized beam located at either sides of the poles, $\chi = 0$ being linearly polarized otherwise light is elliptically polarized, and in the azimuth is often found in literature as the angle of linear polarization (AoLP), returns the linear relation with of the observed beam since it only concern about the linear polarization components in the Stokes vector in the equation.

A Poincaré sphere, first presented by Henri Poincaré[Ref - Poincaré], aids to visualize these angular relations (Figure 4). Figure 5, the two samples are shown, where the points in the sphere represent the normalized Stokes vector, i.e. the vector is normalized respect to the intensity value of S_0, and the points angular distribution shows how it differs from the different materials optical properties.



(a) (b)

Figure 5. Poincaré representation of a (a) non-plastic coated paperboard and (b) plastic coated paperboard. The RGB nature of the camera is represented by the colors in the sphere.

From the obtained polarimetric parameters, we ranked them by its importance for classification using a metric named distance correlation, where we looked for the pair of features ranking the highest. Sample correlation functions are one measure of the linear and nonlinear relationships between variables and combining the distance metric from the observations, one can measure the correlation between the features. The distance correlation function presented by Szekely *et.al.* [Ref - Szekely2007] has demonstrated to be useful in feature selection for relative small datasets with high dimensionality [Ref - Brankovic etal]. This metric does not require any *a priori* knowledge associated to the features distribution and it has advantages over more common used statistical tests, like Pearson correlation, which cannot detect nonlinear relationships between the evaluated variables. An extract from Szekely articles implementation[Ref - Szekely2007] follows. We first obtained the define distance covariance for each combination of features,

$$dCov(Fx, Fy) = \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} D(fx_i, fx_j) \cdot D(fy_i, fy_j) \tag{4}$$

where $D(fx_i, fx_j)$ and $D(fy_i, fy_j)$ are the centered Euclidean distances of the features Fx and Fy. Finally, we can calculate the distance correlation *(dCor)*,

$$dCor(Fx, Fy) = \frac{dCov(Fx, Fy)}{\sqrt{\nu_{Fx}^2 \cdot \nu_{Fy}^2}} \tag{5}$$

where ν_{Fx}^2 and ν_{Fy}^2 are the positive feature variances. If Fx and Fy are independent features then the dCov will be equal zero. We are interested in features with the largest distance correlation, beneficial in the class separation and classification accuracy of the supervised learning algorithm. The result is a selection of the best features and it is meant to be a preprocessing step in our machine learning approach. This reduces the computational cost without compromising accuracy in the resulted classification model.

2.1. SVM algorithm

Here you have a lot of theory about SVM. I don't know if it is important for the reader to understand the algorithm or a short Wikipedia type summary is enough. You should consider who will read your article. Also here there are no details about the features that you used, the number of samples, the training method etc. Basically your important contribution is in figure 8 - the flow

diagram, and you should explain that more, how the SVM applied in your particular case, rather than just tell the user the general theory. Support vector machines (SVM) are a set of supervised learning algorithms used for classification and regression problems. In the event of classification, where two or more classes are required to be unequivocally distinguished, an SVM will map the output of two classes into an n-dimensional space, where n is defined by the number of parameters of both classes. The classification task in an SVM is to find the maximum margin separable of two classes. Linear functions are used to make a hypothesis and a manifold, also known as hyperplane, will divide the two classes with a maximum margin for the classification of novel data. The introduction of this ML algorithm was developed on the basis of optimization theory for training the algorithm that implements a learning bias resulting from a linear classifier ([Vapnik]). Most real-life cases demonstrate a non-linear relation of the features and the data cannot be linearly separable without resulting in non-optimal solutions with large misclassification results. New classification techniques were adopted to solve non-linear classification problems [Doucet].

The most powerful concept found nowadays in SVM is the use of kernel tricks [futureVapnik], where a non-linearly separable problem is input to an algorithm that will find a solution in a higher dimensional (dual) space, that in the functional space will be impossible to determine using a parametric function. If the solution is mapped back to the original dataset space, a decision boundary will be created by the location of the support vectors.

Before the parametrization of the kernel coefficients and penalization parameter, it's possible to obtain a metric to determine how many data points from the original dataset are necessary for optimal training of the machine learning algorithm. It is common practice when using a large dataset to divide it into $80\% - 20\%$ for training and validation, respectively. However, when the dataset is limited and the features are independent it is possible to use a learning curve validation approach. The learning curve provides information about the performance of a specific kernel when trained by different amounts of data. Using the learning curve, we obtained an optimal ratio of $50 - 50\%$ for training and validation, respectively. In figure 6 can be observed that around half of the datapoints the performance of the algorithm reach plateau. This lead to the division of the dataset for training and validation into equal parts for use in a subsequent experiment.

