Visual Parameter Space Analysis for Optimizing the Quality of Industrial Nonwovens

This paper was downloaded from TechRxiv (<https://www.techrxiv.org>).

LICENSE

CC BY 4.0

SUBMISSION DATE / POSTED DATE

29-09-2021 / 30-09-2021

CITATION

Victor, Viny Saajan; Schmeisser, Andre; Leitte, Heike; Gramsch, Simone (2021): Visual Parameter Space Analysis for Optimizing the Quality of Industrial Nonwovens. TechRxiv. Preprint.
<https://doi.org/10.36227/techrxiv.16701841.v1>

DOI

[10.36227/techrxiv.16701841.v1](https://doi.org/10.36227/techrxiv.16701841.v1)

Visual Parameter Space Analysis for Optimizing the Quality of Industrial Nonwovens

Viny Saajan Victor
Fraunhofer ITWM

Dr. Andre Schmeißer
Fraunhofer Center for Machine Learning and Fraunhofer ITWM

Prof. Dr. Heike Leitte
Technical University Kaiserslautern

Prof. Dr. Simone Gramsch
Fraunhofer ITWM and Frankfurt University of Applied Sciences

Abstract—The quality of the technical textiles, in particular nonwovens, depends on the process parameters involved in producing them. It is not practical to investigate the effect of these parameters in real-time. Hence, a digital twin has been developed to analyze nonwoven production processes. The quality of the virtually produced nonwovens from the digital twin is then analyzed and mapped back to the process parameters. Different visualization techniques can be incorporated to guide engineers in finding optimal combinations of these process parameters for obtaining the desired product quality. This would eliminate the traditional trial and error strategies and reduce the domain expertise required in such an analysis. In this paper, we present a visual analytic tool that aids the engineer in analyzing and optimizing the quality of the nonwovens, thereby helping to make beneficial decisions in the industrial production process.

■ INTRODUCTION

Nonwoven products become more and more important in our daily lives. According to EDANA, the European trade association for the nonwovens industry, the annual average growth rate of nonwoven production in the last decade was 4%. In 2020, the nonwoven production growth in Greater Europe was even 7.2% reaching 85.9 billion square metres of nonwovens (see [1]). One obvious reason for the significant increase in production in 2020 is the increased demand for medical products such as medical masks

(FFP2, FFP3) or medical protective clothing.

Medical masks consist of several layers of nonwovens. The outer layer are produced by so-called spunbond processes. In a spunbond process a polymer is extruded to thousands of very thin fibers that are twirled by air and stochastically laid down to a conveyor belt. Finally, the nonwoven fabric is calendered, pleated and seamed (see [2] for details of the manufacturing process).

Controlling the spunbond production process is a big challenge for industry due to its stochastic nature. Typically, the manufacturing takes place

Figure 1. Flow chart showing the incorporation of visual analysis in optimizing the quality of industrial nonwovens.

24/7 which limits longer experimental series. In practice, manufacturers aim for high belt speeds. The higher the belt speed, the greater the production capacity, but the worse the homogeneity of the produced nonwoven fabric. In particular, the filtration properties, but also mechanical properties like tensile strength, suffer with higher belt speeds. Therefore, all process parameters must be adjusted to each other. The challenge for engineers is that is that we have a multidimensional space of process parameters that affect product quality. We are therefore presenting for the first time a visualization tool designed to support the textile industry in order to optimize the nonwoven homogeneity under certain process conditions. Hereby, the visualization tool is based on a digital twin for spunbond processes and leads the designer to optimal settings of process parameters for homogeneous nonwovens.

Problem Setting

We construct our visualization tool on a digital twin for spunbond processes. Using a digital twin has the advantage that the influence of varying process parameters on product quality can be systematically analyzed in an ideal machine. As depicted in Figure 1, we collect simulation data of the digital twin. Since running simulations with the digital twin is very time consuming, we train a machine learning model mapping the process parameters to a measure of the product quality. Finally, we establish multiple visual parameter space analysis techniques for optimizing the nonwoven production process. In the fol-

lowing, we briefly summarize the mathematical model of digital twin. The input features (process parameters) and output features (product quality) are explained in more detail. Additionally, the training of the machine learning model is briefly outlined.

