

A New Distance Metric Learning Algorithm for Hand Posture Recognition

Liu Qiongli^{1, a}, Xu Dajun^{1, b}, Li Zhiguo^{2, c}, Zhou Peng^{1, d}, Zhou Jingjing^{1, d}
and Xu Yongxia^{2, c}

¹ Wuhan Mechanical Technology College, Wuhan, Hubei, China

² Taiyuan Satellite Launch Center, Taiyuan, Shanxi, China

^a416087321@qq.com, ^b83950280@qq.com, ^cdugufei_2000@163.com,

^dqiuchuanfei@hotmail.com

Keywords: Distance metric learning. Nearest classify. Hand posture recognition.

Abstract. A new distance metric learning algorithm for hand posture recognition is proposed in this paper. By adopting the Fourier descriptor of the hand shape as the sample feature, the algorithm exerts new constraints on the distance of inter-class pairs, and learns a Mahalanobis distance metric matrix by minimizing the distance of intra-class pairs and maximizing the distance of inter-class pairs simultaneously. Then k-nearest neighbor classifier is used to test the introduced method. The experimental results on multi-class hand posture recognition by k-nearest neighbor classifier show that the presented algorithm has less error rate than other distance metric learning algorithms.

Introduction

With the development of computer technology, new means of interactive study in line with interpersonal communication habits has become one of the major trends in the development of human-computer interaction. These studies include face recognition, gesture recognition, and hand posture recognition. Among them, as the most common way of communicating, gestures of daily life have the advantages of easy operation, high efficiency and others. Compared with other motion capture mode (such as data gloves), gesture recognition technology based on the computer vision is of convenient operation, interactive non-intrusive, etc., with broad prospects in the application. However, due to the influence of the background and illumination changes, vision-based gesture recognition approach technically still faces many problems to be solved.

For gesture recognition problem under complex background, reference[1] proposes to use Fourier descriptors to express the gesture characteristics, to identify the test specimen based on Euclidean distance. Reference[2] also proposes to use Fourier descriptors to represent gesture characteristics, to recognize gesture based on BP neural network. Reference [3] are also related to gesture recognition technology research. All these methods usually classify gesture based on Euclidean distance in hand shape feature space. However, due to the influence by viewpoint changes to human hand, the same kind of static gesture will generate different hand shapes in viewpoint change, causing small-scale changes in gesture characteristics to affect the accuracy of the classification results. Utilizing the chosen distance metric matrix to evaluate the distance between samples could make the same kind samples gather more, heterogeneous sample disperse more, improving the accuracy of classification. The goal distance metric learning is to learn a good distance metric matrix, making the same kind of samples as close as possible, heterogeneous ones as far as possible.

K-nearest neighbor classifier (KNN) is a commonly used one, its idea is to find K samples which is nearest to test samples from the training samples, classifying test samples by using the most likely categories in K samples. To reduce the error rate of KNN classifier to gesture classification, the paper proposes to learn a Mahalanobis distance matrices based on the distance metric learning, using it to calculate the distance between samples, and gives an improved local oversight distance metrics learning Algorithm. Reference[4] reviews various distance metric learning algorithms, classifies and compares a variety of learning algorithms. While, if similarity evaluation is increased, it will further affect the accuracy of KNN classification. Different from the existing algorithms, the algorithm proposed in this paper added a priori information of heterogeneous gesture difference to the constraint optimization to

get improved, and through the comparison with other algorithms to verify that the improved algorithm can increase the accuracy of classification.

Hand Posture Recognition System Based on Distance Learning and KNN

Fig. 1 is a hand posture recognition system based on distance learning and KNN classification. In the application the system consists of two phases: (1) Mahalanobis distance learning matrix; (2) hand posture. The main objective of Mahalanobis distance matrix learning phase is to learn a good matrix through the sample. In hand posture recognition phase, first the hand shape features(Fourier descriptors) will be extracted from the image, and then find K nearest neighbor samples from the training sample based on the Mahalanobis distance, using KNN classifier to operate gesture recognition.

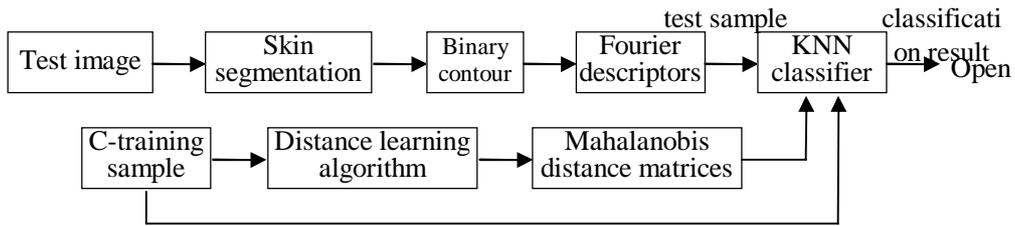


Fig. 1 Hand posture recognition system based on KNN classifier

Local Distance Metric Learning Algorithm

Suitable distance measure has a significant impact on the performance of KNN classifiers. The most commonly used distance measures are mainly the Euclidean distance and Mahalanobis distance. Suppose that $\mathbf{x}_i, \mathbf{x}_j$ are samples in the dataset, the Euclidean distance between them can be expressed as:

$$d(\mathbf{x}_i, \mathbf{x}_j) = \sqrt{(\mathbf{x}_i - \mathbf{x}_j)(\mathbf{x}_i - \mathbf{x}_j)^T} \quad (1)$$

