*Science in China Series F: Information Sciences*

© 2009 SCIENCE IN CHINA PRESS

Springer

www.scichina.com
info.scichina.com
www.springerlink.com

# A survey on virtual reality

ZHAO QinPing

State Key Laboratory of Virtual Reality Technology and System, Beihang University, Beijing 100083, China
(email: zhaoqp@vrlab.buaa.edu.cn)

**Virtual reality (VR) is a scientific method and technology created during the exploration of the nature by human beings to understand, simulate, and better adapt and use the nature. Based on the analysis on the whole process of VR, this paper presents different categories of VR problems and a type of theoretical expression, and abstracts three kinds of scientific and technical problems in VR field. On the basis of foresaid content, this paper also studies current major research objectives, research results and development trend of VR in the aspects of VR modeling method, VR representation technology, human-machine interaction and devices, VR development suites and supporting infrastructure, as well as VR applications. Finally, several theoretical and technical problems that need to be further studied and solved are addressed.**

virtual reality, modeling, rendering, human-machine interaction, development suite

Based on computer science and relevant scientific technologies, virtual reality (VR) produces a digital environment in which visual perception, sense of hearing, and sense of touch are highly similar to those of actual environment within a certain range. With the help of necessary equipment to interact and interfere with objects in a digital environment, users may have feelings and experiences corresponding to those in the actual environment. VR is a scientific technology for understanding or simulating the nature.

With the constant development of social productivity and scientific technologies, the VR technology has found more and more applications in various industries. Now VR has become a new science field. This paper presents an overview of the developing process of VR. First we introduce the thoughts and the related research direction of VR, then we discuss the major contents and current situation of VR research in the aspects of VR modeling, VR representation technology, human-machine interaction and equipment, VR development suite and supporting environment, as well as VR applications, and indicate several problems of the theory and technology.

## 1 Formation of VR research field

### 1.1 Formation of VR thoughts and research objectives

The VR concepts, thoughts and research objectives have close relationship to some relevant scientific

Received June 10, 2008; accepted February 13, 2009
doi: 10.1007/s11432-009-0066-0
Supported by the National Basic Research Program of China (Grant No. 2009CB320805), the National Natural Science Foundation of China (Grant Nos. 60533070, 60503066), and the National High-Tech Research & Development Program of China (Grant Nos. 2006AA01Z333, 2006AA01Z311)

technologies, especially the computer science.

In 1929, Edwin A. Link invented a type of flight simulator for making the passengers experience the feeling of flight. It was the first try that human beings simulated or emulated physical reality. With the constant development of control technology thereafter, various simulators came into being in succession. In 1956, Morton Heileg invented the Sensorama—a motorcycle emulator that showed 3D display and stereophonic effects and produced vibration feeling. He advanced some basic thought of VR technology[1] in the *Sensorama Simulator* patent in 1962. The development of electronic technology and the miniaturization of computers facilitated the development of simulation technology. Finally computer simulation technology appeared.

In 1965, the significant founder of computer graphics, Dr. Sutherland[2] published a piece of essay *The Ultimate Display,* portraying a type of new display technology through his sharp insight and abundant imaginations. He assumed that, supported by this display technology, observers may be surrounded by a virtual environment controlled by a computer, just like daily life in the real world. Meanwhile, observers may also interact with the objects in virtual environment by natural means, like touch perception, control of virtual objects, etc. From the angles of computer display and human-machine interaction, Sutherland's essay put forward the thoughts of simulating the real world, facilitated the development of computer graphics and image technologies, and cultivated the research on a new type of human-machine interaction equipment such as head mounted display (HMD) and data gloves.

In 1966, Sutherland began to develop a device of HMD, and later put the feedback device of simulation force and sense of touch into the system. In 1973, Krueger nominated the phrase of "Artificial Reality"—the VR words appearing during the early stage. Due to the constraints of computer technologies, generally speaking, the development of these types of technologies did not go very fast. During the 1960s and 1970s, the related thoughts, concepts and technologies still stayed in the brewing and formation stage.

During the 1980s, with the development of computer technologies, especially the update of PC and computer network, VR technology made much headway. Several typical VR systems came up during this stage. The implementation of SIMNET (SIMulation NETworking) Plan by US Army and DARPA in 1983 initiated the research and applications of distributed interaction simulation technology. Some successful technologies and experience of SIMNET significantly influenced the development of distributed VR technology. In 1984, M. McGreevy and J. Humphries developed a visual display for virtual environment and built up a 3D virtual surface environment of Mars by inputting the data from Mars probe to the computer. The development of VIDEOPLACE and VIEW systems facilitated the research on VR theories and technologies.

In 1986, Fisher et al.[3] published *The Virtual Environment Display System* concerning VR systems. In 1987, Foley[4] published a paper on *Scientific American*. The journal also published some articles about data gloves that attracted the attention of the public. In 1989, Jaron Lanier, the founder of VPL Co., put forward the phrase of "Virtual Reality", which was generally accepted by researchers and became the specific title of this scientific technology field.

During the 1990s, with the breakthrough and rapid development of computer technology and high performance computation, human-machine interaction technology and equipment, computer network and communication, as well as huge demands in the significant application fields such as military drill, aeronautics and astronautics, and complicated equipment research, VR technology came into a rapid development stage.

In 1990, VR technology was discussed in SIGGRAPH conference held in Dallas, USA, where the main contents of VR technology research were settled—formation technology of real-time 3D graphics, interaction technology of multi-sensor, high resolution display technology, etc. In 1993, Heim[5] portrayed seven characteristics of VR in *Metaphysics of Virtual Reality*: simulation, interaction effect, artificial reality, immersion, telepresence, general immersion and network communication. In 1994, Burdea and Coiffet[6] published their book, *Virtual Reality Technology,* in which they used 3I (Immersion, Interaction, Imagination) to generalize the basic characteristics of VR.

A batch of software platforms and modeling languages for VR system development appeared. In 1989, Quantum 3D Co. developed Open GVS. Sense8 Co. put forward WTK in 1992. On the 1st Annual WWW Conference held in Geneva in March 1994, VRML was brought forward for the first time, initiating the establishment of relevant international standards.

The development of computer modeling and simulation started as early as the beginning of the 1970s in China, mainly concentrating in the fields of aeronautics and astronautics. At the beginning of the 1990s, the researchers of some universities and scientific institutions in China started the VR research[7] from different points of view. The Ministry of Science & Technology of the People's Republic of China and the National Natural Science Foundation of China started to support the research on VR field. The National 863 Program formulated the "Distributed Virtual Environment" as a key program and implemented DVENET Plan[8] in 1996.

In the past decade, Beijing University of Aeronautics & Astronautics University, Zhejiang University, Tsinghua University, Peking University, National University of Defense Technology, Beijing Institute of Technology, Wuhan University, Shandong University, Beijing Normal University, University of Electronic Science & Technology of China, Institute of Computing Technology & the Chinese Academy of Sciences, Institute of Automation, the Chinese Academy of Sciences, the Second Academy of China Aerospace, and scientific researchers of many other application departments and institutions conducted many research efforts with various backgrounds and characteristics[9−16], and achieved remarkable performances in the aspects of VR theory research, technology innovation, system development, application promotion, etc. This science field came into a brand new development stage in China. Due to the integrity and specialty of VR science, and increasing application demands of economic, social and military fields, the *Essentials of National Medium and Long-Term Science and Technology Development Plan* granted by the Central People's Government of the People's Republic of China in 2006 listed the VR technology as one of the prior developing frontier technologies of information field. The Ministry of Science and Technology of P.R. China approved the establishment of State Key Lab of Virtual Reality Technology and Systems in Beijing University of Aeronautics & Astronautics University in 2007.

### 1.2 Scientific and technical problems of virtual reality

The objectives of VR are to use computer technologies and other relevant technologies to duplicate and emulate the real world (hypothetic world), and build a virtual world similar to the real world. Users experience the corresponding real world and even influence the real world through the interaction with virtual world. This procedure could be described as the following forms:

Set $W$ as the aggregate of all states in the real world (hypothetic world), and divide $W$ into the aggregate $T$ of subaggregate without intersection to let different VR modeling methods correspond to the states of real world; $C$ is the aggregate of computer state sequence; $E$ is the aggregate of human-machine interaction equipment; $e \in E$ and $S(e)$ are the aggregate of $e$ state sequence. We define an optional operation "see" first, which reflects the states in $W$ to the divisions it selects:

$$\text{see}: W \to T.$$

To indicate the simulation of real world, the simulation function "in" is defined, which reflects the states in real world to the computer state sequence set:

$$\text{in}: T \to \rho(C) \ (\rho(C) \text{ is the power set of } C.$$

To indicate the effect of object in virtual world on the outside world, define virtual object to present function "show" that reflects the computer state sequence set to the state sequence set of human-machine equipment as show: $\rho(C) \to S(E)(S(E)$ is the state sequence set of all interaction equipment).

To indicate the human control of virtual environment, define the control function "do" as

$$\text{do}: C \times S(E) \to C.$$

Define VR system as an octet at last:

$$\langle W, T, C, E, \text{ see, in, show, do} \rangle$$

The functions of "in", "show", and "do" are the three major scientific and technical problems of virtual reality. "in", the modeling of VR, could be an algorithmic or axiomatic system, or a manual computer input process with structures; the presentation of VR object "show" could be an algorithmic or data conversion, or a data output flow with structures; "do" is the control of virtual environment by humans or the outside world, which is the data input flow with structures and related to interaction equipment closely.

The following sections of this paper will discuss these three major problems.

### 1.3 Classification of virtual reality system

The research object, research objective, and application demands of VR all determine this scientific and technical field with comprehensive and crossed subjects, extensive application areas, and various system categories. Different classifications can be determined by different points of view of VR system.

1.3.1 Classification by system functions. Considering the functions and effects of system, which are actually the applications, VR system could be divided into the following categories: training & drill, planning & design, and presentation & entertainment.

The Training & Drill System can be extensively used for training and drill of technical operation in various dangerous circumstances (such as nuclear facilities and submarine facilities) where operation objects cannot be obtained easily (such as medical operation and spacecraft maintenance) and cost a large amount of money (such as military drill); the Planning & Design System can be used for demonstration and verification of newly-built facilities and equipment, thereby greatly reducing design cost, shortening design cycle and increasing design rationality, such as the planning and design of cities, communities and buildings, and virtual designs and virtual assemblies of equipment and products; the Presentation & Entertainment System will realize the digitalization of the scenes in the real world or the hypothetical world, and provide users with realistic enjoying experiences, such as virtual landscape, digital museum, various game manufactures, movie and television manufacture, etc.

1.3.2 Classification by immersion experience. VR system can be divided into non-interactive experience, human-virtual environmental interactive experience and group-virtual environmental interactive experience by immersion experience.

The above point of view emphasizes the interaction between users and virtual environment. With regard to non-interactive experience, users' experience of virtual environment is passive and the content of experience is completely planned. Some systems provide visual points for users to make selections to a certain degree so as to guide data dispatching of virtual scenes. But no such material interactions will happen as scenic wandering, 4D cinema, etc. In a human-virtual environmental interactive experience system, users may interact with virtual environment through interaction equipment (such as data gloves, digital weapons, digital scalpels, etc.). Scenes in virtual environment will respond to the interaction in time, so that users may feel the change in virtual environment and experience corresponding real world, such as flight simulator, individual-soldier training environment, etc. The group-virtual environmental interactive experience system is the multi-machine and network of human-virtual environmental interactive experience system. Multiple or group users may share one virtual environment, and interact with virtual environment that includes user avatars as well so as to experience the interaction with virtual environment and other users, such as military emulation drill, on-line interaction games, etc.

1.3.3 Data flow in virtual reality system. VR system is a complicated information system possessing various types of data, which can be described by the modes and trends of different types of data flow during operation period. According to this point of view, VR system is an information system that has $m$ types of data interacting with relevant environment and converting, accumulating and flowing constantly in terms of certain rules during $n$ stages.

The data in VR system could be divided into four categories according their origins and effects: Platform Data (PD) includes metadata produced by computer system, network system and public platform that support specific operation of virtual reality system, and administrative data produced by the all-purpose parts of specific VR system; Model Data (MD) is the master data of VR system, and also the reflection of scenes in the real world to digital space, which can be produced by scanners, and model algorithm, etc.; Sense Data (D), or the rendering data, is produced by various rendering algorithms applied to all types of output equipment to let users have visual, auditory, and force feelings; Control Data (CD) means the data produced through input equipment by users, applied to the control of and effects to virtual environment.

Figure 1 shows the data flow in a scenic wandering VR system. Virtual scenes are established by developers using data acquisition equipment and modeling software such as video, image, and spectrum instruments and scanners. It can use the mouse and control sticks, thus inputting control data to dispatch model data of the scene during wandering so as to select the visual point of wandering.

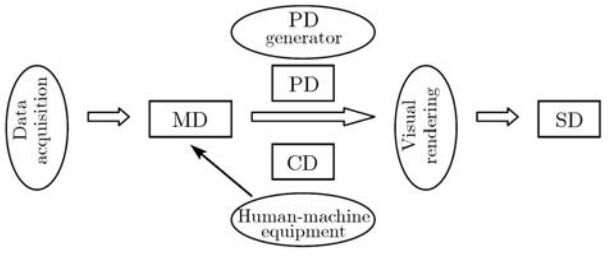


**Figure 1** Scenic wandering VR system.

Figure 2 shows a data flow in individual-soldier training system. Users produce control data and interact with objects in VR scenes through digital weapons. Combining with the platform data formed by system rules, they also produce new model data and then generate new perception data through rendering algorithm.

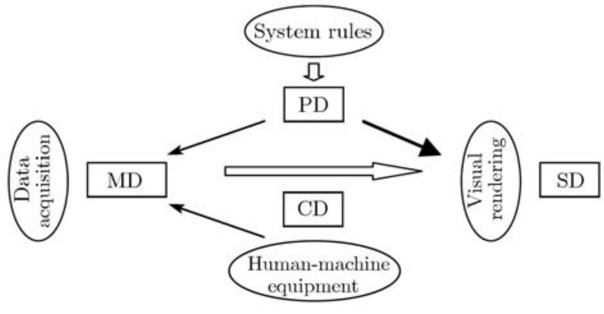


**Figure 2** Individual-soldier VR training system.

Multi-person rivalry-drill VR system can be developed and established based on the individual-soldier training system, and operated under the support of the operative and supportive environment of distributed VR system, such as DIS, RTI, etc.

## 2 Modeling methods in virtual reality

In order to simulate the objects and states of real world in digital space, the regulations followed by the objects, object relationships, interactions between objects, and development and change in the real world should be reflected as various data in digital space for presentation. This process is called modeling, as in function "in" mentioned in VR Octet in section 1. This section mainly introduces the methods and development trends of modeling.

### 2.1 Classification of model data types and modeling methods

2.1.1 Origins and types of model data. Modeling refers to model data first. Due to the multiplicity of objects and states in the real world, there are various model data. But with regard to data origins, there are mainly three categories of model data, namely actual measurement, mathematical generation, and artificial construction.

1) Actual measurement. It refers to the data acquired through 2D or 3D scanners, cameras, motion capture equipment, data clothes, data gloves

and various special data acquisition equipment (such as CT, B-ultrasound, nuclear magnetic resonance, satellite remote sensing, etc.) or artificial measurement.

2) Mathematical generation. It refers to the formation of mathematical models such as physical equation, analogue algorithm, and axiomatic system through experimental analysis, physical emulation, induction, abstraction, etc. The computation of mathematical models can generate model data, such as position data of objects acquired with equation of motion.

3) Artificial construction. It refers to model data generated artificially or through hypothesis and imagination by artificial modeling software (CREATOR, 3D Studio, AutoCAD) such as 3D models of objects, cartoon, movie and television manufacture, etc.

During the process of actual modeling, the complexity of simulated objects in the real world and the expectation degree of model fidelity should be taken into consideration to combine multiple methods to generate model data.

In VR system, according to different aspects presented by the objects in the real world, the model data could also be divided into the following categories—spatial structure data, physical property, behavioral property, dynamic data and motion data. Spatial structure data is to indicate geometric state of objects, such as point cloud, grid, voxel, etc. The data are presented as solid, liquid and gas. Typical spatial structure data is the 3D model of object. Physical properties and parameters are used for describing the properties shown by objects taking part in physical process, such as weight, size, material, force friction, lamellation, deformation parameter; behavioral properties and parameters are to describe the properties represented during behavioral process; dynamic and motion data are to describe the motion, deformation, collision, capture, zooming, and breakage of objects, such as the data of an object's position, speed and acceleration; field data is used for describing the states and properties of various field composition, such as electromagnetic field, gravitational field, fluid field, sound field, and wind field. Field data could be divided into scalar field, vector field, tensor field, etc.

2.1.2 Classification of modeling methods. VR modeling methods are related to object category required to be simulated, as well as the application fields. Therefore, same as the classification of VR system, it could also classify VR modeling methods based on different starting points. For example, based on users' perception, modeling methods include visual modeling, auditory modeling, haptic modeling, etc.

According to different aspects of simulated objects, the modeling methods can be divided into scene appearance modeling, physics-based modeling, behavior modeling, virtual-real combining modeling, etc. Scene appearance modeling focuses on the appearance of the scenery, mainly including those modeling methods based on geometry, image and material and illumination information; physics-based modeling represents object's physical properties, making the dynamic and static sceneries in virtual environment more vivid, mainly referring to the simulation of physical process, such as simulation of dynamics, collision, and deformation; at present, behavior modeling in VR mainly indicates the modeling of autonomous entity, referring to artificial intelligence. Computer generated forces (CGF) is a typical application of behavior modeling or VR; the virtual-real combining modeling is a type of significant modeling methods, which allows the virtual sceneries generated by the computer to naturally interfuse with actual environment of the real world so as to effectively improve the efficiency, flexibility and verisimilitude of VR modeling. The virtual-real combining modeling is the major research content to enhance reality at present.

The following part of this chapter will introduce some normal VR modeling methods used nowadays from the foresaid four aspects.

## 2.2 Scene appearance modeling method

Scene appearance model is used for displaying the spatial structures and appearances of virtual objects, including point cloud model, grid model, voxel model, etc. The major modeling methods include those based on range image, image and voxel. Object's surface illumination model is included as well.

2.2.1 Modeling based on depth range image. Each pixel of range image saves the range information of first intersection point of related light and the scene shot, the modeling process of which involves the acquisition, registration, surface reconstruction and repair of range image.

(1) Acquisition and registration of range image. The most direct way to acquire range image is to use 3D scanning equipment. 3D scanning equipment includes remote type, medium range type, and short range type. All are applicable to those sweep objects with different sizes and different ranges. The contact type and laser type are available at present. 3D laser scanner adopts the round trip longimetry or the triangular longimetry to acquire range information, which has rapid speed of data acquisition, high accuracy, and wide range of applications.

In order to acquire all the information of 3D object's surface, we need to conduct scanning and sampling from different visual points, and then put these range images together to form complete point cloud data of object surface. This process is called the registration of range image. Because the data acquired from different visual points are located at independent coordinate system of their own, we need to transform them to the same coordinate system. The registration of range image is to look for the move of these range images to solve the transformation matrix between these coordinate systems.

At present, the usual registration procedure of range image involves four steps: every two coarse registration, every two precise registration, adjacent relation determination, and multiple visual points registration. In the aspects of every two coarse registration, Chen[17] adopted the Global Search Algorithm to realize the registration; however, this method was slow and unstable; Huang[18] put forward the registration method based on the characteristics to abstract the characteristics of range image to be registered first, to find out reliable coincidence relation between characteristics, and to obtain candidate results of corresponding coarse registration, then to combine the precise registration to conduct consistency inspection on candidate registration results, and finally to acquire the optimal registration results. Based on algorithmic theory of ICP (iterative closest point), the every two precise registration is mainly to estimate the relative position of two images, and then gradually approach to precise registration results through iteration. Relative relation of image is determined through the overlap between images and the results of every two registration[19]. The objective of multiple visual points registration is to optimize the overall error and eliminate the accumulation of every two errors.

(2) Surface reconstruction of 3D model. Assume there is a data set $D$ of sampling point coming from surface $S$. Then surface reconstruction means to construct a surface $S'$ similar to $S$ through $D$. Surface reconstruction algorithm can be divided into two types in general: proximal surface (implicit surface) and interpolation surface (explicit surface). The zero iso-surface of (oriented) distance field from proximal surface's construction point to object surface is the reconstruction surface, such as Marching Cubes Algorithm. The interpolation surface applies Voronoi Drawing to conduct Delaunay triangularization to sampled data, and finds out the topological relation between sampled points so as to obtain a surface covering the original sampled points[20].

(3) Repair of 3D model. Repair methods of 3D model mainly include those methods based on grid and volume data. Repair method of model based on grid[21] is to find out the edge of defective part of a model surface first, then look for appropriate grid faces to fill up these defects, and conduct refinement and smoothness treatment on these grid faces, so that the additive parts can keep consistent with the existing grids in the aspects of grid structure and consistency. This method is applicable to repairing those cavities with relatively small areas, sensitive to data noise, but unable to guarantee the completeness of results. The repair method based on volume data[22,23] first conducts voxelization treatment on polygonal grid, convert

the polygonal grid data into volume data, then repair volume data, and use the method like iso-surface acquisition to regenerate polygonal grids.

