

基于信息共享策略的加权逐点预测动态多目标优化

包全磊¹ 陈红星²

¹(太原学院 山西 太原 030032)

²(山西大学计算机与信息技术学院 山西 太原 030006)

摘要 为了有效实现 Pareto 最优解,并且提升对于输入误差的鲁棒性,提出一种基于信息共享策略的加权逐点预测动态多目标优化算法。引入一种信息共享策略,该策略允许每个点利用其相邻解决方案中的信息有效提升模型的鲁棒性。引入一个相似性度量,并对其与一些常用的相似性度量进行对比分析,进一步提出一种加权逐点预测方法,逐点特性极大增强了捕捉各种模式的能力。在 EC2018 DMO 测试套件上的实验结果验证了该方法的有效性。

关键词 信息共享 鲁棒性 多目标优化 加权逐点预测

中图分类号 TP181 TP3 文献标志码 A DOI:10.3969/j.issn.1000-386x.2024.06.040

WEIGHTED POINT-WISE PREDICTION DYNAMIC MULTIPLE OBJECTIVE OPTIMIZATION BASED ON INFORMATION SHARING STRATEGY

Bao Quanlei¹ Chen Hongxing²

¹(Taiyuan University, Taiyuan 030032, Shanxi, China)

²(College of Computer and Information Technology, Shanxi University, Taiyuan 030006, Shanxi, China)

Abstract In order to achieve the Pareto optimal solution effectively and improve the robustness to the input error, a weighted point-wise prediction dynamic multiple objective optimization algorithm based on information sharing strategy is proposed. An information sharing strategy was introduced, which allowed each point to make use of the information in its adjacent solutions to improve the robustness of the model. A similarity measure was introduced, and by comparing with some common similarity measures, a weighted point-wise prediction method was proposed, which greatly enhanced the ability to capture various patterns. Experimental results on EC2018 DMO test suite show the effectiveness of the proposed method.

Keywords Information sharing Robustness Multiple objective optimization Weighted point-wise prediction

0 引言

动态多目标优化(Dynamic Multi-objective Optimization, DMO)问题是指当一个问题环境随时间变化时,多个目标被优化的一类问题^[1]。因为许多实际问题都受到动态变化和冲突的影响,DMO 在过去的十年中得到了越来越广泛的应用。例如控制系统设计、任务规划、废水处理和车辆路径规划等^[2-3]。

DMO 可以根据它们如何处理问题的动态性质来分类。一类修改所采用的单目标优化方法(Single Multi-objective Optimization, SMO)的算子,使其适应动态问题。文献[4]通过在优化过程中保持高度的多样性,使得算法能够对任何变化和先前收集的信息的丢失迅速作出反应。文献[5]提出一种新的调整邻域大小分配的分解多目标进化算法,以平衡算法的收敛性和多样性。然而,保持高度的多样性不利于信息融合,且满足条件十分苛刻,因为每个时间步的有限评估预

算需要快速收敛。此外,会极大地增大用于优化的计算成本。

另一类 DMO 方法为无干扰 SMO 的操作符,引入了独立运行的其他操作符。这类方法通常由三个模块组成:(1) 静态多目标优化(SMO)方法,用于优化两个连续变化之间的平稳问题;(2) 变化检测机制,用于检测变化发生;(3) 变化响应策略,该策略在变化后立即激活,主要用于为新的时间步提供种子种群。文献[6]提出了一种基于神经动力学系统的分布式多目标优化算法,该算法基于动态权重的神经网络模型,可以解决不可分离的不等式约束问题。文献[7]提出了一种基于决策变量分类的动态多目标算法,避免了分布性对收敛性的影响,使得收敛子种群能集中资源去优化个体的收敛性,加快算法的收敛速度。文献[8]提出了一种基于决策变量分类的混合预测策略以响应环境变化,该方法无论在收敛性还是在多样性上,均具有不错的性能。但是上述 DMO 的模块化允许单独地对每个组件进行探索,并对单个方法的优点得出更可靠的结论。但是上述方法很难找到相应的 Pareto 最优解,另外对于错误输入十分敏感。

为了解决上述问题,提出一种基于信息共享策略的加权逐点预测(WPPT)动态多目标优化算法。引入的信息共享策略允许每个点利用其相邻解决方案中的信息从而有效提升模型的鲁棒性,相似性度量以及加权逐点预测提升了算法的求解能力。实验结果验证了本文方法的有效性。

1 相似性度量指标

1.1 决策变量空间中的相似度量

相似性度量参照文献,它基于决策变量空间中的欧氏距离。对于每个 $\mathbf{x}_{FP_i}^{(t)} \in X_{FP}^{(t)}$,用 $\mathbf{x}_{FP_i}^{(t-1)}$ 表示映射集 $X_{FP}^{(t-1)}$ 中的对应解如下:

$$\mathbf{x}_{FP_i}^{(t-1)} = \arg \min_{\mathbf{x}_{FP_i}^{(t-1)}} (\|\mathbf{x}_{FP_i}^{(t)} - \mathbf{x}_{FP_i}^{(t-1)}\|_2) \quad i=1,2,\dots,N \quad (1)$$

尽管此方法很直观,但它有一个很大的缺陷。当 $X_{FP}^{(t)}$ 和 $X_{FP}^{(t-1)}$ 的质心不重合时,得到的结果很有可能是错的。 $t=3$ 和三个解决方案 $\mathbf{x}_{FP_1}^{(3)}, \mathbf{x}_{FP_2}^{(3)}, \mathbf{x}_{FP_6}^{(3)} \in X_{FP}^{(3)}$ 的情况如图 1 所示,图 1(a) 描述了可以通过将 $\mathbf{x}_{FP}^{(2)}$ 旋转和平移来获得 $\mathbf{x}_{FP}^{(3)}$,图 1(b) 中描绘了这些样本解的正确对应,使用最小欧几里得距离会导致产生错误对应,如图 1(c) 所示。

