Research on Gesture Recognition and Denoising Technology of Electronic Information Teaching Demonstration Based on Kinect V2

Hongyu Yu
Yuefeng Li
Xingjian Jiang

Motion sensor is a kind of sensor which is commonly used in the field of human-computer interaction. This article uses Microsoft's Kinect sensor to get in-depth information about gesture recognition in electronic instructional demonstrations. Here, we present the Kinect V2 sensor for use in demonstrating gesture capture and recognition in electronic information instruction. In this study, depth information is used to dynamically capture electronic information teaching demonstrative gestures. In addition, a combined denoising method is also proposed, which can effectively remove the interference and noise compared with the single denoising method. Then, the researchers programmed and moved the dynamic recognition of demonstrative gestures in electronic information instruction. Compared with some traditional denoising methods, the combined denoising method can effectively remove the blur and boundary burr. This study can be applied to the field of human-computer interaction in electronic information teaching to further improve the accuracy of information.

Keywords: Kinect V2, gesture capture and recognition, depth information, combined-denoising

Tob Regul Sci.™ 2021;7(5-1):4133-4143

DOI:doi.org/10.18001/TRS.7.5.1.189

INTRODUCTION

Nowadays, the information processing becomes more important than ever before. In this paper, we suggest using MicrosoftKinectV2 sensor to obtain the information of dynamic demonstration gesture recognition (HGR) in electronic information teaching.. Microsoft designed the first generation of Kinect in 2010, which has functions such as motion capture, image recognition, voice recognition and recognition, etc. Therefore, a large number of researchers use Kinect for research. For example, Su B et al. [1] studied Kinect's recognition and classification of human skeleton. Neto LB et al. [2] used Kinect for face recognition. Ding J et al. [3] used Kinect's microphone array for speech recognition. In 2013, Microsoft improved Kinect and created Kinect V2, which has more advantages, such as higher resolution, deeper depth and color field of view. It can accommodate six people, identify 25 joint points (V1 is 20 joint points), track up to 6 people (V1 can achieve 2 people) and more available data. In order to better study gesture recognition in electronic information teaching demonstration, we use KinectV2 sensor to study. Gesture recognition of electronic information teaching demonstration based on Kinect is mainly divided into two categories, RGB color image and depth data image. Because the amount of data in color image is too large and the calculation time is too long, there is color information that interferes with the image characteristics and data. Therefore, this paper uses depth images to demonstrate gesture recognition in electronic information teaching.

Hongyu Yu Harbin University of Science and Technology Rongcheng Campus, Rongcheng 264300, China, Yuefeng Li Harbin University of Science and Technology Rongcheng Campus, Rongcheng 264300, China, Xingjian Jiang Harbin University of Science and Technology Rongcheng Campus, Rongcheng 264300, China, *Corresponding author: Yuefeng Li Harbin University of Science and Technology Rongcheng Campus, Rongcheng 264300, China (E-mail: Nicholaslouisfeng@outlook.com)*

On the other hand, electronic information teaching demonstration gesture recognition can be divided into dynamic gesture recognition and static gesture recognition. In the current research, there are many methods to study static gesture recognition, for example, Krisandria KN [4] used Kinect for gesture recognition and used dynamic time warping (DTW) as a classifier for recognition. Alonso DG, Teyseyre A, Berdun L et al. [5] performed deep learning based on LMC's complete data gesture library for gesture recognition. Guo H et al. [6] used LSTM-RNN algorithm for gesture recognition and the accuracy and Robustness is high. Regarding the dynamic gesture recognition, Kane L et al. [7] performed dynamic gesture recognition based on a depth matrix and an adaptive Bayesian classifier. Dynamic gesture recognition is mainly divided into hand detection, finger recognition, gesture recognition and denoising. For a gesture recognition system, its practical range is between 1.2m-3.5m. The deep data obtained during dynamic gesture recognition is continuous and continuously monitored. Most of the current research is written in C# or C++. The program is relatively cumbersome. In contrast, an image acquisition toolbox has been developed in MATLAB, which provides functions and modules that can interactively detect and configure camera properties. This includes the Image Acquisition Toolbox Support Package for Kinect For Windows Sensor support package for Kinect so that the obtained data can be directly imported into MATLAB and used. Therefore, we use MATLAB [8] to write programs.

