Journal of Computing and Electronic Information Management

ISSN: 2413-1660

Kinect sensor principle and Application

Shuwen Wang, Lu Zhang *, Xiang Xie
College of Electrical and Information Engineering, Quzhou University, Quzhou Zhejiang 324000

Abstract: Through Kinect's understanding and step-by-step study of Kinect for Window SDK development, the hardware construction, development requirements and other configurations of Kinect are discussed, and through Kinect's functions and application areas, let us have a new understanding of human-computer interaction, physical operation, The continuous development of virtual reality technology has brought the development of human-computer interaction technology to a new height. In the future life, human-computer interaction technology will become more and more mature.

Keywords: Kinect; Bone Tracking; Human-Computer Interaction; Body Technology.

1. Introduction

With the continuous progress of science and technology, the way of human and computer interaction has entered a new stage of development, more and more people began to pay attention to research Human-computer interaction technology, and human-computer interaction has become an important research topic for researchers in recent years. With the continuous development of human-machine interface technology, early perforated paper bags, panel switches and other interactive ways gradually developed to the present voice, movement, image recognition and other devices with a variety of interactive functions. In today's globalization, all countries in the world have invested a lot of human and material resources in human-computer interaction technology, human-computer interaction technology has gradually become the core competitiveness of information technology.

With the rapid renewal of science and technology, we have ushered in an unprecedented era of information intelligence. The rapid development of computer technology and the wide application of computer make the way of human-computer interaction much anticipated, at the same time, the study of human-computer interaction method has gradually become one of the hot topics in this field. In our

daily communication, often in different ways of communication, such as spoken language, expressions, gestures and so on to enhance each other's communication and understanding skills. However, the computer and human "communication" but no action, expression and it seems very unnatural, so, let the machine receive and understand the human language, which is the level of human-computer interaction and strengthen the practicality of human-machine interface and so on has a significant significance.

2. Introduction to Kinect

2.1 Kinect Overview

In 2010, Microsoft introduced Kinect, a human-computer interaction device based on body-feeling interaction, to meet users' needs and give them a better gaming experience. Kinect was originally released as an external device for Xbox360 consoles, mainly using motion capture technology, speech recognition technology, synchronous tracking technology to allow the experiencer to always follow the characters or animation changes in the game, the experiencer can use their own body movements to complete the game must be played through the gamepad before, really achieve an immersive feeling. Four years later, Microsoft worked tirelessly to develop the second generation of Kinect for Windows,as shown in Figure 1. Kinectv2 has been a big hardware improvement compared to the first generation. As the needs of users grow,the Kinect for Windows SDK is constantly upgrading core technologies with user feedback. So far,Kinect v2 is the best device for combining deep images, bone images, and color images, truly embodying many of the features of a device that enables video image processing.

Figure 1 KinectV2

2.2 Kinect hardware architecture

Kinect has three lenses, the middle of which is an RGB color camera that captures color images. The left and right lenses are 3D structured light depth sensors consisting of infrared transmitters and infrared CMOS cameras, respectively, to capture depth data (the distance from objects in the scene to the camera). Color cameras support up to 1280 x 960 resolution imaging, and infrared cameras support up to 640 x 480 imaging. Kinect is also paired with focus-chasing technology, with the base motor turning as the focusing object moves. Kinect also has an array microphone built in, with four microphones simultaneously, which eliminates noise and uses it to capture sound for speech recognition and source positioning

2.3 Kinect SDK organization architecture

Kinect for Window SDK makes it easier for developers to implement human-computer interaction. First, Kinect takes image and audio data from the target scene, and then transmits the data information of the target object to the NUI. Then THE is calculated in the NUI library based on the raw data obtained, and the converted data is transferred to the NUI application interface, through which third-party software can interact with Kinect, so that researchers can complete the later research Job. NUI is a shorthand for The User Interface, which represents the natural user interface. It is an important part of Kinect for Window SDK, which processes the various information captured by Kinect while also managing the device.

3. Kinect hardware requirements

3.1 System requirements

Windows 8 (x64)
Windows 8.1 (x64)
Windows 8 Embedded Standard (x64)
Windows 8.1 Embedded Standard (x64)

3.2 Hardware Requirements:

64-bit processor
More than 4 GB of memory
CPU: I7 3.1 GHz (or higher)
USB 3.0 (Intel or Renesas chip), external graphics card must support win8
Graphics card: NVIDIA AMD supports DX11 graphics cards are available, win8, 8.1 recommends using video cards that support Direct X11 Win10 recommends using video cards that support Direct X12

4. Kinect features

4.1 Bone imaging technology

Skeleton tracking relies on Kinect v2's infrared sensor, which is primarily a unique black-and-white spectroscopic technology to sense changes in the outside environment and to represent the distance between the target object and the camera in different colors. When the Kinectv2 device is operating normally, the entire visual range of the Kinect v2 camera will be fully covered by the laser emitted by the infrared transmitter, and then use optical coding technology to encode the visible range of space, optical coding is in fact to the space to be measured, generally using the method of lighting, in essence, is still a structural light technology.

