一种加权的核 Fisher 鉴别分析在人脸识别中的应用

刘晓亮，王福龙，黄　诚，曾爱华

（广东工业大学 应用数学学院，广东 广州 510006）

摘要：在非线性空间中采用加权的最大散度差鉴别准则函数，该方法不仅有效地抽取了人脸图像的非线性特征，而且在特征空间 H 中，使用权函数重新构造了类间散度矩阵和类内散度矩阵，从而优化了核的最大散度差准则函数. 最后在 ORL 和 Yale 人脸数据库上的实验结果验证了本文方法的有效性.

关键词：人脸识别；Fisher 非线性鉴别分析；最大散度差鉴别准则

中图分类号：TP391　　**文献标识码**：A　　**文章编号**：1007-7162(2009)04-0065-05

近几年来，人脸识别技术得到了很大的发展，许多优秀的方法被提出. 其中，Fisher 线性鉴别分析(Fisher Linear Discriminant Analysis，FLDA)被认为是较好的方法之一[1]. 此方法的基本思想是在 Fisher 准则下，先求解最佳鉴别矢量，然后将高维的样本投影到最佳鉴别矢量张成的空间，使投影后的样本在低维空间有最大类间距离和最小类内距离，这样在低维空间中样本将有最佳的可分性，然而传统的 Fisher 鉴别分析是通过对样本进行线性变换，得到的是线性鉴别特征. 考虑到人脸模式常常受光照、视角、表情等因素影响而体现出高度的非线性特征，因此该方法在抽取样本更具鉴别力的非线性特征时无法取得良好的效果. 为了解决这个问题，核方法已经成为当前模式识别领域中一个迅猛发展的新方向，前人用核方法将 Fisher 线性鉴别分析推广到非线性空间，提出了核 Fisher 鉴别分析(Kernel Fisher Linear Discriminant Analysis，KFDA)方法[2-3]，其基本思想是首先将原始输入空间的训练样本通过非线性映射 ϕ 映射到某一高维(甚至是无穷维)特征空间 H 中，然后在该空间中应用经典的 Fisher 鉴别准则获得一组最优鉴别矢量，用于特征抽取. 虽然该方法用"核技巧"巧妙地避免了"维数灾难"问题，抽取了有效的非线性鉴别特征. 但是由于该方法中构造核样本向量的维数即为训练样本的个数，所以当训练样本很大时，通常存在核类内散度矩阵奇异的问题，从而无法直接求解最佳鉴别矢量集.

本文从鉴别准则函数本身入手，提出一种加权的核最大散度差鉴别分析方法. 该方法通过对核 Fisher 鉴别函数进行了修正，采用散度差作为鉴别准则函数，使其物理意义在没有发生任何改变的前提下，从理论上消除了核 Fisher 中存在的因核类内散度矩阵奇异而无法直接求解最优投影轴的困难. 另外，本文借鉴文献[4-5]，引入简单加权函数重新构造了类间散度矩阵和类内散度矩阵，从而对核的最大散度差鉴别准则函数做了进一步修正，提出了一种加权的核最大散度差准则鉴别分析方法. 该方法集中了核最大散度差鉴别分析与近似加权方法的优越性. 通过在 ORL 和 Yale 人脸数据库上的实验证明本文方法的有效性和适应性.

1　Fisher 鉴别分析(FLDA)

设有 C 个类别，第 i 类有 N_i 个样本，共有 N 个样本，$N=\sum_{i=1}^{c}N_i$，$\boldsymbol{x}_{i,j}\in\boldsymbol{R}^d$ 为第 i 类，第 j 个 d 维的样本，对人脸图像而言，通常是将图像按列或行拉直成向量，因此 d 通常较大. 类间散度矩阵 $\boldsymbol{S}_b$、类内散度矩阵 $\boldsymbol{S}_w$ 和总散度矩阵 $\boldsymbol{S}_t$ 定义如下：

类间散度矩阵：

$$\boldsymbol{S}_b=\sum_{i=1}^{c}\frac{N_i}{N}(\bar{\boldsymbol{x}}_i-\bar{\boldsymbol{x}})(\bar{\boldsymbol{x}}_i-\bar{\boldsymbol{x}})^{\mathrm{T}};\qquad(1)$$

