Proceedings of the IMProVe 2011
International conference on Innovative Methods in Product Design
June 15th – 17th, 2011, Venice, Italy

Shape Similarity: Methods and Industrial Applications

D. Francia [(a)], *A., Ceruti* [(a)], *G., Caligiana* [(a)], *A., Liverani* [(a)]

[(a)] University of Bologna

Article Information

Keywords:
Shape Similarity
Computer Aided Design
Computer Aided Manufacturing
Standardization

Corresponding author:
Daniela Francia
Tel.: +39.051.2093352
Fax.: +39.051.2093412
e-mail: d.francia@unibo.it
Address: v.le Risorgimento, 2 40137 Bologna

Abstract

Mass production, reduced production investment, reduced response time and variability, globalization and increased customization are the most responsible for the need of optimizing the manufacturing process. As companies provide higher levels of customization, the number of products offered increases. Minimizing the cost of providing variety is possible by exploiting shape similarities amongst parts and products. Another reason for utilization of similarity assessment is the cost estimation. Total cost of any part to machine results from material costs, setup costs, tooling costs, and operation costs. Furthermore in very small productions the total cost of new components is affected more than 70% by process planning and CAM (Computer Aided Manufacturing) programming. A relevant part of this time may be saved by modifying the process plan of an old similar part and this is a well known practice in the manufacturing industry.
Shape similarity search in a CAD system is an answer to reduce process plan preparation because it finds similarities in geometry by extracting shape signatures from the 3D models and then comparing these signatures exploiting distance functions. This paper presents a review of the methods adopted to generate shape signatures and an analysis of the distance functions suitable for the comparison. The first part of this work investigates the models commonly proposed to extract the shape signature from a solid part. The second part of the paper discusses the choice of the distance function. At the end of this discussion the most appropriate method is proposed to assess the shape similarity, once the part to be manufactured has been assigned: results are shown in terms of accuracy of comparison and computational time, seeking for a trade-off between them.

1 Introduction

Within the past decade, the engineering, design and drafting world has been experiencing a shift from 2D to 3D CAD. The use of 3D CAD models is replacing the use of 2D CAD drawings just as the use of the drawing board due to the advent of 2D CAD software.

The growing need for 3D CAD models is driven both by the need to create better and more efficiently products, because 3D CAD is more accurate than 2D CAD, and by reasons driven by technological aspects.

Even if 2D drawing is fast and easy, the output 2D drawing does not readily work with purchasing and manufacturing systems. In the most common case the machines used to manufacture parts need 3D CAD files and do not read 2D CAD drawings because 2D drawings do not contain all information to develop a three-dimensional product. In prototyping, for example, a 3D model has to be made because most of the prototyping machines require 3D data. It is so evident that through the use of 3D CAD modelling, engineers are able to create even better designs that meet client requirements in a relatively short amount of time. Also from an aesthetic point of view, a 3D design is more realistic and the engineer has a better ability to make a design more attractive.

Recently, the development of 3D modelling and digitizing technologies has made the model generating process much easier so that many manufacture enterprises can collect 3D parts in digital libraries containing tens of thousands of CAD models to be archived, analyzed and used. These libraries have to be organized to make the traceability of such models immediate: CAD models classification becomes an issue. Globalization, mass production and increased customization are the most responsible for the need of optimizing the manufacturing process. Minimizing the production costs, even of the simplest parts, is possible making use of existing parts, whenever possible, that have to be rapidly found in part libraries.

The cost saving associated with parts reuse is not only due to the reuse of the existing designs, but also to the existing manufacturing processes. Reuse of parts is possible by exploiting shape similarities amongst products. Geometric and topological complexity of a 3D mechanical part can be determined by assessing the similarity between its design properties and the design properties of the closest parts already stored in the database. Moreover, using 3D shape searching early in the design cycle can recommend the optimum manufacturing process and also provide the cost estimation of manufacturing of parts.

Shape similarity search in a CAD database means to find similarities in geometry by extracting shape signatures from the 3D CAD models and then comparing these signatures exploiting distance functions.

To enable the savings of time and money associated with design reuse it is necessary to develop a 3D CAD automatic retrieval system able to find similar parts for a given query, mainly according to its shape. One of the main challenges in this context is the mapping of a 3D model into compact canonical representations referred to as descriptor or feature vector, which serve as search keys during the retrieval process.

Commonly 3D CAD models are indexed by alphanumeric tags with syntax specific to each company; in the area of group technology, various part coding schemes have been proposed but such manual classification schemes are subjective and limited to standard or general mechanical parts, so they cannot work automatically with computers [41].

Since the descriptor decisively influences the performance of the search engine, an appropriate method is necessary to assess the shape similarity, seeking for a trade-off between computational efficiency and relevance of the results.

Several researches on algorithms for 3D model retrieval in various kinds of applications, including mechanical components, are focused on a shape-based 3D model that extracts the feature from geometrical information. Different approaches have been proposed such as the topology-based retrieval, the image-based retrieval and the surface-attributes-based retrieval, as shown in fig.1.

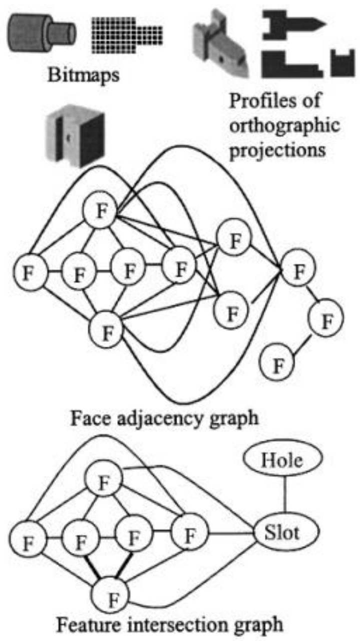

Fig. 1. Three different approaches to the features extraction [23].

Unlike common 3D models, 3D mechanical models usually have semantic information embodying their design intention and deciding their manufacture techniques, which can also be adopted for matching. So the retrieval for mechanical models can use not only the methods for common 3D models above mentioned, but also the structure-based ones. Two or more types of model to extract the feature can even be combined to further enhance the performance.

Similarity assessment between two 3D parts involves two main steps: first to compute the shape signature of the object and second to compare the signatures by a suitable distance function. This paper presents a review of the methods to generate shape signatures and an analysis of the distance functions suitable for the comparison.

2 Overview of techniques

The first part of this work investigates the commonly proposed models to extract the shape signature from a solid part.

Different techniques can be chosen to assess shape similarity of parts: they can be classified on the basis of the type of shape signatures being used, distinguishing among Features, Shape Histograms, Section Images, Topological Graphs or Shape Statistics.