Digital twin of nonwoven fabrics: SURRO

The digital twin for spunbond processes is split into two parts according to the twirling of the fibers in the air and to the stochastic fiber laydown. Here, we focus on the superposition of the fibers on the belt forming the nonwoven fabric. An overview about the mathematical theory can be found in [3], while the simulation software is presented in [4]. Due to the stochastic nature of the fiber laydown, a single fiber on the belt is simulated by a surrogate model based on a system of stochastic differential equations as described in [5]:

$$\begin{aligned} d\xi_1 &= (\cos(\alpha_s) + v) ds, \\ d\xi_2 &= \sin(\alpha_s) ds, \\ d\alpha_s &= \left(\frac{\xi_1}{\sigma_1^2} \sin(\alpha_s) - \frac{\xi_2}{\sigma_2^2} \cos(\alpha_s) \right) ds + AdW_s. \end{aligned}$$

Hereby, ξ_1 describes the fluctuations of the fiber laydown in machine direction, i.e., in the direction of the moving conveyor belt, while ξ_2 goes in cross machine direction. The third equation expresses the stochastic buckling behavior of the fiber laydown. With the angle α_s we denote the change in direction. The surrogate model parameters σ_1 , σ_2 , A , and v correspond to process parameters and are explained in the next paragraph. In order to simulate the full nonwoven fabric multiple fibers are superposed according to the spinning positions of the real plant. Please note that we abbreviate this model as SURRO.

Input Parameters

The input parameters σ_1 , σ_2 , A , and v are extended by a fifth parameter n representing the number of fibers to be superposed. The range of input parameters is based on typical application scenarios for spunbond products and also includes the production of the medical protective masks.

- σ_1 is the standard deviation of the 2D normal distribution of the fibers around the spinning nozzle in the machine direction parallel to the

Figure 2. Input Parameters and Output Feature used for the Visual Parameter Space Analysis. Left: Shows the standard deviations along the machine and cross directions. Middle: Visualizes the effect of parameter A on the nonwovens. Right: Output images of the produced nonwoven for two different resolutions.

Figure 3. Parameter Tuner with sliders for Input Features (left) and corresponding output values (right)

running conveyor belt without belt movement (see Fig. 2). The feature values in our database range from 1 mm to 50 mm.

- σ_2 is the standard deviation of the 2D normal distribution of the fibers around the spinning nozzle in the cross direction without belt movement (see Fig. 2). The feature values range from 1 mm to 50 mm.
- A is the feature that contains all random effects of the production process, e.g., the influence of the turbulent flow during the fiber spinning, fiber-fiber contacts, and laydown. This feature describes how deterministic (corresponding to value 0) or stochastic (value ∞) the simulated fiber lays down. The feature values range from

1 to 50.

- v is the ratio of spinning speed and belt speed. The feature values range from 0.01 to 0.25.
- n is the number of spin positions per meter, which is a fixed value during production, but can be varied at the design time of the machine. The feature values range from 200 to 10 000.

Output Feature

Using the given database of input features, we have performed simulations producing virtual nonwovens with a fully resolved nonwoven microstructure. To assess the quality of the resulting nonwovens, we then discretize the data into a

raster image, where each pixel value maps to the fiber mass in the corresponding cell. To quantify inhomogeneities at different scales, we perform this at seven different resolutions. Finally, the output corresponds to the coefficient of variation (CV) values of the image matrix at the different resolutions obtained from the simulation tool for each row of input feature.

The coefficient of variation (CV) is a statistical measure of the relative dispersion of data points in a data series around the mean μ . With the standard deviation σ it is defined as

$$CV = \frac{\sigma}{\mu}.$$

It determines the homogeneity of the nonwoven web. A lower CV value generally corresponds to a more homogeneous nonwoven, which can affect properties like filter quality and tensile strength.