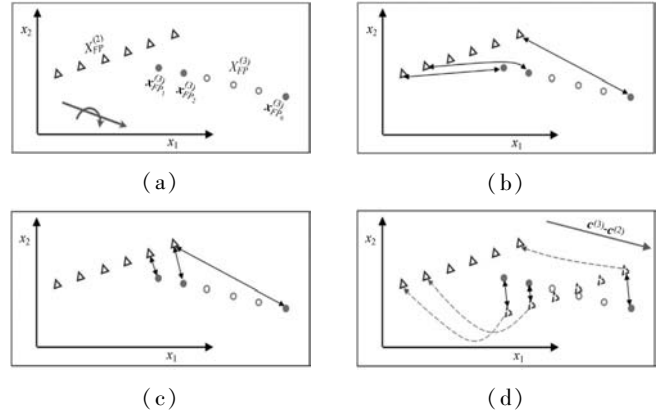


图 1 相应的流程

为了克服上述缺点,本文解决思路如下。首先,通过平移 $\hat{X}_{FP}^{(t-1)} = \{\hat{\mathbf{x}}_{FP_1}^{(t-1)}, \hat{\mathbf{x}}_{FP_2}^{(t-1)}, \dots, \hat{\mathbf{x}}_{FP_N}^{(t-1)}\}$ 来定义一组辅助解 $X_{FP}^{(t-1)}$,使 $\hat{X}_{FP}^{(t-1)}$ 的质心与 $X_{FP}^{(t)}$ 的质心重合:

$$\hat{\mathbf{x}}_{FP_i}^{(t-1)} = \mathbf{x}_{FP_i}^{(t-1)} + (\mathbf{c}^{(t)} - \mathbf{c}^{(t-1)}) \quad i=1,2,\dots,N \quad (2)$$

式中: $\mathbf{c}^{(t)}$ 是 $X_{FP}^{(t)}$ 中解的质心。然后使用这些辅助解 ($\hat{\mathbf{x}}_{FP_i}^{(t-1)}$) 代替 $\mathbf{x}_{FP_i}^{(t-1)}$ 来计算式(2)中的欧几里得距离,并确定每个 $\mathbf{x}_{FP_i}^{(t)}$ 的 $X_{FP}^{(t-1)}$ 中的对应解,相应的流程如图 1 所示。

1.2 目标空间中的相似度量指标

目标空间中最直观的相似性度量指标最小化归一化目标空间中的欧氏距离,对应 $\mathbf{x}_{FP_i}^{(t)}$ 的解决方案如下:

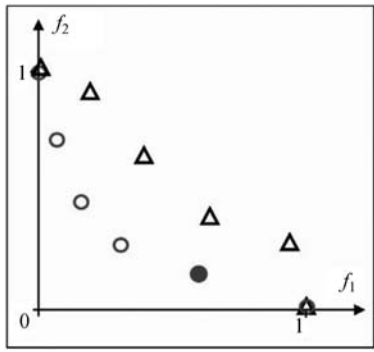
$$\mathbf{x}_{FP_i}^{(t-1)} = \arg \min_{\mathbf{x}_{FP_j}^{(t-1)}} (\|\tilde{\mathbf{f}}_{FP_i}^{(t)} - \tilde{\mathbf{f}}_{FP_j}^{(t-1)}\|_2) \quad i=1,2,\dots,N \quad (3)$$

式中: $\tilde{\mathbf{f}}_{FP_i}^{(t)}$ 和 $\tilde{\mathbf{f}}_{FP_j}^{(t-1)}$ 分别是 $\mathbf{x}_{FP_i}^{(t)}$ 和 $\mathbf{x}_{FP_j}^{(t-1)}$ 的归一化目标值。还可以基于参考方向定义相似性度量,根据该参考方向,相似性度量是连接原点到 $\tilde{\mathbf{f}}_{FP_i}^{(t)}$ 和 $\tilde{\mathbf{f}}_{FP_j}^{(t-1)}$ 的线段之间的平面角。使用点积可以很容易地计算平面角:

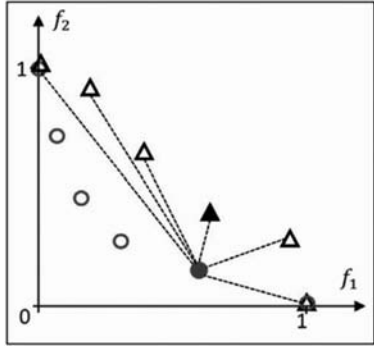
$$\beta_{ij} = \cos^{-1} \left(\frac{\|\tilde{\mathbf{f}}_{FP_i}^{(t)} \cdot \tilde{\mathbf{f}}_{FP_j}^{(t-1)}\|_2}{\|\tilde{\mathbf{f}}_{FP_i}^{(t)}\|_2 \|\tilde{\mathbf{f}}_{FP_j}^{(t-1)}\|_2} \right) \quad i,j=1,2,\dots,N \quad (4)$$

式中: β_{ij} 是平面角。根据该度量,对应 $\tilde{\mathbf{x}}_{FP_j}^{(t-1)}$ 的 $\tilde{\mathbf{x}}_{FP_j}^{(t-1)}$ 是使 β_{ij} 最小化的一个。

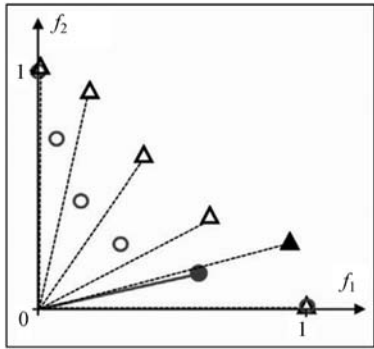
这两个基于目标的相似性度量如图 2(a) 所示。目标是给定 $\tilde{F}_{FP}^{(3)}$ (三角形) 和 $\tilde{F}_{FP}^{(4)}$ (圆圈) 中的指定解在 $\tilde{F}_{FP}^{(3)}$ 中找到相应的解。图 2(b) 为如何使用目标函数空间中的最小欧氏距离来确定相应的解:计算绿色实心圆和所有三角形之间的欧氏距离(虚线段),对应的三角形是距离最小的三角形。图 2(c) 为使用最小平面角相似度量确定的相应解决方案,平面角是将实心圆连接到原点的线段(实实线段)与其中一个三角形连接到原点的每个线段(虚线)之间的角度。可以看出,使用这种相似性度量得到的相应解与以前的不同。



(a) 目标的相似性度量



(b) 目标函数空间中的最小欧氏距离



(c) 最小平面角相似度量

图 2 相应的度量

2 信息共享策略

单模型预测方法分析 $X_{FP}^{(t)}$ 质心的运动,该质心对输入误差并不敏感。在多模型预测方法中,每个模型分析一小部分 $X_{FP}^{(t)}$ 的运动。由于每个模型的数据量很小,因此每个模型以及随后的整个多模型方法可以较好地检测出输入误差。在逐点预测方法实现过程中,每个模型只提供一个解。例如,基于再初始化的种群进化算法 (Prediction-based Population Re-initialization for Evolutionary, PRE) 的逐点预测方法,除去随机变化,计算预测的平移向量 $\tilde{\mathbf{v}}_i^{(t+1)}$ 的表达式为:

$$\mathbf{v}_i^{(t)} = \mathbf{x}_{FP_i}^{(t)} - \mathbf{x}_{FP_i}^{(t-1)} \quad (5)$$

$$\tilde{\mathbf{v}}_i^{(t+1)} = \mathbf{v}_i^{(t)} \quad i = 1, 2, \dots, N$$

这里的相似度量是决策变量空间中的欧氏距离。

时间步长 $t + 1$ 的解如下:

$$\mathbf{x}_{FP_i}^{(t+1)} = \mathbf{x}_{FP_i}^{(t)} + \tilde{\mathbf{v}}_i^{(t+1)} \quad i = 1, 2, \dots, N \quad (6)$$

