

基于 GA-BP 神经网络的双摄像机位姿视觉调节方法

杨风开 程素霞

(华中科技大学电气与电子工程学院 武汉 430074)

摘要 文中提出了双摄像机模组位姿调节参数计算的 GA 优化 BP 神经网络模型,根据目标模板上的特征点在双摄像机上的成像坐标,计算两个摄像机之间的位姿偏离参数。为弥补 BP 神经网络的不足,采用 GA 算法对 BP 神经网络进行了优化。利用训练样本数据集对所提出的模型进行了训练,并利用测试样本数据集对模型进行了测试;最后将训练好的模型用于双摄像机模组位姿调节的实际生产中。实际应用结果表明,基于所提出的方法设计的双摄像机模组位姿调节装置,调节精度和调节时间都能满足实际生产的要求。

关键词 双摄像机模组,位姿,BP 神经网络,GA 算法,视觉调节

中图分类号 TP3-0 **文献标识码** A

Method for Visual Adjustment of Two-camera Position Based on GA-BP Neural Network

YANG Feng-kai CHENG Su-xia

(School of Electrical and Electronic Engineering, Huazhong University of Science and Technology, Wuhan 430074, China)

Abstract The target template was designed and the BP neural network model was proposed, which can calculate the position deviation parameter between the two cameras according to the image coordinates of the feature points on the target template on the dual camera. GA algorithm is used to optimize BP neural network to compensate the shortcomings. The training sample data set is used to train the proposed model, and the model is tested with the test sample data set, and finally the training model is used for the actual production of the two-camera module. The actual application results show that the calibration precision and time can meet the requirements of actual production based on the proposed method.

Keywords Two-camera module, Position, BP neural network, GA algorithm, Visual adjustment

在双摄像机模组生产的封装过程中,需要进行摄像机的位姿调节,以保持两个摄像机的位姿一致性。为此,需要基于两个摄像机对同一个目标模板上的特征点进行视觉识别,测量两个摄像机之间的位姿参数。生产制作流程中,要求测量的位姿参数精确,计算速度较快。

现有的双目视觉定位研究一般都是基于两个标定调节好的摄像机来确定目标物和摄像机模组之间的位姿关系^[1-3],缺乏对两个摄像机之间位姿调节方法的研究。但是在现有的关于摄像机位姿计算和标定的研究成果的基础上,可以以目标模板为参照物分别确定两个摄像机的位姿参数,进而计算出两个摄像机之间的位姿关系。文献[4]离线采集目标物体不同角度的图像生成目标模板图片库,利用基于灰度相关的区域匹配算法进行三维重建实现目标定位,但是采集不同角度的图像对于双摄像机模组调节工艺来说,需要专门的操动机构,增加了设备成本和工艺复杂度。文献[5]提出了一种基于灭点一致性约束的摄像机标定算法,利用目标的双目图像立体匹配结果,通过改进迭代最近点算法实现目标位姿估计,但是计算精度不能满足双摄模组位姿调节的要求。

双摄像机模组的位姿调节系统是一个多变量、非线性、强耦合系统,已有大量研究采用人工智能控制方法,有效地提高了非线性、耦合性系统的控制性能。文献[6]针对磨机负荷控

制存在大惯性滞后、参数耦合性强和时变性等问题,将模糊控制与人工神经网络控制相结合,增加了磨机运行的稳定性,但是迭代的次数较多,计算时间较长。文献[7]提出了一种基于双目视觉导航的仿生机器人鲁棒控制算法,利用光学 CCD 双目视觉动态跟踪系统进行仿生机器人的末端位姿参数测量,建立被控对象的运动学模型;以机器人转动关节的 6 自由度参量为控制约束参量,建立机器人的分层子维空间运动规划模型;采用双目视觉跟踪方法实现仿生机器人的位姿自适应修正。该方法具有良好的动态跟踪性能,但是建模的过程比较复杂,而双摄像机模组位姿调节系统的建模则更加困难。

本文在文献[7]研究成果的基础上,提出了计算位姿参数的 GA 优化 BP 神经网络模型,根据专门设计的目标模板上的特征点在双摄像机上的成像坐标,计算两个摄像机之间的位姿偏离参数,用于控制摄像机的位姿调节,不需要对被控对象进行精确的数学建模,也不需要精确测量目标模板上特征点的世界坐标系坐标。为弥补 BP 神经网络的不足,采用 GA 算法对 BP 神经网络进行了优化。利用训练样本数据集对所提出的模型进行了训练,并利用测试样本数据集对模型进行了测试,最后将训练好的模型用于双摄像机模组位姿调节的实际生产中。

