Terahertz scattering microscopy for dermatology diagnostics

Alfonso Alessandro Tanga1, Valeria Giliberti2, Francesco Vitucci3, Domenico Vitulano4,5, Vittoria Bruni4,5, Andrea Rossetti1, Gabriele Carmine Messina1, Maddalena Daniele2, Giancarlo Ruocco2 and Michele Ortolani1,*

1 Sapienza University of Rome, Department of Physics, Rome, Italy
2 Istituto Italiano di Tecnologia, Center for Life Nano Science, Rome, Italy
3 Crisel Instruments srl., Rome, Italy
4 Sapienza University of Rome, Department of Basic and Applied Sciences for Engineering, Rome, Italy
5 National Research Council of Italy (CNR), Rome, Italy
* Author to whom any correspondence should be addressed.
E-mail: michele.ortolani.roma1.infn.it

Keywords: point spread function, microbolometer camera, dermatology diagnostics, terahertz microscopy

Abstract
We explore the possibility of detecting anomalous structures buried under the skin surface by studying the deviations from the ideal Airy pattern of the point-spread function (PSF) of a terahertz microscope that includes the skin as one of the reflecting surfaces of the optical system. Using a custom terahertz microscope with a monochromatic point source emitting at 0.611 THz, we record the PSF images with a microbolometer camera. Skin simulants based on collagen gel, with and without artificial buried structures, have been analyzed. The geometrical features characterizing the PSF deformations have been extracted automatically from the PSF images. A machine learning algorithm applied to these geometrical features produces a reliable classification of targets with or without buried structures with error below 5%. It can even classify targets with anisotropic buried structures according to their different orientation.

1. Introduction
Terahertz (THz) imaging holds great promise for medical diagnostics in pathology cases where buried structures have to be localized with sub-millimeter precision in both depth and lateral position. A typical example is provided by skin cancer detection, which has been proposed in the early days of THz time-domain spectroscopy [1]. THz instrumentation has progressed significantly, reaching in vivo real-time imaging capabilities [2] and sub-surface detection capabilities [3–5] but a terahertz instrument for detection of buried anomalies in the skin has not yet been approved. Such an instrument would be for example the key to an early diagnosis of the basal cell carcinoma (BCC) [6, 7] and/or dysplastic skin nevi [8]. Beyond early skin cancer diagnostics, other medical diagnostics application of THz imaging include the characterization of burns [9], scars [2], breast cancer [10, 11] and corneal hydration levels [12, 13]. All of these approaches are based on probe beam spectra typically covering a broad bandwidth of hundreds of GHz in the sub-THz range [14].

In this work, we explore a novel approach to THz imaging for dermatology diagnostics. Our approach is based on the analysis of the modification of the point-spread function (PSF) of an optical system, of which the skin surface becomes one of the reflecting components. The PSF can be precisely imaged using a monochromatic THz point source and a terahertz camera. The optical system employed in this work is a custom-built terahertz microscope with monochromatic continuous-wave illumination at 0.611 THz. We have obtained many thousands of images of skin simulants produced by us with and without buried structures. The geometrical features of the microscope PSF, closely approximated by the two-dimensional Airy pattern made of a central disk and a succession of dark and bright rings, have been automatically extracted from the images. A machine learning algorithm has been employed to classify the images as anomalous or not, depending if a buried structure was detected or not. In a final experiment, we also...
Figure 1. Optical path of the THz radiation within the microscope. THz radiation is emitted through a horn antenna (R1) and refocused on the sample with a parabolic mirror pair. The second mirror can be displaced to change the sample area under analysis. An objective made of two TPX lenses (R2) focuses the reflected spot on the camera for imaging the PSF.

demonstrate the capabilities of our technique to determine the orientation of an anisotropic structure buried under collagen.

