Assemblability Analysis with Adjusting Cost Evaluation based on Tolerance and Adjustability

†‡Sukhan Lee †Chunsik Yi §Raúl Suárez

†Dept. of Computer Science, University of Southern California Los Angeles, CA 90089-0781

‡Jet Propulsion Laboratory, California Institute of Technology Pasadena, CA 91109

§Instituto de Cibernética, Universidad Politécnica de Cataluña Diagonal 647, 08028 Barcelona, Spain

Abstract

Tolerances affect the assemblability of a product, which in turn affects the cost of the product because of the scrap cost, and wasted time and energy. In this paper, we propose a method to analyze the product in terms of assemblability based on tolerances and adjustability of the parts. More specifically, the product is considered assembled if the real parts can be fit regardless of the deviations the real parts may have from their nominal, because some parts can be adjusted in position after it is assembled due to clearances between the parts. Moreover, there is a cost associated with adjusting the parts because of the adjusting time, fixturing costs, and additional operations. A method is proposed to evaluated the adjusting cost by counting the minimum number of parts that must be moved or adjusted in order to complete the assembly.

1 Introduction

The designers specify the dimensions and tolerances of the parts of a product. These dimensions denote the ideal geometry of parts, whereas the tolerances describe the permitted variations of the specified dimensions or geometries. The tolerance specification is necessary because the manufacturing processes are inherently imprecise and produce parts that vary.

The tolerances may not allow the parts to be assembled successfully because the small deviations can propagate and accumulate in the product. More specifically, an assembly is typically composed of several to many parts that are assembled together through some type of interconnection, such as loose-fit, tight-fit, attachment, etc., between parts' surfaces, called *mating features*. Small deviations can propagate and accumulate through these mating features such that a new part may not fit into the already assembled parts, thus, failing the assembly operation.

Whether or not a part can be assembled into the already assembled parts depends not only on tolerances but also on clearances between the parts. Clearances between two parts may provide the adjustability, called an *adjustable displacement set* or *adjustable zone*, which is defined as a set of permitted displacements of one part with respect to its mated part. The adjustable zone is a function of clearance of two mating feature and the functionality of the parts in the assembly. For instance, a cylindrical peg with its diameter smaller than the diameter of a cylindrical hole can adjust its pose by as much as the difference of the two diameters. However, because of a functional requirement, the bottom of the peg must be in contact with the floor of the hole when they are mated. These permitted adjustable zones provide the adjustability to the parts for successful assembly operations because the deviations due to tolerances can be compensated.

Although the adjustability may compensate tolerances, it comes with a cost associated with moving or adjusting the parts such as the adjusting time, fixturing cost, and additional operations. Thus, it is necessary to compute the minimum number of objects (parts) that must be adjusted in order to perform a successful assembly operation, given a set of toleranced parts and a valid assembly sequence. It must be noted that in some cases, adjusting the parts may not be enough to assemble them successfully because the deviations caused by tolerances may be too large to be compensated by given adjustable zones.

Following assumptions ae used: 1) Parts are rigid components. 2) The clearance between two mating features must be zero or larger. 3) It is always possible to assembly the parts when they are nominal. 4) Real parts are within the tolerance specification.

Related Works

Turner [1] showed that a tolerance specification can be expressed as an in-tolerance region (established by the tolerance limits) of a normed vector space. He developed the methods for tolerance analysis based on finding the relationship between in-tolerance region and in-design region, which is established by the design constraints. Bjorke [2] has proposed statistical approaches to a tolerance analysis based on functional (sum) dimensions of simple and interrelated tolerance chains. His goal was to derive a set of tolerance chain equations which can solve the functional dimensions. Then, these functional dimensions were checked against the given confidence limits.

