# Investigation on the scattering characteristics and unsupervised clustering of 3D printed samples

Mostafa Elsaadouny ⓘ, Jan Barowski, Jochen Jebramcik and Ilona Rolfes

Institute of Microwave Systems, Ruhr University Bochum, Germany

## Abstract

In this work, the scattering characteristics of 3D-printed samples are being investigated by using a single-polarized and a cross-polarized radar system. The 3D-printed technology participates in a wide range of applications nowadays. The idea of synthetic aperture radar (SAR) has been utilized to investigate the reflected electromagnetic energy from the 3D-printed samples by setting each of the radar systems in a fixed position and the mounting sample on an $x$-$y$ positioning table which has been used to achieve rectangular-scan mode for SAR. The data have been ported and processed by the matched filter approach. For better image interpretation, the data have been further processed by the median filter in order to reduce noise level while preserving the main image details. Afterwards, the data have been further investigated for determining and classifying any possible defects. This process has been accomplished by deploying the unsupervised learning concept to cluster the SAR responses into two groups, compromising the defected positions responses and the non-defected responses. The obtained results of both radar sensors have been compared and evaluated using different quality assessment factors. Moreover, unsupervised learning techniques have been investigated and the obtained results show a high degree of efficiency in clustering the SAR responses.

## Introduction

The 3D-printed technology has a wide range of applications such as prototyping and producing complex objects [1], medical purposes [2], industrial applications such as rapid prototyping of RF components, and antenna fabrication of synthetic aperture radar (SAR) applications [3]. Due to the great dependency on this technology during the recent few years, there has been a great interest in investigating these objects to detect defects without destroying them [4]. The idea of imaging presented in this paper is based on SAR processing which is considered as one of the advanced technologies nowadays regarding electromagnetic imaging. It is based on fixing a simple antenna to a moving platform to cover a large distance [5]. Instead of using many antennas at different positions, this configuration allows for the synthesizing of an effective aperture that is much larger than what would be achieved by using stationary antennas. In this research, two radar sensors have been used to exposure 3D-printed samples to investigate its co-polarized and cross-polarized reflection coefficient. Each of the radar sensors has been set in a fixed position concerning the $x$-$y$ positioning table which has been controlled to achieve rectangularly scanned area. The obtained raw data have been then processed using the matched filter technique and further processed by the median filter to attenuate noise level and improve the acquired results [6]. After obtaining the SAR images of the sample layers, these images have been further directed to a clustering process using the concept of the unsupervised learning [7]. The unsupervised learning is a machine learning branch that includes different techniques which allow the model to deal with unlabeled data and discover information based on its own [7]. These techniques analyze the data to find all unknown features and use these features for categorization. Compared to supervised learning, these techniques are considered to be more efficient when dealing with unknown data [7]. This research is organized as follows. In section "Radar systems specifications and image formation", the technical specifications of the utilized radar systems and the image formation algorithm are presented. The experimental setup and the obtained measurement results are illustrated in section "The experimental setup and imaging results", while the unsupervised learning strategies and the clustering results are declared in section "The unsupervised learning and clustering results". Finally, a conclusion is given in section "Conclusion".

## Radar systems specifications and image formation

In this research, two different radar systems have been mounted and used to scan the targets of interest. The radar specifications and the image formation algorithm are presented in this section.

### Technical specifications of radar systems

The first radar sensor used in this work is a 80 GHz radar system with a co-polarized transmit and receive channel shown in Fig. 1. This radar is designed to transmit and receive electromagnetic waves with similar polarization. For an additional evaluation of the cross-polarized scattering factor of the 3D-sample, another radar sensor has also been utilized which is a cross-polarized 80 GHz radar system shown in Fig. 2. This radar system has separate transmit and receive channels, which can be combined using an orthogonal mode transducer. This system provides transmission and reception in different polarization channels. The technical specifications for both radar systems are presented in Table 1, and more information can be found in [9].

### Image formation algorithm

There are different image formation algorithms which have been widely used during recent years such as polar format [10] and back-projection [11]. These two common algorithms have been widely used during the recent decades despite the phase approximation made in the case of the polar format, which makes it less effective, and the computational complexity of the backprojection. In this research, the image has been reconstructed by using the matched filter algorithm [12]. The matched filter response $I$ can be calculated for a received signal $S$ at given time samples $n$ as follows:

$$I = \frac{1}{NK} \sum_{n=1}^{N} \sum_{k=1}^{K} S(f_k, n) e^{+j4\pi f_k \Delta R(n)/c}, \tag{1}$$

where $K$ represents the frequency samples and $\Delta R(n)$ is the differential range given by the following equation

$$\Delta R(\tau_n) = d_{ao} - R_a, \tag{2}$$

where $R_{ao}$ corresponds to the distance from the antenna phase center to the origin and $R_a$ is the distance between the antenna and the target. The main disadvantages of the matched filter arise as the equation (1) has to be applied to each pixel, therefore the process required too much time. To overcome this issue, the data have been processed on a tesla P100 GPU unit from NVIDIA having 3584 CUDA cores, instead of the standard central processing unit due to the great ability of GPUs to process the data in parallel and speed up the overall process [13]. For better image interpretation, the median filter has been used to attenuate the noise level. The median filter is a spatially invariant filter, which replaces the value of the pixel by the median of the neighborhood pixels. This filter is a global filter and has the advantage of smoothing the whole image, therefore keeping image edge information and not losing many details [6].

