Dynamic gesture recognition and human-computer interaction

Jiali Zhang[1, a], Guixi Liu[1, b]

[1]School of Mechano-Electronic Engineering, Xidian University, Xi'an, 710071, China

[a]email: jlzhang1314@163.com, [b]email: gxliu@xidian.edu.cn

Keywords: Dynamic Gesture Recognition; Feature Fusion; Human-Computer Interaction (HCI).

Abstract. In this paper, the recognition algorithm of dynamic gestures, which based on depth information, is proposed. A new human-computer interaction controlled by dynamic gestures is also achieved. The proposed algorithm extracts Histogram of Oriented Gradients (HOG) and weighted Hu invariant moments from RGB and depth map of the interested region of gestures respectively, and fuses weighted Hu invariant moments into HOG. Then, we establish Support Vector Machine model (SVM) for static hand gestures recognition. Base on this, an HMM toolbox supporting multi-dimensional continuous data input is adopted to do the training and dynamic gesture recognition. Finally, the visual simulation software Vega Prime is applied to simulate a virtual UAV flight controlled by dynamic gestures. Experimental results demonstrate that the proposed algorithm is effective and accurate in complex background and practical real-time applications.

Introduction

Recently, with the rapid development of computer vision technology, the growing popularity of the human-computer interaction has led to a revolution in gesture recognition, and Human-computer interaction based on dynamic gesture recognition is also becoming more popular. In this paper, the recognition algorithm of dynamic gestures, which based on depth information, is proposed. The algorithm mainly consists of feature extraction, fusion and HMM classifier. Finally, the three-dimensional modeling software Creator is applied to build 3D models and visual simulation software Vega Prime to simulate, achieving a visual simulation of flight system controlled by dynamic gestures based on the proposed algorithm.

Dynamic gesture recognition

In this paper, two features combined with the SVM classifier is applied to identify the different types of static gestures, and the dynamic gesture recognition is achieved by HMM training. The two features are HOG and weighted Hu invariant moments extracted from RGB and depth map of the interested region of gestures.

Gesture feature extraction and fusion. Unlike other image geometric features, HOG does not consider the image characteristics from a whole, but subdivides image into several small cell units, and then calculates the gradient or edge direction histogram of each pixel in all cells. To improve performance, a number of cells making up a block, thus the image becomes a connected graph consisting of many blocks, then normalizes the gradient of all cells in these blocks and obtains the final direction of gradient vector. A brief flowchart of extracting the HOG is given in Fig.1.

Input image → Calculate the gradient of the horizontal and vertical direction → The gradient direction histogram statistics within the Cell → The gradient direction histogram statistics within the Block → HOG feature vector

Fig.1 A brief flowchart of extracting the HOG

The basic idea in Ref. [1] is used in this paper, and makes improvements according to gestures itself characteristics. As shown in Fig. 2, considering the differences mainly in the finger part, RGB image of gesture obtained through kinect is divided into five regions, this system only extracts HOG feature in (2), (3), (4) area shown in Fig. 2.

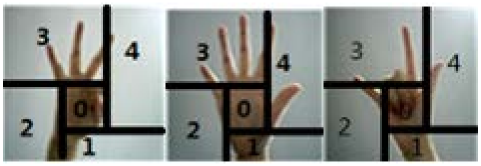


Fig.2 Gesture area images

In order to get the overall shape and contour characteristic, the Hu invariant moments of the binary image is selected as a gesture feature parameter in this page. Hu invariant moment is a good solution to solve the scale changes, image translation changes, rotation of the coordinate transformation in the process. Due to different Hu moment's contribution to the behavior criterion value is different in the feature extraction, so should take full account of the contribution values of different moment to construct a weighted Hu moment. In this paper, nine extended Hu moments are modified in linear programming to find the optimal weights for each Hu moment, and then construct a weighted modified Hu moments.

Set a new weighted fixed Hu structure as:

$$\Phi = \sigma_1\Phi_1^{''} + \sigma_2\Phi_2^{''} + \cdots + \sigma_9\Phi_9^{''}, \sum_{i=1}^{9}\sigma_i = 1, 0 < \sigma_i < 1 \tag{1}$$

The specific solving process of the weighted Hu moments is presented in Ref. [2].

In order to better describe the gesture feature, this page refers to Ref. [3,4], and propose fusing HOG features and weighted Hu moment feature to describe gesture image. Before the feature fusion, HOG vectors F_{HOG} and weighted Hu moments vector F_{Hu} are normalized to $F_{HOG}^{'}$ and $F_{Hu}^{'}$ with the largest-minimum principle. After normalization, we fuse two kinds of features with corresponding weights represented as F in

$$F = [w_1 F_{HOG}^{'}, w_2 F_{Hu}^{'}], \qquad w_1 + w_2 = 1. \tag{2}$$

A whole process of the proposed algorithm is shown in Fig. 3.