图 3 展示了一种典型情况,其中使用排除随机变化的预测模型来初始化时间步长 $t + 1$ 的种群。此处变化包括平移和旋转。当没有输入误差时,预测方法可以在 $S^{(t+1)}$ 上初始化 $\mathbf{x}_{FP_i}^{(t+1)}$,如图 3(a) 所示;尽管如此,当甚至存在微小的输入误差时,如图 3(b) 所示, $\mathbf{x}_{FP_i}^{(t+1)}$ 的初始化与 $S^{(t+1)}$ 之间的距离会较远,从而会产生相当大的预测错误。

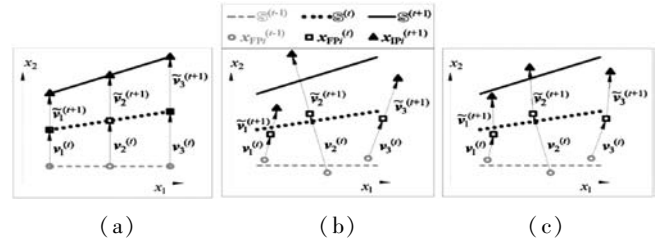


图 3 PRE-47 预测方法对输入误差的敏感性

为了提高逐点预测模型对输入误差的鲁棒性,提出了一种信息共享策略。根据该策略,每个预测的平移向量 $\tilde{\mathbf{v}}_i^{(t+1)}$ 为:

$$\mathbf{v}_i^{(t)} = \mathbf{x}_{FP_i}^{(t)} - \mathbf{x}_{FP_i}^{(t-1)}$$

$$\tilde{\mathbf{v}}_i^{(t+1)} = \frac{\sum_{j=1}^N w_{ij} \mathbf{v}_j^{(t)}}{\sum_{j=1}^N w_{ij}} \quad i = 1, 2, \dots, N \quad (7)$$

式中: $w_{ij} \geq 0$, 指定了 $\mathbf{x}_{FP}^{(t)}$ 提供的信息对 $\mathbf{x}_{FP_i}^{(t)}$ 的预测平移向量的贡献,在本文例子中是 $\mathbf{v}_j^{(t)}$ 。从逻辑上讲,权重的定义应确保更接近 $\mathbf{x}_{FP_i}^{(t)}$ 的解对 $\tilde{\mathbf{v}}_i^{(t+1)}$ 的贡献更大。

图 3(c) 描述了用来自相邻解的信息如何提高逐点预测对输入误差的鲁棒性。对于这种情况, $\tilde{\mathbf{v}}_1^{(t+1)} = (2/3)\mathbf{v}_1^{(t)} + (1/3)\mathbf{v}_2^{(t)}$, $\tilde{\mathbf{v}}_2^{(t+1)} = 0.25\mathbf{v}_1^{(t)} + 0.5\mathbf{v}_2^{(t)} + 0.25\mathbf{v}_3^{(t)}$ 和 $\tilde{\mathbf{v}}_3^{(t+1)} = (1/3)\mathbf{v}_2^{(t)} + (2/3)\mathbf{v}_3^{(t)}$ 。

2.1 权重

为了使得 $\mathbf{x}_{FP_i}^{(t)}$ 的解在 $\tilde{\mathbf{v}}_i^{(t+1)}$ 的计算中的贡献更大,选择式(8)中的权重。因为两个相互接近的解更有可能具有相似的转化向量。因此,应使用距离的降序函数为 $\mathbf{x}_{FP_i}^{(t)}$ 分配权重。存在两种极端情况:

(1) 最大程度地利用全局信息:在这种极端情况下, $w_{i1} = w_{i2} = \dots = w_{iN} = \frac{1}{N}$, 预测的平移向量 $\tilde{\mathbf{v}}_i^{(t+1)}$ 利用来自每个解的信息,而不管它们距离 $\mathbf{x}_{FP_i}^{(t)}$ 有多远。

(2) 最大程度地利用局部信息:在这种极端情况下,结果为 $\tilde{\mathbf{v}}_i^{(t+1)} = \mathbf{v}_i^{(t)}$ 。在这种情况下,预测的平移向量 $\tilde{\mathbf{v}}_i^{(t+1)}$ 仅利用来自 $\mathbf{x}_{FP_i}^{(t)}$ 的最近平移向量的信息,因为其他解决方案的性能较差,不能提供任何有用的

信息。

理想情况下,权重函数应该是一个参数函数,这样通过控制单个参数,就可以在这两种极端情况之间进行衡量。权重服从正态分布,其概率密度函数为:

$$\hat{w}_{ij} = \exp\left(-\left(\frac{d_{ij}}{\bar{d}}\right)^2\right), w_{ij} = \frac{\hat{w}_{ij}}{\sum_{j=1}^N \hat{w}_{ij}} \quad (8)$$

$$d_{ij} = \|\mathbf{x}_{FP_i}^{(t)} - \mathbf{x}_{FP_j}^{(t)}\|_2 \quad i, j = 1, 2, \dots, N$$

式中: $\bar{d} > 0$ 是一个基准距离,用于标准化解之间的距离。此函数具备三个特性:

(1) 数据越大,全局信息的利用率就越高。

(2) 对于 $\bar{d} \rightarrow 0$ 的极端情况,该方法最大程度地利用了局部信息。

(3) 对于 $\bar{d} \rightarrow \infty$ 的极端情况,该方法最大程度地利用了全局信息。

\bar{d} 与决策参数的搜索范围成正比是合理的。然而,POS可能位于搜索空间的一个小区域中。理想情况下,如果解是线性缩放的,则权重不应改变。为了实现这一目标,本文提出一种自适应方法。在每个时间段结束时,计算 $\mathbf{x}_{FP_i}^{(t)}$ 之间的所有成对距离,并将 \bar{d} 设为所有这些成对距离的 q_w 百分位数($0 \leq q_w \leq 100$),其中 q_w 是用户自定义的控制参数。如果 $\bar{d} = 0$,令 $\bar{d} = 10^{-10}$,以防止除数为0。