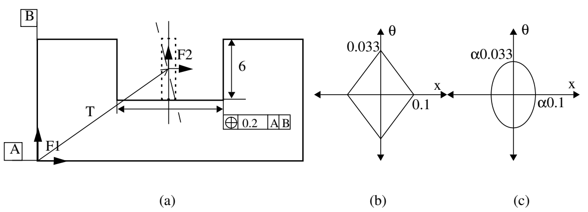

Figure 1: *(a) A position tolerance (ANSI) of a hole feature, (b) the real tolerance boundary in a deviation frame, and (c) an approximated tolerance ellipse.*

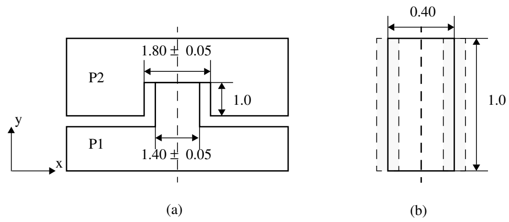

Figure 2: *(a) Assembly of parts P1 and P2 and (b) the nominal adjustable zone and its minimum and maximum boundaries.*

2 Product Assemblability

Tolerances affect the probability of a successful assembly due to the permitted pose and size deviations, as well as the surface and form irregularities, of the parts. Although such deviations are small relative to the the dimensions, they propagate and accumulate in assemblies making them difficult or even impossible to assemble. However, they can be compensated by small clearances between parts which provide the adjustability to already assembled parts to help bring the mating parts to their mating positions, and provide a small zone of possible mating positions for the mating parts instead of a single possible mating position. This probability of assembly is called *assemblability.*

Definition: *The assemblability of a product for a given dimension and tolerance specification is the probability of successfully assembling a set of part instances manufactured within the tolerance specification.*

2.1 Assemblability Analysis

Assemblies consisting of a number of parts have a network of part interconnections, which represents the complex relationships among the parts. More specifically, the parts in an assembly are interconnected through some type of mating conditions such as loose-fit, fight-fit, attachment, etc., using some surfaces or features of parts, called *mating features.* For instance, a cylindrical-peg feature of part P_1 has a loose-fit mating relation with a cylindrical-hole feature of part P_2, assuming the diameter of peg is always smaller than hole.

It is possible that certain combinations of part instances may fail to assemble due to the accumulation of deviations which even the adjustable zones cannot compensate. Thus, it is important to predict the probability of successful assembly of the parts so that the tolerance specifications can be reevaluated and modified if necessary in order to increase the probability and lower the production cost associated with the assembly costs.

A set of parts forming a loop may fail to assemble because of the closed-loop constraint. More specifically, every part in this set must be assembled to its two mating parts, except the base parts which may be fixed, e.g., on a fixture, or attached to other loops. This set of parts forming a loop is called a *parallel chain.* A parallel chain fails to assemble if the last part that completes the loop cannot be assembled because the accumulated deviation of individual parts cannot be compensated by the adjustable zones.

A network of an assembly is usually composed of a set of parallel chains, called a *multi chain.* A multi chain fails to assemble if any one of the parallel chains fails to assemble. Therefore, the assemblability can be computed statistically by computing the probability of assembly of every parallel chains in the product. The computational methods are discussed next.

2.2 Computational Model

It has been already studied by [3] that most of the tolerance types described in ANSI [4] can be statistically approximated by an ellipsoid in a coordinate frame of kinematic parameters, called a *deviation frame* in this paper. In general, a deviation frame has six axes (i.e., three translation and three rotation), and the origin indicates that the feature is in its nominal pose. Centered at the origin of a deviation frame, the ellipsoid approximates the real tolerance. For instance, a point in this ellipsoid represents a permitted deviation of feature from its nominal pose.

