Received: 9 March 2020 | Revised: 1 May 2020 | Accepted: 8 May 2020

DOI: 10.1002/jbio.202000083

FULL ARTICLE

JOURNAL OF BIOPHOTONICS

Depolarization metric spaces for biological tissues classification

Albert Van Eeckhout[1*] | Enric Garcia-Caurel[2] | Razvigor Ossikovski[2] | Angel Lizana[1*] | Carla Rodríguez[1] | Emilio González-Arnay[3,4] | Juan Campos[1]

[1]Grup d'Òptica, Physics Department, Universitat Autònoma de Barcelona, Bellaterra, Spain

[2]LPICM, CNRS, École Polytechnique, Université Paris-Saclay, Palaiseau, France

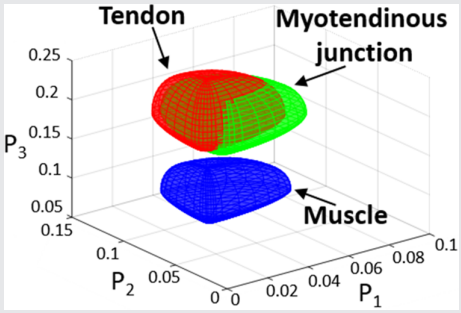
[3]Departamento de Anatomía, Histología y Neurociencia, Universidad Autónoma de Madrid, Madrid, Spain

[4]Servicio de Anatomía Patológica, Hospital Universitario de Canarias, Santa Cruz de Tenerife, Spain

***Correspondence**
Albert Van Eeckhout and Angel Lizana, Grup d'Òptica, Physics Department, Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain.
Email: albert.vaneeckhout@uab.cat (A. V. E) and
Email: angel.lizana@uab.cat (A. L.)

Funding information
Agència de Gestió d'Ajuts Universitaris i de Recerca, Grant/Award Number: 2017-SGR-001500; Ministerio de Economía y Competitividad, Grant/Award Numbers: Fondos FEDER, RTI2018-097107-B-C31

Abstract



Classification of tissues is an important problem in biomedicine. An efficient tissue classification protocol allows, for instance, the guided-recognition of structures through treated images or discriminating between healthy and unhealthy regions (e.g., early detection of cancer). In this framework, we study the potential of some polarimetric metrics, the so-called depolarization spaces, for the classification of biological tissues. The analysis is performed using 120 biological ex vivo samples of three different tissues types. Based on these data collection, we provide for the first time a comparison between these depolarization spaces, as well as with most commonly used depolarization metrics, in terms of biological samples discrimination. The results illustrate the way to determine the set of depolarization metrics which optimizes tissue classification efficiencies. In that sense, the results show the interest of the method which is general, and which can be applied to study multiple types of biological samples, including of course human tissues. The latter can be useful for instance, to improve and to boost applications related to optical biopsy.

KEYWORDS

biological tissue, biomedical, depolarization, imaging, Mueller matrix, polarimetry

1 | INTRODUCTION

Identification and classification of tissues is a recurrent topic in the biomedical domain.[1–8] Successful tissue classification allows, for example, the discrimination between healthy and malignant cancers and the early detection of cancerous regions.[5–8] Under this scenario, optical-based techniques are powerful tools to classify tissues as they are noninvasive techniques and they can achieve significant classification effectiveness.[6–9]

Polarimetry appears to be a promising optical technique for biomedical applications because it can

be combined with other optical techniques, as regular imaging or multispectral imaging, providing a new complementary way of characterization.[10–16] In past studies, it has been shown for instance that polarimetry allows for the grading of skin diseases,[14–17] the determination of the organization and density of fibers in tissues,[16, 18, 19] the discrimination between healthy and malignant cancer tissues,[20–25] and the description of scattering mechanisms in biological samples,[1, 26] among others.

Polarimetric analysis is usually based on two groups of techniques: polarization gating (PG) techniques[27–29] and Mueller (**M**) matrix-based methods.[20–26, 30–33] It has been proved that PG analysis can be derived from the **M** matrix.[34] Moreover, the **M** matrix contains more quantitative information than polarization gating techniques, and thus, **M**-based methods are better suited to classify samples. In this way, the Mueller-Stokes formalism is chosen to conduct the present work.

The **M** matrix is 4-by-4 real matrix which contains the intensity, diattenuation, retardance, and depolarization information of the sample.[30–33]

$$\mathbf{M} = m_{00}\begin{pmatrix} 1 & \mathbf{D}^T \\ \mathbf{P} & \mathbf{m} \end{pmatrix}. \quad (1)$$

Some of the information can be directly derived, such as the sample irradiance, provided by the m_{00} coefficient, the diattenuation and the polarizance characteristics of the sample, the latter provided by the three-dimensional (3D) vectors **D** and **P**, respectively, in Equation (1). Conversely, the retardance and the depolarization information are mixed and encoded in the 3-by-3 **m** submatrix, and to be obtained, they require further mathematical treatment.

The depolarizing content reveals structural information of biological tissues because it is related to scattering process produced by sample constituents,[1, 2] and therefore, it stands as an ideal candidate for the classification of different tissues. The analysis of depolarizing content in biological samples is usually restricted to the depolarization index P_Δ metric.[9, 33, 35] However, other parameters called indices of polarimetric purity (IPP, composed by P_1, P_2 and P_3 parameters),[36, 37] which are calculated from the eigenvalues (λ_i) of the covariance matrix (**H**),[38, 39] present further physical interpretation, synthetization of the depolarization content, and differences between tissues, and in some cases, provide visualization of structures that are hidden by using the depolarization index.[40–42] In a recent review paper, Ossikovski and Vizet presented a compendium of existing depolarization parameters which define different depolarization metric spaces thus, giving different ways to describe the loss of polarization of light.[43]

The increasingly amount of depolarization metrics raises questions about the equivalence of different depolarization metric spaces, their mutual redundancy, and also, the possibility to select a particular depolarization parameter space, better suited than others, to classify a given type of biological tissues. This work addresses these questions for the first time, by comparing the efficiencies of different depolarization metric spaces (and other commonly used depolarization metrics) to classify different organic tissues. To this aim, we measure the Mueller matrix of a collection of ex vivo samples, from which different depolarization metrics, including those related to the depolarization spaces, are calculated. Among those metrics, we analyze their classification potential and viability to discriminate between different tissues by representing the experimental data into the so-called depolarization metric spaces. We also discuss the capability classification models based on the different depolarization metric spaces to discriminate between different tissues. According to the classification results, the best suited depolarization model to treat the data used in the discussed example is proposed.

2 | MATERIALS AND METHODS

2.1 | Depolarization metric spaces

In this work, five depolarization metric spaces (the detailed theoretical background is provided in Data S1) and the depolarization index P_Δ have been used to classify different tissues. Each space is constituted by three parameters. The spaces can be divided in two groups: the ones based on **H** matrix eigenvalues (the Natural space, the IPP space and the high-order depolarization indices space) and the ones based on the type I canonical depolarizer parameters (type I canonical space and the type I Lorentz space). These polarimetric spaces, recently provided in the literature,[43] were selected in this work to be analyzed for the first time in the framework of biophotonics, as they have the potential of provide larger insight on polarimetric properties of biological tissues than other more commonly used metrics.[43] In this context, the well-known and standardly used depolarization index P_Δ is selected as a reference, to be compared with results obtained by using the different polarimetric spaces highlighted in this work.