类内散度矩阵：

$$\boldsymbol{S}_w=\frac{1}{N}\sum_{i=1}^{c}\sum_{j=1}^{N_i}(\boldsymbol{x}_{i,j}-\bar{\boldsymbol{x}}_i)(\boldsymbol{x}_{i,j}-\bar{\boldsymbol{x}}_i)^{\mathrm{T}};\qquad(2)$$

总散度矩阵：

$$\boldsymbol{S}_t=\boldsymbol{S}_b+\boldsymbol{S}_w\qquad(3)$$

其中：

$$\bar{\boldsymbol{x}}_i=\frac{1}{N_i}\sum_{j=1}^{N_i}\boldsymbol{x}_{i,j},\bar{\boldsymbol{x}}=\frac{1}{N_i}\sum_{i=1}^{c}\sum_{j=1}^{N_i}\boldsymbol{x}_{i,j}$$

收稿日期：2009-06-16

作者简介：刘晓亮(1983-)，男，硕士研究生，主要研究方向为信号与图像处理.

从式(1)、(2)、(3)定义易知 S_b、S_w 和 S_t 均为对称非负定阵. 令 W 为最优鉴别矢量, Fisher 准则函数定义如下:

$$\max J(W) = \frac{W^T S_b W}{W^T S_w W}, W \neq 0^d. \quad (4)$$

2　核函数的基本概念

定义1　核函数[5]对所有 $x, y \in X, X \subset R^n$ 若函数 k 满足 $k(x,y) = \langle \phi(x) \cdot \phi(y) \rangle$ 则称 k 为核函数. 式中 ϕ 为从输入空间 X 到内积特征空间 H 的映射,表示内积.

定义2　核函数矩阵[5]对输入空间的向量集 $S = \{x_1, \cdots, x_n\}$, 核函数矩阵 K 定义为一个 $n \times n$ 的矩阵,且其矩阵元素为

$$K_{ij} = \langle \phi(x_i) \cdot \phi(x_j) \rangle = K(x_i, x_j).$$

显然, K 为一对称矩阵. 任何满足如下 Mercer 条件:对任意的 $\phi(x) \neq 0$, 且 $\int \phi^2(x) dx < \infty$, 对称核函数 $k(x,y)$ 满足 $\iint \phi(x) k(x,y) \phi(y) dx dy > 0$. 都可作为核函数,常用两种简单核函数有:

1) 径向基核函数

$$k(x,z) = \exp\left(-\frac{\| x - z \|^2}{2\sigma^2}\right),$$

式中, σ 为核参数.

2) 多项式核函数.

$$k(x,z) = (x^T z + d)^e,$$

式中, d, e 为核参数.

3　核的 Fisher 鉴别分析(KFDA)

基于核的 Fisher 非线性鉴别分析的基本思想是,先对样本进行非线性变换,将其映射到特征空间 H 中,然后在特征空间进行 Fisher 线性鉴别分析. 则经过非线性变换后类内散度矩阵 S_w 和类间散度矩阵 S_t 总散度矩阵 S_t 可改写为如下的矩阵形式:

$$S_b^\phi = \sum_{i=1}^{c} \frac{N_i}{N} ((\bar{x}_i)^\phi - (\bar{x})^\phi)(\bar{x}_i)^\phi - (\bar{x})^\phi))^T; \quad (5)$$

$$S_w^\phi = \frac{1}{N} \sum_{i=1}^{c} \sum_{j=1}^{N_i} (\phi(x_{i,j}) - (\bar{x}_i)^\phi)(\phi(x_{i,j}) - (\bar{x}_i)^\phi)^T, \quad (6)$$

$$S_t^\phi = S_b^\phi + S_w^\phi \quad (7)$$

由式(4)、(5)、(6)易知 S_b^ϕ、S_b^ϕ、S_t^ϕ 均为对称非负定阵. 与线性 Fisher 鉴别分析相同,非线性 Fisher 鉴别分析是在特征空间 H 中求解最佳鉴别矢量 W_ϕ, 使得如下的 Fisher 准则最大

$$\max J(W_\phi) = \frac{W_\phi^T S_b^\phi W_\phi}{W_\phi^T S_w^\phi W_\phi}, W_\phi \neq 0^d. \quad (8)$$