2.2 加权逐点预测

本文提出的预测方法采用加权逐点模式,并以信息共享策略和变化方向来强化。信息共享策略使用式(7)计算预测的平移向量,然后应用方向变化来改善总体的多样性:

$$\mathbf{x}_{IP_i}^{(t+1)} = \mathbf{x}_{FP_i}^{(t)} + N(1, \sigma_d^2) \times \tilde{\mathbf{v}}_i^{(t+1)} \quad i = 1, 2, \dots, N \quad (9)$$

式中: $N(1, \sigma_d^2)$ 是从正态分布中抽样的随机数,平均值为1,标准偏差为 σ_d ,本文令 $\sigma_d = 1$ 。如果 $\mathbf{x}_{IP_i}^{(t+1)} \notin \Omega$,则把它重新分配到 Ω 范围内最近的点。

算法1是一种新的预测方法,称为加权逐点预测法(WPPM)。当时间段 t 刚刚结束时,WPPM初始化 $X_{IP}^{(t+1)}$ 。

算法1 WPPM

输入:相似性策略 q_w ,搜索空间 Ω , $\tilde{F}_{FP}^{(t-1)}$, $\tilde{F}_{FP}^{(t)}$, $X_{FP}^{(t-1)}$, $X_{FP}^{(t)}$, $N = |X_{FP}^{(t)}|$ 。

输出: $\mathbf{x}_{IP_i}^{(t+1)}$, $i = 1, 2, \dots, N$ (时间步长的种群种子 $\#t + 1$)。

1. 计算每个 $\mathbf{x}_{FP_j}^{(t-1)} \in X_{FP}^{(t-1)}$ 和每个 $\mathbf{x}_{FP_j}^{(t)} \in X_{FP}^{(t)}$ 之间的欧氏距离
2. **if** 相似性度量在目标空间内 **then**
3. **if** 相似性度量是参考方向 **then**
4. 根据式(5)计算成对平面角 β_{ij}
5. **else**
6. 根据式(4)计算在每个 $\tilde{f}_{FP}^{(t-1)} \in \tilde{F}_{FP}^{(t-1)}$ 和每个 $\tilde{f}_{FP}^{(t)} \in \tilde{F}_{FP}^{(t)}$ 之

间的成对距离

7. **end**

8. **end**

9. **for** $i \leftarrow 1 : N$ **do**

10. 用本文定义的相似性策略找到 $\mathbf{x}_{FP_j}^{(t-1)}$,在 $X_{FP}^{(t-1)}$ 内的解符合 $\mathbf{x}_{FP_j}^{(t-1)}$

11. **end for**

12. 根据式(9),计算信息共享权重

13. **for** $i \leftarrow 1 : N$ **do**

14. 根据式(8),计算预测平移矩阵 $\tilde{\mathbf{v}}_i^{(t+1)}$

15. 根据式(10),计算 $\mathbf{x}_{IP_i}^{(t+1)}$

16. **if** $\mathbf{x}_{IP_i}^{(t+1)} \notin \Omega$ **then**

17. 将其重新定位到距离 Ω 最近的区域

18. **end**

19. **end for**

算法1中WPPM的时间复杂度计算步骤如下:

第1行计算成对欧几里得距离的复杂度为 $O(DN^2)$ 。

在归一化目标空间(第6行)中计算平面角(第4行)或成对距离,计算复杂度为 $O(MN^2)$ 。对于每个 $\mathbf{x}_{FP_i}^{(t)}$,找到相应的解(第10行)的计算复杂度为 $O(N)$ 。由于每个解都使用一个独立的模型,因此总体的计算复杂度为 $O(N^2)$ 。给定成对欧氏距离,计算信息共享权重(第12行)的复杂度为 $O(N^2)$ 。对于每个 $\mathbf{x}_{FP_i}^{(t)}$,计算预测的平移向量(第14行)的计算复杂度为 $O(ND)$ 。生成 $\mathbf{x}_{IP_i}^{(t+1)}$ (第15行和第17行)的计算复杂度为 $O(D)$ 。对所有解(第13-18行)执行所有这些操作的计算复杂度为 $O(DN^2)$ 。因此,所提出的预测方法的时间复杂度为 $O(DN^2)$ 。如果相似度量在目标函数空间中且满足 $M > D$, M 和 D 表示相应的复杂度系数,则时间复杂度为 $O(MN^2)$ 。

3 实验设置

3.1 测试问题

DMO检验问题在开发预测方法中起着关键的作用,其最主要的特征是POS的变化,其原因是当变化遵循简单的模式时,一些预测方法可能会有效,例如纯平移。当变化的模式复杂时,这些方法性能较差。

传统数学定义的DMO测试问题仅涉及帕累托最优集(Pareto Optimal Set, POS)的平移,或者至少一个纯平移过程可以描述变化模式的大多数方面。其中包括FDA测试,它是DMO研究领域中最常用的测试方法之一。对于此类问题,由于其鲁棒性,采用单个模型的预测方法实际上优于多模型预测方法。由于CEC2018测试方法集成了POS中简单和复杂的测试

问题。本文使用此测试方法进行数值分析和比较。

3.2 优化方法

对于连续问题,NSGA-III 使用模拟二进制交叉(SBX),其发生概率为 P_{cross} 。后代的传播由交叉指数(η_{cross})控制^[9]。使用具有突变指数 η_{mut} 的基于多项式的突变,以 P_{mut} 的概率进行突变。使用参考方向在父代和子代的结合上进行选择。NSGA-III 使用基于超平面的启发式算法估算优化过程中的理想点和最低点,以对目标值进行归一化。采用不同的启发式方法来估计理想点和最低点,这些点是秩等于或小于临界秩的所有目标值的最小值和最大值。临界秩定义为父代和子代按秩排序,取第 N 个解的秩,其中 N 是种群大小。这种方式修改的原因是 NSGA-III 中原来的超平面归一化方法可能会产生负交集。修改后的 NSGA-III 的其他参数设置如下: $N = 100$;SBX 中 $P_{\text{cross}} = 0.9$, $\eta_{\text{cross}} = 5$;基于多项式的突变, $P_{\text{mut}} = 1/D$, $\eta_{\text{mut}} = 5$, D 代表的是突变因子,因子越大,突变概率越小。

这些参数值通常选择较小的参数值,因为它们通过远离后代来提高探索能力。根据初步观察,这可能有助于解决 DMO 问题,因为它有助于种群在每次变化后迅速进行多样化。 N 的值大致等于所需的权衡解数量,权衡解数量是由决策者或问题设置确定的参数。 $N = 100$ 的选定值符合 CEC2018 DMO 竞赛中规定的实验设置,这表明权衡解决方案的数量应设为 100。