### The experimental setup and imaging results

#### The experimental setup

The experimental setup consists of the radar systems, the $x$-$y$ positioning table, and the 3D-printed sample. Each radar system has been mounted in a fixed position with respect to the sample under test, which has been placed on the surface plane of the $x$-$y$ positioning table as shown in Fig. 3. The vertical distance between the radar sensor and the sample under test is set to 22 cm. The sample has been placed on a white Rohacell structure.
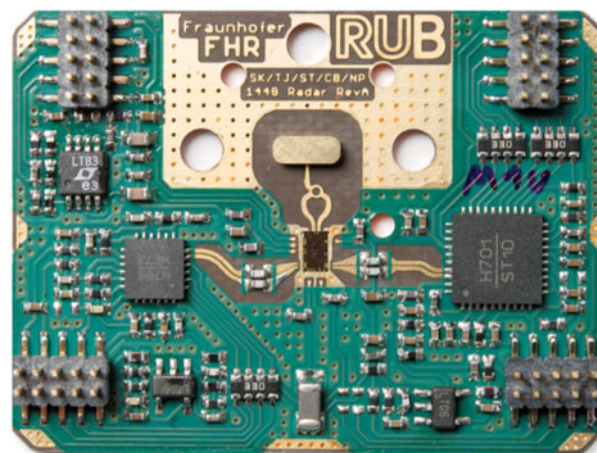


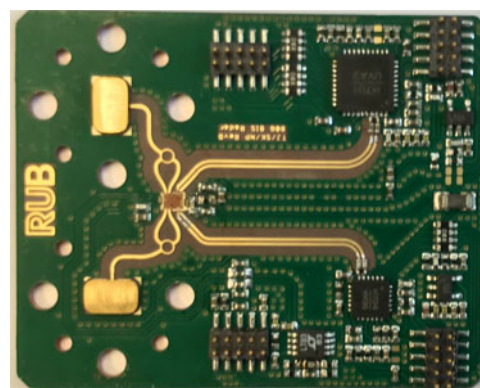**Fig. 1.** The frontend of the co-polarized radar sensor [8].



**Fig. 2.** The frontend of the cross-polarized radar sensor [8].

**Table 1.** Technical specifications of radar sensor

| Radar type | PLL stabilized W-Band FMCW |
|---|---|
| Frequency range | 68–93 GHz |
| Bandwidth | 25 GHz |
| Phase noise | $\leq -80$ dBc/Hz (100 kHz offset) |
| Power consumption | 3.1 W Single channel |

Rohacell is a polymethacrylimide-based structural foam which participates in a wide range of applications. The value of the dielectric constant of Rohacell is 1.04 which is very close to the dielectric constant of space. The control of the movement of the $x$-$y$ positioning table has been set to achieve a step size of 1 mm for each measurement while scanning an appropriate rectangular area which covers the area of the sample placed on Rohacell.

#### The imaging results

The 3D-printed sample has been illuminated by both radar sensors to investigate and differentiate between the co-polarized and cross-polarized reflection factors of this sample. The sample shown in Fig. 4 has been placed on the surface of the positioning
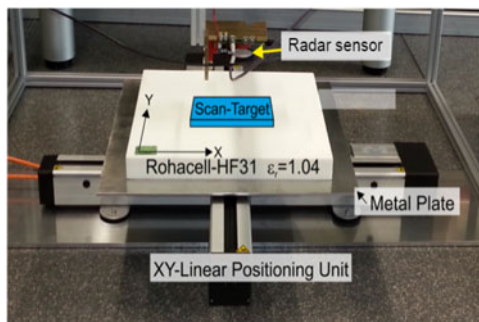
**Fig. 3.** The *x-y* positioning table [8].