由于核方法的特点是非线性映射的具体形式未知,因此无法直接计算式(7). 但根据再生核理论[5],在特征空间 H 中,由于最佳鉴别矢量 W_ϕ 在所有训练样本 $\phi(x_1), \phi(x_2), \cdots, \phi(x_N)$ 所张成的空间内,因此可以写成所有训练样本的组合,即有

$$W_\phi = \sum_{k=1}^{N} a_k \phi(x_k) = X_\phi \alpha, \quad (9)$$

其中 $X_\phi = (\phi(x_1), \phi(x_2), \cdots, \phi(x_N))$, $\alpha = (a_1, a_2, \cdots, a_N)^T$

令 K_i 为 N 维列向量则

$$K_i = (k(x_1, x_i), k(x_2, x_i), \cdots, k(x_N, x_i))^T \quad (10)$$

将上式 K_i 和式(5)、(6)和(9)代入 Fisher 准则函数可得:

$$K_b = \sum_{i=1}^{c} \frac{N_i}{N} (u_i - u_0)(u_i - u_0)^T, \quad (11)$$

$$K_w = \frac{1}{N} \sum_{i=1}^{C} \sum_{j=1}^{N_i} (\varphi(x_{i,j}) - u_i)(\varphi(x_{i,j}) - u_i)^T, \quad (12)$$

$$K_t = K_b + K_w, \quad (13)$$

$$J(\alpha) = \frac{W_\phi^T S_b^\phi W_\phi}{W_\phi^T S_w^\phi W_\phi} = \frac{\alpha^T K_b \alpha}{\alpha^T K_w \alpha} \quad (14)$$

其中 u_i 为第 i 类输入样本的核样本向量均值, u_0 为总体核样本的均值, $\varphi(x_{i,j})$ 为第 i 类 j 列的核样本. K_b、K_w、K_t 分别称为核类间散度矩阵、核类内散度矩阵以及核总散度矩阵. 由式(11)、(12)、(13)可知它们均为对称非负定阵. 这样求解核 Fisher 最佳鉴别矢量 W_ϕ 就转化为求解使式(13)的 Fisher 准则函数达到最大值的最优解 α 的问题. 由矩阵论中的广义 Rayleigh 商定理[6]知,求解此最优解 α 等价于求解如下与广义特征方程式的非零特征值对应的特征矢量

$$K_b \alpha = \lambda K_w \alpha. \quad (15)$$

由于样本可以独立取样得到,因此 S_b^ϕ 的秩为 C-1, 若 S_w^ϕ 非奇异,则此广义特征方程有 C-1 个非零特征值,其所对应的 C-1 个特征向量 $\alpha_k (k = 1, 2, \cdots, C-1)$ 即为满足式 Fisher 准则的最优解. 假设 $\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_k)$ 是 $J(\alpha)$ 一组特征最优解,由 α 和式(8)即可得到核 Fisher 的非线性最佳鉴别矢量 $W_\phi = (w_1, w_2, \cdots, w_k)$, 其中, $w_i = \sum_{k=1}^{C-1} \alpha_k^i \phi(x_k), i = 1, 2, \cdots, N$.

4　加权的核散度差鉴别分析(WKMSDA)

鉴于传统的鉴别式(13)鉴别函数存在的"病态性",即由于人脸图像的训练样本总数太大,而造成的类内散度矩阵 S_w^ϕ 奇异性问题.本文采用散度差[7]作为鉴别准则.在非线性空间散度差准则即为:

$$J(W_\phi) = W_\phi^T S_b^\phi W_\phi - W_\phi^T S_w^\phi W_\phi = W_\phi^T(S_b^\phi - S_w^\phi)W_\phi. \tag{16}$$

该方法从理论上根本解决了核类内散度矩阵 K_w 奇异导致无法求解的问题.

加权的核最大散度差鉴别分析(WKMSDA)的基本思想是:假设有一类别离中心较远,称之为边缘类,在这种情况下,通过最大化核散度差准则函数得到的最优投影轴 W_ϕ 很可能分开边缘类与其他类别,因为在该方向上类别的方差最大,在最大化散度差函数值的过程中,那些离中心具有较大距离的类别起到了主导作用,这样的投影轴不但不能够分开除边缘类的其他类别,而且很可能会造成其它类别的重叠,针对这个问题重新定义了类间散度矩阵,即

$$S_B^\phi = \sum_{i=1}^{c}\frac{N_i}{N}\tilde{\omega}(d_i)(\phi(\bar{x}_i)-\phi(\bar{x}))\cdot(\phi(\bar{x}_i)-\phi(\bar{x}))^T. \tag{17}$$

