Available online at www.sciencedirect.com

ScienceDirect

Procedia CIRP 136 (2025) 147–152

www.elsevier.com/locate/procedia

35th CIRP Design 2025

# Enhancing Design for Additive Manufacturing Through 3D Objects Clustering Using HICA and Segmentation Techniques

Kim Da Rosa[a], Elise Gruhier[a], Robin Kromer[b]*

[a]Arts et Metiers Institute of Technology, University of Bordeaux, CNRS, Bordeaux INP, INRAE, I2M Bordeaux, F-33400 Talence, France
[b]University of Bordeaux, CNRS, Arts et Metiers Institute of Technology, Bordeaux INP, INRAE, I2M Bordeaux, F-33400 Talence, France

* Corresponding author. *E-mail address:* robin.kromer@u-bordeaux.fr

**Abstract**

A novel approach for 3D objects clustering to enhance detail design phase is developed with the Hierarchical Clustering Algorithm (HICA). This bigdata analytic extract features like vertex count, genus or convexity. The aim is to classify a 3D objects database and their compatibility with different manufacturing technologies (casting, milling and additive manufacturing), thus facilitating more informed decision-making for designers. The result is the identification of main features based on Variance Inflation Factor (VIF) that enables to evaluate clusters that share similar characteristics. Mean dihedral angles, minimal thickness, betti numbers, accessibility score are found. Then principal components are employed to offer specific information and enable interpretation. This latter can be directly applied to refine and adapt their designs for various manufacturing technologies or use specific segmentation tools.



*Keywords:* Decision-Making ; Design For Manufacturing ; Clustering ; Process Manufactruing

## 1. Introduction

The emergence of Additive Manufacturing (AM) process has completely changed the situation for engineers. Designers can now imagine complex geometries. Indeed, all kind of products can be manufactured with all the process on the market. However, this great variety of process goes along with lots of constraints to respect. For instance, in AM the dimension of the part is limited to the length of the plate. In machining, no intern geometry is allowed without creating a gate for the tool access. In casting, gate system must be designed to deliver the molten metal into mould cavity [1]. The constraints can be about dimensions, geometries, maximal elastic strain and so on. Therefore, it is necessary to aid designers in their tasks so that they could focus on modelling the product and not being distracted with the process to choose. They need information about the most reliable manufacturing process that could be used to produce the part most easily.

The research work proposed in this paper presents a novel approach to cluster 3D models using HICA approach. This algorithm enables the automation of pattern recognition (cluster analysis in bid data research and data mining). HICA's clustering mechanism operates by analysing high-dimensional feature vectors that characterize 3D model attributes (i.e. surface complexity, internal structure, and spatial configuration). These features are then grouped into clusters that reflect process feasibility, based on specific constraints associated. Beyond simply aiding in process selection, HICA also provides quantitative insights for potential automatic feature adjustments, such as modifying dimensions or wall thicknesses to align with optimal manufacturing parameters in order to use all potential of the different processes [2].


10.1016/j.procir.2025.08.027

The integration of the approach through the design process is first going to be presented. Then identification of features and clustering of a 3D model dataset with HICA algorithm will be described. Finally, the results are presented and discussed for segmentation algorithm.

## 2. State of the art

The designer needs to think about many constraints when designing the product. For instance, the customer requirements have to be conformed and the design product has to be manufacturable [1]. To aid him in his activities and minimize his responsibilities, methods such as DFM (Design For Manufacturing) have been developed. The idea of DFM is to ensure and simplify the manufacturing of products to reduce manufacturing costs [3]. Markus et al. [4] et al. describe the different kinds of solutions to aid designers in their tasks by introducing manufacturing. For them, the current DFM methods mainly support the design engineer in adapting the concept to the limitations of the manufacturing process. For example, Bikas et al. [5] present a framework enabling designers to evaluate part manufacturability across both additive and subtractive methods.

Advancements in clustering algorithms have led to more automated approaches for classifying 3D models based on manufacturability. One example is Ortek et al.'s [3] semi-automated approach, which utilizes the HICA (Hierarchical Clustering Analysis) algorithm to group models according to their geometric and structural characteristics. While partially unsupervised, the algorithm presumes a pre-defined classification into additive or traditional manufacturing families. This pre-classification aids in simplifying the clustering process, but it also introduces limitations by constraining the flexibility of the model groupings.