2. Experiment

2.1. Microscope

In our microscope the Gaussian beam of the monochromatic radiation source emitting at 0.611 THz is focused on the sample, and the reflected beam is refocused onto the camera surface, i.e. the focal plane detector array is not located at the image plane, but at the aperture plane. In this way, the camera does not record a real image of the skin surface, rather it returns the PSF of the optical system that comprehends the sample under test [15]. The microscope scheme is shown in figure 1. The THz source (TeraSchottky 600 by Lytid s.a.s., France) is based on a Schottky diode multiplier chain emitting 2.5 mW in continuous wave at 0.611 THz through a pyramidal horn antenna. The horn antenna functions as an optical point-source [16] located in the focus of the first parabolic mirror in figure 1. A second parabolic mirror (diameter $D = 50$ mm, focal length $f = 100$ mm) focuses the beam onto the sample with an incidence angle of 45 degrees. Following [17], the diffraction-limited spot diameter $A$ at the sample position for a THz Gaussian beam is obtained from the free-space wavelength $\lambda = 491 \mu$m and the focusing mirror parameters $f, D$ as:

$$A = 1.13 \cdot \lambda \cdot (f/D)$$

where the cylindrical truncation factor due to the aperture stop $k \sim 1$ [17] is set here by the first parabolic mirror diameter of 25 mm, leading to $A \approx 1.1$ mm. $A$ sets the smallest possible target area on the skin that could be analyzed by our microscope. Due to the incidence angle of 45 degrees, the spot diameter on the skin surface is elongated in the incidence plane by a factor $(1/\cos(45°))$ to approximately 2 mm, a lateral resolution which is in line with the literature on THz systems for cancer detection [10, 11]. The reflected radiation is then re-focused on the THz camera surface by a THz objective formed by two TPX lenses with magnification of 0.5, chosen to increase the power density for each pixel of the camera. The THz camera (TZcam by i2s, France) is a 320 $\times$ 240 antenna-coupled microbolometer focal plane array (FPA) developed at CEA-LETI (Grenoble, France) whose response has been optimized for the THz range [18]. The FPA is processed above a CMOS read-out integrated circuit with standard silicon technology. A special pixel architecture separates the THz radiation absorber (antennas with resistive loads) and the thermometer (amorphous silicon film). Crossed double bow-tie antennas (formed by a direct-coupling antenna and a capacitive-coupling antenna) collect the THz flux at any polarization. The thermal radiation background at frequencies far from 0.611 THz is suppressed by the use of a band-pass metamaterial filter as the sensor window (a patterned high-resistivity silicon wafer, also working as vacuum seal). The side of the square-shaped pixels of the FPA is 50 $\mu$m, which is almost ten times smaller than $\lambda$. This oversampling condition is required in studies of the geometrical structure of the PSF [15]: indeed, if the pixel side would have been significantly larger than $\lambda$ (as in conventional digital microscopy), the entire PSF would have contributed to the signal from only one pixel. In this work, the Airy disk diameter approximately fills five
pixels. The electric-field polarization direction is vertical (p-polarized, or parallel to the incidence plane). No polarization-sensitive optics were employed.

2.2. Skin simulants

In order to simulate the optical behavior of the human skin at 0.611 THz, we use a gel composed of hydrated collagen of porcine origin used for food processing [19]. Collagen is the main component of the epidermis and it can be reversibly hydrated at widely different levels. Previous THz spectroscopy studies [20] has shown that a solution of collagen in phosphate-buffered saline and the human skin have very similar optical response around 500 GHz, with a normal-incidence reflectivity around 5%, slightly increasing with hydration level. From [21] we retrieve the values for fresh human skin at 0.611 THz of the refractive index

\[ n(\omega) = 2.1 \pm 0.2, \]

the absorption coefficient \( \mu_a = 125 \pm 10 \text{ cm}^{-1} \) and the optical penetration depth \( \delta = 80 \pm 5 \mu\text{m} \), which we used for optical simulations. Fourier-transform infrared (FTIR) spectroscopy of our skin simulants in transmission mode approximately confirmed these values, although care must be taken when using FTIR spectroscopy at 0.611 THz due to the large experimental errors.