其中,$d_i = \sqrt{(\phi(\bar{x}_i)-\phi(\bar{x}))^T(\phi(\bar{x}_i)-\phi(\bar{x}))}$, $\tilde{\omega}(\cdot)$ 为加权函数,d_i 为第 i 个类别到中心的距离,既然要消弱边缘类在优化过程中的主导作用,直觉的考虑是 $\|\bar{x}_i - \bar{x}\|$ 越大给它们的权值越小,不妨定义加权函数为 $\tilde{\omega}(d_i) = d_i^{-3}$. 在类内散度矩阵 S_w^ϕ 前加一个权 $a = 1/(1+\sum_{i=1}^{C}\sum_{j=1}^{N_i} d_j^i)$ 其中 $d_j^i = \sqrt{(\phi(x_{i,j})-(\bar{x}_i)^\phi)^T(\phi(x_{i,j})-(\bar{x}_i)^\phi)}$,则有

$$S_W^\phi = \frac{a}{N}\sum_{i=1}^{C}\sum_{j=1}^{N_i}(\phi(x_{i,j})-(\bar{x}_i)^\phi)(\phi(x_{i,j})-(\bar{x}_i)^\phi)^T. \tag{18}$$

这样可以将每个类的训练样本向其中心靠拢,避免其它类的重叠.特征空间 H 中加权的核散度差鉴别函数可定义为:

$$J(W_\phi^*) = W_\phi^{*T}(S_B^\phi - S_W^\phi)W^*, \tag{19}$$

其中

$$W_\phi^* = \sum_{K=1}^{N}\alpha_k^*\phi(x_k) = X_\phi\alpha^*, \tag{20}$$

α^* 为在特征空间 H 中的最优投影向量,根据再生核理论以及式(9)至(12)及式(16)可得

$$J(\alpha^*) = W_\phi^{*T}(S_B^\phi - S_W^\phi)W_\phi^* = \alpha^{*T}(K_B - K_W)\alpha^*, \tag{21}$$

$$K_B = \sum_{i=1}^{c}\frac{N_i}{N}\tilde{\omega}(d_i^*)(u_i - u_0)(u_i - u_0)^T, \tag{22}$$

$$K_W = \frac{a}{N}\sum_{i=1}^{C}\sum_{j=1}^{N_i}(\varphi(x_{i,j}) - u_i)(\varphi(x_{i,j}) - u_i)^T. \tag{23}$$

$$K_T = K_B + K_W \tag{24}$$

其中,$d_i^* = \sqrt{(u_i-u_0)^T(u_i-u_0)}$,任然取 $\tilde{\omega}(d_i^*) = (d_i^*)^{-3}$,$K_B$、$K_W$、$K_T$ 分别称为加权核类间散度矩阵、加权核类内散度矩阵和加权核总散度矩阵,由式(22)、(23)及(24)可知它们均为对称非负定矩阵.考虑到式(16)等价与如下形式

$$\begin{cases} J(\alpha^*) = \alpha^{*T}(K_B - K_W)\alpha^*, \\ \alpha^{*T}\alpha^* = 1. \end{cases} \tag{25}$$

因为 $\frac{\alpha^{*T}(K_B-K_W)\alpha^*}{\alpha^{*T}\alpha^*} = \frac{\alpha^{*T}(K_B-K_W)\alpha^*}{\alpha^{*T}I\alpha^*}$,其中 I 为单位矩阵.则根据广义 Rayleigh 商定理知,最优解 α 为 $(K_B - K_W)\alpha^* = \lambda\alpha^*$ 对应的单位特征矢量.假设 $\alpha^* = (\alpha_1^*, \alpha_2^*, \cdots, \alpha_k^*)$ 是 $J(\alpha^*)$ 一组最优解,由 α^* 和式(20)即可得到核 Fisher 的非线性最佳鉴别矢量

$$W_\phi^* = (w_1^*, w_2^*, \cdots, w_k^*,),$$

其中,$w_i = \sum_{k=1}^{C-1}\alpha_k^{*i}\phi(x_k)$, $i = 1,2,\cdots,N$.