The evaluation of suitable shape signatures and distance functions is a very hard issue: some basic properties such as positivity, identity-self-similarity, symmetry, triangle inequality, invariance, robustness and sensitivity, computational efficiency can guide the researcher through the comparison of the different approaches [1].

In this section the shape signature proposed in the most suitable shape similarity approaches are first described; at the end of this background the distance functions proposed in literature for shape similarity evaluation are analyzed.

2.1 Shape Signatures

Similarity assessment in 3D cases is usually carried out by generating shape signatures from the 3D models, that have to be compared by means of distance functions. These signatures should describe the features of the 3D model needed for similarity assessment. A shape signature could be an image, a graph, a vector, an histogram or an ordered collection of numeric values, depending on the motivation for performing similarity analysis.

Originally, shape similarity measurement techniques have been inspired from the techniques for image retrieval. Examples of similarity measures for images are image outlines, moments and content [2,3].

Regli et al. first observed that pictorial and multimedia information are not directly applicable to digital libraries of 3D solid models where engineering information, such as inter-part relationships function, are more significant [24].

Some example of part representation used for part comparison are axi-symmetric parts compared by means of bitmaps [4,5] a vector representation of profiles of orthographic projections of prismatic parts [6] and face adjacency graphs along with the areas and directions of faces [25].

The evaluation of similarity is strongly affected by the choice of the signature to be extracted from the model. A classification on the basis of the type of shape signatures, as presented in literature, is detailed as follows.

2.1.1 Shape Statistics Based Signatures

The shape statistics comparison techniques use basic geometric properties in order to perform coarse comparison between solids. Properties used for the comparison include volume, surface area, convex hull volume, these numerical values representing statistical properties of the shape of the solid. Such signatures do not carry any topological information and cannot provide sufficient discrimination power for detailed comparison, but they may be used to reduce the search space in the similarity assessment.

The technique described in [26] uses global shape metrics such as surface area/volume ratio, number of holes, compactness, and crinkliness to perform similarity assessment. These metrics are orientation independent and are extracted from a STL file. Compactness is the non-dimensional ratio of the square of the volume over the cube of the surface area while crinkliness is the surface area of the model divided by surface area of a sphere having the same volume. They are calculated for all the solid models and are stored as searchable entries in a database. To analyze the performance of the search engine, similarity matrices based on human perception of similarity have been generated. In [27], four new filters for shape matching have been proposed. These are based on the coefficient of surface area and convex hull of the solid model.

2.1.2 Shape Histograms Based Signatures

The key idea is to transform an arbitrary 3D model into a parameterized function that can be easily compared with others. The CAD industry is dominated by Constructive Solid Geometry (CSG) and Boundary Representation models (B-rep); these formats makes it very difficult to compare CAD models for indexing across. The shape histograms based signature technique, that matches shape distributions, allows to compare CAD models regardless of their model representation. In this case the shape function is a probability distribution, measuring geometric properties of the 3D model: the shape distribution is a 2D characterization of a 3D shape and it represents the distribution of distances between pairs of randomly selected points on the surface of the 3D model. The first step of this methodology is to obtain a shape model, by means of triangularizations, voxelizations, or meshes generation that approximates the solid model. Common file formats for this translation are VRML (Virtual Reality Modelling Language) and STL (Stereolithography), both based on triangularization.

The second step is to compare models by means of the shape distributions: in order to obtain a shape distribution a shape function and a set of sample random points have to be chosen and the relative histogram have to be calculated associated with the shape distribution function.

These shape functions are easy to compute and produce distributions that are invariant to rigid motions. They are invariant to tessellation of the 3D polygonal model, since points are selected randomly from the surface, they are insensitive to small perturbations due to noise, cracks, and insertion/removal of polygons, since sampling is area weighted.

Fig. 2 shows the shape distribution of some geometric shape: the horizontal axis represent the distance between points while the vertical axis represent the probability of that the distance between two points on the surface.

The last step is the comparison between shape histograms. Osada et al. [28] first proposed an approach that worked directly on the original polygons of a 3D model. They empirically tested five shape functions and concluded that the D2 function results the best shape classification method. D2 measures the distance between two random points on the surface of the model. They executed a series of shape matching experiments with a wide database of 3D models.

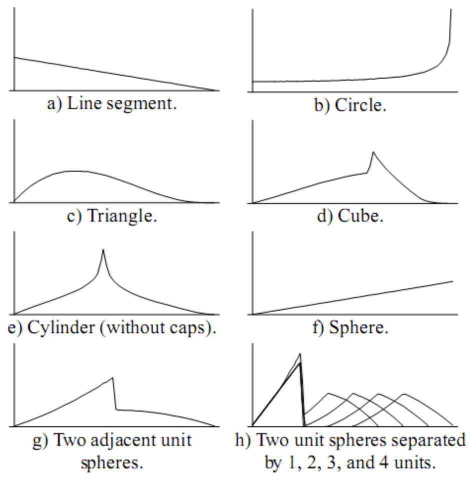

Fig. 2. Shape distributions for standard shapes [28].

Starting from a solid model S, the triangularization T is not unique for S. They assumed that the maximum distance from any point t_i to the nearest point on the model is $\leq \varepsilon$. The facets of T could be obtained from a mesh generation algorithm, or from common exporters, such as STL or VRML, of the model S. The use of a triangularization T occurs to avoid problems related to the underlying representation used in the solid model. Their results demonstrated that shape distributions can be effective at discriminating between groups of 3D models, with the 66% accuracy. In the basic approach the histogram maps distance probability versus measured distance by counting how many distances fall into fixed size bins. This way it is possible to discriminate between shapes that have similar gross shape properties, but it could be very difficult to manage detailed shape properties. Regli et al. [29] observed that over the entire set of sample points three non-intersecting groups could be introduced, besides the ALL points group: the IN group, the OUT group and the MIXED group, depending on the line connecting the 2 points that lies completely inside the model, or completely outside, or both partially inside and partially outside respectively. The shape distance can be calculated for each group of sampled points. They demonstrated that D2 distance, calculated from two solid models very different in shape, could significantly change depending on the sample points group considered. Further developments have been proposed by Ohbuchi [30] who proposed two enhanced shape descriptors: the AD (Angle and Distance) that measures both distance between pairs of points and the inner product of the surface normal vectors of the triangles on which the pair of points are located and the AAD (Absolute Angle and Distance) which considers the absolute value of the same inner product. A recent paper has been presented by Tang et al. [7] in which a shape feature is proposed that considers not only geometry but also texture for similarity assessment of collections, as minerals for example, that have the same geometric characteristics but different appearances. The feature descriptor in this case is a 2D descriptor that considers as first variable the distance between a pair of random points and as second variable the grey level difference between these two points, converting previously the colourful pictures into grey level picture.

