

Research on Application of Perceptive Human-computer Interaction Based on Computer Multimedia

Lanchuan Wang^a

School of Design and Art, Shenyang Jianzhu University

Shenyang, Liaoning, China

^ae-mail: 3575474@qq.com*Yuxin Fang^b

School of Design and Art, Shenyang Jianzhu University

Shenyang, Liaoning, China

^be-mail: 932258188@qq.com

Abstract—Objective, To explore the characteristics and principles between computer multimedia and perceptual human-computer interaction, analyze the connection between computer multimedia and perceptual human-computer interaction from multiple perspectives, and further understand the basic characteristics of perceptual human-computer interaction through the study of human-computer interaction. Methods Computer multimedia and human-computer interaction methods were analyzed and explored by computer multimedia principle and human-computer interaction principle, and the two were combined to study the case of computer multimedia and human-computer interaction to analyze the interaction principle. Interpret the trend and development of multimedia and human-computer interaction. Conclusion The human perception ability is analyzed by the research of computer multimedia and perception human-computer interaction. Human-computer interaction adhere to the principle of "people-oriented", therefore, the machine should also have abilities of visual, tactile and auditory perception. There are specific Settings of human-computer interaction between computer multimedia and perception system, and the future development space is broad.

Keywords—computer; multimedia; human-computer interaction; perception;

I. INTRODUCTION

Human beings depend on their minds invented computer multimedia. In the process of continuous development, voice recognition, computer vision, virtual reality, touch, space, sound and other technologies to join, so computer multimedia has a powerful interaction with human. Using TV, cell phone, camera, projector, DVD and other multimedia devices connected with computer multimedia, giving some specific "instructions" to the interactive multimedia can make the machine and human interactive operation, to achieve human-computer interaction.

II. COMPUTER MULTIMEDIA

A. Principles of Computer Multimedia

Computer multimedia refers to the digitization of text, sound, video, images, animation and other media information by the computer. "Fig. 1", these media information are integrated into the computer through the interactive interface, so that the computer has the function of displaying different forms of media.

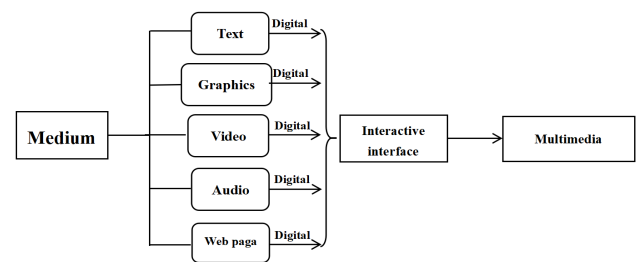


Figure1. The principle of computer multimedia.

B. Computer Multimedia Applications

- *Games*, Interactive games based on electronic media platforms can be divided into computer games, arcade games, console games, mobile games etc..
- *Education*, Computer multimedia learning is the process of learning through multimedia instructional information that is communicated using text, such as pictures, animation, narrative and video, designed to facilitate learning[1].
- *Film and television*, "Fig. 2", taking movie special effects as an example, integrates the production, painting and other functions of computer multimedia into the ideas of the creator, and combine imagination with reality to create a stunning special effects picture. In the special effects production, 3d animation modeling will be used, such as 3Dmax or Maya, then details will be processed according to the materials and finally present the desired effect.

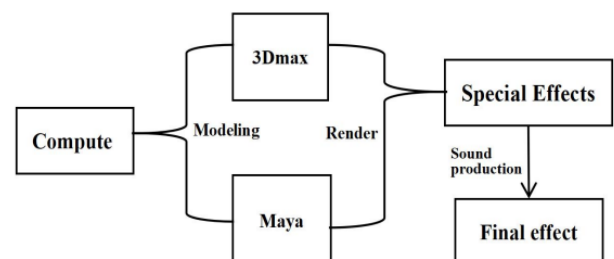


Figure2. Flow chart of film and television special effects.