In recent developments, advanced clustering techniques and feature extraction methods have enhanced the precision of model classification. Based on features in an unsupervised hierarchical clustering framework, clusters of models with shared manufacturability characteristics are formed, enabling a streamlined selection of suitable manufacturing methods for each cluster. Techniques like PCA (Principal Component Analysis) are employed for dimensionality reduction, improving clustering efficiency and interpretability. This paper proposes a clustering method capable of identifying nuanced geometric and topological patterns within high-dimensional datasets. By adopting a fully unsupervised approach, this methodology eliminates the need for pre-assigned manufacturing classifications, allowing for a more flexible analysis that can dynamically reveal suitable manufacturing processes. The framework developed enables the identification of critical geometric and functional features.

## 3. Integration in the design process

The research focuses on manufacturing processes and needs integration into a comprehensive design methodology. The approach is applicable to both routine and innovative designs and can be used in the detailed design stage of product development, provided the volume is modelled in CAD (Computer-Aided Design) tools. Based on Favi et al.'s [6] DFM approach, our work can be integrated between the design concept phase and the best design choice phase. The designer begins by developing a design concept that aligns with the initial requirements. After selecting a design, both DFA (Design For Assembly) and DFM are applied to optimize the chosen concept. DFM is used to streamline the assembly process, focusing on reducing the number of components, simplifying assembly steps, and minimizing assembly time and costs. Complementarily, DFM is executed to ensure the design's manufacturability, which includes optimizing dimensions, tolerances, and material choices to fit specific manufacturing processes. Our approach further complements DFA and DFM by offering a structured framework for manufacturability assessment through clustering techniques, such as HICA (Figure 1). Jiang et al. [7] proposed machine learning integrated design for AM framework. They studied the relationships between the design and performance spaces.

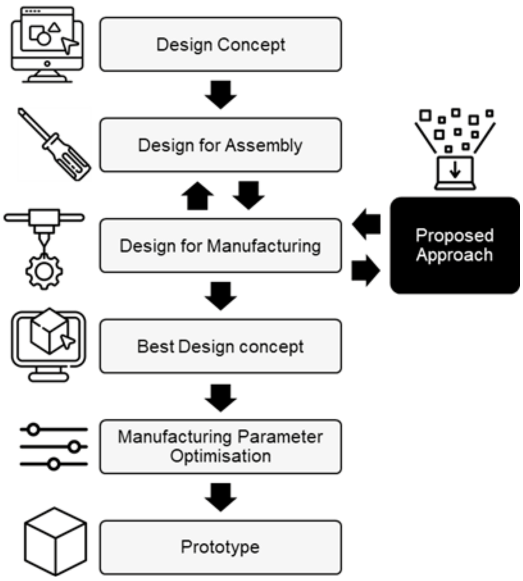


Figure 1: Positioning of the proposed approach in the design workflow

To apply the proposed approach, some inputs are needed. Descriptors based on global features, especially those derived from geometric properties, are typically the most straightforward to extract. In [8], a search engine for filtering 3D models based on geometric attributes was proposed, utilizing features such as volume, surface area, crinkliness, and compactness. Also, Corney et al. [9] expanded these descriptors with hull-related features, reducing shape descriptor variability by smoothing out minor features. Additional characteristics like bounding-box aspect ratio, face count, and hole count were suggested. In [10], a similarity factor based on the average ratio of properties, such as edges, faces, and hollows, was used effectively on industrial parts. Other algorithms leverage moments [11,12] and spherical harmonics [13], which are computationally intensive but allow comparisons across 3D scans and surface meshes.

Graph database are also popular, where a 3D model is represented as an attributed graph with faces as nodes and edges as links [14]. Later, You and Tsai [15] advanced this with simulated annealing to optimize common subgraph matching, detecting global and local similarities at the cost of high computation times.

