A New Perception Mechanism for Virtual Environments

Cheng Cheng[1], Duan Qingling[1]
[1]*Department of Computer Science, Beijing Institute of Technology, Haidian Zhongguancun South Street No.5, Beijing , China*
cc@bit.edu.cn

Keywords: Human-computer interaction, percpetion, direct manipulation, collision detection, intent, virtual assembly.

Abstract: The current virtual environment uses a unique collision detection sensing mechanism. This kind of perception is a kind of super-high-dimensional perception. The high time complexity of the algorithm can't meet the needs of human-computer interaction in the desktop virtual environment. A more important issue is that it does not provide high-level human-computer interaction semantics. This paper studies a new sensing mechanism with distance calculation as the kernel. Firstly, it is oriented to the highly demanding virtual manufacturing environment, but it is not limited to manufacturing virtual environment applications. The composition and operation mechanism of the new perception mechanism are given. Preliminary experiments in its own desktop virtual assembly system verify its feasibility and efficiency. Some theoretical comparison results are obtained.

1 INTRODUCTION

The experience of virtual assembly tells us that there are many shortcomings in the current status and research of virtual assembly, which is far from industrial application and popularization. The reasons are manifold. From our perspective, the important point is that the perception mechanism is insufficient. Perception is a key infrastructure for future virtual assembly and even the entire virtul environment. At present, the virtual assembly and the entire virtual environment adopt the classic collision detection algorithm as the sensing mechanism. It has several technical problems, such as the short analysis given below.

The first point is that the algorithm for collision detection is too complex, and it has to be run on high-performance workstations, and it also needs to optimize the collision detection of multiple scenes. The collision detection algorithm detects whether there is an overlap in the space between two objects in the space. If so, where is the contact point? The purpose of collision avoidance is to avoid any collision between objects. Collision detection exists and is used as the only way to perceive it. It is discussed in many papers and systems (K and M, 2004) (C, 2004).Secondly, the most important problem existing in perception is not the problem of high complexity of the algorithm, but the problem of lack of semantics.From the patch model used for collision detection to the content and structure of the result information of the collision detection, it is difficult to meet this requirement. Therefore, a new sensing mechanism is urgently needed to meet the needs of such semantic information acquisition. Those patches for rendering have not semantic information genes. Collision detection is an object behavior oriented and rendering oriented computation.It is done for the machine. While perception is user-oriented, it is done for human perception and cognition.Third, the result of collision detection is a collision event, which does not meet the requirements of the highly demanding manufacturing application of virtual assembly. For example, we have to do a certain gap assembly during the assembly process. There is a certain distance between the assembly features. There is no collision. We need to perceive this distance. This is distance perception rather than collision perception. For another example, we need an interference fit, which can't be done with collision detection. It is impossible to assemble parts under these conditions. Therefore, it is urgent to propose a new cognitive basic mechanism to replace the existing collision detection to deal with various complex assembly situations.

The human intent in the virtual assembly interaction can be captured, which helps us to obtain real-time semantic-rich perception. In the multi-channel interaction condition, the intent is reflected by multi-channel input. The research in this area is:(Varol H A and M, 2010) The user's intent is obtained by patterning and inferring the behavior data of the collected user. (D, 1999) This is a paper on the study of eye movement channels. It acquires eye gaze tracking through the hidden Markov model to analyze techniques for inferring user intent.(P and S, 2005) also introduced how to obtain user intent through eye movement mode to serve WIMP interface applications.(Schnell et al., 2009) Introduced the cognitive pilot helmet of the University of Iowa study. The prototype system acquires the current state and intent of the pilot through a number of human sensors and eye tracker acquisition information analysis.

The researches on semantic construction are: The semantics of the virtual environment is constructed by establishing the ontology of the virtual environment and the three-dimensional objects(L and Y, 2008).(Tutenel T, 2008) said that, The current virtual environment lacks semantic constructs. Adding rich semantic data to systems and objects will bring unprecedented energy to virtual environment interaction and simulation. Semantic construction automatically performs scenario conversion, which plays a huge role in the construction of virtual environment software.(Xiaoming Z and Z, 2008) uses a lot of static semantics in virtual assembly applications. The article is about how to construct the important application logic of assembly planning with semantics.(L and M, 2007) is a very classic article, it describes how to combine artificial intelligence and computer graphics to construct a smart virtual reality.

The feature model is the basis of semantic generation for a critical system. (Hamidullah and A, 2006) studied the classification and identification methods of assembly features, which is very meaningful analysis and induction.(S and K, 2006) is a good review of the research results of extracting design features. This paper fully demonstrates the availability of the feature model, and the feasibility ,the huge requirements and the potential of the extremely high virtual manufacturing system semantic construction .(H and C, 2007) did a very useful job. Their work shows that XML-based product design systems can enable the semantic transfer and sharing of heterogeneous platforms.