2.1.1 | Depolarization metric spaces based on H matrix eigenvalues

Natural space is composed by the λ_2, λ_3 and λ_4 eigenvalues of the **H** matrix[33]:

$$\mathbf{H} = \frac{1}{4}\sum_{k,l=4}^{4} m_{kl}(\sigma_k \otimes \sigma_l), \tag{2}$$

where m_{kj} are the elements of the **M** matrix, σ_k are the four Pauli spin matrices and the symbol $\otimes$ corresponds to the Kronecker product.

The IPP (P_1, P_2, P_3) are defined as a linear combination of **H** eigenvalues[37]:

$$P_n = \sum_{k=1}^{n} k\Delta\lambda_k, \text{where } \Delta\lambda_k = \lambda_k - \Delta\lambda_{k+1}, \text{and provided that } \sum_{i=1}^{4}\lambda_i = 1, \tag{3}$$

while the high-order depolarization indices (P_Δ, $P_\Delta^{(3)}$, $P_\Delta^{(4)}$) are a nonlinear combination of this eigenvalues,[43]

$$P_\Delta^{(m)} = \sqrt{\frac{1}{4^{m-1}-1}\left(4^{m-1}\sum_{k=1}^{4}\lambda_k^m - 1\right)}. \tag{4}$$

Note that the depolarization index is, in fact, the second-order depolarization index ($P_\Delta = P_\Delta^{(2)}$).

2.1.2 | Depolarization metric spaces based on the type I canonical depolarizer parameters

There is a second group of depolarization metric spaces, not based in the eigenvalues of the **H** matrix but in model matrices called canonical depolarizers. The type I canonical space is composed by three canonical parameters d_1, d_2 and d_3 obtained from the type I depolarization matrix:

$$\mathbf{M}_{\Delta d} = \mathrm{diag}(1\ d_1\ d_2\ d_3), \tag{5}$$

which is derived from the symmetric decomposition.[44]

The type I Lorentz space, described by L_I, $L_\mathrm{I}^{(3)}$ and $L_\mathrm{I}^{(4)}$, is a generalization of the Lorentz parameter.[45] The high-order Lorentz parameters are defined by a nonlinear combination of the type I canonical depolarizer parameters

$$L_\mathrm{I}^{(m)} = \sqrt{\frac{1}{4^{m-1}-1}\left[4^{m-1}\frac{1+d_1^{2m}+d_2^{2m}+d_3^{2m}}{\left(1+d_1^2+d_2^2+d_3^2\right)^m} - 1\right]}. \tag{6}$$

Note that the second order is, in fact, the Lorentz parameter.

2.2 | Experimental methodology

The experiment starts by measuring the Mueller matrices of 120 samples (40 tendons, 40 muscles and 40 myotendinous junction regions) at three different wavelengths, such as 625, 530 and 470 nm. All the samples were obtained from 20 different chicken thighs, and they were measured by using a complete Mueller matrix image polarimeter based on parallel-aligned liquid crystals retarders[41, 42, 46] (the image polarimeter is described in Data S1). In order to ensure similar tissue decomposition conditions, all tissues were submitted to the same measurement procedure (described in Data S1). Moreover, for each wavelength, the experimental matrices are analyzed and a region of interest (ROI) of 150×150 pixels, for each sample is selected to perform the classification (the ROI selection method is described in Section 2.4 of Data S1). Note that each pixel of the selected ROI image corresponds to a particular **M** matrix resulting into 2.7 million of **M** matrices. This important amount of depolarization information of a sample describes the polarimetric behavior of tissues and the results are studied by calculating the different depolarization metrics above mentioned. As above stated, we carried out the analysis by using data obtained from a collection of chicken thighs, the latter illustrates the interest of the above-mentioned polarimetric metrics to be used in the framework of tissues classification based on optical data. We selected to apply our approach on chicken thighs tissues for the ease of handling. The thighs were acquired in a local market of fresh meat and vegetables, and they were initially sold for human consumption purposes. Despite of standard conservation procedures in a fridge for a limited period, there is no particular restriction or hazard concerning their manipulation. However, note that the method could be tested on any biological sample, including human tissues which can be found in hospitals or specialized research institutions.

2.2.1 | Analyzed tissue description

The polarimetric response of the analyzed tissue strongly depends on its structure at different scales (milimeter, micrometer and nanometer). Tendon is a noncontractile mesodermal tissue that connects muscle to bone and it is prepared to resist tensions. It is mainly composed by type I collagen clustered into fascicles showing the same orientation as muscle bundles.[47-50] Muscle is a soft tissue composed by contractile myofibrils organized in bundles surrounded by sheets of connective showing an arrangement that is very similar to the one present in tendons. In the experiment, we measured skeletal muscle, which

has a structure closely related to the one present in tendon, but containing different subtypes of muscle fibers instead of type I collagen fibers.[51] The myotendinous junction shows a combination of the other two studied tissues. It is composed by fibers clustered into fascicles, as the previously described tissues, but in this case, fascicles of contractile (muscle) and collagen (tendon) are intermingled.[52, 53]

Further detailed description of tissues is provided in Data S1.

2.2.2 | Supervised classifiers

After the experimental measurements, all the depolarization information is used to build a supervised classifier using one of the three following methods: tree classifier,[54] linear discriminant classifier,[55] and k-nearest neighbors (kNN) classifier.[56] Tree classifier is a decision model based on a tree-like scheme[54]; linear discriminant classifier is a probabilistic method which finds a linear combination of different parameters capable of separating the data in different classes[55]; and the idea behind the kNN classifier is the selection of the k training data points, in the N space dimension, which are the nearest (in a smaller distance) to a given test data.[56] The three supervised classifiers methods are implemented using the "Statistics and Machine Learning Toolboox" in the MATLAB language and they are widely described in Data S1.

The whole of the results cannot be applied to build (train) the classifier, but the efficiency test of depolarization metric spaces requires the use of a fraction the original data for this task. In our study, the supervised classifier is built using 1% of all the measured data and the remaining 99% were used to use the classifier.

3 | RESULTS AND DISCUSSION

In this section, we qualitatively and quantitatively analyze the classification potential of the different depolarization metrics (described in Data S1) that are calculated from the experimental ROIs of the different measured tissues. The qualitative study is discussed in Section 3.1 by analyzing the depolarization distributions of measured tissues. The quantitative analysis is based on the use of supervised classification models (described in Data S1) to classify the experimental measurements into different tissues. We show how the classification efficiency depends on the application of different depolarization metric spaces. The obtained results are presented and discussed in Section 3.2.

