Dynamic gesture recognition based on deep information feature fusion

Pan Ou, Qingfeng Yu*, Moran Chen
School of Instrument Science and Optoelectronic Engineering, Beijing University of Aeronautics and Astronautics, Beijing, 100191, China.
*Corresponding author e-mail: 739960607@qq.com

Abstract. Gesture recognition, as one of the mature fields in the field of human-computer interaction, determines whether it can be applied in the field of contactless remote control. This paper focuses on the evolution of low level deep features to high level high recognition ability, but also has a concise descriptor, and from the depth map to improve the recognition rate of dynamic gesture recognition.

1. Dynamic gesture development analysis based on depth information
Along with the intelligent breakthrough, gesture recognition application carrier has been transformed from color image to depth sensor in recent years. The depth sensor solves the problem of illumination condition, complex background and lack of information in 3d space. Thus, depth map replaces color image in gesture recognition. Depth map is the grayscale of pixels, which only represents the space distance from the object's surface to its camera. Therefore, the spatial information of objects can be clearly depicted, and the lack of three-dimensional spatial information can be completed. Therefore, the depth map can be effectively extracted from the illumination and background for dynamic gesture recognition. In its gesture recognition, two points should be considered. Firstly, it is necessary to find the undefined areas of depth missing in the depth map to ensure the continuity of depth time and space. Secondly, depth maps provide surface information of objects and can provide rich geometric features, which can be recorded as descriptors of engraved scenes. Kinect depth camera can effectively recognize dynamic gestures based on the above two points, and the recognition process is developed as follows, as shown in figure 1.
2. Application of dynamic gesture recognition algorithm based on depth information

2.1. Algorithm application in the recognition process of Kinect depth camera
Kinect depth camera can provide more convincing feature vectors for motion recognition in the process of dynamic gesture recognition. For the change of depth, the following formula is adopted for calculation:

\[
\Delta x = \frac{D(x+1,y) - D(x-1,y)}{2} \\
\Delta y = \frac{D(x,y+1) - D(x,y-1)}{2}
\]

(1)

\(\Delta x\) and \(\Delta y\) represents the amplitude and direction of the change of depth value.

2.2. Overall feature extraction algorithm based on depth map sequence
In the process of dynamic gesture recognition, all its characteristic data are the most convincing. In the calculator all feature data, will be applied to the overall feature extraction algorithm to calculate. In the process of evolution, its overall characteristics involve the application and calculation of various algorithms. For a clearer and more intuitive presentation, please refer to the following table 1 for details.

Table 1. Gesture recognition algorithm of depth graph involves image features

| Category          | Characteristic          |
|-------------------|-------------------------|
| Local characteristics | STIP (Space time interest point) |
|                   | Dollar features         |
|                   | 3d BoP features         |
|                   | Characteristics of bone |

Figure 1. Application process of Kinect depth camera in dynamic gesture recognition
Spatial global characteristic

| HDD (Depth difference histogram) |
|----------------------------------|
| HONV (Histogram feature of normal vector direction) |
| ROP (Random occupancy mode) |

Global characteristics of time-space domain

| DMM (Depth motion mapping feature) |
|-----------------------------------|
| HON4D (Histogram characteristics of four-dimensional normal vector) |

Along with the development of now, in research depth map sequence integrity in vector histogram recognition algorithm is applied to the four dimensional method, is characterized by the gradient histogram features complement to expand further, will use depth map in the process of dynamic gesture recognition is applied to the depth of the information, spatial coordinate information and temporal information fusion into four dimensions of the normal vector. The depth graph sequence of a gesture is denoted as \( I_1, I_2, \ldots, I_n \), The depth value is expressed as \( z = f(x, y, t) \), The plane in its four-dimensional space is expressed as:

\[
S(x, y, z, t) = f(x, y, t) - z = 0
\]

If the normal vector on the surface, it can be expressed as:

\[
\mathbf{n} = \nabla S = \left( \frac{\partial z}{\partial x}, \frac{\partial z}{\partial y}, \frac{\partial z}{\partial t}, -1 \right)
\]

