Available online at www.sciencedirect.com
ScienceDirect
Neurocomputing 69 (2006) 2017–2025
NEUROCOMPUTING
www.elsevier.com/locate/neucom

Adaptive multiple sets of CSS features for hand posture recognition

Chin-Chen Chang*

Department of Computer Science and Information Engineering, National United University, Miaoli 36003, Taiwan, ROC

Received 1 April 2005; received in revised form 7 October 2005; accepted 4 November 2005
Available online 10 February 2006
Communicated by D. Barber

Abstract

In this paper, an adaptive feature extraction approach based on curvature scale space (CSS) is presented for translation, scale, and rotation invariant recognition of hand postures. First, hands are segmented from hand posture images into binary silhouettes and then binary hand contours are computed. CSS images are then used to represent the contours of hand postures. In particular, adaptive multiple sets of CSS features are extracted to address the problem of deep concavities in the contours of hand postures. Finally, 1-nearest neighbor techniques are used to perform adaptive multiple sets of CSS feature matching for hand posture identification. Results indicate that the proposed approach performs well in the recognition of hand postures. And, the proposed approach is more accurate than previous methods which were based on conventional features. The proposed technique could be useful in improving the recognition of hand postures.
© 2006 Elsevier B.V. All rights reserved.

Keywords: Gesture recognition; Hand posture recognition; Curvature scale space; Feature extraction; Feature matching

1. Introduction

Due to the mass use of computers, the use of gestures in human–computer interaction has been extensively researched in recent years [2,3,6–9,12,13,17,18,20,22,23]. There are two major techniques for analyzing gestures: (a) model-based methods; (b) appearance-based methods. Model-based approaches are based on the 3D spatial description of hands, and appearance-based methods are based on the appearance of hands in images. Generally, gestures can be classified into two types: (a) hand postures; (b) dynamic gestures. 'Hand postures' are described in terms of hand shapes; 'dynamic gestures' refer to certain hand movements. Since hand postures have certain meanings and act as transitions for dynamic gestures, recognizing hand postures is one of the most important aspects of gesture recognition. However, since there are large variations in hand postures, reliable descriptors for characterizing hand postures are critical.

Schlenzig et al. [18] proposed an appearance-based hand gesture interpretation using recursive estimation based on moments [10,11,21] in hand postures. However, the recognition performance would be affected by unstable centroids of the hand shape. Su et al. [20] presented a static hand gesture recognition system based on the ten fingers' flex angles using a composite neural network. However, the control flowchart is too complicated. In addition, a training phase is needed to get the classification rules. Banarse and Duller [2] developed a three-stage self-organizing neural network architecture to perform static gesture recognition and used 2D plane cells as features of hand postures. However, the cost of extracting entire plane cells would be too high. Huang and Huang [7] presented a sign language system and used Fourier descriptors [24] to characterize hand postures. Wu and Huang [23] presented a view-dependent hand posture recognition approach to achieve natural interaction in virtual environments and used Fourier descriptors to represent hand postures. However, although direct extension from 1D to 2D FFT is easy, it would be inaccurate for translating invariant shape description. Lee and Chung [12] proposed an algorithm that extracts features to recognize sign language

*Corresponding author. Tel.: +886 37 381893; fax: +886 37 354326.
E-mail address: ccchang@nuu.edu.tw.

0925-2312/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.neucom.2005.11.002

based on orientation histograms of hand postures. However, the features of the same hand shapes would be affected by rotation. Triesch and von der Marlsburg [22] presented a hand posture recognition system based on elastic graph matching against complex backgrounds. However, the proper node positions in a large number of images have to be edited by hand for constructing the model graphs.