## 4. Novel approach for 3D objects clustering

Clustering requires quantitative data for each component to obtain families that have similarities. Below are listed the different steps done to cluster 3D objects:

- The input data need to be formatted. The 3D model from the CAD tool is meshed and retrieved. The Thingi10k dataset contains 10,000 3D models in STL format.
- Geometric property extraction follows, defining and calculating features for each 3D model. Key properties like *face/edge/vertices count, manifoldness, genus, betti numbers, volume, surface area, minimal thickness, convexity* and *dihedral angles* are extracted. These features provide a foundation for distinguishing different shapes and structures. In addition, different scores are computed such as *draft angle score* (angles on vertical surfaces relative to the parting line), *complexity score* (overall geometric complexity), *accessibility score* (analyzing tool reach across different axes), *symmetry/axisymmetry* score and *undercut score* (trapped area).
- A mesh analysis is used to systematically extract each model's geometric properties. Standard scaler from ScikitLearn library then standardizes these features, ensuring comparability across different numerical scales. Variance Inflation Factor (VIF) measures multicollinearity among the features in your dataset, indicating how much one feature can be explained by others. High VIF values (>10) suggest that a feature is highly correlated with others, introducing redundancy and potentially leading to instability in models (like linear regression or clustering) by distorting the importance of each feature. This reduced collinearity improves the clustering quality by ensuring that each feature contributes unique information about the dataset. For instance, if features like surface area and volume were found to have high VIF scores, removing one would reduce redundancy because both provide size-related information. The refined set of features leads to more interpretable and stable clusters since each dimension in the clustering process reflects a distinct aspect of the model's geometry or complexity.
- Principal Component Analysis (PCA) reduces dimensionality, allowing easier clustering and improved interpretation. Agglomerative Clustering is applied to the PCA-transformed data to create hierarchical clusters. Cluster quality is assessed using the silhouette score, which helps refine clustering parameters and ensures that clusters are distinct. PCA scatter plots illustrate cluster separations, while bar plots or histograms display the distribution of models within each cluster. Additionally, silhouette plots help confirm well-defined clusters and identify any that require further refinement. This structured approach provides an intuitive understanding of model groups based on geometric properties.

## 5. Results

After a standardization, several features have extremely high VIF values such as *vertices, faces* and *edges number,* as well as complexity. This is likely because these features are either directly or linearly related (e.g., more *vertices* and *edges* often mean higher complexity and more *faces*). Such multicollinearity can lead to redundancy, where each of these features doesn't provide unique information. Other features, including *inertia trace, surface area* and *mean edge length*, also have high VIF values, suggesting moderate multicollinearity. These features may share variance with other features related to the overall size or shape of the 3D models. After filtering out high VIF features, the following features with VIF values below 5 were retained: *mean dihedral angles, minimal thickness, betti 0 and 1 numbers, accessibility score* and *genus*. These selected features are less correlated with each other and are therefore more likely to contribute unique, independent information in subsequent analyses (like clustering). A correlation map provides that the *mean dihedral angle* correlates positively with *betti number 1*, while showing a strong negative correlation with *accessibility score*. *minimal thickness* and *betti number 0* also exhibit a moderate negative relationship.

Then, an "Elbow Method" is used. WCSS (Within-Cluster Sum of Squares) measures the sum of squared distances between each point and the centroid of its assigned cluster [16]. Lower WCSS values indicate that points are closer to their cluster centers, implying tighter, more cohesive clusters. The "elbow" represents a point where the rate of decrease in WCSS slows significantly. This point is typically chosen as the optimal number of clusters because adding more clusters beyond this point results in only marginal improvements in WCSS, indicating diminishing returns. After, three clusters, the rate of decrease in WCSS slows, meaning additional clusters offer less benefit in terms of reducing within-cluster variance. Selecting three clusters balances the trade-off between cohesion and complexity. Therefore, clustering results revealed distinct model groups, each with unique geometric attributes. The heatmap provides a detailed comparison of various topological and geometric features across three distinct clusters, offering insights into the structural complexity and manufacturability of each group (Figure 2). Cluster 0 shows a remarkably high *betti number 0*, indicating a larger number of disconnected components or separate regions within these models, suggesting designs that may have multiple parts or segments. This characteristic could imply a need for assembly considerations or potential challenges in maintaining structural cohesion if manufactured as a single piece. In contrast, Cluster

