High Correlation of Double Debye Model Parameters in Skin Cancer Detection

Bao C. Q. Truong[1], H. D. Tuan[1] *Member, IEEE*, Anthony J. Fitzgerald[2], Vincent P. Wallace[2], and H.T. Nguyen[1]
Senior Member, IEEE

Abstract—The double Debye model can be used to capture the dielectric response of human skin in terahertz regime due to high water content in the tissue. The increased water proportion is widely considered as a biomarker of carcinogenesis, which gives rise of using this model in skin cancer detection. Therefore, the goal of this paper is to provide a specific analysis of the double Debye parameters in terms of non-melanoma skin cancer classification. Pearson correlation is applied to investigate the sensitivity of these parameters and their combinations to the variation in tumor percentage of skin samples. The most sensitive parameters are then assessed by using the receiver operating characteristic (ROC) plot to confirm their potential of classifying tumor from normal skin. Our positive outcomes support further steps to clinical application of terahertz imaging in skin cancer delineation.

I. INTRODUCTION

Intensive studies of terahertz (THz) radiation (T-ray) have continuously presented the potential of terahertz pulse imaging (TPI) and terahertz pulse spectroscopy (TPS) in medical applications, especially cancer detection [1]–[3]. The THz wave, which is non-ionizing, only utilises a low power level within safety guide [4]. As THz frequencies ($0.2 - 10$ THz) lie in the excitation range of torsional and vibrational motions in molecular systems, TPI and TPS are capable of providing spectroscopic information of biological tissue [5]. The significantly high sensitivity of waves in the terahertz regime to water is a well-known feature contributing to advantages of the imaging technique as biological tissues comprise a large proportion of water [6].

Contrast images between healthy and abnormal tissues such as skin cancer and breast cancer have been previously studied. A THz imaging system based on reflection was especially used to non-invasively detect cancerous regions of human skin while *ex vivo* images of this skin cancer also proved the ability of TPI to distinguish between basal cell carcinoma (BCC) and normal tissue [2], [7]. THz imaging might be also beneficial to breast-conserving surgery as [8] investigated the possibility of using THz pulses to correctly map breast tumor margins. Recent advances in THz technology triggered applications in other cancers occurring at less accessible areas such as cervix and colon [9].

Non-melanoma skin cancer (NMSC), especially basal cell carcinoma, is the most popular cancer occurring within Cau-

casian population [10]. Mohs's micrographic surgery (MMS) is currently the treatment method providing the highest cure rate. However, this technique is intensively time-consuming and its cure efficiency depends on experiences of pathologist in determining the histological margin of cancerous regions. Therefore, accurate delineation of NMSC is essential for not only saving time and effort in terms of recurrent surgeries and biopsy but also simplifying the treatment modality [11]. Good correlation between terahertz images and their histology sections suggested the application of the THz imaging technique *in vivo* to delineate tumor margins pre-operatively [12].

Preliminary studies demonstrated that the dielectric properties of human skin in terahertz frequencies can be modeled by two Debye relaxation processes, known as the double Debye model [13]. The capability of the double Debye model to specify the pathology of tissue was initially investigated by [14]. However, its outcomes were limited because the fitting procedure to extract the model parameters from measured optical properties was recognised mathematically difficult [15]. To overcome this challenge, a further study by [16] proposed a global optimization-based approach to extract the double Debye model parameters which provide the optimal fit to measured complex permittivities of human skin, both healthy and cancerous. Additionally, contrast values in these parameters between the two types of skin tissue suggested the possibility for skin cancer classification. Thereupon, the aim of the present paper is to identify the potential applicants for skin cancer classification from the double Debye parameters via statistical analysis of pathological correlation and the receiver operating characteristic (ROC) plots of these parameters. The global optimization-based approach proposed by [16] is employed to extract the parameters of the double Debye model corresponding to a variety of skin samples. To the authors's best knowledge, this paper is the first study in providing a specific statistical analysis of these parameters. Its outcomes support further steps towards applying the double Debye model for THz imaging in order to assist the recent tumor-removal surgeries.

The paper has the following structure. Section II present the double Debye model, experiment data and applied parameter extraction procedure. Section III describe the statistical analysis of correlation and classification possibility. Discussion on the results are demonstrated in Section IV. Finally, we summarise our contributions in Section V.