3.1 | Depolarization metric spaces to classify experimental biological tissues: Qualitative analysis

The qualitative analysis starts by calculating the different depolarization metrics, described in Section 2.1, for every pixel in the corresponding polarimetric images of measured tissues. Taking the ensemble of pixels as a whole leads to a large amount of data to be treated and interpreted, that is, for each parameter we have 120 images of (150×150 pixels each) which gives 2.7 million of realizations. To better apprehend such amount of information, we represented them as a collection of points in the different 3D spaces. The kind of visual information that can be obtained in this way is illustrated in Figure 1. The figure represents the polarimetric information expressed in terms of the five parametric spaces that we compare in this study (Figure 1A-E), each space contains 2.7 million of points. Note that in the image different colors represent different tissues (red for tendon, blue for muscle and green for myotendinous junction tissue). The purple lines in each space show the limits of the corresponding physically realizable zone. Since we decided to use the same scale to represent the five spaces, it is possible to see that some of them occupy much more volume than others. The consequences of the latter will be discussed in detail further on. In order to no overextend the length of this manuscript, only results obtained for a wavelength of 625 nm are plotted and discussed as a representative case (530 and 470 nm results are provided in Data S1). The only difference between the results obtained at different wavelengths, in terms of depolarization response, is that muscle and myotendinous junction response is slightly less depolarizing for 530 and for 470 nm measurements than for 625 nm case.

By observing the representation in Figures 1A-E, we can differentiate two types of spaces depending on volume and shape. The linearly shaped spaces (natural, Figure 1A; IPP, Figure 1B) and type I canonical Figure 1D) occupy more volume than the nonlinear spaces (high-order depolarization index, Figure 1C) and type I Lorentz (Figure 1E). However, most of the available space of the nonlinear 3D spaces (Figure 1C,E) is filled by the experimental data, whereas some physically realizable regions in the linear 3D spaces are empty (Figures 1A,B,D). Although is clear that data points tend to spread in the allowed space, they are not randomly mixed, that is, points belonging to each particular class tend to group together.

Although Figure 1 gives relevant information of the sample response, it can lead to errors in the analysis due to the data dispersion and because certain dots can be superposed. To better visualize and describe the data

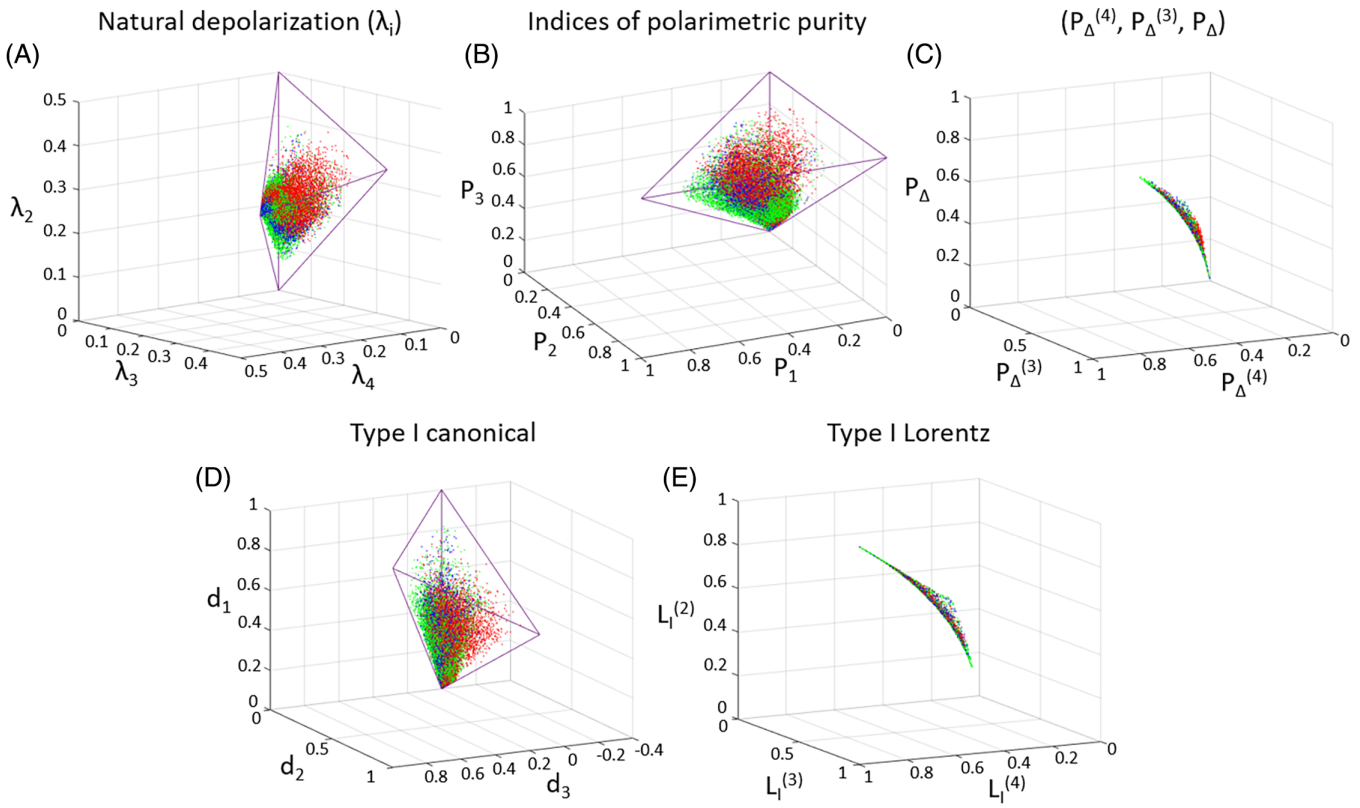


FIGURE 1 Three-dimensional representation of the experimental depolarization data collected at 625 nm for tendons (red), muscles (blue) and myotendinous junction tissues (green). The depolarization data are represented in the depolarization metric spaces: A, natural depolarization space; B, indices of polarimetric purity space; C, higher-order depolarization index space; D, type I canonical depolarization space; E, type I Lorentz depolarization space

distributions, we have calculated nonsymmetric ellipsoids representing data of different tissues at different spaces. These nonsymmetric ellipsoids are defined by a central value and six different semiaxis values, depending on their direction with respect to the center of the ellipsoid. The center of the ellipsoid corresponds to the mean value of the specific tissue data collection, and the semiaxes are calculated as the SD associated with the positive and negative values with respect to each of the three specific parameter's mean (for instance, SDs of λ_2, λ_3 and λ_4 for the natural depolarization space, or SDs of P_1, P_2 and P_3 for IPP space) (mean and SDs data can be consulted in Table S1 in Data S1). We call the ellipsoids nonsymmetric because for a given ellipsoid and a given axis, the length of the corresponding positive and negative semiaxes, do not need to be identical. Under this description, approximately 68% of tissue data are located inside of the volume of the corresponding nonsymmetric ellipsoid. Figure 2 shows the described nonsymmetric ellipsoids obtained from the raw points of the same spaces shown in Figure 1.

Note that for the sake of visualization, the axis range used in Figure 2 does not represent the full space range, but it represents the ≈10% of the physically realizable one. In other words, images in Figure 2 do not show the full space but a reduced one, where data are concentered, this particular zone being associated with a highly depolarizing behavior. Spreading of data shown in Figure 1 may be explained by different polarimetric interaction of light with tissues and nonuniformities of samples themselves. In particular, such spreading may be attributed to sample not-flatness and roughness, and to a less extent, to the different handling of chicken tissue (possible variations in defrosting times, possible pressures during dissection process, etc.) as well as biological differences between dissected chickens. However, points in Figure 1, which are located farthest from ellipsoids in Figure 2, are mostly a consequence of the measurement's noise of a discrete number of pixels, which may be produced by sample irregularities and/or direct reflections.