Regarding the distance calculation, we can see that: (Tang M and J, 2009) uses Hausdorff Distance to measure the similarity of two object shapes and calculate the penetration depth in physics animation. (Sud A, 2006) used a concept of distance field. For a set of objects in a deformable space, the distance field can be used to calculate the proximity between a hard and deformable model.(Hsu D, 1997) (Gilbert E G, 1988)(S, 1994) (E and P, 1992) are research papers on distance calculation in robot path planning. These are typical works of perceptual calculations using object geometry information.

The work on the thesis is that, based on the above technical review, the authors present a perceptual algorithm with distance as the core. At the same time, a new set of perception mechanisms is given. The mechanism was verified and validated by our virtual assembly system construction and virtual assembly experiments. Here are some example descriptions and some analysis results. Finally, some technical summaries and prospects for future work are given.

2 DISTANCE CENTERED PERCEPTION ALGORITHMS

This paper proposes a finite dimensional perception mechanism as a new perception mechanism for virtual environments and virtual assemblies.Polygon's position and normality in space is disorganized and has no semantic information. This is like an infinite dimensional measure, so we can think of collision detection as a relative infinite dimensional perception mechanism. Can we turn this problem into a finite dimension of perception?The following explanation is now given for finite dimensional perception: A perceptual dimension is a collection of perceptual items with the same attributes. Any one of all perceived dimensions is irrelevant to other dimensions. Both dimensions are independent. None of them can be expressed by other combinations of dimensions. Based on the feature representation of the object, we determine the perceived dimensions based on the auxiliary features and surfaces of the feature of the part object. Regarding the preliminary definition of the perceived dimension, the determination of the actual perceived dimension is based on the total number of assisted feature pairs between the objects, and is not determined solely by the characteristics of one object. With regard to the definition of auxiliary features, the auxiliary features are some parameter descriptions of the features, including the source point of the feature, the three local axes of the feature Major, Minor and 3thAxis, and the reference points and normals of the key surface and end face.

First, the strategy of generating perception dimension is given. As mentioned above, each of the auxiliary feature on a matched feature pair will be a potential perception item. When we determine the matching relationship of a set of assembly features, we generate a set of perceptual items for each matched feature on the current part.Each perceptual item has a directed line segment to represent it in the space on the target feature. All the same perceptual directed line segments of a feature on the part form a perceptual dimension. All dimensions together form the entire percetion dimensions of the current part. Whenever we add a matching feature pair, we need to add a set of perceptual items to this feature pair. Whenever a matching feature disappear , the corresponding set of perceptual items will be subtracted from the perceptual dimensions.

Since there are many features on a part, it is necessary to clearly identify the perception dimensions from all possible dimensions. This is the dimension reduction strategy. This is done in such a way that three elements are used for dimension reduction. One is the situation, the second is the movement constraint or the direction of movement, and the third is the interaction intent . First, the potential perception dimension is determined according to the scenario, then the dimensions are further reduced according to the motion constraint; finally, the perception dimensions are minimized according to the user's intent.

Each dimension is measurable in the perception process. For a dimension, we create a coordinate rule for it in space and calculate the distance between two auxiliary features in real time. We will project different perceptual items in the same perceptual dimension onto the perceptual coordinate scale corresponding to the perceptual dimension. There are multiple distance descriptions of the perceptual items of matching feature pairs on the scale on one dimension. The distance descriptions of these perceptual items constitute the perceptual distribution in the perceptual dimension. We define a concept of primary perception in one dimension and then determine other perceptual content as secondary perception. One of the many perceptual items in a dimension is critical, and this is the perception of a pair of assembly features in an assembly. After the primary perception is determined, the more secondary perceptions, the better it is. Each perception corresponds to a pair of matching features. The more matching pairs, the better it is. The more subperceptions that coincide with the main perception in the same dimension, that is, zero distance, the better it is. If there is a negative distance when another coincidence occurs, the penetration phenomenon occurs and needs to be expressed.

Then discuss the laws of perception in different dimensions. First, the percetion representations in multiple dimensions are independent, and the perception of each dimension does not affect each other. The second characteristic of the percetion dimension is time invariance. That is to say, in a short period of time, the percetion dimensions do not change, and the composition of the dimensions is constant. The third characteristic of the percetion dimension is having a major dimension. Not all dimensions of perception have the same importance. Only one dimension is the major dimension at any one time period, and the others are minor dimensions. The importance of these dimensions is reflected in the difference of degree of user's attention to each dimension. The major dimension is the current focus of the user, and the other dimensions play the role of confirmation and verification. The fourth characteristic of the perception dimension is spatial contraction. During the interaction process, the distance distribution in the dimension combinations will shrink continuously in space. That is to say, the perceptual items represented by the distance distribution in the dimensions will spatially converge to a few geometric features. This is the meaning of spatial contraction.

