Digitalized Modeling of Human Hand through Contour Analysis in Hand Gesture Recognition

Wenzhen Yuan,
Dept. of Mechanical Engineering
Tsinghua University
Beijing 100084, China
yuan-wz08@mails.tsinghua.edu.cn

Wenzeng Zhang
Dept. of Mechanical Engineering
Tsinghua University
Beijing 100084, China
wenzeng@tsinghua.edu.cn

Abstract—This paper introduced a method of building and fixing a digitalized model of human hand gesture through contour analysis of the hand’s orthographic projection. Via contour analysis the fingers and palm on the projection are located and their quantized characters are calculated. The model is used in a gesture recognition method, where the articulation is of core concern. Thus the method is adaptive to different subjects with no require of pre-built gesture library or training, and its high efficiency enable it to work in real-time condition. The primal object of the recognition is to control robot hands.

Keywords—hand gesture recognition; contour analysis; human-computer interaction; control of robot hands

I. INTRODUCTION

Hand gesture recognition has drawn great attention as a promising human-computer interaction (HCI) method below[1]. It is a natural and convenient way for human to express and the information it conveys is abundant. Hand gesture recognition is expected to greatly facilitate processes such as machine manipulation, sign language interpretation and computer inputting.

Currently the dominant hand gesture recognition methods are through template matching[2][3]. A gesture library should be built through learning before doing the recognition, and during recognition the model of gesture is matched with the models in the library. Accordingly only the gestures in the library could be recognized. Additionally the library models may not adapt to different people, so the subject is restricted, and the application of gesture recognition as a HCI method is restrained. Some researchers tried to build the real model of the gesture based on shape and kinetic information[2], thus the recognition is not restrained to the subject or the gesture library. Mostly the realization is based on multi-camera system, and the efficiency is not satisfactory.

This paper aims at building a digitalized model of hand gesture based on the hand’s shape information, and monocular camera system is taken as the input. The kinetic characters of human hand[4] is applied in the modeling in order to estimate information of higher dimension from the low dimension input. The model is expected to show no difference between different subjects. Similar attempt has been made by Wu’s group[5]. The focus of this paper is analyzing the gestures feature for model building through contour analysis. Some measures of fixing the model through contour analysis when the input is not favorable are also introduced in the paper. Contour analysis is a commonly used method in shape analysis[6], and some gesture modeling methods such as that introduced in [7] also use contour as an important basis in modeling.

The primal object of the recognition introduced in this paper is to control robot hands. The control of robotic hand is a tough problem as the mechanism has massive degrees of freedom (DOF). Attempts of control by gesture recognizing have been made such as in [8], but refined control of each finger is hardly tried. This recognition project aims at control the separate motion of each finger on a robot hand by making it following the manipulator’s gesture. Experiments proved the method successfully allowed different manipulators to control a robot hand, and real-time recognition and control can be achieved. The recognition method is also promising in other HCI areas with its convenience, ample information available, little restriction on operator or environment and high efficiency.

II. THE DIGITALIZED MODEL FOR GESTURE

The hand-gesture recognition remained a tough issue because a human hand has so many DOFs, thus a gesture has so many possible forms. But only part of the information conveyed by a gesture is meaningful according to the recognition aim. So the first step to scheme a recognition algorithm is to select the target characters. As the primal object of the recognition introduced here is to control robot hand, the states of each finger joint and the position of the hand are the core objective. This paper focuses on finger state recognition in particular.

The input of the system is an image of subject’s hand gesture. As a requirement the hand is orthogonal to the camera, and the fingers are approximately apart. To estimate the 3-D information of a finger from 2-D input image, the knowledge of hand motion characteristics is applied. As also mentioned in [2] and [5], the motion of finger joints are constrained, and when conducting normal tasks in natural ways the joints on a finger bent in a coupled way. So giving the bending degree of a finger, the states of joints on the finger can be estimated. For control a robotic hand, this is enough; and the information is also enough to convey other demands in many other HCI areas.