Figure 2 shows how the depolarization response of measured tissues is distributed over the depolarization metric spaces. Separated closed volumes without overlapping are associated to completely distinguishable tissues. Therefore, the ideal scenario in our study would correspond to three completely separated volumes (chicken muscle, myotendinous junction and tendon), that

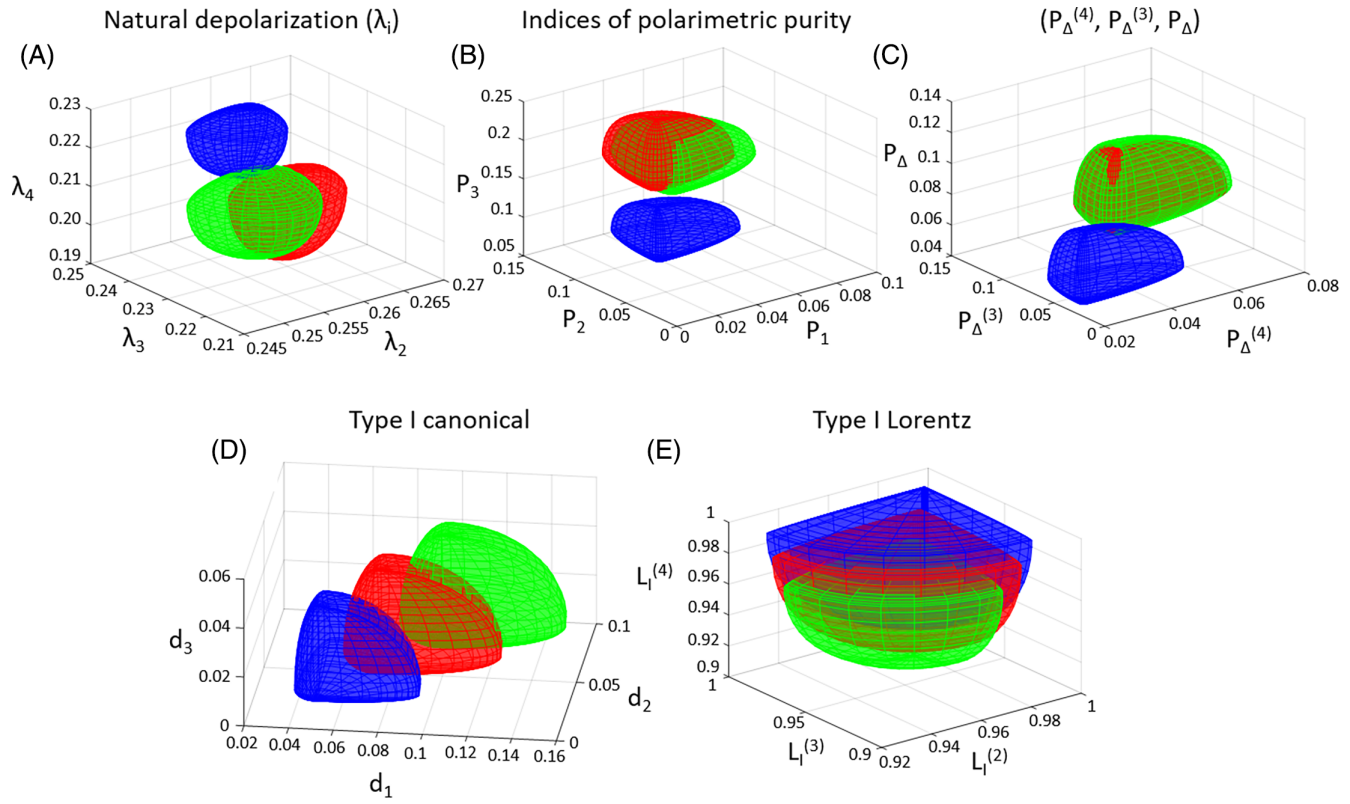


FIGURE 2 Three-dimensional representation of the experimental nonsymmetric ellipsoids (described by the mean value and the SD values) obtained for data measured at 625 nm. Data corresponding to different tissues are represented by different colors: tendon (red), muscle (blue) and myotendinous junction tissue (green). The data representation is conducted in the depolarization space: A, natural depolarization space; B, indices of polarimetric purity space; C, higher-order depolarization index space; D, type I canonical depolarization space; E, type I Lorentz depolarization space

would lead to sensitivity and specificity values for tissues discrimination of 100%. The experimental result (Figure 2) indicates that muscle gives a very distinguished response, in the eigenvalue-based spaces (Figure 2A-C), with respect to the tendon and myotendinous junction tissue. According to this result, the muscle is candidate to be well classified by using such spaces. In contrast to this, a significant uncertainty is likely to happen in the classification process between the tendon and the myotendinous junction tissue, since their response overlaps in all the spaces. From the latter, we also expect that the uncertainty in classification will be maximal in the high-order space, where the response of tendon and myotendinous junction tissue fully overlaps (Figure 2C). On the contrary, it can be expected that discrimination between tendon and myotendinous junction tissue will be optimal in the type I canonical depolarization space (Figure 2D) because the overlap between the corresponding nonsymmetric ellipsoids is minimal. At this point we can anticipate that the capacity to discriminate among different types of tissue is in connection to the tissue structures and the relative chemical composition of them. Although the structure of the analyzed tissues is very similar (all are based on arranged fibrils), tendons and myotendinous junction are both composed by collagen fascicles. Collagen fibrils provides a differentiated and characteristic polarimetric response,[15, 17] and the absence of these fibrils in muscles may produce the differentiated response observed in eigenvalue-based spaces.

Last but not least, by analyzing information in Figure 2, we can clearly see that when illumination at 625 nm is used, the muscle is the structure with larger depolarization. In particular, in Figures 2B-D, the blue ellipsoid is the one closer to the (0, 0, 0) coordinate (pure depolarizer), as well as in Figure 2E it is the ellipsoid closer to the (1, 1, 1) coordinate (note that in the type I Lorentz space, pure depolarizers are located at the position (1, 1, 1)[43]). By following the same reasoning, we note that the second tissue with larger depolarizing capability is the tendon (red ellipsoid), being the myotendinous junction tissue (green ellipsoid) the one presenting the lowest depolarizing capability. Note that although myotendinous junction response is expected to be intermediate, because its structure is a mixture of tendon and muscle fibers, the depolarization response associated to these structure lead to lower depolarization values than those exhibit by tendons and muscles.

3.2 | 3D depolarization metric spaces to classify experimental biological tissues: Quantitative analysis

Once the qualitative analysis of the results is done, a quantitative analysis is performed to identify the best suited depolarization parameters to properly classify different depolarizing tissues. The quantitative analysis is carried out by applying three classification models, tree,[54] discriminant[55] and kNN[56] (described in Data S1), to the multiwavelength results. The percentages of well-classified tissues as a function of the different depolarization metric spaces are presented in Table 1. The percentages are the mean probability to achieve a proper classification of the three tissues and they are calculated using the different models with data taken at the wavelengths 625, 530, and 470 nm, respectively. The probability results have an error of ±1% owing to the randomness of the test data selection.

