Smart Home System Based on Deep Learning Algorithm

Yanfei Peng1,a, Jianjun Peng2,b*, Jiping Li3,c*, Ling Yu4,d

1School of Information Science and Engineering, Dalian Polytechnic University, Dalian Liaoning 116034, China +86-18366881989
2School of Information Science and Engineering, Dalian Polytechnic University, Dalian Liaoning 116034, China +86-13942061732
3College of Mathematics and Informatics, South China Agricultural University, Guangzhou Guangdong 510642, China +86-13138692766
4Network Information Center, Dalian Polytechnic University, Dalian Liaoning 116034, China +86-15524599728

a343626363@qq.com, Corresponding author: bpengjj@dlpu.edu.cn, Corresponding author: cli-jiping@163.com, dyuling@dlpu.edu.cn

ABSTRACT This paper presents a smart home control system on the strength of human body point cloud data attitude recognition technology. Wherein, the human body point cloud image is obtained by Kinect, and the point cloud image is extracted, and the human body attitude is recognized by Convolutional neural network (CNN) algorithm. The Arduino microprocessor will process the received data to realize the intelligent control of residential electrical appliances in the process. The experimental results show that the system algorithm is able to achieve effective recognition of human attitude and effective control of household appliances and other functions through the analysis and processing of human point cloud images, which features a certain degree of innovation, reliability and practicality.

1. INTRODUCTION
With the development of society and the improvement of living standard, the smart home industry is developing rapidly correspondingly. An attitude recognition technology based on depth sensor has gradually entered the field of vision. In November 2010, the launch of Microsoft Kinect depth sensor provided hardware support for the design and development of smart home based on attitude recognition technology [1]. It is simpler, more natural, and humanized compared with the traditional human-computer interaction system using mouse, keyboard and other modes of operation, and human attitude recognition technology with these new features is playing an important role in human-computer interaction.

Smart home is a kind of living environment for human being, integrating intelligent control and management of various subsystems related to home life through the network using advanced computer technology, network communication technology and other integration of personality needs to make home life more convenient and comfortable. There are lots of literatures recoding the studies of theory and technology of smart home, wherein, literature [2-5] refers to the interactive design pattern theory of somatosensory, that is, obtaining the key points of human skeleton data through Kinect and using to recognizing human attitude by various algorithms. However, the skeleton data is extremely unstable, easy to lose, and the algorithms are mostly threshold judgment, recognition rate need to be improved.
thereby. Literature [6-7] connects all kinds of smart devices in the home by using wireless Bluetooth technology, through which users are able to directly control the smart devices in real time by using mobile phone client. Nonetheless, the problem of poor user experience still exists. In literature [8], the integration of smart grid and smart home was studied, however, it mainly merely involved the control of household appliances, whereas the research of interactive technology was not mentioned.

In order to solve the problem that the user unable to interact with the machine in a natural way or the low recognition accuracy in the current smart home system, this paper integrates the human body point cloud data attitude recognition technology into the design of the smart home human-computer interaction system, which uses Kinect to obtain the human body depth image and point cloud image. The algorithm is used to extract the features of human point cloud image, and the Convolutional neural network model is imported to recognize and classify human attitude. The processing result was then transmitted to Arduino microprocessor through XBee wireless communication module to realize intelligent control of home appliances.

The contributions of this paper are as follows: Firstly, the human point cloud images obtained by Kinect are not affected by external illumination, and it is capable of effectively protecting the privacy of users. What’s more, Kinect point cloud data has higher confidence than skeleton data. Secondly, a more diversified set of human attitude point cloud image data was created, and a Convolutional neural network model was built for training and testing. Thirdly, it improves the rigidity of human-computer interaction and the single control object of the original equipment, contributing to offering more natural human-computer interaction and improving the attitude recognition rate so that the "intelligent" component of smart home will be significantly improved.

2. SYSTEM STRUCTURE AND DESIGN

Smart home control system based on human body point cloud data attitude recognition technology is combined with software Processing and Kinect to obtain depth images in the scene, and three-dimensional point cloud images were generated therein. An algorithm was used to extract image features, which were assigned to trained classifiers and then used to judge human attitude. The determined data is transmitted to the Arduino microprocessor via the wireless communication module XBee to control various household appliances. The overall structure of the system is shown in Figure 1.

![Figure 1. Overall structure diagram of the system.](image)

2.1 Kinect Sensor

The somatosensory device used in this system is a kind of somatosensory peripheral Kinect issued by Microsoft. Kinect is comattitude of color camera, infrared device, microphone array, logic circuit, motor and other parts. Kinect has three cameras, the RGB color camera in the middle, the left and right lenses are infrared transmitters and CMOS infrared cameras comattitude of 3D depth sensors, can obtain color images and depth images at the same time. The microphone array, consisting of four downward-facing built-in microphones. Figure 2 depicts the depth image and the corresponding point cloud image taken by Kinect at night. As we have observed, the system can be used late at night or even at night, because Kinect can extract the point cloud image [10] from a lightless room.
2.2 Arduino
Arduino mainly consists of two parts: the hardware part is Arduino circuit board, and the other is Arduino IDE, which is a program development environment stored in the computer. Arduino is an open source hardware development platform with 8 bit ATMEGA328 microprocessor as the core, which provides 14 digital input and output pins and 6 analog input pins. It can support USB data transmission. Humans can connect different electronic devices on the I/O port [9], as shown in Figure 3.