The next step in the validation is to use statistical score validation. Since SVM does not account for statistical score interpretation, the use of cross-validation is recurrent to find the statistical significance of the model's accuracy. One way of doing this is to repeat the training of the algorithm k number of times and summarize the scores statistically, named *k*-fold cross-validation. To reach the validation, the training of the algorithm requires a new training dataset each time. In *k*-fold cross-validation, the training dataset selected previously can be divided into *k*-number of parts. Each *k*-th part is called a fold, and it is completed with a validation dataset that is formed by the rest of the non-use training dataset. The implementation uses a recursive algorithm that measures the score of the *k*-th training. Finally, all the scores can be interpreted with a correlation function like the mean value and standard deviation of the scores. Performing *k*-fold cross-validation for each kernel function can be used for comparing performance and kernel parametrization.

Computing an SVM algorithm from the basis of optimization requires the use of some constants, named parameters, to find the hyperplane in the dual space. The kernel trick will attempt to find a hyperplane based on the parameters associated to the kernel function. In the case of a Gaussian kernel the parameter γ is optimized and in the case of a polynomial kernel the degree associated to it (e.g. 2, 3 and so on depending on the polynomial degree). In case of the Gaussian kernel, larger values of γ will generate a stronger bond of the decision boundary around the support vectors which can lead to a poor generalization of the algorithm. Similar is to the degree of the polynomial associated to the polynomial kernel. A penalization parameter C is used as well in all SVM algorithm and as its name suggest it will provide as a trade-off between stability and missclassification. In other words, a low C will not penalize as hard missclassification as a larger C will do, for example. One can chose to perform *k*-fold cross-validation together with hyper-parameter search. A group of parameters for γ and C are