2.1.3 Section Images Based Shape Signatures

These techniques use as signatures the section views of solids. It classifies part drawings into groups, based on characteristics such as L/D ratio, number of holes, etc.., using group technology. Parts based on bitmap of the part drawings can be classified by means of 2D similarity assessment techniques or using neural networks. Lippmann [8] and Khanna [42] introduced the theory of neural nets. Neural network methods are useful to solve numerous problems associated with manufacturing operations.

Stated that environmental conditions such as noise or brightness can affect strongly the correct recognition of crisp input, Lee [31] first utilized the fuzzy set theory in combination with the ANN (Artificial Neural Network). Later, Kuo et al. [9] presented a novel fuzzy neural network (FNN) for clustering the parts into several families, based on the image captured from the vision sensor. The proposed network, which possesses the fuzzy inputs as well the fuzzy weights, integrates the self-organizing feature map (SOM) neural network and the fuzzy set theory. The process of similarity assessment follows five steps:

1) image acquisition, by means of a charge-coupled-device camera;

2) image processing to transform the image brightness to binary values, as shown in fig.3;

3) feature extraction that consists in grouping and segmenting the image in several blocks, calculate the fuzzy normalized interval for each block, calculate the fuzzy number and then to transform the data;

4) pattern recognition, in order to cluster the parts with fuzzy features into several families;

5) the part clustering by means of the fuzzy SOM neural network.

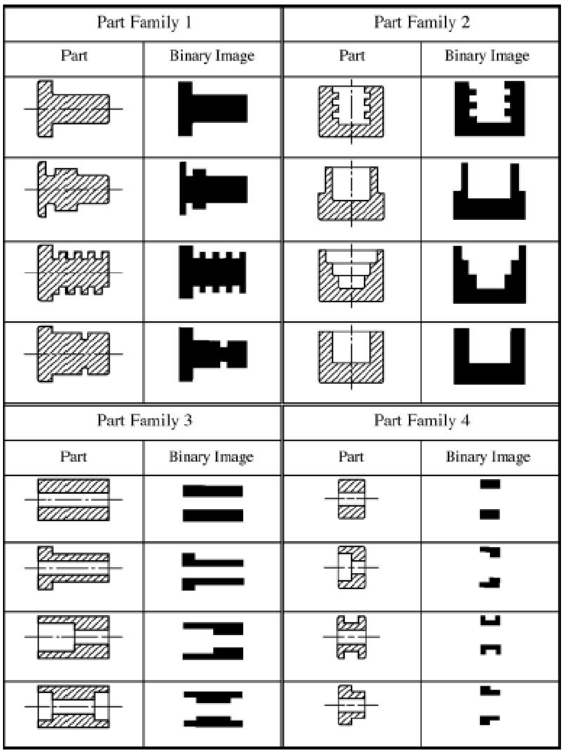

Fig. 3 Some example of transformation the image brightness into binary values [10].

The model evaluation results showed that the proposed FNN can provide accurate decisions. In a more recent study Kuo et al. [10] developed a further novel fuzzy neural network which integrates the fuzzy set theory and the adaptive resonance theory 2 (ART2) neural network for clustering parts into several families and demonstrated that this fuzzy neural network is able to provide more accurate results compared to the fuzzy self-organizing feature maps neural network previously proposed.

2.1.4 Feature Based Shape Signatures

Computer Aided Process Programming (CAPP) is the way the CAD and CAM processes have to be integrated. Starting from CAD data of a part, the goal of CAPP is to generate a sequenced set of instructions to manufacture the part. In this sense the CAPP has to interpret the part in terms of features [11]. In fig.4 a features interpretation is shown: the part is classified by means of the features hole, slot, pocket. In this sense features are here considered machining features.

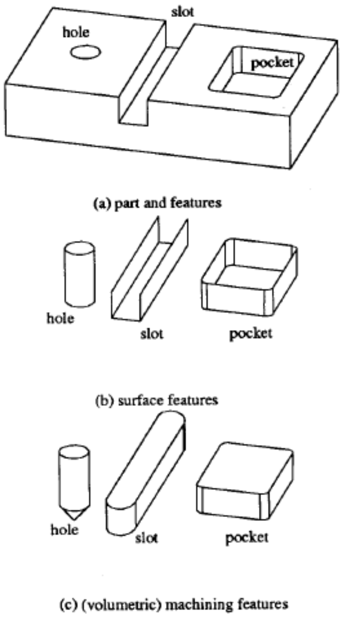

Fig. 4 The part classification by machining features [11].

The feature model is the representation of a part in terms of features: this representation is obtainable either by features recognition, that requires the generation of the machining features starting from customary solid modelling operations, or by feature based design, that uses features already in the design phase. Many CAD systems already use parametric machining features as design features, but the design by manufacturing features approach is not ever the most natural design approach. So, in a lot of cases, it is necessary to convert a design feature model into a machining feature model, even though the difference between them is very fine.

There is a wide state of the art in feature recognition; in this section only the most prominent investigated approaches are proposed: the graph based approach, the hint based approach and the volumetric decomposition approach.

1) Graph based

In the graph based approach, a B-rep model of parts is converted into a graph. The Model Dependency Graph (MDG) is an intermediate data structure through with to model feature interactions and dependencies. The MDG of a query part has then to be compared to the MDG of a part stored in a database and the largest common sub-graph between them has to be determined in order to assess similarity. After performing feature extraction, the Model Dependency Graph representing the features and their interactions is defined. The nodes of this graph correspond to features and store attributes of the features. A graph G=(V,E) is comprises of a set of nodes V= {f0 , . . . , fn}, where fi is a machining feature of the solid, and of a set of edges E. An edge E between two nodes exists if the corresponding features fi and fj have non-zero intersection between them, so:

$$E = \{\{f_i, f_j\} \neg vol(f_i) \cap vol(f_j) \neq 0\} \qquad (1)$$

To compare two solids, the largest common sub-graph (LCS) between the two model dependency graphs needs to be determined. However the problem of exactly determining the largest common sub-graph is NP-complete [43].

In [32] Cicirello et al. propose to solve this problem by means of an iterative improvement algorithm.

The machining feature extraction is first performed to map the solid model to a set of STEP AP 224 machining features; then a UMDG (Undirected Model Dependency Graph) is constructed from the set of machining features, not taking care for the order of the machine features; finally the nearest neighbours to the query graph are found using an algorithm that iteratively searches across a database of other models.