**Fig. 4.** The sample under test [8].

**Table 2.** PLA characteristics

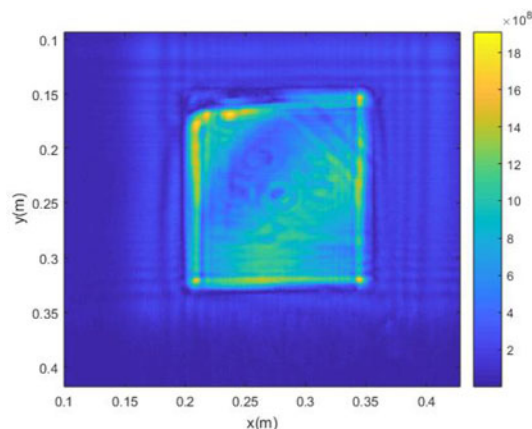| Property | PLA |
|---|---|
| Dielectric constant | 3.2107 |
| Melt temperature | 157–170℃ |
| Tensile strength | 8840–9500 PSI |



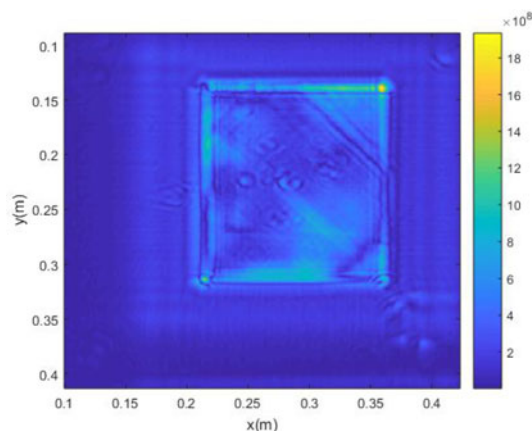**Fig. 5.** The co-polarized imaging result [8].



**Fig. 6.** The cross-polarized imaging result [8].

table and the table has been controlled to achieve rectangular SAR scan wide enough in both *x* and *y* directions. The sample has dimensions of 18 cm by 15 cm and thickness of 0.5 cm. It has been printed using the fused filament fabrication technique. This technique uses continuous filament of thermoplastic materials. In this research, the polylactic acid (PLA) material has been used for fabricating the sample. A review about some of the PLA characteristics is given in Table 2. As the sample has dimensions of 18 cm by 15 cm, the scan has been set to travel 28 cm in both directions, which results in a total number of 78400 measurements, as the movement has been controlled to be 1 mm for each measurement step. The step size plays an important role in the horizontal resolution of the SAR profile.

The effective step size is determined by dividing the wavelength by 4. In these experiments, the step size is set to 1 mm to achieve sufficient horizontal resolution of the SAR profile and detect any possible variations within the area under scan. After finishing the whole scan, the received data matrix has been processed on a GPU by the matched filter technique. For better imaging results, the data have been subjected to a median filter for noise attenuation. The obtained results for both radar sensors are shown in Figs 5 and 6.

According to visual inspection of the obtained results, it can be noted that the sample has sufficient reflection factors for both co-polarized and cross-polarized; however, the co-polarized reflections seem to be higher. Moreover, some small circular shapes can be noticed in both images which are caused by the surface of Rohacell. This surface is not perfectly flat and minor

deviations and space holes exist within it. These deviations have affected the SAR data calculation, and as a result, the images are as shown in Figs 5 and 6. Obviously, the plate, which in the ideal case should not have trans-polarizing properties, has a trans-polarizing character which might be caused by the printing process. For better image interpretation, quality assessment has been investigated by measuring some factors such as peak signal to noise ratio (PSNR), noise variance (NV), and the equivalent number of looks (ENL). PSNR is the most used performance evaluation metric [14]. PSNR can be calculated as follows:

$$PSNR = 10 \log_{10} \frac{A \times A}{MSE}, \tag{3}$$

where $A$ is the maximum value of the image matrix and $MSE$ is the mean square error.

The NV shows the amount of speckle in image and lower NV values indicate better results [14]. The NV can be calculated by:

$$NV = \frac{1}{N} \sum_{i=0}^{N-1} I_i^2 \qquad (4)$$

where $N$ is the size of the image. The ENL is used for estimating the smoothness in homogeneous areas. The high value of ENL shows good quantitative performance [14]. The ENL can be computed as follows:

$$ENL = \left(\frac{\mu}{\sigma}\right)^2, \qquad (5)$$

where $\mu$ and $\sigma$ represent the mean and the standard deviation of the data samples. The SAR image has been divided into segments of [32 × 32] block pixels in a non-overlapping fashion and the ENL has been calculated for each block, then the average has been taken. The obtained results for both radar systems are presented in Table 3. The resultant calculations from the obtained measurements by the co-polarized radar sensor have higher PSNR than that obtained by the cross-polarized radar sensor, which indicates a better quality of image construction. Moreover, the lower values of NV represents a lower amount of noise in the co-polarized measurements. Regarding the ENL, the co-polarized measurements have much higher ENL values than the cross-polarized calculations, and this indicates better quantitative performance. By further examination of the resultant image of the co-polarized radar sensor shown in Fig. 5, it can be depicted that it is not smooth at the left upper corner. This was an expected result as the sample shown in Fig. 4 has some minor deviations at the left upper corner which affected the smoothness and flatness of this part of the sample. This conclusion might be an interesting result regarding non-destructive imaging and faults detection. Investigation of the obtained measurements of both radar sensors by visual inspection and quality assessment calculations shows that the 3D-printed sample has a higher co-polarized scattering factor; therefore, this property can be very beneficial regarding the fabrication of the co-polarized systems, such as fabricating the dielectric lenses of the RF antennas. However, there is a significant unexpected cross-polarized scattering factor which should be very low in the ideal case compared to the co-polarized factor. The main possible reason for this result is the effect of the printing process. In these measurements, the printing has been accomplished using a resolution of 0.1 mm and the object has been printed using PLA. In the next sections, different 3D printed samples have been tested and the printing effect could be further evaluated by inspecting the results.

## The unsupervised learning and clustering results

Machine learning aims to understand the structure of the given data and fit it into understandable models [15]. This field is continually developed and participates in many applications including speech recognition [16] and data classification and regression [17]. The common machine learning methods are the supervised and unsupervised learning [7, 18]. In the supervised method, the learning process is accomplished by providing the machine with input data and associated labels for these data [18]. The algorithm then compares its output with the correct outputs to find errors, and modify the model accordingly. After finishing the training phase and reducing the error to the lowest

**Table 3.** The quality assessment factors for performance evaluation

| Quality assessment factor | Co-polarized | Cross-polarized |
| --- | --- | --- |
| PSNR | 31.956 | 28.464 |
| NV | 3.2083 | 4.1022 |
| ENL | 42.297 | 31.147 |

possible value, the trained algorithm can then be used to predict label values of additional unlabeled data. Unlike supervised learning, unsupervised learning focuses on discovering the hidden patterns within the dataset without any interference and supervision from humans. This means that the unsupervised algorithms use unlabeled datasets, which are complex and unrelated, and organize them in meaningful ways [7]. Unsupervised learning includes various techniques and some of these techniques are presented in the next sections and their performances have been evaluated for the NDT of printed samples.

### K-means algorithm

The *k*-means is considered as a very popular and important unsupervised machine learning algorithm [19]. This method aims to cluster the data points into different groups by defining the initial centroid value by following an iterative process which finds the highest value for each iteration. The initial center value becomes the basis for the clustering process. To summarize the *k*-means process, we start first by selecting random centroids, as starting points for our clusters, and then perform iterative calculations to adjust the position of these centroids. This step can be done by calculating the Euclidean distance between the observed data points and the cluster centroid, as follows:

$$d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}. \qquad (6)$$

At the beginning of the process, the centroids tend to have a big change every iteration and assigning the data points to the clusters depends on the value of the Euclidean distance. After reaching the optimum solution, the centroids tend to have a slight or even no change of their positions which means that they have stabilized [19]. For investigating the performance of the *k*-means, further 3D printed samples have been fabricated with minor defects. These designed samples have dimensions of 7 cm × 7 cm and a thickness of 1 cm. Some of these samples have been designed with space holes of diameter 6 mm centered at the middle layer, while others have been designed with cones or squares placed above their surface as shown in Fig. 7. Each of these samples has been placed on the *x-y* positioning table and scanned using our co-polarized radar sensor. The resultant SAR data have been processed and Fig. 8 shows an example of a defected sample where a space hole was placed in the middle layer of the sample. The resultant images of the different layers have been prepared and divided into small sections, each of size 25 × 25 pixels. Some of these sections correspond to the locations of the defects, while the remaining represent the remaining non-defected parts of the sample. These resultant sections are not labeled and have been used to force the algorithm into detecting hidden features and clustering these sections into different groups. After the data have been prepared, the *k*-means algorithm has been
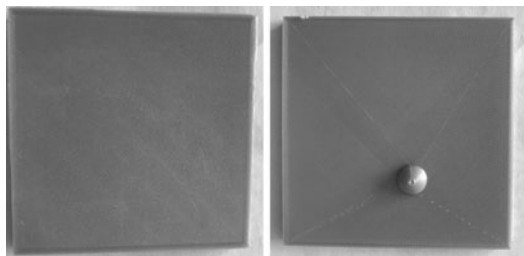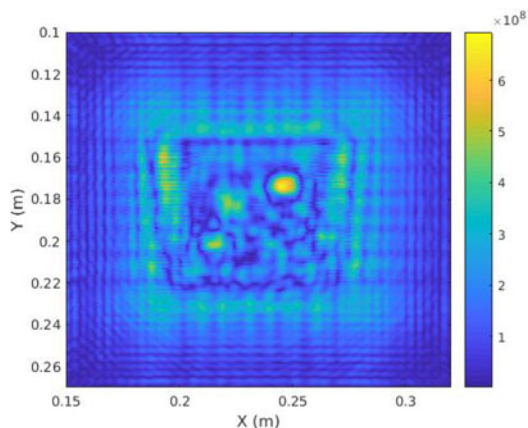
**Fig. 7.** Examples of the 3D printed samples.