1 双摄像机模组的位姿调节系统

1.1 位姿调节参数

分别以两个摄像机成像平面的角点作为摄像机三维坐标系的原点,使成像平面位于摄像机坐标系的XY平面上,如图1所示,则两个摄像机的位姿关系即为两个摄像机坐标系之间的位姿关系,可以用3个坐标轴上的旋转量和偏移量来表征。

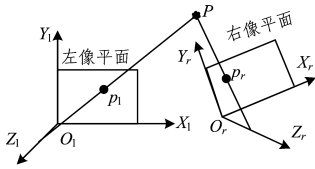


图1 两个摄像机之间的位姿关系

以左摄像机为基准,两个摄像机之间的位姿关系可以表征为右摄像机相对于左摄像机在3个坐标轴上的旋转量 $\theta_x, \theta_y, \theta_z$ 以及3个坐标轴上的平移偏移量 $\Delta x, \Delta y, \Delta z$,共有6个位姿参数。测量这6个位姿参数,然后调节右摄像机,使得 Δx 等于两摄像机之间的预定间距 m ,其余5个参数为0(误差在规定的范围内),即实现了两个摄像机之间的位姿一致性。

1.2 摄像机线性成像模型

理想摄像机线性成像模型如图2所示,包括世界坐标系、摄像机坐标系、图像坐标系和像素坐标系。由于像素坐标系、图像坐标系都可以转换到摄像机坐标系中,因此摄像机的标定就是求取世界坐标系与摄像机坐标系之间的关系。

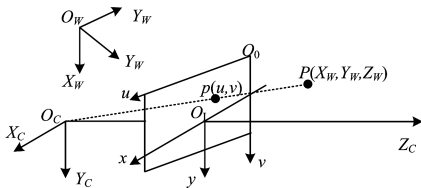


图2 摄像机成像线性模型

设主点在图像坐标系中的坐标为 $O_I(u_0, v_0)$,像素点在 x 轴、 y 轴的物理尺寸分别为 dx, dy ,特征点在世界坐标系中的坐标为 $P(X_w, Y_w, Z_w)$,摄像机坐标系原点到图像坐标自原点的距离为 Z_c ,摄像机的焦距为 f ,则可得某特征点在世界坐标系和摄像机坐标系的对应关系^[5]为:

$$Z_c \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{dx} & 0 & u_0 \\ 0 & \frac{1}{dy} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & t \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \quad (1)$$

其中, R 为单位旋转矩阵, t 为平移矩阵。 R 为 3×3 单位矩阵,可以归一化为3个待求参数。平移矩阵 t 中有3个待求参数,再加上摄像机内参数 f, u_0, v_0 ,共有9个待求未知参数。在空间中取不少于9个点,根据每个点的像素坐标和世界坐标值,分别代入式(1)可得到一个超定方程组,然后采用最小二乘法解此超定方程组即可求出9个待求参数,从而得到摄像机的全部内参数和相对于目标模板的外参数,所得到的外参数即为摄像机的位姿参数。

因此,基于单目摄像机线性成像模型,利用空间中目标模板上的9个以上特征点的成像坐标,就能够以目标模板所在的世界坐标系为基准,求出单目摄像机的位姿参数。分别求

取两个单目摄像机相对于世界坐标系的位姿参数,即能够计算出两个摄像机之间的相对位姿参数。

1.3 双摄像机位姿调节系统

由于摄像机难免存在成像畸变,成像点的坐标与待求解的6个自由度调节量之间为非线性的关系,要想利用几何方法求解,需要采用优化算法来逼近^[8-9],计算过程非常复杂,常见的计算方法需要花费较长的时间,有些算法还需要辅助的位姿测量设备或机械操纵机构。而BP神经网络非常适合求解非线性关系,因此本文采用优化BP神经网络来求解两摄像机之间的各位姿偏移量,然后进行位姿调节,所设计的调节系统的结构如图3所示。

