-- CONFIDENTIAL --

Performance Matrix for Shape-Search Applications

Numerous methods have been developed to simplify and represent 3-dimensional shape for use in shape-search applications, and the strengths and weaknesses have been well-studied. This document summarizes that information to aid in the evaluation of shape-search applications.

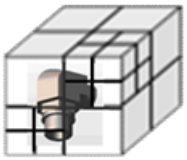
Useful Background Information and Jargon

- Shape-search applications do not compare CAD models directly, but instead create a **shape-representation** through abstracting processes, such as:

Parametric

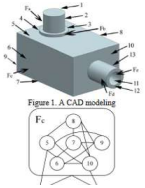
Shape Vector	Part A	Part B	Match in Tolerance
Surface Area	147.75	147.75	Yes
Volume	325.69	325.67	Yes
Moments	4283	4283	Yes
Report as Match?			Yes

Sub-division



Projection

Graphing



Method	Description	Primary Advantages	Primary Disadvantages
Parametric-Substitution	The method uses multiple parametric values such as surface area, volume and moment of inertia as surrogates for shape, using that logic that if two models have the same parametric data, then the models might be identical.	- Simple - Real-time search	- No assemblies - Low accuracy for similarity - Data not always available
CAD Volume Sub-Division	Volume sub-division partitions the CAD volume into smaller cubes, and a parameter like surface area is measured for each cube. The parametric 'maps' are then compared in a pair-wise fashion to determine matches which are recorded in the binary index.	- Potential improved accuracy	- No real-time search for new models - No neutral formats - No assemblies - Low accuracy for similarity
Projection	2D images are projected from the model, and the collection of images are compared in a pair-wise fashion to determine matches.	- Any CAD format	- No real-time search for new models - Low accuracy similarity. - Large footprint for database
Graphing	A graph records the edge and loop relationships of the CAD model. Graphs are then compared in a pair-wise fashion to determine matches.		- No assemblies - No real-time search for new models - Only simple models
Pattern Recognition	Rather than use a surrogate to represent shape, pattern recognition captures the full geometry of the model directly. The geometry is then encoded into a long string of numbers (a **shape-signature**), and the more similar any two signatures the more similar the shape.	- Real time search with new models - High Accuracy for similarity - Similarity scoring - All CAD formats - Parts and assemblies - Full geometry - can even recognize differences in scale of curvature - Highly scalable	

Symbol	Meaning
◎	Yes. Plus symbol (+) indicates excellent performance
◯	No
-	Poor Performance
Δ	Manual Operations Required
?	Uncertain if feature is available

Function & Performance	Description	Volume Sub-division Geolus	Graphing with Parametrization Geosearch	Projection with Parametrization Bingo!	Parametrization ExaLead and many others	Pattern Recognition CADseek
Applications	The applications listed at right are believed to utilize the technique indicated in the column header based on information found in patents and published papers.					
	The ratings for the categories list below are based on well-studied capabilities and limitations of each method. Nevertheless, it is possible a provider using the method has found a technique to overcome a limitation, making evaluation necessary if claimed performance differs.					
Indexable	Shape representation is in Indexable format. (Note 1)	◯	◯	◯	◯	◎
Matching Type	Matching index is of Continuum type. (Note 1)	◯	◯	◯	◯	◎
Geometry	Ability to capture all facets of shape from CAD model	◯	◯	◯	◯	◎
Speed	Average time per model or assembly to capture shape	?	?	?	?	<1 second
Geometric Classification	Ability to classify (group) a CAD database using only geometry	◯	◯	◯	◯	◎
Automated Classification	Ability to automatically group parts by geometry (Note 2)	◯	Δ	◯	◯	◎
Native CAD	Applicable to many popular CAD formats	◯	◯	◯	◯	◎
Assemblies	Applicable to assemblies	◯	◯	◯	◯	◎
Neutral CAD	Applicable to neutral formats like IGES or STL	◯	◯	◯	◯	◎
Accuracy	Identical	◎	◎	◎	◎+	◎+
	Similar scoring (Note 3)	-	-	-	-	◎+
	Ability to score the degree of similarity	◯	◯	◯	◯	◎+
Size	Small data requirement for each shape representation	◯	◯	◯	◎+	◎+
Search Speed	Search response time in a million-plus item dataset.	?	?	?	?	< 2 seconds
Scalability	Ability to deploy in very large datasets	◯	◯	◯	◎+	◎+