Machine Learning to aid the visual analysis

The mapping from input parameters to the output quality using our simulation requires much computation time and resources. To speed up the process, we can utilize the stochastic parameters of the simulation to train a machine learning model that predicts the final product quality. Then, the predictions can be used by our visual analytics tool for faster computation. This approach is computationally efficient and aids in analyzing the influence of the input parameters on the product quality using VSPA.

Our input data contains five continuous input features, which are used to predict seven continuous output values. Hence we used multiple-input, multiple-output regression as our machine learning method. In recent times, extensive research is being done in machine learning which has led to several learning algorithms. To determine which learning algorithm best fits our data, we evaluated linear regression [6], polynomial regression [7], random forests [8], Bayesian regression [9], and neural networks [10]. The model with the least mean absolute percentage error [11] was chosen for the visual analysis.

Theoretical Background

Parameter space analysis (PSA) is the systematic variation of model input parameters, generating outputs for each combination of parameters,

Figure 4. 3D Surface Plot showing CV values for full range of σ_1 and σ_2 while other parameters kept constant ($A=1$, $v=0.01$, $n=200$) for the resolution 50mm showing the local minimum

and investigating the relation between parameter settings and corresponding outputs [12]. A typical motive of the parameter space analysis in the industrial production process is to find a desirable combination of process parameters that optimize product quality. Traditional approaches involve trial and error methods where domain experts set specific input values and run the process. If the resulting output is not desired, the process is repeated by changing the input values until the desired quality of the product is obtained. This approach is time-consuming and requires domain expertise. Incorporating interactive visualization into this process is known as "Visual Parameter Space Analysis (VPSA)" [12]. This approach would significantly reduce both the time involved in the analysis and the required domain expertise, while enabling the engineer to be visually guided towards the desired optimal quality of the product. Sedlmair et al. [12] provide a conceptual framework for VPSA, independent of the application domain. The visualization approaches inspired from the above framework that are used in our application tool can be categorized into navigation strategies and analysis techniques. Relevant for our setting are the following methods.

Navigation Strategies Used for VPSPA

The data which maps the input parameters to the output quality needs to be presented to the engineer for exploratory analysis. We classified three distinctive strategies of how this data was made available for navigation.

- **Informed Trial and Error:** If the computation of the output values takes a short amount of time, the engineer can be provided with an option to try out different combinations of parameter settings for the analysis. We use the machine learning model for the predictions to obtain the results in real-time.
- **Local to Global Navigation:** The local-to-global strategy starts with displaying one specific output and letting the engineer explore alternatives from there. The output for a specific input parameter setting is shown, and the engineers can interactively change the input parameters from that setting, thereby updating the output.
- **Global to Local Navigation:** In this strategy, the engineer is first presented an overview of the whole range of data, from where zoom and filter options can be used to show details on demand.

Analysis Techniques Used for VSPA

Once we have characterized how to present data to the engineer for navigation, the next important goal is to understand the tasks engineers want to engage in when doing visual parameter space analysis. After understanding the requirement, we utilized the following analysis techniques to guide the design and engineering processes.

- **Data Partition:** In this strategy, the output data space is partitioned into clusters to find different types of model behaviors. These behaviors are then mapped back to the input parameter settings to analyze their influence on the input data space.
- **Sensitivity Analysis:** Sensitivity is termed as the uncertainty of the input parameter value. Thus, the variation of outputs is observed against the change in inputs. This analysis helps distinguish optimal parameter settings and the non-sensitive parameter settings, as explained in the later sections.

- **Handling statistical uncertainty:** Statistical uncertainty indicates the difference in the output for multiple non-deterministic runs of the model while keeping the parameter setting constant. Thus, if we have higher statistical uncertainty, it needs to be reduced for our model to be trustworthy.