**Fig. 9.** *K*-means clustering result using two centroids.



**Fig. 8.** The resultant SAR image of the middle layer.



**Fig. 10.** The elbow result.

initialized with $k = 2$, and the associated result is shown in Fig. 9. In this problem, we perform clustering to different data samples which belong to two groups representing the defected and non-defected samples. However, the clustering algorithm usually does not know any information about the labels of the dataset, so the algorithm has to be evaluated and investigated to check whether it could detect the optimized number of centroids. To achieve this purpose, the *k*-means plot has been calculated for different values of *k* starting from 2, and then evaluate the process by incorporating the elbow method [19, 20]. The idea of this method is to run the *k*-means algorithm using a range of values of *k* and calculate the sum of the squared errors (SSE) for each value. For several cluster centroids $C_i$ and data $X_i$, the *SSE* can be calculated as follows:

$$SSE = \sum_{k=1}^{K} \sum_{x_i \in s_k} \| X_i - C_k \|^2 \qquad (7)$$

After calculating the *SSE*, we start searching for the value of *k* which corresponds to a very low value of SSE, but usually SSE tends to 0 as we are increasing *k* till it equals the number of data points, so we select the value of *k* where SSE starts to flatten out and forming an elbow. As shown in Fig. 10, it can be included that $k = 3$ is a good choice; however, the elbow method does not always provide accurate results, so another clustering metric has to be checked to confirm the obtained result. To give intuition about *k*, the silhouette analysis has been implemented [21]. This method is used to determine the degree of separation between clusters [21]. This method is generally implemented by calculating the average distance from all data points in the same cluster $d_i$ and
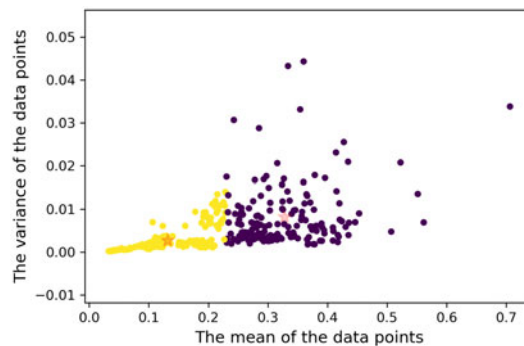


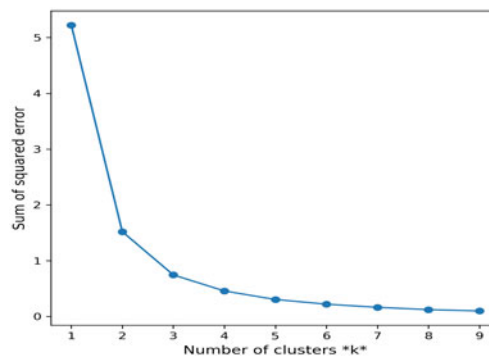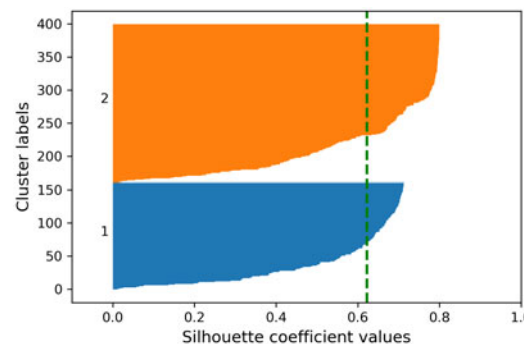**Fig. 11.** Silhouette result for $k = 2$.

in the closest cluster $c_i$. Afterwards, these two factors are used to calculate a coefficient *S* as follows:

$$S = \frac{d_i - c_i}{\max{(d_i, c_i)}} \qquad (8)$$

The values of this coefficient lie within the range $[-1, 1]$, and the clusters are considered good if the value of this coefficient is high or close to 1. The silhouette coefficient has been calculated for *k* within the range $[2, 4]$, and according to the results shown in Figs 11–13, we can conclude that the optimum number of clusters is 2. This decision differs from the elbow result; however, it
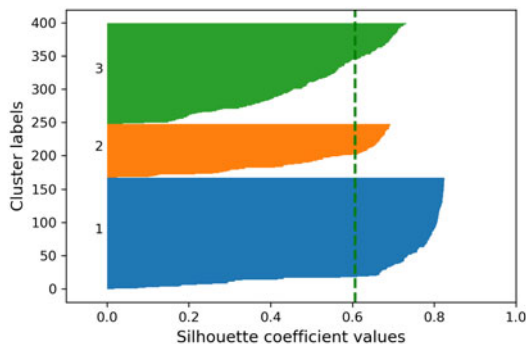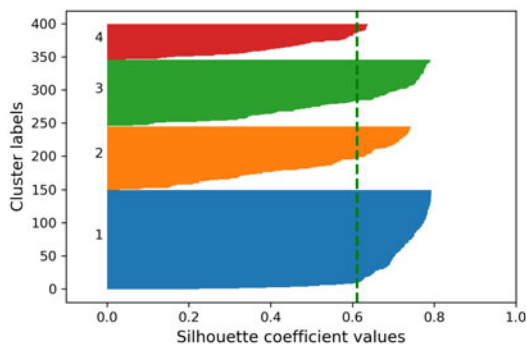
**Fig. 12.** Silhouette result for $k = 3$.



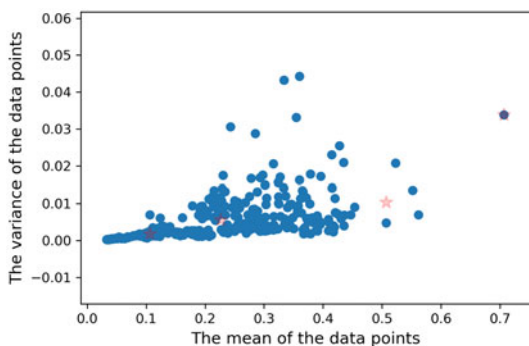**Fig. 13.** Silhouette result for $k = 4$.



**Fig. 14.** Mean-shift clustering result.

seems to be the optimum selection as the silhouette analysis is more reliable and accurate compared to the elbow method.

### Mean-shift clustering

Mean-shift (MS) is a clustering algorithm that assigns the data points to the clusters by repetitively shifting samples toward the highest density of the samples. MS has applications in the field of image processing and computer vision [22]. Given a set of samples, the algorithm iteratively assigns each datapoint toward the closest cluster centroid. The direction to the closest cluster centroid is determined by the location of the closest group of points. Therefore, each data point will move closer during each iteration to the cluster center. When the algorithm stops, each point is assigned to a cluster. Unlike the K-Means algorithm, MS does
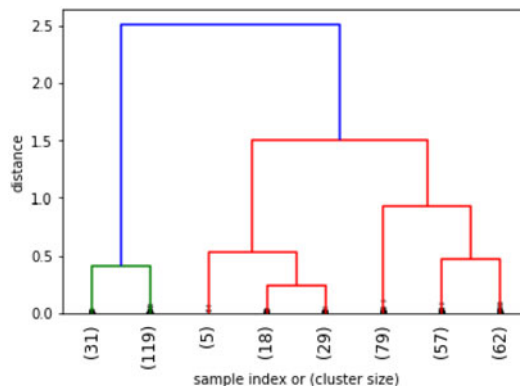


**Fig. 15.** Dendrogram result.

not require specifying the number of clusters in advance. The number of clusters is determined by the algorithm concerning the data. According to the result shown in Fig. 14, it can be concluded that the algorithm has assumed the existence of four clusters as marked by the four cluster centers. This result contradicts the result of the $K$-means algorithm; therefore, further investigation needs to be accomplished to provide more certainty about the data clusters, as presented by the next method.

### Hierarchical clustering

This method provides a stable and efficient estimation of the number of clusters. It includes two major types, known as the agglomerative and the divisive clustering [23]. These techniques are opposite to each other. In the agglomerative method, which is implemented in this research, each point is considered as a single cluster and then compared to the adjacent points, and similar clusters merge till reaching the optimum number of clusters. On the other hand, the decisive method performs the opposite operation. In this method, the whole data points are considered as a single cluster and then, this cluster is divided into different clusters by separating the data points which are not similar. To calculate the proximity of the data points, different techniques have been proposed for this purpose. The authors have focused on the ward's method. In this method, the proximity of the data points is obtained by calculating the sum of the square of the distances between the data points $X_i$ of the first cluster $C_1$ and the data points $X_j$ of the second cluster $C_2$. This process can be expressed mathematically as follows:

$$P(C_1, C_2) = \frac{1}{m_1 m_2} \sum_{i=1}^{m_1} \sum_{j=1}^{m_2} D(X_i, X_j)^2 \qquad (9)$$

where $m_1$ and $m_2$ represent the total number of the data points in the adjacent clusters $C_1$ and $C_2$. This operation is calculated for each adjacent clusters, and the similar clusters are merged. The output of this method can be graphically represented using the dendrogram, as shown in Fig. 15. The main purpose of the dendrogram is to precisely allocate objects to clusters. The interpretation of this method is accomplished by focusing on the height at which the objects are combined. In Fig. 15, the dendrogram indicates a big difference between the group at the left, which consists of 150 of our samples, and the combined remaining samples. This leads to conclude the existence of two clusters as confirmed by the $K$-means method.