The presence of anomalous structures under the skin surface has been simulated by placing plexiglas disks in the collagen matrix, as shown in figure 2. The plexiglas disk is meant as a buried structure with high terahertz-index contrast [3, 22], with no specific resemblance to the skin cancer refractive index value in the terahertz range. The surface of the different plexiglas disks was patterned by high-resolution 2D laser-ablation writing performed with a CO\(_2\) laser. Different two-dimensional patterns, such as linear gratings, square arrays, hexagonal arrays, disordered arrays, with depth of approximately 0.5 mm were obtained on the plexiglas disk surface. The refractive index of plexiglas at 0.611 THz is around 1.4, quite different from that of hydrated collagen of 2.1, so a refractive index contrast is expected. This value of the refractive index is thought to be similar to that of lipid inclusions that can occasionally form in the skin. A close geometrical relation between the patterns engraved in the plexiglas disks and the typical structures found in BCC or in other tumours has not been established at the present stage of our research.

The skin simulant preparation protocol is: 1 g of collagen dissolved in 50 ml of water, heated to bubbling threshold (approx. 70 °C); solution cast into a 3 mm thick hard plastic box without a cover, which can contain a plexiglas disk glued to its bottom, or not; gelification by water evaporation during cooling to room temperature for 10 minutes.

3. Data acquisition and analysis

The images to be processed were the average of ten subsequent frames taken at 10 Hz frame rate. The blackbody contribution from the environment at 0.611 THz has been removed by subtracting a dark frame (image with source off) acquired after each bright frame. In total, the time required for the acquisition of an image corresponds to the collection of 20 frames (2 s), an image acquisition time that we consider compatible with real-time dermatology diagnostics on live patients. Throughout this article, the intensity is given in analog-to-digital-converter (ADC) units, assuming linearity of the readout circuit in the small dynamic range of less than two orders of magnitude encompassed by the present data.

3.1. Definition of the THz spot

Approximately a quarter of the camera pixels display nonzero values corresponding to a THz signal, but most of them are close to the noise floor, so the first step of our algorithm is to select, as the region of interest, a
smaller number of pixels around the pixel of maximum intensity $I_{\text{max}}$. We defined the THz spot as the ensemble of the pixels with values above a threshold $I_{\text{max}}/e$. A 2D-smoothing method was applied to the raw images in order to obtain only THz spots made of contiguous pixels. Two different smoothing methods were alternatively applied: a 2D Gaussian smoothing and a Wavelet-transform-based smoothing. Indeed, as it can be seen in figure 3, the observed THz spot has the form of an Airy disk with a first dark ring and a first bright ring also visible in the figure. Although the analysis of the entire Airy pattern could lead to interesting information, we decided to restrict our algorithm to the central Airy disk and reject the rings, whose visibility is barely above the noise level. This is mostly due to the correction to the Airy pattern expected from Gaussian beam illumination instead of plane-wave illumination, which suppresses the intensity of the external rings. We also applied a flat field correction (FFC), which resulted in the dark ring between the central maximum and the first bright ring becoming a region of quasi-zero signal. In the FFC, the output values of the left part of the camera field (80 columns, corresponding to one quarter of the total field of 320 columns) were averaged and the result was subtracted from the rest of the camera field so as to obtain zero values far from the THz spot. The result of this FFC base-value removal is shown in figure 4: the application of the contiguity criterion now results in the definition of the THz spot as the region corresponding to the central Airy disk only. After the FFC, the peak noise value was just below 3 ADC units. Due to low signal-to-noise (S/N) ratio, in fact, several pixels far from the THz spot ended up to be above threshold, especially for weakly reflecting samples with $I_{\text{max}} \sim 20$ ADC units. The contiguity criterion allowed us to reject these pixels.
3.2. Extracted features

3.2.1. Maximum intensity

The maximum intensity $I_{\text{max}}$ is taken as the pixel value at the center of the reflected spot. Its precise value can depend on the 2D-smoothing method applied (2D-Gaussian or Wavelet-transform), but we have eventually found that this difference does not impact on image classification significantly (see below). $I_{\text{max}}$ is employed as input in an automated routine for sample-plane pre-alignment, which ensures that all investigated samples lie in the same geometrical plane relative to source and camera.