)

The Mahalanobis distance between them can be expressed as:

$$d^M(\mathbf{x}_i, \mathbf{x}_j) = \sqrt{(\mathbf{x}_i - \mathbf{x}_j)M(\mathbf{x}_i - \mathbf{x}_j)^T} \quad ($$

2)

Among which, M is a symmetric half definite matrix, called distance matrix. When M is a unit matrix, the Mahalanobis will degenerate into the Euclidean distance. The main objective of distance matrix learning is to learn an appropriate one, so that the distance between similar samples as close as possible, the distance between heterogeneous sample as large as possible.

Problems Description. Given n samples of training sample set and their corresponding category set, wherein sample, is the dimension of the samples, is the number of categories. Let represent homogeneous sample couple collection, and represent heterogeneous sample couple collection. The definition of distance between samples is Mahalanobis distance, learning goal of distance metric is to learn a Mahalanobis distance metric matrix, so that the distance between heterogeneous samples will be as close as possible, the distance between heterogeneous samples will be as far as possible.

Hand Posture Fourier Descriptors. Hand shape Fourier descriptors is got through Fourier transform of hand shapes to get Fourier coefficients, which will be calculated by normalization after. Using Fourier descriptors to show hand shapes features there are advantage of remaining unchanged in rotation, translation and scaling, that is to say, Fourier descriptors does not vary with hand shape in the above changes, and not sensitive to noise, be is with good robust as well. Here is the usage of one-dimensional Fourier descriptors to describe the hand shape.

Referring to Fig. 2, let C represent hand shape, $r(t)$ represents the radial distance from the sampling point Pt of t on C to Pc, the center of C , using the equal arc distances to perform N sampling on C , and then the discrete Fourier transform of shape C radial distance can be expressed as:

$$a_n = \frac{1}{N} \sum_{t=0}^{N-1} r(t) \exp\left(\frac{-j2\pi n t}{N}\right) \quad (3)$$

)

Wherein, $n = 0, 1, \dots, N-1$, shape C Fourier descriptors can be expressed as:

$$b_n = \left| \frac{a_n}{a_1} \right|, n = 1, 2, \dots, N-1 \quad ($$

4)

represented by the collection as: $\{b_n, 0 < n < N\}$

Improved Distance Metric Learning Algorithm. Ref [5] studied distance metric learning algorithm which applies to KNN classifier, the basic idea of these methods is to find local K training samples that related to testing samples, and in accordance with the correlation strength of the testing samples to give different weights between K training samples. Ref [5] proposed using local linear discriminant analysis to achieve neighbor classification, determining the direction of a partial LDA according to the discriminant analysis results of the query points. Ref [6] proposed a method of classifying a neighbor maximum edge (LMNN) similar to SVM, through distance metric learning to learn a Mahalanobis distance matrix, so that after the change of distance, this sample will gather more with similar samples in K neighborhood component, while disperse more with heterogeneous samples in K neighborhood component. This paper applies this algorithm to realize the hand posture recognition and improvement of the algorithm. In the optimization of the objective function class difference priori information is embedded to restrict. Let $X = \{x_1, x_2, \dots, x_n\}$ as the integration of n training samples, $x_i \in R^d$ as single sample, using d Fourier descriptors to demonstrate hand posture, $y_i \in \{1, 2, \dots, J\}$, J Is the category number, z_{ij} ranges from 0 or 1, 1 represents sample x_i and x_j are the same kind, 0 represents they are heterogeneous samples. The value of h_{ij} ranges from 0 or 1, indicating whether the sample is K neighborhood component of x_j , let $D(\mathbf{x}_i, \mathbf{x}_j) = (\mathbf{x}_i - \mathbf{x}_j)^T M (\mathbf{x}_i - \mathbf{x}_j)$ as the square of Mahalanobis distance, $M \in R^{d \times d}$ as Mahalanobis distance matrix, for the distance metric learning objectives, this problem can be described as an optimization problem.

Experiment Result

The proposed algorithm was validated on Matlab7.0. First select Logitech cameras to capture 8 hand postures("Sign Language 8", "fist", "grip 1", "grip 2", "open", "pinch take", "pointing", "confident V") from images on VC2003.NET platform, then, based on the OpenCV library to process and extract sample characteristics from images, and finally, verify the distance metric learning algorithm on Matlab7.0 proposed in this paper, in the experiment, select 30 Fourier descriptors to represent hand-shaped samples, each type of gesture were collected 100 samples, p % samples are selected from each type randomly as training samples, and the remaining (100-p)% samples as test samples, 3 is taken as K values of KNN classifier. The results shown below, Figure 3 is a hand-shaped profile of "sign language eight" samples, Fig. 4 is using 30 Fourier descriptors to reconstruct the hand shape in Figure 3. Figure 5-6 uses the features of the first three to visualize the distribution of distance

metric learning samples before and after, by comparison, it can be seen that after distance metric learning the distribution of the similar samples gathered more and the heterogeneous samples are more dispersed.