This diffraction light point when the laser reaches an uneven surface object or shines through the foggy glass is laser speckle, these scattered speckle graphics will not be

exactly the same, only in space used such structural light, according to the characteristics of these spots, you can know the position and distance between different objects, so in the obtained depth image, you can use the depth of different objects to make a specific split of the image, Then the infrared receiver is used to receive diffraction spots in space, so that the object is mapped to the coordinate system of Kinect v2, the depth image can be obtained, and the bone point map can be obtained by using bone sensing technology.

Figure 2 Laser spots

Kinect v2 can detect bone point information for six people at the same time, including the identification of two people's movements, each with a detailed record of 20 bone points, and the camera can capture data ranging from limbs, torso, fingers, and more to build a full-body operating system. In order to be able to see the user's movements more carefully, the system uses machine learning methods, on the basis of which the Kinect v2 camera can identify the user's behavior information to the maximum extent possible.

4.2 Bone Tracking:

In the SDK, each bone point is represented by the Joint type, and each frame consists of 20 bone points that form a collection based on the Joint type. This type contains three properties, as shown below. JointType: The type of bone point, which is an enumeration type that lists specific names for 20 bone points, such as "HAND_LEFT" that indicates that the bone point is a left-handed node.

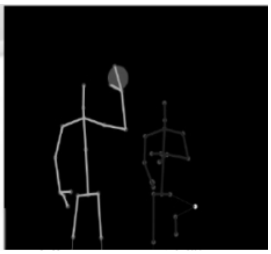

Figure 3 A diagram of a human joint

Position: The SkeletonPoint type represents the location of bone points. SkeletonPoint is a structure that contains three data members, X, Y, and Z, to store the three-dimensional coordinates of bone points. TrackingState: The JointTrackingState type is also an enumeration type that represents the tracking state of the bone point. Where Tracked indicates that the bone point was captured correctly, NotTracked says it did

not snap the bone point, and Inferred indicates that the state is uncertain.

Bone tracking data is obtained in the same way that an application gets the next frame of bone data, as it does by calling callback functions and passing a cache, and getting bone data is called by the OpenSkeletonFrame() function. If the latest bone data is ready, it is copied to the cache, but if the new bone data is not ready when the application makes the request, you can choose to wait for the next bone data until it is ready, or immediately return to send the request later. For the NUI Bone API, the same bone data is provided only once.

4.3 Air mouse :

The air mouse is designed to use Kinect's bone tracking features to track the displacement of a user's left or right hand on a spatial XY coordinate platform, and then call the Windows Native API to simulate mouse cursor movement. By tracking the event through bone, the spatial coordinate PST of the user's left or right hand is captured first. Second, map the XY plane coordinate portion of the PST coordinates to the screen coordinates. This process can be done using the Coding4Fun.Kinect toolkit. First add the Coding4Fun.Kinect.Wpf .dll and Microsoft.Expression.Drawing .dll to the Solution's Reference reference, and declare the use Coding4Fun.Kinect.Wpf in the namespace reference. References are made in Coden4Fun Kinect Toolkit

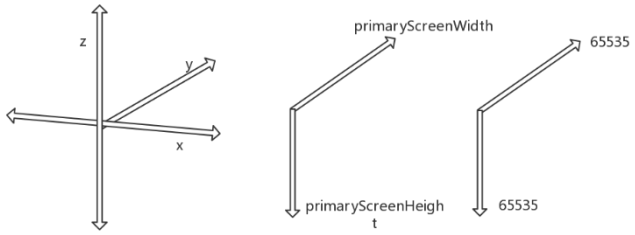

Figure 4 Coordinate conversion

The extension method of Joent.ScaleTo, which maps bone coordinates to screen coordinates. Because the absolute coordinate system and screen coordinate system of mouse movement are not exactly the same - consistent, the absolute coordinate system of the mouse X-axis, Y-axis point coordinates are Int type, the value range is (0 to 65535), so the middle also need to do a conversion, as shown in the figure.