3. DESIGN AND IMPLEMENTATION OF SYSTEM ALGORITHM
In this paper, Kinect was used to acquire the point cloud image of human body, and the attitude recognition instructions of various households were defined. The Convolutional neural network model was established, the sample images were captured and labeled, and each kind of attitude training samples was input into CNN to train the model, through which adjusting the model until it turns to be astrigency; the output layer was transformed into a soft Max classifier. Input test images to identify the accuracy of validation results. Through this model, human attitude was recognized accurately and efficiently, and finally the desired results were achieved. The algorithm box is shown in Figure 4.
3.1 Implementation of Kinect Control Instruction

This novel input method of Kinect provides a new idea for the control of smart home. Firstly, it is necessary to personalize the attitude recognition instructions of all kinds of home, as shown in Table 1. Taking body attitude as input element, the system can understand the command and realize the control of smart home.

| Electrical type | Operation instruction | Attitude operation |
|-----------------|-----------------------|--------------------|
| Lamps           | Create lamps          | Raising hands above the head |
|                 | Selection lamps       | Unbending right arm |
|                 | Adjust luminance      | The distance between the left hand and the head |
| Air conditioner | Open the air conditioner | Body presenting "X" shape |
|                 | Turn off the air conditioner | Presenting "T" shape |
|                 | Refrigeration mode    | Raising left hand highly, right hand on waist |
|                 | Heating mode          | Arising right hand highly, left hand on waist |
| Window          | Open the window       | Left hand on waist, laterally raising right hand |
|                 | Suspend window        | Both hand on the waist |
|                 | Close the window      | Right hand on waist, laterally raising left hand |

3.2 Convolutional Neural Network Model Design

Convolutional neural network is a kind of feedforward artificial neural network, which is developed on the basis of traditional neural network. Nowadays, Convolutional neural network has been widely used in image and video recognition, and it has been an appropriate and effective method for many computer vision problems.

In the first few layers of the Convolutional neural network, each node is organized into a three-dimensional matrix. Each node in the first several layers of Convolutional neural network is only connected with some nodes in the upper layer. Plus, a Convolutional neural network is comprised of input layer, Convolutional layer, pooling layer, full connection layer and Softmax layer as shown in Figure 5. The CNN recognition model designed in this paper has a total of seven layers. Firstly, it comes to the input of the network. CNN network is able to learn the features of two-dimensional image independently. The original image is point cloud image, which can be used as the input of the network directly. The Convolutional operations create 64 features, the first full connection layer will have 384 hidden nodes, the second connection layer will connect 384 hidden nodes to 192 hidden nodes, and finally output 10 classification results.
3.3 Dataset
In this paper, 720 times of image shooting of 6 people and 10 attitudes were collected to build attitude training database. Wherein, the 10 attitudes are raising hands above the head; unbending right arm; body presenting “大” shape; presenting “T” shape; raising left hand highly, right hands on waist; arising right hand highly, left hand on waist; left hand on waist, laterally raising right hand; both hand on the waist; right hand on waist, laterally raising left hand and other attitudes, grouping the images collected to attitude1- attitude10. The image effect of the point cloud image part of collected body attitude by shooting is shown in Figure 6.
Guide the dataset into a convoluted neural network model for training and testing. Firstly, the network is initialized, and each group of attitude samples in the dataset is input into the CNN network model separately, whereby one attitude is trained in one round. With the increase of the number of iteration rounds, the training and testing accuracy of the model will be improved. Then, the loss function value will be reduced, and it converges and tends to be stable after 800 iterations, as shown in Figure 7, whereby the parameter values of the network is determined.

4. EXPERIMENT AND ANALYSIS

In this paper, the system is tested experimentally. The system designed is deployed in a real home environment, and the Kinect is installed in a suitable location to cover the whole home detection environment. The intelligent home system designed in this paper is capable of processing the collected information and completing the control instructions simultaneously without causing the fluctuation of processing speed. Performance degradation. When someone enters the Kinect sensor's field of vision, the system will automatically start, through a specific operation to open attitude command input, real-time collection of human body information and complete control commands, point cloud image to protect the privacy of users.