The discrimination efficiency results presented in Table 1 are ordered in such a way that eigenvalue-based spaces results are presented in columns 3 to 5, and, canonical parameters-based spaces are presented in columns 7 to 8. The results obtained by using the index of depolarization are also provided (column 6) for comparison because it is one parameter widely extended in the literature.[9, 33, 35]

The efficiency results in Table 1 clearly indicate how the tree classification model is the less efficient in discriminating data whereas the kNN is the one which discriminates data the best. Moreover, for a particular wavelength, if we order the discrimination efficiencies for different depolarization metric spaces (different columns) from highest to lowest, we obtain the same result independently on the classification model selected. Therefore, kNN is selected to discuss the quantitative analysis of our results, because it is the best and because the conclusions for this particular classifier are also true for the other two.

In the following, we first compare the results obtained for the eigenvalue-based metrics (columns 3-6). Secondly, we compare the results obtained by using type I canonical parameters-based spaces (columns 7 and 8). Finally, a global conclusion, by taking into account all the spaces analyzed, is also provided.

On the one hand, by considering the efficiencies obtained by using eigenvalue-based metrics (columns 3-6), we see how the depolarization index P_Δ used alone, provides the worst results in all the cases. Therefore, it is directly deduced from the above results that the use of depolarization metric spaces instead of the depolarization index is not only better for visualization, as provided by Van Eeckhout et al. [41, 42], but also for classification. In particular, the combination of higher-order indices $P_\Delta^{(3)}$ and $P_\Delta^{(4)}$ (column 5) to the depolarization index (column 6) gives an increment of 2% to 3% (depending the wavelength) of classification success. If we compare the **H** matrix eigenvalue (column 3) or the IPP spaces (column 4) with respect to P_Δ (column 6), the efficiency increment is even more significant, with an improvement of 5% (for red channel), 8% (for blue channel) and 9% (for green channel). We would like to emphasize here that, this is a particular experiment involving three particular tissues, and the classification efficiency, as well the best depolarization metric space, will depend on the tissues under consideration. However, the obtained results clearly indicate that, regardless of the efficiency obtained for studied tissues, the use of depolarization metric

TABLE 1 Percentage of well-classified tissues as a function of the used wavelength, the classification model and the depolarization parameters

Wavelength	Classifier	Eigenvalue-based depolarization metrics				Canonical-based depolarization metrics	
		$\lambda_2, \lambda_3, \lambda_4$	P_1, P_2, P_3	$P_\Delta, P_\Delta^{(3)}, P_\Delta^{(4)}$	P_Δ	d_1, d_2, d_3	$L_I, L_I^{(3)}, L_I^{(4)}$
625 nm	Tree	58	58	55	53	52	50
	Discriminant	59	59	60	57	57	55
	kNN (50)	**65**	**65**	**62**	**60**	**60**	**55**
530 nm	Tree	51	51	47	46	55	50
	Discriminant	54	54	51	49	62	49
	kNN (50)	**59**	**59**	**53**	**50**	**64**	**50**
470 nm	Tree	52	52	48	46	52	48
	Discriminant	56	55	52	51	60	47
	kNN (50)	**60**	**59**	**54**	**52**	**62**	**49**

Bold values indicate the best classification results for each depolarization space and wavelength measurement.

spaces can improve the classification potential obtained by P_Δ.

Once we realize that eigenvalues-based depolarization spaces (**H** matrix eigenvalues, IPP and high-order depolarization index) provide better results in terms of tissue classification than P_Δ, we focus on determining which one of these three spaces is more suitable for classification. By checking the efficiencies in columns 3 to 5, we see that the worst results are obtained by the high-order space (column 5) whereas **H** matrix eigenvalues and IPP provide the best (and identical if taking into account 1% associated error) results.

We think, although we do not have to date a fundamental proof, the fact that the high-order depolarization space (column 5 in Table 1) provides not as good results as the **H** matrix eigenvalues or the IPP spaces arises from the nonlinear mathematical origin of such space. Such nonlinear combinations reduce the volume occupied by physically valid data in the parametric space, which automatically compresses the information (see Figure 1). The compression of information implies that the differences between tissue responses are reduced, thus reducing the classification efficiency. We would like to highlight that this conclusion arises from a heuristic approach, because we cannot infer the observed loss of information related to nonlinear spaces directly from the mathematical relations itself. However, we have clearly observed this tendency in the case of all the analyzed tissues and for all the wavelengths. As will be further discussed, this situation is also observed in the canonical case.

At this point, the **H** matrix eigenvalues space (column 3) and IPP space (column 4) appears to be the best options to represent depolarization information for tissue classification. Recent papers show that the IPP space can be used to provide a physical interpretation of the studied samples[36, 37, 40] together with an improvement sample visualization with respect to other parameters (**H** matrix eigenvalues or depolarization index).[41] The improvement in visualization quality and contrast is related to the increment of volume. However, as it can be seen in Table 1, the classification potential of both blocks of parameters is always the same. We think that this is because the IPP are a linear combination of eigenvalues, λ_i. Taking this in mind, as the used classification models improve the classification efficiency by conducting linear combinations of data, from the point of view of the classifiers, either IPP or λ_i are equivalent. Therefore, taking into account that **H** matrix eigenvalues space and IPP space provide identical classification results but the latter provides larger tissues visualization, then the use of the IPP space to get depolarization information for classification issues, is recommended.

On the other hand, analogously to the discussion provided above, we can perform a comparative study between results obtained using canonical-based spaces (columns 7 and 8 in Table 1). We clearly observe that the type I canonical parameters (d_i for $i = 1, 2, 3$) space provides significantly better results than the type I Lorentz parameters space. In analogy with our previous discussion, we also consider that the non-linear mathematical origin of the type I Lorentz space leads to a compression of information, and therefore to a loss in classification efficiency. Under this scenario, if canonical-based spaces have to be used, then the type I canonical parameters should be used instead of the Lorentz canonical parameters.

Finally, according to the conclusions to which we have arrived in the previous discussions, we compare, in terms of classification potential, the best space based on eigenvalues (IPP) with the best space based on canonical metrics (type I canonical parameters). By comparing these two spaces, it can be shown that the best efficiency results depend on the wavelength used. However, the type I canonical parameters space presents the best efficiency values for the 470- and 530-nm channels, the IPP space presents the best efficiency for the 625 nm case.

This efficiency dependence with the wavelength can be understood by taking into account the metrics associated to these two spaces. When the **M** matrix has no diattenuation and polarizance content, d_i parameters can be described as a linear combination of λ_i.[43] However, if the **M** matrix contains diattenuation or polarizance response, then, d_i parameters and λ_i are no more linearly related. In theory, the symmetric decomposition[44] needed to obtain the d_i parameters fully separates the diattenuation and retardance from the depolarization content, but in the case of **H** matrix eigenvalues, λ_i, this content is not completely separated. Therefore, the λ_i values are influenced by the diattenuation or polarizance.[33] Therefore, the diattenuation and polarizance of the sample can directly affect the values of λ_i and thus, their potential for classifying different tissues when using the IPP space (note that IPPs depend on λ_i; described in Data S1).