Feature	Description	Volume Sub-division	Graphing with Parametrization	Projection with Parametrization	Parametrization	Pattern Recognition
On-the-fly	Generation of shape representation quickly	?	?	?	◎+	◎+
Real-time	Ability to use a **new** model as a search target, quickly	◯	◯	◯	◎	◎
Text Search	Attribute searching to find an initial search target.	◎	◎	◎	◎	◎
Text Filtering	Ability to filter shape-search results with attributes.	◎	◎	◎	◎	◎
Analytics	Ability to automatically sort entire dataset by similarity	◯	◯	◯	◯	◎
	Analytics across two datasets	◯	◯	◯	◯	◎
Exporting	Exporting analytics data to tab-delimited format	◯	◯	◯	◯	◎
Live Search	Monitor a CAD design in process for similarity	◯	◯	◯	◎	◎
Dashboard	New models report showing similarity	?	?	?	?	◎
Rough Sketch	Search by only geometry, not scale (Note 4)	◯	◯	◯	◯	◎
Assembly Extract	Extract components from assembly for search	◯	◯	◯	◯	◎
CAD Search	Search with the model open in CAD	◯	◯	◯	◎	◎
File Open	Opening a model from a local drive for search	◯	◯	◯	◎	◎
Expanded Search	Multi-tiered search	◯	◯	◯	◎	◎+
Visualized Browsing	Browse data by automatic classification of shape families	◯	◯	◯	◯	◎
CAD Tools	Inspect tools with rotation, clipping, explode, etc.	?	?	?	?	◎
Compare	Compare two models visually	?	?	?	?	◎
Hand-sketch	Search with hand-drawn sketch	?	?	?	?	◎
Photo search	Search with photo of part	?	?	?	?	◎
	Search with photo of assembly	◯	◯	◯	◯	◎
QR or RFID	Search with Quick Read label or RFID tag	?	?	?	?	◎
3D Printing	Match photo or CAD to STL format	◯	◯	◯	◯	◎
Updating	Interval for updating with new models (note 5)	?	?	?	?	◎

Use Cases

Not all shape-search applications are capable across all use cases. For example, the vast majority of applications can identify duplicate models, but the abilty to to provide sales quotations requires accuracy for similarity and the ability to perform real-time search. The table below lists common use cases, and the key capabilities required. CADseek provides the capability and features for all use cases listed.

		Required Capability				
Use Case	**Description**	**Accuracy Identical**	**Accuracy Similarity**	**Assemblies**	**Real-Time**	**Other**
General Search	Shape-search in the existing search database	◎	◎	◎		Text / Browsing
Data Cleansing	Ability to find identical models	◎		◎		Neutral Formats
Part Reuse	Ability to find similar models		◎	◎		Similarity Scoring
Live Search	Enabling part reuse with live search capability		◎		◎	On-the-fly
Should-Cost	Estimating cost of new part based on existing cost		◎	◎	◎	On-the-fly
Monitoring	Management monitoring new models for similarity		◎	◎		
Assignment	Selecting designer based on prior similar work		◎	◎		
Vendor	Selecting potential vendors based on parts supplied		◎	◎	◎	On-the-fly
Substitution	Finding substitute parts based on similarity		◎	◎		
Standardization	Finding pools of similar parts	◎	◎	◎		Analytics
Quotation	Quoting prices based on similarity to existing parts		◎	◎	◎	On-the-fly
Identification	Use photos to identify physical parts		◎	◎	◎	Photo Search
Tool Path	Reuse of tool path and manufacturing fixtures		◎			
3D Printing	Finding STL models					Neutral Formats
Price Validation	Analyzing vendor cost		◎	◎		Analytics
Merging	Finding duplication for acquisitions	◎	◎	◎		Analytics

Summary

Most organization have intended use cases that require the capability to: search with **assemblies,** provide **real-time** search, and retrieve **similar** models. These three requirements eliminate all methods except parametric and pattern recognition methods. Parametric techniques, however, struggle with assemblies because it can be difficult or impossible to get data parametric data for assemblies unless it is computed manually. In addition, parametric methods have long been criticized for their lack of ability to differentiate models. A paper by TDWI stated:

"Yet most engineering departments are still limited to out-of-date parametric search tools that produce either thousands of hits or no hits at all."