This technique provides means for determining objects having similar machining features. However, the UMDG generated using this method is not unique for a given solid. This is because the features can be constructed in multiple different ways and in different order.

In [33] a graph representation of the input 3D models is presented as the shape signature for the model. For each model a graph called design signature is constructed. The nodes are labelled with a number of parameters, such as type of feature and machining direction. The edges are labelled depending on the type of intersection occurred. An equivalence hierarchy is assigned as the two objects are equivalent with respect to the characteristic considered.

The equivalence relation considered in this application is isomorphism between two graphs: in this case the computation is made easier by the labelling of nodes and edges that allows easier matching of sub graphs.

2) Volumetric decomposition based

In this section, the similarity assessment is reached by the decomposition of the input object into a set of intermediate volumes and then by manipulating the volumes to produce features. As follows, two main procedures are described: the convex hull decomposition and the cell-based decomposition.

Convex hull decomposition, as investigated by Kim [40], consists of multiple steps:

- Alternating Sum of Volumes with Partitioning (ASVP) decomposition. The convex hull of a polyhedron is the smallest convex point set containing *P*. The convex hull difference *CHD(P)* is the regularized set difference (-*) between *CH(P)* and P. If *P* is convex, *CHD(P)* is empty and the decomposition terminates. Otherwise, the decomposition is applied recursively to *CHD(P)*. Fig. 5 shows an example part obtained from the convex hull decomposition. Kim proposed ASVP decomposition and proved its convergence in his dissertation.

- Recognition/Generation of Form Features: in [12] Kim proposed to use the ASVP decomposition to generate form features (Form Feature Decomposition). He named as original the faces of the original part, thus an ASVP component that contains at least two original transitively connected faces is a form feature. Recognized components are further classified on the basis of accessibility. For example, in Fig.5a, the ASVP component P has three original faces transitively connected. It is recognized as a form feature and classified as a slot. Similarly, P is recognized as a rib. However, an ASVP decomposition may return unrecognized components especially when they have at most one original face or separated original faces.

(1)

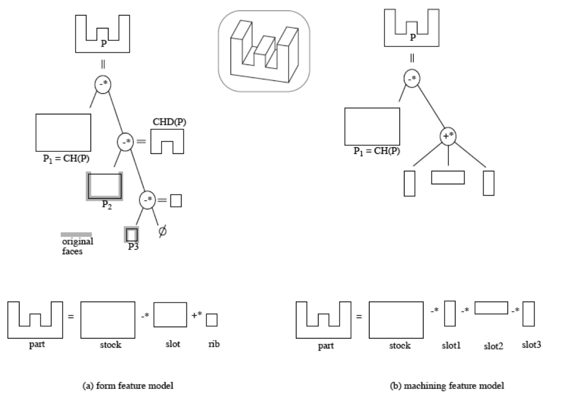

Fig. 5: Volume decomposition: the convex hull [40].

- Generation of primitive machining features and aggregation of machining features; Kim and Wang in [34] proposed a method to generate machining features by rewriting the Boolean expression of every positive FFD using the halfspaces determined by its original faces. The machining features are all negative in the sense that they are subtracted from the workpiece. For example, P in Fig. 5(a) is considered a positive form feature: three halfspaces are created from its original faces. For machining applications, positive-to-negative conversion is applied to convert the FFD into a Negative Feature Decomposition (NFD), where negative features represent removal volumes that provide information about machined faces and tool accessibility. The negative counterpart of each halfspace is intersected with P to generate a new negative component. As shown in Fig. 5(b), three new negative features are obtained and all of them are classified as slots.

Kim et al. in [13] presented a summary of the results conceived by applying the ASVP to obtain a FFD to be further converted into a NFD for machining or cast-then-machined applications.

The cell-based decomposition approach for feature recognition essentially consists of the delta volume decomposition into cells, cell composition and feature classification. Once given a part and its convex hull, the delta volume, that results from the convex hull volume minus the part volume, is decomposed into the cells. These cells are composed in order to generate a volume to be removed by a machining operation, and the resulting volume is classified as a machining feature.

In this approach all possible combinations of cells have to be evaluated to find the best solution. To generate all possible features is very computationally hard, so heuristics are used to prune unpromising compositions, even if this cannot avoid exponential time complexity. In order to reduce the number of interpretations, which may reach N! for N maximal convex cells, Sakurai et al. [14,15] proposed a method to decompose a delta volume into maximal volumes.

The method decomposes an object having planar and/or curved faces into minimal cells and then composes them into maximal volumes which are similar to maximal convex volumes. Maximal volumes are then recognized as features with graph matching.

The cell-based decomposition in feature-based shape similarity assessment by Ramesh et al. [23] describes in details the extraction of features from a B-rep model. The method they proposed consists of two steps: in the first step the part is decomposed into simpler shapes, i.e. cells. They preferred to decompose the part into maximal cells, highlighting that maximal cells output is the minimum possible cells number, so it is closer to a unique decomposition. In the second part cells are mapped to standards machining features. The final step concerns the similarity assessment of two different models.

Seven characteristics are used for comparison. These are feature existence, feature count, feature direction, feature size, directional distribution, size distribution and relative orientation. Feature existence represents the number of different classes of features present in the object and is expressed as a binary vector of dimension n where n is the total number of feature classes in the two objects being compared. Each element in the vector assumes a value of 1, if the corresponding type of feature is present in the object, otherwise it is 0.

Feature count represents the number of instances for every class of feature in a given 3D object. It is expressed as a vector of dimension n. Each element denotes the number of instances of the corresponding feature.

Feature direction represents the number of Translational groups (T-group) for every class and is expressed as a vector of dimension n. T-group is a set of feature instances that differ only by translation. Each element indicates the number of T-groups for the corresponding class. Feature size is similar to feature direction and represents the number of Size groups (S-groups). S-group is a set of feature instances that have the same critical dimensions.

For every class of features, directional distribution represents the number of instances of features within a T-group belonging to the class considered.

Size distribution is similar to directional distribution and is defined for S-groups.

Finally, relative orientation represents the relative orientation between T-groups over all the different classes of features.

A weighted distance is used to compare two objects. The characteristics considered in the comparison have to be independent of each other. Only planar and cylindrical surfaces are considered. Objects where the cylindrical features intersect other faces non-orthogonally are ruled out.