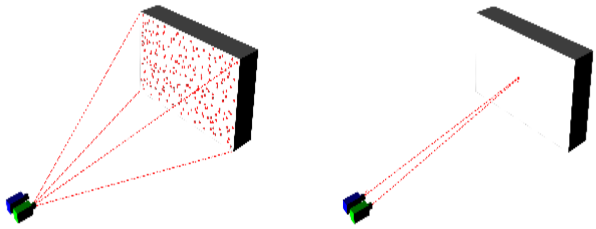
For gesture recognition in electronic information teaching demonstration, image processing is very important, especially in information extraction. There are various denoising methods, such as bilateral filtering [9], median filtering [10], mean filtering [11], Wiener filtering [12], morphological filtering [13] and so on. Demonstration of Electronic Information Teaching of Kinect Depth Image Gesture recognition mainly involves image clarity and boundary burr. Therefore, we propose a joint de-noising method to carry out de-noising. The principle of joint denoising is as follows: the first step is processed by morphological filtering, then the mean filtering is followed by wiener filtering, and finally the final image is obtained by de-marginalizing. The results of each denoising are analyzed through experiments. At last, KinectV2 is used to capture the dynamic demonstration gesture of electronic information teaching.

GESTURE RECOGNITION IN ELECTRONIC INFORMATION TEACHING DEMONSTRATION

Kinect Sensor

Kinect mainly obtains images through two methods. One is to obtain color images through RGB cameras and the other is to obtain deep images through Kinect's depth camera. The method adopted in this paper is to obtain deep images to demonstrate gesture recognition in electronic information teaching. At the same time, Kinect needs to be connected to a computer that meets the configuration to download and install Kinect Windows SDK 2.0 to complete the hardware configuration process. Kinect's deep information is mainly detected and captured by the deep sensor. The core technology of Kinect V1 is to obtain and process depth information through infrared emitters and infrared cameras, and finally determine the distance of the target object. The deep information uses light coding technology, which is different from traditional structured light because its light source is not a periodically changing two-dimensional image coding, but a three-dimensional deep "volume coding". The Kinect V2 obtains deep information through the time reflected from the projected infrared pulses. The matching depth sensor has also been changed. The deep sensor uses Time of Flight (TOF) technology. The basic principle of TOF is to transmit modulated light pulses through an infrared transmitter. After encountering an object reflection, the receiver receives the reflected light pulses based on the light pulses. The round trip time calculates the distance to the object. Compared with optical coding technology, the accuracy of TOF technology during measurement will not be reduced due to the increase of distance. Because sunlight is not modulated, it can be simply considered that it has no effect on phase. Therefore, TOF is also robust to outdoor strong light environments. Working principle of Kinect V1 and V2 is as follows Figure 1.

Figure 1.
Working principle of Kinect V1 and V2.



Gesture Recognition in Electronic Information Teaching Demonstration

To capture the hand motion state during gesture demonstration in electronic information teaching, firstly, Kinect support package in MATLAB should be used to preprocess KinectV2, so as to realize the recognition, capture the hand position and read the related data of the call. Then we use the threshold to remove the interference of the environment. And then we collect and extract the deep data matrix around the right hand. Finally, we denoise the deep image and compare the matrix data with the target gesture matrix data to complete the recognition.

First, construct the data matrix around the right hand. Since Kinect itself can locate and track the position of the hand, we use its palm tracking function to perform gesture recognition. The deep image pixel is 424×512, which is composed of pixels. Therefore, the data is expressed as a 424×512 order matrix. The rows and columns of the matrix respectively correspond to the two-dimensional coordinates of the XY plane of the depth map image. The value corresponding to the row and column is the deep value, which means the distance of each point of the image on the Z axis. The matrix is expressed as:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1512} \\ a_{21} & a_{22} & \cdots & a_{2512} \\ \vdots & \vdots & \cdots & \vdots \\ a_{4241} & a_{4242} & \cdots & a_{424512} \end{bmatrix} \quad (1)$$

If gesture recognition is needed in electronic information teaching, redundant influencing factors must be removed first. Therefore, it is preliminarily envisaged to establish a rectangular frame around the palm in the deep image to determine the segmentation range of the gesture and prepare for the subsequent data extraction. The target range can be established directly by collecting the real-time coordinates of the Kinect palm tracking.