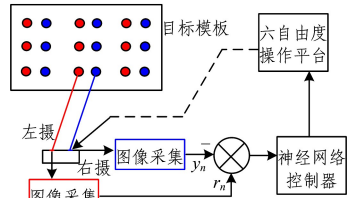


图3 双摄像机位姿调节系统结构

左、右两个摄像机的图像经图像采集处理后,得到目标模板上两个特征对点在各自摄像机中的成像坐标。求取每一特征对点在两个摄像机成像坐标的误差,作为神经网络控制器的输入。神经网络的输出为需调节的6个自由度偏移量,用于控制六自由度操作平台,从而调节通过夹具连接在操作平台上的右摄像机位姿。

2 优化BP神经网络

BP神经网络是一种多层前馈型神经网络,应用非常广泛,其主要用作分类和识别工具,适合处理非线性关系。采用BP神经网络可以不用计算目标模板上的特征点世界坐标,直接利用双摄像机上的成像坐标计算六自由度偏移量,计算量很小。

2.1 BP神经网络结构

BP神经网络模型的拓扑结构包括一个输入层、一个或多个隐含层和一个输出层,本文采用3层神经网络,只有一个隐含层,其拓扑结构如图4所示。BP神经网络的学习过程是由信号的正向传播与误差的反向传播两个阶段组成的^[10]。正向传播时,输入样本值到输入层,经隐含层逐层处理后传向输出层,若输出层的实际输出值与期望输出值不符,则进入反向传播误差的阶段。反向传播误差时,是将输出误差通过隐含层向输入层逐层反传,并分摊误差到各层的所有节点,获得各层节点的误差信号,作为修正各节点权值的依据。信号的正向传播与误差的反向传播循环地进行,期间各层权值不断得到调整,该过程是网络的学习训练过程。学习过程直到网络输出的误差可以接受,或者进行到预先设定的循环学习次数为止^[11]。

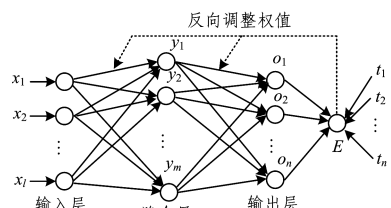


图4 BP神经网络模型的拓扑结构

图 4 中,输入层有 l 个节点,其输入向量为 $X = (x_1, x_2, \dots, x_l)^T$;隐含层有 m 个节点,其输出向量为 $Y = (y_1, y_2, \dots, y_m)^T$;输出层有 n 个节点,其输出向量为 $o = (o_1, o_2, \dots, o_n)^T$ 。与输出层对应的有 n 个输出期望值,其向量为 $t = (t_1, t_2, \dots, t_n)^T$; E 为输出层和期望值之间的误差信号。图 4 的 3 层结构中,上一层的输入作为下一层的输出,节点处的箭头代表输入量和输出量,各层节点的输入与输出之间的数学关系如下:

1) 对于输入层,输入=输出,即第一层的输出也为 x_i ;

2) 对于隐含层,有:

$$y_j = f\left(\sum_{i=1}^m \omega_{ij} x_i - \theta_j\right) \quad (2)$$

其中, ω_{ij} 为连接权值, θ_j 为第 j 个神经元的阈值, $f(\cdot)$ 为激励函数,通常采用 Sigmoid 函数:

$$f(x) = \frac{1}{1 + e^{-x}} \quad (3)$$

$f(x)$ 具有连续可导的特性,即:

$$f(x') = f(x)[1 - f(x)] \quad (4)$$

3) 对于输出层,有:

$$o_k = f\left(\sum_{j=1}^n \omega_{jk} y_j - \theta_k\right) = f\left(\sum_{j=1}^n \omega_{jk} f\left(\sum_{i=1}^m \omega_{ij} x_i - \theta_j\right) - \theta_k\right) \quad (5)$$

4) 误差信号:如果网络的实际输出与期望输出不等,其误差函数为:

$$\begin{aligned} E &= \frac{1}{2} \sum_{k=1}^n (t_k - o_k)^2 \\ &= \frac{1}{2} \sum_{k=1}^n [t_k - f\left(\sum_{j=1}^n \omega_{jk} f\left(\sum_{i=1}^m \omega_{ij} x_i - \theta_j\right) - \theta_k\right)]^2 \end{aligned} \quad (6)$$

