Identification of Body Types Using KFDA in Non-Contacted Body Measurement System

Yuxiu Wang[1], Xiaojiu Li[1], Hao liu[2], Xiaozhi Li[1]

[1] *School of Art and Clothing, Tianjin Polytechnic University, Tianjin, China*

[2] *School of textile, Tianjin Polytechnic University, Tianjin, China*

Email: *yuxiu_wang@yahoo.com.cn*

Abstract: Fast measurement and automatic classification of body types are critical parts in the automatic garment manufacture and made-to-measure (MTM). The calculating accuracy can be upgraded by using the clustering function to classify the body types and the clustering method is presented for classifying the body types without prior knowledge. This paper is focused on using the algorithm of kernel Fisher discriminant analysis (KFDA) to recognize automatically body type. The F-statistic is utilized for determining the optimal class number. KFDA, backpropagation neural network (BPNN) and radial basis function neural network (RBFNN) are compared in the Iris dataset and performed in the body dataset. Experimental results exhibit the KFDA is optimal method for identifying the body data. The result shows that calculating accuracy can be improved by using the discriminant analysis.

Keywords: kernel fisher discriminant; classification; discriminant analysis; body type

1. Introduction

It is an important part to classify and identify the body types objectively in the process of developing non-contacted body measurement system. Because of physical properties of measuring device, stability of lights, and influence from measurement environment, the system may be instability to cause the lack of body data, the distortion and the data error. Establishing the perfect virtual body database would improve measurement accuracy and upgrade robust of the whole system. The body database is built objectively based on data classification according to a certain age and region range, rather than simply according to "chest-waist difference". In this paper we classify the existing body data and built predicting models, then use the result of clustering analysis as the training samples to establish discriminant function of the new samples. Thus, when the system scans the new sample, the new body data can be predicted, judged and modified.

2. Discriminant Analysis of Body Type

2.1. Kernel-Fisher Criterion

Suppose a training set $X=\{x_1,x_2,\cdots,x_n\}$, each sample has p index values. m is the classification number of the samples. $l_i(i=1,\cdots,m)$ is the number of the ith classification. The training set X is standardized by $x_i=(x_i^0-\overline{x}^0)/\delta$, where x_i^0, $\overline{x}^0$ and δ are the initial data, mean value and standard deviation of the ith classification. Input data space R^p can be mapped to feature space by nonlinear mapping function:

$$\phi: R^p \to F, x \mapsto \phi(x) \tag{1}$$

The choice of kernel function determines the mapping process from input space to feature space. The feature space is a high-dimensional space which may even be infinite-dimensional. The function of Fisher criterion in feature space can be defined as follows:

$$J(v)=\frac{v^T S_b v}{v^T S_t v} \tag{2}$$

In (2), S_b and S_t are discrete matrixes between classifications and within classification, v is a vector in feature space. They can be described as follows:

$$S_b=\frac{1}{n}\sum_{i=1}^{m} l_i(\overline{\phi}_i-\overline{\phi}_0)(\overline{\phi}-\overline{\phi}_0)_i^T \tag{3}$$

$$S_t=\frac{1}{n}\sum_{i=1}^{n}(\phi(x_i)-\overline{\phi}_0)(\phi(x_i)-\overline{\phi}_0)^T \tag{4}$$

Where $\overline{\phi}_i$ is the mean value of the sample of the ith classification and $\overline{\phi}_0$ is mean value of sample set.

2.2. Calculating the Optimal Projection Direction Vector Set of Sample Set in the Feature Space

Choose a suitable kernel function to map the data of input data to feature space F. Calculate the kernel matrix K' in feature space where $k'_{ij}(x_i,x_j)=\phi^T(x_i)\phi(x_j)$. Centralize K' to get matrix K:

$$K=K'-BK'-K'B-BK'B \tag{5}$$

Where matrix B is $(1/n)_{n\times n}$.

In order to maximize (3), the characteristic solution of (6) need to be solved:

$$\lambda S_b v = S_t v \tag{6}$$

978-1-935068-40-2 © 2010 SciRes.