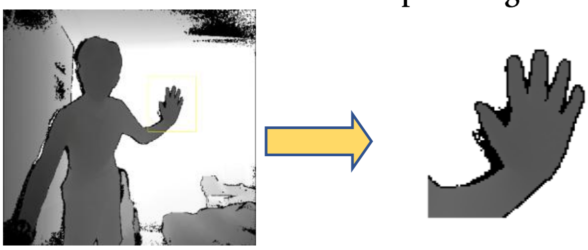
The range of the rectangular frame is established according to the palm can be represented by a matrix. The matrix is included in the above matrix, refer to matrix (2).

$$\begin{bmatrix} a_{(x-50,y-50)} & \cdots & a_{(x-50,y+50)} \\ \vdots & a_{hand(x,y)} & \vdots \\ a_{(x+50,y-50)} & \cdots & a_{(x+50,y+50)} \end{bmatrix} \tag{2}$$

Where $a_{hand(x,y)}$ represents the position of the palm in the matrix, that is, the position coordinates in the X-Y plane of the palm.

After the rectangular frame is established, the deep image can be displayed by only reading the data matrix. The image in the rectangular frame (Figure 2a) can be initially segmented. It can be observed from the image that after the initial segmentation. There are still a lot of interference factors, such as arms and environment, and the image has shadows. Regarding the formation of shadows, we believe that due to the blockage of the human body during reception after infrared rays are emitted, some light signals cannot be transmitted back. Therefore, shadows appear. Under this assumption, the shadow can be eliminated by removing the environmental interference from the back of the hand (Figure 2b).

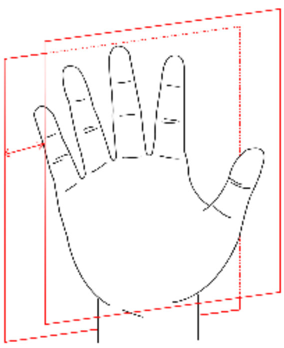
Figure 2.
Gesture extraction depth image.



(a) Depth image after segmentation (b) Deep image after removing interference

The purpose of this paper is to obtain the coordinate information of palm and finger of electronic information teaching demonstration gesture. Therefore, in order to further optimize the image, the interference of other data still needs to be removed. Therefore, the palm is separated from the environment by setting a depth threshold, and the position coordinates of the wrist are introduced to remove the arm (Figure 3).

Figure 3.
Schematic diagram of threshold processing



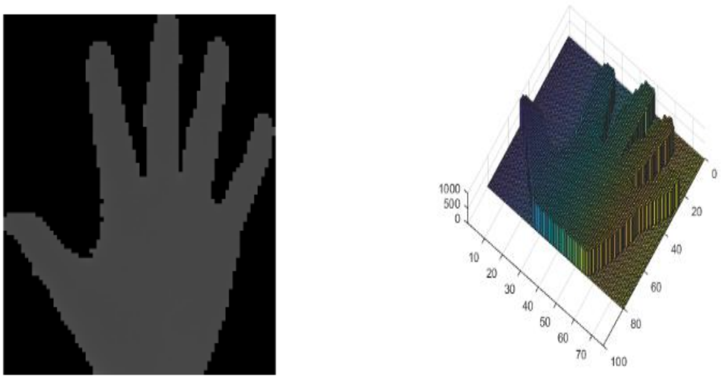
The steps are as follows:

1. After the demonstration gesture of electronic information teaching is presented, the background interference is great. In addition to setting the deep information of the required palm, the deep values of other factors are all 0 and other useful information is retained. Find the pixel point with the smallest deep value on the hand. During the normal operation, the fingertip leans forward slightly and we set this point as the fingertip point.

2. Based on the depth of the fingertip, we read the corresponding deep data to determine the approximate range of the palm at the recognition position. The choice of threshold is an empirical value. The final threshold added after the program test in this article is 1300. Clear all depth values beyond a certain range.

