

三维深度学习网络的几何差异感知能力

许煜濠¹, 刘石坚¹, 康朝明¹, 吴连杰¹, 邹峥²

(1. 福建省大数据挖掘与应用技术重点实验室, 福建 福州 350118;

2. 福建师范大学 计算机与网络空间安全学院, 福建 福州 350117)

摘要: 在使用深度学习技术处理三角网格等三维数据时, 如果网络不具备感知数据位置、朝向、尺寸等几何属性差异的能力, 可能导致模型泛化能力不足、准确率偏低的后果。为解决该问题, 在变换网络 T-Net 的基础上, 提出名为几何差异感知 (geometric difference perception, GDP) 的网络模块。其核心思想是通过多样化的样本训练, 学习到一个变换矩阵, 对高维特征进行规范化。通过以牙齿网格分割为任务的多项实验表明, GDP 能够有效应对三维数据的几何差异问题, 避免其对模型性能造成的不良影响, 对于网格分割等三维任务性能的提升具有重要意义。

关键词: 三角网格; 深度学习; 几何变换; 牙齿分割

中图分类号: TP301

文献标志码: A

文章编号: 1672-4348(2023)06-0592-06

Geometric difference perception capabilities of 3D deep learning networks

XU Yuhao¹, LIU Shijian¹, KANG Chaoming¹, WU Lianjie¹, ZOU Zheng²

(1. Fujian Provincial Key Laboratory of Big Data Mining and Applications, Fuzhou 350118, China;

2. College of Computer and Cyber Security, Fujian Normal University, Fuzhou 350117, China)

Abstract: When using deep learning techniques to process three-dimensional data like triangular meshes, if the network cannot perceive the differences of geometric properties, such as position, orientation and size of the data, the trained model may have poor generalization ability and low accuracy. To solve this problem, a network module called geometric difference perception (GDP) is proposed based on the transformation network named T-Net. The core idea of GDP is to learn a transformation matrix through training with diverse samples and use it to standardize high-dimensional features. Multiple experiments regarding tooth mesh segmentation tasks demonstrate that GDP can effectively address the geometric differences of three-dimensional data, and avoid its adverse effects on the performance of the model, which is of great significance for improving the performance of 3D tasks such as mesh segmentation.

Keywords: triangular meshes; deep learning; geometric transformation; tooth segmentation

位置、朝向、尺寸是点云、三角网格等三维 (three dimensional, 3D) 数据的基本几何属性, 其差异性普遍存在。例如, 图 1(a) 中的两个牙颌网格数据的位置和朝向即存在明显差别。使用深度学习技术处理 3D 数据时, 如果网络模型不具备几何差异感知能力, 则可能导致泛化能力不足、准确率偏低的后果。

通常有两种策略解决上述问题: 修正数据或者修正模型。修正数据是通过人为干预, 消除几何差异, 避免网络模型面临该挑战。具体方法是:

收稿日期: 2023-11-10

基金项目: 福建省科技厅自然科学基金项目 (2022J01932); 福建省教育厅科研项目 (JAT210283; JAT220052)

第一作者简介: 许煜濠 (1999—), 男, 福建泉州人, 硕士研究生, 研究方向: 深度学习, 网格处理等。

通信作者: 刘石坚 (1983—), 男, 湖南长沙人, 副教授, 博士, 研究方向: 计算机图形学, 深度学习, 机器视觉等。

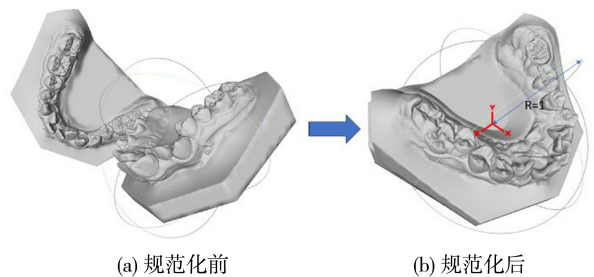


图 1 牙颌网格数据规范化

Fig.1 Standardization of dental meshes

将数据规范化至统一的位置、朝向、尺寸,再进行训练或预测。例如,图 1(b)即为对图 1(a)中数据进行规范化处理后的效果。修正数据方法的局限性在于:(1)手工交互难以避免,耗时费力;(2)并非从源头解决模型的几何感知问题。修正模型的思想则是着眼于模型网络结构优化,使其从多样化的数据中学习感知几何差异的能力。