Because feature vector v is the linear combination of the elements in feature F, there will always be a coefficient $\alpha_i (i = 1, 2, \cdots, n)$ to make

$$v = \sum_{i=1}^{n} \alpha_i \phi(x_i) \quad (7)$$

Suppose $W = diag(W_1, W_2, \cdots, W_c)$, where the element $(W_i)_{l_i \times l_i}$ is a square matrix $1/l_i$. Based on W and K, (3) can be converted to (8). So there is no need to solve the element in high-dimension when maximize the criterion function. By calculate the kernel function, the conversion from input space to feature space can be achieved and the best projection vector can be found.

$$J(\alpha) = \frac{\alpha^T KWK\alpha}{\alpha^T KK\alpha} \quad (8)$$

Resolve K for simplifying the calculation. Suppose Λ is the diagonal matrix of $K's$ non-zero eigenvalues and its diagonal elements are arranged by descending order. So Λ^{-1} can be calculated. P is a characteristic vector corresponding Λ and $P \times P^T = I$, where I is identity matrix. Then K can be described as $K = P\Lambda P^T$ and (8) can be turned into (9):

$$J(\alpha) = \frac{(\Lambda P^T \alpha)^T P^T WP(\Lambda P^T \alpha)}{(\Lambda P^T \alpha)^T P^T P(\Lambda P^T \alpha)} \quad (9)$$

Suppose

$$\beta = \Lambda P^T \alpha \quad (10)$$

Then

$$J(\beta) = \frac{\beta^T P^T WP\beta}{\beta^T P^T P\beta} \quad (11)$$

the maximization of (11) can be summed as calculating the eigenvalues of (12).

$$\lambda\beta = P^T WP\beta \quad (12)$$

Calculate the characteristic vector β of matrix $P^T WP$. From (10),for a given β, there is at least one α to satisfy $\alpha = P\Lambda^{-1}\beta$. From (7), the optimal projection direction vector set V can be got about (2).

$$v_i = \alpha_i \Big/ \sqrt{\alpha_i^T K \alpha_i} \quad (i = 1, 2, \cdots, r) \quad (13)$$

Where $r = \min(rank(K), m-1)$ and $rank(K)$ is the rank of K.

2.3. The Establishment of Kernel-Fisher Discriminant Function

(1) obtaining the projection vector of the training sample set in the projection direction

For the selected kernel function, the kernel matrix K'' of training sample set can be got by (14).

$$k''_{ij}(x_i, x_j) = \phi^T(x_i)\phi(x_j) - \frac{1}{n}\sum_{i=1}^{n} k''_{ij}$$
$$(i = 1, 2, \cdots n; j = 1, 2, \cdots, n) \quad (14)$$

The projection vector P of training sample can be got by (15).

$$P = K''V \quad (15)$$

(2) getting the central projection vector of training sample set

P is $n \times r$ matrix and can be described as $P = (p_1, p_2, \cdots, p_m)$, where $p_i = (p_{i1}, p_{i2}, \cdots, p_{il_i})$. The classification of training samples can be got by clustering analysis, probabilistic neural network or self-organizing neural network. The central vector of the training sample set in the projection direction can be calculated by (16).

$$c_i = \sum_{j=1}^{l_i} p_{ij} \quad (16)$$

(3) obtaining the testing samples of projection vector

Suppose testing sample set $T = (t_1, t_2, \cdots, t_q)$, the testing samples can be standardized by mean value $\overline{x}^0$ and standard deviation δ of training samples. Calculate the kernel matrix $k''_{ij}(t_i, x_j) = \phi^T(t_i)\phi(x_j) - \frac{1}{n}\sum_{i=1}^{n} k''_{ij}$ of training samples X and testing samples T of based on the selected kernel function, where $i = 1, 2, \cdots q; j = 1, 2, \cdots, n$. Then obtain the projection vector set $P' = (p'_1, p'_2, \cdots, p'_q)$ of T by (15). The discriminant function group used to class the testing sample t_j is

$$d_i = \| p'_j - c_i \| \quad (i = 1, 2, \cdots, m; j = 1, 2, \cdots, q) \quad (17)$$

Based on the principle of nearest distance, if $d_k = \min(d_i)$ $(k = 1, 2, \cdots, m)$ for t_j, it should be the kth classification.