3.2.2. Area

The area is simply defined as the number of pixels above the threshold $I_{\text{max}}/e$. This value is quite robust against different 2D-smoothing methods, and so are all other features defined below, therefore we kept the distinction between \textit{Max Gaussian} and \textit{Max Wavelet}, and \textit{Area Gaussian} and \textit{Area Wavelet}, but we neglected this distinction for all other features.

3.2.3. Perimeter

The perimeter is calculated with the Matlab \textit{regionprops} function as well as the sum of the distance between each adjoining pair of pixels around the region boundary. This value is ill-defined for non-contiguous spots, which allows one to easily detect such situations.

3.2.4. Integral

The integral is given by the sum of all values taken by pixels belonging to the THz spot. Because the functional form of the THz spot is bound to be an Airy disk, the integral is strongly correlated to the maximum value.

3.2.5. Eccentricity, major and minor axis

The eccentricity is calculated with the Matlab \textit{regionprops} function that calculates the ellipse that has the same second-moments as the region above threshold and extrapolates the eccentricity as the ratio of the distance between the foci of the ellipse and its major axis length. This value is bound between 0 and 1 (an ellipse whose eccentricity is 0 is actually a circle, while an ellipse whose eccentricity is 1 is a line segment). The Major and Minor Axis are simply the length in pixels of the two axes of the ellipse.

3.2.6. Asymmetry

Some spots are asymmetric along the X axis, as shown in figure 5. This can be due to the presence of multiple overlapping Airy disks, because the bilayer reflections at non-normal incidence at the air-collagen and collagen-plexiglas interfaces imply a spatial shift of the center position of the different reflected beams that end up in different center pixels at the camera plane. To define the asymmetry, we used the ratio between two terms: the first is the distance between the center (maximum) and the farthest pixel above the threshold on its row, while the second is the major axis of the above-defined ellipse that has the same second-moments as the THz spot.
3.3. Description of decision tree algorithms

In order to automate the analysis and to use all features simultaneously, we decided to use a machine learning approach that allows a multivariate classification analysis with almost any number of pre-defined features. To this specific aim, we used two machine learning models: the decision tree (DT), and its evolution, the boosted decision tree (BDT). The DT is a non-metric algorithm, based on the classification of each sample in a different cluster according to a series of questions. In figure 6 we show a generic DT and the classification of samples in clusters that results from it. Every division of the initial dataset in clusters leads to an increase of the purity of the new level of the tree, which is related to the entropy of each generated cluster. In fact, the information gain at each step is given by the difference between the entropy of a ‘parent’ node and the sum of the entropies of its ‘children’ nodes \[23\]. One of the main advantages of this algorithm is its simple interpretation (2D graphic representation) and implementation: the DT provides a recursive partitioning of the ensemble in regions of increasing purity. Although the DT concept can be generalized to multi-variable trees, we have limited our classification to two-variable trees (only two possible answers to a question, e.g. ‘is a given parameter higher or lower than a given threshold?’).

A problem of the DT algorithm is its low stability: small variations in the training ensemble can have large effects on classification of the test ensemble. A solution is given by cross validation algorithms in which many different training sub-ensembles are chosen in the full training ensemble. Here we used the \(k\)-fold cross validation algorithm where the samples in the training sub-ensemble are divided in \(k\) folds, and the DT algorithm is run \(k\) times, changing the composition of the training and the test ensemble every time. The final DT is obtained as the average of all different trees created by the different combinations of training and test ensembles.

The principle of using a recursive method to obtain higher discriminant power naturally leads to a more advanced machine learning algorithm called the BDT. In a BDT, the same classification algorithm is trained several times, each time using a reweighed (boosted) ensemble of training events. The events that have been mis-classified during a classification run (growth of one tree) receive a greater weight in the growth of the subsequent tree, until a set of a given number of trees is obtained (a ‘forest’). Each sample in the test ensemble is then classified on the basis of the weighed average of the response given by all trees in the ‘forest’. In the first experiment 4500 images were employed, while in the second experiment we had 1500 images. We have set aside one (or more, see above the description of the different algorithms) sub-set of samples for each experiment for training the algorithms. The training:test ratio was 30:70.