[1]Centre for Health Technologies, University of Technology Sydney, Ultimo 2007, Australia; Email: cao.q.truong@student.uts.edu.au, tuan.hoang@uts.edu.au, hung.nguyen@uts.edu.au.

[2]School of Physics, University of Wertern Australia, Crawley 6009, Australia; Email: vincent.wallace@uwa.edu.au.

Table I
THE DOUBLE DEBYE PARAMETERS

Sample	BCC percentage (%)	ε_s	ε_2	ε_∞	$\tau_1 (ps)$	$\tau_2 (ps)$
Normal	0	27.13 ± 10.00	4.64 ± 0.18	2.89 ± 0.19	4.30 ± 2.27	0.11 ± 0.02
BCC	< 30	44.79 ± 32.37	4.83 ± 0.26	3.06 ± 0.13	6.56 ± 5.34	0.12 ± 0.02
BCC	$30 - 50$	53.91 ± 34.69	4.77 ± 0.02	2.92 ± 0.06	8.25 ± 5.67	0.12 ± 0.01
BCC	> 50	73.11 ± 55.17	5.04 ± 0.15	2.97 ± 0.11	9.71 ± 7.66	0.12 ± 0.02

II. PARAMETER EXTRACTION

A. The double Debye model

References [11], [13], [16] confirm the possibility of using the double Debye theory to predict the dielectric response of human skin in the terahertz regime. This approach is based on the molecular interaction between water molecules of skin tissue and THz radiation. The double Debye model incorporates two Debye relaxation processes that represent the impact of an external electric field on water molecules. Under the excitation caused by incident T-rays, natural tetrahedral structures of water molecules are reoriented by the breakage and reformation of multiple hydrogen bonds surrounding central water molecules in the slow relaxation process (on a time scale of picosecond) [11], [17]. The fast relaxation process is attributed to the reorientation of the single central water molecules with a short moment (hundreds of femtoseconds). The double Debye model used for analytically describing the relative complex permittivity of human skin is [5]

$$\tilde{\varepsilon}_r(\omega) = \varepsilon_\infty + \frac{\varepsilon_s - \varepsilon_2}{1 + \jmath\omega\tau_1} + \frac{\varepsilon_2 - \varepsilon_\infty}{1 + \jmath\omega\tau_2}. \quad (1)$$

Here ε_s is static permittivity at low frequency, ε_∞ is the limiting permittivity at high frequency, and ε_2 is the transitional dielectric constant between two relaxation process. $\varepsilon_s - \varepsilon_2$ and $\varepsilon_2 - \varepsilon_\infty$ respectively represent the dispersion in amplitude of the slow and fast relaxation processes corresponding to their relaxation time constants τ_1 and τ_2.

B. Experimental Data and Parameter Extraction

The data used in this study were published by [11]. In particular, tissue samples were excised from mainly head and neck of ten patients undergoing MMS. Each sample includes a certain proportion of BCC and adjacent normal tissue. There are totally 23 samples which include 13 containing tumor and 10 without tumor. A TPI system was used for spectroscopy measurements in transmission mode. The time-domain signals measured from this device are transformed into frequency-domain spectra through Fourier transformation. Amplitude and phase recovered from these spectra facilitate directly measuring frequency-dependent refractive index of tissue n and absorption coefficient α. Further details of this experiment procedure is fully described in [11]. The measured frequency-dependent complex permittivity is calculated through the following relationship:

$$\tilde{\varepsilon}(\omega) = (n(\omega) - \jmath\kappa(\omega))^2. \quad (2)$$

The fitting procedure is based on minimizing the total square error function $\sum_{i=1}^{N} |\tilde{\varepsilon}_r(\omega_i) - \tilde{\varepsilon}(\omega_i)|^2$ over N sampled frequencies in order to obtain the best fit to the permittivity data

from (2). The global optimization-based approach by [16] is employed to find the optimal solution. According to recent studies, τ_1 is reasonably constrained in the range $0.1 - 20$ picoseconds for all samples while the constraint of τ_1 is $[50 - 150]$ femtoseconds [5], [18]. The extracted parameters of the double Debye model corresponding to all tumorous samples are grouped by tumor percentage. The average values of the double Debye parameters together with their standard errors in three groups including samples with less than 30%, $30 - 50\%$ and more than 50% of tumor are recorded in Table. I.