2 has the lowest *betti number 1*, which reflects a simpler structure with fewer loops or holes, potentially making it more straightforward to manufacture with minimal internal complexity. *minimal thickness* values vary notably across clusters, with Cluster 1 showing the highest value, indicating thicker components that could contribute to better structural integrity but might require higher material usage and extended production time. *Genus* scores for Clusters 1 and 2 suggest that models in these clusters have multiple holes, adding to their topological complexity and possibly influencing design considerations for support structures in additive manufacturing. Lastly, the *accessibility score* is higher in Cluster 2, which may indicate that models in this cluster have more accessible surfaces, facilitating easier post-processing and inspection. These detailed insights into each cluster's structural and topological characteristics offer valuable guidance in selecting manufacturing techniques and highlight potential challenges related to material usage, assembly, and complexity.

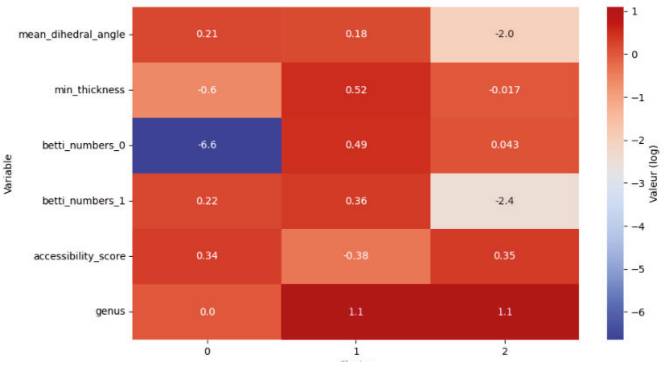


Figure 2: Heatmap displays the log-scaled averages of various topological and geometric features across three clusters

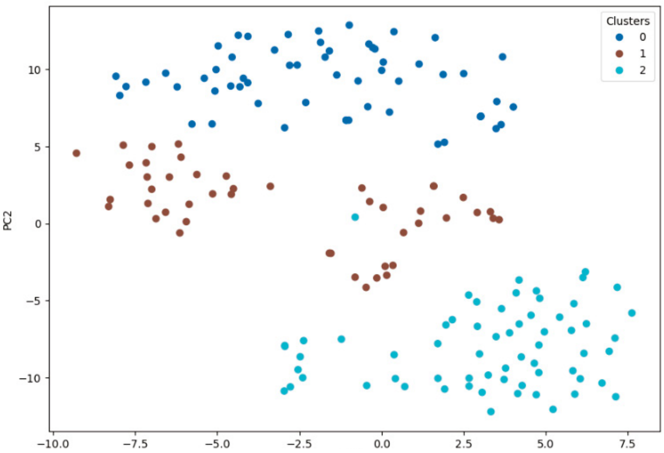


Figure 3: Scatter plot illustrates the distribution of data points across the three clusters

Different representations are used for interpretation. First, the t-SNE (t-Distributed Stochastic Neighbor Embedding) plot provided here visualizes the clustering results by reducing the high-dimensional data into two components (PC1 and 2). Each color represents a different cluster, allowing us to visually inspect the separation and cohesion of clusters. Figure 3 illustrates the distribution of data points across clusters: Cluster 0 (dark blue), Cluster 1 (brown), and Cluster 2 (light blue). Second, Figure 4 visualizes the principal components (PC1 and PC2) along with feature vectors representing key variables: *betti numbers 0, betti numbers 1, min thickness, mean dihedral angle,* and *accessibility score*. *Betti numbers 0* and *minimal thickness* contribute strongly along PC2, suggesting that variations in connectivity and thickness are prominent factors in this dimension. *Mean dihedral angle* shows a clear influence along PC1, indicating its significance in differentiating samples along this axis, potentially related to surface angularity. *Accessibility score* has a moderate contribution along both axes, showing its role in defining overall shape characteristics across clusters. The alignment and direction of these vectors indicate correlations between features and their effect on the clustering structure.

