The Study on Virtual Input Method Based on Machine Vision

Xu Yi

Department of Information Engineering，Hunan Industry Polytechnic，Changsha，Hunan

12936394@163.com

Keywords: Virtual Reality; Machine Vision; Image Processing; Gesture Recognition; Template Matching

Abstract. Currently, virtual reality technology has been increasingly used in different industries. In the virtual and reality environment, the human-computer interaction device is challenged by expensive cost and restrained using environment. With the development of computer technology, using computer vision technology to solve the problem of human-computer interaction in virtual reality has become a thought used for the study the man-machine interface. This article uses the Gaussian model based on skin color in combination with improved optical flow field tracking algorithm, realizing the real-time rapid gesture tracking under complicated situations. The method can ensure high algorithm execution efficiency and detection accuracy while ensuring high robustness.

Introduction

In the environment of the current virtual reality, the man-machine interaction is mainly divided into three categories. The first category, interaction based on data glove. Using the sensor which is connected to the gloves to convert finger movements to electrical signal so that the gestures can be determined; the relative position of the finger is determined by the additional sensor that is usually connected to magnetic or acoustic related sensors attached to data glove. Using the data glove and location tracker have high recognition rate of in measuring the trajectory of gestures in space motion y and temporal information, but the system requires users to wear complex data gloves and position tracker, causing some inconvenience to users, and currently the input device is expensive, posing difficulties upon massive promoting; The second category is mouse, keyboard, touch screen based interaction. The hand movement outside the devices is deemed as input, which is non-contact type. The third class: the man-machine interface based on vision. It has the advantage of non-contact without gloves, fitting the “nature” requirements of human-computer interaction. The fast computing power nowadays has made it possible to achieve real-time visual calculation, thus becoming the main direction of current research. There have been many scholars at home and abroad that have conducted study on this topic from different perspectives and levels. Applying machine vision system to man-machine interface that imitates that of information accepted by human can realize direct and natural human computer interaction mode.

Realization of Matching Algorithm under Complicated Scene

There are large different conditions of two-dimensional projection images under different occasions, this is mainly caused by the following reasons: the sensor noise, the image change caused by the change of perspective during imaging, target movement and deformation, light or

image changes resulting from the change of environment and the use of a variety of sensors, etc. In order to solve the match difficulty caused by above figure, people put forward many matching algorithm that all composed of the followings:

(1) The feature space: the feature space is composed of attend matching image features. Choosing good features can improve the matching performance, reduce the search space, and reduce the influence of noise and other uncertainties that exert on matching algorithm. Matching process can use the global or local characteristics as well as the combination of both.

(2)Similarity measurement; the similarity measurement refers to what measurement is used to determine the similarity between the characteristics that needs to be determined; it is usually defined as a cost function or in the form of distance functions. The classical similarity measurement includes correlation function and Minkowski distance; people proposed Hausdorff distance as the matching measures in recent years.

(3) The type of image matching change the, image geometric transformation is used to solve the geometric position of the difference between the two images, including the rigid body transformation, affine transformation, projection transformation, polynomial transform etc. Here we use affine transformation and set a single pixel as the image coordinates, the coordinates of the after affine transformation is:

$$\begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} t_x \\ t_y \end{pmatrix} + s \times \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$$

Whereas, t_x and t_y is the amount of translation of the x and y direction, θ is the rotation Angle, s is the scale factor. By solving the four parameters can determine an affine transformation.

(4) the search of transformation parameters , the search strategy is to use a suitable search method to figure out optima estimation in the search space, such as translation, rotation transformation, making the similarity of images become largest after transformation. The search strategy includes end search, hierarchical search, simulated annealing algorithm. The genetic algorithm uses the traversal optimization search strategy, which can guarantee the optimization of search results with global optimality, whose the required amount of calculation is far less than the traversal type of search.