1) Hint based:

The hint based approach is driven from the suggestion that parts should be designed in terms of manufacturing features, allowing this way designers to work in terms of functional features, best concerned to specific processes, thus restricting the ability to select the best manufacturing methods. The hint based approach is not only based on the recognition of machinable features from solid models of parts but also from additional data such as design features, tolerances, and surface attributes (e.g., a thread attribute may be taken as a hole hint). It is important to notice that it does not require that a part be entirely designed by features. In [16] Vandenbrande et al. presented an automatic feature recognizer that decomposes the total volume to be machined into solid features that satisfy stringent conditions for manufacturability, and correspond to operations typically performed in 3-axis machining centers. As machinable volumetric feature it is considered a solid that can be removed in a single machining operation, in a single setup. A single machining operation, however, may consist of several passes by one or more cutters. In [16] the technique proposed takes into account features with intersecting volumes, only conventional machining are considered and non rotational parts that are typically manufactured in 3-axis machining centers, with no care for the process and tool selection, setup planning, cutter-path generation or other important issues in the machining process planning.

In literature several feature finder strategy based on the hint based technique have been proposed; we focus mainly on the Integrated Incremental Feature Finder (IF2) description, giving only some mention to the other Object-Oriented (OOFF) and F-Rex feature finders.

IF2 can recognize holes, slots and pockets. To explain the recognition algorithm a slot example is illustrated. In IF2, a slot trace is generated from nominal geometry when a pair of parallel opposing planar faces is encountered, which corresponds to the slot walls. Given the part shown in fig. 6(d), the vertical inner faces constitute a slot trace.

The geometric completion procedures of IF2 follow a generate-test-repair paradigm [35]. The generate step first finds the slot floor. Only the space between the wall faces is considered, and the part faces that are planar and perpendicular to the wall faces are taken as floor candidates. In fig. 6(d), several floor candidates can be found and the heavily shaded face is an example of it.

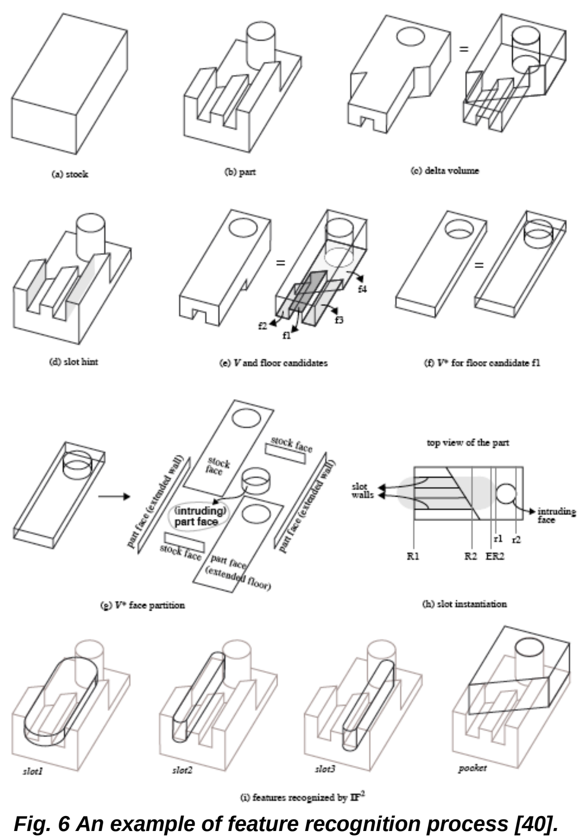

Fig. 6 An example of feature recognition process [40].

Then, the portion of the delta volume between the walls and above the floor, shown in fig. 6(e), is proposed as a volume to be removed by a slot machining operation.

The test step checks the boundary of the proposed volume. The boundary is partitioned into 'stock faces', which originate from the stock, and 'part faces', which originate from the part. 'Stock faces' are those to be removed by feature machining operations, and 'part faces' are those to be created by feature machining operations. For a slot, the proposed removal volume is not machinable as a whole if its boundary contains any 'part faces' besides the walls and floor. This is because such 'part faces' will be removed by the parameterized slot feature volume which completely covers the proposed removal volume. Note that 'part faces' are those to be created. The cylindrical face depicted in Fig. 6(g) is such a 'part face.'

If the test step determines that the volume proposed by the generate step is not machinable as a whole, the repair step removes a subset of the early proposed removal volume, such that the machining operation does not intrude into the 'part face.' This is a geometric fitting problem, and in the example IF2 finally produces the parameterized slot volume shown in fig.6(i).

A problem for hint-based approaches arises when there are to recognize more traces than those to be considered as good features. A trace or hint is nothing but an implication for the possible existence of a feature, and therefore a significant number of traces may not lead to valid features.

IF2 tackles this problem by assigning every trace a heuristic strength. The assigned value is a combined measure of preference for such a feature over alternative feature interpretations and belief that the trace will lead to a valid machining feature.

Once assigned a 3D volumetric model, all the ranked traces are stored in a priority queue. IF2 selects the strongest trace from the priority queue and fires a geometric completion procedure on it. If geometric completion fails to construct a valid machining feature from the trace, the trace is discarded and the next highest-ranked trace is extracted. If completion succeeds, two tasks are done before selecting the next highest-ranked trace: 1) priority queue updating and 2) termination test.

The priority queue is updated to reflect the new feature's influence on other traces, as for example, in cases in which more than one slot can be assembled to create just one bigger slot, so the single previous ones have to be considered as redundant slots.

The termination test is aimed to check for null solids when subtracting a new feature volume from the original delta volume. Once a valid machining feature is found, the priority queue is updated and IF2 updates the material to be removed from the delta volume, recursively. Once the result is null, the process stops because the delta volume is fully decomposed. Otherwise, IF2 takes the new top-ranked trace and repeats the same process.

2.2 The Distance Function

To assess similarity between the two parts, one set of feature vectors is transformed in space using rigid body transformations with respect to the other set such that the distance between them is minimized. The distance between the two sets of feature vectors is used as a measure of similarity between the two parts. The higher the value of the distance, the more dissimilar are the corresponding CAD models.

A *metric space* is a collection of objects along with a distance function δ(), known as the metric, which computes a distance between any two elements in the set. The distance function *δ(x,y)* must satisfy the conditions of positivity, identity, symmetry, and triangle inequality expressed in eq.2, eq.3, eq.4, and eq.5 respectively.

$$\delta(x,y)=0 \Leftrightarrow x=y \quad (2)$$

$$\delta(x,y)\geq 0 \quad (3)$$

$$\delta(x,y)=\delta(y.x) \quad (4)$$

$$\delta(x,y)+\delta(y,z)\geq \delta(x.z) \quad (5)$$

Once the choice of the most suitable shape signature has been made, the main characteristics are determined and they are organized in different forms such as vectors, histograms, matrixes, etc. In order to store data for an efficient search and retrieval, a large amount of research has been put forth to the suitable metric spaces. Depending on the form that the shape signature description assumes, the distance between the two solid models can be computed using LN norm distances, usually L2 norm, such as the Earth Mover's distance [36], the Match distances [17, 18], the Housdorff distance [37], the Minkowski distance [23].