For example, a feature (hole) of a part shown in Fig. 1(a) has a position tolerance of 0.2. The meaning of this tolerance is that the axis of the hole (shown by a solid vertical line) is permitted to deviate within the tolerance zone (shown by a dotted rectangle with the width of 0.2.) That is, the coordinate frame attached to the nominal pose of the hole can deviate from its nominal as much as the axis is allowed to deviate from its nominal. This maximum boundary in a deviation frame of two axes (θ and x) is shown in Fig. 1(b), which shows that the coordinate frame attached to the nominal hole can rotate (θ) approximately at most by ± 0.033 and can translate in x-axis (x) at most by ± 0.1. Fig. 1(c) shows an ellipse approximation of the real tolerance boundary. The details of the approximation algorithm using ellipsoid are described in [3] and [5].

An example of the adjustable zone of a peg and hole is shown in Fig. 2(a), which is a rectangular zone

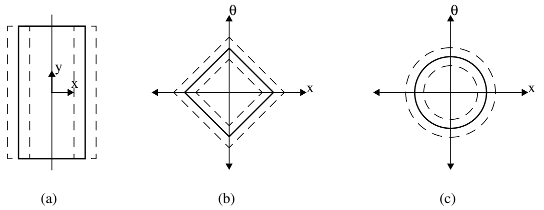

Figure 3: *(a) An adjustable zone of P1 and P2, (b) the real boundary in a deviation frame, and (c) an ellipse approximation.*

of 0.40 by 1.0 when the parts are nominal (i.e., the part dimensions without the tolerances). However, due to the tolerances (± 0.05) of the diameters of the peg and hole, the adjustable zone can vary accordingly within the maximum and minimum limits as shown by dotted lines in Fig. 2(b).

An adjustable zone and its limits are represented by a nominal ellipsoid and its minimum and maximum limits in a deviation frame. For instance, the coordinate frame attached to the center of the axis of the hole of P2, which is aligned with the axis of the peg, is permitted to be adjusted as much as the adjustable zone allows, as shown in Fig. 3(a). The nominal and the limits of the adjustable zone in a deviation frame is shown in Fig. 3(b), where the solid line is the nominal adjustable boundary, whereas the dotted lines indicate the minimum and maximum limits. Fig. 3(c) shows an adjustable ellipse and the limits statistically approximating the real adjustable boundary and the limits of Fig. 3(b). In brief, the distribution of the simulated adjustable zones (the distribution of diamond areas) is optimally approximated by an analytical distribution that can closely approximate the simulated distribution of the real adjustable zones. This analytical solution (Gaussian-Sigmoid Function [5]) has a flat-top bell shape whose minimum (the top) and maximum (approximately $\pm 3\sigma$) can approximate the limits of the adjustable boundary limits. The details are described in [5].

2.3 Parallel Chain Computation

The statistical ellipsoids of the tolerance and adjustable zone of a parallel chain, as well as its assemblability, can be computed in the following way: (1) Identify the last part to be assembled in the parallel chain. (2) From the base(s), compute the ellipsoids of tolerance and adjustable zone of each chain, called a *serial chain*, formed by cutting the parallel chain at the last part. This cut assumes that the last part is been assembled to only one of the serial chains, and computes the accumulation at the mating features of each serial chain. (3) Numerically compute the new ellipsoids of the tolerance and adjustable zone, as well as the assemblability, of the parallel chain.

The computation of ellipsoids of tolerance and adjustable zone of a serial chain in step (2) requires the addition and propagation of ellipsoids of individual mating features in the serial chain. This addition operation is called a *tolerance sweep operation*, which adds all combinations of randomly and normally generated samples (points) of the two ellipsoids. Then, the distribution of the addition is optimally approximated by an ellipsoid using an analytic solution, i.e., Gaussian function. The tolerance ellipsoid of a serial chain is the addition of all the individual ellipsoids in the chain.

The sweep operation of two adjustable ellipsoids approximates the distribution of the sweep area of two randomly selected adjustable ellipsoids from each given nominal adjustable ellipsoids and their limits. More specifically, (i) Randomly generate an ellipsoid from each nominal adjustable ellipsoids and their limits. This can be done by randomly and normally generating a value within the limits of the ellipsoids axes, so that randomly generated ellipsoid lies within the limits. (ii) Compute the sweep volume of the two randomly generated ellipsoids. (i) Repeat the above two steps many times, and accumulate the sweep volumes. (iv) Optimally approximate the distribution of the accumulated sweep volumes using an analytical solution describe in the previous subsection.