5　实验结果

在 ORL 和 Yale 人脸数据库中进行人脸识别实验,以检验本文方法的性能. ORL 数据库由40个人,每人10幅 112×92 pixels 的图像共400幅人脸图像组成,这些图像有的拍摄于不同时期,有的人脸的表情、细节、视觉和尺度均有不同程度的变化.例如有笑与不笑、睁眼或闭眼、戴或不戴眼睛. Yale 人脸库容量较小,只包括15人,每人11幅 320×243 pixels 的图像,包括光照、表情和细节的变化,但视觉和尺度没有变化.实验之前,为降低计算量,先对人脸图像通过2级小波变换[8]进行降维,把得到的低频分量作为特征,用 $k(x,y) = (x^Ty+1)^2$ 作为统一的核函数,采用的分类器是最小距离分类器[9-10],由于试验中的训练、测试样本都是随机选择的,因此为了减少实验结果的波动,每个实验将重复30次,本文给出的结果是30次试验的平均结果.图1和图2分别是 ORL 和 Yale 人脸库中部分人脸图像.

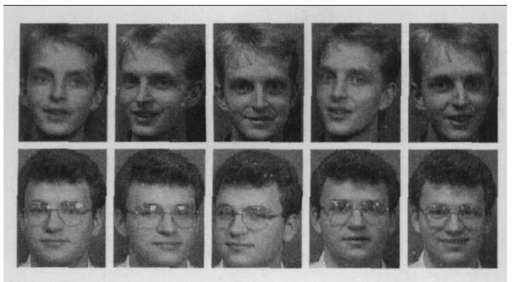

图1 ORL 人脸库上部分人脸图像

图2 Yale 人脸库上部分人脸图像

采用 WKMSDA 与传统的 KFDA、特征脸(KEIAENface)及 KPCA 方法分别在 ORL 和 Yale 人脸库中试验,对各种方法的性能进行了比较.图3显示了在 ORL 人脸库上取不同数目的训练样本时各种方法的识别率(特征均为39维),可以看出,当训练样本数目变化时,WKMSDA 方法具有最高的识别率,正确识别率对照如表1.图4显示了在 Yale 人脸库

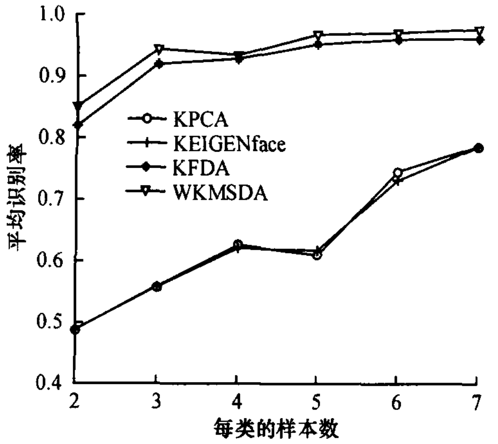

图3 在 ORL 人脸库中几种方法的比较

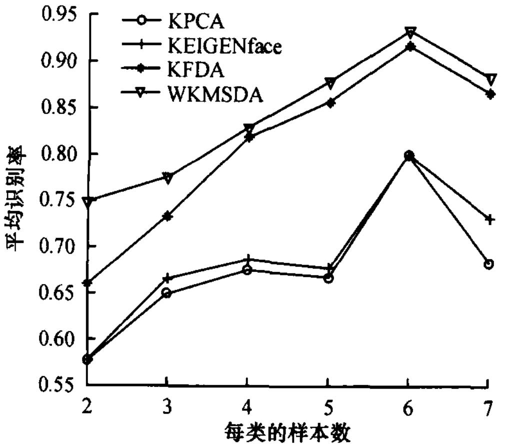

图4 在 Yale 人脸库中几种方法的比较

中取不同数目的训练样本时各方法的识别率(特征均为14),可以看出在 Yale 人脸库上 WKMSDA 仍然高于另外3种方法,总之,WKMSDA 在 ORL 和 Yale 人脸库中都显示出良好的性能.