Due to the fixed template image, the template matching algorithm has poor self-adaption in the target motion state and the illumination change; it may get lost when there are great changes of target motion. In order to solve the problem, after the characteristics of sequence images and the target shape feature and gray characteristic differences between background are analyzed, an improved algorithm is proposed. The algorithm based on optimization and training determines the grayscale characteristics of the target, and then search target according to the characteristics of gray level to, the image sequences of the following several characteristics are mainly studied :

Gray average $m_1 = \frac{1}{JK}\sum_{x=1}^{j}\sum_{y=1}^{K} f(x,y)$, whereas, J and K are the dimension of target sample.

$$\sigma = \sqrt{\frac{1}{JK}\sum_{x=1}^{J}\sum_{y=1}^{K}(f(x,y)-m_1)^2}$$

Gray mean square error $\sigma = \sqrt{\frac{1}{JK}\sum_{x=1}^{J}\sum_{y=1}^{K}(f(x,y)-m_1)^2}$

Gray histogram. The same texture primitive features same histogram, but the same histogram may have different texture primitives, so other characteristics shall be added when at conducting comparison and analysis using histogram.

The torque characteristic is one of THE widely used shapes in pattern recognition. Some of the most basic two-dimensional shapes are directly associated with moments. $r(x,y)$ represents the distance from pixel (x,y) and the target center pixel (x_0,y_0), namely:

$$r(x,y)=\sqrt{(x-x_0)^2-(y-y_0)^2}$$

Define $M(x_0,y_0)=\sum f(x,y)r(x,y)$, considering the practical characteristics of moving target, these characteristics can keep invariance in the respect of target size, direction, translation, rotation and image background, contrast, brightness change; using them to identify the target has good robustness and adaptability to change of the target motion.

Design and Realization of Algorithm

(1) Collect and analyze the statistical analysis; select a certain number of target images as the training sample and calculate the average grayscale value m_1 and variance σ, the histogram and characteristic quantity M.

(2)Judge the image to see if there are images to be determined. If it does not exist, then the identification is not required, If any, then continue with the processing.

According to the target size and the certain search subgraph search order. Using template matching method, when the neutron image differencing image pixel is less than a certain threshold, then the graph in the original image is deemed as a background to proceed with a new search, or calculate the grayscale average value. If the absolute value of difference between grayscale average grayscale value and the target difference between the absolute value is greater than a certain threshold, then search again, otherwise, the following step should be continued.

Calculate the gray variance of subgraph. Similarly, if the absolute value of difference between gray variance and the target gray is greater than a certain threshold, then search again, otherwise, the following step should be continued.

Calculate the grayscale histogram and feature vector M of subgraph. Compare grayscale average value and variance, histogram and characteristic quantity M of subgraph with target samples, and endow different weights to find the overall optimal point of whole image.

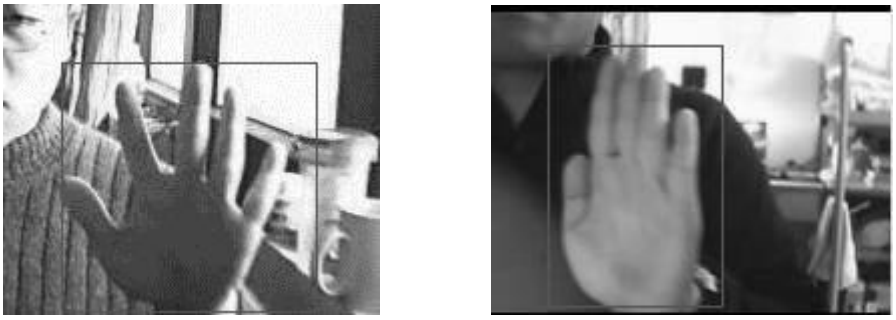
The experiment and the experimental results of image matching

The template image is shown in figure 1; we define two kinds of gesture templates, of which one is positive palm open, with the merging of positive hands.

（a） (b)

Fig.1 Template of Two Gestures

When conducting the experiment, the experiment is conducted both to identification of static images and capturing of dynamic videos, the results are shown in figure 2:



(a) (b)

Fig.2 Template Matching Experimental Result

The Realization of Gesture Recognition Based on Template Matching

We use the method based on template matching to extract the gestures and analyze the specific meanings that they express through the picture that each frame pass through the above processes. The algorithm based on shape matching is mainly aimed at small areas of interest to build a template; it also creates templates of the whole images, but this must be that the object accounts for large proportion of the whole image, such as the upper body during video conference; this is object tracking experiment aimed at the whole picture; it is often sacrifice of matching speed. The basic flow is so: the HALCON is used in combination with Visual C++.Net to develop, with the process as follows:

(1) First determine the rectangular area of the ROI, this only require determination of the coordinates of upper left and right points , and using gen_rectangle () function will help you to generate a rectangle and using area_center () to find the center of the rectangle.