本课题采取修正模型的策略,在分析已有相关研究的基础上,提出名为“几何差异感知(geometric difference perception,GDP)”的模块,以提升模型的几何差异感知能力;以牙齿分割为应用背景,通过主流模型进行对比实验,验证 GDP 的可行性和有效性。

1 相关工作

1.1 点云数据

由于点云的复杂性比三角网格要低,当深度学习方法从二维应用场景(例如处理图像数据)扩展到三维时,前期研究多关注其在点云数据上的应用。例如,Qi 等人^[1]所提出的 PointNet 即为其中的典型代表,用于实现点云分类和分割。

在 PointNet 中,作者提出名为变换网络(transform net,T-Net)的局部结构来解决输入模型的点云特征以及高维特征的对齐问题。如图 2 所示,当通道数为 3(即三维坐标)的点云数据输入网络之后,将首先进行一个输入变换。该变换中的 T-Net 网络会输出一个 3×3 的变换矩阵,与原始数据进行矩阵相乘。其本质即使用上述几何变换,对输入点云进行规范化。

此外,网络中还有另一个包含 T-Net 结构的特征变换,第二个 T-Net 与第一个的区别在于:其输入为通道数等于 64 的高维特征,输出为 64×64 的变换矩阵,即实现高维特征空间中的规范化。

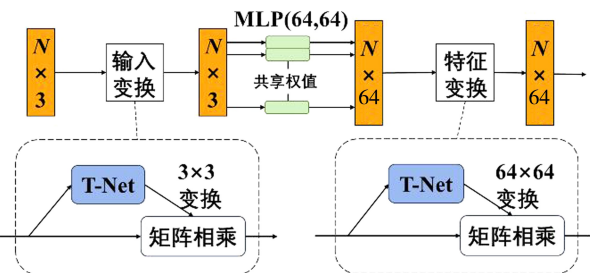


图 2 PointNet 中的 T-Net 迷你网络

Fig.2 T-Net mini networks in PointNet

文献[2]所提出的 DGCNN 模型同样采用点

云变换模块来估计仿射变换矩阵,将输入点集对齐到一个规范空间。与 PointNet 不同的是,DGCNN 弃用了高维特征变换模块。

为便于区分,本研究将作用于低维特征的 T-Net 记为 T-Net- I ,作用于高维特征的 T-Net 记为 T-Net- II 。

1.2 网格数据

在对三角网格进行深度学习时,目前主流的方法是以面片为样本基本单元进行处理。由于一个面片可以通过 3 个顶点和 1 个中心点共 4 组特征予以表征,在使用三维坐标作为原始特征的情况下,一个样本的特征尺度则为 $N \times 12$ (N 为网格数据的面片数)。如果再加上法向量特征,则特征尺度将变为 $N \times 24$ 。TSGCN^[3]、MGFL^[4]是这类方法的代表。由于使用比点云方法更多的特征数据,基于面片的方法通常性能更优,但计算资源需求更大。为此,它们仅使用 T-Net- I 来应对输入数据的几何差异问题,从而降低计算量。

1.3 T-Net 模块

综上可知,在应对数据几何差异挑战时,已有方法的核心思想主要是通过监督学习得到一个变换矩阵,将特征数据对齐到一个规范空间。具体是采用 T-Net 迷你网络实现,区别在于 T-Net 的数量和位置,详见表 1。表 1 中 PointNet ++^[5]是 PointNet 的升级版,它与基于面片的网格方法 GACNet^[6]一样,都没有采用相关措施。若输入数据存在几何差异,则 PointNet ++ 和 GACNet 均无法达到理想的预测效果。