表1 在 ORL 人脸库中几种方法识别率的对照

方法	样本数					
	2	3	4	5	6	7
KPCA	0.4875	0.5571	0.6250	0.6100	0.7438	0.7833
KEIGENface	0.4875	0.5571	0.6208	0.6150	0.7312	0.7833
KFDA	0.8187	0.9179	0.9250	0.9500	0.9563	0.9583
WKMSDA	0.8500	0.9429	0.9333	0.9650	0.9688	0.9750

表2 在 Yale 人脸库中几种方法识别率的对照

方法	样本数					
	2	3	4	5	6	7
KPCA	0.5778	0.6500	0.6762	0.6667	0.8002	0.6833
KEIGENface	0.5778	0.6650	0.6857	0.6767	0.8004	0.7312
KFDA	0.6593	0.7333	0.8190	0.8556	0.9167	0.8667
WKMSDA	0.7481	0.7750	0.8286	0.8778	0.9333	0.8833

6 结束语

本文在散度差方法的基础上,针对核 Fisher 准则函数在处理多类模式识别问题中的不足,提出了核的加权的散度差鉴别分析方法.该方法在高维特征空间 H 内使用加权函数重新估计了类间散度矩阵和类内散度矩阵,通过给予边缘类较小的权值,而给那些离中心距离较小的类别较大的权值.减轻边缘类对投影方向结果的主导作用,同时给类内散度矩阵一个特殊的权值使得每类的样本都向中心靠拢,这样在一定程度上避免其它类可能出现的重叠,从而使得本文方法的识别率得到显著的提高.

参考文献:

[1] Belhumeur P N, Hespanha J P, Kriegnan D J. Eigenfaces vs Fisherfaces: recognition using special linear projection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997, 19(7): 711-720.

[2] Mika S, Ratch G, Weston J, et al. Fisher discriminantAnalysis with kernels[C]// proceedings of IEEE International Workshop on Neural Networks for Signal proceedings IX. USA: Madison, Wisconsin, 1999: 41-48.

[3] Volker Roth, Volker Steinhage. Nonlinear discriminant analysis using kernel functions[C]// Denver: Proc of Neural Information Processing systems, 1999: 186-198.

[4] Liang Yi-xiong, Li Cheng-rong, Gong Wei-guo, et al. Uncorrelated linear discriminant analysis based on weighted pairwise Fisher criterion[J]. Journal of Pattern Recognition Society, 2007(40): 3606-3615.

[5] 成新民,蒋云良,胡文军,等. 基于核的 Fisher 非线性最佳鉴别分析在人脸识别中的应用[J]. 中国图像图形学报,2007,12(8):1394-1400.

[6] 程云鹏. 矩阵论[M]. 西安:西北工业大学出版社,1989.

[7] Li Hai-feng, Jing Tao, Zhang Ke-shu. Efficient and Robust Feature Extraction by Maximum MarginCriterion[J]. IEEE Transactions on Neural Network, 2006, 17(1): 157-165.

[8] 耿则勋. 小波变换理论及在遥感影像压缩中的应用[M]. 北京:测绘出版社,2002.

[9] Richard O Duda, Peter E Hart, David G Stork. Pattern Classification[M]. 北京:机械工业出版社,2002:196-201.

[10] 孔锐,张冰. 基于核 Fisher 判决分析的高性能多类分类算法[J]. 计算机应用,2005,25(6):1327-1329.

The Application of a Weighted Kernel Fisher Discriminant Analysis Applied in Face Recognition

Liu Xiao-liang, Wang Fu-long, Huang Cheng, Zeng Ai-hua

(Faculty of Applied Mathematics, Guangdong University of Technology, Guangzhou 510006, China)

Abstract: A weighted kernel maximum scatter difference discriminating criterion is developed for extraction of nonlinear features. The proposed method can be used to extract nonlinear features for faces effectively anf to reconstruct between-class scatter matrix and within-class scatter matrix by weighted schemes. Therefore, it can modify the kernel maximum scatter difference discriminating criterions. Experiments performed on ORL and Yale face database verify the effectiveness of the proposed method.

Key words: face recognition; Kernel Fisher Linear Discriminant Analysis(KFDA); maximum satter difference criterion

简讯

我校二刊双获全国高校科技期刊优秀编辑质量奖

由中国高等学校自然科学学报研究会组织开展的2009年全国高校科技期刊优秀编辑质量奖评比活动已经结束,由我部编辑出版的《广东工业大学学报》和《工业工程》双双获得全国高校科技期刊优秀编辑质量奖(奖状见封二). 本次全国获奖期刊共279种,评比分2阶段进行,第1阶段由审读员给出编辑质量审读结果,差错率在3/万以下者参与评奖;第2阶段,由评委会评出优秀编辑质量奖期刊. 期刊获此奖项,说明该刊编辑质量达到较高水平.

《广东工业大学学报》编辑部

《工业工程》编辑部