4. Results

4.1. First experiment: detection of buried structures

In the first series of experiments, we demonstrate the discrimination between samples made of pure collagen and samples with a plexiglas disk immersed in a collagen matrix. In properly prepared skin simulants, when the collagen gel retracts after water evaporation, a layer of collagen of thickness \(t\) remains above the plexiglas. For \(t < 300\ \mu\text{m}\), the values of many features are extremely different from those extracted from samples made of pure collagen, leading to clear discrimination. For \(300\ \mu\text{m} < t < 600\ \mu\text{m}\), the difference in the THz spots of pure-collagen and buried-disk samples is unnoticeable when the images are seen at naked eye (see figure 7) but we have obtained meaningful classification results with our multi-feature machine learning algorithms (see below). Samples with \(t\) up to 1 mm were investigated: for \(t > 600\ \mu\text{m}\) we seldom obtained meaningful classification.
Our assumption is that all features, taken individually, are heavily influenced by external variables like hydration (sweating in live patients), environmental temperature (body temperature in live patients) or sample roughness (skin texture in live patients). The use of a machine learning approach may help the classification based on simultaneous analysis of many features, instead of one or two. The selection of the most discriminant features was obtained a posteriori by comparing their weights in the output of an image classification algorithm. In order to obtain a very large number of distinct acquisitions, we measured the same sample at different times during water evaporation and consequent collagen contraction. This resulted in around 500 acquisitions of geometrically different samples, all aligned in the exactly same sample plane, for each prepared skin simulant. Also, in the skin cancer simulant (buried-disk) samples, the depth \( t \) at which the disk was buried changed for each acquisition.

Ensembles of samples (PSF images) were generated by merging the datasets of around 500 images (sub-ensembles) taken on each healthy skin simulant (pure collagen) and on each skin cancer simulant (buried plexiglas disk with its own pattern geometry). In total, we obtained an ensemble of roughly 10,000 images of distinct simulants, which is sufficient for a reasonable application of machine learning algorithms. While generating the ensembles, we have defined normalized values of the features as the ratio of the raw feature value to the average value calculated on the sub-ensemble of images produced on a single skin simulant so as to reduce the feature variability between sub-ensembles.

Normalized intensity values oscillate in an interval between 0.5 and 1.5 for both types of simulants, pure-collagen and buried-disk. The pure-collagen simulants return a higher average value than the buried-disk ones, as shown on the x-axis of figure 8. We speculate that the presence of an optical interface, although buried, could lead to a decrease of the specular reflection due to interference effects.

The strong correlation of the normalized maximum intensity with the normalized area of the spot is highlighted in the scatter plot of figure 8, where the intensity is reported on the X-axis and the area is reported on the Y-axis. The area ranges between 0.8 and 1.6, and it is larger on average for the buried-disk samples. The scatter plot shows that all samples are distributed on an hyperbola, probably due to the electromagnetic energy conservation: if the normalized energy is taken as unity, the normalized maximum intensity should be inversely proportional to the area because energy is either specularly reflected or scattered at non-specular angle. As a result, the normalized integral is approximately the same for all samples, and it is proportional to the product of the area times the maximum intensity. This in turn generates the hyperbolic dependence of the area on the maximum intensity.

It can be also noted that there is a predominance of pure-collagen samples at the center of the plot (close to point \((1;1)\)). Also, there is a cluster of buried-disk samples in the same region, corresponding to data taken in the first part of the evaporation process, where the thickness of the collagen layer above the plexiglas disk is higher and the buried-disk samples resemble more the pure-collagen samples. This observation clarifies that a multi-feature analysis is required for a meaningful discrimination of buried-disk samples. A manual analysis of the features, performed by generating scatter plots with all possible pairs of features on the X- and Y-axes, has confirmed that the asymmetry and the ratio between major and minor axis of the above-defined ellipse have good discriminatory power, but an automated multi-feature protocol is certainly to be preferred.
4.2. First experiment: machine learning results