4 volunteers participated in the evaluation of the algorithm and system in this paper. Each volunteer performed the above nine attitudes to control different household electrical appliances, each attitude was performed 25 times. Through the training of Convolutional neural network as well as the commissioning of Arduino control system, the experimental results obtained are shown in Table 2. It can be seen that the recognition accuracy reaches up to above 93%, and the expected accuracy is achieved.
Table 2. Accuracy of smart home control

| Attitude operation | Operation instruction | Test times | Correct times | Accuracy |
|--------------------|-----------------------|------------|---------------|----------|
| 1. raising hands above the head | Create lamps | 100 | 96 | 96% |
| 2. unbending right arm | Selection lamps | 100 | 97 | 97% |
| 3. body presenting “T” shape | Open the air conditioner | 100 | 97 | 97% |
| 4. presenting “I” shape | Turn off the air conditioner | 100 | 95 | 95% |
| 5. raising left hand highly, right hand on waist | Refrigeration mode | 100 | 94 | 94% |
| 6. arising right hand highly, left hand on waist | Heating mode | 100 | 95 | 95% |
| 7. left hand on waist, laterally raising right hand | Open The window | 100 | 94 | 94% |
| 8. both hand on the waist | suspend | 100 | 95 | 95% |
| 9. right hand on waist, laterally raising left hand | Close the window | 100 | 93 | 93% |

This paper designs a smart home system based on Kinect using depth learning algorithm and a more natural human-computer interaction control system, which is able to meet the user's demand for intelligent home environment. The detection accuracy of this method is higher than that of the Kinect skeletal tracking method adopted in [2-5], addressing the problems of unstable data of human skeletal joint points and the situation of easy to get lost. The Convolutional neural network constructed and trained in this paper also witnessed some progresses compared with reference [14]. In addition, compared with the traditional intelligent home control method in literature [6-8], it improved the rigidity of the original equipment and the unitary control object, realizing more natural human-computer interaction, and protecting the privacy of users at the same time.

5. CONCLUSION
This paper introduces a reliable and more natural human-machine interaction smart home system. The point cloud images of human living environment are captured by Kinect, and the features of different human attitude images are extracted. Plus, the Convolutional neural network model was built to recognize human attitude accurately and efficiently. The Arduino microprocessor would conduct real-time processing upon receiving the control instructions, and control the work of household appliances by wireless a well. The experimental results showed that the system has certain practical value. During the operation of the system, there is no need wearing any devices or controllers, instead, it contributes to objects interacting with the machine in a simpler and more natural way; using the Kinect point cloud image for color spectrum detection, which is able to effectively protect privacy; meanwhile, the Kinect infrared camera will not be affected by external light. It can work at any time, improving the detection efficiency.

ACKNOWLEDGMENTS
The authors wish to express their gratitude towards the financial supports from Science and Technology Department of Liaoning province (No. 20170052), Education Department of Liaoning province (No. 2017J047) and China Science Technical Department (No. 2017YFC0821003).

REFERENCES
[1] Qu Chang, Sun Jie, Wang Junze, et al. 2016. Automatic detection of falls in elderly people based on Kinect somatosensory sensor Journal of [J]. sensing technology, 2016,29 (3): 378-383.
[2] Shen Muqi, Qi Shunlan. 2017. Internet of Things smart home system based on somatosensory interaction technology [J]. Modern information technology, 2017, 1 (4): 42-44.

[3] Su Benyue, Wang Guangjun, Zhang Jian. 2013. Smart home system based on somatosensory interaction technology under the Internet of Things [J]. Journal of Central South University (Natural Science Edition), 2013, 44 (s1): 181-184.

[4] Yu Zesheng. 2017. Research and Design of Smart Home System Based on Kinect attitude Recognition [D]. Liaoning University of Science and Technology.

[5] Guo Zhe, Chen Peitou, Hu Mengkai, et al. 2016. Kinect-based Smart Home System [J]. Modern Electronic Technology, 2016, 39 (18): 149-152.

[6] Hou Yuyi, Yang Dongtao, Liu Yan, et al. 2016. Smart Home Life and Security System Based on Wireless Bluetooth Technology [J]. Journal of Jiaying University, 2016, 34 (5): 36-40.

[7] Li Tao. 2014. Design and implementation of Android based smart home APP [D]. Suzhou: Soochow University.

[8] Naglic M, Souvent A. 2013. Concept of Smart Home and Smart Grids integration [C]. Energy, International Youth Conference on. IEEE, 2013:1-5.

[9] Smisek J, Jancosek M, Pajdla T. 2013. 3D with Kinect[J]. Advances in Computer Vision & Pattern Recognition, 2013, 21(5):1154-1160.

[10] Kepski M, Kwolek B. 2013. Human Fall Detection Using Kinect Sensor[J]. 2013, 226:743-752.

[11] Klemenjak C, Egarter D, Elmenreich W. 2016. YoMo: the Arduino-based smart metering board[J]. Computer Science - Research and Development, 2016, 31(1-2):97-103.

[12] Barbon G, Margolis M, Palumbo F, et al. 2016. Taking Arduino to the Internet of Things: The ASIP programming model[J]. Computer Communications, 2016, s 89 – 90:128-140.

[13] Krauss R. 2016. Combining Raspberry Pi and Arduino to form a low-cost, real-time autonomous vehicle platform[C]/ American Control Conference. IEEE, 2016:6628-6633.

[14] Dai XiGuo. 2017. Research on human attitude recognition based on Convolutional neural network [D]. Chengdu University of Technology.