In this project, the bending degree of a finger is estimate from the length of the finger’s orthographic projection on the camera view. The overview of the method is stated in Fig. 1.
In this paper the contour analysis method is proposed to segment the hand projection and get the finger skeleton. It proved to be efficient and reliable. The calculated finger length is compared with the full length of the finger-the length of the finger when it erects fully, and the shorter the current length is, the more the finger is bending. The identification of the finger is matched by the angle of the finger’s root to the palm’s center. An initialization frame is required in the project, where the subject is asked to pose his hand in the center of the camera’s view with five fingers spread. Then the full length and root angle of each finger, as well as the subject’s hand color, are recorded. These features vary between different people, and the initialization procedure help the model to fit different subjects.

III. BUILDING THE DIGITALIZED MODEL

A. To get the contour of the hand

The method of segmenting gesture from a relative complicated background in this paper is introduced in [9]. After the process, the orthographic projection of the current gesture is got, as well as the edge. The texture information of the gesture is elided in this paper although it contains abundant information. Then the points on the contour are saved in series according to clockwise order.

B. Getting the incircle of the gesture projection

A most distinct character of the points on the finger contour is that the distance between the points and the center of the palm is relatively longer. In this paper the palm center could not be defined as the centroid of the whole projection, in which case the character may be not so distinct. Instead, the incircle center is chosen as the base point in the following contour calculating steps, since experiments shows that the incircle center best defines the center of the contour.

Another reason for finding the incircle of the projection is to make the lengths of fingers into dimensionless lengths. The quantitative characters of the projection, the fingers’ lengths for example, vary due to not only the difference in gesture, but also the distance between the hand and the camera. These characters should be transformed into dimensionless forms, which require a factor to carry out the dimensionless process. The incircle radius is a proper factor and is adopt in this analysis. Another possible symbol is the square root of the palm’s projection area. Both the symbols are relatively stable among different gestures compared to other symbols. TABLE I. compares the two features.

<table>
<thead>
<tr>
<th>Image got from camera</th>
<th>Incircle of the hand projection</th>
<th>Palm part of the hand projection</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>incircle radius</td>
<td>√(Palm Area/π)</td>
</tr>
<tr>
<td></td>
<td>82.2</td>
<td>85.8</td>
</tr>
<tr>
<td></td>
<td>93.5</td>
<td>94.2</td>
</tr>
</tbody>
</table>

The dimension information can also help evaluating the distance from the hand to the camera. Similar practice on calculating the depth information is introduced in [11]. Fig. 2 compares the incircle radius and the square root of palm area in different situation. In Fig. 2 (a), the hand gesture remains but the distance of the subject’s hand moves evenly towards or backward to the camera; in Fig.2 (b), the hand remains approximately static but the gesture changes randomly. The comparison shows both features could indicate the depth information and projection dimension while the incircle radius is steadier.

![Incircle radius and palm area’s square root in experiments.](image)
C. Finding the fingers through histogram analysis

The two characters of contour used in the analysis are the distance and angle derivative, with the center of incircle as the base point. The two characters have also been used in [10] and [12] to indicate the geometrical features of shapes. In gesture projection analysis, the contour points of the fingers are likely to have large distance and small angle derivative. In Fig.3, peaks are signs of the fingers, while the steep slopes on their sides indicate the contour of the fingers. So by finding the peak points and valley points of big slopes on the histogram we can find most finger tips and roots on the contour.

![Image of histograms and contours](image)

The angle derivative histogram represents the derivative of angle of the line between the contour point and the circle center. Theoretically, in palm contour the angles of the points change at a relative steady and fast speed; in finger area, as the contour is longer, the derivative is smaller or even changes sign. But in practice it is not a distinct character to define fingers, so it is taken as an auxiliary character in defining fingers. The angle histogram is essential in defining the thumb, especially when it is of certain angles, because the thumb may not be as distinct on the distance histogram due to its special position. On the angle derivative histogram the thumb contour is very distinct and easy to recognize.

![Figure 3](image)  
**Figure 3.** The histograms of different gestures. Column 1 are original images got from the camera; Column 2 are the projections abstracted from the original images; Column 3 are the contour’s distance histograms calculated from the projection; Column 4 are the contour’s angle derivative histogram. The finger tips on the contour are marked with green circle, and the finger root points on the contour are marked with red circle.