3. Experiment and Data Validation

In order to test the feasibility and accuracy of the discriminance, the Iris datebase are utilized. The Iris data contains 150 examples with 4 dimensions and 3 classes, the front 4/5 samples of Iris dataset are regarded as training samples, and the rest samples are regarded as testing samples on Iris dataset. We compare the recognition effects of the BPNN, RBFNN and KFDA algorithms. First, analyze the data by BPNN, RBFNN and KFDA. BPNN is composed of three layers: input layer, hidden layer and output layer. The choice of neurons in hidden layer is based on $n_1 = \sqrt{i+o} + a$, where n_1, i, o are the nodes of the hidden layer, input layer and output layer and a is between 1 to 10. The transfer function in middle layer is “tansig”. The transfer function in output layer is “purelin”. The training algorithm is based on the Levenberg-Marquardt algorithm of numerical optimization. Its output difference is 0.05 and the training error is 0.001. The supervision signals of these two neural networks are “0-1” type. The transmission of RBFNN is 1.0. The mapped kernel function in KFDA is $k(x,y) = \exp(-\|x-y\|^2/\sigma)$, where σ is 0.7. The first three quarters Iris data of every classification in Iris is used as training samples and the last one quarter data as testing samples.

The methods of BPNN, RBFNN and KFDA are used to discriminate the data and the discrimination ratios are 93.33%, 86.67% and 96.67% which indicate the method of KFDA much better.

Analyzing the Iris data and measured body data by the methods of KFDA, BPNN and RBFNN, KFDA for discriminating the samples is the best discrimination efficiency. The method of KFDA effectively resolves the dimension problems by mapping method when the data is calculated in high-dimensional feature space and also overcome the deficiency of disriminant function group that Fisher discriminant analysis can not construct linear indivisible in input space. So the method of KFDA to establish the disriminant function of new samples is success.

4. Conclusion

KFDA is a new machine learning algorithm based upon the method that the kernel technology is used in Fisher linear discriminate analysis. By using kernel technology, the deficiency of Fisher linear discriminate analysis is improved and it will make the multicriteria data in the feature space more linearly divisible than in the input space. In this paper, KFDA is used to identify the human body and the result is excellent. Unsupervised fuzzy clustering analysis and KFDA constitute a complete system to classify and identify the human body more objective and accurate and it improves the efficiency of analyzing multicriteria data and reduces the difficult of processing the data. After identifying the human body, the result is used to predict and determine the body size and this improves the stability of the system.

References

[1] Wang Y. X.; Li X. J. & Liu H.: Fuzzy clustering analysis of body type in 2D non-contacted body measurement system, Journal of Textile Research, Vol. 28(2007) No.2, pp.100-103, ISSN 0253—9721.

[2] Li X. J.; Liu H. & Wang Y. X.: Classification and identification of body type in non-contacted body measurement system, *Journal of Tianjin Polytechnic University,* Vol. 26(2007) No.5, pp.33-35, ISSN 1671—024X.

[3] Baudat G. & Anouar F.: Generalized discriminant analysis using a kernel approach, *Neural Computation*, Vol. 12(2000) No.10, pp.2385-2404, ISSN 0089-7667.

[4] Yang J.; Jin Z. & Yang J. Y.: Essence of kernel Fisher discriminant: KPCA plus LDA, *Pattern Recognition*, Vol. 37(2004), pp.2097-2100, ISSN 0031-3203.

[5] Abdallah F. & Richard C.: An improved training algorithm for nonlinear kernel discriminants, *IEEE Transactions on Signal Processing*, Vol. 52(2004) No.10, pp.2798-2806, ISSN 1053-587X.

[6] Liang Z. Z. & Shi P. F.: An efficient and effective method to solve kernel Fisher discriminant analysis, *Neural Computation*, Vol. 61(2004), pp.485-496, ISSN 0899-7667.