HICA method enable the generation of a dendrogram to

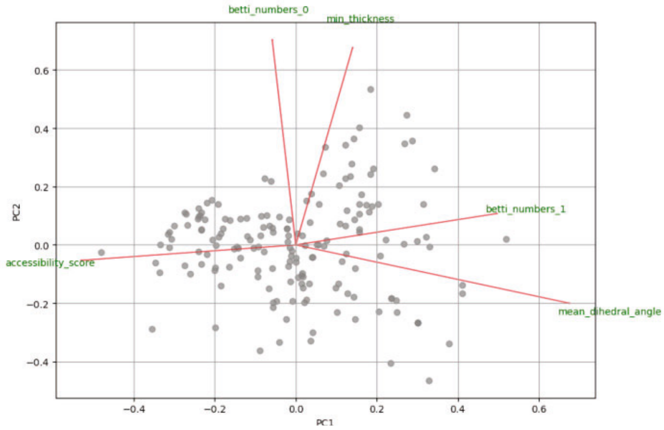


Figure 4: biplot visualizes the principal components (PC1 and PC2) along with feature vectors representing key variables

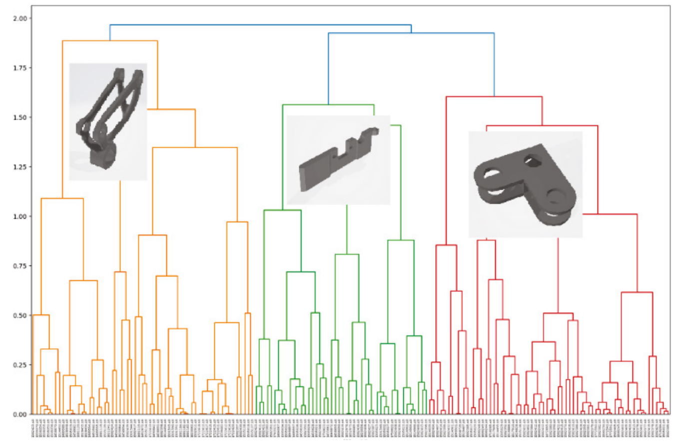

Figure 5: Dendrogram of the 3 clusters based on HICA – 3D model examples in each cluster

visualize clustering structure of the dataset, illustrating how samples are grouped based on their similarity across the selected features. Each branching point, or node, represents a merge between clusters, with the height of the node indicating the distance (or dissimilarity) between the clusters being combined. The dendrogram displays three primary clusters (color-coded as orange, green, and red), suggesting a natural grouping within the dataset (Figure 5). The height at which branches merge indicates cluster cohesion; lower branches reflect higher similarity within the merged clusters, while taller branches suggest more distinct groups. The large separation between the top-level branches shows significant dissimilarity among the three main clusters, possibly reflecting distinct structural or geometric characteristics in each group. This hierarchical clustering visualization aids in understanding how samples progressively group at different levels of similarity, providing a potential cutoff point to define the optimal number of clusters for downstream analysis. The chosen clusters can then be examined for unique characteristics to guide manufacturing decisions based on the specific features that define each group.

For example, a component with a high genus, such as a porous scaffold used in biomedical applications, may be challenging to manufacture using conventional processes like milling but well-suited for additive manufacturing. The first component, characterized by its cylindrical conduit transitioning to a rectangular flange, exhibits a mean dihedral angle in the approximate range of 108 degrees, reflecting the predominantly prismatic faces and the potential presence of large fillets or chamfers at critical junctions. Betti 1 takes a value of 1 due to the principal longitudinal channel. Furthermore, an accessibility score in the vicinity of 86% underscores the relatively unimpeded approach paths for potential tooling or inspection probes. Conversely, pump housing with high accessible score may be more compatible with traditional manufacturing techniques such as casting. Larger mean dihedral angle of roughly 112° are computed driven by the multitude of short edges formed around each circular perforation. With a minimal thickness likely lying 1mm, the lightweight construction achieves a high surface-area-to-volume ratio but requires careful consideration of local stress concentrations. Betti 1 climb into 44, given the numerous through-holes distributed around the structure. The accessibility score 46% reflects the inherent difficulty.