In a DT algorithm it is possible to use any number of features at the same time, even if some of them are strictly correlated, without the risk of overtraining the model. Before using the algorithm with all features, however, we have run it with two features only, but with all possible pairs of features, in order to understand which pair of features had the highest discriminatory power. In the table shown in figure 9 (left) the classification error is reported for all feature pairs (and for single features along the diagonal) for the DT algorithm. The classification error is simply given by the ratio (given in percentage) of the number of wrong classifications over the total number of samples (recall that we actually know the correct classification for each image). The columns with lower error indicate the features with higher discriminatory power, e.g. the perimeter. Interestingly, features that have low discriminatory power when used alone, become highly discriminant when used in conjunction with another low-discriminatory power feature (e.g. the asymmetry and the minor axis). This fact hints to the simultaneous analysis of all features together in a machine learning approach. In figure 9 (right) it is reported the classification error for the BDT algorithm, which shows a trend similar to that of the DT algorithm but with lower error percentages.

When all ten features were used at the same time, we obtained the decision tree shown in figure 10, with a final classification error of less than 5%. The Area was found to be the feature with the highest discriminatory power. The results with the BDT did not vary significantly from those of the simple DT when all ten features were used.
4.3. Second experiment: discrimination between different buried structures

In order to make a significant step towards skin cancer detection and not simple anomaly detection under the skin, we now demonstrate the capability of our microscope and algorithm to discriminate between two different buried-disk geometries, a situation that simulates two different kinds of sub-surface anomalies without specific reference to any tumour structure for the moment. We fabricated a plexiglas disk with a linear grating pattern (grating period of 1.2 mm, grating duty cycle of 0.5). The disk was buried under hydrated collagen with thickness $t \sim 500 \mu m$, decreasing with time due to water evaporation as in the above described experiment. The sample was then rotated under the microscope without changing the sample-plane orientation in the microscope, de facto obtaining many different samples with identical geometry apart from the orientation of the linear grating pattern with respect to the incidence plane, as shown in the top panel of figure 11. For each value of the thickness $t$ obtained during water evaporation, we performed one acquisition with the lines parallel to the incidence plane and one with lines orthogonal to it. We used the same data acquisition and analysis methods as for the previous experiment.

We extracted all feature with the aim of providing classification of buried disks into sub-categories. As expected, the normalized values of the features are similar to the buried-disk samples of the previous experiment, therefore the scatter plot with the normalized maximum intensity on the $X$-axis and the normalized area on the $Y$-axis shown in figure 11 (bottom) is quite similar to the buried-disk cluster of figure 8. One may glimpse, however, that the samples with lines parallel to the incidence plane (hereafter named parallel samples) are more often located in the high-$Y$, low-$X$ region along the hyperbole if compared to the perpendicular samples: the THz spot reflected by the parallel samples is on average broader than that reflected by the perpendicular samples. This behavior corresponds to that of the normalized eccentricity, which is in the range 0.4–1.4 for the perpendicular samples compared to 0.8–1.8 for the parallel samples. The major axis of the ellipse also shows significant differences in the total range, while the minor axis is on average identical for the two sample groups, because it is already limited by the diffraction limit of equation (1). The major axis of the perpendicular samples ranges from 1 to 1.2, while that of the parallel samples can reach up to 1.6. The major axis has high correlation to both the eccentricity and the area.

4.4. Second experiment: machine learning results

We applied the same DT and BDT algorithms used for the previous analysis, both for all the combinations of pairs of features and for the use of all of them at the same time. In the table shown in figure 12 on the left, reporting the classification error (in percentage) for all pairs and the single features along the diagonal, it is possible to see that there are three features which have averagely a value lower than the others: the eccentricity, the major and the minor axes. The eccentricity seems to have the higher discriminant power regardless of the feature with which it is associated, and it has also the lowest loss when used as singular feature. The major and the minor axes, when taken on their own account do not have an high discriminant power, but when they are used together with features like the area, the perimeter, the maximum or with each other, they have a great increase in discriminant power, showing the great improvement obtained by the use of two or more features at the same time. The maximum, the area and the perimeter show average values, with some peak of high discriminant power. The integral and the asymmetry appear to be the less discriminant. When the BDT algorithm is applied, the results are similar and change of only few percentage points, as shown in figure 12 on the right.
Figure 11. Top: pictorial scheme defining the perpendicular and parallel directions of the line pattern engraved on the disk surface with respect to the incidence plane, represented by red cones. Bottom: scatter plot of all samples of the second experiment in the area vs. maximum intensity plane. The circle color defines the known (training) pattern direction: parallel (green) and perpendicular (pink).

