Chapter 1
Haptics: General Principles

1.1 Introduction

Our senses are physiological tools for perceiving environmental information. Humans have at least five senses (as defined and classified by Aristotle). These senses are: sight or vision, hearing or audition, smell or olfaction, touch or taction, and taste or gustation. They are perceived when sensory neurons react to stimuli and send messages to the central nervous system. We actually have more than five senses. For example, Gibson has stated that we have both outward-orientated (exteroceptive) senses and inward-orientated (interoceptive) senses [127]. The sense of equilibrium, also known as proprioception, is one example of these other senses. Each of the sense modalities is characterized by many factors, such as the types of received and accepted data, the sensitivity to the data in terms of temporal and spatial resolutions, the information processing rate or bandwidth, and the capability of the receptors to adapt to the received data.

1.2 Human Senses

Typically, it is believed that vision and audition convey the most information about an environment while the other senses are more subtle. Because of this, their characteristics have been widely investigated over the last few decades by scientists and engineers, which has led to the development of reliable multimedia systems and environments.

1.2.1 Vision

The visual sense is based on the level of absorption of light energy by the eye and the conversion of this energy into neural messages. The acceptable wavelength range for

human eyes is between 0.3 and 0.7 μm (1 μm $= 10^{-6}$ m). The temporal resolution sensitivity of the human visual system is biologically limited and not sufficient to detect the presentation of sequential video frames past a certain speed. This is the reason why we do not perceive a digital movie as a series of still images, but rather as moving pictures. Similarly, our spatial resolution is limited and does not allow us to resolve individual pixels. The spatial resolution is determined by the density and type of photoreceptors in the retina. Several factors limit the retina's functionality, such as the size of the pupil, the stimulated area of the retina, the eye movement, the background light, and the exposure time of the target.

1.2.2   Audition

The human auditory system transmits sound waves through the outer, middle, and inner ears. This sound wave is transformed into neural energy in the inner ear. It is then transmitted to the auditory cortex for processing. The audible frequency of humans ranges from 16 to 20,000 Hz and is most efficient between 1,000 and 4,000 Hz. A sound can also be described in terms of the sound wave's direction (or relative position of the emitter to the receiver since each ear has a nonuniform directional sensitivity), frequency, intensity, or loudness (which ranges from 0 to 160 dB), and duration.

1.2.3   Touch

Indeed, the sense of touch is distributed over the entire body, unlike the other conventional four senses, which are centralized around specific parts of the body. The sense of touch is mainly associated with active tactile senses such as our hands. Such senses can be categorized in several ways, and they have a link to the kinesthetic senses. Humans are very sensitive to touch, but different parts of our body have different sensitivities. These sensitivities vary because the skin is an interface that centrally discriminates four modalities of sensation, namely touch (including both light touch and pressure), cold, heat, and pain. Furthermore, a combination of two or more modalities can be used to characterize sensations such as roughness, wetness, and vibration. A human would not be able to sense and respond to the physical environment without these tactile receptors located over the entire body. To appreciate the sense of touch more fully, consider the following facts: according to Heller and Schiff [155], touch is twenty times faster than vision, so humans are able to differentiate between two stimuli just 5 ms apart; Bolanowski et al. [44] found that touch is highly sensitive to vibration up to 1 KHz, with the peak sensitivity around 250 Hz; and skin receptors on the human palm can sense displacements as low as 0.2 μm in length [197].

1.2.4 What Does the Sense of Touch Do for Us?

Robles de la Torre [309] states that losing the sense of touch has catastrophic effects such as impairment of hand dexterity, loss of limb position perception, and the inability to walk, just to name a few. Every day we use human–computer interfaces to interact, communicate, or perform various tasks, e.g., sending e-mails, downloading a video, controlling a process in an industrial plant. It seems that audio and visual feedback is dominant for these types of interactions; however, there is considerable importance in developing and applying sophisticated touch-enabled interfaces to perform similar tasks or improve the performance of existing tasks. Therefore, the following question may arise: what level of realism can be achieved upon enabling touch interactions with virtual environments? To answer this question, the haptic modality must be more fully explored [310].

1.3 Haptics Exploration

Haptics, a term that was derived from the Greek word "haptesthai" meaning "of or relating to the sense of touch," refers to the science of manual sensing (exploration for information extraction) and manipulation (for modifying the environment) through touch. It has also been described as "the sensibility of the individual to the world adjacent to his body by the use of his body" [127]. This word was introduced at the beginning of the twentieth century by researchers in the field of experimental psychology to refer to the active touch of real objects by humans. In the late 1980s, the term was redefined to enlarge its scope to include all aspects of machine touch and human–machine touch interaction. The 'touching' of objects could be done by humans, machines, or a combination of both, and the environment can be real, virtual, or a combination of both. Also, the interaction may or may not be accompanied by other sensory modalities such as vision or audition. Currently, the term has brought together many different disciplines, including biomechanics, psychology, neurophysiology, engineering, and computer science, that use this term to refer to the study of human touch and force feedback with the external environment.