Visual Parameter Space Analysis

This chapter explains details about the interactive visualization tool built to assist the textile engineers in parameter space analysis. We provide them with four primary parameter space exploration analyses, which also help them understand the effects of the input parameters on the product quality.

Real-time Navigation through the Parameter Space

We provide a parameter tuner for the engineers to navigate through the parameter space in real-time, see Figure 3. Engineers can analyze the output for all resolutions, compare two or more parameter settings against the output and analyze the output for individual resolutions separately.

In practice, manufacturers are interested in a nonwoven with a specific base weight, as this is the most important sales feature. Filtering the database by base weight is therefore a must-have in order to find optimal process conditions for the targeted products.

Analyzing the influence of process parameters

We designed an interactive 3D surface plot to provide engineers a better understanding of the influence of the input parameters by freezing a combination of two parameters, see Figure 4 for an example. Although the engineers are not allowed to change these parameters, they act as free parameters, while the other three parameters that are allowed to be varied act as frozen parameters for generating the data samples.

For the construction, one hundred uniform samples are generated for the two free parameters in the allowed range of these parameters. The combination of the free parameters produces 10 000 samples. The three frozen parameters are appended to each of the above combinations to form a data set with 10 000 input rows. In the 3D surface graph, the output for the selected

resolution along the z-axis is plotted with the free parameter values along the x and y axes. The 3D Surface for the free parameters σ_1 and σ_2 is shown in Figure 4.

The 3D Surface graph helps the engineers to select the parameter setting of their interest and lets them explore alternatives from here. It also allows them to observe the influence of the free parameters with respect to the output. This type of analysis is called local to global navigation. The 3D surface graph also helps the engineers to visually identify the local minima and maxima.

The engineers are provided with an option to select a sub-range for the free parameters which lie in the specific region of interest on the surface of the graph. Upon selecting this sub-range, the graph will be updated with new values for the given sub-range of input parameters. This shows a detailed on demand analysis of the desired range of values.

Sensitivity Analysis and Visual Guidance for reaching Optimal and Non-Sensitive Parameter Settings

Sensitivity, in our context, pertains to the robustness of the output values against small changes in the input parameter settings. To examine this effect, we perform a sensitivity analysis ([13]).

The sensitivity analysis contains three major components, the input space graph, the output space graph and the parallel plot (cf. Figure 5). The input space graph allows the engineers to select areas of interest, for which the corresponding prediction values are plotted in the output space graph and the parallel plot.

For the construction, we create multiple Latin hypercube samples [14] with dimension five (number of input features) which are subjected to dimensionality reduction via Principle Component Analysis [15]. The first two principal components with the highest variance are chosen to transform the data into two dimensions. Then, the predictions of these inputs are shown in the same way in the output space graph. Finally, the parallel plot shows the individual input and output values for the selected data point.

For the sensitivity analysis, the selected input parameter for the analysis is called the focal point. The steps to carry out the analysis are as

follows:

- **Selection of the focal point:** The engineers select the data point of their interest either from the input space graph (by clicking on it) or entering it manually. The selected focal point gets highlighted on all three components.
- **Selection of area and population of the local neighborhood around the focal point:** The engineers can generate N data points around the focal point with a radius r . Here we set $N = 50$ samples, and r equal to 10% of the feature range.
- **Updating the local neighborhood:** The two types of updates provided to the engineers are global and local updates. In the global update, the neighborhood of the focal point is populated via uniform sampling by altering all the input features. In the local update, only the user-selected input features are altered for populating the neighborhood.
- **Apply the changes:** The three components are updated with the new neighborhood values around the focal point.

Additionally, the tool visually guides the engineer to find the optimal solution. The engineers start at a specific parameter setting of their interest and generate points around it. They then select an interesting point around this new neighborhood as the next focal point. The process is repeated until they find the parameter setting of their interest.

Patterns, Anomalies and Outlier Examination using Clustering

Clustering [16] is an unsupervised machine learning algorithm that involves the automatic discovery of natural grouping in data. Clustering algorithms only interpret the data and find natural groups or clusters in the feature space.