表 1 相关方法信息表			
Tab.1 Information of related methods			
方法	来源	发表时间	几何差异问题 解决方案
PointNet	CVPR	2017	T-Net- I 和 T-Net- II
DGCNN	TOG	2019	T-Net- I
TSGCN	TMI	2021	T-Net- I
PointNet++	NIP	2017	无
GACNet	PRL	2021	无

图 3 展示了 T-Net- I 在 TSGCN 中的网络细节。图中虚线框内的结构即为 T-Net 网络,它由 3 个卷积层、1 个最大池化层和 3 个全连接层依次

拼接而成。就特征数据而言,其输入为 $N \times 12$ (4 组 $\times 3$ 通道/组 = 12 通道) 的原始特征。输入数据首先通过 3 个卷积层,通道数逐步提升至 64、128 和 512;然后通过最大池化操作,特征尺寸变为 1×512 ;最后,通过 3 个全连接层的作用,输出为一个 12×12 的仿射变换矩阵。将该矩阵与输入特征相乘,即可实现特征的规范化。

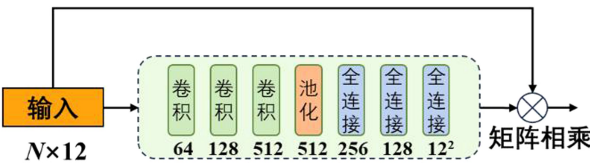


图 3 TSGCN 中的 T-Net 模块细节
Fig.3 Details of T-Net adopted in TSGCN

2 几何差异感知模块及应用范例

本研究的动机是基于以下两个问题:(1) 是否有必要使用多个 T-Net 结构;(2) 如果出于计算成本考虑只选用一个 T-Net,要如何进行配置。

2.1 GDP 模块

GDP 模块的结构细节如图 4 所示,其核心思想是:使用 K 个维度为 L 的多层感知机 (multilayer perceptron, MLP, 如图中虚线框所示) 将 $N \times L$ 的低维特征转化为 $N \times H$ 的高维特征之后,再输送给后续的 T-Net 网络,其中 N 为样本基本单元的个数, $L < H$ 。

GDP 本质上是一个作用于高维特征的 T-Net。相较图 3 中的 T-Net 结构,GDP 中的 T-Net

同样由三层卷积、一层池化、三层全连接层组成,区别在于:最后一层的维度为 H^2 ,即确保输出为 $H \times H$ 的变换矩阵,从而实现 $N \times H$ 尺寸高维输入特征的规范化。

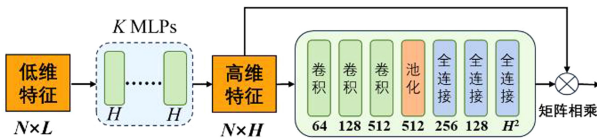


图 4 GDP 模块
Fig.4 GDP module structure

2.2 牙齿分割应用

3D 牙颌网格数据记录了患者口腔中牙齿的形态信息,通过深度学习技术将单颗牙齿从牙颌网格数据中准确分割出来,对于牙齿疾病诊断、手术规划等具有重要意义^[7]。

由于从不同渠道采集获取的牙颌网格数据,其尺寸、朝向、位置存在一定差异,为了验证 GDP 的有效性,选择未考虑几何差异问题的牙齿分割模型 GACNet^[6] 作为改进范例。

GACNet 是一种基于面片策略处理网格数据的深度学习方法。如图 5 所示,该网络整体上是一个双分支结构,其中包含 3 个图注意力卷积 (graph attentional convolution, GAC) 的分支用于局部特征提取,另一分支则用于全局特征提取。以牙颌网格数据上提取到的 $N \times 24$ 的原始特征作为输入,GACNet 最终将给出 $N \times C$ 的预测结果,其中 N 为面片数, C 表示预设的类别数。