## 6. Discussion

Clusters that differ significantly along PC1 are likely distinguished by structure-related features, such as *minimal thickness* and *mean dihedral angle*. These robust models may be more suitable for manufacturing methods that favor thicker sections, such as casting or milling, where structural integrity is critical. Clusters separated along PC2, however, are differentiated by topological complexity (*betti number 0* and *betti numbers 1*), which suggest a higher number of disconnected regions and loops or holes. These characteristics may align with designs requiring modular assembly or support structures in additive manufacturing. Models with high *accessibility score* align closely with the origin, suggesting moderate influence across PC1 and PC2 without strongly distinguishing clusters. Integrating design rules for manufacturing within this framework opens up significant possibilities for optimizing part design by aligning it closely with specific manufacturing process requirements. For instance, parts with high *minimum thickness* and *betti numbers 0* values could trigger design rules tailored to casting processes, such as ensuring gradual thickness transitions to reduce stress concentration or minimizing sharp internal corners to enhance mold flow. Similarly, models with high *mean dihedral angle* and *accessibility* score could be flagged as suitable for machining. Design rules for machining, like avoiding undercuts or ensuring open access to features, can be recommended based on the specific cluster attributes. In additive manufacturing (AM), intricate features (indicated by low *minimal thickness* and/or high *betti numbers 1*) might suggest adjusting part orientation to minimize the need for supports. The dendrogram (cf. Figure 5) enables a systemic approach to reduce trial-and-error in the design stage and helps designers make informed choices. Representative parts embedded within each of the main clusters (orange: AM, green: milling, and red: casting) illustrate the types of geometries and structures grouped together. The three groups are listed below:

- The AM group primarily includes parts with intricate, elongated geometries and minimal thickness, as shown by the part with a slender, arm-like structure.
- The parts in the Casting cluster, exemplified by the middle part, generally exhibit simpler, flatter geometries with linear features.
- The parts within the milling cluster, as represented by the rounded component with through-holes, tend to have robust, interconnected features with greater thickness and rotational symmetry.

Using Product Data Management (PDM), it is possible to align functional requirements and manufacturing constraints, enhancing efficiency and reducing production costs. HICA clustering can be integrated into an automated framework that classifies new 3D models into clusters and applies corresponding guidelines. For example, when a new model is uploaded, it could automatically be assigned to one of the three clusters based on its geometric and structural features, with predefined adjustments suggested based on cluster characteristics. Neighbor models within the cluster can be identified and provided as examples, assisting the designer in refining their part to fit manufacturability standards. The AM model (in Figure 5) is assigned to the cluster associated with slender (elongated structures). The system might then prompt DFAM (Design For Additive Manufacturing) guidelines, such as ensuring a minimum wall thickness for structural integrity and avoiding unsupported overhangs. Neighboring models within the same cluster would be displayed, showcasing similar designs that have already been optimized for AM, helping the designer understand practical adjustments. For instance, if other models in the cluster include support struts or slight modifications to reduce overhang angles, the system could suggest these adaptations to improve manufacturability. This automated classification and guidance streamline the design process, reduce the need for trial-and-error adjustments, and improve the final part's alignment with production requirements.

Finally, segmentation technic can also be adapted depending on clusters. For milling, which is a subtractive process using rotary cutters to remove material, segmentation can involve feature-based approaches where standard machining features like pockets, slots, holes, and bosses are identified. They need assigned specific machining strategies, optimizing toolpaths and cutting parameters to enhance efficiency and precision. Accessibility analysis is also crucial, segmenting regions based on tool reachability to prevent collisions and ensure all surfaces are machined effectively. Casting, involving pouring liquid material into molds, benefits from segmentation techniques that address mold design considerations; draft angle segmentation identifies surfaces with insufficient draft relative to mold removal direction, using angle measurement tools and color-coded visualization to adjust designs for easy ejection, thereby reducing the risk of the casting sticking to the mold. In 3D printing or AM, segmentation techniques address considerations like support structures, print orientation, and material properties. For AM model segmentation with geometric recognition provides a powerful approach to analyzing and optimizing complex 3D meshes. AM models can be decomposed into functional zones with distinct design and manufacturing needs. For instance, regions can be marked for rotation-optimized printing, overhangs could be reinforced or adjusted to minimize support structures [17]. This approach creates a comprehensive segmentation strategy that categorizes a model's geometry into actionable segments, tailored to AM constraints, enabling more efficient and high-quality additive manufacturing through automated, feature-based model adjustments. Integrating segmentation into the design workflow involves using CAD software with unified platforms that allow for real-time analysis, automation tools for feature recognition, user-friendly visualization with interactive interfaces and immediate feedback. However, it requires a large dataset with all specific features based on manufacturing processes.