The goal of Cluster Analysis in our context is to find possible clusters in the output space to capture different types of behaviors. Then, these behaviors are analyzed by mapping the clusters found in the output space to the corresponding range of input space parameters. This analysis includes the following steps:

- **Partition of output space:** In this step, the output data points (CV values at all resolutions) are clustered based on two unsupervised machine learning algorithms. First, Principal

Figure 5. Sensitivity Analysis showing the different components. (A) Input Space Graph, (B) Output Space Graph, (C) Parallel Plot, and (D) User Input Interface.

Figure 6. (Left) Shows the five clusters obtained by partitioning the output dataspace, (middle) Average CV values of the selected cluster for all the resolutions, (right) The distribution of the feature "SpinPositionsPerMeter" responsible for the corresponding cluster.

Component Analysis (PCA) is performed on the entire output data to reduce dimensions from seven to two. The selected two principal components explain more than 99% of the variance in the data, and hence we will not lose much information by selecting these two components. These principal components are a linear combination of original dimensions. Then, we use Gaussian Mixture Models (GMM) clustering [17] to cluster the reduced output space. The more popular K-means clustering [18] was not chosen for this analysis as it works only if the clusters are circular and hence will not consider the variance in our data. Thereby, we chose GMM Clustering, which considers the variance in the data and handles clusters of different shapes. The optimal number of clusters was chosen using Bayesian information criterion [19]. The Figure 6 (left) shows the five clusters obtained in our analysis.

- **Assign a quality measure to the partitioned**

data: The five clusters obtained while partitioning the output space corresponds to five major behaviors. A quality measure is assigned to each of the clusters to differentiate them. We chose average CV values in all seven resolutions as the quality measure for each cluster as shown in Figure 6 (middle). The clusters with lower average CV values are desired over the ones with higher average CV values.

- **Mapping the clusters to the input space:** We map each cluster to input space to know about the range of values that generated these clusters. This gives a better overview of the desired set of values for obtaining the output of interest. Hence, all five clusters are mapped to the input space. Next, we provide a box plot for the distribution of each feature leading to the selected cluster as shown in Figure 6 (right). Box-plot displays variations in the sample of the statistical population without making any assumptions of the underlying data. It also shows the degree of dispersion, skewness, and

Figure 7. Local minima in the direction of σ_2 while other parameters are kept constant at [1,25,0.01,200] for different resolutions

outliers in the data.

Inferences from the Visual Parameter Space Analysis

Applying the visualization tool gives the following first insights into spunbond processes.

Local Minima in Optimizing the Process Parameters Customized for Engineers

During our analysis, we realized that there were local minima when there was a maximum of two free parameters while other input parameters are kept fixed. During the physical production of the nonwovens, sometimes the engineer does not have the flexibility to change all the process parameters. Our analysis helps the engineer to find local minima while keeping some parameters fixed. Hence our tool provides customized analysis for engineers based on their requirement. For example, if the engineers are interested in improving the quality where they are allowed to optimize only the parameter σ_2 while keeping the other parameters constant (here $[\sigma_1 = 1, A = 25, v = 0.01, n = 200]$), they can find the local minimum as shown in Figure 7.

Desired Resolutions for Differentiating Nonwovens

Different resolutions correspond to the same virtual material, but are treated as different metrics to judge the quality. During our analysis, we observed that the higher resolutions often do a better job classifying nonwoven materials as "good" or "bad" compared to lower resolutions.

Figure 8. Comparing two nonwoven materials using the visual tool (left) and rastered images of actual generated samples (right)

As an example, we consider two nonwoven materials with the following parameter settings:

- 1) Material_1 = [1, 1, 1, 0.01, 200].
- 2) Material_2 = [1, 5, 1, 0.01, 200].