Figure 12. Classification error table for all possible feature pairs used in two-feature trees (left: DTs, right: BDTs) for the second experiment. The error is expressed as the percentage of wrong discrimination between two orthogonal orientations of the same disk with a line pattern on its surface.

When all the features are used at the same time, we obtain the tree shown in figure 13. It has 12 splits and has a percentage of error of 12%. As expected from the tables of the pairs, the first node has the eccentricity as discriminant feature. Instead the following nodes were not expected, the integral of the spot and its asymmetry. It appears that they become more relevant when used with more than another feature, but it could not be foreseen without trying all the other possible combinations with three or more features at the same time. The axes and the perimeter are used at different levels. It is interesting to note that the variable we
Figure 13. Decision tree obtained using all the features for the discrimination between two orthogonal orientations of the same sample with a strip-pattern on its surface. The classification error using this tree is below 12%.

Figure 14. Variation of the minimum, mean and maximum errors as a function of the number of folds created.

relied more on along all the measurements, because of their easy interpretation, the maximum intensity and the area, do not appear in this tree. However they are strictly correlated to other features like the integral of the spot or its perimeter, so this could be simply a coincidence.

In order to find out if the results depended on the choice of the training and the test data, we tried to change the number of folds used for the k-fold cross validation algorithm described in section 3.3. Starting from the minimum value of $k = 2$ we arrived to $k = 40$, an extreme case. In figure 14 are shown the maximum, minimum and mean value of the error of all the trees generated per iteration, that are $k$ at the $k$th iteration. As we can see the mean value has only a slightly decreasing trend, that can be approximated to a constant trend. The maximum and minimum value move away from the mean value when $k$ increases, meaning that if we did not use this cross validation method, we could have had results very different from the mean ones, the most reliable.

5. Conclusions

A custom terahertz reflection microscope with monochromatic illumination at 0.611 THz has been designed and built to image the deformations of the point-spread function due to non-ideal reflecting surfaces with a terahertz camera. The samples analyzed here were collagen-based skin simulants with buried plexiglas disks conceived as simulants of anomalous sub-surface structures of the skin. Several numerical features have been
extracted by image analysis of the central Airy disk of the point-spread functions. Classification of images as related to samples in which buried structures are present or not has been performed with two machine learning algorithms (DT and BDT) applied to the extracted numerical features. To achieve statistical significance, the image features have been extracted automatically from several thousands of different images, partly used for training the algorithms, eventually leading to classification errors smaller than 5%. The results are relevant for future applications of terahertz imaging to medical diagnostics, including early skin cancer detection.

Data availability statement

The data that support the findings of this study are available upon reasonable request from the authors.

Acknowledgment

This work was funded by Regional Funds POR FESR Lazio 2014–2020, LIFE2020 program, project MICOTED.