Touch is a unique human sensory modality in contrast with other modalities. As previously mentioned, it enables bidirectional flow of energy due to the sensing and acting activities performed, as well as an exchange of information between the real or virtual environment and the end user (see Fig. 1.1). This is referred to as active touch. For instance, to sense the shape of a cup, one must run his/her fingers across its shape and surfaces to build a mental image of the cup. Furthermore, in a manipulation task, for instance sewing with a needle, the division between "input" and "output" is often very sharp and difficult to define. This co-dependence between sensing and manipulating is at the heart of understanding how humans can so deftly interact with the physical world.
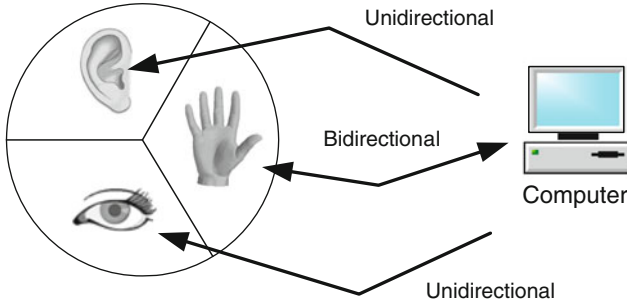
Fig. 1.1 A distinguishing feature of haptics is the bidirectional flow of information
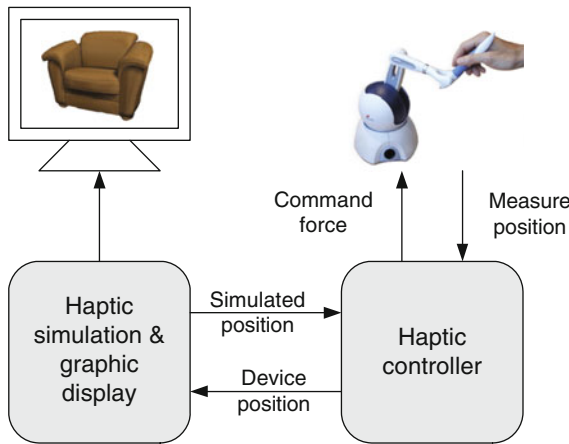
Fig. 1.2 Force representation in a virtual world

The initial sense of contact when one's hand interacts with an object is provided by the touch receptors (nerves endings) in the skin. The receptors provide information on the geometry, texture, slippage, etc. of the object surface. This information is tactile or cutaneous. When the hand applies force, trying to hold this object, kinesthetic information (force feedback) comes into play by providing physical information about the position and motion of the hand relative to the object (see Fig. 1.2).

From Fig. 1.2, one can see how we can make objects that populate the virtual environment touchable. The basic principle behind haptic interaction is simple. When the human user manipulates the generic probe (sometimes referred to as end-effector) of the haptic device, the position sensors of the device convey its tip position to the computer. At every time interval – say every 1 ms – the computer that controls the device checks for collisions between the simulated stylus and the virtual objects populating the virtual environment. If a collision has occurred, the haptic rendering system calculates the reaction forces/torques that must be

applied at the human–device interaction point and controls the actuator (a computer controlled electric DC motor) attached to the device, leading to a tactual perception of the virtual objects. In the case that no collision is detected, no forces will be computed/applied, and the user is free to move the stylus as if exploring empty space. In the simplest case, the magnitudes of the reaction forces are assumed proportional to the depth of indentation, and the forces are applied immediately following surface penetration.

1.4   Concepts and Terminology

We rely on our sense of touch to do everyday tasks such as dialing a touch-tone phone, finding first gear in a manual transmission car, or playing a musical instrument. We rely heavily on the tactile and kinesthetic cues we receive from the environment. Tactile cues include textures, vibrations, and bumps, while kinesthetic cues include weight, impact, etc. In the following section, we present some crucial concepts and terminology related to haptics:

Haptic: the science of applying tactile, kinesthetic, or both sensations to human–computer interactions. It refers to the ability of sensing and/or manipulating objects in a natural or synthetic environment using a haptic interface.

Cutaneous: relating to or involving the skin. It includes sensations of pressure, temperature, and pain.

Tactile: pertaining to the cutaneous sense, but more specifically the sensation of pressure rather than temperature or pain.

Kinesthetic: relating to the feeling of motion. It is related to sensations originating in muscles, tendons, and joints.

Force Feedback: relating to the mechanical production of information that can be sensed by the human kinesthetic system.

Haptics or haptic technology: an emerging interdisciplinary field that deals with the understanding of human touch (human haptics), motor characteristics (machine haptics), and with the development of computer-controlled systems (computer haptics) that allow physical interactions with real or virtual environments through touch.

Haptic communication: the means by which humans and machines communicate via touch. It mostly concerns networking issues.

Haptic device: is a manipulator with sensors, actuators, or both. A variety of haptic devices have been developed for their own purposes. The most popular are tactile-based, pen-based, and 3 degree-of-freedom (DOF) force feedback devices.

Haptic interface: consists of a haptic device and software-based computer control mechanisms. It enables human–machine communication through the sense of touch. By using a haptic interface, someone can not only feed the information to the computer but can also receive information or feedback from the computer in the form of a physical sensation on some parts of the body.

Haptic perception: the process of perceiving the characteristics of objects through touch.

Haptic rendering: the process of calculating the sense of touch, especially force. It involves sampling the position sensors at the haptic device to obtain the user's position within the virtual environment. The position information received is used to check whether there are any collisions between the user and any objects in the virtual environment. In case a collision is detected, the haptic rendering module will compute the appropriate feedback forces that will finally be applied onto the user through the actuators (see Fig. 1.2). Haptic rendering is, therefore, a system that consists of three parts, a collision detection algorithm, a collision response algorithm, and a control algorithm.

Sensors and Actuators: a sensor is responsible for sensing the haptic information exerted by the user on a certain object and sending these force readings to the haptic rendering module. The actuator will read the haptic data sent by the haptic rendering module and transform this information into a form perceivable by human beings.

Tele-haptics: the science of transmitting haptic sensations from a remote explored object/environment, using a network such as the Internet, to a human operator. In other words, it is an extension of human touching sensation/capability beyond physical distance limits.

Tele-presence: the situation of sensing sufficient information about the remote task environment and communicating this to the human operator in a way that is sufficient for the operator to feel physically present at the remote site. The user's voice, movements, actions, etc. may be sensed, transmitted, and duplicated in the remote location. Information may be traveling in both directions between the user and the remote location.