After further optimizing the image, a deep data matrix with a smaller range around the hand is obtained and then scan the matrix. After the threshold processing, the deep information retained in the image range is only the palm(Figure4a). Taking the leftmost, topmost and rightmost non-zero depth values in the matrix scan results as the boundaries, respectively, the smallest matrix containing gesture information can be obtained to complete the segmentation of the target gesture (Figure 4b).

Figure 4.
The processed images



(a)Gesture segmentation effect image (b) Depth value 3D image

A dynamic gesture recognition method for electronic information teaching demonstration based on information data processing is proposed. After obtaining the gesture information matrix, first calculate the left, right, top and bottom distances of the hand, respectively. It can be preliminarily judged whether the gesture matches the target gesture. When the distance is within a certain error range, the preliminary recognition screening is completed and the specific recognition of the gesture information can be further carried out, that is, by scanning the gesture information matrix. Therefore, we scan the pixels with non-zero values to get the position information and determine the shape of the gesture. And then we compare the result of the scan with the gesture information matrix of the target gesture. When the similarity rate exceeds 90%, we can determine that the recognition is successful. If the similarity is low, continue to identify until it is successfully stopped.

JOINT DENOISING

Because there are many irregular burrs and noise points in the gesture recognition image of electronic information teaching demonstration, it is necessary to further process the collected image. The effect of noise is not very great, so the burr of the boundary needs to be removed. Therefore, the term 'Joint denoising' is used here to refer to remove the interference information. The principle of joint denoising is to be processed by the expansive and corroded method at the first stage. Then the mean filtering is carried out after the boundary corrosion. And then the wiener filtering is used to carry out the filling for the deburring processing. Finally, the edge blur is removed, and the final gesture recognition image of electronic information teaching demonstration is obtained.

Based on Morphology----'Swelling and Corrosion Denoising'

In the field of 'Swelling and Corrosion Denoising' in morphology, two ways of it are found. The one is opening operation, which refers to corrosion and then expansion. The other one is a closed operation, which refers to expansion and then corrosion. The opening operation is used to set the image processing, which refers to use the imerode (X, B) in Matlab. Firstly, the idea of corrosion is carried out, and a set of 4*4 square structure elements and the set of covered target elements are used to do the logical "and" operation. If the last two sets are both 1, the element is 1. If not, it is 0. If A is the target element and B is the structural element, the corrosion principle is as equation (3).

$$A \otimes B = \left\{ x \middle| \left(B_y \subseteq A \right) \right\} \tag{3}$$

Expansion is also a logical "or" operation with a set of structural elements of size 4*4 and a set of overridden target elements. If the last two sets are 0, the element is 0. If not, it is 1. If A is the target element and B is the structural element, the corrosion principle is as equation (4).

$$A \oplus B = \left\{ x \middle| \left(\left(B_y \right) \cap A \right) \neq \varnothing \right\} \tag{4}$$

At last, MATLAB is used to program to get the corroded image (Figure5a) and we compares the corroded image with the original depth image (Figure 5b).

Figure 5.
Compare corroded image with the original depth image.

(a) original image (b) Corroded image

It can be observed that the palm was obviously corroded off the edge in a circle according to the corrosion elements. However, the subsequent expansion treatment in this paper does not adopt the way of expansion elements, but takes the mean value of edge elements to carry out expansion.