3.3 性能分析

作为一种替代方法,本文使用 IGD + 指标来定义 DMO 方法的平均 IGD +。IGD + 与 IGD 相似,只是用一个新的距离度量指标替换了欧几里得距离。设 $P = \{p_1, p_2, \dots, p_{|P|}\}$ 是均匀分布在 POF 上的一组参考点,其中符号 $|\cdot|$ 为计算一组的基数。最小化目标值问题为 $F = \{f_1, f_2, \dots, f_{|F|}\}$ 的非支配解集的 IGD + 的定义如下:

$$IGD^+(F, P) = \frac{1}{|P|} \sum_{i=1}^{|P|} \min_{f_j \in F} (f_j, p_i) \quad (10)$$

$$d(f_j, p_i) = \sqrt{\sum_{k=1}^M (\max\{f_{jk} - p_{ik}, 0\})^2}$$

式中: f_{jk} 是解决方案 j 的第 k 个目标值; p_{ik} 是第 i 个参考点的第 k 个坐标。

此处小修改使 IGD + 成为一个弱 Pareto 兼容指标,同时保持了其计算效率。考虑到 IGD + 的优势,本文提出将 IGD + 作为 DMO 方法性能的评估:

$$M_{\text{IGD}}^+ = \sum_{t=1}^{N_{\text{TS}}} IGD^+(F^{(t)}, P^{(t)}) \quad (11)$$

式中: $P^{(t)}$ 是 $F^{(t)}$ 上均匀分布点的子集; N_{TS} 是时间步数。在计算 IGD + 时,将真实理想点和最低点进行归一化。本文排除了式(11)中第 0 个时间段的结果,因为预测方法在第一次改变之前对优化过程没有影响。与传统的 MIGD 不同, $MIGD^+$ 是弱 Pareto 兼容的,而这两个指标在计算上都是有效的^[10]。本文利用这一指标对 DMO 进行性能评估和比较。

4 实验结果分析

4.1 相似性度量的影响

考虑两种情况:(1) 理想情况下,假设 SMO 能够准确地识别前一时间步中的 POS。这意味着对于时间步长 $(t+1)$, $S^{(0)}, S^{(1)}, \dots, S^{(t)}$ 的解的重新初始化,将 POS 运用到预测过程中。(2) 实际情况下,其中每个时间段结束时的总体 ($X_{FP}^{(t)}$) 用作实际 POS 的近似值。

理想情况通过提供模型的实际位置来抑制输入误差。虽然实际情况并非如此,但可以建立一种预测方法在处理问题变化中的复杂情况。即使在理想情况下,所提供的 POS 在变量空间中的分布也是不均匀的,因为 SMO 的目标在目标函数空间中具有最大化多样性。

为了探索上文讨论的相似性度量的影响,本文测试了 WPPM 四种变体的性能:

(1) WPPM-X: 采用决策变量空间中的欧氏距离作为相似度量。

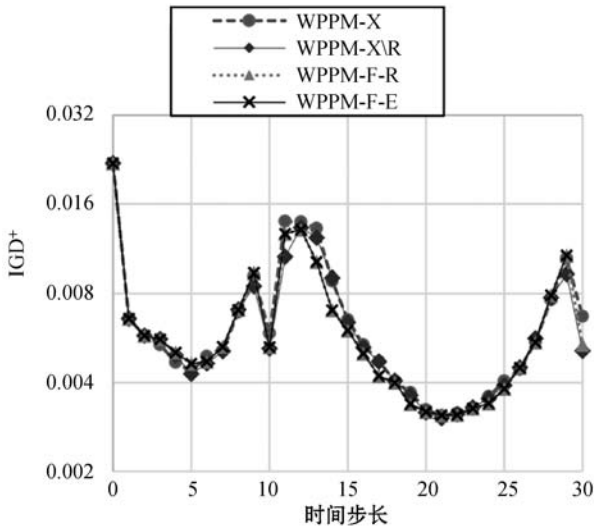
(2) WPPM-X\R: 它类似于 WPPM-X,但在找到相应的解决方案之前并不对 $X_{FP}^{(t-1)}$ 进行重新定位。

(3) WPPM-F-R: 采用归一化目标空间中的参考方向作为相似性度量。

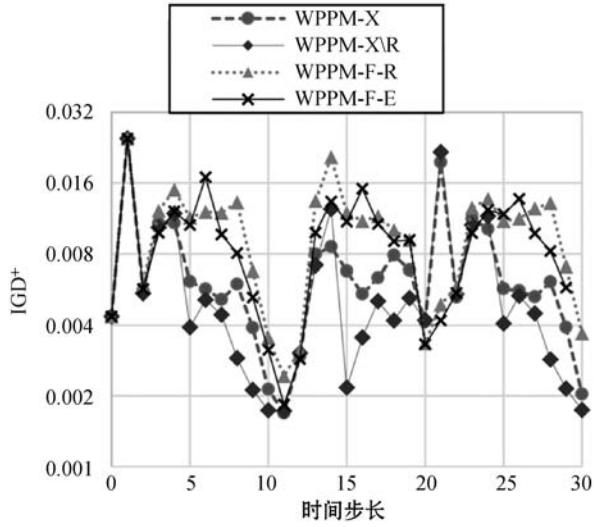
(4) WPPM-F-E: 采用归一化目标空间中的欧氏距离作为相似性度量。

在 $q_w = 0$ 的理想情况下对这些变体进行了测试,以排除信息共享策略的影响。在 10 维空间中考虑了两个测试问题 DF7 和 DF8。在这两个问题中,POS 生成一条曲线。主要区别在于 DF7 的 POF 形状变化不大,而 DF8 的 POF 形状变化较大,特别是曲率变化较大。

在 DF7 和 DF8 的每个时间段结束时计算的 IGD + 如图 4 所示,将实验运行 30 次,并求平均值。对于 DF7,所有四种变体在某些时间段上都表现出相似的性能。相比之下,对于 DF8,这种相似性度量是最差的选择。