To further analyze this fact, we have calculated probability histograms for the diattenuation and polarizance present in our experimental samples (three tissues types) for the three wavelengths studied. The histograms show the proportion (probability density) of pixels with different $|\mathbf{D}|$ and $|\mathbf{P}|$ (see Equation (1)) values. Results are shown in different graphics in Figure 3 ordered in two rows and three columns. Each column of graphics in the figure corresponds to a given wavelength. Graphics in the first row of the figure correspond to diattenuation data, and graphics in the second row of the figure to

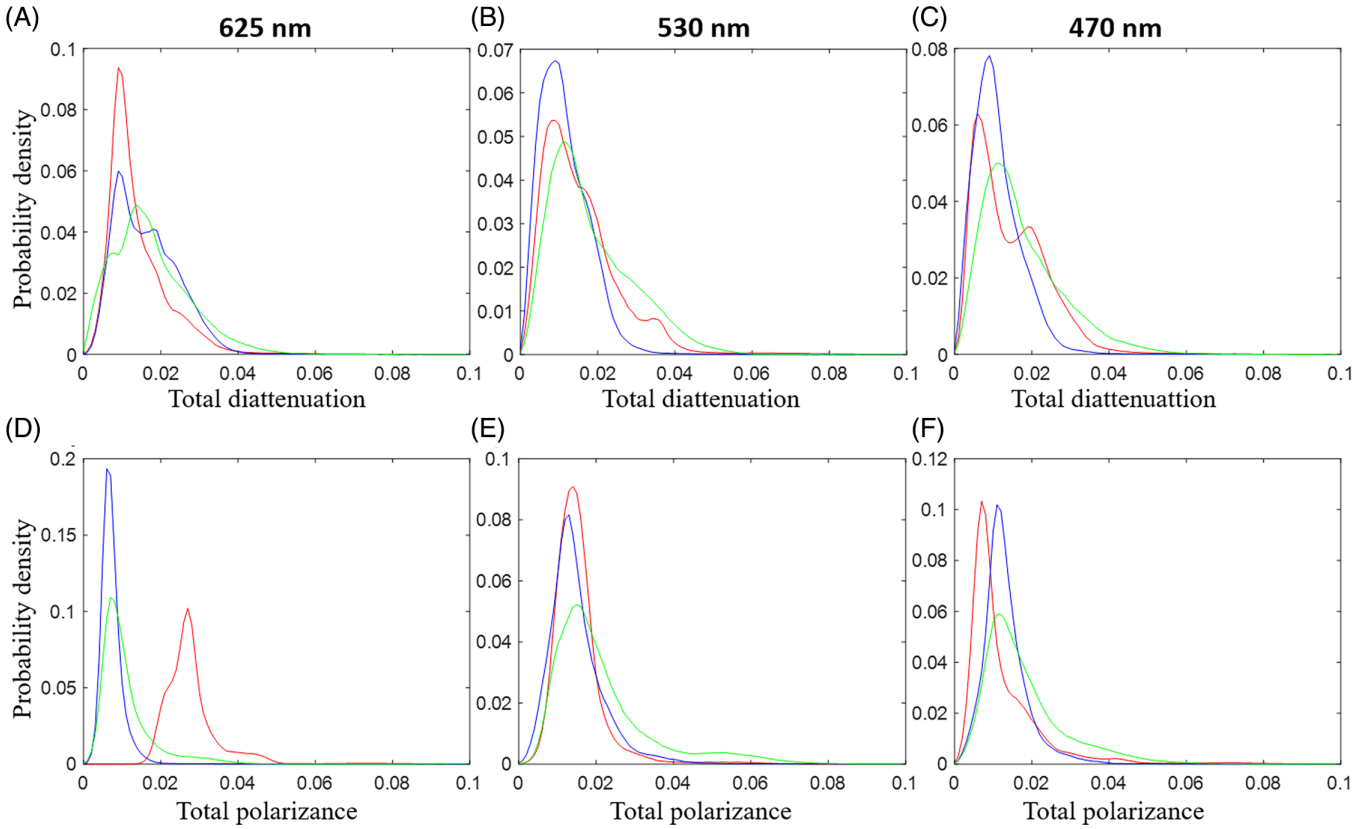


FIGURE 3 Histogram of the total diattenuation response of the different tissues using, A, 625-nm wavelength; B, 530-nm wavelength; C, 470-nm wavelength. Total polarizance response using the same, D, 625-nm wavelength; E, 530-nm wavelength; F, 470-nm wavelength. Tendon is represented in red, muscle in blue and the myotendinous junction tissue in green

polarizance data. To separate the contributions of different tissues, tendons are represented in red, muscles in blue and the myotendinous junction tissues in green.

The diattenuation histograms show that the diattenuation content of the different tissues is almost the same for all tissues and for all the wavelengths. In consequence, diattenuation cannot be used to distinguish among tissues. In this particular case, the diattenuation channel can be considered as a source of "information noise" which may hinder the performance of the classification algorithm. On the contrary, the polarizance response of tendons (red color) is significantly different from that of the other two tissues for the 625-nm channel. Moreover, the central polarizance of the tendon case decreases with the wavelength of light. In contrast, the polarizance of muscle and myotendinous junction tissue increases when wavelength shortens. This particular polarimetric behavior produces three different scenarios depending on the wavelength used to illuminate the samples.

First, distinguishable polarizance results are produced illuminating with 625 nm (Figure 3D). The mean polarizance of the tendon is 0.03, while the same characteristic of the other two tissues is approximately 0.01. Thus, polarizance gives relevant information to classify the tissue and this information is present in the λ_i. As this classification information is relevant, the polarizance improves the classification efficiency of the IPP space (with polarizance contribution) with respect that of the type I canonical parameters space (note that d_i have not polarizance contribution). Therefore, in Table 1, IPP space provides better efficiencies than type I canonical parameters space (65% and 60%, respectively) for the red channel.

Second, by using the 530-nm wavelength, the polarizance and diattenuation responses are almost the same for all the tissues studied, and thus, they do not provide relevant information. Since this irrelevant information appears in the **H** eigenvalues, λ_i, as a source of "information noise," but not in the d_i parameters, the classification efficiency of the IPP space is reduced with respect to of the type I canonical parameters space. Consequently, canonical space provides better classification efficiencies than IPP space in the 530 nm case (64% vs 59%, respectively, in Table 1).

Finally, when illuminating the tissues with 470 nm, an intermediate case between the two previous ones is observed. The polarizance response provides some new

valuable information in terms of differences between tissues, but it is not enough to counter the irrelevant information given by the diattenuation content. As a result, the IPP space classification efficiency appears to be slightly smaller than the efficiency of the type I canonical parameters (59% vs 62%, respectively, in Table 1).

Summarizing the discussion provided in Section 4.2, we can state that among the five spaces compared in this study, the IPP space and the type I canonical parameters space perform the best. In absence of diattenuation or polarizance, the IPP space and the type I canonical parameters space provide are comparable. If diattenuation and polarizance information are present and show a discrimination ability, then the IPP space can perform better than the type I canonical parameters space, However, if the samples present diattenuation or polarizance but this information does not discriminate between tissues, then, the type I canonical parameters space will lead to better classification efficiencies as it is not affected by such information. Note that the **D** and **P** content in samples may depend on wavelength, as proved in our analysis, so a multichannel analysis is recommended when possible as a good practice in the analysis of polarimetric images of tissues.