In addition, there are repair methods based on radial basis function, finite element and existing model samples[24]. These methods are effective to small cavities, but will cause nonconformity from geometric characteristic of original object when repairing large cavities. Therefore, Sharf[25] put forward a repair method based on the surrounding grid information, which is to use the geometric characteristic information of known area to rebuild defective surface of object and acquire favorable repair results.

2.2.2 Image-based modeling method. Image-based modeling method uses photos to construct vivid 3D models, which can be divided into active type and passive type. Active type obtains 3D information of object actively through the control of illumination of the scene, which has high reconstruction accuracy and an easy algorithm. But it requires artificial reconstruction of clues, whose operating process is complicated. Passive type receives intensity information of sceneries passively, and then conducts 3D reconstruction through analyzing passive clues of image such as tone, shadow, focal length, texture, and parallax. This method barely restricts the scale and position of modeling sceneries, but has poor accuracy and complicated algorithm.

(1) Reconstruction of geometric model based on single image. People often use tone, texture and focal distance to conduct geometric reconstruction of single images in the aspect of computer vision. As this type of method strictly requires the shape, reflection property and exposure degree of image in the scene, it is only applicable to the 3D reconstruction in certain special sceneries. There are mainly two types of methods of geometric reconstruction based on single image for VR applications at present. One is based on plenty of interactions, e.g. using conventional image editing method to specify depth value of each point in the image; the other is to introduce model base, e.g. using the basic structural model base of architecture to reconstruct the model of architecture based on single image[26].

(2) Reconstruction of geometric model by adopting stereo vision and structured light method. Stereo vision method is to rebuild geometric model of the object by using two or more images. The most important part of stereo vision method is to seek and match the corresponding points between images. By introducing polar line constraint, the scope of searching for image corresponding points can be reduced from two dimensions to one dimension[27] to obtain the matching with higher stability and accuracy. Scharstein[28] has made classification and comparison of all corresponding matching algorithms. The method used to get the corresponding point between light source and image by the known light source coding information is called structured light method. The common light source codes include checker, black and white stripes, sine light, etc. Salvi has summarized the common coding method of structured light.

(3) Scenery reconstruction by using prior knowledge. As it is difficult to make good effect by directly recovering geometric information of scenery only depending on image information, some methods for rebuilding scenery with the aid of prior knowledge were proposed. Wei[29] firstly used the "accurate and strict" method to recover geometric information and then combined it with prior knowledge to finally recover hair and plant models. Tan adopted the image-based method in combination with prior knowledge to rebuild a tree model. Compared to not using prior knowledge, the rebuilding effect is greatly improved[30].

(4) Reconstruction of geometric model based on silhouette contour line. Silhouette contour line of object on the image is an important clue to understand geometric shape of an object. Silhouette contour line and the corresponding projection center determine a conical shell of 3D space together. The set of all these conical shells is composed a spatial envelop covering this object. This spatial envelop is called the visual shell of object, which is an approximation of this object. Since generation algorithm of visual shell usually requires reference images of multiple visual points, we need to conduct image calibration. And computation of get-

ting intersection for related visual shells consumes much time too. Matusik[31] adopted the limit geometric technology of computer vision to accelerate the computation of visual shells. The method based on silhouette contour line enables rapid reconstruction speed and high robustness. However, due to the constraints of contour information, this method is only applicable to some objects with simple shapes.

2.2.3 Material illumination modeling method. Computer graphics normally uses texture mapping technology to map the 2D image of surface reflection property of reflection object into the 3D model surface. This method only represents object's color information under certain specified illumination and visual points, but a lot of surface reflection properties of objects possess a large amount of high gloss information. Simple texture mapping technology is unable to realistically present plentiful light reflection and refraction effects of objects. Therefore, the research on acquisition and modeling technology of surface material properties has significant meanings in realism modeling field.

Object's material reflection properties mainly occur in the interaction between the light and object surface. According to photon physics, the interaction between the light and object can be described as a 12-dimension function through the position, direction, and wavelength of incident light, and the transmission time of light inside the object. In the aspect of practical applications, simplifying this function may yield an 8-dimension BSSRDF (bidirectional surface scattering distribution function)[32]. BSSRDF indicate the general process of the interaction between the light and material, including reflection and secondary surface scattering. To further simplify BSSRDF, we need to obtain different reflection models including BRDF (bidirectional reflectance distribute function)[33], SVBRDF (spatial varying bidirectional reflectance distribute function)[34], BTF (bidirectional texture function)[35], and TVBRDF (time and spatial varying DRDF)[36] as well. Kalyan[37] conducted the collection, modeling and mapping of BRDF change on material surface during spontaneous change process, such as paint pigment, wet and rough concrete surface and dust accumulation.

2.2.4 Volume illumination model. Volume illumination mode is the mathematical description of physical phenomena like light intensity change when the light goes through voxel, serving as the basis for direct volume rendering. The current normal volume illumination models are active-attenuation model, varying density emission model, and material classification and hybrid model.

Active-attenuation model treats each voxel in volume data field as a particle light source, and distributes source intensity and an attenuation coefficient[38]. After attenuation along a distance in data field, the light generated is projected to the visual plane to form resulting image. This type of model derives light intensity computing formulae applicable to 3D volume data with different properties according to relevant physical laws and equations. With reliable physical and mathematical basis, this method is the most frequently used illumination mode for volume rendering.

Varying density emission model treats any object as a continuously distributed particle light source system. Object space is filled with particle clouds. Each particle can emit lights. Illumination is computed through accumulating the contribution from particle past through by the light towards light intensity. Although there are different microscopic explanations, this type of model has the same result as that of active-attenuation model.

Material classification and hybrid model considers the voxel as consisting of several types of materials. The density distribution, light intensity and opacity are calculated using different materials. The lightness and opacity of this voxel are mixed by combinational formulae according to the percentage of each type of material, and the overall light intensity is calculated according to the emitted light and the surface scattered light of assembly.

2.2.5 Field modeling method. Physical fields like electromagnetic field and gravitational field are all invisible. However, some VR applications often require the distribution, range of action and effects of certain fields in virtual environment (such as the scanning of radar beam) to be delivered to users through visualization. Therefore, the research on field modeling could be divided into visualization and appearance modeling. Fields like vector field and scalar field usually use grid for modeling. The grids mainly refer to regular grid, rectangular grid, curvilinear grid, nonstructural grid, scattered data, etc.

(1) Vector field. Data mapping in vector field includes three types in general, namely icon-based method, method based on stream line, stream surface, and fluid, and texture-mapping-based method. The texture-mapping-based method mainly includes spot noise texture synthetic method and line integral convolution method (LIC). Because LIC-based method can overcome the defects of the two former methods, it has developed rapidly. Stalling put forward a curve B approximation method based on structural vector field and a rapid LIC thought; Helgeland[39] implemented rapid LIC treatment according to seed points of vector field in the interested area, which increased the expressive force of structure in dense vector field by reducing the inter-screen degree between sampled points; Wegenkittl displayed the direction of vector line through special input texture of structure; Taponecco[40] put forward an input texture generation method based on Markov random field and relevant LIC treatment strategies.

(2) Scalar field. Data mapping in scalar field mainly refers to grid subdivision and iso-surface construction. Gird subdivision subdivides the data field into several voxels, making the field value point locate the peak of grid cell. Field value distribution inside grid cell is determined by the field value at each peak. To meet the application demands, a great number of grid subdivision algorithms have been proposed, e.g. quadtree/octree method, advancing front method and Delaunay Triangulation.

The early iso-surface construction algorithm was a simple iso-line connection method put forward by H. Fuchs and others in 1977. In 1987, W. Lorensen et al. brought forward the Marching Cubes method (MC) that could generate high resolution iso-surface. Akio advanced a method for subdividing a hexahedron into a tetrahedron, and assumed that the density was shown with a quadratic distribution along the edge of tetrahedron. This method is called the Marching Tetrahedral (MT). MT was developed based on MC method, which can avoid the ambiguity in MC method.

### 2.3 Physically based modeling method

Virtual environment and object's verisimilitude are related to appearance modeling level, especially the movement, interaction, and the extent of conforming to physical laws during changes, which owes a lot to the modeling of physical laws. On SIGGRAPH meeting in 1987, Ronen Barzel and Alan Barr put forward the concept—"Physically Based Modeling" in *Topic in Physically Based Modeling*. Thereafter, more and more physical methods like fluid model and combustion phenomenon have been introduced into VR research field. Physical models have become more and more precise and diverse, and the range of application has become more and more extensive.

According to different modeling objects, the physically based modeling methods include those methods for rigid bodies, flexible objects, nonfiguration objects, and human motions. In addition, in order to simulate certain natural sceneries and random variations, people also adopt particle system and process methods in physically based modeling.

2.3.1 Rigid body motion modeling method. If each point of object cannot be translated or rotated correspondingly to any other point of this object, then this object is called the rigid body. This modeling method only has to consider the position and direction change of rigid body motion modeling in environment. The object's deformation can be neglected. The content referred to mainly includes the motion simulation, collision detection, connection and constraint of rigid body.

(1) Simulation of rigid body motion. The methods for realizing realistic movements of an object in virtual environment include key frame method, kinematics method, and dynamics meth-

od. The latter two build mathematical models of physical movements through kinematics and dynamics, respectively. Dynamics method requires less specified parameters, and has more realistic simulation of complicated motion process than kinematics mathod. However, dynamics method requires a large amount of calculation and is difficult in motion control.

There are two types of strategies to solve the problems of motion control in dynamics method. One type is the pre-treatment, which is to transfer the required constraints and control to proper force and moment, and then introduce dynamical equation. The other type adopts constraint equation to present constraints. It can use normal sparse metrix method to rapidly solve the problem that the number of constraint equation is equal to the unknown (i.e. full constraint). But for the circumstance of under constraint, the problem of constraint resolution is relatively complicated.

Witkin used the Lagrangian dynamics equation and the space time and energy constraint equation to carry out motion emulation of a rigid body. Kaufman[41] put forward a type of algorithm that can simulate a large quantity of non-convex rigid body motions. By doing so, the extrusion model and friction model between rigid bodies were constructed. According to the quality, position and speed parameters of an object, the extrusion model constrains the speed when a rigid body is bearing maximum extrusion. Friction model is used for speed calculation and response treatment of rigid body when friction exists.

(2) Collision detection. The collision detection methods based on spatial structure mainly refer to spatial subdivision method and hierarchy bounding box method. The spatial subdivision method is to subdivide the virutal space into small cells with equal volumes. During the process of collision detection, intersection test is carried out on geometric objects in the same cell or adjacent cells. The strategies of spatial subdivision refer to equal subdivision, BSP tree, $k-d$ tree and octree, etc. The hierarchy bounding box method is to use bounding boxes with simple shapes to wrap complicated geometric objects. During the process of collision detection, first the intersection test is carried out between bounding boxes, and then gradually the detection particle size is thinned according to test resutls. The structures of bounding box refer to hierarchy bounding ball tree, AABB (Aligned Axis Bounding Box) hierarchy tree, OBB (Oriented Bounding Box) hierarchy tree, k-Dop (Discrete Orientation Polytope) hierarchy tree, convex shell hierarchy tree, hybrid hierarchy tree, etc. Spatial subdivision method is normally used for collision detection between geometric objects that are thin and equally distributed. Hierarchy bouding box method has found a wide range of applications, such as collision detection in complicated environment.

In order to treat continuous multi-object contact (i.e. continuous collision detection), the constrained force of contact point and the instantaneous time have to be calculated when contact stops, and then contact points that have stopped through combinatorial search have to be determined. Continuous collision detection has been extensively used for flexible objects, especially cloth simulation. Wong[42] put forward a type of method to randomly label triangular faces of object surface, and introduced characteristics filter layer through continuous collision detection process to increase algorithm efficiency. Govindaraju[43] adopted GPU to accelerate the continuous collision detection process.

The methods of collision response treatment are mainly based on dynamics and elasticity. The core problem lies in the equation resolution and formula calculation. Rigid body normally cannot be penetrated by each other. To avoid the occurrence of this phenomenon, these methods must respond to the collision, adjust object speed, and then calculate and apply contact force till the objects are separated.

(3) Connection and constraint modeling. The motion problems of hinge objects (such as doors, windows, rotary handles, etc.) under circumstance of multiple constraints are attributable to the research category of correlation motion. Generally speaking, correlation motion refers to forward correlation motion and reverse correlation motion.

Researches on forward correlation motion deal with the angle and length of each joint of correlation motion, and locate the position the joint end can reach. Main application area refers to robot and automatic control. Researches on reverse correlation motion deal with certain given position, and determine the accessibility of known models. If it is accessible, then calculate the angle of each joint. The main application areas are robot, automatic control, and computer animation, etc.

2.3.2 Flexible object modeling method. In the real world, a lot of objects have no rigid body characteristics. They deform during movement or interaction, namely they are flexible objects. The flexible object modeling mainly refers to geometrically based method, physically based method and collision detection of flexible object. According to the shape and appearance, the modeling of geometrically based method usually adopts catenary method and B spline method, resulting in limited effect of realistic simulation. This section mainly discusses the physically based flexible object modeling method and collision detection.

Physically based modeling method divides the flexible objects into particles in cells, and treats flexible object motion as a particle motion under various physical quantity effects such as force and energy. Grid structures of flexible objects are of discrete type and continuous type. There are point particle model, particle model, and spring-point particle model dealing with former type, while there are spring deformation model, air aerodynamic model, wave propagation model, finite element model, etc. dealing with latter type.

(1) Discrete particle model. Discrete particle model treats flexible objects as a set of discrete particles. The action between particles is simulated using deformation energy method and spring mechanical method. According to the energy minimization principle, the energy method considers the energy as an objective function and flexible object as a constraint condition. Combining the effects generated by external force, we use optimization method to solve flexible object shape with minimum energy. This type of method is more suitable to static simulation of flexible objects. Mechanical method conducts force analysis on particles. It uses dynamics principles to establish differential equation of particle motion, and determines motion tracks of particles in time series through numerical solution of differential equation so as to further determine the space configuration of flexible object. This method is suitable for dynamic simulation of flexible objects.

Feynman adopted the principle of energy minimization to simulate overhang effect of fabrics. Based on particle grid model, he derived energy formula of grid particles according to the principle of elastic sheet, and solved the problem of energy minimization using the steepest descent method. Breen put forward particle model, which uses the warp and filling to divide up fabrics, considers the intersection point of warp and filling as the particle, and determines the energy expression of particle interaction through tests. This method conducts more realistic simulations of fabrics. Its only deficiency is complicated computation.

Provot put forward the classical particle—spring model. Flexible fabrics are dispersed to regular four-sided grids under model constraints. Grid intersection point is the particle. Particles are connected by springs. Then it uses Newtonian mechanics to get motion equation of particle. Provot's method is really effective for the simulation of simple fabrics. Its deficiency is that the model is not accurate enough, thus resulting in low emulation efficiency of complicated fabrics.

(2) Continuous particle model. Continuous particle model considers the flexible object as a continuous medium, and uses continuous medium mechanical theory to calculate the deformation. Terzopoulos and others adopted the continuous elastic theory to simulate the deformation and motion of object, and simulated dynamics response of flexible objects to external force according to the object's physical properties, such as quality and elasticity. A lot of deformation effects could be simulated, such as complete elastic deformation, non-complete elastic deformation, plastic deformation, fracture, etc. They also invoked Lagrangian equation and heat equation to simulate the melting process of deformable object.

Witkin put forward the method to solve geometric constraint in parameterization model, by which constraint was described as an energy function. Sorkine[44] brought forward the principle of energy minimization during deformation process. Botsch et al.[45] advanced several deformation methods to keep the details, the main thoughts of which were to form hierarchy simplified grid according to relevant detailed parameters, and the research on self-intersection due to the models containing a lot of geometric details during deformation process.

According to the simulation of fabric folds, Thomaszewski[46] put forward a type of co-rotating coordinate algorithm based on the refinement of finite elements that could obtain realistic representation effect. Müller et al.[47] brought forward a type of methods based on position operation usable for the real-time fabric simulation in games.

Pentland described the dynamics status and geometric deformation of objects through vibration mode and volume mode respectively, and combined the above two by means of polynominal deformation mapping. Zhou et al.[48] tried to keep volume characteristics by subdividing the internal and surface space. Huang et al.[49] put forward a subspace method, which enhanced the effectiveness and stability during deformation process by establishing subspatial structure. Sorkine et al.[50] conducted researches on nonlinearity problem during deformation process.

(3) Collision detection of flexible object. People usually apply the continuous collision detection method to flexible objects in consideration of their characteristics. The main problem is how to precisely detect possible collisions. In addition, collision detection of flexible objects normally adopts hierarchy bounding volume tree structure[51].

Khulief and colleagues used the law of conservation of momentum to research collision problem of flexible objects, trying to solve the dynamics response of system after collision of flexible objects. Goldsmith put forward a method of treating flexible object collision based on Kelvin-Voigt model which made use of linear spring-damper model. Khulief extended this method into the multi-body system and constructed a type of linear spring-nonlinear damper models. Yigit brought forward a type of nonlinear spring-damper model. Lankarani adopted Hurtz collision method to analyze the collision procedure of two spheres. The required parameters could be solved according to the geometric characteristics, material features and recovery coefficient of two spheres. Wu divided an object into several substructures, adopted fixed interface main modal & modal coordinate, and constraint modal & modal coordinate to describe node deformation, and calculated the modal speed of substructures after collision in the light of the law of conservation of momentum.

2.3.3 Virtual human motion modeling method. Virtual human is an important component of virtual environment. The research on virtual human motion modeling mainly deals with the acquisition and treatment of motion data and control of motion.

(1) Acquisition of motion data. The acquisition of motion data requires capture equipment of motion data, such as data clothes, data gloves, etc. The use of acquired motion data for virtual roles can effectively increase the realism in virtual environment. Sensor is the key part of data acquisition equipment. According to different principles of sensors, the present motion capture equipment mainly consists of a mechanical system, an electromagnetic system, and an optical system.

Mechanical system is composed of a potentiometer that measures the position and direction of human joint points. The main defect is that the realism of acquired motions depends on the operators to a high degree. Electromagnetic system uses electromagnetic sensor to capture position data and angle data of human joint points. The defect is that it is sensitive to metalized products, and the naturalness of human motions is limited by sensor system; low data sampling rate can hardly satisfy the requirements of special application such as the athletic sports analysis. Optical system sticks special labels on human bodies, uses several cameras to film the video stream with a high contrast, and abstracts the human motion data from it. Optical system can provide an extremely high sampling rate of motion data, but the system is complicated and expensive.

(2) Treatment of motion data. The treatment of motion data mainly refers to motion editing and motion synthesis. A few changes of motion fragments make the motion editing compatible with certain constraints of time and space so as to synthesize human motion that satisfies certain requirements.

Because the motion capture data is the sampling of motion signals of each and every human joint, the technology in signal treatment area could be applied to the adjustment of motion capture data. Bruderlin put forward a type of deviation mapping method that could adjust local shape of the signal under the premise of keeping the overall shape and continuity of signals. Witkin[52] put forward a type of motion deformation method based on one motion fragment; it generated a series of similar motion fragments by using motion deformation technology.

Motion retargeting is one important technology of motion editing, which is to apply one same set of motion data for the manikin with the same framework but different skeleton length. Gleicher put forward the motion retargeting algorithm based on the constraints of time and space, which adopted the optimal theory to move and solve within the overall range, then adjusted each and every joint status of framework according the solution, and obtained satisfactory constraint conditions and kept new motions with original motion characteristics. The problem of this method lies in a big amount of calculation.

Motion synthesis is to synthesize several motion fragments into one new motion fragment according to time transformation weight value. Motion synthesis technology may synthesize a simple motion to a complicated one. Perlin put forward a type of real-time generation motion systems based on the process, which could use synthesis technology to generate new motions and the connection and transition between motions after manual construction of basic motions by users. Park and colleagues[53] adopted the motion synthesis algorithm to conduct continuous control of motions.

(3) Motion control. Through controlling virtual skeleton of the manikin, motion control may generate continuous frames of human motions. There are two difficulties in the control of virtual human motion: One is that the high freedom degree of human motion (usually over a hundred) makes motion control very complicated; the second is that human beings are very sensitive to the motions of virtual human beings, thus increasing the difficulty in motion control. At present, the methods of virtual human motion control are mainly the key frame method, kinematics control, dynamics control and process-based motion control.

The key frame method is the conventional method for virtual human motion control. Users control the kinematics characteristics of virtual human by adjusting the parameters of interpolation function, such as speed, acceleration, etc. Almost all animation software such as 3DMAX, MAYA provide the key frame method. This type of method is easy to be made and understood, but it requires a large amount of workload and high qualities of employees.

Two representatives of kinematics control are the JACK system developed by Badler and colleagues of the University of Pennsylvania, and the PODA system developed by Girard and colleagues of Ohio State University[54].

Kinematics methods are based on forward kinematics and inverse kinematics. Forward kinematics determines the position and direction of human body joints by setting their status information so as to control the motion of virtual human beings. Inverse kinematics determines the status information of other joints of human bodies according to the position and direction of each and every terminal joint in the world coordinate system, which relieves the fussy work of forward kinematics method to a certain degree. Although the kinematics-based control method of human motion control is intuitive, it lacks completeness and does not respond to basic physical quantities, such as gravity and inertial force.

Dynamics method solves each kinematics parameter according to the force and moment conditions of human bodies, such as position, direction, speed, and acceleration so as to determine the overall motion process of human bodies. Dynamics methods are based on the forward dynamics and inverse dynamics as well. Forward dynamics solves human motion according to force conditions of human bodies, while inverse dynamics solves the force conditions of human bodies by knowing the results of human motions. Dynamics methods can generate more complicated and realistic motions, and require less specified parameters than kinematics methods. However, dynamics methods required a large amount of calculations and are hard to control. Therefore, the foresaid two methods are usually combined for practical uses.