III. STATISTICAL ANALYSIS

A. Correlation Analysis

Reference [11] analysed the correlation between the terahertz absorption coefficients and tumor content of tissue samples for the purpose of evaluating the senility of TPI for detecting tumor. The author found the highest correlation $r = 0.75$ of the absorption at 0.5 THz for each BCC samples with the percentage of tumor. Pearson correlation between each parameter of the double Debye model and tumor percentage for all samples were also presented in Table II. These parameters only provide the correlation up to $r = 0.62$ that is still far lower than $r = 0.75$. However, it is worth considering the fact that the correlations of τ_1 with two parameters $\varepsilon_s, \varepsilon_2$ are significant $(0.99, 0.65$ respectively). This suggests certain combinations of these parameters could improve the correlation with the BCC percentages of skin samples. Therefore, several empirical non-linear relationships of $\varepsilon_s, \varepsilon_2, \varepsilon_\infty, \tau_1$ were considered as follows,

$$\delta_1 = \frac{\varepsilon_s - \varepsilon_2}{\tau_1}, \quad (3)$$

$$\delta_2 = \frac{\varepsilon_s - \varepsilon_2 - \varepsilon_\infty}{\tau_1}, \quad (4)$$

$$\Delta_1 = \frac{\delta_1^{BCC} - \Delta_1^{normal}}{\delta_1^{normal}}, \quad (5)$$

$$\Delta_2 = \frac{\delta_2^{BCC} - \Delta_2^{normal}}{\delta_2^{normal}}, \quad (6)$$

Here, δ_1^{BCC} and δ_2^{BCC} respectively stand for the values of δ_1 and δ_2 corresponding to a BCC-contained sample while δ_1^{normal} and δ_2^{normal} are values of δ_1 and δ_2 for the healthy sample of the same patient with that tumor. δ_1 and δ_2 are continuously calculated for all 23 skin samples before applying (5)-(6) to find Δ_1 and Δ_2 of each BCC sample. (5)-(6) are proposed based on the fact that the hydration profile of skin varies between patients while the double Debye model represents the dielectric response of water content inside skin

tissue. The tumor is generally believed to be more highly hydrated than normal tissue but this context may be only available when both the former and the latter belong to the same patient and maybe even the same examined body part. As a result, the subtraction in (5)-(6) facilitates removing the effect of different skin hydration backgrounds between patients on the changes due to the difference in tumor content. The correlation values of $\delta_1, \delta_2, \Delta_1, \Delta_2$ with BCC percentages are provided in Table II.

Table II
THE PEARSON CORRELATION VALUES OF MODEL PARAMETERS

ε_s	ε_2	ε_∞	$\tau_1(ps)$	$\tau_2(ps)$	δ_1	δ_2	Δ_1	Δ_2
0.44	0.63	-0.1	0.34	0.2	0.74	0.79	0.92	0.85

B. Classification Analysis

The applied Pearson correlation in III-A, which only examine the linear dependence between two variables, is readily not a complete measure for identifying potential indicators in classification. Thus, the receiver operating characteristic (ROC) plot is commonly employed to explore the classification accuracy in medical community [19]. The ROC plots also demonstrate the trade-offs between the sensitivity and specificity. A threshold value chosen by averaging two consecutive instances of original data is used to split the data into two separate sets which are compared with pathology to determine the correspondent sensitivity and specificity. Fig. 1 demonstrate the ROC curves of $\varepsilon_s, \varepsilon_2, \varepsilon_\infty, \tau_1, \tau_2, \delta_1, \delta_2$ and the absorption coefficient α at 0.5 THz respectively. Furthermore, the area under the ROC curve (AUC) of a classification feature is an independent index from any particular threshold value and commonly used to assess the general performance of this feature. An AUC closer to 1 indicates a better classification ability. Specific AUC values of all ROC curves in Fig. 1 can be seen in Table III. According to Fig. 1, it is possible to

Table III
THE AREA UNDER ROC CURVES (AUC).