When we compare these two materials, as seen in Figure 8 (left), the metric for resolutions 0.5 mm, 1 mm and 2 mm is better for Material_2 than for Material_1, while for the remaining lower resolutions 5 mm, 10 mm, 20 mm and 50 mm Material_1 is better than Material_2. However, when we generate actual samples for both settings, as we can see from the Figure 8 (right), Material_1 is bad nonwoven with horizontal stripes when compared to Material_2 which shows no such obvious artifacts.

This shows that higher resolutions can be better for comparing two materials compared to the lower resolutions. We could also realize this in our individual resolution plot Figure 7 for the whole range of input feature (σ_2). There is a clear separation in the behavior of the materials between higher and lower resolutions.

Optimal vs Non-sensitive Parameter Settings

The input features do not directly affect the production process, but they decide the values of the process parameters, affecting the product quality. We cannot ascertain the process parameters to exactly have the desired values in reality because of various factors in production such as manufacturing error, human error, modeling error, etc. If the input parameter setting is sensitive to small changes, it might result in a bad quality

Figure 9. (A) Comparing the CV values of two parameter settings, (B) Input Space Graph showing small changes around the settings and (c) Output Space Graph showing corresponding changes in the outputs.

of the nonwovens. Hence, during the parameter space exploration, along with finding optimal parameter settings for the desired output, the engineer needs to make sure that these settings are non-sensitive to small changes around them. In our sensitivity analysis, we found that some parameter settings are optimal but are very sensitive to small changes.

From figure Figure 9 (A), we can observe that the input setting [21.7, 44.44, 13.22, 0.08, 952] gives better CV values when compared to the setting [21.2, 14.46, 31.01, 0.2, 1985]. However, the second setting is more robust, as small variations to both settings (Figure 9 (B)) lead to smaller changes Figure 9 (C).

Conclusion

In the paper, we demonstrated the successful visualization approaches in the industrial production of nonwovens in the context of parameter space analysis. We presented a visual analytical tool that aids engineers visually in finding the optimal parameters of interest while navigating through the parameter space. We also performed visual analysis, helping engineers making sound design and engineering decisions while reducing the need for domain expertise. Furthermore, we reported the significant inferences made with the help of our tool, which contributes to the decision-making process. Future work can involve the inclusion of virtual nonwoven images to bring human perception into the analysis.

Acknowledgment

This work was developed in Fraunhofer Cluster of Excellence Cognitive Internet Technologies. Furthermore, we acknowledge the support of the High Performance Center Simulation and Software Based Innovation.

REFERENCES

1. EDANA, "European nonwoven production grows by over 7 % to exceed 3 million tonnes in 2020" *Press release*, 30/03/2021, <https://www.edana.org/about-us/news/2020-nonwovens-market-insights>, accessed: 2021-07-30
2. Albrecht, Wilhelm and Fuchs, Hilmar and Kittelmann, Walter, *Nonwoven Fabrics: Raw Materials, Manufacture, Applications, Characteristics, Testing Processes*, WileyVCH, 2002.
3. Klar, Axel and Marheineke, Nicole and Wegener, Raimund, "Hierarchy of mathematical models for production processes of technical textiles", *ZAMM Zeitschrift für Angewandte Mathematik und Mechanik*, vol. 89, no. 12, pp. 941–961, 2009.
4. Gramsch, Simone and Hietel, Dietmar and Wegener, Raimund, "Optimizing spunbond, meltblown and airlay processes with FIDYST" *Melliand International*, vol. 96, no. 2, pp. 96–98, 2015.
5. Bonilla, L.L. and Götz, T. and Klar, A. and Marheineke, N. and Wegener, R. "Hydrodynamic Limit of a FokkerPlanck Equation Describing Fiber Lay-Down Processes" *SIAM J. Appl. Math.*, vol. 68, no. 3, pp. 648–665, 2007.
6. Seber, George AF and Lee, Alan J, *Linear regression analysis*, John Wiley & Sons, vol. 329, 2012.
7. Ostertagová, Eva, "Modelling using polynomial regres-