- *Medical science*, One of the powerful features of interactive computerized multimedia reporting is that it creates a strong link between written imaging reports and image data within PACS [2]. For example, "Fig. 3", medical pathology slides can be scanned to a computer via a digital scanner, turned into "digital" slides, transferred to the physician's computer, and diagnosed remotely. The diagnosis is completed through the diagnosis of "digital" slices. This technology reduces the distance between cities, also, making online diagnosis not only convenient for communication between doctors, but also convenient for patients to seek medical treatment.

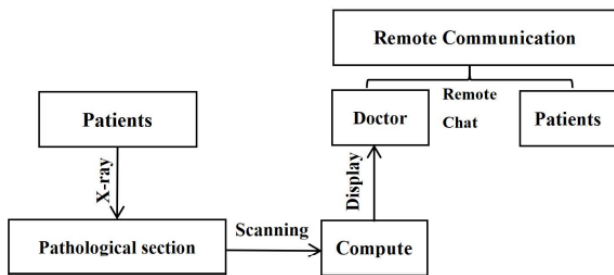


Figure3. Remote diagnosis process

III. HUMAN-COMPUTER INTERACTION

Human-computer interaction refers to the interaction between users and an interactive computer system [3]. Literally, "human" means "human being". Human being experience the world through cognition, appearance, the five senses, context, environment, and physiology, etc. "Machine" refers to "system", which can only experience the external world through system, software, controller, etc. Therefore, human-computer interaction (HCI) is a field that combines various disciplines such as computer science, psychology, human factors, and ergonomics [4].

A. Basic Forms of Human-computer Interaction

- *Commands*: Send instructions to the system through the user and select options. For example, enter commands, gestures and buttons.
- *Dialogue*, When the user talks and interacts with the system that controls the system's menu drivers. For example, in "Fig. 4", the intelligent voice "Siri" in the iPhone, when the user talks to the phone, the voice will be recognized and immediately encoded into a digitized voice message file. This digitized voice information data is sent to a cloud computing server via the Internet network (LSP). The data center recognizes the user's voice and provides feedback via Internet data and local data.

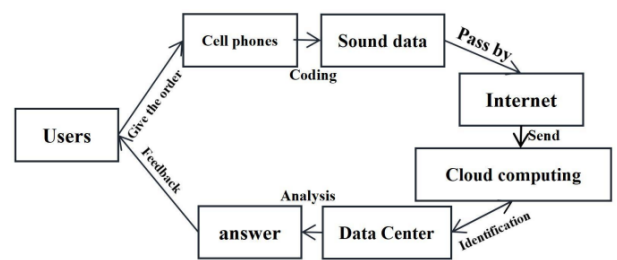


Figure4. Speech recognition flow chart

- *Manipulation*, Refers to interacting with objects in virtual or physical space by manipulating them (e.g., opening, holding, closing, and placing).
- *Exploration*, Refers to moving through a virtual or real space where they can deepen their familiar knowledge by actually moving around.
- *Response*, The System can start a session (initiate an interaction) and the user can choose whether or not to respond. For example, if the user is interested in a notification message from the app, he or she will click on the pop-up message on the phone and interact with it.

B. Basic Mode of Human-computer Interaction

- *WIMP*, Refers to windows, icons, menus, pointers, the mainstream interaction mode reflected by the graphical interface.
- *Direct*, This action command is immediately displayed on the screen, i.e. "what you see is what you get".
- *Natural interaction*, "Fig. 5". It is a new generation of human-computer interaction, which studies the dialogue between human and machine. The natural interaction approach emphasizes more on the way that human and robot communicate with each other in the same ways as human beings [5]. For example, the guiding robots in shopping malls. Therefore, when it comes to human-computer interaction design, attention should be paid to its characteristics such as harmony, naturalness and interactivity.

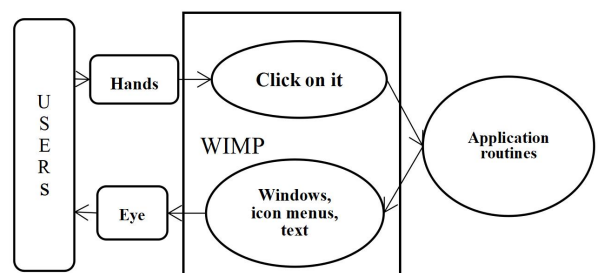


Figure5. WIMP interface interaction flow

- *Multimodal interfaces*, "Fig. 6", modality refers to a source of information, multimodality refers to various information or forms of interaction, and multimodal

interaction refers to people communicating with the computer through channels, such as voice, body language, information carriers, and the environment to simulate human interaction.