The process-based methods are very effective to some human motions with specified types, such as walking and running. Using these types of methods, designers only have to specify a small number of parameters, such as speed and the step length of walking in the calculation of human gestures at every single time. The process-based methods can generate group motions as well.

## 2.4 Behavior modeling method

Behavior modeling is the major content of autonomous object research in virtual environment, which dated from the research on Computer Generated Forces (CGF) in virtual battlefield. According to the definition by the US Department of Defense[55], behavior modeling is the "modeling of human behaviors or representations required to be expressed in military emulation". Behavior modeling technology and CFG have been extensively applied to distributed virtual battlefields like ModSAF, STOW, WARSIM2000, etc.[56], and have been used for the games developed by the US Army based on VR technology, such as FSC (Full Spectrum Command), FSW (Full Spectrum Warrior), etc.[57]. With the development of VR research and application, behavior modeling such as ESP (the Emergency Simulation Program) has been expanded to many fields like public safety, education, culture and entertainment.

During recent years, VR applications depend more and more on intelligent ability of autonomous object's behaviors. New behavior modeling methods have been constantly coming out, making behavior modeling an important field of VR research. But this field relates to the research category of artificial intelligence in general, and is extremely difficult. We only make a brief introduction in this paper.

2.4.1 Main types of autonomous object. Behavior modeling methods are related to the numbers of autonomous objects to be represented. Autonomous objects can be divided into individual and group types. The group type can be further divided into aggregation type and organization type.

Individual object is the autonomous object only containing a single individual. The modeling content of this type of object normally is to establish concept models, architectures, behavior rules, and conditions required of target tasks.

Aggregation object type contains multiple individuals, but at an overall form of multiple analyzing level. That is to say, aggregation object type can be represented as a single object as well as a form of multiple individuals. When the incidents affect the group in the virtual environment, a certain method is needed to calculate the effect on group parameter caused by these incidents so as to conduct the transformation among different analyzing levels. These procedures are called aggregation (high analyzing level transforms to low analyzing level) and deaggregation (low analyzing level transforms to high analyzing level). The research of this type mainly involves such aspects as model calculation, representation method of multiple analyzing levels, and the rules of aggregation and deaggregation.

Organization type has multiple individuals existing independently. They are combined together according to certain rules. Each single individual conducts decision-makings on its own, and obeys the control rules of organization as well. The behaviors of organization refer to individual behavior level and overall organization behavior level. Behavior modeling at different levels shows different requirements and characteristics. In the aspect of individual, in order to support the coordination and interactive control between individual and other autonomous object, we need to model relevant concept of social behavior, and conduct hierarchy of internal organization so as to support individual behavior and social behavior. In the aspect of overall organization, behavior represents the interaction and control among different autonomous objects, so modeling of interaction agreement, and social regulations should be conducted.

Human behaviors are very complicated. The dominating mechanism of a lot of behaviors is not quite clear. According to present cognitive level and based on technical realization, behavioral models could be divided into low-level reaction behavior, high-level cognitive behavior, complicated interaction agreement behavior, aggregation object's behavior, etc.

2.4.2 Modeling method based on finite state automata. The simple reaction behavior of autonomous object can be modeled by using the finite state automata. Every possible response action of autonomous object represents a state. Uncertain incidents can be handled with "task stack". Finite state automata is the normal behavior modeling method for military emulations. Fox example, ModSAF and WARSI2000 use this method to describe vehicle motions, ammunition emission, etc.

2.4.3 Expert-system-oriented modeling method. This type of method considers the autonomous object as an approximate expert system, and treats behavior modeling as the acquisition of knowledge, representation, and the establishment of inference system, and is applicable to higher-level behavior modeling of individuals and groups. Knowledge is of various types. For certain knowledge, we can adopt the representation based on logic, rules, and frames, and relevant inference system, such as the autonomous object of CCTT used for tactical training; for uncertain knowledge, we can adopt fuzzy logic, neural network, example-based inference and the Bayes method, such as the autonomous object of MCSF (Marine Corps Synthetic Forces) of US Marine[58]. Researchers usually combine multiple modeling methods together to construct mixed behavioral models. In addition, reinforcement learning and the theory of games[59,60] have been introduced as well to increase autonomous object's ability in solving complicated problems.

2.4.4 Agent-based modeling method. With constant increase in object types in virtual environment, the models of object had become more and more complicated. Traditional methods can hardly express high spiritual status and complicated behaviors of autonomous objects, such as belief, desire, and intention, nor can they simulate group objects' behaviors like autonomous object organization. Therefore, researchers introduced the Agent method in artificial intelligence field, which formed the Agent-based modeling method by combining with VR behavior modeling. This type of method is capable of individual and group behavior modeling. On the one hand, Agent has descriptions of high spiritual status like belief, promise, intention, and desire, to represent autonomy and intelligence of individual object to a certain degree; on the other hand, the communication, consultation and cooperation between Agents can describe the cooperative characteristics of autonomous object organization. The Agent-based modeling method has been more and more frequently applied to distribute virtual battlefield. For instance, IFOR (Intelligent Forces) has adopted this method to conduct behavior modeling of individual and autonomous object organization[61].

2.4.5 Aggregation and deaggregation models. Complicated objects often have multiple hierarchical tree structures. One model includes many submodels, and a submodel includes another submodel. There are many ways of analyzing level of objects that can be analyzed. For example, the shaping map of a division in the army is actually a multi-branch tree. In a battle simulation system, people might only need to describe the battalion, but in a tactical simulation system, people might need to describe a company, a platoon, or even an individual soldier.

Aggregation simulation is to transform a high analytic object into a low analytic object. Aggregation is a process from separation to synthesis, from complication to simplification, and from dispersion to unification. Deaggregation is just opposite to aggregation in which process the model goes through from synthesis to separation, from simplification to complication, and from unification to dispersion. There are three types of forms of model aggregation, namely aggregation of homogeneous model, aggregation of heterogeneous model, and aggregation of hybrid model. Homogeneous model mainly refers to the model that has the same principle and method of modeling, or the input/output data of model has the same characteristics. Heterogeneous model mainly refers to the model that has different principles and methods of modeling, or the input/output data of model has different characteristics. Hybrid model means that the aggregated model contains both homogeneous model and heterogeneous model. We need to aggregate submodels when solving and analyzing the overall effect of models.

Generally speaking, deaggregation of model is harder than aggregation, because the construction of the model with higher analyzing level requires more information. Through deaggregation of homogeneous model, we can solve the required information and finish deaggregation by analysis; but for heterogeneous model, this process is extremely difficult.

### 2.5 Modeling of virtual-real combining scene

Traditional virtual environment emphasizes scenery modeling and representation of virtual scene, which does a lot of modeling and rendering work. With constant expansion of VR applications, people find that some applications can completely rely on the real environment and it is feasible to only construct a few objects that need to be changed flexibly into relevant virtual objects and make them seamlessly merged in the real scene. This can effectively increase modeling efficiency in virtual environment, and expand VR application areas. This is the main research content in augmented reality at present. The main problem is how to model virtual-real combining scene. This section introduces the modeling technology of virtual-real combining scene for the acquisition and representation of information in real environment, 3D registration between real and virtual, occlusion handling between real and virtual, and combination of real and virtual illumination, etc.

2.5.1 Acquisition and representation of real environment information in contractions of virtual-real combining scene. The virtual-real combining scene based on multiple video sequences mainly refers to video sequences describing the real environment and the geometric models describing virtual object. When entering the 3D virtual-real combining scene, different users often have different viewpoints and interactions in the perception and interaction regions of their own; as a result, the constantly varying multi-view real environment needs to be described in virtual-real combining scene. But the field of view of one video sequence is limited; thus all scene information in the real environment cannot be totally included. Therefore, we need to use multiple video sequences to describe the real environment in virtual-real combining scene. A lot of researches have been carried out according to the acquisition of multi-view information in real environment.

In order to obtain the information in real environment, Regenbrecht[62] not only assembles a head mounted display for each user, but also sets multiple cameras to obtain multi-view information in real environment, allowing users from different views to observe the same 3D virtual-real combining scene, interact with virtual objects in the scene, and deal with various collaborative work.

Based on multiple video sequences, Allard and colleagues[63] developed a markerless real environment construction system. This system sets up several cameras at the edges and top of a cube. First it uses multiple cameras to collect information in real environment and uses the change of background image to segment real objects and acquire relevant texture. Then this system calculates the visual scene area based on visual hull, and completes scene rendering through physical simulation and texture mapping at last. It can provide the real environmental information for the modeling of virtual-real combining scene.

Ercan and colleagues[64] put forward a video sequence selection algorithm based on linear estimation for 2D target tracking. This algorithm can select a group of optimal subsets from $N$ pieces of video sequences, thus allowing the minimum error of 2D target tracking, and ensuing the construction of virtual-real combining scene by optimal video sequences. Fusiello and colleagues[65] used video and audio sensors to describe underwater real environment, and acquired and preprocessed the video and audio information in real environment respectively. They carried out 3D registration and location matching between real and virtual environments via combining these two different types of information, allowing users to perceive the realistic virtual-real combining scene underwater through a multimodal fashion.

2.5.2 Methods of 3D registration between real and virtual. 3D registration of virtual-actual object helps to calculate the mapping position of virtual object in real environment; it settles the foundation for spatial relations between virtual and real environment, and virtual-real combining effect between virtual objects and real environment. The registration method of virtual-real object mainly refers to the hardware equipment based registration method and the computer vision based registration method. The latter is the mainstream of research work in China and other countries, which will be emphatically discussed and analyzed as follows.

The computer vision based virtual-real object 3D registration mainly refers to the methods based on marker, planar tracking, and image matching. As the marker based 3D registration method has good practicability, small cost of calculation, simple markers, strong automaticity, and high accuracy, it has a wide range of applications.

Billinghurst and colleagues[66] realized the marker based 3D registration method by calculating the geometric characteristics and relevant image plane positions of markers, and developed the MagicBook. Readers can not only read the words in the book, but also observe practical sceneries and perceive virtual objects and events generated according to the book content.

Rekimoto designed a type of 2D marker with $2^{16}$ marker patterns, which uses this marker and the principle of stereo vision to calculate the calibration position of real environment, and register virtual objects at that position.

During the process of 3D registration, sometimes we need to use the sensor to obtain camera's parameters and operator's perception orientation. However, the speed of virtual-real registration might be lower than the generation of sensor data, resulting in the problem of time delay. Didier and colleagues[67] put forward the compensation method of time delay based on texture. Their method could construct the 3D virtual-real combining scene with predicted data, and correct scenery data and rendering effect by complementing new predicted data.

Bajura and Neumann put forward a type of dynamic 3D registration and correction method for virtual-real object. Their method could correct 3D registration with different video sequences and maintain the spatial consistency of virtual-real combining scene as best as possible, allowing the users to realistically perceive the depth of field with dynamic changes and determine spatial position of virtual objects in the reference coordinate.

2.5.3 Methods of occlusion handling between real and virtual. The spatial occlusion relationship between the real and virtual objects directly affects the seamless merging of virtual and real environments, which is the foundation for obtaining spatial perception and conducting interaction operation by users in virtual-real combining scene. The spatial occlusion relationship of virtual-real object includes virtual occluding real and real occluding virtual. Virtual occluding real means that the virtual object occludes the real object, and real occluding virtual means that the real object occludes the virtual object. How to calculate the depth of a virtual-real combining scene is generally the key problem in occlusion handling between real and virtual.

The methods of occlusion handling between the real and virtual are: The methods not based on spatial information in real environment, such as the virtual-real object occlusion handling method put forward by Vincent and colleagues; the methods based on spatial information in real environment partially, such as the visual point occlusion handling method based on binocular parallax arrangement; and the methods completely based on the spatial information in real environment, such as the handling method based on voxelized spatial structure put forward by Kanade and colleagues.

Milgram and his colleagues[68] brought forward a visual point occlusion handling method based on binocular parallax arrangement, with which they also put forward the occlusion handling method on the basis of binocular parallax and depth of field that could increase the control accuracy of remote objects operated by users under the circumstance of maintaining low interactive complexity. Huang and Essa[69] marked all real objects according to the principle of computer vision, used the dynamic change of video sequences to calculate relative depth information between objects, and dealt with the occlusion relationships between objects through depth information.

Vogt and Schmidt[70] put forward a type of occlusion handling method for virtual objects based on binocular stereo vision matching, which estimated the position relation between virtual object and real environment using depth information in real scene calculated by video images. This method can deal with not only occlusion relationship between virtual and real objects, but also asymmetric correction and edge information of objects. Berger brought forward the occlusion handling method for virtual-real object based on edge detection, and with it, obtained the edge point of real object image according to edge detection algorithm of video image. They used binocular video sequence to estimate the "front" or "back" relationship between edge points and virtual objects, and determined the occlusion relationship between virtual objects and actual objects in relevant area by virtue of the "front" or "back" relationship.

2.5.4 Methods of combining real and virtual illumination. Virtual object's illumination model and real object's illumination information are two important factors that affect realistic effects of virtual-real combining scene. 3D registration method mainly focuses on the spatial position relationship between real object and virtual object, while virtual-real illumination combining handling method mainly deals with the problem of illumination combining effects of real and virtual objects.

Pilet and colleagues[71] obtained illumination information of real sceneries at 2D plane according to camera calibration parameters. They used the inverse illumination principle to create the environment map, calculated illumination information and shadow information of virtual objects through normal vector calculation of 2D plane, and rendered the virtual objects with light reflection characteristics and illumination shadows.

Jacobs and colleagues[72] put forward the virtual-real shadow combining handling method based on real object's shadow. They used the edge detection of scene image to determine shadow contour of real object, and calculate the color and brightness of shadow area, with which color, brightness and shadow effects of virtual objects are generated. Kanbara and colleagues[73] researched and developed an augmented reality system that supports scene illumination registration. This system sets up a small sphere in real environment, uses the model view matrix and sphere size to calculate the image position of sphere and the distribution of light source, and determines the brightness and illumination colors in real environment through the pixel of sphere image.

Beihang University has implemented a series of work on scene modeling of virtual-real combining as well. For the purpose of acquisition and representation of information in real environment, the perception of users and interaction regions in virtual-real combining scene are analyzed, and a type of multi-view representation method in real environment based on multiple video sequences is presented; a 3D registration method is put forward for virtual objects, a cooperative and complementary method of virtual-real 3D registration information based on the shared scene area is advanced and a cooperative occlusion handling method is established based on multiple video sequences. The university has researched and implemented the virtual-real illumination combining handling method based on video images.

## 3 Presentation technology of VR

VR presentation technology is an important research content in the field of VR; it renders all kinds of virtual scenery models in digital space onto presentation equipment using different VR presentation methods and algorithms, and presents them to the user in the way of immersion. This method belongs to VR function show of the VR octet mentioned in Chapter I.

According to the concept model put forward by Gibson, a person's perceptual system could be divided into five parts, namely vision, hearing, feeling, olfaction/gestation and sense of direction. VR presentation technology is required to provide the user not only with realistic visual perception, but also with a comprehensive vivid performance in terms of hearing, sense of force/feeling, etc.

### 3.1 Visual presentation technology

Realistic rendering of 3D graphics and image in real-time is the most important means to achieve VR visual perception and a core technology to construct virtual environment. Realism and real-time are two important safeguards for sensation of system immersion and interactivity in VR, and a pair of prominent conflicts at the same time. Realistic rendering of 3D graphics and image in real-time is mostly about improving realism and real-time capabilities.

The rendering technologies of graphics and image could be divided into graphics rendering technology, image generation technology, and the combination rendering technology of graphics and image according to different objects processed. In the process of rendering graphics, the current mainstream research directions are realistic lighting, texture mapping and rendering of natural scenery; for the scene represented by polygons, many researches have been carried out on grid simplification and parallel rendering in order to speed up the rendering. Using images as the input data, image-based rendering technology can create new perspective images without geometric 3D reconstruction, including the all-optical functions, optical field calculation, lumen chart, concentric mosaics and panoramic mosaics. The rendering technologies combining graphics and image mainly include geometric consistency and illumination consistency when processing graphics and image fusion.

3.1.1 Realistic illumination calculation. There are two ways of realistic illumination calculation, namely local illumination and global illumination. In local illumination model, when we calculate the brightness of a certain point on illumination target, we only consider all predefined light source target of virtual scene. Global illumination takes the whole environment as light source, considering not only the direct effect of light source of virtual scene to the objects rendered, but also the indirect effect of reflection, refraction and scattering of light to the objects rendered. It could enlarge the reality of rendering results. Therefore, global illumination is the key point of current illumination calculation research.

Classic global illumination phenomenon involves color penetration, shadow/soft shadow, caustics/caustic spectrum, sub-surface scattering, etc. The simulation and reconstruction of every phenomenon could significantly improve the rendering effect.

(1) Ray tracing algorithm. Ray tracing algorithm proposed by Appel and his colleagues in 1968 is the first global illumination algorithm ever appeared. This algorithm considers that light transmits in straight line, simulates camera negative to catch lights, and projects lights to every pixel in the direction starting from focus, and determines the light intensity by the location of scene objects that intersect with the projection lights.

Whitted and Kay dealt with mirror reflection and refraction on the basis of Appel's algorithm. Shadow, mirror reflection and transparent object refraction effects are the most typical applied areas of the early ray tracing algorithm. At this stage, ray tracing algorithm was not an entirely global illumination algorithm. Due to the limitation of the computational capability, the diffuse reflection surface which was difficult to deal with was still processed with algorithms which were similar to local illumination model.

Monte Carlo ray tracing method is no longer subject to this restriction. This method only projects one beam at a time and only selects a certain direction for continuously tracing rays after the reflection or refraction of the light. When a new direction is selected, it randomly takes the bidirectional reflectance distribution function (BRDF) on the object surface as the probability functions. The entire incident light radiance on each surface could be sampled using thousands of rays transmitted from each pixel, thus yielding high-quality rendering results.

Reverse ray tracing is tracing the rays from light source, which is opposite to the calculation direction of ray tracing. The flaw of ray tracing disposing diffusion can be overcome when the above two methods are combined together.

In ray tracing methods, what needs most time-consuming calculation is line-surface intersection, and the main strategy for optimizing the methods is to reduce line-surface intersection times. Bounding volume hierarchy methods and space subdivision methods reduce unnecessary intersections to the greatest extent to raise the speed of ray tracing through space division and organization of the objects in the scene.

Another way to improve the ray tracing calculation is to reduce the image aliasing or enhance the realism of the image. The main improved algorithms are taper tracing, beam tracing, random sampling and distributed ray tracing, etc. Taper tracing and beam tracing change the ray model from non-lateral scale line into tapers or pyramids, and overcome the phenomenon of aliasing caused by point sampling.

Ray tracing algorithms enables high-quality rendering, but the calculation burden is too heavy. Even by using various optimizing methods to reduce unnecessary calculation and the newest hardware equipment of graphics, it still takes a long time to complete the calculation work for one scene and the method is difficult to be applied in a real-time environment.

(2) Radiance methods. The thought of radiance algorithm comes from heat exchange theory which is different from the physical feature of ray tracing using light transmits along with straight line, and its main process originates from the light energy transfer theory. The earliest radiance algorithm was put forward by Goral and colleagues in 1984. First, the algorithm establishes the energy distribution for the whole scene for radiance algorithm, and then renders the scene into one or several different views. In solving radiance algorithm, a huge amount of calculation is needed for iterative solving of shape factor calculation and energy balance equations; therefore, a lot of efforts have been devoted to optimizing it, among which the comparatively important ones include half-cube method, gradually refinement technology and hierarchical structure radiance method. The importance of these methods makes radiance method a practical rendering technology.

As ray tracing method is suitable for solving the smooth surface that engenders mirror reflection and refraction, while radiance method is more suitable for solving diffuse reflection of non-smooth surface, the multi-rounds strategies combining ray tracing and radiance method appears.

Like ray tracing algorithm, the amount of calculation for radiance algorithm is also huge; therefore the real-time application cannot be used basically. A feasible real-time calculation method is calculating the light distribution in advance, and attaching it into real-time rendering scene in a static or dynamic way. The method of static attachment was put forward earlier and has been well used in Quake3. Through preservation of the critical angle, Hoffman and Mitchell[74] calculated the illumination situation of sunlight which was directly radiated to each vertex of the terrain. This is a typical dynamic usage in radiance algorithm.

(3) Photon mapping. Photon mapping method first projects “photon” from light source (different from “ray” in the Monte Carlo method), and tracks to record several reflection conditions of photon in the scene (maybe refraction, etc.). It collects the “photon” around this pixel in a certain radius and accumulates their values to obtain the color of this pixel. This method can relatively easily manifest the rendering effects of sub-surface scattering and caustics, and keeps certain advantages when rendering objects with complex materials.

(4) Pre-computed radiance transfer. Pre-computed radiance transfer proposed by Sloan[75] in 2002 is a kind of lighting technology. This method divides global illumination into lighting

computing and real-time rendering. The lighting computing part is accomplished by pre-computed program in advance, with results compressed and stored, real-time rendering part uses the completed data to obtain the final rendering results combining real-time environment lighting.