由此可见,网络误差是各层权值 ω_{ij} 和 θ_j 与 ω_{jk} 和 θ_k 的函数。

调整权值和阈值使误差不断减小,就是使权值和阈值的调整量正比于误差的负梯度,即:

$$\Delta \omega_{ij} = -\eta \frac{\partial E}{\partial \omega_{ij}}, \Delta \omega_{jk} = -\eta \frac{\partial E}{\partial \omega_{jk}} \quad (7)$$

$$\Delta \theta_j = -\eta \frac{\partial E}{\partial \theta_j}, \Delta \theta_k = -\eta \frac{\partial E}{\partial \theta_k} \quad (8)$$

其中,常数 η 称为学习因子, $\eta \in (0, 1)$ 在训练中反映了学习的速度,也称为学习率^[12]。

2.2 GA 优化算法

经典的 BP 神经网络存在易陷入局部极小点、收敛速度较慢等缺点^[13],为此本文采用遗传算法(Genetic Algorithm, GA)对 BP 网络进行优化。GA 是一种仿照生物“物竞天择,适者生存”的演化规律而提出的算法,其基本算法思想是在基本算子选择、交叉和变异的作用下,产生适应度更高的个体基因,以优化种群基因数据^[14]。

GA 算法的基本步骤如下。

1) 编码:GA 在进行搜索之前,先将解空间的解数据表示成遗传空间的基因型串结构数据,这些串结构数据即为初始个体;

2) 初始群体的生成:在上述生成的初始串结构中随机选取 N 个作为初始种群,进行选择、交叉和变异操作;

3) 适应度评价:适应度函数是评价新产生个体优劣的指标,选取其中适应度高的个体形成新的种群继续进行选择、交叉和变异,直到满足终止条件;

4) 对最后一代群体解码得到最优解。

用 GA 遗传算法辅助 BP 神经网络对数据进行预处理,得到最佳神经网络初始权值和阈值,赋值给神经网络。采用

GA 优化的神经网络算法的流程图如图 5 所示。

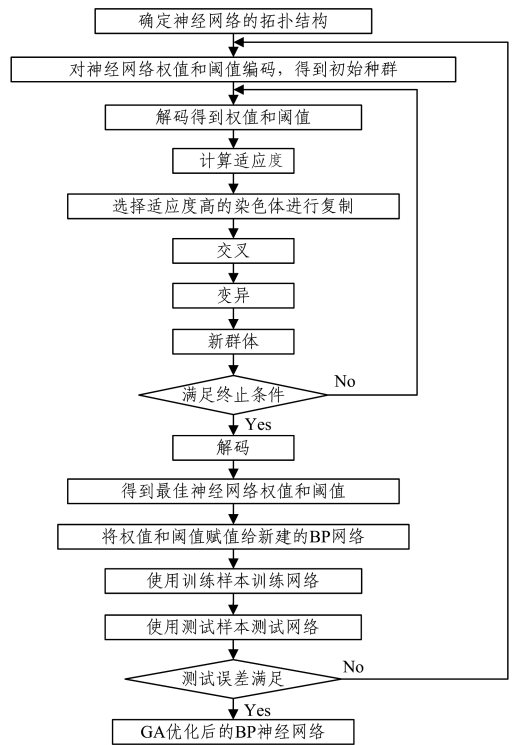


图 5 GA 优化的 BP 神经网络模型拓扑结构

调用 MATLAB 中的神经网络工具箱和遗传算法 GAOT 工具箱,设定遗传代数为 100,学习率为 0.1,最大迭代步数为 1000,学习目标为 0.001,对神经网络进行训练。采用 GA 算法优化的 BP 神经网络,既可以优化 BP 神经网络的初始权值和阈值,又可以提高算法的收敛速度和鲁棒性。

3 双摄像机调节神经网络模型

根据 GA 优化 BP 神经网络的算法原理,按双摄像机调节的实际情况进行具体神经网络模型的设计,然后获取网络训练所需的样本数据集并得到网络模型所需的参数,才能进行网络的训练和计算。