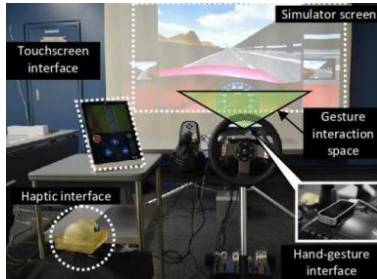


Figure6. Multimodal Simulation,Source network

IV. RESEARCH ONE THE APPLICATION OF MULTIMEDIA AND PERCEPTUAL HUMAN-COMPUTER INTERACTION

A. Perception

From the era of click, touch, sound to the current era of somatosensory, after years of development of human-computer interaction, machine perception technology gradually occupies an important position in human-computer interaction. However, computer multimedia has no life because they are mechanical, a system composed of many parts and with various functions [6]. Human-computer interaction takes human as the main subject, and human cognitive process is the focus of human-computer interaction design. As one of the main sources of human cognition of the world, it includes vision, hearing, and touch. Vision is the most important sensory channel that connects people with their surroundings, and hearing converts the external stimuli involved into neural excitement, which is processed by the human body and transmitted to the brain. The human sense of touch (tactile perception) differs from our other senses in an important way, because it gathers information about the external world mainly through active exploration [7].

B. Analysis of the Application of Computer Multimedia and Perceptive Human-computer Interaction

1) Multimodal experiment

a) *Equipment*: projector, computer, camera, kaleidoscope, multimodal operating table, stereo.

b) *Experimental process*: "Fig. 7", the prepared photos and music were imported into the computer and pictures were placed on the interactive display through the computer. At the same time, the camera took a picture of the kaleidoscope through the computer connected to the projector and placed it on the display. Music is played by the audio system. The computer software recognizes the high and low frequencies of the music and controls the size of the screen picture. "Fig. 8", on the interactive console, a person holds the kaleidoscope and the camera and adjusts the angle of the kaleidoscope for the purpose of real-time interaction.

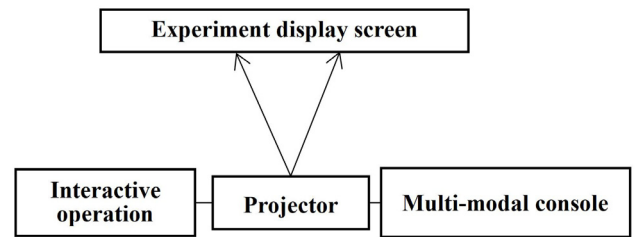


Figure7. Distribution of multimodal experiments



Figure8. Multimodal experimental diagram

2) Virtual reality

"Fig. 9", uses computer multimedia technology to simulate people's visual, auditory, and tactile sensations from space and location, thus achieving an immersive effect.



Figure9. V R game diagram, Source network

3) Network virtual anchor

a) *Network virtual anchor*: "Fig. 10", the network virtual anchor is essentially a computer program, formed by the digitization of the character after the computer simulation of the character's appearance, language, actions, etc. The appearance of the "network virtual anchor" is an image presented by a terminal such as computer 3D technology. Facial expressions and body movements can be operated in real-time by computer "commands". These "commands" are collected by a motion capture system from simulated data of human or computer human-computer intelligence. The 3D displacement data in the commands are matched with the 3D model marks of the network virtual host to realize synchronous control, so as to control the movement of the network virtual host. Its voice can be played through the audio player, the audio source can be recorded by real people, text conversion and other ways to achieve. For example, the virtual anchor "Kumquat 2049" appears in a short video platform.