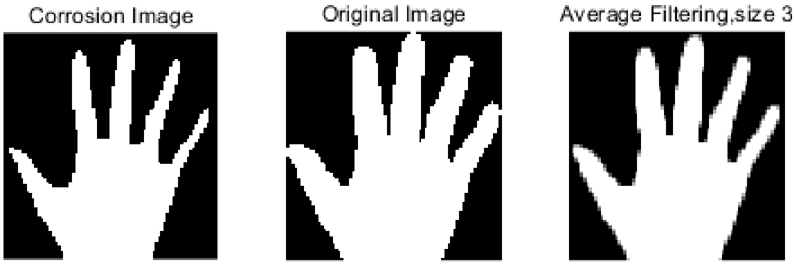
Mean Filtering

Compared with the bilateral filtering and other methods, the edge of the depth map can be expanded by means of the mean filtering compared with the original relatively obscure. In the application of this paper, the edge burrs can be filled by taking the mean value of image elements, so as to eliminate some burrs. Due to the corrosion treatment first, the filling is equivalent to the subsequent expansion. In this way, the original shape and size of the palm can be well maintained at the same time. The mean filtering with template size of 3 is adopted in this paper. The mean filtering is mainly Neighborhood Averaging. Assuming that each pixel point is (x, y), then select a region, which is composed of several adjacent pixels. And then take the mean value of all pixel points in the region to get the image at the pixel point, which refers to g(x, y). The formula is as equation (5).

$$g(x,y) = \frac{1}{N} \sum_{i=1}^{N} f(x_i, y_i) \tag{5}$$

Finally, the comparison between the processed image and the original image is as Figure 6.

Figure 6.
Comparison of mean filtering effects.



It can be observed that after the initial filling, on the basis of maintaining the original hand shape, some burrs are effectively filled. However, the defect of the image with shallow filling degree and after average processing is that the image becomes relatively fuzzy. Therefore, further image processing is needed to improve the clarity of gesture recognition images in electronic information teaching demonstration.

3.3. Wiener Filtering

Wiener filtering [14] is an algorithm for image processing in the frequency domain. It is a very classic image enhancement algorithm, which can not only reduce image noise and eliminate image blurring caused by motion, but also make the image edges smooth. Since the mean value denoising was used before and some fuzzy factors were carried out, the Wiener Filter could be used to effectively solve the fuzzy problem. The filter-based deblurring is filtered by some filter estimators, and the adaptive low-pass Wiener filter in this paper is used to filter the image. Wiener filters can adapt themselves to local image variance. When the variance is large, Wiener2 hardly performs any smoothing processing when the variance is large. When the variance is small, Wiener2 executes more smoothing process. This method usually produces better results than linear filtering. Adaptive filters are more selective than similar linear filters because they can retain edges and other high-frequency parts of the image. The principle of the algorithm is as follows:

Wiener2 estimates the local mean and variance around each pixel as equations (6)-(7):

$$\mu = \frac{1}{NM} \sum_{x,y \in \eta} a(x,y) \tag{6}$$

$$\sigma^2 = \frac{1}{NM} \sum_{x,y \in \eta} a^2(x,y) - \mu^2 \tag{7}$$

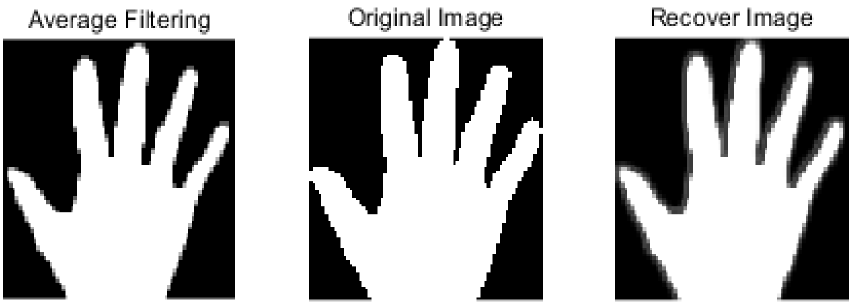
Where η is the local neighborhood of each pixel's $N \times M$ in the image. Wiener2 uses these estimates to create pixel-level Wiener filters, as equation (8).

$$b(x,y) = \mu + \frac{\sigma^2 - \upsilon^2}{\sigma^2} \left[\mathrm{a}(x,y) - \mu \right] \tag{8}$$

Where υ^2 is the variance of the noise. If no noise variance is given, Wiener2 uses the average of all local estimated variances.

Finally, the original image, mean image and Wiener filtered image are compared. These images can be seen in Figure 7.

Figure 7.
Comparison of Wiener filter effects.