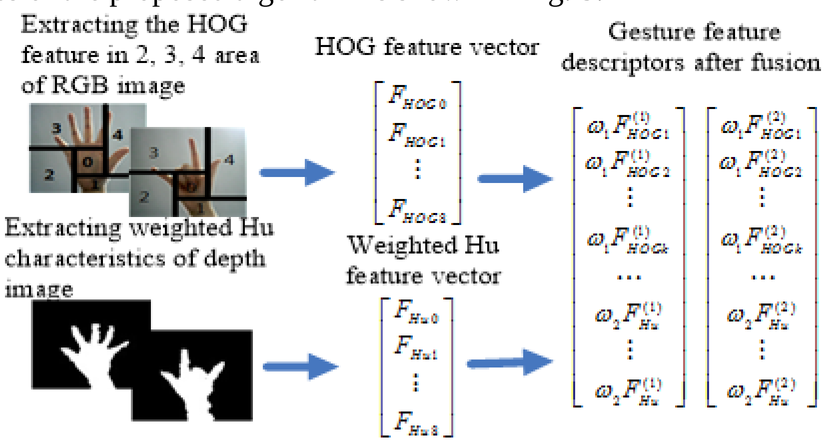


Fig.3 The overall process of feature fusion

Gesture recognition and tracking. Feature after fusion is regarded as the input of SVM. SMO (Sequential Minimal Optimiz-ation) algorithm is selected for training in SVM, and Gaussian kernel function for inner product operations. Then the results of SVM training regarded as the HMM state vector , and select the center of gravity of adjacent palm direction vector as an HMM 's output observation variable, to train and recognize dynamic gestures.

Human-Computer Interaction

With the advanced human-computer interaction technology proposal and development, human-computer interaction using gesture recognition technology has become a hot research topic. In this paper, three-dimensional modeling software Creator builds all entity object models and large terrain model required by visual simulation of flight system. Visual simulation software Vega prime organizes 3D visual database. In the visual simulation of flight system, dynamic gestures are used to control UAV to pitch, turn and roll.

Three-dimensional modeling software Creator use OpenFlight data format to build physical models and large terrain models through texture mapping, detail -level technology. This paper builds the UAV and terrain shown in Fig. 4.

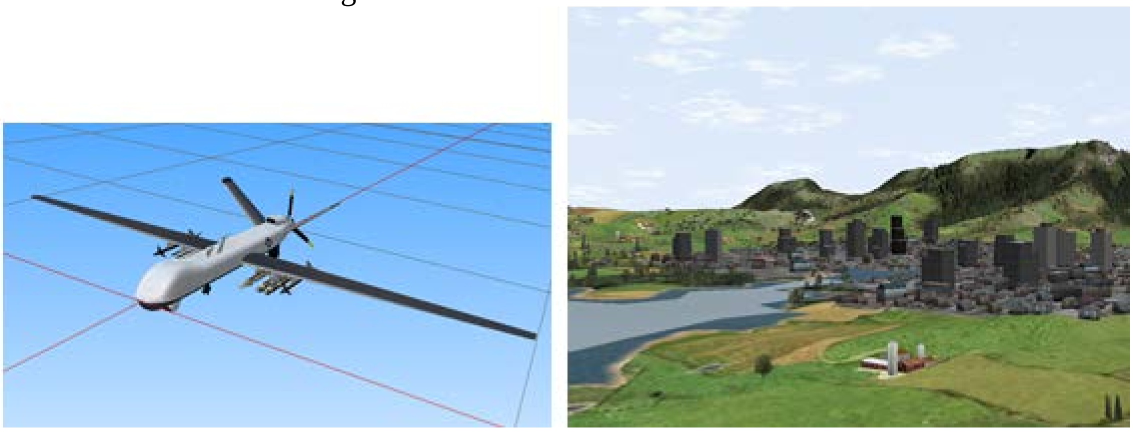


Fig.4 UAV and the terrain model

Physical models and terrain model built previously are added to Vega Prime, and configure the initial state of the scene. Taking into account that dynamic gestures control UAV flight, the program provides dynamic gesture response interface. The overall framework of the system is shown in Fig. 5.