3.1 输入输出向量

设计目标模板示意如图 3 所示,其中含有相互对应匹配的 9 组目标特征点,分为左、右两组,其中特征点采用十字型。摄像机拍摄的图像经过图像采集卡采集到计算机,采用角点检测程序 Harris 对图像进行处理,得到在左、右摄像机拍摄图片中对应的像素点坐标分别为 $X_{r1}(x_{r1}, y_{r1}), \dots, X_{r9}(x_{r9}, y_{r9}), X_{l1}(x_{l1}, y_{l1}), \dots, X_{l9}(x_{l9}, y_{l9})$ 。在如图 4 所示的神经网络中,以右摄像机与左摄像机的各特征点像素坐标的差值构建输入向量,则共有 18 个输入参数,输入向量为: $X = (x_{r1} - x_{l1}, x_{r2} - x_{l2}, \dots, x_{r9} - x_{l9}, y_{r1} - y_{l1}, y_{r2} - y_{l2}, \dots, y_{r9} - y_{l9})^T$; 以需调节的 6 个自由度参数作为输出参数,得到输出向量为: $O = (\theta_x, \theta_y, \theta_z, \Delta x, \Delta y, \Delta z)^T$ 。

3.2 数据的标准化

原始数据的量值相差很大,量纲也互不相同,为了便于计算,需要对原始数据进行标准化处理。由于 BP 神经网络使用的 S 型激励函数对 0 到 1 之间的数据比较敏感,因此本文采用线性函数转换法,将所有数据映射到 $(0, 1)$ 区间内,计算方法如下:

$$x' = \frac{x - x_{\min}}{x_{\max} - x_{\min} + \alpha} \quad (9)$$

其中, x, x' 为转换前、后的值, x_{\max}, x_{\min} 为样本中的最大值和最小值。 α 为常数, 以避免出现 $x' = 1$ 的情况, 本文取 $\alpha = 0.001$ 。

3.3 模型的训练和测试

从标准状态开始, 随机调节 6 个自由度 600 次, 使右摄像机处于 600 种位姿状态, 记录每种状态下的 $\theta_x, \theta_y, \theta_z, \Delta x, \Delta y, \Delta z$ 值, 作为输出期望值; 同时记录每种状态下两个像素坐标系中测得各特征点成像坐标值, 作为输入参数。由此可得到包含 600 组数据的数据集, 随机选取其中 550 组数据作为训练样本数据集, 用于训练 GA 优化 BP 神经网络; 剩下 50 组作为测试样本数据集, 用于验证训练好的神经网络。分别采用 BP 神经网络和 GA 优化的神经网络训练模型, 得到的模型收敛曲线如图 6 所示。

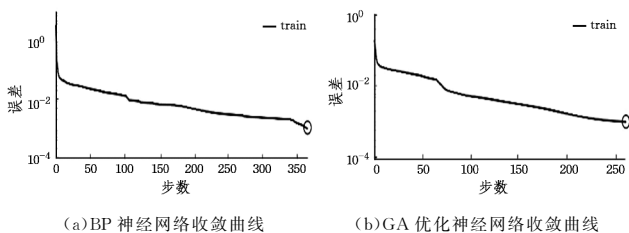


图 6 模型训练收敛曲线

由图 6 可知, BP 神经网络初始误差为 2.5, 且迭代到 358 步达到收敛目标, 收敛速度比较慢; 而 GA 优化的神经网络初始误差仅为 0.1, 且在迭代到 254 步就达到收敛目标, 收敛速度相对较快。通过对比可知, GA 遗传算法可以优化神经网络的初始权值和阈值, 提高神经网络训练的收敛速度和鲁棒性。

对于经过训练后的 GA 优化 BP 神经网络模型, 利用 50 组测试样本数据集进行测试, 位姿偏移量计算的合格率为 100%, 满足双摄像机模组封装生产工艺的要求。

4 应用结果

将训练、测试好的 GA 优化 BP 神经网络模型集成到双摄像机模组位姿调节系统中, 构成双摄像机模组位姿调节装置, 其内部结构如图 7 所示。

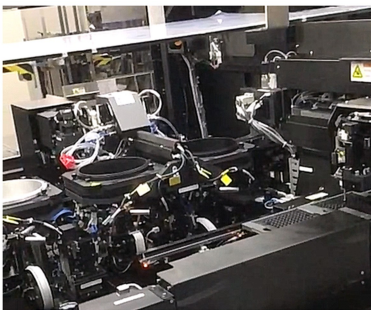


图 7 双摄像机模组调节装置的内部结构

在双摄像机模组封装生产工艺中进行实际应用, 结果表明: 每个双摄像机模组的位姿调节工艺时间不超过 12 s, 三维坐标方向的旋转量误差不超过 0.2° , X 轴、Y 轴坐标方向的平移量误差不超过 8 Pixel (像素点), Z 轴平移量偏差不超过