Pre-computed radiance transfer algorithm compromises the contradiction between the rendering quality and the rendering speed of the global illumination, to provide new methods for photo realistic real-time rendering and thus attracts great attention very quickly. Ng[76] used Haar wavelet function defined in the six planes of cubic graphics to replace the spherical harmonics of the original algorithm as a compression function of radiance transfer matrix $\boldsymbol{T}$. It uses nonlinear wavelet to implement real-time compression for light source vector $\boldsymbol{L}$, and overcomes the deficiency that the original algorithm can only manifest low-frequency lighting and realizes the whole spectrum shadow effects. The method of using matrix to describe pre-computed light radiation transmission function accelerates the rendering of object surface with asymmetric BRDF. Zhou and colleagues[77] solved the radiance transmission problems among many objects using shadow field. Sloan[78] realized deformable local radiance transmission algorithm using regional harmonic function, and this broke the important restrictions of radiance transmission algorithm applied to dynamic objects although the algorithm still has some local restrictions. Ren[79] accomplished low-frequency shadow rendering of a large range through expressing objects as a pair of spheres and carrying on visibility accumulation in logarithmic space.

The current pre-computed radiance transfer method can be carried on real-time global illumination rendering for dynamic scene in certain restrictive conditions. However, this technology is difficult to be applied to large-scale scenes due to the need of pre-computation. The pre-computation will increase along with the scale of scene, and the pre-computation data will be largely consumed by memory and display memory.

3.1.2 Complicated texture mapping. Traditional texture mapping is carried out through attaching the image with color information to geometric object surface to improve the realism of rendering effects without increasing the geometric details of objects. Along with the rendering technology's incessant development, more and more complex texture forms and corresponding mapping methods appear.

(1) Bump mapping. Traditional texture mapping is unable to represent bump details on the surface of objects. In order to solve this problem, Blinn[80] proposed bump mapping in 1978, which uses a perturbation function to disturb the normal vector of object surface to simulate the normal vector change of object surface, thus affecting the change of reflection lightness distribution and producing more realistic and detailed surfaces. The difference between bump mapping and traditional texture mapping is that texture mapping adds color onto polygons, while bump mapping adds rough information onto polygons. And this will produce an attractive visual effect of polygons.

A derivation of bump mapping is normal mapping. Normal mapping pre-computes the normal vector of object surface and stores the results in a normal map, by directly using the vector of the normal map as surface normal vector, which is different from bump mapping by using perturbation function to disturb the normal vector of object surface. Therefore, the speed of this method increases obviously compared to bump mapping, but it takes up more memory at the same time.

(2) Displacement mapping. Bump mapping does not change the actual shape of a model. Therefore, it has no influence on the contour line of the object and cannot exert the effects of occlusion and self-shadow. In 1984, Cook[81] proposed displacement mapping technology. This technology moves the location of the vertices according to the height map. It takes the moving direction as the normal direction and determines the moving size by a height map. Because displacement mapping essentially changes the geometric attributes of object surfaces, the roughness of the object's surface is obvious. However, it cannot be used in real-time rendering due to big amount of calculation.

(3) Parallax mapping. Parallax mapping is an

improvement to displacement mapping[82]. The inputs of parallax mapping include a height map and a normal map. The height map is used to execute texture coordination along with the viewing direction, and thus has more realistic effect than bump mapping technology and the required amount of calculation is reduced at the same time.

Parallax mapping uses pixel by pixel depth expansion technology to displace texture mapping coordinate, grants the texture a high degree appearance through parallax distortion. Taking a smaller amount of calculation as the premise, it could make materials have depth perception without using a large quantity of polygons. Parallax mapping could be used in the process of texture mapping and bump mapping, producing self-occlusion effect but it cannot produce self-shadow effect.

(4) Relief mapping. The principle of relief texture mapping[83] is similar to that of parallax mapping. Both are simulating the 3D parallax through forward mapping of an image, but relief mapping is more accurate than parallax mapping and it supports self-shadow effect and normal texture. This method uses image morphing and material pixel by pixel depth strengthening technology to represent complex geometric details on flat polygons. Depth information is calculated from the distance between reference surface and sampling surface. Relief mapping technology only uses normal texture to produce bump cubic effect on a plane and is usually used to draw object surfaces, but the amount of calculation is relatively huge.

3.1.3 Natural scenery rendering. As natural scenery is far more complex than most artificial objects, the simulation of natural scenery is always one of the most challenging problems in computer graphics. With the development of VR technology, researchers proposed a series of generation methods for natural scenery based on physics, graphics and images. Among these methods, the distinction between modeling and rendering is not so clear and even is combined as a whole. This section will take particle system, water wave simulation, modeling and rendering of hair and plant as examples to illustrate the rendering technology for natural scenery.

(1) Particle system. Reeves[84] proposed particle system in 1983 to simulate unshaped objects. In particle systems, a scenery object is defined as an organization formed by tens of thousands of irregular and random distributed particles. The location, speed and profile parameter of every particle could be set to have a certain life cycle that experience three phases, namely “birth”, “movement and growth” and “death”. Particle system fully reflects the dynamic and random character of unshaped objects. It makes it possible to simulate natural sceneries such as fire, clouds, water, etc.

One key advantage of particle system is the function of database amplification, e.g. Reeves and colleagues proposed a method that uses three basic descriptions to produce forest scenery formed by a million particles. Peachey and Fournier used a particle system to simulate animation of bubbles and splashes caused by winds, and also achieved good results. Recently, Reed and colleagues successfully simulated lightning using a particle system. The special light effects, e.g. firelight and smog, simulated by particle system have been widely used in movie making.

(2) Simulation and rendering of water wave. Water wave simulation has been widely done in the field of computer games, movies, advertisements, etc. Because water has different forms under different circumstances and the representation also has special requirements, water wave simulation is somewhat difficult.

Water wave simulation methods could be divided into three types. The first type is structure-based methods that constructs water wave shape using mathematical functions and then transforms time parameters to generate water wave. This type of methods could exert the visual effect but cannot reflect the rule of water flow. The second type is a physical-based method that starts from physical principle of water wave and obtains the state of fluid particle at each moment by solving a group of hydrodynamic equations. The effect of this kind of method is more realistic but the computational efficiency is low. The third type of method is adopting particle system. It could be used to simulate snowflake, waterfall, spoondrift, etc. Yuksel and colleagues[85] proposed a wave particle model, which could simulate the spoondrift of a fluid surface in real-time and the interaction with floating objects of surface. Adams and colleagues[86] proposed a particle water wave simulation method based on adaptive sampling which could carry out adaptive particle sampling according to the local feature of fluid and visual interest to reduce the number of required particles for fluid simulation. Using this method for adaptive sampling could effectively guarantee the smoothness of interframe change and improve the efficiency at the same time.

(3) Simulation and rendering of hair. The realistic simulation of hair could effectively increase the realism of the scene with animals and human beings. People once used particle system, procedural texture and volumetric texture to render hair in early years. Recently, many 3D animations use the method of geometric modeling to represent hair. They express every hair as a string of triangular patches or cones and carry out antialiasing and complex lighting computing, thereby generating a satisfactory rendering effect, but the rendering speed is too slow to meet real-time requirements.

Gelder and colleagues proposed a simplified geometric method to represent hair as polylines in 1997. Although this method could meet the requirements of real-time and interaction, the rendering effect was not satisfactory and the number of hair could not be too many. In 2000, Lengyel proposed a method of using multi-texture to stand for hair and cut a block of hair volumetric texture into multi-layer 2D texture, mapped them onto multi-grid surface, and then integrated and rendered these grid surfaces to produce hair effects. In 2003, Marschner researched direct lighting and shading rendering of hair. In 2004, Zinke first used multi-scattering lighting, but only transferred Marschner's model into near-field scattering model and used Monte Carlo's path following algorithm. In 2006, Moon[87] proposed the multi-scattering hair simulation method based on photon mapping algorithm. This method first took multi-scattering light of hair as a continuous distribution and the idea was similar to that of the rendering method of participating media. This method has improved the efficiency a lot compared to path following algorithm and has a similar rendering effect.

(4) Simulation and rendering of plant. It is no easy job using the realistic simulating and plant-rendering technology to treat various kinds of appearances with high geometric complexities. Using manual methods to carry out modeling and rendering for plants is very time-consuming and resource-consuming, and there are relatively high technical requirements for art designers. In 1968, Aristid Lindenmayer, an American biologist, proposed L-System whose basic idea is to generate plants using recursive procedure. Now, all 3D plants formed by simulation plant-rendering technologies are indicative.

In recent years, quick image-based 3D plants modeling has received more and more attention. This type of modeling takes several pictures of plant taken from different angles as input and reconstructs the 3D model which looks similar to the real plant in image according to a certain algorithm. The advantages of this method are: 1) easy to input, the user can conveniently use camera to collect plant images; 2) the user only needs simple interaction operation when modeling and rendering. Quan et al.[88] proposed an image-based semiautomatic modeling and rendering technology in view of small flowers, bushes, and large trees. For small flowers and bushes, the method first automatically calculates all shooting parameters of camera according to the images of different angles after inputting many plant images, to generate 3D point cloud structure of plant; and then divides the 3D spot cloud data and two-dimensional image through some simple user operations. The user first divides an independent leaf which could be used as a deformable general model, and matches it to construct other leaves; the user could use data in frame to drive editor to edit plant branches after constructing the leaves, and finally tries to form dimensional plant combining leaves. For trees, the difference from pot plant refers to its leafy profusion and the leaf blade is small. The occlusion relationship is complex in image and difficult to directly model and render for every leaf and therefore it uses a lot of copies of one leaf for assembly to reconstruct the morphology of whole tree. In addition, using the method of using one kind of visible branch figure model to predict the occluded branches, people have effectively reconstructed the branches with complex occlusion structure.

3.1.4 Grid simplification technology. The gradual progress of VR technology and fast development of hardware technology greatly raised the precision of 3D model. As a result, it leads to rapid increase of geometric patches after model grid simplification and has great effect on real-time transmission and processing of the system, so the model is simplified by grid on being processed. Most of the original models are demonstrated with triangle patches and those demonstrated with geometric patches may well be trianglefied. Therefore research on grid model simplification is mainly focused on simplification of triangle grid.

Simplification of triangle grid refers to the use of one simplified model similar to the original one in terms of height retaining the important visual characteristics of it, which contains less triangle patches by deleting less important points and triangles from the original precise model of the object. In 1976, James Clark put forward the thought of model simplification and multiple resolution model demonstration, soon after, many domestic and foreign researchers have carried through extensive and further researches on algorithms. Particularly in the past decade, they have proposed many grid model simplifying algorithms, of which some representative methods are as follows.

(1) Vertex clustering algorithm. The vertex clustering algorithm is to encircle the original model with a bounding box which is then divided into some areas through space partitioning. Thus all the vertices of the original model are left within some small areas. The vertices in each area are merged into a new one. These new vertices get triangulated according to topological relation of the original network to realize the simplified model. This is one of the universal simplified algorithms that do not maintain topology and can deal with any network model of topological types at a quite fast speed. Due to the method is to partition the bounding box of model uniformly; it cannot maintain the characteristic referring to be higher than partition frequency. Meanwhile, the generation of new vertices is only by means of simple weighted average, so the models generated are not qualified enough.

In 2000, Lindstrom[89] proposed an out-of-core simplified algorithm for large-scale polygonal data sets which extents the vertex clustering algorithm with the application of quadric error metric and, better describes the details with lower average geometric errors. He then proposed a memory sensitive technology for large-scale model simplification to improve the algorithm, which enhanced the precision of approximation.

(2) Deletion of geometric elements. Since Schroeder proposed the grid simplification method of vertex deletion in 1992, some geometric element deletion methods such as edge collapse algorithm, triangle deletion method, etc. have been put forward in succession. One thing in common for these methods is to realize simplification through the deletion of geometric elements, i.e. according to geometric topological information of the original models; some geometric graphic elements (point, edge and surface) are under the premise of maintaining some geometric errors. In triangular grid, if a vertex is considered to be coplanar with the triangle patches around it and its deletion does not change topology, then it can be deleted and all the faces connected to it can also be separated from the original model to triangulate the adjacent area and fill the empty holes.

David Luebke and colleagues divided the space with octree. In this way all the vertices of an octree node can be made to pucker together and all the degenerated polygons can be deleted when the volume of the corresponding space of the octree node projecting on the screen is within the designated range.

(3) Visual point correlation method. In 1996 Julie C. Xia brought forward the triangle grid model simplification algorithm based on visual points, which can choose different precision levels in different areas of the same model in real-time. Hoppe[90] defined an elaborate standard based on normal direction of visual cone model and geomet-

ric error of screen space. The standard includes the principles of visual cone, principles of face directivity, and principles of geometric error of screen space. Meanwhile, a model with multiple resolutions was established by the use of this standard for selective edge collapse and vertex split.

(4) Progressive gridding method. In 1996, Hoppe[91] proposed his famous PM algorithm in SIGGRAPH'96. PM algorithm takes edge collapse and vertex splitting as the basic operation, records the locations of original vertices and some new ones and fluctuation information of connection relationship between the vertices in the process of model simplification, and generates one PM representation model composed of the most simplified models of the original model and a series of simplified data. It can present any arbitrary topological network with one highly efficient and non-destructive coding with successive resolutions. In real-time rendering, each piece of simplification information is dealt with vertex splitting reverse operation through reverse tracing simplification information sequence to gradually resume the deleted model details and gain the simplification model of uninterrupted precision of the original model in real-time and actualize smooth transition of LOD model. PM overcomes the drawbacks of the previous models in terms of smooth transition to a great extent supporting real-time generation of grid model with different details. But as for real-time generation of the details of multiple resolutions in different areas of the same grid, PM is still devoid of support of powerful data structure. At the same time it is hard to realize real-time generation of LOD model, for the sequence of edge deletions is independent of geometric topological information of the edges, and thereby it must be judged in the course of model resumption.

3.1.5 Parallel rendering method. Parallel rendering method concurrently operates multiple processors to increase the speed of rendering based on inherent parallelizability of graphic rendering process. One typical graphic rendering assembly line consists of two parts: geometric conversion and rasterization process. The former is to convert geometric graphic elements from physical space into screens space while the latter is to convert the elements into 2D images. There exists a weak data correlation between the elements in front of the graphics assembly line and between the fragments of the pixel bits at the back, and the whole line can be separated into individual modules. So graphic rendering bears functional and data parallelism and is prone to parallel treatment. For instance, in ray tracing algorithm, the conus can be divided into $n$ parts for parallel treatment by emitting $n$ strips of light rays.

Parallel rendering can distinctly accelerate rendering but also introduce extra expenditure that serial algorithm does not have, such as the delay caused by the communication between process and processor, and unbalanced load, as well as extra memory expenditure appropriative caused by data duplication or auxiliary data structure. In order to avoid these overheads as much as possible, such issues as data correlation, division of tasks and data, load balance, etc. shall be concerned completely in designing parallel algorithm.

(1) Parallel rendering classification. Molnar and colleagues divided parallel rendering systems into three patterns: Sort-first, Sort-middle, Sort-last[92] according to the correlation between the stages of parallel task division and typical graphics rendering process.

Sort-first is to distribute graphic elements to each rendering node in the initial stage of rendering and divide the screen into disconnected areas so that each rendering processor takes charge of the rendering of one or more screen areas. It is much shallower intervention in the assembly line, less communication traffic, suitable for software realization, limited expansibility and possible load unbalance. The components may converge in some areas of the screen as the graphic components are distributed between the original processors.

Sort-middle is to distribute graphic elements between geometric conversion and rasterization in the middle stage of rendering. In the method, each geometric processor can process data stream individually, graphic elements are converted into 2D screen coordinate and then distributed to another rasterization processor to be processed after being dealt with geometric process according to the graphic element screen coordinate and lastly all the results of rasterization processor are spliced into the ultimate display image. This method is easy to modularize, conforming to natural shape of graphic assembly line to some extent and fit for realization of hardware. Its defects are that it is easily agglomerated by components so that the load can lose balance and communication traffic increases rapidly as the processors do. Presently, how to solve clustering of components in this kind of hardware/software parallel rendering system it is one of the significant problems.

In Sort-last, rendering processors operate independently from one another till determination of visibility stage, i.e. pixels are distributed once again at the end of rasterization. Similar to Sort-middle, the method is also suitable for realization of hardware and the problem of load balance due to clustering of components can be solved effectively as a result of pixel distribution at the last stage of rendering. However, the capacity of communication is very great in the parallel mode so that generally compression technology is applied to reduce data transmission capacity.

(2) Representative parallel rendering system. WireGL of Stanford University is the first graphics cluster system based on Sort-first system structure. Its main contribution is to advance solutions for all key technologies of such parallel systems of graphics rendering as system structure, task distribution, state tracing, etc. It for the first time, actually realized one Sort-first system structure of mono-layer parallel rendering assembly line. But it has poor expandability and therefore cannot be used in large-scale scenes.

Chromium is the subsequent version of WireGL with the most prominent characteristics. It definitely proposes the conception "stream processing" and raises modularization degree and openness of the system. In the Chromium system, modularization of rendering function is carried out with SPUStream Process Unitas in the rendering process, data stream flows in SPU. Different connections and combinations of various SPU can constitute different system structures like Sort-first, Sort-last or mixture of both of them.

Both Reality Engine and Infinite Reality of SGI apply the structure of Sort-middle. Parallel rasterization processing makes them have great capability of texture processing. The Infinite Reality system has more communication traffics so bus transmission is its main bottleneck.

PixelFlow which is corporately developed by HP and UNC is a typical Sort-last system which applies logic strengthening memory technology to allow a group of rasterization processors to access graphic elements simultaneously and increase the speed of synthesizing images. Meanwhile, it allows programmers to program rasterization processors to guarantee rendering quality. Its defects are that it cannot make sure of exact parallel rendering sequence so the rendering sequence must be designated by the application.

With the continuous development of the parallel rendering technology, composite parallel rendering systems such as Pomegranate and AnyGL present themselves gradually. Pomegranate is a prototype system whose rendering process is divided into five stages: order processing, geometric commutation and illumination processing, rasterization and texture, image synthesis and image display, where each stage can have parallel process, and parallel images can get recombined at all levels of the assembly line. So this system is one of hardware systems of Sort-anywhere. AnyGL is a composite parallel rendering system of Sort-first and Sort-last developed by Zhejiang University which supports large-scale distributing parallel rendering.

3.1.6 Image-based rendering method. The image-based rendering technology applies the existing multi-view object scene image to generate result image under the new viewing angle. In comparison with the traditional geometric graphic-based rendering technology, it is unnecessary to carry through complicated geometric modeling and it has access to less calculation resources. The source image it depends on can be either a real digital image shot or a virtual image generated through geometric modeling of the computer. The complex degree of operation of the method will not change as the complexity of scene does. The following are some representative image-based rendering methods.

(1) Plenopic function and plenopic modeling. To obtain a complete image scene description, Adelson et al.[93] proposed plenopic function to describe the radiation information of rays into the eyes of viewers at any time. Plenopic function is an ideal modeling thought. In fact, because we have no complete access to all the seven parameters of the plenopic function, we have to simplify the parameters required by adding or subtracting parameter restrictions on different application occasions. Optical field and lumen diagram are used for reducing the dimensions of plenopic function, i.e. it degrades from 7D to 4D through some conditional limits.

The optical field approach makes some restrictions for the location of viewpoint by building up a greater convex bounding box outside the small convex bounding box of the object to gain a 4D optical field plenopic function. First of all, gridding is applied on the plane of the video camera to confirm images beneath places of network video cameras, and these images constitute optical plane field whose function is to have access to pixel information corresponding to the optical plane on the basis of the theory of light straight line propagation. As each grid corresponds to an image, the optical field method needs to save a plenty of structuralized images, and the requirements of data memory increase greatly in comparison to other methods. The optical field function for original non-structuralized image collection must do with view choice with high calculation capacity and view regulation.

Lumen diagram is a subset of complete plenopic function to describe optical flow in all directions in the place where it is. Supposing the atmosphere is transparent, the radiation along a ray through space will remain constant. If it is further limited in a ray out of convex closure committing the object, only the value of plenopic function on some surface enclosing it shall be taken into account. In this way, plenopic function formed by the object can be simplified into a 4D function. By virtue of lumen diagram, a new image of the object can be produced quickly without correlation with geometric forms of scene or object or illumination complexity.

(2) Panaroma. The image representing an overall scene is called panaroma, which is regarded as scene independent of visual point. Presently, there are three panaromas: sphere panaroma, polyhedron panaroma, and cylinder panaroma, which respectively look on view space as sphere, polyhedron and cylinder. Both sphere panaroma and polyhedron panaroma can reflect scenes at any direction in space with great processing difficulties. Cylinder panaroma is the simplified form of the former two panaromas actually without the two scenes of top cover and bottom cover, thus limiting the observation angle in the vertical direction, but in the horizontal direction it is a 360 degree angle of view, which can meet the needs of most applications and it is much simple to deal with and therefore has been used widely.

What a panaroma deals with is mainly image stitching and the core is to seek a switch to align overlapped parts between images and connect them into the view of a new bigger frame which involves corresponding relation of matching parts between images and switches to align the corresponding parts.

Concentric stitching method[94] is an extention of cylinder panaroma. Shooting points of source image needed by it is a series of concentric circles and the longitudinal projection to be shot must get tangent to the circle. Its advantages are that the position of a user can move arbitrarily within the range of concentric circle. The concentric stitching method is based on the hypothesis that the scene is infinitely far from the center of the circle. If the hypothesis cannot be satisfied well, the image at the longitudinal axis will deform. What is more, the concentric stitching method is short of longitudinal parallax, which results from equal level shooting. The virtual scene generated by concentric stitching can only roam horizontally, limited by field of view.