ε_s	ε_2	ε_∞	τ_1	τ_2	δ_1	δ_2	α
0.89	0.69	0.46	0.727	0.46	0.95	0.93	0.91

determine the ROC point which provides the highest correct classification accuracy (CCR) with respect to each parameter. Table IV only demonstrates specific threshold values responsible for the best CCRs of $\varepsilon_s, \varepsilon_2, \delta_1, \delta_2$ and α together with their sensitivity and specificity due to the prominent AUC values of these parameters.

Table IV
STATISTICAL MEASURES OF THE BEST CLASSIFICATION PERFORMANCE BASED ON ROC ANALYSIS.

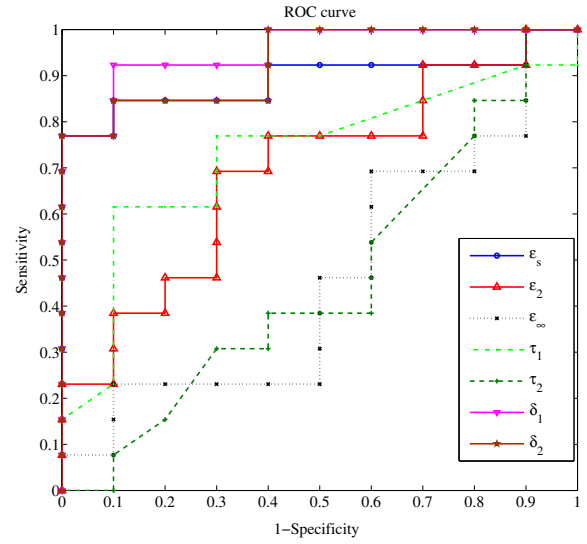| Parameter	Threshold	Sensitivity	Specificity	CCR
ε_s	28.23	84.62	90	86.96
ε_2	4.71	69.23	70	69.57
δ_1	5.80	92.31	90	91.30
δ_2	5.24	84.62	90	86.96
α	119.88	84.62	90	86.96

Fig. 1. ROC curves of $\varepsilon_s, \varepsilon_2, \varepsilon_\infty, \tau_1, \tau_2, \delta_1, \delta_2$ and the absorption coefficient α at 0.5 THz.

IV. DISCUSSION

As can be seen in Table I, the impact of increased tumor proportion inside skin samples on the variation of the double Debye parameters generally results in the increases of these values. In particular, average values of ε_s consistently increase due to the increase of BCC percentage while the similar trend is also observed in the remaining cases of $\varepsilon_2, \varepsilon_\infty$ and τ_1. The high variance of ε_s, τ_1, which are well-correlated, could be due to overfitting. Constraining the variation range of τ_1 will be helpful to reduce the variance of these two parameters. The differences in $\varepsilon_2, \varepsilon_\infty$ are far smaller as compared to those in ε_s, τ_1 and τ_2. In fact, high frequency parameters like $\varepsilon_2, \varepsilon_\infty$ are more dependent on refractive index and low frequency parameters including ε_s, τ_1 and τ_2 are more sensitive to absorption coefficient [5]. Thus, the more prevalent difference between BCC and normal skin in absorption coefficient than refractive index in THz regime can be reflected using the double Debye model.

Despite that Pearson correlations of the five double Debye parameters with tumor percentage are quite low, the combinations of these parameters even demonstrate far higher correlations than the previous study by [11]. Table II particularly demonstrates the higher sensitivity of δ_1 and δ_2 to BCC percentage than that of absorption coefficient at 0.5 THz, which indicates prominent potential of these combinations for skin cancer delineation. The very high standard errors of the Debye parameters as can be seen in Table I limits the sensitivity of each individual parameter to BCC content of samples. However, the dependencies of the double Debye parameters on tumor content may be non-linear whilst the Pearson correlation test is only based on a linear regression analysis. Thus, there may require further studies to clarify this assumption. On the other hand, the aforementioned correlations between the parameters are beneficial to the combinations of the double Debye parameters. The exceptional correlations provided by

Δ_1 ($r = 0.92$) and Δ_2 ($r = 0.85$) suggest that the different nature of water profile between patients' skin tissues does considerably affect the sensitivity of the double Debye model to skin, especially tumorous tissue. Furthermore, the approach proposed in (5)-(6) is worth considering for further applications in order to separate the cancer-related contrast from the natural differences between patients' bodies.