The assemblability and the ellipsoids of tolerance and adjustable zone of a parallel chain can be computed from the solutions of the two serial chains. Note that the ellipsoids of tolerance and adjustable zone of each serial chain denote the probability distribution of tolerance and adjustable zone of the serial chain. Therefore, the assemblability can be computed by randomly and normally selecting (generating) a sample point from each tolerance ellipsoid, and checking whether or not the randomly selected adjustable ellipsoids located at the tolerance sample point intersect. The intersection implies that the two serial chains can be assembled for the selected samples or instances. The assemblability of the parallel chain is the total number of success divided by the total number of trials. For all the intersection volumes, the distribution of the centers and the distribution of the volumes, whose centers transformed to an origin of a adjustable deviation frame, form the tolerance and adjustable zones of the parallel chain. These distributions can be approximated by the ellipsoids of tolerance and adjustable zone as described in the previous section.

2.4 Multi Chain Computation

A multi chain represents a network of two or more parallel chains. An example of a multi chain is shown in Fig. 4, where there are six parts, P1, $\cdots$, P6, and seven parallel chains, 1, $\cdots$, 7. The parallel chains are shown by dotted loops in the figure.

The assemblability of a multi chain is computed by iteratively solving each parallel chain of the multi

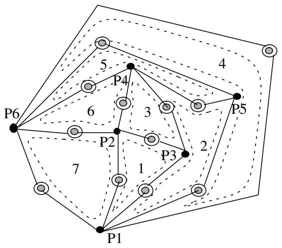

Figure 4: *A multi chain with seven parallel chains*

chain. The order of parallel chains to be solved can be determined from the given assembly sequence, or by the users (designers), as long as the next selected parallel chain is connected to the already processed (solved) parallel chain. When the solution of a new parallel chain is computed, at a selected part, its effect is propagated back (back-propagation) to other nodes of the already solved parallel chains. This is because the completion (or assembly) of a new parallel chain can affect the previously computed solutions of other nodes. That is, when a new parallel chain is assembled to the currently assembled parts, the tolerances and adjustable zones of the parts may no longer be the same. The algorithm ends when the last parallel chain in a multi chain is solved.

In Fig. 4 for example, when the P-chain 1 is selected, the pool of nodes under consideration is {P1,P2,P3}, as shown in Table 1. The parallel chain is solved at P3, which is underlined and bold faced in the table. After the computation of tolerance and adjustable zone of P3, the result back-propagates to P1 and P2. When the P-chain 2 is considered, the part nodes attached to it, e.g., P4 and P5, are added to the pool of nodes to consider. There are P1, P2, P3, P4, and P5 which are attached to the future P-chains, i.e., 3,4,5, 6, and 7. P4 is selected as a node for which P-chain 2 is solved. When the solution of P4 is calculated, its solution back-propagates to P1, P2, P3, and P5. This process continues, and can be traced as shown in Table 1.

The computational complexity of the algorithm is provided in terms of the number of Intersection Operations (**IOs**), as in the parallel chain, and Sweep Operations (**SOs**). In the worst case, the number of **IOs** is $P * n$, where P is the number of part nodes that join two or more parallel chains, and n is the number of P-chains in the multi chain (M-chain.) The number of Tolerance Sweep Operations (**TSOs**) performed by the algorithm is $n(P-1)+2M$, and the number of Adjustable Displacement Sweep Operations (**ASOs**) is $n(P-1)+M$, where M is the number of mating nodes.