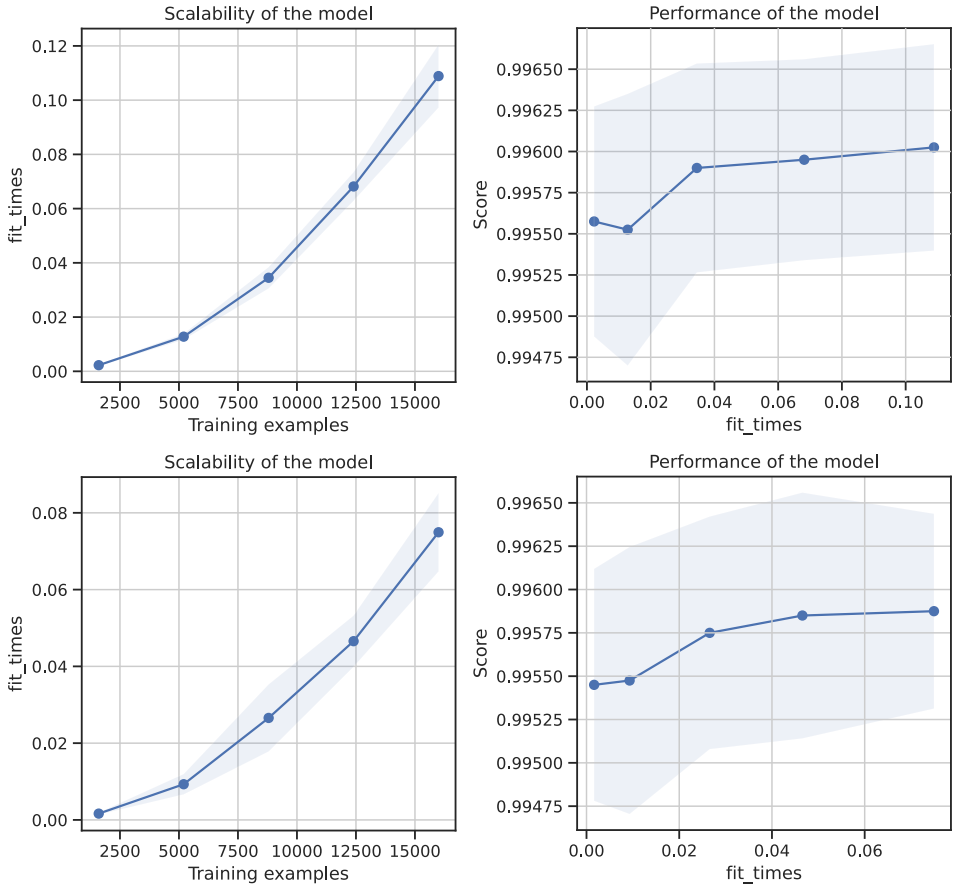


Figure 6. Performance of two kernels when trained with different amounts of data

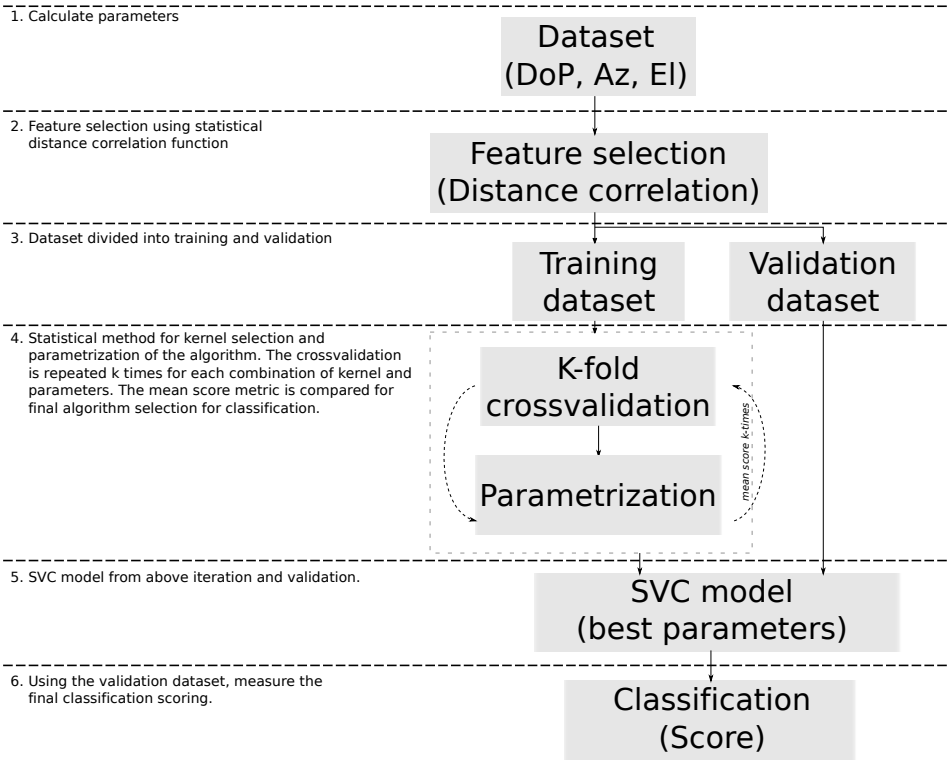


Figure 7. Flow diagram showing the strategy for this article

recursively evaluated and the final scores are compered resulting in the best set of parametrization and statistical validation for choosing the SVM algorithm for an specific task. The summary of the strategy used to validate our results is shown in figure 7.

3. Results and discussions

3.1. Feature selection and classifier parametrization

Pixel-wise polarimetric values of the degree of polarization (dop), azimuth (az) and ellipticity (el) were calculated from the measured Stokes parameters. We use the scattering matrix as visualization tool for the relation between these parameters.

We evaluated the distance correlation between these three independent parameters in order to asses the best of these parameters for the classification problem. As feature selection step, we measured the correlation between these polarization parameters. Metric for the distance correlation function (*dCorr*[5]) is based on the distance covariance (*dCov*[4]) and it was used as a feature selection tool. Figure 8 shows the pairwise relation of these three parameters, DOP and angular variables θ and χ, in form of scatter matrix. The diagonal of the matrix represents the parameters distribution, the top corner displays the correlation between features in form of scatter plot, and the bottom corner of the figure shows the values of distance covariance (DC) metric for each pair of features.

From the scatter matrix the combination of Dop and the angular parameters could be tested for classification, while the angles with themselves shows a greater overlapping. Just by using the scores from the DC metric within the features, it showed that the degree of polarization (*Dop*) and the ellipticity (*el*) scored the highest in distance correlation while minimizing the overlapping. These suggested to be good candidates for the classification of PE and non-PE coated paperboard, which were further selected for training of the classifier.