For example, in the graph based signature approach the graph isomorphism is computed to compare feature-graph signatures. Two graphs are isomorphic if and only if the distance between them is zero. As a result, asymptotic computing time will be related to the complexity of algorithms to compute graph isomorphism. It has been classified as an NP-hard problem. In order to solve the graph distance metric problem an approximation algorithms has been proposed to efficiently compute distances between MSG (Model Signature Graphs). In [38] it is shown how it is possible, through the use of graph invariants, to form groups or clusters of potentially isomorphic graphs, or graphs that are close enough in similarity to satisfy a query.

First the eigenvalue spectrum of the graph is computed, then this spectrum is considered to be a projection of the graph from the space of graphs to R^n; thus this space can be used as a basis for most distance computations by means of metrics.

Distance computations between the image of graphs in R^n can be done using any one of the vector-based metric norms. In particular, they investigated the use of L_2 and L_p norms, which obey the positivity, identity, symmetry, and triangle inequality properties. [39]

In other cases the distance function has to be performed ad hoc to better fit the similarity assessment.

For example, in the machining feature based approach, the distance from the query part can be used as a basis for estimating the cost of machining the new part. In this case the key drivers for the machining cost of a prismatic part are identified in the number of setups, tool changes and machining operation cost. In [19], assuming that $p \in P$ and $q \in Q$ are two sets of Reduced Feature Vectors (RFVs) corresponding to parts *P* and *Q,* the distance function is defined in eq.6, as follows.

$$d(p,q)=A+[B\times C]+D \quad (6)$$

where:

$$A=(x_p-x_q)^2+(y_p-y_q)^2+(z_p-z_q)^2$$
$$B=(1-\delta(p,q))$$
$$C=\left[w_V(V(p)+V(q))^2+w_\varepsilon(\varepsilon(p)-\varepsilon(q))^2+w_c(n(p)-n(q))^2\right]$$
$$D=w_T\delta(p,q)$$

Each RFV is represented by using six components, that are x_p, y_p, z_p, V(p), ε(p), n(p). The first three components x_p, y_p, z_p, represent the orientation of the RFV p, and are transformation-dependent. The other three components V(p), ε(p) and n(p) are transformation-invariant and represent the normalized volume of the RFV, the normalized dimensional tolerance and the group cardinality followed to rank ordered the machined parts. Each term in Equation (6) account for some specific parameter that relate to the number of tool setups. Or to the machining operation cost.

In [20] the output of the system, to measure similarity between 3D mechanical parts, is a similarity factor calculated between most similar parts. Entities derived from a STEP (STandard for the Exchange of Product model data) format are mapped into a mathematical model called face-edge relation matrix. STEP format is recognized as the standardized means for product data exchange between different CAD systems to avoid the problems underlying the representation of the solid model. STEP Application Protocol (AP) 224 defines machining features as classes of shapes representing volumes to be removed from a part by machining operations.

The matrix of new design is compared to matrix of the candidate designs in the database and the goal is achieved by finding matrix of a candidate design with maximum matching surfaces corresponding to matrix of new design. Similarity factor between two matrices is calculated as follows:

$$SF\%=\left[\left(2\times F_M\right)/F_{TOTAL}\right]\times 100 \qquad (7)$$

being SF the similarity factor between new design and candidate design. F_M is the number of matched faces in the new and candidate matrices F_{TOTAL} is the sum of number of faces in the both new and candidate matrices.

3 Comparison

After the discussion about the most common methods presented in literature for the shape similarity assessment, in this section a comparison between the different approaches is proposed.

All methods we dealt in this paper have some merits and some drawbacks: the final choice of a suitable method depends upon the relative importance of each functional requirement. Efficient algorithms design for simple topology features often suffer from computational complexity, on the other hand to gain generality efficiency is usually sacrificed. As follows, a detailed analysis is offered in terms of advantages and disadvantages that of each technique and the evaluation of a quite good suitable approach is proposed, in particular centered on the machining features recognition.

Techniques based on Shape Statistics use basic geometric properties in order to perform comparison between solids. These properties are used to represent the signature of the solid. Therefore, such signatures do not carry any topological information about 3D solid models. They are robust but not sensitive to feature location and cannot provide sufficient discrimination power for detailed comparison: on the other hand, they can be useful as quick and efficient filters to reduce the search space.

Histograms based techniques are robust and have no restrictions on the type of solid models to be applied on; they don't satisfy the properties of identity and symmetry because of the random choice of points on the model surface. The accuracy of these signatures depends on the number of points used. Obviously, large number of points result in higher accuracy. On the other hand, the efficiency of these signatures varies inversely as the number of points. Thus with an increase in the accuracy, the computational efficiency decreases. Besides, as objects tend to become more complex, shape histograms tend to become more similar each other.

Image based technique practically classifies 2D part drawings, thus not taking into account rotation neither translation of the solid model.

In [21] Han et al. proposed a detailed analysis of feature recognition between the same parts by means of all three different approaches: the graph based, the volume decomposition based and the hint based.

The main advantage of graph based recognition method is to be applicable to a big variety of domains and not just to machining. Moreover, it allows the user to add new feature types without changing the code. It is robust, but not sensitive to changes in features location and, furthermore, it does not create a unique graph for a solid model because features can be constructed in multiple ways. Anyway, the main problem of this approach is the inability to recognize features intersections inside the model. When features intersect, many of the faces of a features may be entirely absent or partially missing so, stated that in a complex part the possible types of feature intersections can be unlimited, the resulting pattern is week in recognizing intersecting features. Several attempts to restore the missing arcs into a part graph have been made; some good result have been presented in [22] with the implication of restricting the geometries of parts, constrained to be polyhedral and iso-oriented. Last but not least, because of the computational complexity of general subgraph matching is exponential [32], graph based methods can be slow unless some kind of partitioning is done or hint is used as a starting point.

On the contrary, volume decomposition-recomposition approaches can handle all types of interactions in a general way. The main advantage of these methods is that several candidate feature sets can be generated to suit the requirements of an application.