Figure10. Webcast of virtual human, Source network

b) Motion capture technology: motion capture technology mainly includes: mechanical type, acoustic type, electromagnetic type, inertial type, optical type, where optical is one of the common forms of motion capture technology, etc.. It consists of three elements: camera, marker point and data processing system. Optical motion capture is achieved by tracking and monitoring the position points on the target. It is mainly applied to the principles of computer vision, where images taken by depth cameras are correctly processed by computer vision algorithms designed to discriminate human motion from the background in a scene [8]. If a point in space is captured by two cameras at the same time, then the position of the point can be calculated from the parameters of that point by both cameras. If the camera's capture rate is high then the motion trajectory of the point can be captured.

c) Motion Capture technology: Optical motion capture systems generally use six to eight cameras arranged in a wrap-around fashion. In order to capture the movement more accurately, people who need to be collected are generally required to wear dark clothes. The venue should be selected in a place where the light must be uniform, preferably from above, and the surroundings must be clear and uniformly colored [9]. "Fig. 11", In the key parts of the body with special reflective dots, the computer vision system will recognize these "dots", the system determines the marker, the character moves, the camera captures the person's action trajectory continuous shooting, these images will be saved for analysis and processing. Similarly, if we put "dots" on the key points of facial expressions, we can achieve expression capture.



Figure11. Motion capture process, Source network

V. DEVELOPMENT TREND OF COMPUTER MULTIMEDIA AND HUMAN-COMPUTER INTERACTION APPLICATIONS

In the future, multimedia-based perceptual human-machine interaction applications will more fully meet people's needs; more refined biometric systems; and more comprehensive

perception systems; All kinds of interaction methods will have a deep development, the use of intelligent devices will naturally react and communicate with human beings, including thinking, vision, touch, hearing, etc., and even in the future, human-machine interaction will rise from the external interaction of machinery to the level of emotion and attitude, so that the machine and human beings can communicate more freely.

VI. CONCLUSIONS

Through the research on computer multimedia, human-computer interaction, and human-computer perception, it is concluded that computer multimedia is an important research direction in the future, while people are more concerned about the natural communication between human and machine, and then "perception" will become the primary task of human-computer interaction research. In the age of the Internet, with the computer vision technology that simulates human vision; Motion capture technology that simulates human action; Speech recognition technology simulating human language is developing rapidly, and more perceptual human-computer interaction will become the basic mode of human-computer interaction in the future, such as virtual anchor. In education, medicine, military, life and service, the production and innovation of human society will be realized.

REFERENCES

- [1] Cwp A, Dgk B, Sc C, et al. Adoption of multimedia technology for learning and gender difference - Science Direct[J]. Computers in Human Behavior, 2019, 92:288-296.
- [2] Steven D. Beesley, James T. Patrie, Cree M. Gaskin. Radiologist Adoption of Interactive Multimedia Reporting Technology[J]. Journal of the American College of Radiology, 2019, 16(4).
- [3] Costa S D, Barcellos M P, Falbo R, et al. A core ontology on the Human-Computer Interaction phenomenon[J]. Data & Knowledge Engineering, 2022, 138:101977-.
- [4] Alkathairi Mohammed S.. Artificial intelligence assisted improved human-computer interactions for computer systems[J]. Computers and Electrical Engineering, 2022, 101.
- [5] Sharma R P, Verma G K. Human Computer Interaction using Hand Gesture[J]. Procedia Computer Science, 2015, 54(3):721-727.
- [6] Al-Ma'aitah Mohammed, Alwadain Ayed, Saad Aldosary. Application dependable interaction module for computer vision-based human-computer interactions[J]. Computers and Electrical Engineering, 2022, 97.
- [7] A.J Bongers. Interaction in multimedia art[J]. Knowledge-Based Systems, 2000, 13(7).
- [8] Bortolini M, Gamberi M, Pilati F, et al. Automatic assessment of the ergonomic risk for manual manufacturing and assembly activities through optical motion capture technology[J]. Procedia CIRP on SciVerse ScienceDirect, 2018.
- [9] Daniele Regazzoni, Giordano de Vecchi, Caterina Rizzi. RGB cams vs RGB-D sensors: Low cost motion capture technologies performances and limitations[J]. Journal of Manufacturing Systems, 2014, 33(4).