It is necessary to propose a new object representation model for human-computer interaction.This is an important geometric basis in terms of perception. A real-time calculation of various complex relationships in space based on feature geometry information is the perceptual algorithm we need. With regard to data, we organize it in a multi-resolution data structure based on the inherent structural characteristics of the part. The so-called multi-resolution is the different granularity, from a visual point of view. We use different resolution data in different interaction phases, and the finest granularity is the polygon data of the geometric features of the virtual object. This together with its parameter data constitutes the data structure content of the virtual object.We obtain the representation of the feature tree of the part as one input , and then obtain the parameter set of each feature, including the feature source $Origin$,the three axes of the feature, $Major$, $Minor$, $3thAxis$. Features refer to specific geometries, which can be summarized into more than a dozen geometric shapes, including convex features such as cylinders, cuboids, etc., concave features such as holes, grooves . For each feature, we also need to obtain some of their feature faces . The other is to get the geometric parameters of the features.

The processing strategy is feature matching. The so-called feature matching strategy is to determine a pair of assembly feature pairs between two parts, and then apply constraints to the parts so that the matched feature pairs can be aligned for assembly. The strategy of feature matching is to repeatedly apply the feature matching method to the part to adjust its direction. In this way, the overlap and collision problems between the two features without the assembly relationship will not occur after the constraint is gradually added. The only thing left is how to complete the assembly of the parts in place through the face-fitting perception. This problem has not been completely solved. The reason is that when the design is unreasonable or even wrong, how can it be discovered and exposed during the interactive assembly process

. In other words, it may be reasonable to assume that the design of the part is reasonable, so that the movement of the object does not occur under the given strategy. Thus, the final step is the face-fitting perceptual calculation in the case of various constrained motions. Therefore, we say that face-to-face perception is the core algorithm in our proposed perceptual mechanism.

The distance-calculated core perception has such a structure, gives the distance calculation formula of each relevant perceptual dimension, and then gives the semantic judgment formula of the distance of each perceptual item for the corresponding semantic object type, and finally gives the complex advanced spatial semantics. The logical judgment formula is formed by the logical synthesis of the distance judgment formula according to the semantic content and the domain requirement.

Now we give the corresponding distance calculation method for the features in the virtual assembly as the core content of the new calculation perception. Let's look at the situation in the virtual assembly interaction. The perception of the bolt screw hole feature. This is a typical matching feature pair in a machine. We further analyze the blind holes and through holes and discuss the case with and without gaskets. This is an example of a pair of convex and concave features. Select such a set of points on the geometric feature, the point A_1 is a point on the inner surface of the bolt cap, the point A_2 is a point on the plane of the bolt head,and the point B_1 is a point on the outer plane surface of the box part, for example, the hole feature, B_2 is A point on the bottom surface of the blind hole. The formula for distance perception calculation is as follows:

$$\begin{aligned} Dist_{FaM} = Min(((A_1 \cdot proj \cdot Major_{TF} - B_1 \\ \cdot proj \cdot Major_{TF}) \cdot Major_{TF}, ((A_2 \cdot proj \cdot \\ Major_{TF} - B_2 \cdot proj \cdot Major_{TF}) \cdot Major_{TF})) \end{aligned} \tag{1}$$

The judgement formula is :

$$FaceMating := 0 < Dist_{FaM} \leq \varepsilon_0 \tag{2}$$

In the above formula, $Dist_{FaM}$ is a distance between two surfaces which will mate when the parts are assembly in place. $Major_{TF}$ is the major axis of the hole feature . This is the calculation during the motion after the two parts have been aligned. So each point is projected onto the main axis of the target hole feature before the calculation is conducted. If the above distance is less than a specified threshold ε_0 , then we say that the bolt assembly is in place. This is the perceptual calculation of the face mating phase . The feature matching perception of bolt and hole is to determine whether to match in real time in the interaction. In fact, the fundamental purpose is to find and select the corresponding hole as the target feature, that is, the hole feature, according to the current feature of the bolt when the bolt part is grabbed for assembly. Find the appropriate matching hole feature on the target part in the current scene. The matching of feature types is determined based on a feature mapping table generated by the assembly feature relationship.This feature searching is also done by distance calculation. When the distance is small, the matching degree between the bolt and the hole is high. When the distance is large, the matching between the feature pairs will be poor or even not matched at all. What is needed is the following two main distance formulas. Of course, in the actual assembly simulation system, it is necessary to add a number of formulas related to the technic requirements. For the sake of brevity, they are omitted here.