Among the volume decomposition-based techniques, both the convex hull decomposition and the cell-based feature recognition approach are based on multiple-step reasoning: cell decomposition, cell composition and feature classification and each step is completely separated from the others. For example, the initial steps, i.e. ASVP decomposition in the convex hull decomposition algorithm and delta volume decomposition into cells in the cell-based decomposition algorithm, are done independently of features and manufacturing process rationale. The consequence is that no robust method, justifiable from a manufacturing point of view, has been developed to manipulate the intermediate volumes created by these initial steps.

More in particular, in the convex hulls approach it is difficult to compute curved bodies; this problem has been overcome by using polyhedral conversion of curved surfaces in the decomposition. Moreover, alternative feature interpretations are generated by growing and aggregating the basic features recognized, but no success in finding a suitable machining feature model is guaranteed.

The same way, the cell-based approach has the difficulty of how to combine the cells to produce suitable features. Besides, the cell combination, that offer a large amount of possible features, can be interpreted as a power set as well as responsible of exponential time complexity.

However, a big shared disadvantage is that both methods involves Boolean operations and thus they are computationally intensive and limited to analytical surfaces.

The *Hint-based approach* was devised to avoid the computational complexity of other methods by by-passing the cell decomposition phase and going directly to produce maximal volumes that had a better chance of being machining features. This approach is more efficient for a small feature library and a small set of hints. Because rules are hard-coded specific to each feature type, extensibility of the system requires changes to the source code. It is not clear how this method will scale up to more complex features and how efficient it will be with large number of features and hints. Certainly the use of feature finder strategies can enhance the methodology capabilities. In fact, for example, the Integrated Incremental Feature Finder (IF^2) avoids unnecessary reasoning as much as possible by focusing on promising traces, even if it does not always generate a desirable interpretation. A big issue in the hint based approach is to the deal with multiple interpretations; the implementation of IF^2 shows an effort for handling the problems of completeness and multiple interpretations.

Therefore usually first using a signature that is computationally efficient but produces few false positives, followed by a signature that is computationally less efficient but able to eliminate false positives could be a good strategy.

4 Conclusion and developments

Once observed that performing evaluation of existing shape signatures is a difficult task, some preamble to conclusion of this dissertation is necessary:

1. Shape signatures are abstractions of 3D shapes so, in any case, their adoption reduce discrimination capabilities.

2. In practice, to select a particular shape signature, an average case run-time for the computation and matching algorithms is necessary. Therefore, the performance of both computation and matching algorithms needs to be evaluated by running them on a wide variety of models of differing sizes and complexity.

3. In order to offer a complete analysis of the comparison, all shape signatures considered have to be tested on a particular application, so their effectiveness has to be assessed based on their success in a particular application.

4. In a given application, usually a single signature will not be good enough to provide both accuracy of comparison and computational performance. A signature that performs high degree of abstraction usually has limited discrimination capabilities. On the other hand, a signature that performs lesser degree of abstraction is usually computationally not very efficient.

From what discussed in the previous sections we can conclude that all ‘‘pure’’ techniques have some advantages and limitations. These considerations have led several researchers to consider hybrid feature recognition algorithms that combine several basic techniques.

Up to now, hybrid methods have been proposed that combines conventional graph based with hint based feature recognition in order to extract first candidate face sets by a graph-matching technique and then to use constraints and rules to determine more closely the recognized feature type.

In the next future works dealing with similarity assessment, our idea is to combine a method that, owing to test the shape similarity between two CAD models that could be very different in shape and dimensions, can give as first step output a reduced research space of gross suitable candidate solutions. This can be reached through the employment of robust techniques able to assess basic geometric properties in order to perform comparison between solids, even if they do not carry any topological information about 3D solid models. To this purpose, an advisable technique is the Shape Statistics based that is a quick and efficient tool to reduce the search space.

Once reduced the features library research space, the hint based technique could occur, in order to assess the features recognition in a more accurate way. The IF^2 is a valid tool that can *generate* good solutions to the features recognition issue. The *test-repair* paradigm could be enhanced by the integration of some heuristic, such as Genetic Algorithms, Neural Networks, Particle Swarm Optimization. Once the features recognition phase is terminated, a model completely characterized by machining operation is ready to be compared to the models collected in the database. Hence, the similarity assessment can be achieved by means of a distance function, that measure the differences among the features previously recognized, or by means of a similarity factor ad-hoc formulated.

After a quite general discussion about the features recognition process, we focused our conclusions on the machining features recognition because this paper was aimed to deal with the optimization of the manufacturing process: with this purpose we are quite confident that the hybrid approach can be the best tradeoff in order to perform results at a good level of accuracy and in reasonable computational time.

References

[1] A. Cardone, S.K. Gupta and M.V. Karnik. *A survey of shape similarity assessment algorithms for product design and manufacturing applications*. Journal of Computing and Information Science in Engineering 3, 2, pp. 109-118.

[2] S. Marshall. *Review of shape coding techniques*. Image and Vision Computing 7, 3, (1989) pp.281-294.

[3] B.M. Mehtre, M.S. Kankanhalli & W.F. Lee, *Shape Measures for Content Based Image Retrieval: A Comparison.* Information Processing and Management 33, 3, (1997) pp. 319-337.

[4] S. Kaparthi and N.C. Suresh. *A Neural Network System for Shape-Based Classification and Coding of Rotational Parts*. International Journal of Production Research 29, 9 (1991) pp1771-1784.

[5] Y.K. Chung and A. Kusiak. *Grouping Parts with a Neural Network*. ASME Journal of Manufacturing Systems 13, 4, (1994) pp. 262-275.

[6] M.C. Wu, S.R. Jen. *A neural network approach to the classification of 3D prismatic parts*. International Journal of Advanced Manufacturing Technology 11 (1996) 325-335.