(a) DF7

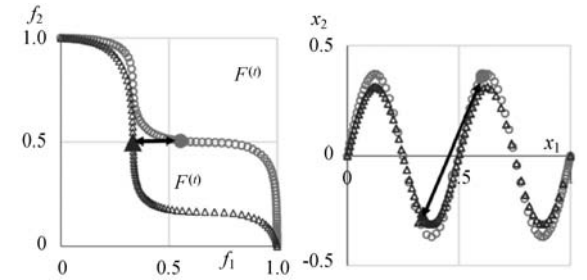


(b) DF8

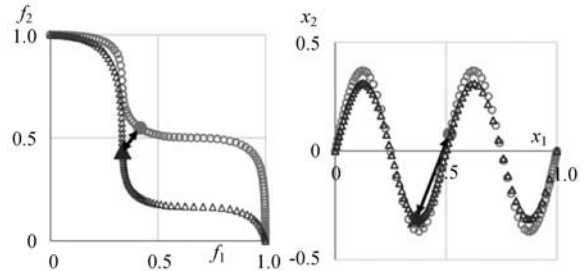
图4 超过30次独立运行IGD平均值

对于DF8, WPPM-X\R 的性能比 WPPM-X 要好。对于此问题, POS 流形的质心不会改变, 但是 WPPM-X 会计算 $X_{FP}^{(t)}$ 的质心的非零平移, 因为 $X_{FP}^{(t)}$ 中的解的分布会改变。尽管图1中指明该策略通常是有益的, 但在该特定情况下, 重定位策略会对算法性能产生不良影响。

DF8 基于目标的相似性度量失败的主要原因是 POF 的曲率和形状的显著变化。为了深入探讨这个问题, 进一步分析了二维 DF8 的时间步长的对应关系, 图5说明了当 $t=5$ 时使用基于目标的相似性度量确定的对应关系。黑色双箭头线段表示 $X_{FP}^{(t-1)}$ 中的对应解, 对于 $X_{FP}^{(t-1)}$ 中的选定解, 使用它们的标准化值。在决策变量空间中检验这些解, 可以发现 WPPM-F-E 和 WPPM-F-R 在决策变量空间中的解差距较大。这是基于目标的相似度量的一个基本缺点: 目标空间中两个解的相似性不能保证它们在变量空间中的相对相似性; 然而, 预测模型是应用于解决方案, 而不是处理对应的值。



(a) WPPM-F-E



(b) WPPM-F-R

图5 相似性度量确定的对应关系

4.2 信息共享策略性能

为了验证利用相邻解决方案的好处, 用不同的 q_w 值对 WPPM 进行了测试。图6说明了使用 WPPM-X 在 DF3、DF10 和 DF12 的理想和实际场景中, 作为 q_w 函数的 MIGD+。该图说明在理想情况下, 增加 q_w 可能会降低性能, 相比之下, 在实际情况下, 增加 q_w 会显著提高 MIGD+ 的性能, 直到 $q_w=10$, 尽管 $q_w=5$ 之后, 改善百分比会迅速下降, 这表明所提出的信息共享策略可以有效地提高点式方法对输入误差的鲁棒性。理想情况下 MIGD+ 与实际情况下 MIGD+ 的差异说明了输入误差对预测误差的影响。

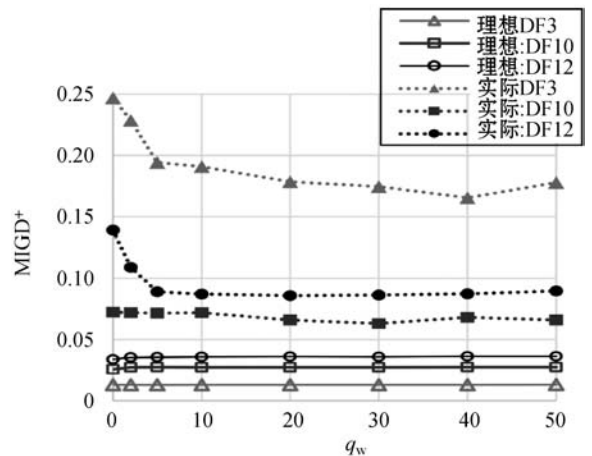


图6 理想和实际场景中 MIGD+

5 与现有预测方法的比较

5.1 性能比较

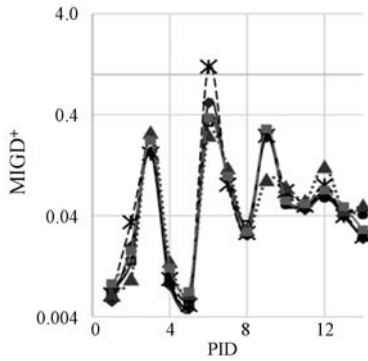
本节将 WPPM-X 与最近一些热门的预测方法进行

比较。本节不使用与问题相关的参数调整。令 $q_w = 10$, 通过以下预测方法进行比较: (1) HM: D-NSGA-II-B 中采用的基于超突变的重新初始化策略。(2) P-SGEA: 在文献[11]中提出的 SSGEA 预测方法。(3) DDM: 在文献[12]中提出的基于分解的差分模型。(4) PRE: 在文献[13]中搭建预测方法。(5) MDP: 在文献[14]中搭建多向预测方法。

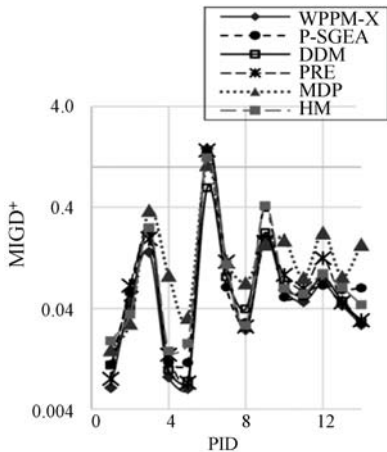
采用了 CEC2018 DMO 测试,用于 DMO 的 CEC2018 测试方法的实验设置使用两个值 10 和 30 表示变化频率(τ_i),而变化严重性参数(n_i)与变化的显著性成反比,设为 10,令 $\tau_i \times n_i = 200$,并将 τ_i 设为不同值。所有问题都是决策变量 10 维空间。表 1 给出了本文中考虑的变化严重性的三个级别。第一个变化发生在第 50 代之后,每次模拟用不同的随机种子,并重复 30 次,且计算 MIGD+。这 30 次运行的平均 MIGD+ 如图 7 所示。

表 1 变化严重性的三个级别

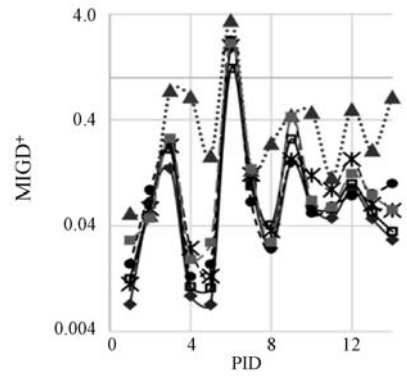
变化程度	τ_i	n_i	N_{change}	$N_{generations}$
严重程度	40	5	15	650
中等程度	20	10	30	650
最小程度	10	20	60	650