5. Kinect application area

5.1 AR photography

Make photography smarter with Kinect V2 real-time HD high-quality photography and synthesis into virtual scenes, switching background images (instead of traditional screen virtual buttons) with "right-hand swipes" and right-handed lifting trigger-taking (replacing traditional screen virtual buttons).

5.2 Body games

At present, the popular game system is mostly physical, virtual separation, participants in the process of experience single interest, very narrow income, and even in some ways will have a negative effect. Smart Eye launched a Kinect-based simulation game system, through face recognition, dynamic tracking, gesture recognition and other technologies can be a good combination of physical games, virtual games advantages, fully bring participants a new gaming experience. In the course of the game participants need to use the hand and brain to manipulate the machine, but also to think about the rules of the game, cooperation with teammates and other factors, compared with the traditional game single, limited, simulation game system really achieve the purpose of the game, and kinect,human-computer interactive robot and other scientific and technological fields are very good trial application, in guiding young people to pay attention to the development of science and technology has a certain significance.

5.3 Cast screen

Work with Kinect with a projector to enable human-machine interaction on any wall based on projection technology. In view of the problem of inaccurate positioning and disturbance interference on the traditional large screen, the affine transformation algorithm is used to achieve accurate location of interactive area coordinates. Use Kinect infrared reception positioning to capture interactive movements and analyze them to determine the real intent of the interacting person.

5.4 Gesture control

With a depth-of-fieldcamera and an RGB camera, you can place physical 3D images on the screen. At the same time, take color and infrared images, capture the user's gesture movements, according to the data given, and then convert these gesture language into the corresponding commands and execute

5.5 Fall monitoring

Fall monitoring system will Kinect v2 for indoor human body detection, the use of bone tracking technology and RGB combined methods can be 24h real-time detection, combined with Kalman algorithm to track the movement of the elderly, can overcome the impact of changes in external lighting conditions, improve the accuracy of fall detection, can be timely alarm treatment of fall behavior, to ensure the safety of the elderly.

5.6 Sitting position detection

Whether it's student parties or office workers, more and more people are becoming "sedentary", myopia sufferers and people with spinal diseases are increasing, and those who use wearable physical orthotics or sensor-based electronic correctors can place an additional burden on users.

In view of this problem, the sitting position detection technology is studied comprehensively. Compared with traditional wearable physical orthosis or simple electronic correctors based on acceleration sensors and horizontal sensors, the method of identifying irregular sitting positions based on machine vision techniques has the advantages of being more precise and flexible.

5.7 Smart Home

Using Kinect body sensing technology, a character identification and posture recognition system based on human bone structure is designed, and combined with Internet of Things technology, it can make up for the shortcomings of the existing identification system and can be used as the second choice of the character recognition system. At the same time, the posture recognition is used in intelligent control home appliances, according to human bone information recognition, for different home use recorded custom posture control, users do not need mobile phones and other third-party media can complete the control of home appliances. System design of face recognition and bone matching door, stranger alarm, children anti-falling building and other security measures to ensure family safety, and truly realize the intelligent home life.

Acknowledgments

This work was partially supported by Zhejiang Province Science and Technology Innovation Program for College Students (No. 2020R471005).

References

[1]Nie Hongliang. Research on the development of human-computer interaction in China under the view of scientific and technological innovation. Yanbian University, 2018.

[2]Wu Xia, Zhang Zi, Xu Yanxu. Summary of the Development of Gesture Recognition Research . Electronic Technology, 2013, 26 (6): 171-174.

[3]Yu Chao. Visual-based gesture recognition research. China University of Science and Technology, 2015.

[4]Tao Xuejun. Study on the History and Trends of Human-Computer Interaction Development. Technology Communications, 2019, 11 (22): 137-139.

[5]Zhang Feng, Chen Shuo. Review and Prospects of Multi-Touch Interactions Ergonomics, 2010, 16 (04): 76-78.

[6]6. Beret. Research on robot control technology based on speech recognition. Southwest Petroleum University, 2014.

[7]Li Zuhe, Yu Yu. Summary of Visual-Based Emotional Analysis Research. Computer Application Research, 2015, 32 (12): 3521-3526.

[8]HeBei, Wang Guijin, Lin Xinggang. Fast video keying algorithm combined with Kinect depth map. Journal of Tsinghua University: Nature Science Edition, 2012, 52 (4):561

[9]LuKaiyang, Ye Huamao, Li Xiaoguang, etc. The application and prospect of Kinect's body-

feeling technology in the experimental teaching of animal surgery. China Medical Education Technology, 2012, 26 (2): 171-173.