CADseek by iSEEK Corporation is the only shape-search application that utilizes pattern recognition technology. CADseek is universal across all types of CAD formats including assemblies, and provides search results for new models in real-time. CADseek is a surface-based approach, allowing it to work at the same high resolution of the CAD model. The benefit is the ability to 'see' similarity across all facets of shape, and to accurately score the similarity of models. The overwhelming majority of use cases, such as part reuse, substitution, standardization and analytics rely on the ability to accurately match and score models for similarity. It is important that evaluators don't view the ability to identify identical models as representative of the performance of the application because identifying identical models is relatively simple.

NOTES:

Note 1

Background: Shape-search applications do not actually compare CAD models directly, but instead use a variety of methods to create a shape-representation for each CAD model through a process called abstraction resulting in a **shape representation**, such as a map or graph of the model. Performing a search involves selecting one representation (the 'target') in order to find others like it. While an application's user interface may show images of models, it is important to understand that the image and the underlying shape representation are two very different things. For example, the search interface may show a 3D image of a bracket, but the basis for search may be a parametric value or 2D image of the bracket.

'Shape is composed of multiple **facets:** topology, constraints and scale. The amount of geometry that's embodied in a representation varies depending on an application's sophistication. For example, the parametric method (surface area, volume, etc.) only represents the facet of scale. 2D projections of models simplify geometry so that features such as holes, grooves, curvature and internal geometry are diminished or lost depending on perspective.

While there are several types of shape representation, each falls into one of the following types:

'**Non-indexable** Representations do not have an absolute value (they are not numerical), and include such things as 2D projected images of a CAD model, or a surface area map of a model. These types of representations have no intrinsic basis for comparison. Instead, the comparisons are made in a **pair-wise** fashion and the binary match / no-match result for each pair is recorded in the search database.

'Indexable Representations have an absolute value and can naturally be placed along a **continuum,** such as how any number can be placed along the numerical continuum. In theory, the closer the numerical value of any two representations the more similar the models, but this depends on the technology. For example, two models can have surface area values that vary by only 2%, but that does not mean the models are 98% similar, or have any similarity at all.

Note 2

Automated Classification: The graphing with parametrization method requires the manual creation of a class structure, and then the development of rules for each class. If an object fits the rules for a class (for example: Side 1, Side 2, contact edge between Side 1 and Side 2, angle of contact, etc.), then the model will be automatically placed into that particular class, but many objects will not fit a class and will require manual determination. The classification method is not applicable (NA) to the concept of classification because the representations are based on pair-wise comparisons.

Note 3

Similarity and Similarity Scoring: The graphing with parametrization can report similar but is limited to reporting all models of a particular class, without individual scoring for similarity within the class. In practice, there is a practical limit on the number of classes that can logically be created, and many similar models will be placed into different classes, such as two models that are identical but with different scale, models that are identical except with a different number of holes, etc. In these cases, the similar models will be omitted from the search results because they are in a different class.

Note 4

Rough Sketch vs. Live Search: Rough sketch differs from Live Search in that the designer is not intending to design a part, but instead wants to quickly create a search target, but without spending time for careful dimensioning or feature development. CADseek's unique ability to independently consider shape and scale allows the user to focus on the primary shape without concern for dimensions or minor feature development. Rough sketch search requires the ability to provide expanded search settings that can consider shape and scale independently, as well as on-the-fly generation of a shape-representation in an indexable format in order to provide search results in real-time.

Note 5

Updating of Search Database for New Models: The database that holds the shape search data must be updated periodically for any new or modified models that have been created. The frequency of the updating process is primarily of function of whether shape signatures are indexable or non-indexable. Non-indexable representations require the pair-wise comparison process to be completed for each new pair of models, which may make updating the database impractical to perform with high frequency. The CADseek search index is typically updated on a daily basis.