## 7. Conclusion

The article presents a framework that leverages HICA to classify 3D models based on geometric and topological features. The framework includes a feature analysis based on VIF. The identification of 3D model features can thus be done and provides a big data analysis for clustering. Five features were retained: *mean dihedral angles, minimal thickness, betti 0 and 1 numbers, accessibility score* and *genus*. It suggests variations in connectivity and thickness, differentiation samples along axis and definition overall shape characteristics across clusters. Dendrogram from HICA analysis enables displaying neighboring models within the same cluster. Three clusters have been highlighted and features characteristics of clusters enabled to define best suited manufacturing technology. This framework aligns functional design requirements with manufacturing constraints, offering new design making solution.

## Acknowledgements

This research was supported by the French Research National Agency (ANR) Grant No. ARN-24-CE10-3374.

## References

[1] Swift KG, Booker JD. Manufacturing process selection handbook. Butterworth-Heinemann;2013.

[2] Ordek B, Borgianni Y. HICA: A MATLAB-based hierarchical image clustering algorithm for classifying parts suitable for additive and traditional manufacturing technologies. SoftwareX 2024;27.

[3] Rolls-Royce. How to design for manufacture and assembly – Value chain competitiveness;2019.

[4] Markus D, Lorin A, Thomas N, Sven M. Identifying an opportunistic method in design for manufacturing: an experimental study on successful a on the manufacturability ans manufacturing effort of design concepts. 31st CIRP Design Conference 2021;100:720-725.

[5] Bikas H, Koutsoukos S, Stavropoulos P. A decision support method for evaluation and process selection of additive manufacturing. 52nd CIRP Conference on Manufacturing Systems 2019;81:1107-1112.

[6] Favi C, Germani M, Mandolini M. Design for Manufacturing and Assembly vs. Design to Cost: toward a multi-objective approach for decision-making strategies during conceptual design of complex products. 26th CIRP Design Conference 2016;50:275-280.

[7] Jiang, J, Xiong, Y, Zhang, Z, Rosen, DW. Machine learning integrated design for additive manufacturing. J Intell Manuf. 2022; 33, 1073–1086.

[8] Rea HJ, Corney JR, Clark DER, Pritchard J, Breaks ML. Part-sourcing in a global market. Concurrent Engineering. 2002;10(4):325–333.

[9] Corney J, Rea H, Clark D, Pritchard J, Breaks M, MacLeod R. Coarse filters for shape matching. IEEE Computer Graphics and Applications 2002; 22(3):65–74.

[10] Cicconi P, Raffaeli R, Germani M. An approach to support model-based definition by PMI annotations. Computer-Aided Design and Applications 2016;14(4):526–534.

[11] Cybenko G, Bhasin A, Cohen KD. Pattern recognition of 3D CAD objects: towards an electronic yellow page of mechanical parts. International Journal of Smart Engineering Systems Design 1997;1(1):1–13.

[12] Li Z, Zhou X, Liu W. A geometric reasoning approach to hierarchical representation for B-rep model retrieval. Computer-Aided Design 2015;62:190–202.

[13] Kazhdan M, Funkhouser T, Rusinkiewicz S. Rotation invariant spherical harmonic representation of 3D shape descriptors. ESGP 2003;pp156–164.

[14] El-Mehalawi M Miller RA. A database system of mechanical components based on geometric and topological similarity. Computer-Aided Design 2003;35(1):95–105.

[15] You CF, Tsai YL. 3D solid model retrieval for engineering reuse based on local feature correspondence. International Journal of Advanced Manufacturing Technology 2010;46(5–8):649–661.

[16] Shi, C., Wei, B., Wei, S. et al. A quantitative discriminant method of elbow point for the optimal number of clusters in clustering algorithm. J Wireless Com Network 2021, 31 (2021).

[17] Mbow M, Grandvallet C, Vignat F, Marin FR, Perry N, Mathematization of experts knowledge: example of part orientation in additive manufacturing. Journal of Intelligent Manufacturing, 2021, pp.1-19