Virtual Reality (VR): can be described as the computer simulation of a real or virtual (imaginary) world where users can interact with it in real time and change its state to increase realism. Such interactions are sometimes carried out with the help of haptic interfaces, allowing participants to exchange tactile and kinesthetic information with the virtual environment.

Virtual environment (VE): is an immersive virtual reality that is simulated by a computer and primarily involves audiovisual experiences. Despite the fact that the terminology is evolving, a virtual environment is mainly concerned with defining interactive and virtual image displays.

Collaborative virtual environments (CVE): is one of the most challenging fields in VR because the simulation is distributed among geographically dispersed computers. Potential CVE applications vary widely from medical applications to gaming.

Simulation engine: is responsible for computing the virtual environment behavior over time.

Collaborative haptic audio visual environment (C-HAVE): in addition to traditional media, such as image, audio, and video, haptics – as a new media – plays a prominent role in making virtual or real-world objects physically palpable in a CVE. A C-HAVE allows multiple users, each with his/her own haptic interface, to collaboratively and/or remotely manipulate shared objects in a virtual or real environment.

1.5   Roadmap to Multimedia Haptics

In a virtual environment, a real scenario is simulated by a computer generated application where some of the user's senses are ingeniously represented in order for them to interact and perceive stimuli that are very similar to the real environment. Traditionally, human–computer interfaces have delivered types of stimuli that are based on two of our senses, namely vision and sound. However, with the addition of the sense of touch through tactile and force feedback, the computer-based applications become richer in media content through better mimicry of real-life situations and tasks or remote real environments.

The sensing of forces is tightly coupled with both the visual system and one's spatial sense; the eyes and hands work collectively to explore and manipulate objects. Moreover, researchers have demonstrated that haptic modality reduces the perceived musculoskeletal loading that is measured through pain and discomfort in completing a task [92]. Therefore, there is a trend in the design of interfaces toward multimodal human–computer interaction that incorporates the sense of touch.

Our perceptions of the world arise as a combination of correlated inputs across several of our senses. With this in mind, we might ask whether it is possible to increase our sensory perception by simultaneously coupling visual cues to the haptic modality in a haptic-based application. In literature, it is found that most haptic-based applications, with the exception of those designed for the visually impaired, seem to be augmented by visual feedback. Many researchers have shown that the interaction with stimuli arriving in more than one sensory modality can increase the realism of a virtual reality. However, the keyword here is "perception", so if the cross-modal information is not well synchronized and consistent, the added sensory information might corrupt the intended stimulus. For instance, researchers have found that when conflict between sensory cues (for instance, between the hands and eyes) arise, the brain effectively splits the difference to produce a single mental image, and the overall perception experienced by the subject will be a compromise between the two senses. Therefore, visual cues must be synchronized with haptic interactions to increase the quality of perception.

It would be easier to extract shape information through visual means than to collect this information haptically. Exploring an object to perceive its shape using the sense of touch places large demands on the observer's memory for the exploration and integration of spatial and temporal signals. In contrast, the optimal exploratory procedures for texture – pressure and lateral motion – are simple and quick to perceive using a haptic modality. Therefore, visual cues help us anticipate the haptic sensation resulting from the interaction with an object. Imagine pressing your hand against a pillow: the visual cues have already prepared you to feel a soft object. In this case, we can say that the visual image has influenced our haptic perception.

On the other hand, many researchers have acknowledged the importance of everyday listening as: "the act of gaining information about events in the world by listening to the sounds they make" [124]. Therefore, the human auditory modality contributes intensively to the perception of the ambient environment.

In the early stages of audio-haptic inter-modal perception, it was shown that auditory stimuli do not significantly influence haptic perception [258]. Later, researchers found that sound cues that are typically associated with tapping harder surfaces were generally perceived as stiffer [95]. These studies suggest that coupling audio and haptics could help create more sophisticated perceptions of solidity, shape, location, and proximity. We believe, however, that the addition of sound to augment the perceptual capabilities of a haptic-based application is constrained by many requirements. For instance, the sound needs to be generated in real-time based on the user's interaction, and it must respond to continuous input data (such as continuous contact force). Furthermore, the synthesized sound must reflect the auditory properties of the contacting objects.

The roadmap toward Haptics Audio Visual Environments (HAVE) comprises three different paths; human haptics, machine haptics, computer haptics, and one roundabout called multimedia haptics (as shown in Fig. 1.3). Notice that the knowledge is cumulative by nature. For instance, to design a proper haptic device, one needs to understand the human haptic road, which investigates human haptic system capabilities and limitations. To develop a proper haptic rendering algorithm,
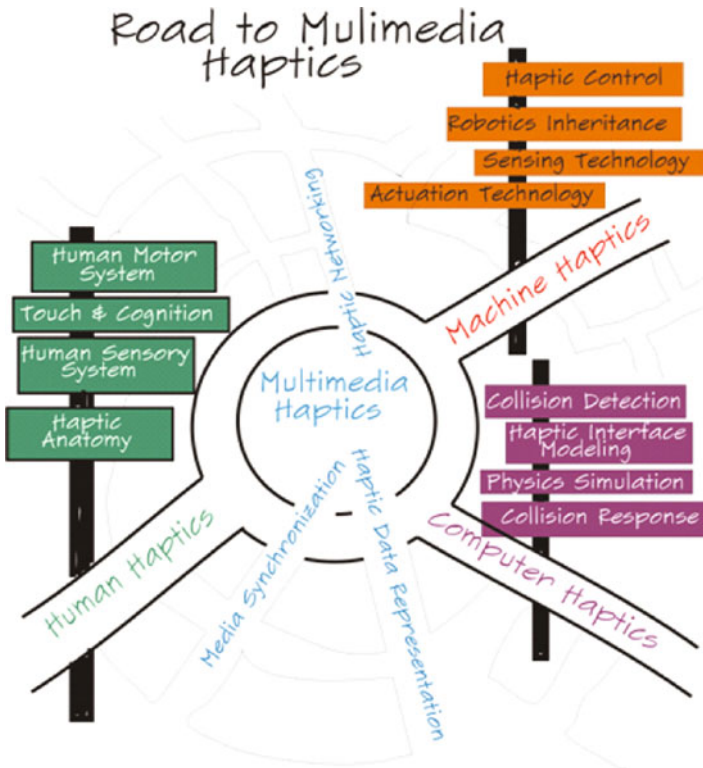