Table 1: *Multi chain execution steps for the example in Fig. 4*

P-chain Order	Pool of Nodes under Consideration	Pool of Nodes for Future P-chains
1	{P1, P2, P3} = S1	{P1, P2, **P3**}
2	{P4, P5} ∪ S1 = S2	{P1, P2, P3, **P4**, P5}
3	S2 = S3	{P1, **P2**, P4, P5}
4	{P6} ∪ S3 = S4	{P1, P2, P4, P5, **P6**}
5	S4 = S5	{P1, P2, **P4**, P6}
6	S5 = S6	{P1, P2, **P6**}
7	S6 = S7	∅

3 Cost Evaluation

This section presents a method to deal with a parallel chain of parts that can be systematically extended for the case of a manipulated object that links a larger number of parallel chains.

Let M be the manipulated part, the last part that will link two serial chains, and P_1 and P_2 be the two last parts of the two serial chains that have to be assembled with M. The number of parts N={0,1,2,3} that has to be adjusted to be able to perform the assembly as well as which are these parts (P_1, P_2, and M), allows eight different solutions according to the different sets of parts to be adjusted: $\emptyset$, $\{P_1\}$, $\{P_2\}$, $\{M\}$, $\{M, P_1\}$, $\{M, P_2\}$, $\{P_1, P_2\}$ and $\{M, P_1, P_2\}$. The algorithm to compute the statistical occurrence of each one is described below:

Algorithm: Adjustability Cost

1. Compute the ellipsoids of tolerance (T_i) and adjustable zone (D_i) of two serial chains, i=1,2.
2. Set num_α=0 for $\alpha = \emptyset$, $\{P_1\}$, $\{P_2\}$, $\{M\}$, $\{M, P_1\}$, $\{M, P_2\}$, $\{P_1, P_2\}$, and $\{M, P_1, P_2\}$.
3. Repeat n times (n large enough):
   (a) Randomly generate an adjustable ellipsoid L_i between M and P_i.
   (b) Randomly generate a sample $\mathbf{p}_i$ from the T_i and D_i. $\mathbf{p}_i$ denotes an possible pose of a serial chain i.
   (c) Randomly generate a sample $\mathbf{m}_i$ from the mating feature i of M.
   (d) Compute the minimum number of parts to be adjusted.
   (e) Increment num_α according to the result returned in previous step.
4. Return $\frac{\text{num}_\alpha}{n}$ for each α.

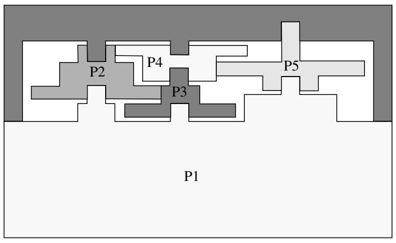

Figure 5: *Example assembly with six parts*

Table 2: *Simulation results of the multi chain*

	Par1	Par2	Par3	Par4	Par5	Par6	Par7	A
Sim1	0.997	1.0	1.0	0.277	1.0	1.0	0.513	0.1417
Sim2	0.997	1.0	1.0	0.085	0.922	0.999	0.618	0.0482
Sim3	0.996	1.0	1.0	0.184	0.945	1.0	0.949	0.1644
Sim4	0.997	1.0	1.0	0.126	0.895	0.999	0.708	0.0795
Sim5	0.989	1.0	1.0	0.451	0.993	1.0	0.585	0.2591
Average								0.1386

4 Simulation of Assemblability

An assembly of six parts, shown in Fig. 5, is used in the simulation. The parts are two dimensional. We used two dimensional coordinate frames with x and θ axes, where x denotes the translation in x axis and θ denotes the rotation about the axis perpendicular to the plane of the paper.

The graph of the assembly is shown in Fig. 4. It has seven parallel chains, 1-7. This multi chain is solved by solving the parallel chains in the sequence of 1 to 7. As explained before, the i_{th} parallel chain depends on the solution of the sub-multi chain composed of parallel chains 1 through $i-1$.