ORCID iD

Michele Ortolani  https://orcid.org/0000-0002-7203-5355

References

[1] Woodward R M, Wallace V P, Arnone D D, Linfield E H and Pepper M 2003 Terahertz pulsed imaging of skin cancer in the time and frequency domain J. Biol. Phys. 29 257–9
[2] Wang J, Stantchev R, Sun Q, Chiu T, Ahuja A and Pickwell MacPherson E 2018 THz in vivo measurements: the effects of pressure on skin reflectivity Biomed. Opt. Express 9 6467–76
[3] Fitzgerald A J, Tie X, Hackmann M J, Cense B, Gibson A P and Wallace V P 2020 Co-registered combined OCT and THz imaging to extract depth and refractive index of a tissue-equivalent test object Biomed. Opt. Express 11 1417–31
[4] Gibson et al 2018 An assessment of multimodal imaging of subsurface text in mummy cartonnage using surrogate papyrus phantoms Science Hist. 6 7
[5] Ciano C, Flammini M, Giliberti V, Calvani P, DelRe E, Talarico F, Torre M, Misiori M and Ortolani M 2018 Confocal imaging at 0.3 THz with depth resolution of a painted wood artwork for the identification of buried thin metal foils IEEE Trans. Terahertz Sci. Technol. 8 390–6
[6] Wallace V P, Fitzgerald A J, Pickwell E, Bye R J, Taday P F, Flanagan N and Ha T 2006 Terahertz pulsed spectroscopy of human Basal cell carcinoma Appl. Spectrosc. 60 1127–33
[7] Yaroslavsky A N, Joseph C, Patel R, Muzikansky A, Neel V A and Giles R H 2017 Delineating nonmelanoma skin cancer margins using terahertz and optical imaging J. Biomed. Photon. Eng. 3 010301
[8] Zaitsev K I, Kudrin K G, Karasik V E, Reshetov I V and Yurchenko S O 2015 In vivo terahertz spectroscopy of pigmented skin nevi: pilot study of non-invasive early diagnosis of dysplasia Appl. Phys. Lett. 106 053702
[9] Taylor Z, Singh R, Culjat M, Suen J, Grundfest W, Lee H and Brown E 2008 Reflective terahertz imaging of porcine skin burns Opt. Lett. 33 1258–60
[10] Ashworth P C, Pickwell-MacPherson E, Provenzano E, Pinder S E, Purushotham A D, Pepper M and Wallace V P 2009 Terahertz pulsed spectroscopy of freshly excised human breast cancer Opt. Express 17 12444–54
[11] Cassar Q et al 2018 Pilot study of freshly excised breast tissue response in the 300–600 GHz range Biomed. Opt. Express 9 2930–42
[12] Sung S et al 2018 THz imaging system for in vivo human cornea IEEE Trans. Terahertz Sci. Technol. 8 27–37
[13] Chen A, Osman O, Harris Z, Abazri A, Honkanen R and Arbab M 2020 Investigation of water diffusion dynamics in corneal phantoms using terahertz time-domain spectroscopy Biomed. Opt. Express 11 1264–97
[14] Smolyanskaya O A et al 2018 Terahertz biophotonics as a tool for studies of dielectric and spectral properties of biological tissues and liquids Prog. Quantum Electron. 62 1–77
[15] Popescu D C and Hellicar A D 2010 Point spread function estimation for a terahertz imaging system EURASIP J. Adv. Signal Process. 2010 1
[16] Kazemipour A, Hudlicka M, Dickhoff R, Salhi M, Kleine-Ostmann T and Schrader T 2014 The horn antenna as Gaussian-source in the mm-wave domain J. Infrared, Millim., THz Waves 35 720–31
[17] Ahi K 2017 Mathematical modeling of THz point spread function and simulation of THz imaging systems IEEE Trans. Terahertz Sci. Technol. 7 747–54
[18] Simoens F and Meilhan J 2014 Terahertz real-time imaging uncooled array based on antenna- and cavity-coupled bolometers Phil. Trans. R. Soc. A 372 20130111
[19] Liekhov E, Smolyanskaya O, Popov A, Odlyanitskiy E, Balbekin N and Khodzitsky M 2016 Fabrication and characterization of biotissue-mimicking phantoms in the THz frequency range J. Phys. Conf. Ser. 735 012080
[20] Tewari P, Culjat M O, Taylor Z D, Suen J Y, Burt B O, Lee H, Brown E R, Grundfest W S and Singh R S 2009 Role of collagen in terahertz absorption in skin Advanced Biomedical and Clinical Diagnostic Systems, Proc. of SPIE p 71691A
[21] Wilmink G J et al 2011 Development of a compact terahertz time-domain spectrometer for the measurement of the optical properties of biological tissues J. Biomed. Opt. 16 4
[22] Wang J, Sun Q, Stantchev R I, Chiu T W, Ahuja A T and Pickwell-MacPherson E 2019 In vivo terahertz imaging to evaluate scar treatment strategies: silicone gel sheeting Biomed. Opt. Express 10 3584–90
[23] Witten I H, Frank E and Hall M 2011 Data Mining: Practical Machine Learning Tools and Techniques (Burlington, MA: Morgan Kaufmann Publishers)