Fig. 1.3 Roadmap to multimedia haptics

one needs a knowledge of spatial and temporal attributes of haptic devices, which lies in machine haptics, etc.

1.5.1   Path 1: Human Haptics

Human haptics refers to the study of human sensing and manipulation through tactile and kinesthetic sensations. When a person touches an object, the interaction force or pressure is imposed on the skin. The associated sensory system conveys this information to the brain, which leads to perception. As a response, the brain issues motor commands to activate the muscles, which results in hand or arm movements. Human haptics focuses mainly on this human sensorimotor loop and all aspects related to human perception of the sense of touch. Therefore, human haptics research deals with all the mechanical, sensory, motor, and cognitive components of the body–brain haptic system.

Haptic perception can be defined as the process of interpreting touch information, or the sense of feeling things via the sense of touch, to recognize objects. It involves tactile perception through the skin and kinesthetic perception through the movements and positions of the joints and muscles. Humans explore and identify an object by moving their fingers on the object's surface or by holding and moving the whole object, which is called haptic perceptual exploration, and it is identified as active touch as opposed to passive touch [310].

The journey toward multimedia haptics starts by understanding the human haptic system, including the tactile and kinesthetic perceptual processes and the functioning of the human perceptual system. Researchers in this domain strive to comprehensively understand the human haptic system. This includes research into understanding the human sensory system, haptic perception and cognition in the human brain, and the human motor system (actuation system). This research also provides guidelines for the design and development of haptic interfaces. Chapter 3 of this book thoroughly covers the fundamental concepts and state-of-the-art research in human haptics.

Once the physiological elements needed to reproduce the real world as a virtual scenario have been identified, we turn to the discipline that covers such requirements in practical terms. The discipline of developing haptic technology has been named "machine haptics".

1.5.2   Path 2: Machine Haptics

Based on the knowledge of the capabilities and limitations of the human sense of touch, the second phase is to design and develop haptic interfaces – or what is referred to as machine haptics. Machine haptics involves designing, constructing, and developing mechanical devices that replace or augment human touch. These

devices, also called haptic interfaces, are put into physical contact with the human body for the purpose of exchanging (measuring and displaying) information with the human nervous system. In general, haptic interfaces have two basic functions; first, they measure the poses (positions and/or orientations) and/or contact forces of any part of the human body, and second, they display the computed reaction touch to a haptic scene that populates touchable virtual objects with haptic properties such as stiffness, roughness, friction, etc. Haptic interfaces can be broadly divided into two categories: force feedback devices and tactile devices. Force feedback devices display force and/or torque and enable users to feel resistive force, friction, roughness, etc. Tactile devices present vibration, temperature, pressure, etc. on the human skin and display textures of a virtual object or provide information such as showing direction, reading text, displaying distance, etc.

Force feedback devices behave like small robots that exchange mechanical energy with users. One way to distinguish between haptic interface devices is by the number of DOFs of motion and/or force present at the device–body interface. Devices with three to six DOFs are mostly used because of their mechanical and programming simplicity in addition to their low cost. The users usually grab and move the device, which controls a tool-type avatar in a haptic scene, and when the avatar makes contact with an object in the scene, the contact force and/or torque is displayed to the user's hand through the device. Multi-DOF force feedback devices such as hand-worn gloves and arm-worn exoskeletons can provide more dexterity but are usually bulky and hard to wear. Combining multiple low-DOF force feedback devices provides simplicity and dexterity such as in two-finger grabbing. Another possible classification of force feedback devices relates to their grounding locations. Two examples are ground-based and body-based. Finally, the desirable characteristics of force feedback devices include, but are not limited to, the following: (1) symmetric inertia, friction, stiffness, and resonant-frequency properties, (2) balanced range, resolution, and bandwidth of possible sensing and force reflection, and (3) low back-drive inertia and friction [329].

Tactile devices are arrays of actuators that have direct contact with human skin. Since an actuator module cannot cover the entire continuous surface of the specific human body part, and since human skin cannot distinguish two adjacent stimuli within a certain threshold (two-point threshold) [316], most tactile devices consist of a number of actuator modules that are uniformly distributed. As discovered through human haptics research, the human body has various two-point thresholds across the body, so the density of the actuators is dependent on these thresholds. For example, the fingertip has a very small two-point threshold compared to that of the arm, so fingertip tactile devices have finely distributed actuators compared to armband-type tactile devices. Tactile devices are also broadly categorized by the stimuli that they can generate, whether it is vibration, pressure, temperature, etc., and they can be further categorized by their actuator types, such as pneumatic, motor, hydraulic, shape memory alloy, etc. Since tactile devices provide cutaneous stimuli while force feedback devices provide kinesthetic stimuli, these two types of devices can be combined to provide a very natural haptic feedback.

In a nutshell, this path involves researchers acquiring knowledge about the existing sensory and actuation hardware technologies and the control of such devices. Researchers are concerned about the design and implementation of efficient and effective sensors and/or actuators that make up a haptic device. Furthermore, this domain explores the attributes that define the quality of haptic interfaces that are based on electromechanical technologies. This domain is extensively presented in Chap. 4 of this book along with a taxonomy of haptic interfaces according to the proposed quality attributes.