0.02 mm, 调节精度达到封装工艺预定的要求。

在双摄像机模组位姿调节生产线上, 利用训练好的 GA 优化 BP 神经网络模型测量摄像机位姿参数, 实际生产合格率不低于 99.9%, 能够满足双摄像机模组位姿调节生产工艺的需求。

结束语 根据所设计的目标模板上的特征点在双摄像机上的成像坐标, 基于 GA 优化 BP 神经网络模型进行双摄像机之间的位姿调节。调节工艺的机械结构简单, 不需要外部测量设备, 不需要标定摄像机的参数, 不需要测量目标模板上特征点的世界坐标, 能够一次性得到需要调节的 6 个自由度参数, 从而一次性将 6 个自由度调节到位, 不需要反复测量计算并反复调节, 计算和调节的时间较短, 有利于提高生产效率。

样本数据集训练和测试结果表明, 所提出的 GA 优化 BP 神经网络模型的收敛性好, 具有较高的计算精度, 网络训练的迭代次数较少, 计算效率较高。在双摄像机模组位姿调节生产线上, 利用训练好的 GA 优化 BP 神经网络模型计算摄像机位姿参数, 合格率不低于 99.9%。

参考文献

- [1] 唐庆顺, 吴春富, 李国栋, 等. 移动机器人车载摄像机位姿的高精度快速求解[J]. 山东大学学报(理学版), 2015, 50(3): 32-40.
- [2] 赵霞, 袁家政, 刘宏哲. 基于视觉的目标定位技术的研究进展[J]. 计算机科学, 2016, 43(6): 10-16.
- [3] 龙邹荣, 冯鹏, 杨荣来, 等. 复杂场景双目立体视觉测量精确定位系统研究[J]. 光电子·激光, 2017, 28(11): 1252-1260.
- [4] 王德海, 洪伟, 程群哲. 基于双目立体视觉的目标识别与定位[J]. 吉林大学学报(信息科学版), 2016, 34(2): 289-296.
- [5] 江士雄, 曹丹华, 吴裕斌. 面向机器人抓取的双目视觉系统标定与目标位姿估计算法[J]. 仪表技术与传感器, 2016, 16: 110-114.
- [6] 张杰, 王建民, 杨志刚, 等. 模糊神经网络在磨机负荷控制中的应用[J]. 仪表技术与传感器, 2014, 5: 66-69.
- [7] 李秀娟, 刘伟, 李珊红. 基于双目视觉导航的仿生机器人鲁棒控制算法[J]. 计算机科学, 2017, 44(8): 318-321.
- [8] 陈梦婷, 闫冬梅, 王刚. 基于 Harris 角点和 SIFT 描述符的高分辨率遥感影像匹配算法[J]. 中国图像图形学报, 2012, 17(11): 1453-1460.
- [9] 陈双叶, 王善喜. PTZ 摄像机跟踪运动目标的智能控制算法的研究[J]. 计算机科学, 2015, 42(S2): 135-139.
- [10] 严旭, 李思源, 张征. 基于遗传算法的 BP 神经网络在城市用水量预测中的应用[J]. 计算机科学, 2016, 43(S2): 547-550.
- [11] 陈馈, 牛彦杰, 李阁强, 等. 基于单神经元 PID 的盾构推进系统同步控制研究[J]. 工程设计学报, 2017, 24(3): 330-337.
- [12] 李亚, 刘丽平, 李柏青, 等. 基于改进 K-Means 聚类和 BP 神经网络的台区线损率计算方法[J]. 中国电机工程学报, 2016, 36(17): 4543-4551.
- [13] 姜长泓, 张永恒, 王盛慧. 基于改进粒子群优化算法的 PID 控制器参数优化[J]. 应用科学学报, 2017, 35(5): 667-675.
- [14] 胡宋国峰, 梁昌勇, 梁森, 等. 改进遗传算法优化 BP 神经网络的旅游景区日客流量预测[J]. 小型微型计算机系统, 2014(9): 2137-2140.