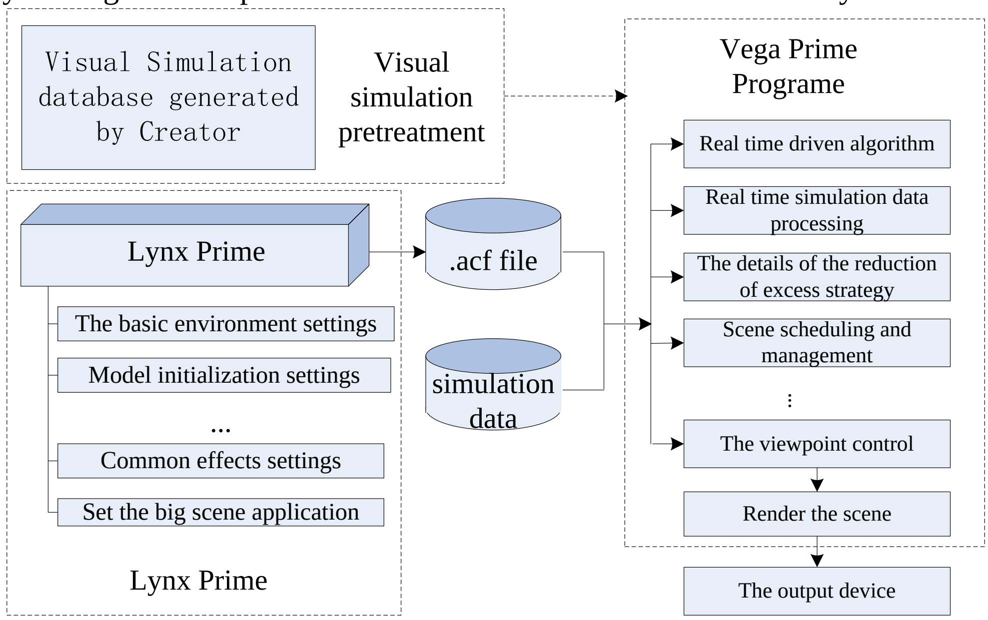


Fig.5 framework of the system

Experimental results and evaluation

Experiment 1. To evaluate the validity of the proposed algorithm, this paper chooses 6 kinds of dynamic hand gestures for SVM and HMM training as shown in Fig. 6.

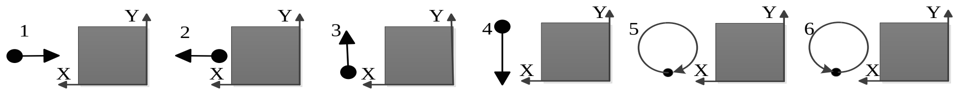
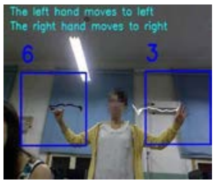


Fig.6 The trajectory training model

The recognition performance of the algorithm in this paper is tested. The results of dynamic gesture recognition are shown in Fig. 7. It can be seen from the Fig. 7 that the algorithm can recognize a variety of different gestures accurately when skin objects, multi-object moving and multiple signal interference exists in the background. It can not only recognize out hand shape or trajectory of single hand, but also the combination of hands shape and trajectory in complex environment.

(a) Single hand trajectory

(b) Two hands trajectory



(c) Combination hand shape and trajectory

Fig.7 The recognition results of hand shape and trajectory

Experiment 2. In this paper, Vega Prime software simulates a UAV flight simulation environment, and receives different dynamic gesture input to control the UAV in flight. In the simulation, set two channels, one channel displays UAV flight, and the other channel displays dynamic gestures. Simulation results show that the accuracy of dynamic gesture recognition is very high and response time is less than 0.5s, meeting real-time requirements. Fig. 8 shows the response of dynamic gestures.

Fig.8 UAV flight controlled by dynamic gestures

Conclusions

In this paper, a dynamic gestures recognition algorithm fusing multiple features is proposed. The algorithm can overcome the influence of complex background. We achieve a visual simulation of flight system controlled by dynamic gestures with the help of Creator and Vega Prime visual simulation software. This Human-Computer Interaction controlled by dynamic gestures can be used in robot control, UAV operator station, somatosensory games and so on.

Acknowledgement

This work was partially supported by National Ministries Research of Twelfth Five projects under grant #Y31011040315, National Ministries Foundation of China under grant # Y42013040181, Fundamental Research Funds for the Central Universities under grant # NSIY191414.

References

[1] Feng K, Yuan F. Static hand gesture recognition based on HOG characters and support vector machines[C].Instrumentation and Measurement, Sensor Network and Automation (IMSNA), 2013 2nd International Symposium on. IEEE, 2013: 936-938.

[2] Trigueiros P, Ribeiro F, Reis L P. Vision-Based Portuguese Sign Language Recognition System. New Perspectives in Information Systems and Technologies, Volume 1. Springer International Publishing, 2014: 605-617

[3] Yun L, Lifeng Z, Shujun Z. A hand gesture recognition method based on multi-feature fusion and template matching[J]. Procedia Engineering, 2012, 29: 1678-1684

[4] Liu K, Chen C, Jafari R, et al. Fusion of Inertial and Depth Sensor Data for Robust Hand Gesture Recognition[J]. 2014