Today almost any electromechanical interface requires a human–machine interface, which enables the user to interact with the simulated or remotely located real world. These devices are mainly products of research and development on computational elements related to computer haptics.

1.5.3   Path 3: Computer Haptics

Once haptic interfaces are developed, we move from the machine haptics path to the computer haptics path. Computer haptics is related to the design and development of algorithms and software that compute interaction forces and simulate physical properties of touched objects, including collision detection and force computation algorithms. Essentially, computer haptics deals with modeling and rendering virtual objects for real-time display by touch, and this computing process is called haptic rendering; it is analogous to graphic rendering. We anticipate rapid improvements in computer haptics as computers become more powerful and affordable and sophisticated software tools and techniques become increasingly available.

Since the term haptic rendering has been widely used in literature with slightly different meanings, we explicitly define it as the following:

"Haptic rendering refers to the set of algorithms and techniques that are used to compute and generate forces and torques in response to interaction between the haptic interface avatar inside the virtual environment and the virtual objects populating the environment."

The above definition has many implications. First, the avatar is a virtual representation of the haptic interface whose position and orientation are controlled by the operator. The avatar's geometry and type of contact varies according to the application and can be point based (3-DOF), object based (6-DOF), multipoint based (multiple 3-DOF), or volumetric based. The point-based haptic interface is the most widely used interface since it is computationally efficient for presenting stable haptic interaction and provides pen-like tool-based interaction that allows the user to perform a variety of tasks. Although object-based avatars give more realistic interaction forces and torques, such tools require computationally expensive algorithms; such computations, if not completed promptly, can cause the device to become unstable. These haptic devices are not easily available due to their bulkiness and high cost. An alternative to this is the multipoint-based representation. This is simply a set of multiple point-based representations that is used to provide grabbing

functionality and allow more dexterous interactions such as two-finger grabbing. Volumetric-based representation is usually found in medical applications to enable cutting, drilling, etc., by sacrificing memory storage for accelerated complex and time-consuming computations.

The second implication is that the ability to find the point(s) of contact is at the core of the haptic rendering process. This is the problem of collision detection, which becomes more difficult and computationally expensive as the complexity of the models increases. However, an important characteristic of haptic interaction, locality, drastically accelerates the collision detection process by building hierarchical bounding volumes. While graphic rendering occurs globally to display the whole viewing area, haptic interaction happens in the vicinity of the haptic interface avatar, so the collision needs to be examined only around the avatar. By hierarchically dividing virtual objects into bounding volumes, the only virtual objects that are examined are the ones included in the bounding volume where the haptic interface avatar is located.

The third implication is the need for a force response algorithm. This is the calculation of the ideal contact forces. Upon detecting a collision in a virtual environment, interaction forces between avatars and virtual objects are computed and transmitted to users via haptic interfaces, generating tactile and/or kinesthetic sensations. The interaction force is generally calculated based on a penetration depth, described as the distance the haptic interface avatar penetrates the object it is acting upon. Due to the mechanical compliance of haptic interfaces and the discrete computation characteristics of computers, the haptic interface avatar often penetrates virtual objects [410]. By introducing an ideal haptic interface avatar that has the same position as the haptic interface avatar in free space and cannot penetrate virtual objects, namely a god-object or a proxy, the penetration depth is calculated as the distance between the real haptic interface and ideal haptic interface avatar. As a result, the interaction force is calculated according to Hooke's law[1] using the stiffness value of the virtual object being acted upon. In order to add surface properties such as friction or roughness to the calculated force, the position of the ideal haptic interface avatar on the virtual object can be modulated.

The final implication is that the interaction between avatars and virtual objects is bidirectional; the energy and information flows both from and toward the user. This means that the virtually generated energy in the virtual environment is physically embodied via haptic interfaces and can potentially injure the user or pose a safety problem. Generally, this can be avoided by keeping the haptic rendering update rate higher than 1 kHz, providing a reasonable amount of stable and smooth force to simulate stiff objects [53]. However, in order to guarantee stability of haptic rendering in low power systems, or to keep high fidelity, haptic control algorithms need to be considered as introduced in [81, 147].

---

[1] $F = kx$, where F is the restoring force, x is the penetration depth, and k is a stiffness value of the closest surface.

Consequently, computer haptics provides software architectures for haptic inter-actions and synchronization with other display modalities. Chapter 5 of this book presents the fundamental concepts of haptic rendering along with some discussions about design and implementation details.

1.5.4 The Roundabout: Multimedia Haptics

The last phase in this journey is multimedia haptics, which considers haptics as a new media channel in a complete multimedia system. We define multimedia haptics as the following:

"the acquisition of spatial, temporal, and physical knowledge of the environment through the human touch sensory system and the integration/coordination of this knowledge with other sensory displays (such as audio, video, and text) in a multimedia system"

Multimedia haptics involves integrating and coordinating the presentation of haptic and other types of media in the multimedia application. Generally, a multimedia system consists of media acquisition or creation, content authoring, and transmission and consumption components. Multimedia haptics research can be categorized based on these components as described below [69].

First of all, haptic content needs to be created before it can be delivered to the audience and consumed. While there are a lot of standard tools to capture or synthesize audio and video (AV) media, such as a camcorder, it is less obvious how the same objective can be achieved for haptic media. Basically, haptic media can be created through three key approaches like what has been done in AV media: data can be recorded using physical sensors; it can be generated using specialized modeling tools; and it can be derived automatically from analysis of other associated media.