## Conclusion

In this paper, the scattering characteristics of a 3D-printed sample have been investigated. The obtained measurements have been processed by using the matched filter, and for further improvement and noise removal, the data have been processed by the median filter for better image interpretation. The obtained results have been investigated by different quality assessment factors such as PSNR, NV, and ENL. The obtained quality calculations show better results of the co-polarized reflected energy which is much better than those of the cross-polarized scattered energy as noticed by visual inspection. These properties can be very beneficial while using this technology in designing different hardware components. For further investigation and testing purposes, multiple printed samples have been scanned and their resultant SAR images have been further processed using various clustering algorithms. The obtained results have been compared to emphasize the functionality of these algorithms in clustering the dataset. The *k*-means and hierarchical algorithms have provided the most accurate results, while the MS algorithms failed to obtain the optimum number of clusters. The results of the former methods have been used to facilitate the non-destructive testing of samples. This has been done by clustering the defected and the non-defected samples into different groups and optimizing the number of the groups using different evaluation metrics, including the elbow method and the silhouette analysis.

## References

1. **Wohlers T** (2013) *Additive Manufacturing and 3D Printing State of the Industry*. Fort Collins, CO: Wohlers Associates.
2. **Ryzhkov A, Matrosov SY, Melnikov V, Zrnic D, Zhang P, Cao Q, Knight M, Simmer C and Troemel S** (2015) 3D printed microfluidics for biological applications. *Lab Chip* **15**, 3627–3637.
3. **Otter WJ and Lucyszyn S** (2016) 3-D printing of microwave components for 21st century applications. In *2016 IEEE MTT-S International Microwave Workshop Series on Advanced Material and Processes for RF and THz Applications (IMWSAMP)*, Chengdu, 2016, pp. 1–3, doi: 10.1109/IMWS-AMP.2016.7588327.
4. **Becquaert M, Cristofani E, Pandey G, Vandewal M, Stiens J and Deligiannis N** (2016) Compressed sensing mm-wave SAR for non-destructive testing applications using side information. In *2016 IEEE Radar Conference*, Philadelphia, PA, 2016, pp. 1–5, doi: 10.1109/RADAR.2016.7485244.
5. **Solimene R, Catapano I, Gennarelli G, Cuccaro A, Dell'Aversano A and Soldovieri F** (2014) SAR imaging algorithms and some unconventional applications: a unified mathematical overview. *Signal Processing Magazine IEEE* **31**, 90–98.
6. **Zhu J, Wen J and Zhang Y** (2013) A new algorithm for SAR image despeckling using an enhanced Lee filter and median filter. In *6th International Congress on Image and Signal Processing (CISP)*, Hangzhou, 2013, pp. 224–228, doi: 10.1109/CISP.2013.6743991.
7. **Dike HU, Zhou Y, Deveerasetty KK and Wu Q** (2018) Unsupervised learning based on artificial neural network: a review. In *Proceedings of the IEEE International Conference on Cyborgs and Bionic Systems (CBS)*, Shenzhen, 2018, pp. 322–327, doi: 10.1109/CBS.2018.8612259.
8. **Elsaadouny M, Barowski J, Jebramcik J and Rolfes I** (2019) Investigation on scattering characteristics of a 3D-printed sample based on SAR processing. In *2019 European Microwave Conference in Central Europe (EuMCE)*, Prague, Czech Republic, pp. 273–276.
9. **Jaeschke T, Bredendiek C, Kueppers S, Schulz C, Baer C and Pohl N** (2016) Cross-polarized multi-channel W-band radar for turbulent flow velocity measurements. In *IEEE International Microwave Symposium 2016*, San Francisco, CA, USA, pp. 1–4, doi: 10.1109/MWSYM.2016.7540256.
10. **Linchen Z, Jindong Z and Daiyin Z** (2013) FPGA implementation of polar format algorithm for airborne spotlight SAR processing. In *2013 IEEE 11th International Conference on Dependable, Autonomic and Secure Computing (DASC)*, Chengdu, 2013, pp. 143–147, doi: 10.1109/DASC.2013.52.
11. **Song X and Yu W** (2017) Processing video-SAR data with the fast back-projection method. *IEEE Transactions on Aerospace and Electronic Systems* **52**, 2838–2848.
12. **Gorham LA and Moore LJ** (2010) SAR image formation toolbox for MATLAB. *Proceedings of the SPIE 7699*, Volume 7699, id. 769906, April.
13. **Zhang F, Wang B-n and Xiang M-s** (2010) Accelerating InSAR raw data simulation on GPU using CUDA. In *2010 IEEE International Geoscience and Remote Sensing Symposium*, Honolulu, HI, USA, pp. 2932–2935, doi: 10.1109/IGARSS.2010.5650737.
14. **Argenti F, Lapini A and Alparone L** (2013) A tutorial on speckle reduction in synthetic aperture radar images. *IEEE Geoscience Remote Sensing Magazine* **1**, 6–35.
15. **Loussaief S and Abdelkrim A** (2016) Machine learning framework for image classification. In *7th International Conference on Sciences of Electronics, Technologies of Information and Telecommunications (SETIT)*, Hammamet, 2016, pp. 58–61, doi: 10.1109/SETIT.2016.7939841.
16. **Zhang Y, Li P, Jin Y and Choe Y** (2015) A digital liquid state machine with biologically inspired learning and its application to speech recognition. *IEEE Transactions on Neural Networks and Learning Systems* **26**, 2635–2649.
17. **Ji S, Xu W, Yang M and Yu K** (2013) 3D convolutional neural networks for human action recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **35**, 221–231.
18. **Shin H, Roth HR, Gao M, Lu L, Xu Z, Nogues I, Yao J, Mollura D and Summers RM** (2016) Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. *IEEE Transactions on Medical Imaging* **35**, 1285–1298.
19. **Syakur M, Khotimah B, Rochman E and Satoto B** (2018) Integration k-means clustering method and elbow method for identification of the best customer profile cluster. In *IOP Conference Series: Materials Science and Engineering*, vol. 336.
20. **Marutho D, Hendra Handaka S, Wijaya E and Muljono** (2018) The determination of cluster number at k-mean using elbow method and purity evaluation on headline news. In *2018 International Seminar on Application for Technology of Information and Communication*, Semarang.
21. **Gupta T and Panda SP** (2019) Clustering validation of CLARA and k-means using silhouette & DUNN measures on Iris dataset. In *2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon)*, Faridabad, India, pp. 10–13, doi: 10.1109/COMITCon.2019.8862199.
22. **Carreira-Perpinn MA** (2015) A review of mean-shift algorithms for clustering. arXiv preprint arXiv:1503.00687.
23. **Nazari Z, Kang D, Asharif MR, Sung Y and Ogawa S** (2015) A new hierarchical clustering algorithm. In *2015 International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS)*, Okinawa, pp. 148–152, doi: 10.1109/ICIIBMS.2015.7439517.