The result of the five simulation runs is shown in Table 2, where Par_i denotes the parallel chain i, Sim_j denotes the j_{th} simulation run, and A denotes the assemblability of the product. This simulation result shows that $Par4$ has very low probability, which leads to very low probability for the whole assembly. The probability of $Par7$ is also low.

The assemblability of the example assembly is low mainly due to the small clearance that was assigned to the mating features between the parts P3 and P4. The peg diameter of P3 is 1.63 and the hole diameter of P4 is 1.67, whereas other peg and hole mating features have diameters of 1.40 and 1.80, respectively.

5 Simulation of Cost Evaluation

The four part example assembly shown in Fig. 6 is used in this simulation. The bottom part (P1) has two holes where the left shaft (P2) and the right shaft (P3) are inserted into. The top part (P4) with two holes is assembled to the two shafts. The holes of part P1 have diameter of 1.0 with size tolerance of ± 0.01 and position tolerance of 0.10. The holes of part P4 have diameter of 1.02 with size tolerance of ± 0.01 and position tolerance of 0.10. Note that the holes of P4 have diameter slightly larger, by 0.2, than the holes of P1. The shafts, P2 and P3, have diameter of 0.97 with size tolerance of ± 0.02. Size and position tolerances are used in this example to make it simple to illustrate simulation results.

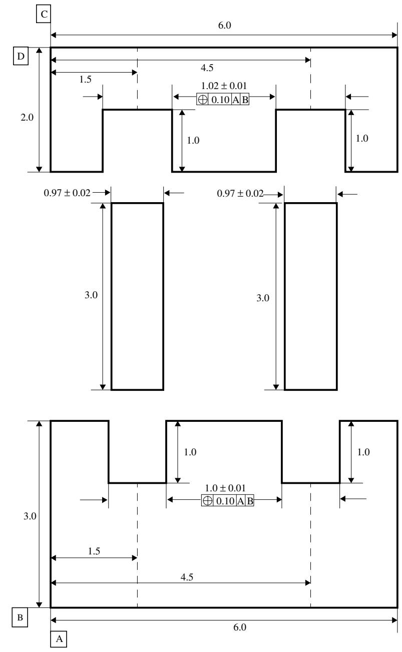

Figure 6: *ANSI dimension and tolerance specification for the four part 2D example assembly: 1 Bottom (P1), 2 Shafts (P2 and P3), 1 COver (P4).*

The simulation counts the number of objects that must be moved in order to successfully assemble the parts. Therefore, it assumes that the cost is the same for moving the same number of objects, although they may be different, e.g., moving two objects {P1,P2} or {P1,M} costs the same. It must be noted that the algorithm described in the previous section accounts for all possible results including different objects but the same number of objects.

Two assembly sequences were tested and compared in the simulation: (1) {P1,P2,P3} $\cup$ {P4} and (2) {P4,P2,P3} $\cup$ {P1}, where {.} denotes a subassembly and $\cup$ denotes the assembly operation. For example, the subassembly {P1,P2,P3} has been already assembled, and the next assembly task is to assemble {P4} to {P1,P2,P3}. It must be noted that the subassemblies {P1,P2,P3} and {P4,P2,P3} are serial chains, thus, the cost of assembling them has been ignored in this simulation. We show the cost of assembling a parallel chain.

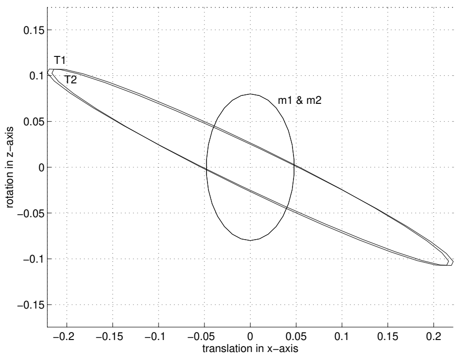

Figure 7: *Accumulated tolerance ellipses at P2 (T1), P3 (T2), left hole of P4 (m1), and right hole of P4 (m2).*