(a) 严重变化



(b) 适度变化



(c) 微小变化

图 7 30 次运行的平均 MIGD+

本文采用了一个基于分数的指标来比较每种方法的总体性能。首先,使用显著性水平为 0.01 的 Wilcoxon 秩和检验来检测某一种方法在特定问题上是否优于另一种方法。如果方法 M1 在问题 P1 上的统计性能优于方法 M2,则方法 M1 获胜,方法 M2 失败,否则认为是平局。对于每个问题,方法的得分为获胜次数加平局次数的一半。每种方法和每种变化严重程度等级的得分如图 8 所示。最后,每种方法的平均得分是所有 14 个测试问题的平均得分,如表 2 所示。

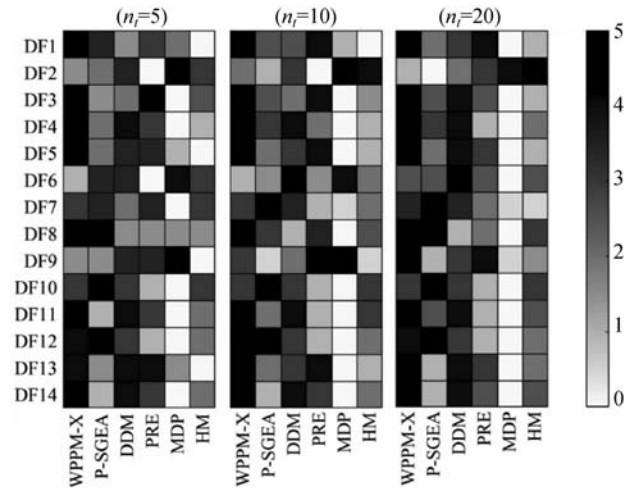


图 8 每种方法和每种变化严重程度等级的得分

表 2 每种方法对每种变化的平均严重程度

方法	变化程度			总体
	严重	中等	最小	
WPPM-X	3.71	4.04	4.18	3.98
P-SGEA	2.68	2.54	2.64	2.62
DDM	3.07	3.07	3.39	3.18
PRE	2.46	2.46	2.32	2.42
MDP	1.43	1.07	0.36	0.95
HM	1.64	1.82	2.11	1.86

如表 2 所示,当考虑所有这 14 个测试问题的总体性能时,WPPM-X 在每一个变化严重性方面都优于其

他方法。当然,对于一些问题,例如 DF2 和 DF6,其他方法也表现出了良好的性能。对比而言次优方法是 DDM。这两种方法的共同特点是保守地使用随机变量。WPPM-X 将随机变量限制在特定的方向上,DDM 完全排除了随机变量。这两种方法性能均较好,结果表明,较高的随机变化可能会对动态变化中具有某些变化方式的问题造成不利影响。相反,采用强等向性随机变化的 MDP 是最差的方法。另外,尽管 DNSGA 证明了某些解的随机变化是有益的,这是 DMO 最早的重新初始化方法之一,但该方法不进行任何预测,而是将一些随机变化应用于已经结束的时间步长的最终解。

在少数情况下,如 DF2 和 DF9 有严重和中度改变时,MDP 是性能最好的方法。这两个测试问题的变化模式很难被测试方法捕捉到。因此,对于这些有严重和中度变化的问题,MDP 的高随机变异是有益的。

当变化不太严重以及较严重时,WPPM-X 比其他测试方法更优越。这一观察结果表明,高度利用变化模式的预测方法在发生微小变化的问题中可能更为有利,即使变化十分频繁。与其他测试方法相比,MDP 应用了更强的随机变化,这减少了关于变化模式的可用信息,有利于使用初始化解方案。

图 7 显示,即使 MIGD + 值之间的差距在统计上是显著的,但这种差距并不显著。这样的原因是所采用的 SMO 和变化检测方法对于所有方法都是相同的。因此,只有预测方法才会导致差异。这种方法的优点是性能差异完全可以归因于预测方法。此外,当变化变得更激烈时,测试方法的性能之间的差异变得不那么显著,并且为每个时间步长分配了更高的评估预算。其原因是更严重的变化会降低预测方法的准确性,同时,每个较大的时间步长的评估预算为 DMO 提供了更多的机会来补偿初始化错误的种群。

5.2 时间性能比较

本节进行数值模拟分析,从而比较测试预测方法所需的计算时间,重点是比较实际时间,而不是忽略任何比例常数的大 O 方法。在进化 DMO 的背景下,模拟了 D 和 N 的合理范围。因此,本文关注的范围仅限于 $D \leq 100$ 和 $N \leq 1\,000$ 具有两个或三个目标的问题。

本文采用经过检验的预测方法对两个目标 DF3 和三个目标 DF10 的 30 个变化重新初始化总体,并计算除前两个变化外所有变化重新初始化总体的平均时间。实验的硬件配置为 Intel(R)Core(TM)i7-3770 CPU 和 16 GB RAM 的台式计算机,编程环境为 MATLAB 2017b。所比较的预测方法的实际时间性能如图 9

所示。

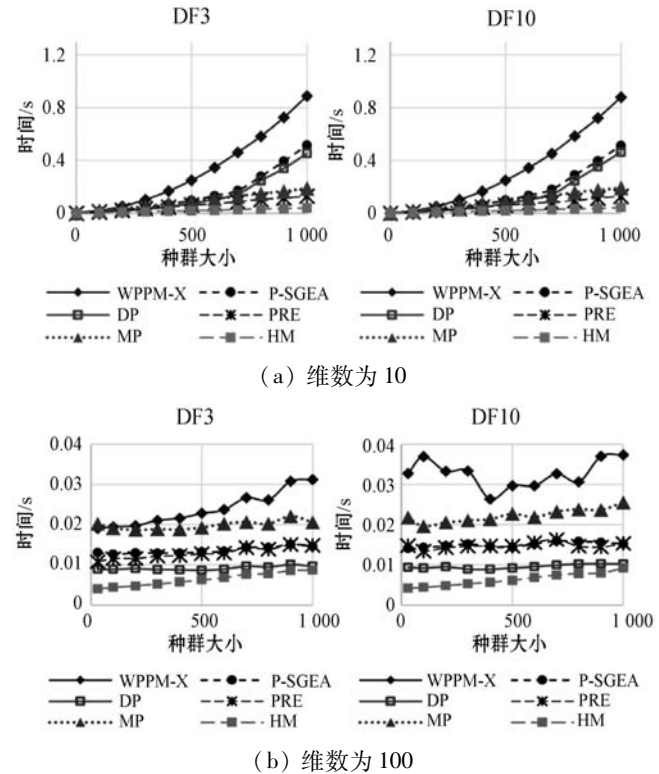


图 9 实际时间复杂度

对于所有测试方法,重新初始化所需的时间都相当小。例如,对于 $D = 10$ 和 $N = 100$,所有测试方法的重新初始化所需时间都小于 0.04 s,这包括存储初始化的解决方案所需的时间。