(3) View interpolation. View interpolation is to rebuild images of arbitrary visual points according to two given input images. The source images

satisfy some change rule generally. It can set up smooth and natural transition between contiguous sampling point images to reproduce changes in scene perspective shifts between adjacent sampling points. Empty areas of new viewpoint will appear due to projecting area of images and visibility changes in rendering new images with this method. The solution is customarily to apply corresponding offset vectors of pixel colors whose interpolations are contiguous to each other or adjacent pixels in source images to fill up empty areas.

(4) Layered depth image. To process artificial influence in 3D deformation, not only the visible parts of input images but also the back parts of visible surface are to be saved, which is Layered Depth Image (LDI). In LDI, each pixel contains depth and color value and there are many depth pixel points along each visual line; in other words, there are pixels with depth information. In the method, usually layered depth pixel values are stored in a 2D array and a layered depth pixel memorizes many depth pixels along a certain visual line in the sequence from front to back. When rendering with LDI, if the new visual point deviates from the original LDI visual point, the image which is invisible in the first layer will be displayed.

(5) Other rendering methods. Source images in the plenoptic stitching method are shot in real-time according to the defined graphic lattice track by a 360 degree camera and saved in the mode of film frame. The graphic lattice track ensures that scene graphs from any arbitrary angle of view at the interactive route can be acquired from source images after being processed. Theoretically, within the graphic lattice track more extensive modeling space can be gained through the plenoptic stitching method, but the track is based on the defined interactive route so image information of the original graphic lattice track will not well satisfy the needs of a new track once the interactive route is altered.

The view deformation method is to deal with part of the source image by such graphic element processings as deviating, distorting, rotating, etc. to generate new virtual images in virtue of correlation of graphic elements according to certain rules. According to a variety of feature primitive elements in view deformation, there are image deformation algorithms based on control point network, free curve/surface, feature, discrete data interpolation and frequency space.

When we have the depth information of an image, the 3D deformation technology can render the visual points around it. By projecting pixels of the original image to the exact 3D location and once again to the new frames, an image can be rendered from the points around it.

3.1.7 Rendering technology of image combined with graphics. The rendering method based on graphics usually has more known information (geometric information, surface attribute information and light source information) for processing of some special effects such as illumination calculation, object deformation, etc., but it is quite hard making the rendering effects exactly the same as the real world. The image-based rendering method takes the videos and photos taken directly as the input and generates result images. It has high fidelity, but not enough flexibility to operate. So in some VR applications, some methods combining both of them were proposed. In such a kind of methods, geometric consistency and illumination consistency processing is the key point of the research.

(1) Geometric consistency process. Geometric consistency means that virtual objects shall consist with images in real scenes and maintains correct sheltering relationship.

Hayashi et al.[95] proposed to realize geometric consistency calculation between video images and virtual objects generated by computers through visual calculation. Fortin et al. divided the distance between the camera to the scene into different areas and proposed two methods of static scene and dynamic images. Edward et al. found the linear relationship between distance and depth on the basis of the egotistic depth perception theory from the point of view of psychology. Pilet et al.[96] rendered virtual objects on the surface of transformable real ones showed the sheltering relationship between real objects and virtual objects, and processed illumination on the surface of virtual objects.

(2) Illumination consistency process. The illumination consistency means that virtual objects and images in real scenes have consistent illumination effects. Bradley and Roth made use of the ARToolKit to develop such a system to demonstrate approximate illumination and shade, and achieved improvement. Pilet et al. applied some optical models to calculate the information of illumination and shade of virtual object and rendered such a virtual object in possession of reflection and shade. Bradley et al.[97] obtained real information of illumination and shade from input frames with the image-based method, attached virtual object as the grid with 2D texture to the surface of textile and finished rendering of illumination effects in the study on video enhancing reality of textile type flexible objects. Jung et al.[98] proposed a global illumination model which consists with real environment. The model applies high precision algorithm to deal with sheltering and shade and renders virtual objects in multilayer by dint of X3D expanding nodes in a multi-path manner to permeate virtual objects with real environment truly.

### 3.2 Auditory presentation technology

Hearing is the second greatest perception source of human beings next to vision and about 15% of human perception information of the objective world depends on hearing. Hearing can not only accompany the sound for visual pictures, but also can complement information out of vision to enhance the senses of space and realism of virtual world. Therefore, the realistic auditory presentation technology plays an important part in VR research.

3.2.1 Virtual auditory realization. In virtual environment, every vocalizing object is regarded as a sound source, and auditory presentation is to render audio characteristics and spatial information of these sound sources, including sound simulation, synthesis, etc. Most of the existing sound materials are stored in digital form; therefore, analysis and synthesis of audio characteristics of sound source can be processed using the spectrum analysis method of digital signals. However, without special information, stereo and reality sense of virtual auditory space cannot be embodied only by audio characteristics. Sound location in Virtual Auditory Space (VAS) is a key technology for realizing verisimilar 3D audio effect.

In the real world, people perceive spatial characteristics of sound with their ears. Some researchers believe that the spatial location of human ears depends on 1) binaural time; 2) the information of acoustic level difference, including filter effect induced by diffraction of sound wave by head, shoulders and outer ear with auricle in it as well as binaural difference change owing to head turning. Therefore, head related transfer function (HRTF) was introduced to describe different spectrum characteristics of sounds from different locations produced by human auditory system, including information of binaural time difference and acoustic level difference. In virtual auditory space, through acquiring and calculating HRTF, people can make sound wave state which is the same as actual sound source artificially to cause human brain to produce corresponding acoustic images in corresponding space locations.

3.2.2 Acquirement and individualization of HRTF.

(1) HRTF acquirement. HRFT can be acquired through theoretical calculation and experimental measurement. The conventional methods in theoretical calculation include boundary element method, finite element method and infinite element method (IFEM), etc. If a method takes too much time to acquire HRFT through theoretical calculation, then it is not fit for real-time applications. Experimental means are concerned much in recent years, mainly including procedures of acquiring auditory measurement data in real space, abstracting characteristics, predicting data values at non-measurement points with known measurement data and building integrated and continuous auditory space, etc.

Measurement of auditory data in real space is the basis for acquiring HRFT. At present, foreign research institutions for auditory data measurement include MIT Media Lab, University of Wisconsin-Madison and Ames Laboratory Center of NASA, etc. The sole domestic institution is HRFT Database for Chinese specimen established by South China University of Technology[99].

As realization of virtual auditory space involves a large amount of calculation of HRFT, some researches are focused on the simplification of auditory model by abstracting characteristics from measured data to degrade dimensions and condense and realize real-time processing. Currently, some principal element analysis methods are adopted to abstract characteristics.

To establish an integrated and continuous virtual auditory space, it needs to predict data value at non-measurement points with known measurement data. E. M. Wenzel first proposed a simple linear interpolation method based on measured data. Thereafter, many researchers at home and abroad studied HRFT interpolation and proposed such interpolation methods as spline interpolation, inner interpolation based on PCA, nonlinear prevailing learning, etc.

(2) HRFT individualization. Everybody has his/her distinctive ears and head structure, and HRFT individualization setup can improve reality sense and location veracity of virtual auditory space. Theoretical calculation based HRFT can be individualized by adjusting obtained mathematical model parameters, and experimental measurement based HRFT needs to be further individualized.

Zotkin et al.[100] proposed a database match method, which starts with collecting hearers' physiological structure parameters (shape and size of head and outer ears, etc.) using video camera, and synthesized virtual hearing with the most matching HRTF data produced using minimum mean square error method and HRTF database containing these physiological structure parameters. Research workers in Nagoya University, Japan, proposed PCA method[101], which decomposes HRFT into weight of base function through PCA, established the relationship between weight values and physiological parameters with regression analysis method and realized HRFT individualization. The result indicates that the method has some feasibility and relatively high precision. Structure model method was put forward by Brown et al.[102], in which HRTF is divided into three particle filters and HRFT individualization is actualized by confirming the relationship between physiological parameters and particle filter coefficients. It has high-speed and is convenient for real-time processing, but these advantages are yet to be validated by further experiments. Frequency scale method was proposed by Middlehrooks[103], whose principle is to realize HRFT individualization by seeking optimum scaling ratio usable for describing proportional relation of HRTF of two hearers in frequency domain. It is also to be validated by further researches.

3.2.3 Distance sense presentation of sound. If the distance from sound source to head center is within 2 m, sound distance information only depends on HRFT; otherwise, distance information will be related to refection parameters of virtual environment such as sound intensity, air absorption, reverberation, refection, etc. Therefore, the sole use of HRFT in auditory location will cause abnormal phenomena such as front back overturn and sound in head, thus losing distance information of sound. Now we can simulate environment sound field by virtue of measured data or geometrical acoustics model.

### 3.3 Force/tactility presentation technology

What vision and audition provide is non-contact perception information. By providing contact perception information of virtual object, the user's sense of reality, immersion can be directly enhanced and the application field of VR can be expanded. The purpose of force/tactility presentation is to enable user to obtain vivid force/tactility experience in contact interaction with virtual objects.

The presentation of force/tactility should be assisted by force/tactility equipment and at the same time requires handling force/tactility information effectively, which mainly involves collision detection and collision response. When users operate force tactility equipment, collision detection algorithm serves to detect collision conditions with virtual objects. If there is any collision, collision response should be given according to the strategies formulated in advance and relevant force/tactility information of virtual objects should be provided. As human tactility is acute, the rendering of force/tactility requires the refresh frequency to be at least 1k Hz so that people can obtain the sense of reality. Therefore, high fresh frequency and rapid calculation becomes the key to the research of force/tactility presentation technology.

3.3.1 Collision detection. In force/tactility presentation, people usually adopt graphical technologies such as spatial partitioning and hierarchy bounding boxes, etc. to detect collision. For detailed information of collision detection, please refer to section 2. Here we will not discuss it any more.

3.3.2 Texture mapping of tactility. Most of the objects in the natural world have their surfaces covered with a certain kind of texture, which can be sensed and distinguished by humans through their tactile system. In VR presentation technology, tactility texture can be simulated by appropriate disturbance of counterforce vector. And the counterforce can be obtained through calculation of objects' geometrical and physical properties. Texture mapping of tactility can generally be divided into image-based texture mapping and process-based texture mapping[104].

If the surface texture of 3D objects is mapped directly from a 2D image, image-based texture mapping should be adopted to map the color and grayscale information in 2D texture. First 2D image is mapped on a simple medium surface like plane, cube, cylinder or sphere; and then the texture is mapped from medium surface onto object surface. After two mappings, people can conveniently get height value of any point and its gradient value through finite difference calculation method.

The goal of process-based texture mapping is to produce synthesized texture by using mathematical function. The inputs of functions are texture coordinates $(x, y, z)$, and the outputs are height and grads of corresponding points. Some researchers have successfully succeeded in producing random tactile texture through noise texture. Besides, fractal technology can also be used to simulate natural texture, because many objects have shown self-similarity to some extent.

3.3.3 Six degrees of freedom. According to the number of optional variables, force/tactility interaction can be divided into three degrees of freedom (DoF) and six degrees of freedom. Variables in 3-DoF represent axes $X$, $Y$, $Z$ respectively; for 6 DoF, besides spatial coordinates, there are such variables as rotation, slope and sway. As 6 DoF can provide more force/tactility information, different rendering methods for 6 DoF force/tactility presentations have been proposed. These methods can be classified generally into direct rendering method and virtual coupling method[105].

Direct rendering method is one type of the geometric graphics methods, which needs no rigid body dynamics simulation and thus is easy to be realized. The penetration depth value might be rather big, that is, there are more mutual penetrations between objects. In this case, the stability of the system will be affected when the frequency of feedback and refreshing is relatively low. Virtual coupling method can reduce penetrating degree between objects and has reliable stability, which can perform force/tactility presentation more naturally. But it needs rigid body dynamics simulation and thus can hardly be realized. If there is no limitation to computing resource, the ideal 6 DoF rendering method is to perform rigid body dynamics simulation based on restriction and do force feedback calculation by virtual coupling method. But limitation of computing resource and requirements of force/tactility feedback refreshing frequency are a pair of conflicts. Therefore, 6 DoF rendering algorithm with complicated models and scenes is a huge challenge.

## 4 VR interaction mode and equipment

VR human-machine interaction is essential for users to obtain vivid perception by operating various virtual objects in virtual environment, which mainly involves information exchange modes and equipment of interaction and mutual influence between people and virtual environment. The content here involves the function do in VR 8-element set described in section 1 to a great extent.

VR human-machine interaction has very close relation with human-computer interaction in computer system. There are some modes in common and also many distinguishing features as well. The human-machine interaction in computer system focuses on information transfer between man and machine while VR human-machine interaction emphasizes perception transfer between man and virtual environment. The vivid virtual scene display, real tactility/force sense perception, 3D spatial azimuth tracking, interactive behavior information exchange and the like have become the important parts of human-machine interaction technology in VR system. Comparatively speaking, VR human-machine interaction requires higher naturalness and sense of reality.

In this section, we analyze the conditions and development trends of VR human-machine interaction technology according to interaction requirements and typical application of VR system and summarize VR human-machine interaction modes and equipment in terms of scene display, tactility/force sense interaction, azimuth tracking and walking interaction, based on the properties of perceptibility, orientability and operability VR human-machine interaction possesses.

### 4.1 Scenes display mode and equipment

It has been shown that 70%–80% information people get is through vision. Therefore, vision system is the most important perception approach of VR. At the same time, virtual scene perceptibility is the prerequisite for human-machine interaction in virtual environment. Therefore, scene display mode and equipment are fundamental components for human-machine interaction in VR system. In this section, we give a brief introduction of the head-mounted, desktop, projection, hand-held and automatic stereo display modes and equipment.

4.1.1 Head mounted display. Since its introduction into the world, head mounted display (HMD) has been used widely to some VR systems to present 3D virtual scenes for users. In a non-perspective head mounted display system, two displays are installed inside the helmet near to eyes. HMD moves along with the head so that HMD position tracker can detect the position and direction of the head, and the computer will render the scenes under current viewpoint on HMD displays according to the position and direction of head. As users can only see the 3D scenes generated by computer on non-perspective displays, it is easy for them to feel immerged, and at the same time it is more likely to lead to eyestrain, dizziness, etc.

Perspective HMD is commonly used human-machine interaction equipment in reality enhancement systems. For video perspective HMD, it can be used to get real environment information by binocular vidicon installed in front of the helmet and the computer can be used for real-time information stacking data, words, graphics, etc. in vidicon video so that users can sense the enhanced real scenes of virtual geometric objects integrated with real environmental videos. Besides, optical perspective HMD can be used to get real environmental information through the binocular optical synthesizer in front of glasses[106] and users can observe the surrounding real environment through optical synthesizer as well as the information such as data, words and graphics generated by computer.

4.1.2 Desktop display. Desktop is a kind of working environment used and required by many users. In a desktop display system, the scene image in virtual environment is projected on horizontally placed display equipment so that users can fulfill interactive operation on horizontal surface of a workbench. Desktop display system is mainly composed of workbench, projector and computer[107]. There are reflector and desktop display screen on the workbench. The scene image generated by computer is projected onto reflector by a projector and further reflected onto display screen by a reflector. The scene image on the display screen can represent 3D virtual objects as well as menu of operable system tools and interface.

Assisted by stereo glasses, multi-users can sense 3D scenes in virtual environment and the position and direction of users' viewpoint can be ascertained by tracking equipment of desktop display system. In view of display characteristics and application effect, desktop display system is suitable for electronic graphs rendering and digital design. And with it, teachers may make explanation and demonstration using virtual objects, and multi-users may realize cooperation in network environment. But the immersive sense of 3D virtual scene generated by a desktop display system is not strong.

4.1.3 Projection display. CAVE (cave automatic virtual environment) projection display system is a typical polyhedral projection virtual scene display system, which can allow multi-users to sense vivid 3D virtual scene simultaneously. In CAVE system, multiple projection displays serve as display "faces" of virtual scene in different locations. Three or more conjoint display "faces" can make up a cave-like cube[108]. Usually, projector is installed outside the "face" so that the virtual scene image generated by computer may be projected to the screens of all "faces". CAVE projection display system can represent 3D virtual scene in directions of front, left, right, up and down respectively so that users can acquire vivid visual sense and "immerge" in virtual environment, while application and popularization of the system are often restricted by its high price and space-killing characteristics.

Comparatively speaking, wall-mounted projection blending display system is a virtual scene display mode that is economic and easy to generalize. Generally, it is composed of computer, large screen, fusion cage and multi-projector, which can represent 3D virtual scenes with a wide field of view by the aid of an annular or cylindrical large screen and two or more projectors. And the fusion cage is the key component to wall-mounted projection blending display system, serving to blend mage information of different display models. Interactive projection display system is a human-machine interaction mode for reality enhancing system. For example, the virtual and real combined snow mountain and grassland are projected onto the ground and step information of the user is obtained by a vidicon. When the user walks across the "snow" ground, there will be distinctive footprint left along the treading line on the ground.

4.1.4 Hand held display. With the quick development of mobile computing device and wireless network technology, Personal Digital Assistant (PDA) and Smart Phone (SP) have already possessed intensive capability in computing, memorizing and transmitting, especially in graphs and video handling. The examples include 3D graphics library of embedded computer system, OpenGL ES, 3D graphics library of mobile computer system, Direct3Dm and 3D graphical interface of mobile computer system, M3G, etc., which enable mobile computing device to represent 3D scene in virtual environment and support multi-model interactions such as GPS location, speech recognition and handwriting recognition, etc. The mobile computing based hand held display is becoming more and more powerful in capability, drawing more and more attention and thus has an extensive application prospect.

The mobile computing based hand held display is also an important human-machine interaction mode for enhancing real systems, especially some application systems having few demands of immersive sense, in that it has great advantages in the aspects of interactivity, portability, mobility and security, over head mounted display. Hand held mobile computing equipment can display quickly enhanced real scene that combines 3D graphics model with real scene video[109]. Holding mobile computing equipment for getting video information in football court, and facing the virtual goalkeeper, the user can control kick direction by adjusting the held equipment's orientation to experience virtual kick shooting.

4.1.5 Automatic stereo display. Whatever head mounted display, desktop display or projection display, stereo display equipment is a necessity (e.g. stereo glasses) for the user to have stereo sense of virtual scene. Considering users would feel uncomfortable in a way when wearing stereo display equipment, people begin to do research on many kinds of "automatic" stereo display modes and equipment, which permit the user to experience the stereo effects of virtual scenes without getting tired of any device. One kind of automatic stereo display shown on SIGGRAPH conference in 2007 can represent stereo effects of 3D model well[110]. The user can experience automatic stereo 3D virtual scene in a 360° range around the display and operate 3D virtual scene represented by automatic stereo display through interaction equipment.

### 4.2 Force/tactility interaction mode and equipment

A lot of VR applications require the user to sense the force and tactility effect generated by objects in virtual environment; therefore, the VR system needs force/tactility interaction modes and equipment, such as force feedback joystick and tactility data gloves.

4.2.1 Force feedback joystick. The mechanical components of a joystick installed with force feedback equipment can produce counterforce responding to the user's operation and relevant computing mode for virtual environment, thus enabling the user to experience human-machine interaction force in virtual environment. The interaction freedoms and accuracies differ from one another according to the force feedback joystick.

A typical force feedback joystick is able to provide 3-DoF position induction and angular measurement in a certain interaction range. Free axial direction of joystick is connected to an electromotor of the equipment through gear or steel wire, and the electromotor generates mechanical movement according to interaction information processed by force feedback chip to drive the high intensity force sensation equipment to provide force feedback of multiple DoFs to make the user have force perception in human-machine interaction process. For example, based on the computing model for virtual vehicles, force feedback joystick can enable the user to have a tossing sense as caused by bumpy road and to experience the counterforce as caused by turning steering wheel.

The main application fields of force feedback joystick are virtual design, virtual assembly, operation training, remote operation, military exercises, etc. Force feedback joystick of different performance is to be provided for different VR systems according to interaction requirements. For example, for comic and animation design, it is appropriate to use lower accuracy force feedback joystick, while for medical operation training, higher accuracy force feedback joystick is required.

4.2.2 Tactility data gloves. Most of human interactions are performed by hands. Data gloves can be used to get palm and fingers information on position, direction, gesture, thus helping the user to fetch, move, assemble and control various objects in virtual environment.

Usually, data gloves are made of light weight elastic materials, like optical fiber and can be easily installed as a sensor. Each figure knuckle is equipped with an optical fiber ring to measure its bending angle; position sensor is used to test the position and direction of hand, holding or stretching state of fingers, upwarping or bending angle of fingers, etc. At the same time, data gloves are equipped with a sensor composed of multi-pair of strain resistance chips to get interaction information of fingers in virtual environment through testing the signal change of strain resistance chips.

The generation and presentation of tactility sense in virtual environment is more difficult than force sense perception. At present, the main methods for developing data gloves with tactility feedback are pneumatic sensation, vibration tactility, electron tactility, and nerve and muscle simulation. Inflatable tactility data gloves fulfill the task of inflating and exhausting for device according to human-machine interaction by making use of bubble of tactility device. In the course of inflation, the bubble expands and presses the ginger skin such that tactility feedback of data gloves is produced. On the basis of voice coil principle of loudspeaker, vibrating tactility data gloves generate tactility feedback of data gloves through stimulating skin by the vibration device on finger back.