**Mostafa Elsaadouny** received his B.Sc. and M.Sc. degrees in electrical engineering from the Military Technical Colleague, Cairo, Egypt in 2010 and 2016, respectively. Since 2012, he worked as a Research Assistant at the Radar and Communication Department at the Military Technical Colleague. In 2017, he joined the institute of the High Frequency Systems at the Ruhr-University Bochum to complete his Ph.D. degree. His main research interests are radar signal and image processing and the design and optimization of machine learning algorithms for images classification.

**Jan Barowski** received the B.Sc. and M.Sc. degrees in electrical engineering from Ruhr-University Bochum, Bochum, Germany, in 2010 and 2012, respectively. Since 2012 he is with the Institute of Microwave Systems, headed by Ilona Rolfes, Ruhr-University Bochum, as a Research Assistant. In 2017, he received the Dr.-Ing. degree in electrical engineering from Ruhr-University Bochum and is now working as a Post-doctoral Research Scientist at the Institute of Microwave Systems. His current fields of research are concerned with radar signal processing, radar imaging, and material characterization techniques.

**Jochen Jebramcik** was born in Essen, Germany, in 1987. He received the M.Sc. degree in electrical engineering and information technology from Ruhr-University Bochum, Bochum, in 2014, where he is currently pursuing the Ph.D. degree in electrical engineering. Since 2014, he has been with the Institute of Microwave Systems, Ruhr-University Bochum, as a Research Assistant. His current research interests include simulation techniques, material characterization, and radar signal processing.

**Ilona Rolfes** received the Dipl.-Ing. and Dr.-Ing. degrees in electrical engineering from Ruhr-University Bochum, Bochum, Germany, in 1997 and 2002, respectively. From 1997 to 2005, she was with the High Frequency Measurements Research Group, Ruhr-University Bochum, as a Research Assistant. From 2005 to 2009, she was a Junior Professor with the Department of Electrical Engineering, Leibniz Universitaet Hannover, Hannover, Germany, where she became the Head of the Institute of Radio frequency and Microwave Engineering in 2006. Since 2010, she has led the Institute of Microwave Systems, Ruhr-University Bochum. Her fields of research concern high-frequency measurement methods for vector network analysis, material characterization, noise characterization of microwave devices, sensor principles for radar systems, and wireless solutions for communication systems.