Eliminate Edge Burrs

The main reason for the observed edge fuzziness is the edge value with small depth due to the calculation of the edge area with small proportion during denoising [15]. Therefore, the edge removal algorithm is finally adopted to eliminate the edge blur, which refers to eliminate the smallest depth value of the edge, so as to obtain a relatively clear edge burr depth map. In order to remove the burrs, the fuzzy weights are obtained to eliminate the fuzzy values with small edge weights. The weight calculation formula is as equation (9).

$$w(i, j) = e^{-\frac{|g(i,j)-g(x,y)|^2}{2\delta^2}} \tag{9}$$

In the formula, the parameter controls the attenuation degree of the brightness similarity factor. When the difference of the depth value between the two pixels points increases [16], the calculated weight decreases accordingly, and the pixel point whose weight is less than a certain value is set back to zero to eliminate edge blurring.

The original image, bilateral filtering, median filtering, mean filtering, Wiener filtering [17], morphological filtering and combined denoising images are evaluated in this paper. The evaluation index is the Edge Strength, and the gradient operator is the Sobel Operator. Edge is a basic feature of image, and Edge Intensity refers to the amplitude of edge point gradient. Setting the first difference in x and y direction in the ith row and j th column of the image as equation (10 11).

$$\nabla x f(i, j) = f(i, j) - f(i-1, j) \tag{10}$$

$$\nabla y f(i, j) = f(i, j) - f(i, j-1) \tag{11}$$

Then taking the gradient value of each pixel of each image, as equation (12).

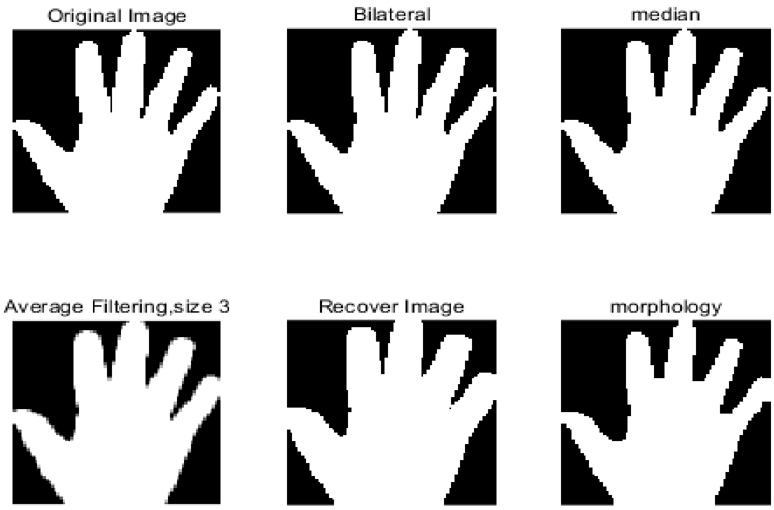
$$G(i, j) = \sqrt{\nabla x f(i, j)^2 + \nabla y f(i, j)^2} \tag{12}$$

Then taking the mean value of each gradient and getting the judged average edge strength as the evaluation index. The edge strength is normalized to comprehensively evaluate the image. The closer the edge strength is to 1, the higher the image evaluation will be. This article uses a program to get the Histogram of edge depth values. Finally, a high-precision image of electronic information teaching demonstration gesture with edge burr removed is obtained.

EXPERIMENT ANALYSIS

In this paper, bilateral filtering, median filtering, mean filtering, Wiener filtering and morphological filtering are used to process and compare the original gesture images of electronic information teaching demonstration. Each method has shortcomings but does not achieve the best denoising effect. For the bilateral filtering, the denoising effect is not obvious and there are obvious boundary burrs; For the median filtering, the denoising effect also has boundary problems and there is a finger joint phenomenon, so it is not recommended to use; And the mean value filtering is to divide the pixels evenly and construct them again. Therefore, the mean value filtering can effectively remove the boundary burrs, but the image will become fuzzy and the finger crack adhesion phenomenon is serious; The Wiener filtering has a certain effect on the image boundary burr and the fuzzification is not very serious, which is a better filtering method. But after Wiener filtering, image recovery will have the situation of thickening, finger joint phenomenon is also heavy; For the morphological filtering, the method of dilatation-corrosion is used for denoising. According to the results, burrs will also be reduced in the image, but there are still some adhesions between the fingers, resulting in a large deformation error of the image. The images after the experiment are as shown in Figure 8.