- sion," *Procedia Engineering*., Elsevier, vol. 48, pp. 500–506, 2012.
8. Breiman, Leo, "Random forests," *Machine learning*., Springer, vol. 45, no. 1, pp. 5–32, 2001.
 9. Bishop, Christopher M and Tipping, Michael E, "Bayesian regression and classification," *Nato Science Series sub Series III Computer And Systems Sciences*., IOS PRESS, vol. 190, pp. 267–288, 2003.
 10. Specht, Donald F and others, "A general regression neural network," *IEEE transactions on neural networks*., Citeseer, vol. 2, no. 2, pp. 568–576, 1991.
 11. De Myttenaere, Arnaud and Golden, Boris and Le Grand, Bénédicte and Rossi, Fabrice, "Mean absolute percentage error for regression models," *Neurocomputing*., Elsevier, vol. 192, pp. 38–48, 2016.
 12. Sedlmair, Michael and Heinzl, Christoph and Bruckner, Stefan and Piringer, Harald and Möller, Torsten, "Visual parameter space analysis: A conceptual framework," *IEEE Transactions on Visualization and Computer Graphics*., vol. 20, no. 12, pp. 2161–2170, 2014.
 13. Berger, Wolfgang and Piringer, Harald and Filzmoser, Peter and Gröller, Eduard, "Uncertainty-aware exploration of continuous parameter spaces using multivariate prediction," *Computer Graphics Forum*., vol. 30, no. 3, pp. 911–920, 2011.
 14. Helton, Jon C and Davis, Freddie Joe, "Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems," *Reliability Engineering & System Safety*., Elsevier, vol. 81, no. 1, pp. 23–69, 2003.
 15. Wold, Svante and Esbensen, Kim and Geladi, Paul, "Principal component analysis," *Chemometrics and intelligent laboratory systems*., Elsevier, vol. 2, no. 1-3, pp. 37–52, 1987.
 16. Xu, Rui and Wunsch, Don, *Clustering*, John Wiley & Sons, vol. 10, 2008.
 17. Yang, Miin-Shen and Lai, Chien-Yo and Lin, Chih-Ying, "A robust EM clustering algorithm for Gaussian mixture models," *Pattern Recognition*., Elsevier, vol. 45, no. 11, pp. 3950–3961, 2012.
 18. Likas, Aristidis and Vlassis, Nikos and Verbeek, Jakob J, "The global k-means clustering algorithm," *Pattern recognition*., Elsevier, vol. 36, no. 2, pp. 451–461, 2003.
 19. Watanabe, Sumio, "A widely applicable Bayesian information criterion," *Journal of Machine Learning Research*., vol. 14, no. Mar, pp. 867–897, 2013.

Viny Saajan Victor is currently pursuing PhD at Technical University Kaiserslautern and Fraunhofer ITWM. He received M.S. degree from Technical University Kaiserslautern. His research interests include

Visual Information Analysis, Scientific Visualization, Computer Graphics and Machine Learning. Contact him at viny.sajaan.victor@itwm.fraunhofer.de.

Dr. Andre Schmeißer is a research scientist at the Fraunhofer ITWM and has received his Ph.D. in computer science from the Technical University Kaiserslautern. His research interests include Scientific Computing, Scientific Visualization, and Machine Learning. Contact him at andre.schmeisser@itwm.fraunhofer.de.

Prof. Dr. Heike Leitte is professor for computer science in visual information analysis department at TU Kaiserslautern. Her research interests include Interactive visual data exploration, Topological data analysis, Machine learning in visualization. Contact her at leitte@cs.uni-kl.de.

Prof. Dr. Simone Gramsch is professor for mathematics in engineering at Frankfurt UAS and research scientist at Fraunhofer ITWM. Her research interests are Modeling, Simulating, and Optimizing nonwoven productions processes as well as Scientific Visualization. Contact her at simone.gramsch@itwm.fraunhofer.de.