The potential Debye-related parameters for skin cancer classification have been found through intensive analysis presented in section III-B. To be more specific, $\varepsilon_s, \delta_1, \delta_2$ possess the highest AUC indices, which represent themselves as the most prominent explanatory variables for classification. The value of AUC corresponding to the absorption coefficient at 0.5 THz is also recorded in Table III for comparison with those of $\varepsilon_s, \delta_1, \delta_2$, which highlights the supremacy of these Debye parameters in terms of classification potential. This superiority is coherent due to the fact that the double Debye model mathematically incorporates the differences of both the frequency-dependent refractive index and absorption coefficient. In reality, these optical properties also can be combined under the form of time-domain impulse functions or frequency-domain power spectra. The contrast in these factors between cancerous and normal tissue may also include either unknown contributors besides water absorption or misleading features [7]. Deeper analysis of ROC curves in Fig. 1 results in the best threshold values in Table IV for detecting the BCC samples. The best CCR with the sensitivity 92.31% and specificity 90% corresponding to δ_1, which has the highest AUC, is obtained using the threshold value $\delta_1 = 5.80$. δ_2 is also a potential feature for classification since it highly correlates with the tumor percentage of skin sample ($r = 0.79$) as well as achieves impressive sensitivity (84.62%) and specificity (90%). In spite of only having very low Pearson correlation with BCC content, ε_s still expresses itself as a key classification feature with its statistical measures of performance similar to those of δ_2. At the bottom line, as the notable classification outcomes are simply based on a very basic method using threshold values, applying more complex classifiers such as support vector machine promise significant improvement of accuracy.

V. Conclusion

We have provided a comprehensive analysis of the double Debye model in terms of its sensitivity to tumor content of human skin and the potential of using its parameters to classify tumors from normal skin tissue. While there is no significant correlation of each Debye parameter with the BCC percentage of skin sample, the proposed combination coefficients $\delta_1, \delta_2, \Delta_1, \Delta_2$ are significantly dependent on the tumor content with the corresponding Pearson correlation up to $r = 92\%$. In addition, both δ_1 and δ_2, our proposed combination of the double Debye parameters, acquire promising values of AUC, sensitivity and specificity which confirm the striking potential of using these parameters to detect skin tumor. Future studies will consider applying advanced classification algorithms together with intensive tests on THz images of skin cancer.

References

[1] S. Nakajima, H. Hoshina, M. Yamashita, C. Otani, and N. Miyoshi, "Terahertz imaging diagnostics of cancer tissues with a chemometrics technique," *Applied Physics Letters*, vol. 90, no. 4, p. 041102, 2007.

[2] R. M. Woodward, B. E. Cole, V. P. Wallace, R. J. Pye, D. D. Arnone, E. H. Linfield, and M. Pepper, "Terahertz pulse imaging in reflection geometry of human skin cancer and skin tissue," *Physics in Medicine and Biology*, vol. 47, no. 21, pp. 3853–3863, 2002.

[3] V. P. Wallace, P. F. Taday, A. J. Fitzgerald, R. M. Woodward, J. Cluff, R. J. Pye, and D. D. Arnone, "Terahertz pulsed imaging and spectroscopy for biomedical and pharmaceutical applications," *Faraday Discussions*, vol. 126, pp. 255–263, 2004.

[4] E. Berry, G. C. Walker, A. J. Fitzgerald, N. N. Zinovev, M. Chamberlain, S. W. Smye, R. E. Miles, and M. A. Smith, "Do in vivo terahertz imaging systems comply with safety guidelines?," *Journal of Laser Applications*, vol. 15, no. 3, 2003.

[5] E. Pickwell, A. J. Fitzgerald, B. E. Cole, P. F. Taday, R. J. Pye, T. Ha, M. Pepper, and V. P. Wallace, "Simulating the response of terahertz radiation to basal cell carcinoma using ex vivo spectroscopy measurements," *Journal of Biomedical Optics*, vol. 10, no. 6, p. 064021, 2005.