### 4.3 Tracking location mode and equipment

Tracking localizer is a tracking device to ascertain 3D spatial location in VR system, providing location information of tracking targets for stereo glasses, head mounted display and data gloves, etc. It enables the user to have free moving interaction space and hence more flexibility for interaction operation. For active tracking location mode, it is equipped with an emitter and a receiver and thus are able to determine the position and state of tracked objects by emitting and receiving physical parameters of signal, while passive tracking location mode has no active signal source, and can only be used to determine the position and state of tracked objects through measuring the changes in received signal.

### 4.4 Walking interaction mode and equipment

Walking interaction mode deals with the cases where the user walks on interaction equipment. The relevant information of the user's walking activity is transmitted to virtual environment by interaction equipment. Walking interaction involves interaction movements like advancing, turning around, going up and down the slope, walking over obstacles and changing poses. We will introduce the following pedal walking, ground walking and walking according to equipment characteristics.

4.4.1 Pedal walking interaction. The bicycle simulator fixed onto the ground is a typical pedal walking interaction, which can simulate pedal resistance by making use of braking frictional force, inertia and viscosity by flywheels and spokes, up and down slope through adjusting pedals state by an electromotor, change in walking direction by handlebars and can measure the scope and direction of interaction movement by position sensors fixed on pedals and handlebars.

We can transplant the pedals of a bicycle to an up-down step walking support[111]. The user stamps on the pedals which is installed with a force sensor and goes forward in virtual environment by making use of pedals' counterforce. He can change walking direction in virtual environment through left and right rotation of pedal support. This support pedal walking interaction equipment enables the user to experience changes in rise and fall and the softness of ground when his walking speed is not too fast.

4.4.2 Ground walking interaction. The ground walking interaction mode does not need pedals; using sensors it measures directly the position and direction of the key points of the user when he walks[112]. The sensors are fixed on the shoe-pads, knees, thighs, waist, etc. Shoe-pads' force sensors are to get walking step information; knees' sensors are to get information of height, speed and direction of relevant points; waist's tracker is to get information of position and direction of relevant points; head mounted display and its sensor are to get information of location of head and viewpoint. Ground walking interaction mode possesses the advantage of free walking, while it is low in accuracy and long in delay at present.

4.4.3 Transmission platform walking interaction. Omnibearing transmission platform[113] walking interaction includes two orthogonal transmission platforms, which are composed of rotating rollers, and transmission belt of each platform has 3400 rollers. By making use of actuating motor and rotatable platform support, rollers can get information about the changes in position and direction of steps through rotating of rollers. The brace system the user wears is connected with a position tracker installed above the user, which can get information about the changes in position and direction of the key body parts as the user is walking. At the same time, it can provide the user with 3D virtual scene through CAVE projection display system or head mounted display. The omnibearing transmission platform walking interaction mode has the advantages of high accuracy, excellent real-time quality and more natural interaction mode although it is quite noisy and not very steady.

## 5 VR development suite and support environment

VR technology has got many successful applications in various fields such as military exercise, aerospace and quick industrial design since it emerged. With large demand pull and application department promotion, many kinds of VR system development tools have come into being since the early 1990s. It plays an important part in lowering VR system development threshold and improving development efficiency, and thus greatly promotes the development of VR application.

There are various tools for developing VR system, some of which support different component module development of VR system, and others support development of VR application systems in a certain field. In this section, some modeling tools, rendering tools, distributing development suites and other specialized development tools are described.

### 5.1 Modeling tools

Modeling is the first step to establish vivid and interactive VR application systems. The existing VR modeling tools, including animation-oriented modeling tools and real-time rendering-oriented modeling tools generally, focus on appearance and physical modeling to support virtual sceneries.

5.1.1 Animation-oriented modeling tools. Animation-oriented modeling, also named 3D geometric modeling design is a fundamental 3D animation tool. The modeling of animation generally involves rendering of basic geometric forms and combining of complicated models. Some published 3D model bases are available to promote development efficiency.

At present, the popular 3D animation tools include Maya developed by Alias Co., Softimage XSL by Avid Co., Houdini by Effects Software Co., 3D Studio Max by Discreet Co., Lightwave 3D by Newtek Co., Photorealistic Renderman by Pixar Co., Mental Ray by Mental Images Co., etc. Maya and Softimage XSL are advanced 3D animation tools, which are widely applied in TV & film production industry. 3D Studio Max is 3D animation software on PC platform, which can be transformed into other model forms conveniently with open interfaces and has many plugs in support from the third party and thus is widely applied in game and industry fields.

5.1.2 Real-time rendering-oriented modeling tools. 3D model data organization format has great impact on real-time rendering, which is not applicable to 3D animation model format. The most typical format is OpenFlight of Multigen, which accepts the model data standard in visual scene simulation field and is supported by most of VR development tools (such as VEGA, OpenGVS, etc.). MultiGen Creator is an interactive 3D modeling software developed by MultiGen-Paradigm Co., USA, which helps to build optimized 3D models with the functions like polygon modeling, vector modeling, precise generation of large-scale terrain, etc.

As heavy workload of manual modeling is required in handling some large-scale terrain modeling, some geographical data-oriented terrain generation tools emerge, such as Creator Terrain Studio, Terra Vista, etc. Creator Terrain Studio supports various vector data forms and has a great number of terrain data supports. Terra Vista is Windows-based real-time 3D terrain generation tool software, applicable to mass data terrain generation. Terra Vista has point, line and plane editors. It can not only enable developers to edit and modify cultural characteristics data of vector interactively but also generate relevant cultural characteristics in a direct and automatic way such as road, forest, vegetation, etc. through customizing attribute information of cultural characteristics data.

Some auxiliary tools for possessing certain functions have been explored to improve modeling efficiency, including format transformation tools such as Polytrans, DeepExploration, etc. These tools can be used to transform animation design-oriented 3D model data format into another format for real-time VR system. 3D model simplification tools of Geomagic Decimate, Action3D Reducer, Rational Reducer and the like can be used for simplifying animation-oriented 3D model to satisfy the requirements of real-time rendering.

Besides, there are also some modeling tools for depth images such as software system attached with video capture devices and laser scanners, which are used specifically to process depth images captured by relevant devices and generate geometrical models.

5.1.3 Web 3D standard and modeling tools. The initial Web 3D standard is VRML (virtual reality markup language) describing geometric dimensions, shape, color, material, illumination, etc. of 3D sceneries. VRML was issued formally as an international standard in December 1997 and approved formally by International Standards Organization (ISO) in January 1998 with VRML97 for short[114]. VRML 97 enables 3D model files delivery in Internet. But due to its excessive interpretation, VRML has low processing efficiency. Many Internet graphic software do not follow VRML 97 standard completely. As a result, many companies develop their own authoring tools and use special file formats and browser plug-ins.

In 1998, VRML organization was renamed Web 3D Consortium and worked out a new standard, Extensible 3D (X3D)[114] at the same time. X3D consolidated the newly developed technologies like XML, Java and stream transmission for the purpose of improving processing ability, rendering quality and transmission speed. In 2002, Web 3D Consortium issued an X3D standard draft. In August 2004, X3D standard was approved by ISO as ISO/IEC 19775 international standard. But until now, X3D standard has not unified Web 3D formats with some powerful rivals like U3D (Universal 3D) Standard[115] supported by 3DIF (3D Industry Forum) Union composed of Intel, Microsoft, Macromedia, Adobe, Boeing, etc.

There are also some image-based modeling tools for Web-oriented application, such as Canoma, Photo3D, PhotoModeler, ImageModeler, etc.

5.1.4 Physics engine. The vividness of virtual environment and objects is dependent on appearance modeling level and physical attribute modeling of virtual objects as well, i.e. the realization for physics engines. Physics engines originated from simulation techniques such as technique for simulating flight attitude, missile trajectory, etc. Because of the complexity of physical computation, the development of physics engines had been progressing very slowly for a long time, compared with that of graphics rendering and graphics engines. In the recent years, the further progress in VR application has placed urgent demand on physics engines and with continuous development of computer hardware at the same time, information processing competency has become faster, providing necessary conditions for developing physics engines as well.

Physics engines are the program for object motion, scene change, interaction and dynamics feature effect of object motion, scene change, object-scene, object-object in computing virtual environment, usually provided in the form of program library that contains several function modules each of which will set aside interfaces for application program. Physics engines define a high-rise API assemblage, which encapsulates all rock-bottom physics computing details so that the developers can focus on high-rise application program development and thus shorten development periods substantially. Table 1 shows a comparison of their main function indexes of several well-known physics engines.

Havok Physics, a physics engine developed by Havok Co. based on rigid body dynamics, can simulate restriction and joint of multi-jointed rigid body. Developers can specify physical parameters of the object, such as quality, density, friction coefficient, etc. Havok Physics introduces continuous collision detection technology, including Rag Doll human body model system, which can exert various dynamic effects on vehicles in virtual environment, including simulation of collision between vehicles and various operations.

Havok Physics has good cross-platform performance, and is applicable to PC mainframe and many game mainframes. Havok Physics is also applied to such software as 3DS Max, Maya to provide vivid physical manifestation of various virtual scenes in the form of plug-in. Havok Physics has been one of the most widely applied physics engines so far, involving massive computing capacity and thus having demanding requirements on hardware like CPU.

In order to solve the heavy physics computing capacity, AGEIA Co., USA developed special physics accelerating hardware. The company put forward the concept of PPU (Physics Processing Unit) at GDC05 Conference in 2005. PPU is an important specialization of processor function after GPU. GPU is used for processing graphics and images, which increases graphics rendering effect by a large margin. PPU further separates physics computation from CPU and provides hardware with accelerating supports and thus enhances physics engine effect obviously. At GDC06 Conference in March, 2006, the first piece of physics processing unit was issued formally and was named PhysX.

**Table 1** Comparison of function index for mainstream physics engine

| Physics engine | Support joint type | Collision primitive | Integrator type | Acceleration mode | Platform |
|---|---|---|---|---|---|
| Havok Physics | ball-socket, hinge, spring, etc. | sphere, plane, convex polyhedron, triangular grid | Eulerian RK approach | use GPU | PC, Xbox/Xbox360, GameCube/Wii, PS2/PS3/PSP |
| PhysX | ball-socket, hinge, slide block, etc. | sphere, cube, capsule, convex polyhedron, plane, triangular grid | | use PPU | PC, Xbox360, PS3 |
| ODE | ball-socket, hinge, slide block, dead | sphere, cube, plane, capsule, triangular grid | Eulerian method | | PC |
| SPE | provide basic joint functions only | sphere, capsule, triangular grid | Eulerian method | | PC |

PhysX and Havok Physics have much in common and the difference between them is that the former is supported by PPU and thus its efficiency increases greatly. AGEIA Co. claimed that PhysX can process 32000—50000 rigid bodies per minute, while other "Physics Engine + CPU" modes can only support hundreds of rigid bodies under the same conditions. So far PhysX has been applied to various games but is rather expensive.

ODE (open dynamics engine) is a kind of open source physics engine, which is easy to use, fast in speed and with all source codes open but it has limited functions available. ODE is applicable to multi-jointed rigid body, which supports joint types such as Ball & Socket, Hinge, Slide Block, Dead Axle, etc. and various collision primitives (e.g. spheres collision and planes collision) but not continuous collision detection. Worse still, it is low in collision detection speed.

Although the research and development of physics engines in China started very late, the progress is very fast. SPE (simple physics engine) is the typical representative of physics engine developed in China. SPE supports rigid body simulation and continuous collision detection under certain conditions and provides liquid simulation system based on particles.

From the above, we can conclude that the present physics engines are mainly focusing on supporting multi-jointed rigid body motions and most of them are provided with abundant joint types and precise collision detection algorithms. But they are weak in supporting deformable objects and flexible objects or even do not support them at all. The research on physics engines is a serious subject in VR field, and at the meanwhile, is bumpy and challenging with heavy responsibilities.

### 5.2 Rendering tools

During the development of VR, many rendering tools have come up with different real-time or application orientations. Basically, they can be classified into three levels from technical viewpoint. The basic 3D graphic rendering library is in the bottom level that supports a series of graphic rendering standards API. 3D graphic engine is in the second level and the visualized development suite of it is in the top level. These three levels possess different characteristics and uses, and they will be introduced as follows:

5.2.1 Basic 3D graphic rendering library. Basic 3D graphic rendering libraries mainly refer to OpenGL, Direct 3D and Java 3D that operate graphic hardware directly and offer the bottom basic API for 3D graphic rendering.

OpenGL is an open 3D graphic software kit possessing seven functions including modeling, conversion, color mode setting, illumination and material setting, texture mapping, bitmap display and image enhancement and double buffer memory animation. OpenGL is independent of window system and other operating systems, and the application program developed based on it can be transplanted between different platforms conveniently.

Direct 3D offers 3D graphic API based on Microsoft COM interface standard, possessing good hardware compatibility and supporting many latest technological achievements of graphics. Now almost every display card with 3D graphic acceleration supports Direct 3D. But the interface is relatively complicated and can only be used on Windows platform.

Java 3D API is defined by Sun, serving to program interfaces of 3D graphics and Web 3D application program. In addition to a part of bottom rendering functions defined by OpenGL and Direct 3D, Java 3D API also provides many high level functions for construction of 3D objects. Therefore, viewed from the levels located, Java 3D has the functions of both basic 3D graphic rendering library and 3D graphic engine.

Direct use of basic graphic rendering library can offer complete and independent control of the bottom details of an operating system. It has good flexibility and can represent the demands of application systems much better. However, due to high technological requirements for system developers, more complexity and longer period of system development, the scale of application system is also limited. Therefore, the present application system development is not that much carried out directly with these basic APIs.

5.2.2 3D graphic engine. 3D graphic engine offers a complete software development support that orients the real-time VR applications, takes charge of data organization and processing for 3D graphic rendering at bottom level, brings the acceleration feature of hardware into play, and provides an effective graphic rendering support for application programs at the upper level. Graphic engine usually includes graphic rendering of third dimension, 3D scenery management, sound management, collision detection, terrain matching, and real-time object maintenance. In addition, the engine also provides high level API related to 3D virtual environment rendering. Graphic engine has its own VR application program frame. Based on the frame and high level API, developers can create their own application programs conveniently. The common 3D rendering engines include OpenGL, Performer, OpenGVS, Vega, OSG, V Tree and WTK.

OpenGL Performer is an expandable software kit for real-time 3D vision development which is structured based on OpenGL graphic library. It provides not only a group of programming interfaces with standard C or C++, but also high performance rendering ability using a flexible 3D graphic toolset. Performer supports multiple CPU systems that can realize the parallel graphics rendering.

OpenGVS is directly structured on the APIs of 3D graphics like OpenGL, Glide and Direct 3D. It includes a group of object-oriented C++ APIs that are packaged with complicated bottom level graphic driving functions. These APIs are classified into different groups of resources including camera, passage, frame buffer, fog, light source, object, scenery, specificity, and tools which are invoked to drive hardware and create required graphics in the real-time by developers.

Vega includes two main types of configuration, namely VEGA-MP (Multi-Process) for multiprocessor and VEGA-SP (Single-Process) for uniprocessor. Both of them support various data inputs as well as CAD data conversions. Vega provides varied optional modules supporting simulation demands of navigation, infrared ray, radar, illumination, animation character, large-scale terrain database management, CAD data input and DIS/HLA distributed application.

OpenSceneGraph (OSG) is an open source 3D graphic development library based on OpenGL which offers a set of C++ API, possesses relatively complete 3D graphic development functions, and optimizes the rendering performance through status conversion, drawing tube and custom operation. OSG mainly includes four parts—scene graphic core, Producer library, Open Thread library and custom plug-in package.

VTree is one of the object-oriented 3D graphic development libraries that includes series functions of or related to C++ VTree creates and connects different nodes with the attached virtual tree structure of the scenery entity which defines rendering and handling method for the entity.

WTK is one of the 3D graphic development libraries based on OpenGL and Direct X which is developed by Sense8 Co. of America, capable of running in operating systems like Windows or Linux. WTK offers a function library based on C/C++ which is classified into more than 20 categories related to 3D graphics like universe, geometry, task, material texture, illumination and others. Through these categories and hardware interface supported by WTK, customers can create or manage virtual scenes and realize interaction between reality and scene.

Web 3D plug-in package for explorer is also a kind of real-time graphic rendering engines which can be used to explain grammar of scene model document or analysis for model document format. It can carry out a real-time rendering for scene model documents from server in customer's web explorer.

5.2.3 Visualization development suite. In recent years, some VR visualization development suites come up, by which most of the VR application systems can be realized with common functions through configuration and editing of GUI. Consequently, these suites lower the technical requirements of developers. However, because of the high level of software, the VR systems developed by those development suites have low operation efficiency and these systems may be used in prototype system creation or other location without high real-time requirement. Recently, popular visualization development suites will include Virtools Dev of Dassault Co. in France, EON Studio of EON Co. and Quest3D of Act 3D Co. Some game development editors developed by the gaming companies also possess similar features.

Virtools Dev is composed of development environment program, behavioral engine, rendering engine, web player and SDK. The video camera, illumination, curve, interface element can be created by clicking the icons easily. Behavioral engine of Virtools offers many behavior modules that can be used repetitively and also offers VSL language as supplements of graphic editor through development by SDK. Similar to Virtools, EON Studio and Quest3D also offers repetitively usable function modules and GUI editors and plentiful simulation physical models. Developer can create a certain VR application system rapidly only via selection and configuration.

Because the graphics rendering tools are expensive, Beihang University has developed a real-time 3D graphic platform, BH_GRAPH[116]. BH_GRAPH consists of 3D vision rendering engine, 3D object modeling tools and 3D scene distribution tools. Vision rendering engine is a complete SDK oriented real-time simulation system including graphic display of third dimension, 3D scenery management, terrain matching, object interaction, real-time object maintenance and other functions. It supplies a set of high level 3D graphic rendering development libraries which are object oriented and mainly contain two parts, i.e. API and running library for real-time scenery management. This platform can handle geometric models with general format like Open Flight, 3DS and others.

5.2.4 Comparison of rendering tools. Graphic tools in three levels possess their own characteristics and application scopes. Tools in lower levels require high professional level of developer with more development difficulties but much higher flexibility and operating efficiency, while tools in upper levels possess less difficulties but lower flexibility and operating efficiency.

Most of 3D graphic rendering tools have not opened their architecture, key technology and other technical details to public. Therefore, comparison and analysis between them can hardly be carried out completely. We only make a comparison between graphic rendering functions in Table 2. Due to the limitation of data resources, some contents may be not all-sided.

### 5.3 Distributed VR development suite

Existing distributed VR supporting tools are mainly of two types: virtual environment server and distributed development suite. Distributed VR systems adopting virtual environment server

**Table 2** Comparison between 3D vision rendering engine functions

| | OpenGVS 4.4 | Vega 3.7 | OSG 0.99 | Virtools 3.0 | BH_GRAPH1.0 |
|---|---|---|---|---|---|
| Model type | multiple | multiple | multiple | single special format | multiple |
| Interface mode | API | API/IDE | API | API/visualized development environment | API/IDE |
| Encapsulation mode | function | function | object | object | object |
| Scenery management | viewing volume cutting, LOD (level of detail) mode | viewing volume cutting, LOD (level of detail) model | viewing volume cutting, LOD (level of detail) model, occlusion culling | viewing volume cutting, LOD (level of detail) model, occlusion culling, entrance culling occlusion culling, image proxy | viewing volume cutting, LOD (level of detail) model,occlusion culling, image proxy |
| Texture capability | monolayer texture, mipmapping | mono- & multi-layer texture, mipmapping, video texture | mono- & multi-layer texture mipmapping, projective texture mapping | mono- & multi-layer texture, mipmapping, video texture, bump mapping procedural texture | mono- & multi-layer texture, mipmapping, video texture, bump mapping |
| Illumination method | pixel illumination | pixel illumination | peak/pixel illumination | hardward acceleration peak/pixel illumination, illumination mapping, HDR | hardware acceleration peak/pixel illumination, illumination mapping, IBL, HDR |
| Real-time shadow | none | none | projected planar, shadow volume | projected planar, shadow volume, shadow mappine | projected planar, shadow volume, all-frequency real-time shadow |
| Special effect | fog, smoke, skybox | Environment mapping, notice board, skybox, mirror surface, water, explosion, flame, weather effect | Environment mapping, notice board, skybox, fog, mirrir surface | Environment mapping, notice board, skybox, fog, lens flare | Environment mapping, notice board, skybox, weather effect, fireworks, mirror surface, lens flare |
| Physical characteristic | none | collision detection and others | collision detection | collision detection and others | collision detection, simple collision response |
| Sound effect | none | 3D sound | none | 3D sound | 3D sound |

have a strong coupling in modules. It means that the program codes of VR server and client should be developed as two parts of a system. Some VR development suites with high integration such as Virtools, EON Studio possess independent virtual environment server module. Distributed VR development suite is formed by abstracting those relevant to distributed process in distributed VR and other supporting services.

IEEE approved the standard of IEEE 1278—DIS[117] (distributed interactive simulation) in 1993. It enhanced the application development of distributed VR. In 2000, another standard of IEEE P1516 HLA[118] (high level architecture) was also approved. Also in the same year HLA 1.3 became a compulsory standard for military simulation systems in US army. HLA is designed to create a general high level simulation system structure that can achieve interoperation and reusability between varied models and simulation which include four aspects: rules, interface specification, object model template and VV & A (verification, validation and accreditation). RTI (run-time infrastructure), the infrastructure software of HLA, provides six types of services, namely federation management, declaration management, object management, time management, ownership management and data distribution management.