D. Fixing the segmented fingers

After the fingers’ tips and roots on the contour are found, the two root points on the finger contour can separate the finger’s projection from the palm. By calculating the distance from the tip point and the root points, we get the general length of the finger’s projection. But in some cases the separation is not accurate while deviations appear, and the calculated length may not be exact. So some fixing measure should be taken in order to get better result.

The first common problem is that the finger’s root point may not locate on the contour of the projection, but in the inland area of the projection. This situation is likely to occur when two fingers are so near together that parts of their projection are cohered. The compensating measure is to move the calculated root point nearer to the palm center. The expected distance between root points and the palm center can be recorded in the initialization frame, and the distance of other root points in the same frame can be reference too. As far as the distance of a specific root point is considered to be too large, this root point is adjusted to a new position with the distance shortened but the azimuth angle remains. Experiments shows that this method can adjust the found root point much nearer to the real finger root point, though not exactly the right place, and the calculated finger length is much closer to the real situation. Fig.4 shows the adjusting result.

![Image of root adjustment](image)

**Figure 4.** The result of root adjustment. (a) is the original image, where the middle finger and the ring finger is too close. (b) is the shade and the separating result after steps in III.C, (c) is the adjusting result.

In other cases, two fingers are too near together and would be considered as one on the projection. Theoretically the problem could be solved by finding the border line between the two fingers, but other interference factors may be introduced. The solution applied in this project is to separate the possible fused fingers by analyzing the contour and projection shape. Longer finger root length is the sign of fused fingers. These finger root lines are divided medially to figure out separate fingers. To locate the tip points of the two fingers, the distance histogram of contour is used. Although on the histogram the division of the two fingers is not apparent or even indistinguishable, the tip points are easy to identify, as they usually appear as local peaks, or at least the derivatives undergo sharp reduction. By finding points with these characters can the two tip points be found, and thus the fingers could be located respectively. Fig. 5 gives an example of the separation of combined fingers, where the two finger tips are apparent peaks on the contour’s distance histogram thus could be easily found.
The bending degree of a finger can be estimate according to its length on the projection. Dimensionless length is adopted to reduce the influence of the depth. The dimension base used here is the projection’s incircle radius. A finger’s dimensionless length is calculated as follows:

\[ \text{length}_r[i] = \frac{\text{length}[i]}{R} \]  

Here the \( R \) refers to the incircle radius. Then the length is compared with the dimensionless length of the finger recorded in the initialization frame. A ratio is calculated:

\[ \text{ratio}[i] = \frac{\text{length}_r[i]}{\text{length}_{o,r}[i]} \]  

The smaller the ratio is, the greater the finger \( i \) is bending. When the finger retracts, it would not appear on the projection contour, and its length, and ratio is 0. In completed experiments the finger states is divided into four stages to matching the bending degree of the finger.

The character chosen to identify the fingers is the angle between the finger’s root and the incircle center and the side line of the palm. It is a relatively reliable and adaptive character and is easily applied.

### IV. EXPERIMENT RESULTS

The system is tested on a personal laptop computer (dual kernel CPU, 2.20GHz) with an image capture system of a CCD camera and an image grabbing card. The recognition program is written in C++ on Microsoft Visual Studio 2008 platform. The average frequency of image acquisition system on the calculating platform is 24.8 fps. When doing recognition the frequency remains generally steady but little variance existed when recognizing different gestures. The average frequency of recognizing random gestures from complex background is around 19.7 fps.

Experiments shows the recognition system can successfully extract states of tester’s fingers from a monocular camera system when the subject poses his hand directly to the camera and the spread fingers are not closed together. However when the input gesture is not favorable enough the program can adjust automatically. Experiments on different operators have been conducted and succeeded.

### V. CONCLUSION

This paper introduced a method of building a digitalized model of human hand-gesture through hand’s contour analysis. By analyzing the contour of the gesture’s projection in monocular camera system the recognition method figures out the states of each finger. Some amending procedure is also introduced to help the recognition fit unfavorable situation. The recognition method proves efficient, convenient and adaptive. The initial intent for building the recognition system is to manipulate robot hand, while it has great potential to be applied in other human-machine interaction fields too.

### ACKNOWLEDGMENT

This paper was funded by NSF of China (No. 50905093) and Foundation of National Research Innovation for University Students (No. 101000316).

### REFERENCES