After selecting the polarimetric features for classification, we proceed to evaluate the classification performance from two SVM kernel functions, i.e. Gaussian kernel, also known as radial basis function (RBF), and a polynomial kernel with polynomial degree of order two. Both kernels are evaluated with

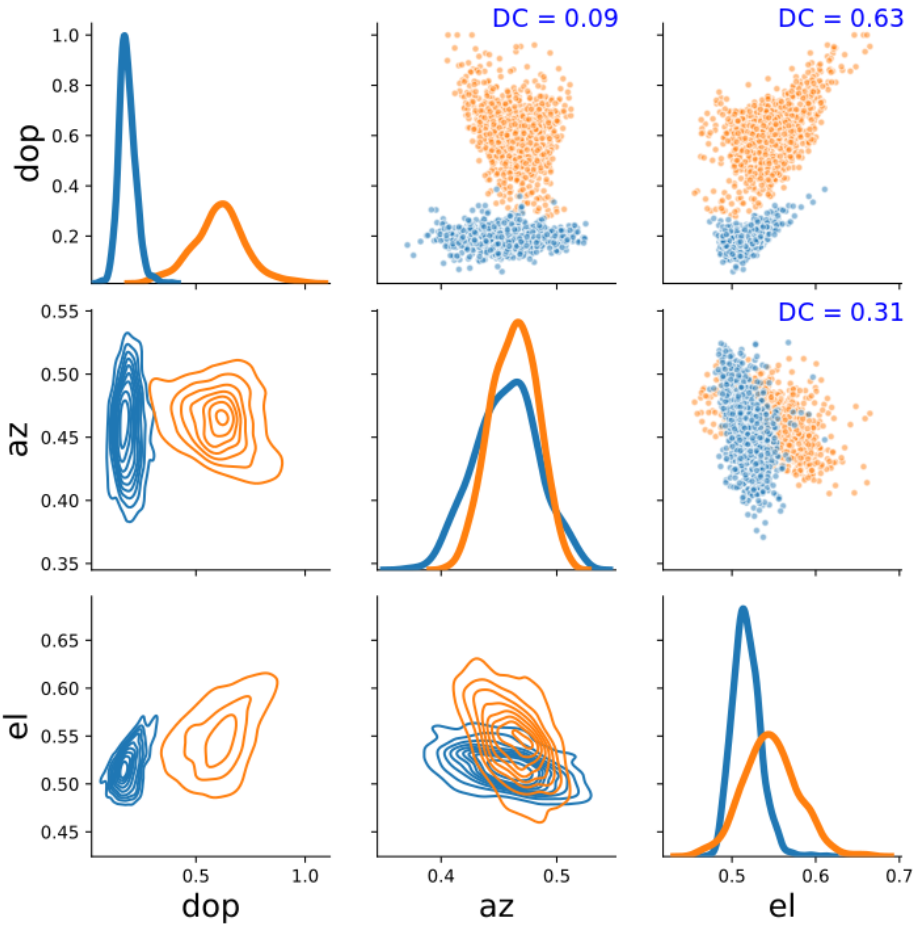
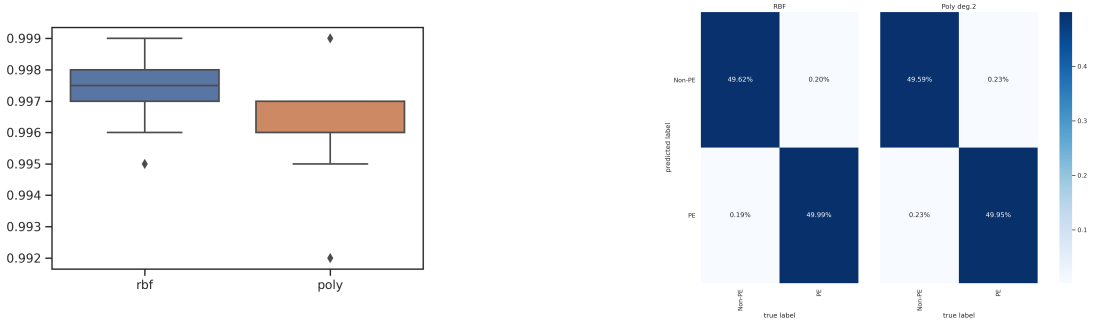


Figure 8. [CORRECT THHE TABLE, VALUES ARE INTERCHANGED]scatter matrix displaying the relation between features dop, az θ, and el χ. The DC value represents the distance correlation between the features. Only 20% of the dataset is been shown to aid the visualization.