RTI has great effect on the scalability and efficiency of distributed VR application systems. USA, UK, France, Japan, Canada, Sweden and other countries give supports to RTI development, and up to now some RTI systems have been established. However, due to lack of some key technologies and complex integrated testing, practical RTIs are not so many. Among them, DMSO RTI, RTI-s, MAK-RTI and pRTI have a relatively long history.

US Department of Defense established DMSO (Defense Modeling & Simulation Office), and supported MIT Lincoln Laboratory to explore the prototype of DMSO RTI. RTIs used by US army in STOW (Synthetic Theater of War) 97 Project was also developed during this period. Due to the special background and a long term free release and use of military RTI, DMSO RTI possesses an important position internationally. DMSO RTI only conforms to HLA 1.3.

pRTI of Pitch in Sweden is one of the commercialized RTIs. pRTI supports two standards of HLA 1.3 and IEEE 1516. It can realize inter operation between different versions of RTI with different standards through linkers, supplies APIs with different types of Java and C++, and at the same time can support platforms like Windows, Redhat Linux, Sun, SGI and others. MAK-RTI of MAK Co. in USA also supports two standards of HLA 1.3 and IEEE 1516 and supplies APIs which can support such platforms as Windows, Linux, IRIX and other platforms.

Some domestic research institutes are also carrying out the related research and development work with respect to RTI. The representative systems include KD-RTI, YH-RTI and StarRTI based on CORBA of NUDT (National University of Defense Technology); SSS-RTI of CASIC (The Second Academy of China Aerospace), BH RTI[119] of the State Key Laboratory for VR Technology and Systems of Beijing University of Aeronautics & Astronautics University and AST-RTI of department of Automation Engineering. All of the RTIs have their own features and are used in different industry fields.

BH RTI conforms to both HLA 1.3 and IEEE 1516 standards. It can support real-time interaction of 5000 federates and nearly 50000 dynamic objects. BH RTI, released on http://www.hlarti.com, also developed some application systems. Table 3 gives a function and performance comparison, in which the scope indications come from existing documents and reports.

### 5.4 Other professional tools

During the VR development process, corresponding professional development standards and tools have been formed in some important fields. Those professional development tools have definite data sources and types, and are supported by professional knowledge, which, combined with 3D rendering tool and others, can be used to solve the practical problems.

In the field of machine design, many professional tools have been utilized frequently, such as virtual prototype analysis software ADAMS (automatic dynamic analysis of mechanical systems), CATIA (computer aided tri-dimensional interface application) in the field of CAD/CAE/CAM and so on. Interactive graphic environment, element library and constraint base are used to create complete parameterized geometric model for mechanical systems, carry out statics, kinematics and dynamics analysis for virtual mechanical systems, output curves of displacement, speed, acceleration and reaction force in ADAMS. ADAMS simulation can be utilized in performance forecast of mechanical

**Table 3** Comparison for functions and performances of RTIs

| RTI | DMSO RTI NG | MÄK RTI | pRTI | BH RTI |
|---|---|---|---|---|
| Development affiliation | (Product Forming of DMSO RTI)SAIC Co., USA | MÄK Co., USA | Pitch Co., Sweden | Beihang University |
| Supporting standard | HLA 1.3 | HLA 1.3; IEEE 1516 | HLA 1.3; IEEE 1516 | HLA 1.3; IEEE 1516 |
| Hardware filtering | multicast | none | none | multicast |
| Support leaguers scale | Approx. 150 (2003 JVB C4ISR Test, USA) | only less than 20 samples are seen in existing documents | no more than 100 in existing documents | 5000 |
| Support scale of dynamic objects | Approx. 20,000 (2003 JVB C4ISR Test, USA) | 200–300 (product white book) | no more than 100 in e xisting documents | nearly 50000 |

systems, moving range, collision detection, finite element calculation of input load and other fields; CATIA can create a work environment for product developing process which includes product concept design, detail design, engineering analysis, application and maintenance of product in entire life cycle. CATIA which started from aeronautics and astronautics possesses an important part in manufacturing industry. Boeing Co. accomplished virtual assembly for Boeing 777 by CATIA which has already become a typical successful application sample.

Virtual instruments possess a bit more complicated human-machine interface which is required in lots of applications like different instrument panel display in aviation cockpit, instrument panel display in automobile, display of communications equipment and other places. The special development tools for virtual instrument include GLStudio, Vaps, GMS, CST and Iocomp. Simulation software of radar can set the scanning, speed, angle, echo delay, interfering noise and visibility of detection for radar.

STK (satellite tool kit) is a software tool kit for satellite developed by Analytical Graphics, a US Company, which can support the entire process involving concept, requirement, design, manufacture, testing, operation, application and others. The core capabilities of STK are to create position and attitude data, to get date and time, and to carry out coverage analysis of remote sensor while the extended functions include additional orbit prediction algorithm, attitude definition and others.

Simulation of virtual human also has some professional SDKs (software development kit), of which the representatives are DI-GUY and Jack. DI-GUY collects the feature data of actual soldier training while possessing total texture model, varied kinds of uniform, weapon and other affiliated equipment for soldier's seven levels LoD details. The definable actions include standing, kneeling, crawling, walking, jumping, prowling, weapon using and other actions. JACK is an SDK for virtual human developed by University of Pennsylvania, USA, through which, we can put the virtual human in 3D scene model and carry out the simulation inspection by controlling its movements.

## 6 VR applications

### 6.1 Main application directions

VR science and technology field possesses combinational features of two aspects. One is formed by cross-link of multiple disciplines to make further innovation and development. The research on VR modeling, rendering, human-machine interaction and others needs to combine research achievements of multi-discipline, including mathematics, physics, electronics, control science, computer science, psychology and AI. The other refers to strong applicability which can be combined with the features and demands of applied fields closely. When VR is applied in different industries and fields, it needs to get combined with the features of those industries or fields, and the method and technology of linking application models with different fields should be explored to meet application demands. Applications will be the driving power for the development of VR technology.

Every scientific technology has its own background as well as a file for its applications and functions. The essential function of VR is to "replace actuality with virtuality" and "replace actual test with scientific calculation". Therefore, some experts think that scientific calculation with VR has already become the third method, following closely behind scientific research, and theory and experiment of engineering practice.

Generally speaking, applications of VR in different fields mainly focus on three aspects, namely training and drill, planning and design, as well as display and entertainment. The characteristic of training and drill system is to create modeling for actual world, and replace actual training environment with virtual environment formed. Operators can join in the virtual environment to take repetitive operation training and collaborative work, to have a sensation similar to that in actual environment. The planning and design system is to make lifelike emulation, forecast and estimation of inexistent object or nonoccurrent phenomenon in reality, thus making the planning and design more scientific and rational. The display and entertainment system is to simulate the actual or virtual objects to accomplish the purposes of appreciation and entertainment by disseminator and fulfil people's participation.

### 6.2 Typical fields and systems

The above mentioned three types of VR applications have successful typical systems in various fields, especially in military, medicine, industry, education and culture.

6.2.1 Military field. Military artificial training and drill is one of the most important application fields of VR technology, as well as the earliest and most frequently applied field of VR technology. US Department of Defense has listed VR technology as one of the seven core technologies that ensure the dominant position for US military in the 21st Century. The application of VR technology to military drill has brought the innovation into the concepts and styles of military drill, and improved the development of military artificial training and drill.

In 1983, US Military and the Defense Advanced Research Projects Agency of US Department of Defense (DARPA) established and implemented SIMNET program[120] together. SIMNET is actually the earliest distributed virtual environment. During the following twenty years, with the rapid development of core technologies required by constructing distributed virtual environment, such as computer technology and network technology, as well as constant improvement of military artificial training demand, distributed virtual battlefield environment used for military emulation training has changed a lot.

6.2.2 Medicine field. The field of medicine has a great deal of application requirements for VR technology; it provides strong traction force on the development of VR technology, and makes VR research face serious challenges. Due to a huge amount of geometric, physical, physiological and biochemical data of a human body, all kinds of tissues and organs are elastic-plastic and various interactive operations such as cutting, suturing and excision also require the change of topological structure of a human body. Therefore, it is rather difficult to establish real-time, immersive and interactive medical-use VR system. At present, VR technology has been primarily used for virtual operation training, remote consultation, operation planning and navigation, remote collaboration, etc. Some of these applications have already become the irreplaceable means and links during medical process[121].

In the aspect of virtual operation training, typical systems include Procedicus MIST System developed by Mentice Co. of Sweden, LapSim System developed by Surgical Science, and Select IT VEST System developed by Forschungszentrum

Karlsruhe of Germany, etc.[122]. In the aspect of remote consultation, Northern Carolina University of America has developed a suite of 3D remote medical consultation system which rebuilds a real-time and online actual environment with a few cameras. Combined with head positions and direction tracking, the system also provides the doctors with continuous dynamic remote images and stereo vision suitable for visual effects. This system overcomes the shortcomings that traditional 2D video system cannot obtain the required camera angle as well as poor lays of graphics. UPC of Spain has developed a remote cooperative virtual environment platform for medical purpose—ALICE[123], which uses the multithreading technology based on P2P topology that allows the users to communicate with each other on this platform, opening a window at client-side to observe the real-time image of another remote user.

In the aspect of operation planning and navigation, there are some virtual operation planning systems which could be used primarily at home and abroad, such as the operation planning system developed by Harvard University for the treatment of slippage femoral head and bone socket, operation system developed by Queen's University of Canada for the treatment of tibial osteotomy, virtual operation planning system co-operated by Tsinghua University and Chinese PLA General Hospital for the treatment of geneogenous infantile hip dislocation, and the robot-aided brain surgery planning and navigation system co-operated by the Robot Institute of Beijing University of Aeronautics and Astronautics and Neurosurgical Center of Navy General Hospital.

In the aspect of remote collaborative operation, the remote medical operation system developed by Stanford International Institute of America amplifies the surgical part using a VR system and let the doctors operate at the amplified range through routine operation. At the same time, VR system reduces the doctor's action to manipulator's remote action in real-time making the microsurgery easier. The vessel suture robot remotely controlled by Tokyo University of Japan and Medical Board of Okayama University accomplishes a remote operation of suturing 1 mm vessel on mouse. However, due to strict requirements for the equipment and the big risk during the robot remote operations, at present, the technology and system of remote collaborative operation is mainly used by experienced doctors to instruct the surgical medical staff to operate in a different place. Real operative process still needs medical staff to finish it on site.

6.2.3 Industry field. In industry field, VR technology is widely used for product demonstration, design, assembly, man-machine work efficiency, and performance evaluation. Many industrial sectors have paid attention to such typical applications as simulated training and virtual prototype technology. In the 1990s, Johnson Space Center in USA used VR technology for the purpose of maintenance and training of Hubble Space Telescope, and Boeing Co. used VR technology in assisting the pipeline design of Boeing 777. All these are typical successful examples.

DEPTH system[124], developed by Armstrong Laboratory of US air force uses visualization and VR technology for analyzing maintainability and indemnification to enable designers to know whether the maintenance task is available or not in their design, and thereby to find out hidden problems of indemnification before setting the design. This system has verified the availability of simulated maintenance action on Plane F-22, F-16 and B-1B, which has become the main supporting tool for virtual maintainability analysis of LOCKHEED MARTIN tactical aircraft system.

Virtual sample machine has been applied to the associated fighter plane JSF project by aeronautical department of LOCKHEED MARTIN and it has provided with full support of the design concept, scheme selection, performance test, processing and assembling, maintenance training and the demonstration of new products of the new machine type. The system supports collaborative design with networking technology. And simulation demonstrates the whole process of design, manufacture, operation, training and maintaining by Intranet and supply interaction and analysis tool of each layer to find out and solve the defects of concept design and manufacture. The statistical data shows that it can reduce 30% of development cost and 40% of development cycle by using virtual sample machine instead of physical prototypes.

Some European developed countries also pay great attention to VR technology and related research on the fields like industrialized application, etc. They funded and finished the project of VIEW (Virtual and Interactive Environments for the Workplaces of the Future) by the end of 2003[125]. VIEW has guided the application and development of virtual environment, supported the conception of movable, portable and immersive virtual environment and developed several portable displays and interaction prototype systems for assisted design link of VR.

6.2.4 Education and culture field. The field of education and culture is another important application field of VR technology. Today, VR has become the core supporting technology for application systems such as digital museum/science museum, simulation of rehearsal for the opening and closing ceremonies of large-scale activities, immersive interactive games, etc. For digital museum/science museum, VR technology is able to be used for digitalization and demonstration of various cultural relics such as literature, scripts, photos, records, films, collections, etc. Highly accurate modeling of relic exhibits also brings higher requirements of VR modeling methods and data acquisition equipment, which in turn promotes the development of VR. Many countries have fulfilled this work actively and many museums like New York Metropolitan Museum, British Museum, Russian Winter Palace Museum, and Louvre have established their own digital museums. China has also developed and established university digital museums, digital science museums, virtual Animation and virtual Imperial Palace as well[126].

VR technology could support immersive and verisimilitude simulating rehearsal for large-scale entertainment performance, opening and closing ceremonies of large-scale activities, thus making the flow and arrangement of the activities more reasonable and artistic effects much better. A typical example for the application is that VR technology was once used to support the creative design for the fireworks and simulation demonstration of visual effect for the opening and closing ceremonies of the 2004 Athens Olympic Games.

VR technology could provide real-time and verisimilitude 3D virtual scenes for games, support active and cooperative participation from users in network environment, and enhance naturalness of interaction for multi-users games by means of high precision interactive equipment. Therefore, VR has been more and more applied in games. Game products constantly pursue larger and larger scale of user access, more and more verisimilitude audiovisual effects and more and more natural man-machine interactions, which will continuously bring about new requirements for the research and development of technology, system and equipment of VR. For example, the virtual society—the Second Life, developed by Linden Lab, USA, has now owned 8 million registered users and 45 transnational corporations. There are virtual embassies and news agencies founded by some countries in the virtual society. The GDP of The Second Life is 0.7 billion in 2007. We can see that digital entertainment based on VR technology has grand influence on the future society.

## 7 Scientific problems for further research

Through years of development, great progress has been made in the research of VR. Some successful application examples have conceivably indicated the unique and irreplaceable position and function of VR technology. However, both the present status and developing trend analysis of VR on the one hand and the research and practice in this field on the other hand show that there are still a series of fundamental theoretical problems and key technical problems yet to be solved in this field. Any breakthrough of these problems will greatly promote VR technology and improve its practicality.

**Question 1. Modeling, model complexity and model dependability**

Can all things in the realistic world be modeled or even digitally modeled? This question is related with computability and model complexity as well as the expecting degree of models approaching the reality. For example, the absolutely real sea wave is difficult to model; however, a simplified conceptual sea wave model can be established; a meteorological model which fully fits the reality is difficult to model; however, a simplified meteorological model can be established. The question lies in to what extent the simplified model can be used and depended. Therefore, the evaluation and measurement of modeling, model complex, model simplification, and model dependability, and the way to balance the modeling and expectancy to modeling, etc. are important theoretical questions.

**Question 2. The similarity measure of image and image quality evaluation**
Image similarity is a typical example of model similarity issues, a basic issue in image identification and image retrieval, and an important issue in VR as well. It is not difficult for a person to compare the similarity degree of two images or evaluate the deformation degree of an image; however, it remains an outstanding issue for computer so far. One of the reasons is that it is difficult to find a general and effective quantitative measure of image similarity. Additionally, it is still an outstanding issue how to find a kind of objective measure method for evaluating image quality (similarity degree with realistic image).

**Question 3. Complex and realistic virtual environment structure and mass data management**
Presently, construction of virtual environment is carried out through mathematical modeling and reality duplication. Both of them possess their own advantages and inherent limits. Generally speaking, the 3D character of the former is superior to that of the latter, while the latter has higher fidelity than the former. The more applications the VR finds, the higher the requirements for the complexity and verisimilitude of virtual environment. This poses a key problem to VR modeling. The way to solve this problem depends on the further improvement of performance of computer and special processing part. We should develop duplication equipment with high precision and efficiency or even intelligentization; and we should develop complete modeling tools with high automaticity; meanwhile, we should also solve this question by seeking for a virtual environment structure combined with the above two ways and embody the advantages of them.

Complex and realistic virtual environment possess mass data of concentrated or distributed memory, and the effective dispatch, retrieval and maintenance of these mass data constitute a challenging problem.

**Question 4. Physical and action modeling of virtual entity**
The modeling of virtual entities includes the modeling of geometry, modeling of physical properties and modeling of actions. As the modeling of geometry started rather early, the theories and tools are mature and the difficulties are relatively few. Presently, modeling of physical properties mainly focuses on the kinematics and dynamics characters of virtual entity. To further improve the fidelity, the material, friction, elastoplasticity and slumpability, physical properties of virtual entities, etc. must be considered. Computational model, high performance algorithm, and effective performance method are three big problems to be researched and solved. Physical property modeling is a permanent direction in VR.

In modeling of actions, computer generated entity and the integrated entity formed therefrom are important research contents. Integrated entity gives prominence to the general characters of group entity and group function, thus reducing the redundancy and complexity in modeling virtual entities. The core problem in modeling of virtual actions lies in artificial intelligence (AI). Presently, the research of AI has made no breakthrough as yet. But if the existing research achievements of AI can be applied to modeling of virtual actions creatively, the intelligence level of modeling virtual entity action will undoubtedly be raised, because low is the degree of AI research achievement introduced and used for reference in modeling of virtual entities.

**Question 5. Confluences of virtual scene and realistic scene**
The confluence of virtual scene and realistic scene refers to virtual scene inosculated in dynamic and static realistic scenery frame and dynamic and

static realistic scene inosculated in virtual environment, involving many theoretical problems such as how to realize data confluence and model confluence, and many key technique problems such as how to make correct 3D registration of virtual & realistic scenes, space shielding relation of virtual & realistic scenes, confluence processing of virtual & realistic lighting effect, etc., as well as the development of a series of special tools and supporting platforms.

**Question 6. Man-machine interaction mechanism and immergence sense**

The similarity of user's operation, information release, information acquisition and transmission, etc. for the whole experience process in virtual environment to those in realistic environment, i.e., the immergence sense, is closely related with man-machine interaction mechanism. Man-machine interaction mechanism is involved in man-machine interaction mode, man-machine interaction equipment, man-machine interaction information processing, etc. The naturalization and intelligentization of man-machine interaction is the goal to be pursued in this field. For the future development, firstly, we should improve the validity and accuracy of the existing man-machine equipment; secondly, we should enhance the interaction mobility of the user to make man-machine interaction better fit to men, e.g., with the help of wearable computer, personal digital assistant, and virtual and realistic confluence environment interaction; thirdly, we should explore man-machine interaction equipment based on new concepts, e.g., stereo display based on hologram information, etc.; fourthly, we should improve the intelligence level of man-machine interaction, e.g., machine identification & understanding, etc. A breakthrough in these questions will further realize the naturalization and intelligentization of man-machine interaction.

The immergence sense of man-machine interaction keeps close relation with the verisimilitude and real-time of virtual environment, as with the case where the algorithmic time keeps certain duality with spatial efficiency, or in other words, the arithmetic product of both measures has a constant value to some level. To be consistent to the actual demand, balance is needed.

**Question 7. Real-time and consistency of distributed virtual reality**

In view of computer network facilities, distributed VRs are divided into two categories: high speed based special network and Internet based network. The former is generally used for the large scale special virtual reality like military simulation training, drill, etc., while the latter is suitable for personal stochastic use of virtual reality, e.g., distance education, entertainment, etc. The problems faced by them are different. However, they both face the problem of real-time of distributed VR, time consistency, event consistency, space consistency in distributed virtual environment. The root cause of time and event consistency lies in the randomicity of time delay caused by the intrinsic and length of internet. There is still no specific answer to whether the problem can be thoroughly solved as yet. To this end, different solutions have been presented, but we still need to do more researches, experiments and analyses.

**Question 8. Assembled structure of VR application system and development policy of VR platform**

The construction of VR application system through assembling modularized reusable parts is a way to improve system development efficiency and system dependability. Hence, problems concerning assembled and reusable software theory, and reusable parts, assembled system structures, standard interfaces, etc. need to be further researched.

Making as much as possible abstraction of public services needed for VR system development and application in different fields and the commonness part of different VR systems and establishing public supporting service platforms and tools are effective approaches to improving the reusability of distributed VR application system and development efficiency. As these platforms and tools are involved in all stages and full life cycle of VR system development and application, and different types of system users with vast categories, it is necessary to choose and determine adequate development policy, method and frame, and to establish large scale development supporting platforms with independent intellectual property, such as graphic platform and RTI.

**Question 9. VR system performances and evaluation on VR application result** Regarding various performances of VR system, like verisimilitude, interactivity, and real-time, particularly the dependability of VR system application, etc., the establishment of reasonable evaluation indexes and the enaction of appropriate evaluation standard are important premises for facilitating the advancement, application, and generalization of virtual reality technology. Presently, as researches on this aspect are relatively scarce and there are many difficulties per se, it is difficult to describe most of the performances quantitatively. To meet the demand of VR scientific and technological development and application, we should lay emphasis on the development of specific VR technology to facilitate the research in this direction.

While developing VR application systems, it is also necessary to explore methods for evaluating the results of using various VR system applications and fabricate evaluation tools for supporting various application indexes. Reasonable application indexes and result evaluation measuring methods will make VR application systems meet practical demands, and predict and control the development processes of system timely in the early stage of system development.