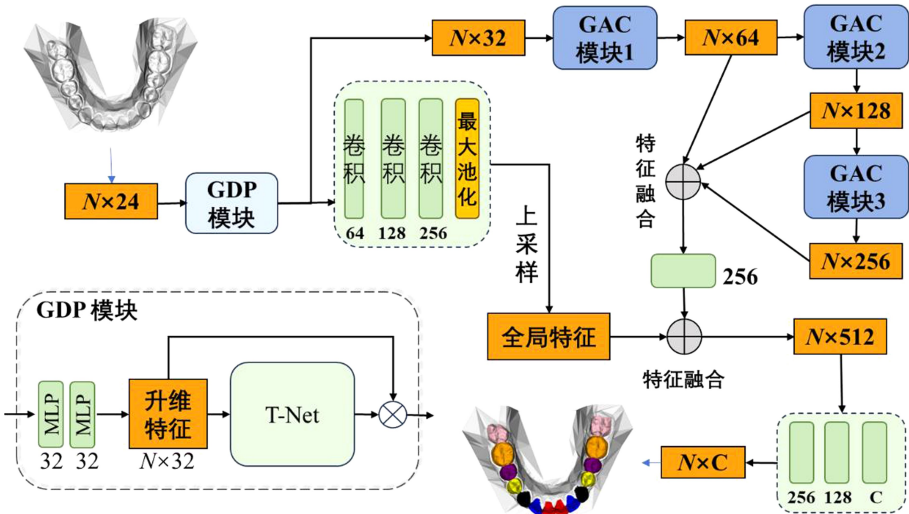


图 5 具有 GDP 的 GACNet 改进网络结构展示
Fig.5 Demonstration of improved GACNet with GDP

按照 2.1 节的思路,将 GDP 模块(细节见图中虚线框)放置于 $N \times 24$ 的输入特征之后,分支结构之前。经实验表明,当 K 取 2、 H 取值 32 时,性能提升效果最佳。

3 实验与讨论

3.1 数据集及实验环境

实验所用的数据来自 MICCAI 挑战赛的公共数据集 3DTeethSeg^[8],每个原始牙颌网格数据由 100 000~300 000 个数量不等的面片组成。为了便于训练,将数据统一为 17 000 个面片数的规模。

由于训练样本的几何差异性越丰富,模型的泛化能力越高,因此对网格数据进行包含随机角度旋转和随机坐标位移的数据扩充操作。扩充后的训练集规模为 2 058,验证集和测试集分别为 588 和 294。

所有实验均运行于一台显卡为 NVIDIA GeForce RTX 3090 (24 GB),CPU 为 Intel Core i9-10920X (3.50 GHz) 的台式计算机上。共训练 120 个 epoch,训练时采用 Adam 优化器,batch_size 设为 2,学习率初始为 1×10^{-3} ,每 20 个 epoch 进行 0.5 倍衰减。

3.2 评价指标

分割性能通过 3 个指标进行定量评估,分别为:Dice 相似系数(DSC)、灵敏度(SEN)和正预测值(PPV)。令 TP、TN、FP、FN 分别表示真阳性、真阴性、假阳性和假阴性,DSC、SEN 和 PPV 的计算方法分别如公式(1)(2)和(3)所示。

$$DSC = \frac{2TP}{FP + 2TP + FN}$$

(1)

$$SEN = \frac{TP}{TP + FN}$$

(2)

$$PPV = \frac{TP}{TP + FP}$$

(3)

上述 3 个指标值与分割性能成正比,即数值越高代表性能越好,区别在于侧重点存在差异。通过对这些差异性进行评估,可以确保模型在不同数据集划分和数据增强条件下都能够稳健地完成任务目标。

3.3 对比实验

为了验证 GDP 的效果,使用原始 GACNet 模

型、加入 GDP 模块后的 GACNet 改进模型(记作 GACNet-GDP)、以及加入 T-Net- I 模块后的 GACNet 改进模型(记作 GACNet-T-Net- I)进行对比实验,结果如表 2 所示。

得益于模型几何感知能力的提升,加入 T-Net 之后的分割准确率相较加入之前大幅提升。另外,相较 GACNet-T-Net- I,包含 T-Net- II 的 GACNet-GDP 的准确率更高。该实验结果说明如果只选用一个 T-Net,T-Net- II 比 T-Net- I 的性能更好。