Therefore, if the classification efficiency is considered as the main metric to guide the use of the polarimetric data, then based on the results of this study for the particular sample of chicken tissues discussed here, we recommend the use of a model based on the type I canonical parameters space. This is because if the sample presents non-discriminating polarizance or diattenuation information, the type I canonical parameters space is not affected by them. On the contrary if diattenuation and polarizance provide non-negligible discrimination capacity, then the IPP-based space is to be preferred. Finally, since the evaluation IPP requires less computation time than the evaluation of the type I canonical parameters, we also recommend the use of the IPP-based space when diattenuation and polarizance are negligible.

4 | CONCLUSIONS

In this work, we discussed the efficiency of different groups of depolarizion metrics to classify tissues. Such groups of observables are composed of three depolarization parameters which are characteristic of different depolarization metric 3D spaces. They are divided into two main groups: derived from **H** matrix eigenvalues (eigenvalue-based spaces) and derived from the type I canonical depolarization parameters (canonical-based spaces).

To this aim, we measured the Mueller matrix of 120 chicken thighs. Three different tissues are studied (40 muscles, 40 tendons and 40 myotendinous junction tissues). To ensure pure tissues (ie, without contributions of other surrounding tissues type), a region of interest selecting method is applied, resulting in a collection of images of 150×150 pixels.

The analysis of the suitability of the depolarization metric spaces to classify samples is conducted from a qualitative (Section 3.1) and a quantitative (Section 3.2) point of view. The qualitative observation of the results is conducted through the visualization of the data at the studied 3D spaces. It shows that muscles are significantly differentiated from the other two tissues by using **H** matrix eigenvalues space, allowing a successful muscle classification. However, tendons and myotendinous junction tissues occupy similar regions in such spaces, being the best visualization obtained by using the type I canonical space. The quantitative analysis is performed by using three different supervised classifications models: tree classifier, linear discriminant classifier and kNN classifier. From those models, and for the particular case of the samples used in the present discussion, the higher classification efficiencies are obtained by using the kNN model, so this supervised model is selected.

Regarding to the depolarization metric spaces, we have conducted a first comparative between **H** matrix eigenvalue-based spaces, and they have also been compared with the standard depolarization index P_Δ. We have shown that depolarization spaces provide better classification efficiencies than the P_Δ, which is commonly used in polarimetric community, so the inclusion of depolarization metric spaces in polarimetric studies instead of P_Δ is highly recommended. From the eigenvalue-based spaces, the highest efficiencies are obtained for the IPP and the **H** matrix eigenvalues spaces, being the former recommended as it also provides physical interpretation of data. We have performed a second comparison between canonical spaces, the type I canonical space providing much better results than the type I Lorentz space, in terms of classification efficiency.

Finally, the two selected spaces, IPP space, and type I canonical space, are compared to determine the most suitable one to build a classification model for the studied tissues. We have observed that the best efficiencies to classify a tissue are reached by one or the other spaces depending on the wavelength illuminating the sample. This situation is explained by the relation between the values of the **H** matrix eigenvalues and the presence of diattenuation and polarizance in the sample. In the case of nonpolarizing and nondiattenuating samples, both spaces are equivalent in terms of discrimination efficiency because in such conditions the two spaces are

related by a linear combination. Moreover, it has been shown that samples presenting nondiscriminating polarizance or diattenuation information, are better classified with the type I canonical space parameters, whereas samples presenting discriminating polarizance or diattenuation information, are better classified by using the IPP space.

The method proposed in the present is general, and we believe that it can be applied to any type of sample. The fact that the use of the method has been illustrated here with a particular case of study does not imply any intrinsic limitation. Taking in account the increasing place that polarimetry is taking in biophotonics; we think that the proposed method has the potential to open new possibilities in the field and to improve the existing approaches providing a means to achieve optimal tissue classification efficiencies. Moreover, since each biological tissue (human, animal or vegetal) shows a singular polarimetric response the proposed protocol can in principle be applied to define the optimal model for the specific application. In this sense an experimental data set (Mueller matrices of the tissues under investigation) must be measured, and then used to feed the statistical data protocol discussed in this manuscript. If properly applied, the method will provide the set of observables which optimize the ability of an automatic classifier to handle the data. The fact that different data sets will eventually result in different sets of optimal observables for subsequent classification, is inherent to the optical properties of the samples themselves, and not a limitation of the method proposed here.

ACKNOWLEDGMENTS

We acknowledge the financial support of Spanish MINECO (RTI2018-097107-B-C31 and Fondos FEDER); Catalan Government (2017-SGR-001500).