[6] Z. D. Taylor, R. S. Singh, D. B. Bennett, P. Tewari, C. P. Kealey, N. Bajwa, M. O. Culjat, A. Stojadinovic, H. Lee, J. P. Hubschman, E. R. Brown, and W. S. Grundfest, "THz medical imaging: in vivo hydration sensing," *IEEE Transactions on Terahertz Science and Technology*, vol. 1, no. 1, pp. 201–219, 2011.

[7] R. M. Woodward, V. P. Wallace, R. J. Pye, B. E. Cole, D. D. Arnone, E. H. Linfield, and M. Pepper, "Terahertz pulse imaging of ex vivo basal cell carcinoma.," *The Journal of Investigative Dermatology*, vol. 120, no. 1, pp. 72–78, 2003.

[8] A. J. Fitzgerald, V. P. Wallace, M. Jimenez-Linan, L. Bobrow, R. J. Pye, A. D. Purushotham, and D. D. Arnone, "Terahertz pulsed imaging of human breast tumors," *Radiology*, vol. 239, no. 2, pp. 533–540, 2006.

[9] K. Wang and D. M. Mittleman, "Metal wires of terahertz wave guilding," *Nature*, vol. 430, pp. 376–379, 2012.

[10] M. Mogensen and G. B. E. Jemec, "Diagnosis of nonmelanoma skin cancer/keratinocyte carcinoma: A review of diagnostic accuracy of nonmelanoma skin cancer diagnostic tests and technologies," *Dermatologic Surgery*, vol. 33, no. 10, pp. 1158–1174, 2007.

[11] V. P. Wallace, A. J. Fitzgerald, E. Pickwell, R. J. Pye, P. F. Taday, N. Flanagan, and T. Ha, "Terahertz pulsed spectroscopy of human basal cell carcinoma.," *Applied Spectroscopy*, vol. 60, no. 10, pp. 1127–1133, 2006.

[12] V. Wallace, A. Fitzgerald, S. Shankar, N. Flanagan, R. Pye, J. Cluff, and D. Arnone, "Terahertz pulsed imaging of basal cell carcinoma ex vivo and in vivo," *British Journal of Dermatology*, vol. 151, no. 2, pp. 424–432, 2004.

[13] E. Pickwell, B. E. Cole, A. J. Fitzgerald, M. Pepper, and V. P. Wallace, "In vivo study of human skin using pulsed terahertz radiation," *Physics in Medicine and Biology*, vol. 49, no. 9, pp. 1595–1607, 2004.

[14] C. B. Reid, E. Pickwell-MacPherson, J. G. Laufer, A. P. Gibson, J. C. Hebden, and V. P. Wallace, "Accuracy and resolution of thz reflection spectroscopy for medical imaging," *Physics in Medicine and Biology*, vol. 55, no. 16, p. 4825, 2010.

[15] B. C. Q. Truong, H. D. Tuan, H. H. Kha, and H. T. Nguyen, "System identification for terahertz wave's propagation and reflection in human skin," in *Communications and Electronics (ICCE), 2012 Fourth International Conference on*, pp. 364–368, Aug 2012.

[16] B. C. Q. Truong, H. D. Tuan, H. H. Kha, and H. T. Nguyen, "Debye parameter extraction for characterizing interaction of terahertz radiation with human skin tissue," *IEEE Transactions on Biomedical Engineering*, vol. 60, pp. 1528–1537, June 2013.

[17] C. Ronne and S. R. Keiding, "Low frequency spectroscopy of liquid water using thz-time domain spectroscopy," *Journal of Molecular Liquids*, vol. 101, no. 13, pp. 199 – 218, 2002. Molecular Liquids. Water at the New Millenium.

[18] J. T. Kindt and C. A. Schmuttenmaer, "Far-infrared dielectric properties of polar liquids probed by femtosecond terahertz pulse spectroscopy," *The Journal of Physical Chemistry*, vol. 100, no. 24, pp. 10373–10379, 1996.

[19] T. Fawcett, "Roc graphs: Notes and practical considerations for researchers," *ReCALL*, vol. 31, no. HPL-2003-4, pp. 1–38, 2004.