表 2 分割准确性对比

Tab.2 Segmentation accuracy comparison

基础模型	DSC	SEN	PPV
GACNet	79.80	79.94	81.52
GACNet-GDP	89.99	90.33	90.30
GACNet-T-Net- I	86.08	85.89	86.97

对应的可视化结果如图 6 所示,其中每行对应着一个典型样本,第 1 列是原始输入网格,第 2 列至 4 列分别为真实值、GACNet-GDP 以及 GACNet-T-Net- I 和 GACNet 的结果。从中不难发现,加入 GDP 之后的效果相较于其他方法更为准确,而 GACNet 没有针对数据的几何差异进行处理,因此存在明显差错,该结论与表 2 相符。

由于在面片方法上测试多个 T-Net 的计算资源要求过高,故改用点云方法 PointNet++作为基础模型进行对比实验。实验对象分别为加入 GDP 模块的 PointNet++(记作 PN++-GDP)和加入 2 个 T-Net 模块的 PointNet++(记作 PN++-T-Net- I &II),实验结果如表 3 所示。从表 3 可见,使用 1 个包含于 GDP 内部的 T-Net 与使用 2 个 T-Net 的效果差别不大,且单个 GDP 计算资源要求更少,因此更适合处理网格数据。

表 3 不同数量 T-Net 模块下的准确率对比

Tab.3 Accuracy comparison with different numbers of T-Net modules

模型	DSC	SEN	PPV
PN++-GDP	87.02	86.71	88.12
PN++-T-Net- I &II	86.98	86.67	88.08

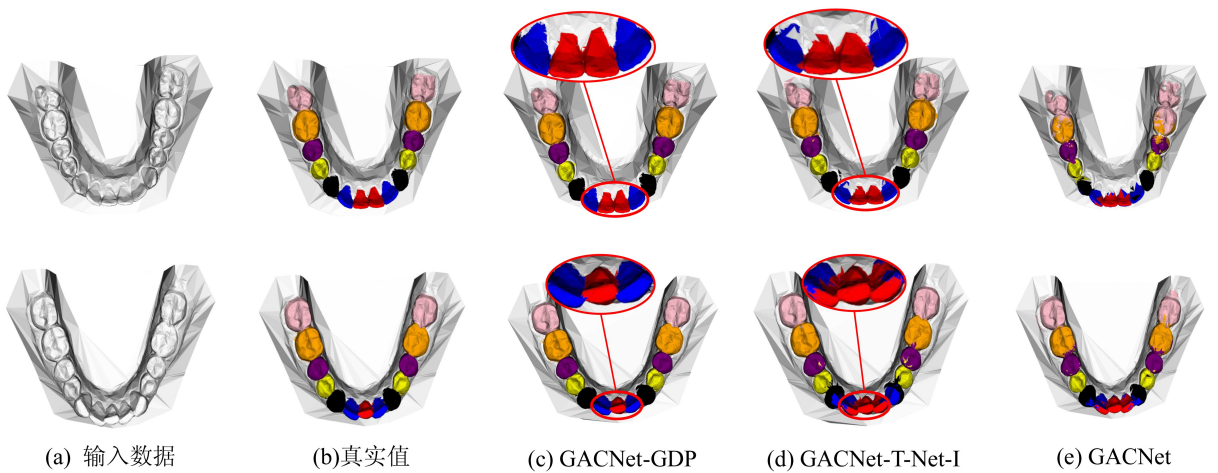


图 6 对比结果的可视化

Fig.6 Visualization of comparison results

3.4 消融实验

GDP 具有超参数 K 和 H , 其中 K 表示 MLP 的个数, H 表示 MLP 的维度。在对 GACNet 的改进中两者取值分别为 2 和 32。为验证该超参数的选择, 本研究对不同取值下的结果进行对比。

3.4.1 MLP 的个数选择

在 H 取值 32 的情况下, 将 K 分别以 0、1、2、3 赋值, 并进行分割性能对比。当 $K=0$ 时, GDP 退化为 T-Net- I, 其余情况本质上是 T-Net- II。如表 4 所示的实验结果表明: (1) 在 T-Net 前加入 MLP 确实能够提升性能; (2) 使用 2 层 MLP 的效果较其他方案效果更优。