The acquired haptic media needs to be represented in a proper format so that it can be stored synchronously with other media. There have been endeavors to add haptic media into existing multimedia representation frameworks such as Reachin API to VRML, H3D into X3D, HAML based on XML, and haptic broadcasting framework based on MPEG-4. Furthermore, MPEG-V Media Context and Control (ISO/IEC 23005) is another framework that deals with sensory information, including haptic/tactile modality in a virtual world. Haptic media can be synchronized temporally and spatially with the multimedia representation to produce meaningful content. This requires haptic authoring tools, which are counterparts of audiovisual media production tools such as video authoring tools, 3D modeling tools, etc. 3D haptic modelers, such as HAMLAT (HAML-based Authoring Tool [103]) and K-Touch [336], provide graphic and haptic user inter-faces for adding haptic properties to existing virtual objects. Tactile editors, such as posVibEditor [322], a tactile movie authoring tool [206], enable the creation and editing of vibration patterns that can be synchronized with AV media. We call the resultant authored content "haptic content" to differentiate it from AV content.

The generated haptic content can be stored in files to be delivered through storage devices such as CD/DVD, USB memory drive, etc. or transmitted through the communication network. Sometimes the haptic content may be acquired and transmitted immediately for real-time interactions in a shared virtual simulation. Traditionally, the implementation of the shared virtual simulation is limited by two problems: latency and coherency in manipulation. Delays in processing haptic media can easily bring the haptic interface to a state of instability. Therefore, intense research has been undertaken to reduce delays and jitters in processing and transmitting force information over long distances. The various techniques that were developed to integrate force feedback in shared virtual simulations must deal with significant and unpredictable delays and synchronization issues. One example of such a system is the Collaborative Haptic Audio Visual Environments (C-HAVE), which allows multiple users with their own haptic interfaces to collaboratively and/or remotely manipulate shared objects in a virtual environment.

To recapitulate, this phase covers concepts such as haptic data capturing and representation, transmission and compression, and the synchronized dissemination of haptic media have been explored. One of the most challenging areas of research in haptics is the on-time communication of haptic data, and currently, extensive research is being conducted in the domain of haptic media transmission (or tele-haptics). Several communication frameworks and protocols for haptic data communication, as well as performance issues and challenges, will be discussed in Chap. 6 of this book.

1.6 Haptic-Audio-Visual Multimedia System

With the technologies that have been developed in human haptics, machine haptics, computer haptics, and the multimedia haptics, conventional multimedia systems have the potential to evolve into a haptic-audio-visual (HAV) multimedia system that brings more interactive and immersive experiences because of the haptic modality. For example, while viewers passively watch TV or movies, or gamers interact with video game characters audiovisually through a game controller, the HAV multimedia users would be able to touch the game characters and feel a physical event, such as an earthquake, happening in a movie. Figure 1.4 shows a diagram of the HAV multimedia system; compared to the conventional multimedia system, haptic sensors and displays are added to capture haptic properties and display haptic interactions through corresponding devices. Since most haptic displays have physical contact with the human body, they are designed based on human haptics and machine haptics to provide a comfortable and safe interaction experience. The interaction response, such as a reacting force or vibration on the skin, is simulated and calculated via the haptic rendering process, which works together closely with the graphic rendering process and other simulation engines. In order to provide a stable and high fidelity interaction force, high update rates need to be maintained,
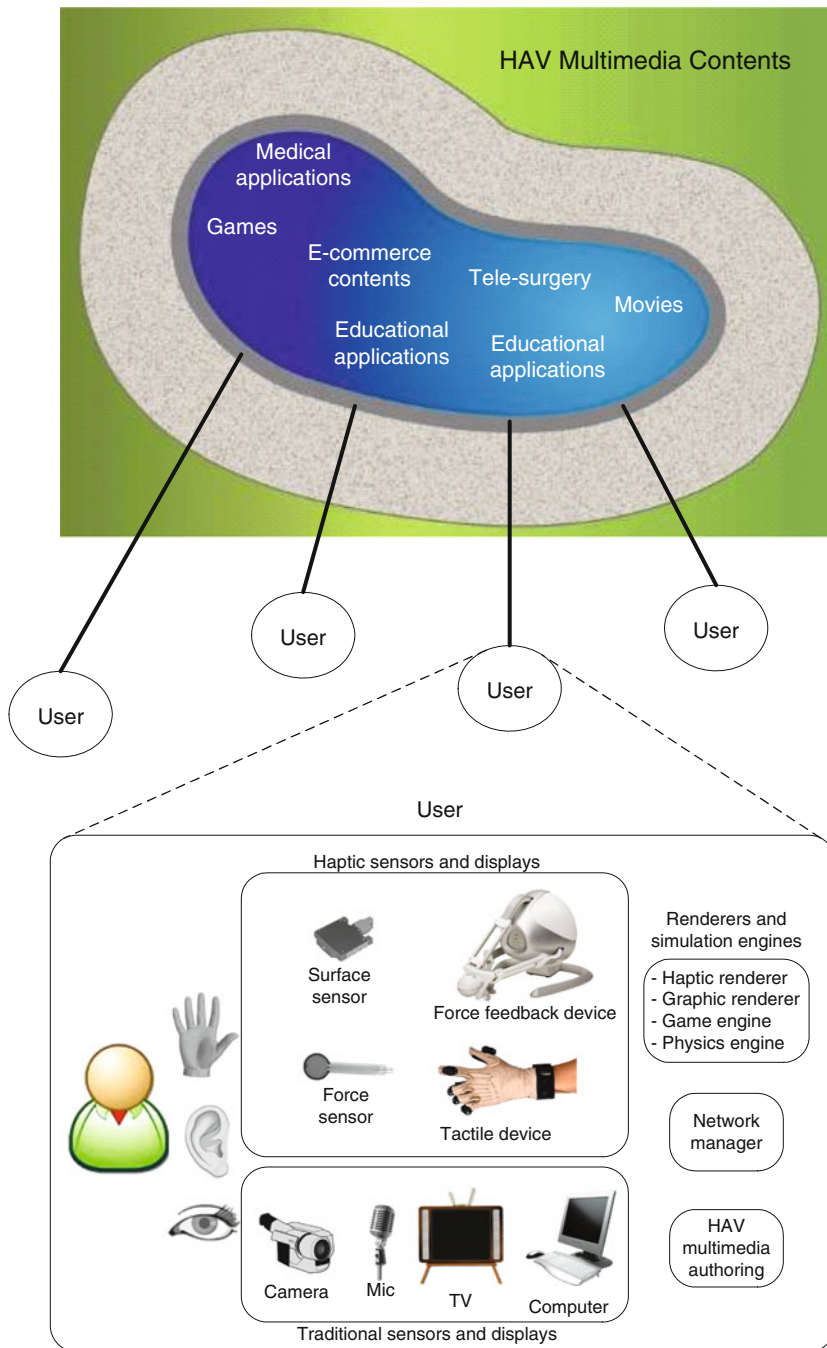
Fig. 1.4 General HAV multimedia system