(2)Then the image in the rectangle image has to be captured from the image, reduce_domain () will get the ROI; After that, this rectangular can be built with templates; while before the template is established, such area can be first processed for convenience of modeling , such as threshold segmentation, mathematical morphology.

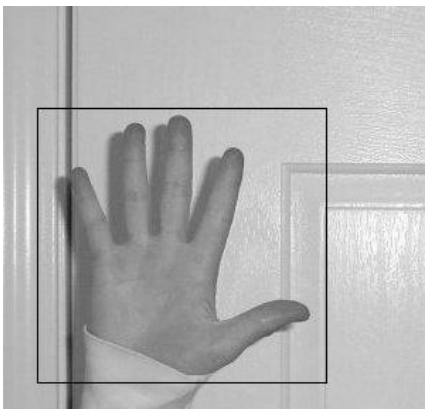
(3) The the template can be created through create_shape_model, such function has many parameters, among which the pyramids series is designated by the Numlevels, the higher the value, the less time fining the object is. The AngleStart and AngleExtent determine possible rotation range, AngleStep specifies Angle range search step length; Attention: in any case, the template should fit

main memory, which can shorten the search time. For particularly large template, using Optimization reduce the number of template points is very useful; MinConstrast extract templates w from the noise of the image, if the gray value range is 10, then MinConstrast shall be set 10; the Metric parameters determines template recognition conditions, if 'use_polarity' is set, the image of the object and template must have the same contrast; after template is created, the templates shall also be monitored, this can be done by inspect_shape_model () to complete, it checks the applicability of the parameters and can also help to find the right parameters; In addition, the outline of this template is also required, which can be used in the following match; get_shape_model_contours () will be helpful for us to find the outline of the template.

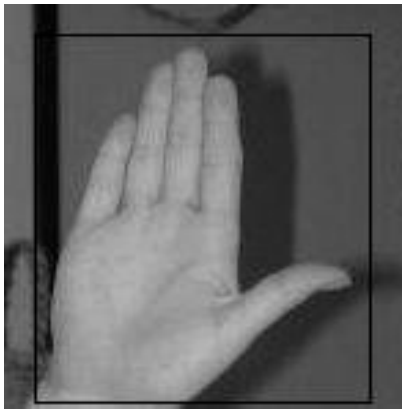
(4) Once the template is established, ther template can be matched by means of the image captured from video straming.

(5)After finding it out, it still needs to be converted to show the two functions vector_angle_to_rigid () and affine_trans_contour_xld () which can play big role here, of which the former one calculates rigid body affine transformation from one point and perspective, this function is very useful to construction of one affine transformation from the function matching result, the reference image is changed to the current image.

The experimental result is shown below:



(a) Gesture a Identification result



(a) Gesture b Identification result

Figure 3. Gesture Recognition Result

Summary

The gesture recognition technology can date back to the 1990 s; it is not practically applied until the end of last century. The virtual reality human-machine interface based on machine vision is an example of gesture recognition practicality; it applies a gesture recognition module to the specific virtual reality environment .This paper focuses on the theory of the gesture recognition and take a specific virtual scene as a demonstration window, and move forward the model identification method to a brand-new level.

Reference

[1] N. K. H. Tang,J. F. Pickering,J. M. Freeman. An Investigation into the Control of Brushed Yarn Properties: The Application of Machine Vision and Knowledge-based Systems. Part II: The Machine Vision System. Journal of the Textile Institute,1993,842:.

[2] Anonymous. Keyence offers machine vision system and machine vision software. Military & Aerospace Electronics,2009,206:.

[3] Anonymous. Machine Vision; Research from Jiangsu University yields new data on machine vision. Computers, Networks & Communications,2010,:.

[4] Anonymous. Machine Vision; Studies from University of Stellenbosch further understanding of machine vision. Computers, Networks & Communications,2010,:.

[5] Anonymous. Machine Vision; Studies from University of Alaska yield new information about machine vision. Technology & Business Journal,2010,:.