Figure 8.
Gesture images with different denoising methods.

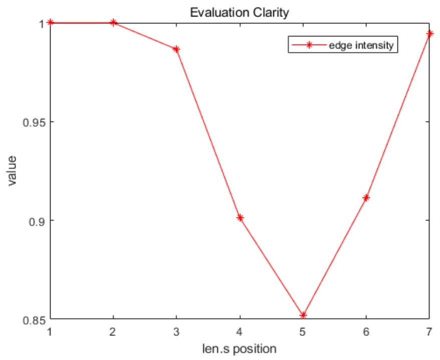


In the end, this paper uses a joint-denoise processing method for the image denoising, which can absorb the advantages of various denoising methods to obtain clear images more effectively. The principle of joint-denoising is to be processed by the expansive and corroded method at the first stage. Then, the wiener filter is used for processing after the mean filtering. Finally, the final gesture image of electronic information teaching demonstration is obtained by marginalization (Figure 9) .

Figure 9.
Joint denoising gesture image.

This article uses edge strength to evaluate the images after each denoising method, and the final image is as shown in Figure 10.

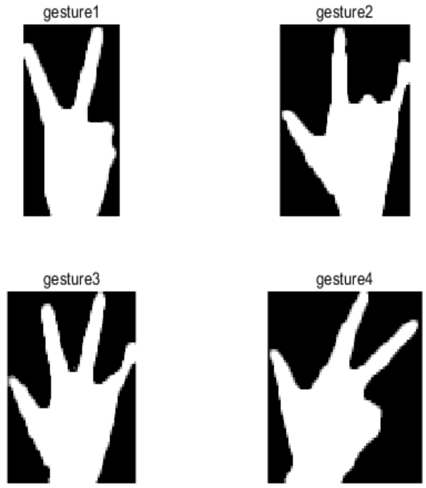
Figure 10.
Edge strength evaluation value of different methods.



The 1-7 abscissa of the image are the original image, bilateral filtering, median filtering, mean filtering, Wiener filtering, morphological filtering and combined de-noising. The ordinate is the value of evaluation. After the threshold processing, a clear depth image can be obtained with less noise interference, so the boundary intensity is the highest. However, the main problem is that there are a lot of irregular burrs on the hand edge. Because the nuclear of the bilateral filter is generated by the function of spatial domain nuclear and range of nuclear, the bilateral filtering is able to protect the edge while smoothly denoising. Although the edge strength is 1, the edge is still similar to the original image due to its protection, which fails to achieve the purpose of removing burrs.

Compared with many denoising methods, the better images can be obtained by joint-denoising method and obviously this mothed can remove the interference and the boundary burrs. On the other hand, in order to capture the dynamic demonstration gesture of electronic information teaching, this paper makes some dynamic changes to the gesture. And the following images are shown in Figure 11.

Figure 11.
Gesture dynamic images.



DISCUSSION AND CONCLUSIONS

In this paper, we use KinectV2 to capture electronic information to teach dynamic demonstration gestures and use depth images for gesture recognition.. The purpose is to avoid the interference of light and color to the image. The overall algorithm is based on matrix scan for dynamic gesture recognition. First, it captures the gesture to remove other interference environments. And then it divides the threshold gesture to remove the interference of the arm and other environmental factors. Finally, the image of gesture recognition is obtained. Because the gesture recognition images have some disadvantages such as noise, boundary burring, etc. In this case, this paper discusses the effects of bilateral filtering, median filtering, mean filtering, Wiener filtering and morphological filtering on image denoising, but none of them is satisfactory. Therefore, this paper proposes a joint-denoising processing method for the image processing, which uses the dilatation-corrosion method for the first step of processing, then uses wiener filtering for processing after mean filtering, and finally gets the final image after de-marginalization. The result of this method is clear image, effective denoising and boundary burr, which makes the boundary smoother. In order to realize the dynamic recognition, this paper proposes a method of dynamic gesture recognition based on the data comparison of gesture matrix. In the circumstance that the recognition rate and speed are guaranteed, the algorithm is still very simple.