(a) (b)

Figure 9. (a) 10 fold cross-validation scores for RBF and Polynomial of 2nd degree. (b) confusion matrix after validation using validation dataset on the trained support vector classifiers, i.e. RBF and 2nd degree polynomial kernels.

optimized hyperparameters under a 10-fold cross-validation strategy. The optimized hyperparameters were obtained using also 10-fold cross-validation, but instead of comparing the results of classification within two different kernels, the results are compared for the same kernel but within a set of parameters and selecting the ones that best scored without losing on performance. Figure 9 shows a box plot for the 10-fold cross-validation for the two different kernels. After this results, the Gaussian kernel was selected for the experiment with hyperparameters C=1000 and gamma=0.1, figure 9. There was no substantial difference between both kernel results, but assuming Gaussian distribution of the features and the widespread use of Gaussian kernel by the ML community due to the good results obtained, we decided the later over the polynomial one.

Figure 10 exemplifies the powerful idea behind using the kernel function to map the original dataset in dual space and trying to use parametric functions to separate the classes. In the original space, the dataset is mapped back and the kernel now separates the two classes by the location of the support vectors.

We can observe that there is an overlapping area from the support vectors that will lead to missclassification.

3.2. Classification of plastic coated and uncoated

In the previous section a cross-validation test provided an independent test for accuracy. We now test the classification model using a new dataset from two different samples. The algorithm can separate the PE coated paperboard from the one without PE coating. Figure 11 shows the resulting classification, with few areas in each sample opposite to the expected class. To further extend the discussion of whether or not this points can be regarded as missclassification, further experiments are required. In the case of PE coated areas, the scattering mechanism within the coating and reflection from the substrate can affect the accuracy of the measurement. There is also possibility that subsurface scattering from irregularities or defects within the extruded material generates the undesired effect. In the case of the base material, missclassification can be the result of geometrical constrains, where the facets of the surface texture can generate multiple reflections before reaching the detector, resulting in similar polarization properties as the PE coated sample. In our understanding these analysis can be done in a later study, where more understanding in the optical properties of each material can be extended for the use of polarimetric techniques. However, for the purpose of large area manufacturing

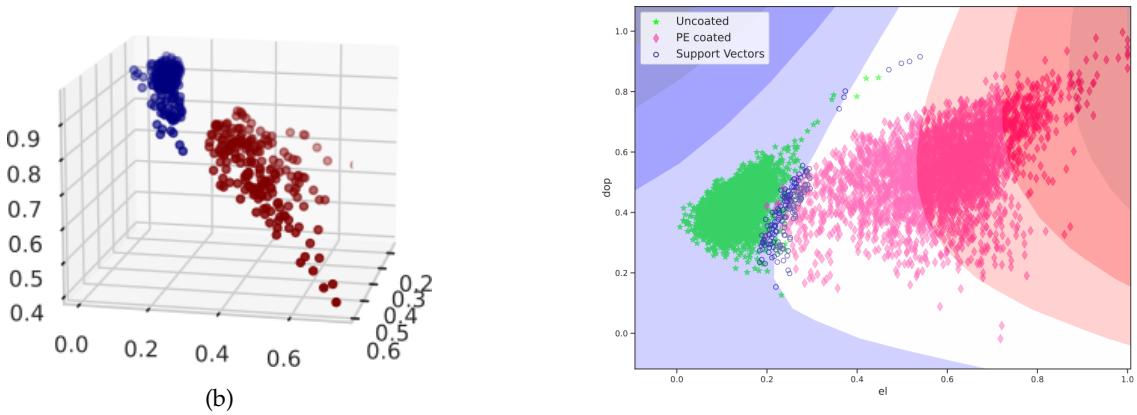


Figure 10. Feature spaces for *(a)* in higher dimensional space, i.e. dual space, and *(b)* original dataset space with applied SVM kernel for maximum class separation (in this case a polynomial kernel degree 2)