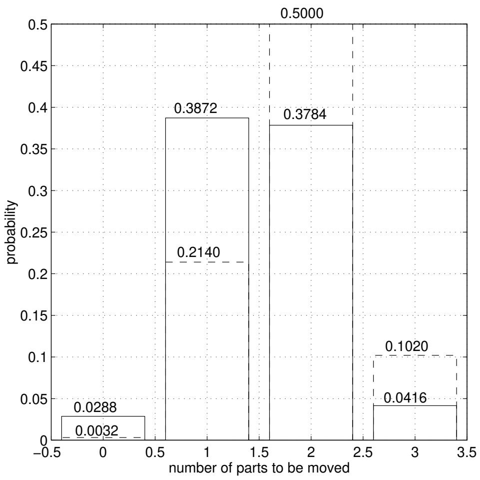

Figure 8: *Bar graphs of probability of moving N = {0, 1, 2, 3}. number of parts for assembly sequence 1: {P1, P2, P3} ∪ {P4} (dashed-lines), and assembly sequence 2: {P4, P2, P3} ∪ {P1} (solid-lines).*

In simulation (1), ellipses of tolerances and adjustable zones, shown in Fig. 7, are computed at the upper end of shafts, P2 and P3, propagated from P1. Note that although the deviation frame requires three axes, y component has been ignored since it affects little the results and illustrates better the example. **T1** and **T2** are the accumulated tolerance ellipses at the tip of the left shaft (P2) and the right shaft (P3), respectively. **m1** and **m2** are the tolerance ellipses of the left and right holes of the top part (P4). The solid-line bar graph in Fig. 8 shows the probabilities of moving N = {0, 1, 2, 3} number of objects for the sequence. In simulation (2), the ellipses of tolerances and adjustable zones are computed at the lower end of the shafts, P2 and P3, propagated from P4. The dashed-line bar graph in Fig. 8 shows the probabilities of moving N number of objects for this sequence.

These results show that the assembly sequence (1) can be assembled without moving an object about 2.88% of the time, and one object about 38.72% of the time. However, the assembly sequence (2) can only assemble 0.32% without moving any objects, and 21.4% by moving one object. By assuming that the cost is directly related to the number of objects to be moved (e.g., fixturing cost), then the assembly sequence (1) is better than the assembly sequence (2).

6 Conclusion

A statistical method of computing the assemblability of a product is proposed based on solving parallel chains iteratively. In addition, the product can be statistically evaluated in terms of adjusting cost. In assembly processes, adjusting parts can increase the cost of the product due to additional fixtures or holding devices, extra time, etc. The predictability of the product assemblability allow the designers to better design the tolerances, and decrease the cost involved in assembly process.

The expected contributions of this paper are: (1) a computer-aided tool for analyzing the product assemblability can be built to help the designers check the tolerance specification, (2) a new cost metric is proposed based on the concept of adjustability of assembly, and (3) assembly sequences can be selected using this metric, thus bringing the assembly sequence evaluation closer to more realistic problem.

References

[1] J. U. Turner. Tolerances in computer-aided geometric design. *Ph.D. Thesis, Rensselaer Polytechnic Institute*, 1987.

[2] O. Bjorke. *Computer-Aided Tolerancing.* ASME PRESS New York, 1989.

[3] D. E. Whitney and O. L. Gilbert. Representation of geometric variations using matrix transforms for statistical tolerance analysis in assemblies. *1993 IEEE International Conference on Robotics and Automation*, pages 314–321, 1993.

[4] Dimensioning and tolerancing. *ANSI Y14.5M, American Society of Mechanical Engineers, USA*, 1982.

[5] S. Lee and C. Yi. Statistical tolerance and clearance analysis for assembly. *Proceedings of the 1996 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 688–695, Osaka, Japan, 1996.

[6] P. Ghosh. A mathematical model for shape description using minkowski operators. *Computer Vision, Graphics, and Image Processing*, 44(3):239–269, 1988.