The distance perception calculations are :

$$\begin{aligned} Dist_{FeM1} &= length_{hole} - length_{bolt} \\ Dist_{FeM2} &= \phi_{hole} - \phi_{bolt} \end{aligned} \tag{3}$$

The judgement formula is :

$$FeaMatching := 0 < Dist_{FeM1} \leq \varepsilon_1 \wedge 0 < Dist_{FeM2} \leq \varepsilon_2 \tag{4}$$

Where $Dist_{FeM1}$ represents the difference of two lengths.$Dist_{FeM2}$ is the difference of distance between the bolt diameter and the hole diameter. *FeaMatching* is the variable of the feature matching perception logic formula based on distance calculation. Note that the feature we are discussing here is a blind hole feature. The distance formulas for through holes are different but very similar, so they are not discussed here.

3 COMPOSITION AND DYNAMICS OF PERCEPTION MECHANISM

An analysis of the perceptual mechanism is given, which gives a clear understanding of the static structure and dynamic operation of the perceptual mechanism. Then give its model and further implement it. We need to model the relationship between scenarios, intentions, and perception types. The scenario includes the background, object composition, and object state. The virtual scenario we define has a specific name and has several specific potential activities. And these activities have a specific timing. The situation and intention constitute a perceptual matrix. The scene is a row, the intent is a column, and the internal values of the matrix are values of different perceptual types.

The perception mechanism is shown from a logical view. Look at the internal structure of the perception mechanism inside the component. First there is an algorithm module that contains a set of distance-aware algorithms for the virtual environment application. There is a script description. The perceptual time series used by each state and is application specific.Through this script description, the current perceived context can be obtained. There is also a console module that is responsible for launching the corresponding perceptual algorithm. In addition, there is a multi-channel perceptual feedback module that correctly feeds back perceived content in different ways.

The perception mechanism is shown from the process view. We explain the parallelism and timing of the perceptual mechanism. First, the execution of the perceptual algorithm accompanied by an interaction has a timing constraint. It is generally presented as a parallel perception of multiple feature pairs, see Figure 1. The central logic starts monitoring and scheduling when the perception is triggered. The program receives the VE current context and user intent, and determines the type of perception currently being executed. The participating object data is received, the specific object of the perception is determined, and perceivable dimensions and the distance formula used in each dimension are also determined . During the interaction process, it through quering script automatically switches to the corresponding perception type and mode. After the specific spatial relationship is sensed, the related content is delivered to the multi-channel perceptual representation module for display.

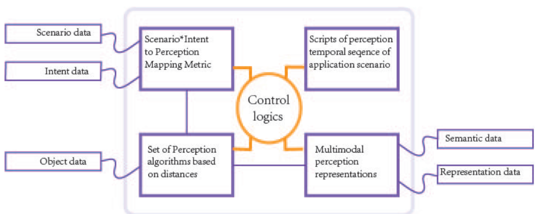

Figure 1: This caption has one line so it is centered.

The perceptual mechanism applied to virtual assembly systems is a convincing example. In the virtual assembly process, the scenarios we set are as below:

$$ScenarioSet_{VA} : \{Initial, FreeManip, FeaMatch, CnManip, Fix, DisAss, AutoAss, Browse, Simulation, TecPlan\} \tag{5}$$

And the intents we used are as below:

$$IntentSet_{VA} : \{Pick, MultiPick, FeaMatch, Align, FaceMate, DisCn, Review, DisAss, MultiDisAss, AutoAss, MultiAutoAss, TechPlan, InterSimulat, AutoSimulate\} \tag{6}$$

4 EXPERIMENTS OF PERCEPTION IN VIRTUAL ASSEMBLY

The author uses the perceptual mechanism introduced in this paper for human-computer interaction in a self-constructed new virtual assembly prototype system. Two car transmission assembles were verified. More than 200 parts have been designed. It involves almost all common assembly features, including a large number of concave features. The main achievement is to generate a lot of assembly semantics. All the desired assembly relationships are automatically acquired through the sensing during the direct operation, which is not available in other virtual assembly systems. We also made virtual hands in the virtual assembly system. A virtual hand is different from a part object. We also made the perception of virtual hands.This is another aspect of the virtual assembly application that can shows the adaptability of this perception mechanism.