and the stability of the mechanical system should be guaranteed through computer haptics.

Another important aspect in a HAV multimedia system is their mode of operation. Most HAV multimedia contents work in a standalone fashion; however, networked collaborative environments have gained interest in a society that is more and more interconnected these days. Thus, a haptic transmission over networks such as a dedicated network or the Internet can be located at different places in the architecture based on computer haptic and multimedia haptics. The network management component is responsible for communicating the multimedia contents (haptic, audio, visual, etc.) over a network (dedicated or nondedicated networks). This component implements all network related algorithms to maintain intramodal (within the same media stream) and intermodal (between different media streams) synchronization for the multimedia contents and compensates for network deficiencies such as network reliability and delay/jitter. Several specific haptic algorithms need to be developed to compensate for information loss and network delays and jitter to maintain the stability of the haptic interactions.

The HAV multimedia authoring component allows users or producers to develop HAV multimedia contents and applications. Similar to conventional authoring tools, it should provide modules to import captured media, edit and compose the media into meaningful contents synchronized temporally and spatially, and store the results for delivery.

1.7 Case Study: HugMe Interpersonal Communication System

In this section, we demonstrate the HAVE architecture by considering a specific haptic application called the HugMe system (haptic interpersonal communication system) [105]. Having all the contributing media types, the HugMe application is an excellent example of a complete HAVE system incorporating haptic, audio, and visual information.

The HugMe system enables a parent and child to communicate over the Internet using multimodal interactions (haptic, audio, and video information). As shown in Fig. 1.5, the child is wearing a haptic suit (haptic jacket) that is capable of simulating nurturing physical stimuli. The parent, on the other side of the network, uses a haptic device to communicate his/her feelings to the child. A 2.5-dimensional (2.5D) camera is used to capture the image and depth information of the child and send it to the parent. The parent can use the haptic device to apply forces to the child representation shown on their screen. The interaction information is calculated and sent back to the child, and the child feels the interaction of the parent via the haptic jacket. Meanwhile, the force feedback of the child's image is conveyed to the parent using a Novint Falcon force feedback device.

The HAVE architecture implementing the HugMe system is redrawn in Fig. 1.6. The visual sensor in the HugMe system is the depth video camera that captures the color information (RGB signal) and depth information (D signal). The depth
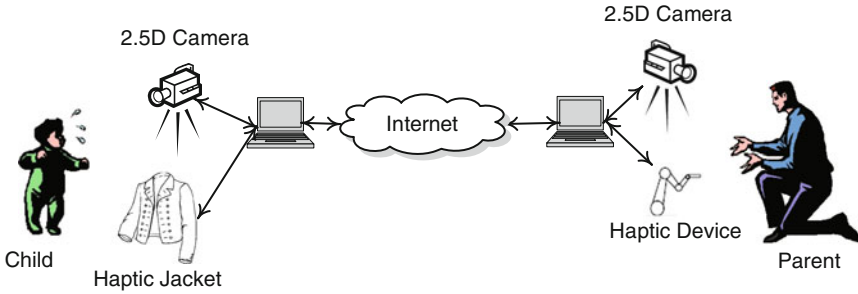
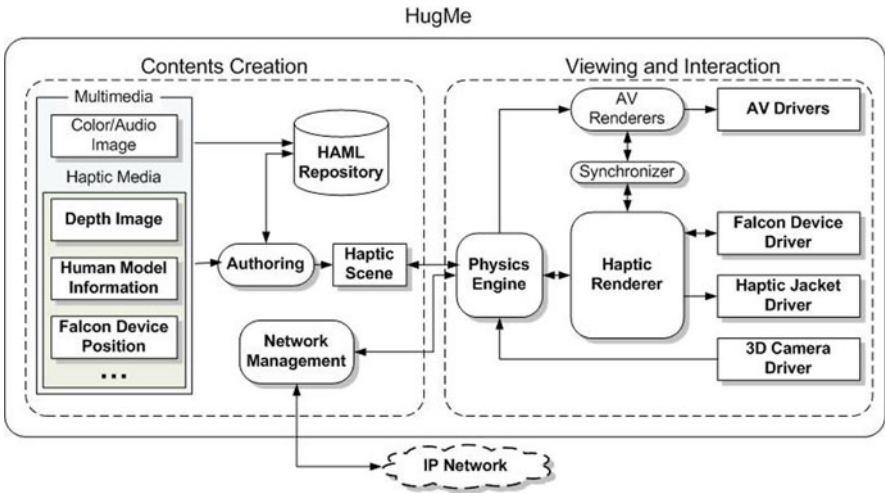**Fig. 1.5** HugMe application scenario

Fig. 1.6 HugMe system architecture

signal is a grayscale bitmap image where each pixel value represents the distance between the camera and the respective pixel in the RGB image. The HugMe system uses a commercially available camera, called the ZCam^{TM} from 3DV Systems, to capture both the RGB and the depth data. Furthermore, special markers are used to track the child's body movements and construct a human model that is used in collision detection. All the captured information is stored in a data repository using the HAML format. The network management component implements Admux (a multimedia communication protocol for synchronous haptic-audio-video communication) [104], which synchronizes the multimedia rendering and adapts it to the network requirements, compensating for any network deficiencies.