MP、PRE 和 HM 的时间复杂度与 N 成正比,而对于 DB、P_SGEA 和 WPPM-X,大致与 N^2 成正比关系。问题维数对重新初始化时间影响不大,这是因为问题维数只影响计算解的距离所需的时间。数值分析表明,在常用的 D 和 N 范围内,所有测试方法的计算量都很小。

6 结 语

为了有效实现 Pareto 最优解,并且提升对于输入误差的鲁棒性,提出一种基于信息共享策略的加权逐点预测动态多目标优化算法。通过分析 EC2018 DMO 测试套件的实验结果可以得出如下结论:

(1) 本文方法能够有效实现 Pareto 最优解,并且无论变化的严重程度如何。与现有的预测方法相比,WPPM 和 PRE 在捕捉不同类型的变化模式方面有更大的灵活性以及更高的精度,因为它们对每个解都采用了单独的模型。

(2) 加权逐点预测方法的附加控制参数允许在逐点预测方法的灵活性和单模型预测方法的鲁棒性之间

进行任意权衡,从而实现保证预测精度与误差鲁棒性。

(3) 较高的随机变化可能会对动态变化中具有某些变化方式的问题造成不利影响,因此,对于这些有严重和中度变化的问题,本文方法的高随机变异是有益的。

参 考 文 献

[1] 王鹤. 基于禁忌搜索的数字微流控生物芯片多目标综合优化算法[J]. 计算机应用与软件, 2019, 36(8): 325 - 333.

[2] 吴耕锐, 郭三学, 吴虎胜, 等. 改进多目标蚁群算法在动态路径优化中的应用[J]. 计算机应用与软件, 2019, 36(5): 249 - 254, 288.

[3] 刘若辰, 李建霞, 刘静, 等. 动态多目标优化研究综述[J]. 计算机学报, 2020, 43(7): 1246 - 1278.

[4] Azzouz R, Bechikh S, Said L B, et al. Handling time-varying constraints and objectives in dynamic evolutionary multi-objective optimization[J]. Swarm and Evolutionary Computation, 2018, 39(6): 222 - 248.

[5] Blank J, Deb K, Roy P C. Investigating the normalization procedure of NSGA-II[C]//International Conference on Evolutionary Multi-Criterion Optimization, 2019: 229 - 240.

[6] Cao L, Xu L H, Goodman E D, et al. Decomposition-based evolutionary dynamic multi-objective optimization using a difference model[J]. Applied Soft Computing, 2019, 76(6): 473 - 490.

[7] Han H G, Zhang L, Liu H X, et al. Multi-objective design of fuzzy neural network controller for wastewater treatment process[J]. Applied Soft Computing, 2018, 67(6): 467 - 478.

[8] Jiang M, Qiu L M, Huang Z Q, et al. Dynamic multi-objective estimation of distribution algorithm based on domain adaptation and nonparametric estimation[J]. Information Sciences, 2018, 435(8): 203 - 223.

[9] Liu X F, Zhou Y R, Yu X. Cooperative particle swarm optimization with reference-point-based prediction strategy for dynamic multi-objective optimization[J]. Applied Soft Computing, 2020, 87(6): 988 - 1001.

[10] 谭玮, 邱启仓, 俞维, 等. 一种基于邻域改进的分解多目标进化算法[J]. 小型微型计算机系统, 2020, 41(12): 2543 - 2549.

[11] 陈伟龙, 陈春良, 刘彦, 等. 抢修任务多目标动态调度的动态驱动策略研究[J]. 计算机测量与控制, 2018, 26(4): 258 - 263, 297.

[12] Rong M, Gong D W, Zhang Y, et al. Multidirectional prediction approach for dynamic multi-objective optimization problems[J]. IEEE Transactions on Cybernetics, 2018, 49

(7): 3362 - 3374.

[13] 周兵凯, 杨晓峰, 张智, 等. 能量路由器中双有源桥直流变换器多目标优化控制策略[J]. 电工技术学报, 2020, 35(14): 3030 - 3040.

[14] 程文, 孙树敏, 李宝, 等. 考虑需求响应的多目标模糊机会约束动态经济调度[J]. 科学技术与工程, 2020, 20(31): 12849 - 12856.

(上接第 262 页)

[9] Kim K J, Kim P K, Chung Y S, et al. Multi-scale detector for accurate vehicle detection in traffic surveillance data[J]. IEEE Access, 2019, 7: 78311 - 78319.

[10] 陈毅, 张帅, 汪贵平. 基于激光雷达和摄像头信息融合的车辆检测算法[J]. 机械与电子, 2020, 38(1): 52 - 56.

[11] 魏子洋, 赵志宏, 赵敬娇. 改进 Faster R-CNN 算法及其在车辆检测中的应用[J]. 应用科学学报, 2020, 38(3): 377 - 387.

[12] Tian Z, Shen C H, Chen H, et al. FCOS: Fully convolutional one-stage object detection[C]//IEEE International Conference on Computer Vision, 2019: 9627 - 9636.

[13] Bodla N, Singh B, Chellappa R, et al. Soft-NMS—improving object detection with one line of code[C]//IEEE International Conference on Computer Vision, 2017: 5561 - 5569.

[14] Tychsen-Smith L, Petersson L. Improving object localization with fitness NMS and bounded IoU loss[C]//IEEE Conference on Computer Vision & Pattern Recognition, 2018: 6877 - 6885.

[15] Rezatofighi H, Tsoi N, Gwak J Y, et al. Generalized intersection over union: A metric and a loss for bounding box regression[C]//IEEE Conference on Computer Vision & Pattern Recognition, 2019: 658 - 666.

[16] Zheng Z H, Wang P, Ren D W, et al. Enhancing geometric factors in model learning and inference for object detection and instance segmentation[EB]. arXiv:2005.03572, 2020.

[17] Zheng Z H, Wang P, Liu W, et al. Distance-IoU loss: Faster and better learning for bounding box regression[J]. Proceeding of the AAAI Conference on Artificial Intelligence, 2020, 34(7): 12993 - 13000.

[18] Lin T Y, Dollar P, Girshick R, et al. Feature pyramid networks for object detection[C]//IEEE Conference on Computer Vision and Pattern Recognition, 2017: 2117 - 2125.

[19] Li D, Yao A, Chen Q F. PSConv: Squeezing feature pyramid into one compact Poly-scale convolutional layer[EB]. arXiv:2007.06191, 2020.

[20] Yu F H, Wang D Q, Shelhamer E, et al. Deep layer aggregation[C]//IEEE Conference on Computer Vision and Pattern Recognition, 2018: 2403 - 2412.