Although these studies have provided many techniques for extracting hand posture features, there are still numerous problems that have not yet been explored. In this paper, an adaptive feature extraction approach is presented for translation, scale, and rotation invariant recognition of hand postures based on the curvature scale space (CSS) [1,15,16]. The CSS image is a multi-scale organization of the zero-crossing curvature points of the closed planar curve as it is evolved. All sets of peaks in the CSS image can be calculated to form a set of coordinate peaks. However, since the human hand is highly changeable, the location of the largest peak in the CSS image will be unstable for the same hand postures. Thus, the conventional CSS features normalized by the highest peak are unreliable and will affect recognition. Therefore, adaptive multiple sets of CSS features are extracted in the proposed approach to address the problem of deep concavities in contours of hand postures. Results show that the proposed approach performs well for recognizing hand postures, and the proposed algorithm is more accurate than previous methods based on conventional features. The proposed technique may significantly improve the recognition of hand postures.

The rest of this paper is organized as follows: In Section 2, related works are briefly discussed. The proposed approach is presented in Section 3. Results are given in Section 4. Analysis and discussion are carried out in Section 5. Finally, conclusions are presented in Section 6.

2. Related works

Mokhtarian and Mackworth [15,16] first proposed the object shape descriptor based on the CSS image of the contour. CSS descriptors are utilized for obtaining a quantitative shape feature description of the hand for recognizing hand postures.

The curvature κ of a planar curve is defined as the derivative of the tangent angle ϕ with respect to the arc length s. The curvature κ is written as follows:

$$\kappa = \frac{d\phi}{ds}. \tag{1}$$

Let Γ be a planar curve defined by

$$\Gamma = \{(x(u), y(u)) | u \in [0, 1]\}, \tag{2}$$

where u is the normalized arc length parameter. Next, a convolution operation is performed with respect to closed curve Γ and a one-dimensional Gauss function $g(u, \sigma)$ for obtaining a smoothed curve defined by

$$\Gamma_\sigma = \{X(u, \sigma), Y(u, \sigma)\}, \tag{3}$$

where σ is the standard deviation; $X(u, \sigma)$ and $Y(u, \sigma)$ are defined as

$$X(u, \sigma) = x(u) * g(u, \sigma), \tag{4}$$

and

$$Y(u, \sigma) = y(u) * g(u, \sigma), \tag{5}$$

respectively; $g(u, \sigma)$ is defined by

$$g(u, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(\frac{-u^2}{2\sigma^2}\right). \tag{6}$$

Generally, a smooth curve has a higher resolution when σ is relatively small, but a lower resolution when σ is relatively large. Hence, the contour drawn by $\Gamma_\sigma = \{X(u, \sigma), Y(u, \sigma)\}$ becomes smoother as σ increases. Different standard deviations σ are used to find a location having zero curvature in $\Gamma_\sigma = \{X(u, \sigma), Y(u, \sigma)\}$ and thus all locations that have zero curvature are drawn under different standard deviations σ. As a result, a CSS image of input hand contour is obtained. That is, the CSS image is a multi-scale organization of the zero-crossing curvature points of a closed planar curve as it is evolved.

After the CSS image is obtained, for each peak in the CSS images, if the value of the peak is larger than threshold T_p, then the peak is detected to form a set of coordinate peaks. Let T_p be defined by

$$T_p = M_p / Thres, \tag{7}$$

where M_p is the largest peak value of the CSS image and *Thres* is a parameter used to adjust threshold T_p.

Then a coordinate with a largest peak in a coordinate-peak set formed by the CSS image is selected as a basis point for alignment. A circular rotation is performed to generate an aligned CSS image according to the basis point for determining feature parameters of the hand posture. The feature parameters of the hand posture are then compared with feature parameters of reference hand postures, thereby determining a hand shape corresponding to the hand posture image.

For a hand posture, the locations of the largest peaks in the CSS image correspond to the deep concavities in the original hand contour. However, CSS images represent the same hand shape, no matter where the largest peak occurs. The hand posture recognizer may produce different results because of different largest peaks. Fig. 1 shows that (a) and (b) are the same hand posture Five; (c) and (d) are the CSS images of (a) and (b), respectively. For (a) and (c), the largest peak of the CSS image of the hand posture occurs at the recessed part between the thumb and index finger. For (b) and (d), the largest peak of the CSS image of the hand posture occurs at the recessed part between the index and middle fingers.

Download English Version:

https://daneshyari.com/en/article/410981

Download Persian Version:

https://daneshyari.com/article/410981

Daneshyari.com