[7] J. Tang, F. Zhang. *Content based 3D model retrieval for digital museum*, Lecture Notes in Computer Science 3942, (2006) pp. 1371-1376.
[8] R. P. Lippmann. *Pattern classification using neural networks.* IEEE Communications Magazine 27, 11, (1989) pp. 47–54.
[9] R.J. Kuo, S.S. Chi, P.W. Teng. *Generalized part family formation through fuzzy self-organizing feature map neural network*. International Journal of Computers in Industrial Engineering 40, (2001) pp.79–100.
[10] R.J. Kuo, Y.T. Su, C.Y. Chiu, K.Y. Chen and F.C. Tien. *Part family formation through fuzzy ART2 neural network.* Decision Support Systems 42, 1, (2006) pp. 89-103.
[11] J.H. Han, M. Pratt and W.C., Regli *Manufacturing feature recognition from solid models: A status report*. IEEE Trans. On Robotics and Automation 16, 6, (2000) pp. 782-796.
[12] Y.S. Kim. *Recognition of form features using convex decomposition.* Computer Aided Design 24, 9, (1992) pp. 461–476.
[13] E. Wang and Y.S. Kim Form feature recognition using convex decomposition: results presented at the 1997 ASME CIE feature panel session. Computer Aided Design 30, 13, (1998) pp. 983 – 989.
[14] H. Sakurai, Volume Decomposition ad Feature *Recognition Part I: polyhedral objects*. Computer Aided Design 27, (1995) pp.833-843.
[15] H. Sakurai and P. Dave. *Volume Decomposition and Feature Recognition Part II: Curved Objects*. Computer Aided Design 28, 6/7, (1996) pp. 519-537.
[16] J.H. Vandenbrande and A.A.G. Requicha. *Spatial reasoning for the automatic recognition of machinable features in solid models*. IEEE Transactions on Pattern Analysis and Machine Intelligence 15, (1993) pp. 1-17.
[17] H. C. Shen and A.K.C. Wong, *Generalized Texture Representation and Metric.* Computer Vision Graphics and Image Processing 23, 2, (1983) pp. 187–206.
[18] M. Werman, S. Peleg and A. Rosenfeld. *A Distance Metric for Multidimensional Histograms*. Computer Vision Graphics and Image Processing 32, 3, (1985) pp. 328–336.
[19] A. Cardone, S.K. Gupta, A. Deshmukh and M. Karnik. *Machining feature-based similarity assessment algorithms for prismatic machined parts.* Computer Aided Design 38, 9, (2006) pp.954-972.
[20] A. Cicek. *STEP based geometric and topological similarity assessment of mechanical parts*. Mathematical and Computational Applications 12, 3, (2007) pp. 141-150.
[21] J.H. Han, M. Pratt and W.C. Regli. *Manufacturing feature recognition from solid models: A status report*. IEEE Trans. On Robotics and Automation 16, 6, (2000) pp.782-796.
[22] M. Marefat and R. L. Kashyap. *Geometric reasoning for recognition of 3-D object features*. IEEE Transactions on Pattern Analysis and Machine Intelligence 12, 10, (1990) pp. 949-965.
[23] M.Ramesh, D. Yip-Hoi and D. Dutta, *Feature-based Shape Similarity Measurement for Retrieval of Mechanical Parts*. Journal of Computing and Information Science in Engineering 1, 3, (2001) pp. 245–256.
[24] W.C. Regli. *Digital Library Support for Engineering Design and Manufacturing*. ASME Design Engineering Technical Conference, 1999, Las Vegas, DETC99/CIE-9080.
[25] T. L. Sun, C. J. Su, R. J. Mayer, and R. A. Wysk. *Shape similarity assessment of mechanical parts based on solid models.* ASME Design for Manufacturing Conference, September 1995, Boston, MA.
[26] H.J. Rea, J.R. Corney, D.E.R. Clark, J. Pritchard, M.L. Breaks and R. A. MacLeod, 2001, *Part-Sourcing in a Global Market*. International Conference on e-Commerce Engineering, September 2001, Xi'an, China.
[27] R. Sung, H.J. Rea, J.R. Corney, D.E.R. Clark, J. Pritchard, M.L. Breaks and R.A. MacLeod 2002, *Assessing the Effectiveness of Filters for Shape Matching.* ASME International Mechanical Engineering Congress & Exposition, November 2002, New Orleans, LA.
[28] R. Osada, T. Funkhouser, B. Chazelle and D. Dobkin. *Matching 3D Models with Shape Distributions.* International Conference on Shape Modeling and Applications, May 7-11, 2001, Genova, Italy.
[29] W. C. Regli, Y.Ip. Cheuk, D. Lapadat and L. Sieger, *Using shape distributions to compare solid models.* 7th ACM symposium on Solid modelling and applications, June 17–21, 2002, Saarbrücken, Germany, pp. 273-280.
[30] R. Ohbuchi, T. Takei *Shape-Similarity Comparison of 3D Models Using Alpha Shapes.* Pacific Conference on Computer Graphics and Applications, 8-10 October 2003, Canmore, Canada, pp. 293-302.
[31] D.G. Lee. *Preliminary results of applying neural networks to ship image recognition.* International Joint Conference on Neural Networks 2, (1989) pp. 576-579.
[32] V.A. Cicirello, W.C. Regli. *Machining Feature-Based Comparisons of Mechanical Parts.* International Conference on Shape Modelling and Applications, May 7-11, 2001, Genova, Italy.
[33] A. Elinson, D.S. Nau and W.C. Regli. *Feature-based Similarity Assessment of Solid Models.* 4th Symposium on Solid Modelling and Applications. May 1997, Atlanta, GA.
[34] D. Waco and Y.S. Kim. *Geometric reasoning for machining features using convex decomposition*. 2nd ACM Solid Modelling Symposium, 1993, Montreal, Quebec, Canada, pp. 323–331.
[35] R. Simmons. *A theory of debugging plans and interpretations*. 7th National Conference on Artificial Intelligence, August 21–26, 1988, Saint Paul, Minnesota, pp. 94–99.
[36] Y. Rubner, C. Tomasi and L. Guibas. *A Metric for Distributions with Applications to Image Databases.* 6th International Conference on Computer Vision, January 1998, Bombay, India.
[37] D. Huttenlocher, and K. Kedem. *On Computing the Minimum Hausdorff Distance for Point Sets Under Translation*. 6th ACM Symposium Computational Geometry, June, 1990, Berkeley, CA.
[38] D. McWherter, M. Peabody, W. C. Regli, A. Shokoufandeh. *Transformation Invariant Shape Similarity Comparison of Solid Models*. ASME Design Engineering Technical Conferences, 6th Design for Manufacturing Conference, September 9-12, 2001, Pittsburgh, Pennsylvania.
[39] H. Shum, M. Hebert and K. Ikeuchi. *On 3d Shape Similarity*. IEEE Conference on Computer Vision and Pattern Recognition, June 1996, San Francisco, CA.
[40] Y.S. Kim. *Convex Decomposition and Solid Geometric Modelling.* Ph.D. dissertation, Stanford Univ., 1990.
[41] Burbidge J.L. *The introduction to group technology*. John Wiley, New York, 1975.
[42] T. Khanna. *Foundations of neural networks*. Addison-Wesley Longman, Inc., Boston, 1990.
[43] M. Garey and D. Johnson. Computers and Intractability A Guide to the theory of NP-Completeness. W.H. Freeman and Co., New York, 1979.