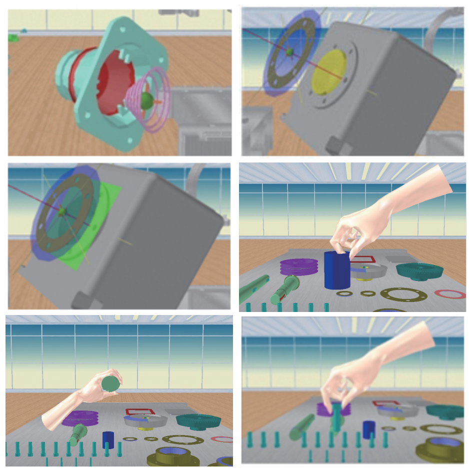

Figure 2: Perception examples in virtual assembly.

5 CONCLUSION

The new perception mechanism for virtual environments has a very important background, that is, the semantic requirements of virtual manufacturing. Under the condition of this new perception mechanism, the virtual assembly environment and the real manufacturing space can be truly integrated to become a real CPS system. The perceptual mechanism proposed in this paper has been verified to a certain extent. The experimental results show that it has great application potential. Later, we will further explore the various perceptual needs of virtual assembly, and we will further refine the theory of perception and various perceptual distance algorithms. And we will verify it in assembly processes of various products .

ACKNOWLEDGEMENTS

REFERENCES

C, E. (2004). *Real-Time Collision Detection*. CRC Press, Inc., Beijing, 1nd edition.

D, S. D. (1999). Inferring intent in eye-based interfaces: tracing eye movements with process models. In *Sigchi Conference on Human Factors in Computing Systems*. ACM.

E, R. and P, B. S. (1992). Efficient distance computation using best ellipsoid fit. In *IEEE International Symposium on Intelligent Control*. IEEE.

Gilbert E G, Johnson D W, K. S. S. (1988). A fast procedure for computing the distance between complex objects in three-dimensional space. *Robotics Automation IEEE Journal of*, 4(2):193–203.

H, S. I. and C, C. S. (2007). Xml-based assembly visualization for a multi-cad digital mock-up system. *Journal of Mechanical Science and Technology*, 21(12):1986–1993.

Hamidullah, B. E. and A, I. M. (2006). Assembly features: Definition, classification, and instantiation. In *International Conference on Emerging Technologies*. IEEE Xplore.

Hsu D, Latombe J C, M. R. (1997). Path planning in expansive configuration spaces. In *IEEE International Conference on Robotics and Automation*. IEEE Xplore.

K, C. E. and M, V. J. (2004). Collision detection and part interaction modeling to facilitate immersive virtual assembly methods. *Journal of Computing & Information Science in Engineering*, 4(2):83–90.

L, C. Y. and Y, L. T. (2008). Using pluggable procedures and ontology to realize semantic virtual environments 2.0. In *ACM SIGGRAPH International Conference on Virtual-Reality Continuum and ITS Applications in Industry*. ACM.

L, L. J. and M, C. (2007). Making sense of virtual environments: action representation, grounding and common sense. In *International Conference on Intelligent User Interfaces*. DBLP.

P, Q. and S, Z. (2005). Conversing with the user based on eye-gaze patterns. In *Conference on Human Factors in Computing Systems, CHI 2005*. DBLP.

S, A. N. E. and K, K. A. (2006). A new methodology for extracting manufacturing features from cad system. *Computers Industrial Engineering*, 51(3):389–415.

S, Q. (1994). Efficient distance computation between non-convex objects. In *IEEE International Conference on Robotics and Automation , 1994. Proceedings*. IEEE Xplore.

Schnell, T., Melzer, J. E., and Robbins, S. J. (2009). The cognitive pilot helmet: enabling pilot-aware smart avionics. In *SPIE Defense, Security, and Sensing*, pages 73260A–73260A–9.

Sud A, Govindaraju N, G. R. e. a. (2006). Interactive 3d distance field computation using linear factorization. In *ACM Symposium on Interactive 3d Graphics Games*. ACM.

Tang M, L. M. and J, K. Y. (2009). Interactive hausdorff distance computation for general polygonal models. *Acm Transactions on Graphics*, 28(3):1–9.

Tutenel T, Bidarra R, S. R. M. e. a. (2008). The role of semantics in games and simulations. *Computers in Entertainment*, 6(4):57.

Varol H A, S. F. and M, G. (2010). Multiclass real-time intent recognition of a powered lower limb prosthesis. *IEEE Transactions on Biomedical Engineering*, 57(3):542–551.

Xiaoming Z, P. D. and Z, Y. (2008). A model based approach to assembly sequence planning. *The International Journal of Advanced Manufacturing Technology*, 39(9):983–994.

APPENDIX