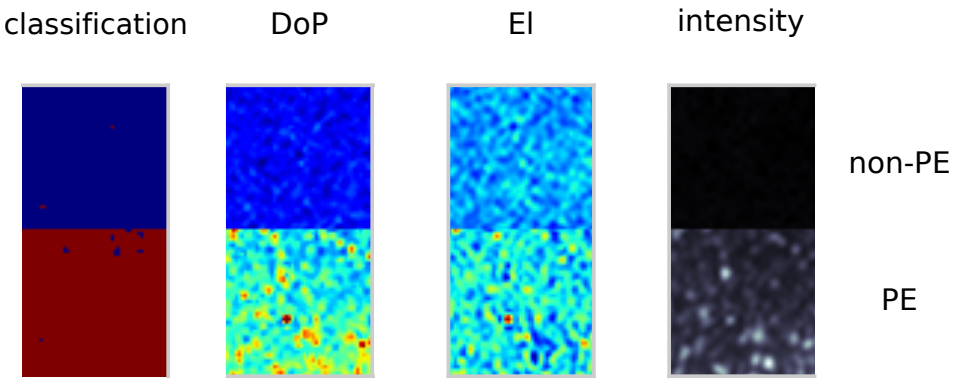


Figure 11. (left) Classification of PE coated (*red color*) and non-PE coated paperboard (*blue color*). (right) Intensity (S0) image representation of the scene, with a left bight PE coated and a less reflective non-Pe coated paperboard.

process, where more than 99% of the samples contains the expected class, this areas can be tracked and investigated with the current off-line production strategy.

4. Discussion

The camera system probes being useful for the detection and classification of PE coated paperboard. The system is used offline while having the potential to being use as online metrology tool. The system can acquire in one shot the images and calculate the parameters. However, it requires several second to calculate the Stokes parameters by using a calibration matrix to ensure correct demosaicing.

The use of faster computational tools can aid to use the sensor in online process control. New techniques like flat optics are also being developed for commercial uses and soon new in plane tools that required less hardware can help with the measurement of complete stokes parameters.

We can show high accuracy because the parameter correlation but there are still some misclassification issues. This can be understood from the geometrical relation of the surface texture with respect to light polarization but also other polarization effects can be considered like subsurface scattering and possible material defects in the samples, i.e. unevenness of the base layer, thickness reduction and other source of defects not visible from the top side of the structure.

5. Conclusions

In this work, we used a novel full-Stokes imaging polarimetry to measure the polarization information from industrially manufactured polyethylene (PE) coated paperboard samples, and we implemented a classification algorithm to separate samples that were coated from those without coating. With high spatial resolution and pixel-wise complete polarization information, we derived a set polarimetric features, i.e. degree of polarization (dop), ellipticity χ and azimuth φ, which later were used for the classification problem. We proposed a robust feature selection method based on distance correlation to reduce the computational cost of the algorithm while not compromising the classification accuracy. From the remaining features, we implemented a support vector machine classifier (SVC) that uses a non-linear kernel trick, i.e. Gaussian kernel function, obtaining a 99% classification accuracy. It is demonstrated that full-Stokes imaging polarimetry with appropriate illumination can be envisioned for in-process metrology, where complete polarization information provides a set of robust features directly related with the material optical properties that otherwise hidden for other machine vision systems. This system has the potential to eliminate some of the actual off-line quality control practices for coating related issues, e.g. pinhole appearances and the sealable metrics, by adopting the polarimetric imaging systems for the in-process control.

In our future work, we will study the use spectral information in combination of the polarization information, which is available from the presented polarimetric imaging system. This can increase the possibilities in finding new correlations for surface parameters or in case of quality control, defects originated along the paperboard production line.

Author Contributions: For research articles with several authors, a short paragraph specifying their individual contributions must be provided. The following statements should be used "conceptualization, X.X. and Y.Y.; methodology, X.X.; software, X.X.; validation, X.X., Y.Y. and Z.Z.; formal analysis, X.X.; investigation, X.X.; resources, X.X.; data curation, X.X.; writing–original draft preparation, X.X.; writing–review and editing, X.X.; visualization, X.X.; supervision, X.X.; project administration, X.X.; funding acquisition, Y.Y.", please turn to the CRediT taxonomy for the term explanation. Authorship must be limited to those who have contributed substantially to the work reported.

Funding: Please add: "This research received no external funding" or "This research was funded by NAME OF FUNDER grant number XXX." and and "The APC was funded by XXX". Check carefully that the details given are accurate and use the standard spelling of funding agency names at https://search.crossref.org/funding, any errors may affect your future funding.

Acknowledgments: To Prof. Liang and Ph.D Xiabo Tian for the support on the experimental data collected at University of Arizona.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

MDPI	Multidisciplinary Digital Publishing Institute
DOAJ	Directory of open access journals
TLA	Three letter acronym
LD	linear dichroism

References

1. Author1, T. The title of the cited article. *Journal Abbreviation* **2008**, *10*, 142–149.
2. Author2, L. The title of the cited contribution. In *The Book Title*; Editor1, F., Editor2, A., Eds.; Publishing House: City, Country, 2007; pp. 32–58.

Sample Availability: datasets in form of poincare sphere can be found in https://art1poincare.herokuapp.com/