ORCID

Albert Van Eeckhout https://orcid.org/0000-0003-2540-2746

REFERENCES

[1] V. V. Tuchin, *Tissue Optics: Light Scattering Methods and Instruments for Medical Diagnosis*, SPIE Press, Bellingham **2007**.
[2] V. V. Tuchin, L. Wang, D. A. Zimnyakov, *Optical Polarization in Biomedical Applications*, Springer-Verlag, Berlin Heidelberg **2006**.
[3] M. Nagendran, D. P. Riordan, P. B. Harbury, T. J. Desai, *Elife* **2018**, *7*: e30510.
[4] S. H. Youn, T. Sim, A. Choi, J. Song, K. Y. Shin, I. K. Lee, H. M. Heo, D. Lee, J. H. Mun, *Comp. Biol. Med.* **2015**, *61*(1), 92.
[5] J. M. Gálvez, D. Castillo, L. J. Herrera, B. San Román, O. Valenzuela, F. M. Ortuño, I. Rojas, *PloS One* **2018**, *13*(5), e0196836.
[6] L. van Manen, J. Dijkstra, C. Boccara, E. Benoit, A. L. Vahrmeijer, M. J. Gora, J. Mieog, *J Cancer Res. Clini. Oncol.* **2018**, *144*(10), 1967.
[7] L. Marcu, S. A. Boppart, M. R. Hutchinson, J. Popp, B. C. Wilson, *J. Biomed. Opt.* **2018**, *23*(2), 021103.
[8] Y. A. Khristoforova, I. A. Bratchenko, O. O. Myakinin, D. N. Artemyev, A. A. Moryatov, A. E. Orlov, S. V. Kozlov, V. P. Zakharov, *J. Biophoton.* **2019**, *12*(4), e201800400.
[9] V. P. Zakharov, I. A. Bratchenko, D. N. Artemyev, O. O. Myakinin, S. V. Kozlov, A. A. Moryatov, A. E. Orlov, *Neurophoton. Biomed. Spect.* **2019**, *1*, 449.
[10] J. Qi, D. S. Elson, *J. Biophotonics* **2017**, *10*, 950.
[11] I. J. Vaughn, B. G. Hoover, J. S. Tyo, *Proc. SPIE* **2012**, *8364*, 83640S.
[12] B. G. Hoover, J. S. Tyo, *Appl. Opt.* **2007**, *46*(34), 8364.
[13] G. Anna, F. Goudail, D. Dolfi, *Opt. Exp.* **2011**, *19*(25), 25367.
[14] A. S. Alenin, L. Morrison, C. Curiel, J. S. Tyo, *Proc. SPIE* **2011**, *8160*, 816014.
[15] L. Graham, Y. Yitzhaky, I. Abdulhalim, *J. Biomed. Opt.* **2013**, *18*, 111403.
[16] T. Yasui, Y. Tohno, T. Araki, *J. Biome. Opt.* **2004**, *9*, 259.
[17] Y. Yitzhaky, L. Graham, I. Abdulhalim, *Proc. of SPIE* **2013**, *8856*, 88562J.
[18] K. Komatsu, L. Mosekilde, A. Viidik, M. Chiba, *Anat. Rec.* **2002**, *268*, 381.
[19] M. Borovkova, V. A. Ushenko, A. V. Dubolazov, O. Y. Vanchulyak, O. G. Ushenko, A. V. Bykov, I. Meglinski, *PLOS One* **2019**, *14*(5), e0214494.
[20] T. Novikova, A. Pierangelo, A. De Martino, A. Benali, P. Validire, *Opt. Phot. News* **2012**, *23*, 26.
[21] A. Pierangelo, A. Nazac, A. Benali, P. Validire, H. Cohen, T. Novikova, B. H. Ibrahim, S. Manhas, C. Fallet, M. R. Antonelli, A. De Martino, *Opt. Exp.* **2013**, *21*, 14120.
[22] E. Du, H. He, N. Zeng, M. Sun, Y. Guo, J. Wu, S. Liu, H. Ma, *J. Biomed. Opt.* **2014**, *19*, 076013.
[23] V. Ushenko, A. Sdobnov, A. Syvokorovskaya, A. Dubolazov, O. Vanchulyak, A. Ushenko, Y. Ushenko, M. Gorsky, M. Sidor, A. Bykov, I. Meglinski, *Photonics* **2018**, *5*(4), 54.
[24] M. Kupinski, M. Boffety, F. Goudail, R. Ossikovski, A. Pierangelo, J. Rehbinder, J. Vizet, T. Novikova, *Biomed. Opt. Exp.* **2018**, *11*(9), 5691.
[25] B. Liu, Y. Yao, R. Liu, H. Ma, L. Ma, *Opt. Commun.* **2019**, *433*, 60.
[26] N. Ghosh, M. F. Wood, I. Vitkin, *J. Biomed. Opt.* **2008**, *3*, 044036.
[27] P. Shukla, A. Pradhan, *Appl. Opt.* **2009**, *48*, 6099.
[28] B. Kunnen, C. Macdonald, A. Doronin, S. Jacques, M. Eccles, I. Meglinski, *J. Biophoton.* **2015**, *8*, 317.
[29] C. W. Sun, C. Y. Wang, C. C. Yang, Y. W. Kiang, I. J. Hsu, C. W. Lin, *Opt. Lett.* **2001**, *26*, 432.
[30] D. Goldstein, *Polarized Light*, 2nd ed., Marcel Dekker, New York, NY **2003**.
[31] E. Garcia-Caurel, R. Ossikovski, M. Foldyna, A. Pierangelo, B. Drévillon, A. De Martino, in *Advanced Mueller Ellipsometry Instrumentation and Data Analysis* (Eds: M. Losurdo, K. Hingerls), Springer-Varlag, Berlin **2013**.
[32] R. A. Chipman, *Polarimetry: Handbook of Optics*, 2nd ed., McGrawHill, New York, NY **1995**.
[33] J. J. Gil, R. Ossikovski, *Polarized Light and the Mueller Matrix Approach*, CRC Press, Boca Raton, FL **2016**.

[34] A. Lizana, A. Van Eeckhout, K. Adamczyk, C. Rodríguez, J. C. Escalera, E. Garcia-Caurel, I. Moreno, J. Campos, *J. Biomed. Opt.* **2017**, *22*, 056004.
[35] J. J. Gil, E. Bernabeu, *Opt. Acta* **1986**, *33*, 185.
[36] J. J. Gil, J. M. Correas, P. A. Melero, C. Ferreira, *Monog. Sem. Mat. G. Galdeano* **2004**, *31*, 161.
[37] I. San José, J. J. Gil, *Opt. Commun.* **2011**, *294*, 38.
[38] S. R. Cloude, *Optik* **1986**, *75*, 26.
[39] S. R. Cloude, *SPIE Proc.* **1989**, *1166*, 177.
[40] A. Van Eeckhout, A. Lizana, E. Garcia-Caurel, J. J. Gil, R. Ossikovski, J. Campos, *Opt. Lett.* **2017**, *42*, 4155.
[41] A. Van Eeckhout, A. Lizana, E. Garcia-Caurel, J.J Gil, A. Sansa, C. Rodríguez, I. Estévez, E. González, J. C. Escalera, I. Moreno, and J. Campos, *J. Biophoton.*, *11*(14): e201700189 (**2018**)x
[42] A. Van Eeckhout, E. Garcia-Caurel, T. Garnatje, M. Durfort, J. C. Escalera, J. Vidal, J. J. Gil, J. Campos, A. Lizana, *PLoS One* **2019**, *14*(3): e0213909.
[43] R. Ossikovski, J. Vizet, J. Opt, *Soc. Am. A* **2019**, *36*, 1173.
[44] R. Ossikovski, *J. Opt. Soc. Am. A* **2009**, *26*, 1109.
[45] R. Ossikovski, *J. Opt. Soc. Am. A* **2010**, *27*, 123.
[46] A. Peinado, A. Lizana, J. Vidal, C. Iemmi, J. Campos, *Opt. Exp.* **2010**, *18*, 9815.
[47] J. Kastelic, A. Galeski, E. Baer, *Connect. Tissue Res.* **1978**, *6*, 11.
[48] D. L. Butler, E. S. Grood, F. R. Noyes, R. F. Zernucke, *Exer. Sports Sci. Rev.* **1978**, *6*, 125.
[49] R. James, G. Kesturu, G. Balian, A. B. Chhabra, *J. Hand Surg. Am* **2008**, *33*(1), 102.
[50] D. J. Hulmes, *Struct. Biol.* **2002**, *137*(1–2), 2.
[51] S. Schiaffino, C. Reggiani, *Physiol. Rev.* **2011**, *91*(4), 1447.
[52] B. Charvet, F. Ruggiero, D. Le Guellec, *Musc. Liga. Tend. J.* **2012**, *2*(2), 53.
[53] H. Asahara, M. Inui, M. K. Lotz, *J. Bone Miner. Res.* **2017**, *32* (9), 1773.
[54] L. Breiman, J. H. Friedman, R. A. Olshen, C. J. Stone, *Classification and Regression Trees*, Chapman & Hall, Boca Raton, FL **1984**.
[55] Y. Guo, T. Hastie, R. Tibshirani, *Biostatistics* **2007**, *8*(1), 86.
[56] T. Mitchell, *Machine Learning*, McGraw-Hill, New York **1997**.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Van Eeckhout A, Garcia-Caurel E, Ossikovski R, et al. Depolarization metric spaces for biological tissues classification. *J. Biophotonics*. 2020;13:e202000083. https://doi.org/10.1002/jbio.202000083