## 8 Conclusions

During the past ten years, VR has made significant progress in theoretical research, technical innovation, system development, application and promotion, and has shown huge potential and broad application prospect. China has entered a new development stage in this scientific and technology field, and VR technology has become a national goal of development of information science and technology. Meanwhile, still a great number of pending problems and difficulties exist in VR research field. Furthermore, new theories and key technical problems continuously occur, the breakthrough of which will lead to further development of VR technology. The next 10 years will be the time to solve and to make breakthrough in these problems, as well as the time to promote and apply VR and bring huge economic and social benefits in China. VR is a scientific and technological field full of energy and challenge. We wish more and more computer science workers join in the VR research and development team and make contributions to enable China to take a leading role in this important field in the world.


1 Heilig M L. Sensorama simulator. US Patent 3 050 870, 1962-08-28
2 Sutherland I E. The ultimate display. In: Proceedings of the International Federation of Information Processing (IFIP) Congress, 1965. 506–508
3 Fisher S S, Humphries J, McGreevy M, et al. The virtual environment display system. In: ACM Workshop on Interactive 3D Graphics. New York: ACM Press, 1986. 77–87
4 Foley J D. Interfaces for advanced computing (cover story). Sci Am Mag, October 1987
5 Michael H. Metaphysics of Virtual Reality. Oxford: Oxford University Press, 1993
6 Grigore B, Phillippe C. Virtual Reality Technology. New York: John Wiley and Sons, 1994
7 Wang C W, Gao W, Wang X R. Theory, Realization and Application of Virtual Reality (VR) Technology (in Chinese). Beijing: Tsinghua University Press, 1996
8 Zhao Q P. DVENET Distributed Virtual Environment (in Chinese). Beijing: Science Press, 2002
9 Huang K D, Liu B H, Huang J, et al. Overview of battle simulation technology (in Chinese). J Syst Simul, 2004, 16(9): 1887–1895
10 Dong S H. Man-machine Interaction (in Chinese). Beijing: Beijing University Press, 2004
11 Pan Y H, Fan J S. Reality and Virtuality of Tunhuang (in Chinese). Hangzhou: Zhejiang Univeristy Press, 2003
12 Li B H, Chai X D. Preliminary research and development of complex product virtual sample machine supporting platform (in Chinese). Comput Simul, 2003, 20(1): 4–8
13 Li D R. Function of digital province and city in state land planning and town construction (in Chinese). J Survey Mapp, 2002, 31(Supplement): 16–21
14 Sun J G. Computer Graphics (in Chinese). Beijing: Tsinghua University Press, 2002
15 Dai G Z. Wearable interaction computation (in Chinese). High-tech Commun, 2001, 7: 51–55
16 Peng Q S. Arithmetic Basis of Computer Photo-realism Graphic (in Chinese). Beijing: Science Press, 1999
17 Chen C S, Hung Y P, Chung J B. A fast automatic method for registration of partially-overlapping range images. In: Proceedings of IEEE ICCV1998, 1998. 242–248
18 Huang Q X, Simon F, Gelfaln N, et al. Reassembling fractured objects by geometric matching. ACM Trans Graph, 2006, 25: 569–578

19 Huber D. Automatic three-dimensional modeling from reality. PhD thesis. Pittsburgh: Carnegie Mellon University, 2002
20 Amenta N, Bern M, Kamvysselis M. A new voronoi-based surface reconstruction algorithm. In: Cohen M, ed. Proceedings of SIGGRAPH'98. New York: ACM Press, 1998. 415–421
21 Liepa P. Filling holes in meshes. In: Proceedings of the 2003 Eurographics/ACM SIGGRAPH symposium on Geometry processing. P. O. Box 16 Aire-la-Ville Switzerland: Eurographics Association, 2003. 200–205
22 Davis J, Marschner S R, Garr M, et al. Filling holes in complex surfaces using volumetric diffusion. In: First International Symposium on 3D Data Processing, Visualization, and Transmission. 2002. 19–21
23 Ju T. Robust repair of polygonal models. ACM Trans Graph, 2004, 23(3): 888–895
24 Pauly M, Mitra N J, Giesen J, et al. Example-based 3D scan completion. In: Symposium on Geometry Processing. 2005. 23–32
25 Sharf A, Alexa M. Context-based surface completion. ACM Trans Graph, 2004, 23(3): 878–887
26 Mueller P, Zeng G, Wonka P, et al. Image-based procedural modeling of facades. ACM Trans Graph, 2007, 26(3): 85
27 Hartley R I, Zisserman A. Multiple View Geometry. Cambridge, UK: Cambridge University Press, 2004
28 Scharstein D, Szeliski R. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. Int J Comput Vision, 2002, 47(1): 7–42
29 Wei Y, Ofek E, Quan L, et al. Modeling hair from multiple views. In: Proceedings of ACM SIGGRAPH, 2005. 816–820
30 Tan P, Zeng G, Wang J D, et al. Image-based tree modeling. Proc ACM SIGGRAPH, 2007, 26(87): 1–7
31 Matusik W, Pfister H, Ngan A, et al. Image-based 3D photography using opacity hulls. In: Proceedings of ACM SIGGRAPH, San Antonio, Texas, 2002. 427–437
32 Nicodemus F E, Richmond J C, Hsia J J, et al. Geometrical Considerations and Nomenclature for Reflectance. Washington: National Bureau of Standards, 1997. 1–67
33 Ngan A, Durand F, Amatusik W. Experimental analysis of BRDF models. In: Eurographics Symposium on Rendering, 2005. 117–226
34 Jason L, Aner B -A, Christopher D, et al. Inverse shade trees for non-parametric material representation and editing. ACM Trans Graph, 2006, 25: 735–745
35 Dana K J, Ginneken B V, Nayar S K, et al. Reflectance and texture of real-world surfaces. ACM Trans Graph, 1999, 18: 1–34
36 Gu J W, Chien I T, Ravi R, et al. Time-varying surface appearance: acquisition, modeling and rendering. ACM Trans Graph, 2006, 25: 762–771
37 Kalyan S, Ravi R, Peter N B, et al. Time-varying BRDFs. IEEE Trans Visual Comput Graph, 2007, 13: 595–609
38 Jaffery S M, Dutta K. Digital reconstruction methods for 3D image visulization. In: SPIE Processings and Display of Three-Dimensional Data II, 1984. 155–159
39 Helgeland A, Andreassen O. Visualization of vector fields using seed LIC and volume rendering. IEEE Trans Visual Comput Graph, 2004, 10(6): 673–682
40 Taponecco F, Alexa M. Vector field visualization using markov random field texture synthesis. In: Proceedings of the Symposium on Data Visualisation. Switzerland: Eurographics Association, 2003. 195–202
41 Kaufman D M, Edmunds T, Pai D K. Fast frictional dynamics for rigid bodies. ACM Trans Graph, 2005, 24(3): 946–956
42 Wong W S -K, Baciu G. A randomized marking scheme for continuous collision detection in simulation of deformable surfaces. In: Proceedings of the 2006 ACM International Conference on Virtual Reality Continuum and Its Applications. New York: ACM Press, 2006. 181–188
43 Govindaraju N K, Kabul I, Lin M C, et al. Fast continuous collision detection among deformable models using graphics processors. Comput Graph, 2007, 31(1): 5–14
44 Sorkine O. Laplacian mesh processing. State of the Art Reports. 2005
45 Botsch M, Sumner R, Pauly M, et al. Deformation transfer for detail-preserving surface editing. In: Proceedings of Vision, Modeling, and Visualization (VMV '06). Ios Pr Inc, 2006. 357–364
46 Thomaszewski B, Wacker M, Straßer W. A consistent bending model for cloth simulation with corotational subdivision finite elements. In: Proceedings of the 2006 ACM SIGGRAPH/Eurographics Symposium on Computer Animation. Switzerland: Eurographics Association, 2006. 107–116
47 Müller M, Heidelberger B, Hennix M, et al. Position based dynamics. J Visual Commun Image Represent, 2007, 18(2): 109–118
48 Shi L, Yu Y, Bell N, et al. A fast multigrid algorithm for mesh deformation. ACM Trans Graph, 2006, 25(3): 1108–1117
49 Huang J, Shi X, Liu X, et al. Subspace gradient domain mesh deformation. ACM Trans Graph, 2006, 25(3): 1126–1134
50 Botsch M, Sorkine O. On linear variational surface deformation methods. IEEE Trans Visual Comput Graph, 2008, 14(1): 213–230
51 Eitz M, Gu L X. Hierarchical spatial hashing for real-time collision detection. In: IEEE International Conference on Shape Modeling and Applications (SMI'07) . Los Alamitos: IEEE Computer Society Press, 2007. 61–70
52 Witkin A, Zoran P. Motion wraping. In: ACM SIGGRAPH. Addison Wesley, 1995. 105–108
53 Park S I, Shin H J, Shin S Y. On-line locomotion generation based on motion blending. In: Proceedings of ACM SIGGRAPH Symposium on Computer Animation. New York: ACM, 2002. 105–111
54 Girard M. Constrained Optimization of Articulated Animal Movement in Computer Animation. Making Them Move: Mechanics, Control, and Animation of Articulated Figures. Badler B, Zelter, eds. San Francisco: Morgan Kaufmann Publishers, 1990. 209–232
55 Pew R W, Mavor A S. Modeling Human and Organizational Behavior: Application to Military Simulations. Washington, DC: National Academy Press, 1998

56 Karr C R, Holbrook R. Modeling Command and Control in WARSIM 2000. In: Proceedings of the 8th Conference on Computer Generated Forces and Behavioral Representation, 1999
57 Commercial Platform Training Aids. Los Angeles: Institute for Creative Technologies, University of Southern California. http://www.ict.usc.edu/content/view/49/103/
58 Hoff B R. USMC individual combatants. In: Proceedings of the Defense Modeling and Simulation Office Individual Combatant Workshop, VA: Defense Modeling and Simulation Office, Alexandria, July 1-2, 1996
59 Cora B E -T, Nicholas R J. Using reinforcement learning to coordinate better. Comput Intell, 2005, 21(3): 217–245
60 Kazakov D, Bartlett M. Cooperation navigation and the faculty of language. Appl Art Intell, 2004, 18(9-10): 885–901
61 Soar/IFOR: intelligent agents for air simulation and control. In: Nielsen P E, ed. Simulation Conference Proceedings, 3-6 Dec. 1995. 620–625
62 Regenbrecht H, Wagner M. Interaction in a collaborative augmented reality environment. In: CHI '02 Extended Abstracts on Human Factors in Computing Systems. New York: ACM, 2002. 504–505
63 Allard J, Ménier C, Raffin B, et al. Grimage: markerless 3D interactions. In: ACM SIGGRAPH 2007 Emerging Technologies. New York: ACM, 2007
64 Ercan A O, Yang D B, Gamal A E, et al. Optimal placement and selection of camera network nodes for target localization. In: Proceedings of Second IEEE International Conference on Distributed Computing in Sensor Systems (DCOSS). Berlin: Springer, 2006. 389–404
65 Fusiello A, Murino V. Augmented scene modeling and visualization by optical and acoustic sensor integration. IEEE Trans Visual Comput Graph, 2004, 10(6): 625–636
66 Billinghurst M, Kato H, Poupyrev I. MagicBook: transitioning between reality and virtuality. In: CHI '01 Extended Abstracts on Human Factors in Computing Systems. New York: ACM, 2001. 25–26
67 Didier J Y, Roussel D, Mallem M, et al. A texture based time delay compensation method for augmented reality. In: 3rd IEEE and ACM International Symposium on Mixed and Augmented. 2004. 262–263
68 Chen X W, Milgram P. Integration of pointed-based interposition with binocular disparity alignment in stereoscopic augmented reality environments. In: The Conference on IRIS, 2002
69 Huang Y, Essa I. Tracking multiple objects through occlusions. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Washington, DC: IEEE Computer Society, 2005. 1051–1058
70 Schmidt J, Niemann H, Vogt S. Dense disparity maps in realtime with an application to augmented reality. In: Sixth IEEE Workshop on Applications of Computer Vision. Washington, DC: IEEE Computer Society, 2002. 225–230
71 Pilet J, Geiger A, Pascal L, et al. An all-in-one solution to geometric and photometric calibration. In: IEEE/ACM International Symposium on Mixed and Augmented Reality. Washington, DC: IEEE Computer Society, 2006. 69–78
72 Jacobs K, Nahmias J D, Angus C, et al. Automatic generation of consistent shadows for augmented reality. In: Proceedings of Graphics Interface 2005. Ontario, Canada: Human-Computer Communications Society, 2005. 113–120
73 Kanbara M, Yokoya N. Real-time estimation of light source environment for photorealistic augmented reality. In: Proceedings of the 17th International Conference on Pattern Recognition. Washington, DC: IEEE Computer Society, 2004. 911–914
74 Hoffman N, Mitchell K. Real-time photorealistic terrain lighting. Game Develop, 2001, 8(7): 32–41
75 Sloan P -P, Kautz J, Snyder J. Pre-computed radiance transfer for real-time rendering in dynamic, low-frequency lighting environments. In: SIGGRAPH '02: Proceedings of the 29th Annual Conference on Computer Graphics and Interactive Techniques. New York: ACM Press, 2002. 527–536
76 Ng R, Ramamoorthi R, Hanrahan P. All-frequency shadows using non-linear wavelet lighting approximation. ACM Trans Graph, 2003, 22(3): 376–381
77 Zhou K, Hu Y, Lin S, et al. Pre-computed shadow fields for dynamic scenes. ACM Trans Graph, 2005, 24(3): 1196–1201
78 Sloan P -P, Luna B, Snyder J. Local, deformable pre-computed radiance transfer. ACM Trans Graph, 2005, 24(3): 1216–1224
79 Ren Z, Wang R, Snyder J, et al. Real-time soft shadows in dynamic scenes using spherical harmonic exponentiation. ACM Trans Graph, 2006, 25(3): 977–986
80 Blinn J F. Simulation of wrinkled surfaces. In: Proceedings SIGGRAPH '78, 1978. 286–292
81 Cook R L. Shade trees. In: Proceedings of SIGGRAPH '84, Computer Graphics. 1984, 18: 223–231
82 Kaneko T, Takahei T, Inami M, et al. Detailed shape representation with parallax mapping. In: Proceedings of the ICAT 2001, 2001. 205–208
83 Oliveira M M, Bishop G, McAllister D. Relief texture mapping. In: Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques, 2000. 359–368
84 Reeves W T, Blau R. Approximate and probabilistic algorithms for shadingand rending structured particle system. In: Proceedings of SIGGRAPH85, Computer Graphics. 1985, 19(3): 313–322
85 Yuksel C, House D H, Keyser J. Water particles. In: Proceedings SIGGRAPH. New York: ACM Press, 2007. 99–106
86 Adams B, Pauly M, Keiser R, et al. Adaptively sampled particle fluids. In: ACM SIGGRAPH, 2007
87 Moon J T, Marschner S R. Simulating multiple scattering in hair using a photon mapping approach. In: International Conference on Computer Graphics and Interactive Techniques, ACM SIGGRAPH'06, 2006. 1067–1074
88 Quan L, Tan P, Zeng G, et al. Image-based plant modeling. ACM Trans Graph (TOG) and Proc. of SIGGRAPH 2006, 2006, 25(3): 599–604
89 Lindstrom P, Silva C. A memory insensitive technique for large model simplification. In: IEEE Visualization '01, 2001

90 Hoppe H. View-dependent refinement of progressive meshes. In: Proc. SIGGRAPH'97, Los Angeles, CA, 1997. 189–198
91 Hoppe H. Progressive meshes. In: Proceedings of the SIGGRAPH'96. 1996, 99–108
92 Molnar S, Cox M, Ellsworth D, et al. A sorting classification of parallel rendering. IEEE Comp Graph Appl, 1994, 23–31
93 Adelson E H, Bergen J R. The plenoptic function and the elements of early vision. Computational Models of Visual Processing. Canbridge: The MIT Press, 1991
94 Shum H -Y, He L -W. Rendering with concentric mosaics. In: SIGGRAPH'99 Computer Graphics, 1999. 299–306
95 Hayashi K, Kato H, Nishida S. Occlusion detection of real objects using contour based stereo matching. In: Proceedings of the 2005 International Conference on Augmented Teleexistence, 2005. 180–186
96 Pilet J, Lepetit V, Fua P. Retexturing in the presence of complex illumination and occlusions. In: International Symposium on Mixed and Augmented Reality, Nara, Japan, November 2007
97 Bradley D, Roth G, Bose P. Augmented reality on cloth with realistic illumination. Mach Vision Appl J, 2007
98 Jung Y, Franke T, Dähne P, et al. Enhancing X3D for advanced MR appliances. In: ACM Special Interest Group on Computer Graphics and Interactive Techniques, Perugia, Italy, 2007. 27–36
99 Xie B S, Zhong X L, Rao D, et al. Head-related transfer function database and its analyses. Sci China Ser G-Phys Mech Astron, 2007, 50(3): 267–280
100 Zotkin D N, Duraiswami R, Davis L S. Victual audio system customization using visual matching of ear parameters. In: Proceedings of IEEE 16th International Conference on Pattern Recognition, 2002
101 Ionue N. HRTF modeling using physical features. In: Proceedings of Forum Acusticum. Budapest, 2005
102 Brown C P, Duda R O. A structural model for binaural sound synthesis. IEEE Trans Speech Audio Process, 1998, 6(5): 476–488
103 Middlebrooks J C. Individual difference in external ear transfer functions reduced by scaling in frequenty. J Acoust Soc Am, 1999, 106(3): 1480–1492
104 Basdogan C, Srinivasan M A. Haptic rendering in virtual environments. In: Stanney K, ed. Virtual Environments Hand Book. NJ: Lawrence Erlbaum Inc, 2001
105 Otaduy M A, Lin Ming C. Introduction to haptic rendering. In: International Conference on Computer Graphics and Interactive Techniques, ACM SIGGRAPH 2005, 2005. A3–A33
106 Wang Y T. Strengthen realistic research progress. In: The 5th China Computer Graphics Conference, 2005
107 Kruger W, Bohn C A, Frohlich B, et al. The responsive workbench: a virtual work environment. Computer, 1995, 28(7): 42–48
108 Neira C C, Danie J S, Thomas A D, et al. The CAVE: audio visual experience automatic virtual environment. Commun ACM, 1992, 35(6): 64–72
109 Rohs M. Marker-based embodied interaction for handheld augmented reality games. J Virtual Reality Broadcast, 2007, 4(5): 793–805
110 Jones A, Ian M, Yamada H, et al. Rendering for an interactive 360 deg; light field display. In: ACM SIGGRAPH 2007 papers. New York: ACM, 2007
111 Chen X W, Milgram P. Integration of pointed-based interposition with binocular disparity alignment in stereoscopic augmented reality environments. In: The Conference on IRIS, 2002
112 Templeman J N, Denbrook P S, Sibert L E. Virtual locomotion: walking in place through virtual environments. Presence: Teleoperat Virtual Environ, 1999, 8(6): 598–617
113 ISO/IEC 14772-1:1997 and ISO/IEC 14772-2:2004 Virtual Reality Modeling Language (VRML)
114 X3D architecture and Base Components Edition 2, ISO/IEC FDIS 19775-1.2:2008, Final Draft International Standard, Dec 2007
115 Universal 3D Specification, http://www.3dif.org/. 3D Industry Forum (3DIF)
116 Zhao Q P, Hao A M, Wang L L, et al. Real-time 3D graphic platform BH_Graph, Comput Research Develop, 2006, 43(9): 1491–1497
117 IEEE Std 1278–1993: IEEE Standard for Information Technology–Protocols for Distributed Interactive Simulation Applications, Entity Information and Interaction, 1993
118 Simulation Interoperability Standards Committee (SISC) of the IEEE Computer Society. IEEE Standard for Modeling and Simulation (M&S) High Level Architecture (HLA) IEEE std 1516-2000, 1516.1-2000, 1516.2-2000. New York: The Institute of Electrical and Electronics Engineers Inc., 2000
119 Zhou Z, Zhao Q P. Reducing time cost of distributed run-time infrastructure. In: Proceedings of the 16th Internal Conference on Artificial Reality and Telexistence, Hangzhou, Nov 29-Dec 02, 2006, Springer LNCS, 2006, 4282: 969–979
120 Arthur R P, Rechard L S. The SIMNET Network and Protocols. BBN Systems and Technologies Corporation, Report No. 7627, 1991
121 Preim B, Bartz D. Visualization in Medicine: Theory, Algorithms and Applications. San Francisco: Morgan Kaufmann Publishers Inc., 2007
122 Seymour N E, Gallagher A G, Roman S A, et al. Virtual reality surgical laparoscopic simulators: how to choose. Surgic Endosc, 2003, 17: 1943–1950
123 Andujar C, Fairen M, Brunet P. Affordable projection system for 3D interaction. In: 1st Ibero-American Symposium in Computer Graphics. University of Minho, Portugal, July 2002
124 Glor P J, Boyle E S. Design evaluation for personnel, training and human factors (DEPTH). In: Reliability and Maintainability Symposium, Proceedings Annual, 1993. 18–25
125 VIEW. http://www.view.iao.fraunhofer.de/
126 Zhao Q P, Shen X K, Qi Y. Some key technology research on digital museum. J Syst Simulat, 2007, 19(Supp. 2): 1–6