The haptic interface used in the HugMe system is the Falcon device, developed and marketed by Novint Technologies, Inc. It provides the parent with the touch feeling whenever the haptic device end-effector collides with an object of the remote environment (in our case the video contents). The Falcon device is both a

haptic sensor (represented by the Falcon device position component) and a force feedback device (shown as the Falcon device driver component), as shown in the HAVE general diagram. At the child's side, the haptic jacket is used to display tactile information to the child. The haptic jacket comprises an array of vibrotactile actuators to simulate continuous tactile feeling to the user. The jacket is connected to the HugMe system using Bluetooth technology to enhance its mobility and wearability.

1.8   Roadmap of the Book

The incorporation of more and more forms of media (from audio and video to touch, smell, and more) is a significant research trend in multimedia systems. The goal is to attain the most natural and intuitive modes of human interaction with a digital world. Haptics is sure to play a prominent role in making virtual objects in these worlds physically sensible and palpable, which increases the realism of these interactions. In particular, the proper use of synchronous haptic interactions results in better quality of experience for the end users. This book provides a basic knowledge of haptics, including current research and commercial potential. The content is spread over four major areas, which are described as follows:

1.8.1   Haptic Applications and Research Challenges

Because the number of possible human activities is unlimited, so too is the number of haptic applications. In Chap. 2, we present a description of application categories showing some of these applications. As a matter of fact, applications of this technology have rapidly been extended to devices used in graphical user interfaces (GUIs), games, multimedia publishing, scientific discovery and visualization, arts and model creation, editing sound and images, the vehicle industry, engineering, manufacturing, tele-robotics and tele-operation, education and training, and medical simulation and rehabilitation. From the literature, one can make several observations as well as recommendations for future research in haptics for multimedia. The literature also helps to pinpoint different research challenges that the haptic community is facing; Chap. 7 does exactly this. These challenges are classified in parallel with the topics covered through the book; many stem from the limitations of haptic device hardware – impractical, expensive, and inaccessible – and the complexity of touch and physical interactions.

1.8.2   General Principles in Haptic Multimedia Systems and Human Haptic Perception

Chapters 1 and 3 explain the basic principles of a haptic system and the foundation and disciplines related to haptics. Chapter 1 introduces the basic concepts and terminology used among the haptic community and provides the "big picture" of what a haptic system is. It also explains the common features of haptic applications; the architecture of a virtual reality application that incorporates visual, auditory, and haptic feedback; and haptic perception and modeling in the virtual environment. Chapter 3 describes the biology of touch in the human body, the classification and measurement methodologies for haptic perception, and perception experimentations and tools.

1.8.3   Haptic Interfaces and Rendering

One of the most important aspects of haptic applications is the haptic interface because it provides a path for perceived stimuli and the human kinesthetic and/or touch channels. This is discussed in Chap. 4. Discussion on haptic rendering is found in Chap. 5 and covers three main topics: collision detection algorithms and their classifications, force response algorithms, and control algorithms. The collision-detection algorithm uses position information collected through sensors to find collisions between objects and avatars and report the resulting degree of penetration or indentation. Next, the force-response algorithm computes the "ideal" interaction forces between avatars and virtual objects involved in a collision. And finally, the control algorithm collects interaction force information from the force-response and applies them on the operator through the haptic device while maintaining a stable overall behavior.

1.8.4   Haptic Audio Visual Environment

In Chap. 6, we study the types and designs of applications that gain access to the virtual object perceptual information through haptic displays. Chapter 6 contains descriptions of the various techniques used in integrating force feedback into shared virtual simulations. This integration requires dealing with significant and unpredictable delays, haptic information representation, synchronization, haptic APIs, existing haptic software frameworks, such as Reachin and Novint e-Touch, and haptic programming toolkits. The chapter elaborates on the discussion of networked haptics (commonly referred to as the Collaborative Haptic Audio Visual Environment (C-HAVE)). Some characteristics, such as quality of experience and security, are also highlighted.

1.9   Further Reading about Haptics

In recent years, there has been extensive research literature on all aspects of haptic systems. Some journals, such as IEEE Transactions on Haptics, IEEE Transactions on Instrumentation and Measurement, ACM Transactions on Graphics, ACM Transactions on Applied Perception, ACM Transactions on Multimedia Computing, Communications and Applications, MIT Press Presence, Springer Multimedia Tools and Applications, Springer Multimedia Systems Journal, and the Electronic Journal of Haptics Research "Haptics-e" frequently publish haptics-based research results. Many other journals, such as Journal of Robotics and Mechatronics, have special issues on this subject.

In addition, a good number of international conferences and workshops are either dedicated to haptics or have special sessions on haptics. Some examples are: the IEEE International Symposium on Haptic Audio Visual Environments and Games (HAVE), IEEE Haptics Symposium, Eurohaptics, Worldhaptics, ACM Multimedia (ACM MM) and Human Computer Interaction (ACM HCI) Conferences, and the IEEE International Conference on Robotics and Automation.

Springer