Summary

For many types of hand posture recognition problem, the paper proposes a measure based on the combination of distance metric learning algorithm and KNN classifier, and gives an improved LMNN algorithm to learn the distance between samples. In the experiment, hand shapes in the image are extracted through color segmentation, using Fourier descriptors as samples (hand shape features). Mahalanobis distance matrix is calculated by utilizing the collected multi-class hand posture samples to optimize the cost function. Through comparison with other local distance metric learning algorithm, the one proposed in this paper can reduce the error rate of hand posture classification.

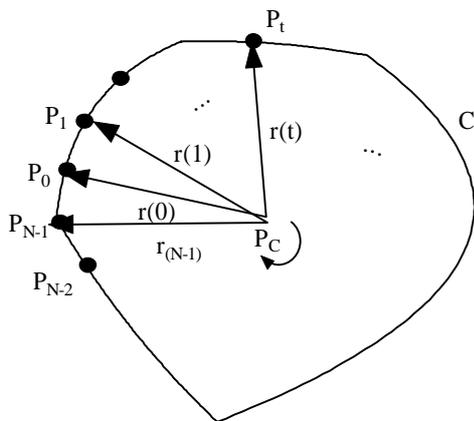


Fig.2 The radial distance sample from the hand shape

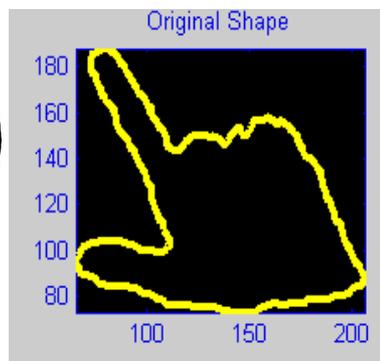


Fig. 3 an "eight" type hand posture originally

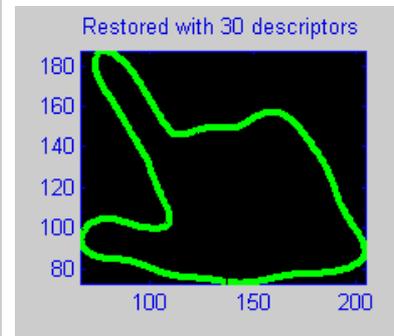


Fig. 4 using 30 Fourier descriptors to reconstruct the hand shape

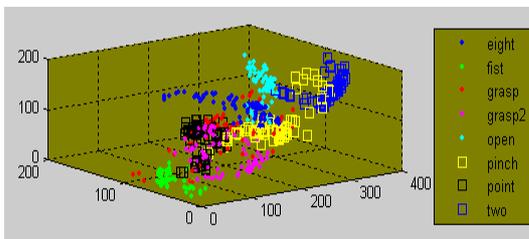


Fig. 5 using Fourier descriptors (first three) to show three-dimensional distribution of various types of gesture data before distance metric learning

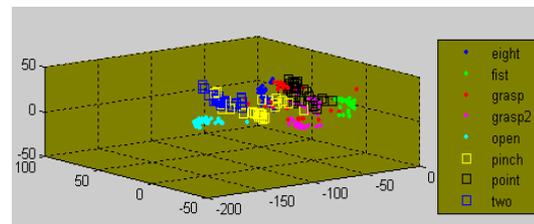


Fig. 6 various types of gesture distribution after distance metric learning

References

- [1] Zhang Shaoli. Fourier Descriptor hand posture recognition method based on Vision [J]. Software Guide. 2008(06).
- [2] Liu Yan, Teng Xiaolong, Liu Chongqing. Hand Posture Recognition Under Complex Background Based on Fourier Descriptors [J]. Computer Simulation. 2005(12).
- [3] Yuan Xiangrong, Tian Xincheng. Gesture Recognition Based on Neural Networks [J]. Shandong Jiaotong University Journal. 2006(02).
- [4] L. Yang and R. J. and. Distance metric learning: A comprehensive survey. Technical report, Michigan State University. http://www.cse.msu.edu/~yangliu1/frame_survey_v2.pdf.
- [5] T. Hastie and R. Tibshirani, "Discriminant adaptive nearest neighbor classification," IEEE Pattern Analysis and Machine Intelligence, vol. 18, no. 6, 1996.

[6] Weinberger, K., Blitzer, J., & Saul, L, “Distance metric learning for large margin nearest neighbor classification,” In Y. Weiss, B. Schölkopf and J. Platt (Eds.), Advances in neural information processing systems ,2006.