3. Application of gesture decomposition in dynamic gesture recognition

Nowadays, there are some problems in gesture decomposition in dynamic gesture recognition distance. The problem is: First, gesture sequences have redundant information in time domain, which has limitations in the thinking normal vector histogram algorithm and is not suitable for full application. Second, the same gesture in the depth map has the difference of speed transformation, which will lead to the recognition error in the algorithm. Thirdly, for its undefined large area depth information, the gesture area will be denoted as 0 in its algorithm, and this area has no computational significance at this time. In view of the above problems, the decomposition algorithm of gesture sequence is carried out to discard the disadvantages brought by difference and redundant information. The recognition feature is extracted from the gesture again to remove the undefined large area depth information and add markers, so as to calculate the four-dimensional normal vector histogram algorithm. The algorithm is shown in figure 2.
4. Experimental comparison results
MSRGesture3D gesture database can conduct experiments on gesture decomposition sequences to identify dynamic gestures. For dynamic gesture recognition, five experimental tables are sampled, and their gestures are processed into gesture sequences. Its training set and test set are denoted as $C_{10}^5 = 253$, In this set, the recognition rate of 253 different cases is represented, and the details are shown in table 2 below. The experimental methods in the database are the same, and the number represents the percentage of the recognition rate, and the dark color represents the increase of the recognition rate, while the light color represents the small recognition rate.
| Recognition rate | 1.50 | 2.50 | 3.50 |
|-----------------|------|------|------|
| 0.80            | 87.63| 86.42| 84.29|
| 0.82            | 87.71| 87.34| 86.05|
| 0.84            | 88.63| 87.74| 87.16|
| 0.86            | 89.23| 88.31| 87.43|
| 0.88            | 88.84| 88.55| 88.56|
| 0.90            | 88.41| 89.61| 89.52|
| 0.92            | 88.32| 89.23| 90.50|
| 0.94            | 88.32| 88.56| 89.49|
| 0.96            | 88.32| 88.41| 88.70|
| 0.98            | 88.32| 88.32| 88.70|
| 1.00            | 88.32| 88.32| 88.32|

As can be seen from the above table, the similarity between sequences has been limited in time domain. When the sequence approaches 1.00, it is closest to the original data.

5. Conclusion
Dynamic gesture recognition technology can be widely used in human-computer interaction and outstanding advantages, compared with the traditional algorithm based on the color map, based on the depth of information dynamic gesture recognition background factors such as light, can be ignored, the three dimensional space further promoted to the four-dimensional space to improve performance, the algorithm is improved to some extent in 4 d algorithm, so as to realize good user experience.

References
[1] xiao, han jing,zhang yuelong. li Gesture recognition based on geometric distribution characteristics of gestures [J]. Computer science,2019,46(S1):246-249+262.
[2] yan, Qiang linjia, dong yuan-juan, zhaofast gesture recognition method based on stack sparse self-coding multi-feature fusion [J]. Journal of Beijing institute of technology, 2019, 39(06):638-643.
[3] zhengwei,Zhu lei, zhu peng rao . Research on gesture recognition technology based on bs-hmm and bar-type distance [J]. Computer application and software,2019,36(06):163-166+253.
[4] longjiao,Zhang xiaoqin. zeng SEMG gesture recognition based on deep neural network [J/OL]. Computer engineering and application :1-12[2019-07-31].
[5] Yuqing Peng xiao-song zhao, TaoHuiFang, Liu Xianzi, tie-jun li. Under complicated background based on the deep learning of gesture recognition [J/OL]. Robot: 1-9 [2019-07-31]. https://doi.org/10.13973/j.cnki.robot.180568.
[6] guangyuFeng, wenjun, hou hu,zhou you zhenwei. Volley gesture recognition based on chaos dynamics [J/OL]. Computer integrated manufacturing system :1-12[2019-07-31].