AUTHOR CONTRIBUTIONS

The programmed algorithm was conceived and developed by Hongyu Yu. The design and analysis of experiments was done by Hongyu Yu and Yuefeng Li. The manuscript was written by Yuefeng Li and Hongyu Yu. The work presented in the paper was conceived within research projects designed and led by Yuefeng Li. All authors have read and agreed to the published version of the manuscript.

FUNDING

This research was supported by a grant from the Undergraduate Innovation and Entrepreneurship Training Programs funded by the Heilongjiang Province (No. 201910214065) [Design of Bionic Manipulator Based on Fluidic Muscle and Fuzzy PID]

REFERENCES

1. Su B, Wu H, Sheng M, et al. Accurate hierarchical human actions recognition from kinect skeleton data[J]. IEEE Access, 2019, 7: 52532-52541.
2. Neto L B, Grijalva F, Maike V R M L, et al. A Kinect-based wearable face recognition system to aid visually impaired users[J]. IEEE Transactions on Human-Machine Systems, 2016, 47(1): 52-64.
3. Ding J, Shi J Y. Kinect microphone array-based speech and speaker recognition for the exhibition control of humanoid robots[J]. Computers & Electrical Engineering, 2017, 62: 719-729.
4. Krisandria K N, Dewantara B S B, Pramadihanto D. HOG-based Hand Gesture Recognition Using Kinect[C]//2019 International Electronics Symposium (IES). IEEE, 2019: 254-259.
5. Alonso D G, Teyseyre A, Berdun L, et al. A Deep Learning Approach for Hybrid Hand Gesture Recognition[C]//Mexican International Conference on Artificial Intelligence. Springer, Cham, 2019: 87-99.
6. Guo H, Yang Y, Cai H. Exploiting LSTM-RNNs and 3D Skeleton Features for Hand Gesture Recognition[C]//2019 WRC Symposium on Advanced Robotics and Automation (WRC SARA). IEEE, 2019: 322-327.
7. Kane L, Khanna P. Depth matrix and adaptive Bayes classifier based dynamic hand gesture recognition[J]. Pattern Recognition Letters, 2019, 120: 24-30.
8. Terven J R, Córdova-Esparza D M. Kin2. A Kinect 2 toolbox for MATLAB[J]. Science of Computer Programming, 2016, 130: 97-106.
9. Paris S, Kornprobst P, Tumblin J, et al. Bilateral filtering: Theory and applications[M]. Now Publishers Inc, 2009.
10. Cao G, Zhao Y, Ni R, et al. Forensic detection of median filtering in digital images[C]//2010 IEEE International Conference on Multimedia and Expo. IEEE, 2010: 89-94.
11. Coupé P, Yger P, Prima S, et al. An optimized blockwise nonlocal means denoising filter for 3-D magnetic resonance images[J]. IEEE transactions on medical imaging, 2008, 27(4): 425-441.
12. Jin F, Fieguth P, Winger L, et al. Adaptive Wiener filtering of noisy images and image sequences[C]//Proceedings 2003 International Conference on Image Processing (Cat. No. 03CH37429). IEEE, 2003, 3: III-349.
13. Yu-Qian Z, Wei-Hua G, Zhen-Cheng C, et al. Medical images edge detection based on mathematical morphology[C]//2005 IEEE engineering in medicine and biology 27th annual conference. IEEE, 2006: 6492-6495.
14. Lim, Jae S., Two-Dimensional Signal and Image Processing, Englewood Cliffs, NJ, Prentice Hall, 1990, p. 548, equations 9.26, 9.27, and 9.29. Lim J S. Two-dimensional signal and image processing[J]. ph, 1990.
15. Lin Lin, Wang Lizhen. Image segmentation based on edge strength[J]. Journal of Yunnan University Natural Science Edition, 1996: 53-56.
16. Liu H, Deng Z, Xu C. Learning spatiotemporal features with 3D DenseNet and attention for gesture recognition. The International Journal of Electrical Engineering & Education. December 2020.
17. Teju V, Bhavana D. An efficient object detection using OFSA for thermal imaging. The International Journal of Electrical Engineering & Education. August 2020.