表 4 不同 MLP 个数下的准确率对比

Tab.4 Accuracy comparison with different numbers of MLPs

K	DSC	SEN	PPV
0	86.08	85.89	86.97
1	89.49	89.68	89.95
2	89.99	90.33	90.30
3	86.92	86.72	88.02

3.4.2 MLP 的维度选择

在 K 取值 2 的情况下, 测试不同维度 MLP 对性能的影响。由于输入数据的通道数是 24, 分别以 12(降维)、24(不变)、32(升维)、64(升维) 对 H 进行赋值, 并将分割准确率记录如表 5

所示。

实验结果表明, 通过 MLP 对输入特征维度降低会影响模型的分割性能, 导致准确率下降。保持特征维度不变或升高维度对模型则有提升作用, 且当 MLP 的维度取 32 时模型精度最优。

表 5 不同 MLP 维度下的准确率对比

Tab.5 Accuracy comparison with different dimensions of MLPs

H	DSC	SEN	PPV
12	81.12	81.24	81.93
24	87.55	87.48	88.39
32	89.99	90.33	90.30
64	89.23	89.76	89.23

4 结束语

本研究针对三维深度学习模型的几何差异感知能力进行研究, 提出名为 GDP 的网络模块。其核心思想是通过作用于高维特征的 T-Net 来对特征进行规范化。将 GDP 应用于 3D 牙齿分割, 实验结果表明: GDP 模块能够有效应对 3D 数据的几何差异问题, 确保模型的泛化能力。

未来将针对牙齿分割准确率提升方法进行研究, 并在新方法中进一步验证 GDP 模块的效果。

参考文献:

[1] CHARLES R Q,HAO S,MO K C,et al. PointNet:deep learning on point sets for 3D classification and segmentation[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu,HI: IEEE,2017:652-660.

[2] WANG Y,SUN Y B,LIU Z W,et al. Dynamic graph CNN for learning on point clouds[J]. ACM Transactions on Graphics,38(5): 1-12.

[3] ZHAO YE,ZHANG L M,LIU Y,et al. Two-stream graph convolutional network for intra-oral scanner image segmentation[J]. IEEE Transactions on Medical Imaging,2022,41(4):826-835.

[4] MA TA,YANG Y Z,ZHAI J C,et al. A tooth segmentation method based on multiple geometric feature learning[J]. Health-care,2022,10(10):2089.

[5] QI C R,YI L,SU H,et al. PointNet++:deep hierarchical feature learning on point sets in a metric space[C]//Proceedings of the 31st International Conference on Neural Information Processing Systems. New York:ACM,2017:5105-5114.

[6] ZHAO Y,ZHANG L M,YANG C S,et al. 3D Dental model segmentation with graph attentional convolution network[J]. Pattern Recognition Letters,2021,152:79-85.

[7] 张雅玲,于泽宽,何炳蔚. 基于 GCNN 的 CBCT 模拟口扫点云数据牙齿分割算法[J]. 计算机辅助设计与图形学学报,2020,32(7):1162-1170.

[8] BEN-HAMADOU A,SMAOUI O,REKIK A,et al. 3DTeethSeg'22:3D teeth scan segmentation and labeling challenge[EB/OL]. 2023:arXiv:2305.18277.

(责任编辑:方素华)

(上接第 572 页)

参考文献:

[1] 李中哲,陈勇,武一民. 基于多体动力学的电动客车变速器噪声预测研究[J]. 机械设计与制造,2020(2):292-295.

[2] 梁伟智,贾爽,李慎龙,等. 混合动力重卡行星传动系统振动噪声分析[J]. 机械设计与制造,2021(08):221-224.

[3] 邢宏福,杨朝会,于楠,等. 基于 Romax 的斜齿圆柱齿轮的微观参数优化[J]. 机械传动,2021,45(6):85-89,138.

[4] 游宇,曹雪梅. 基于 MASTA 仿真的承载传动误差对弧齿锥齿轮振动特性的影响分析[J]. 机械传动,2021,45(9):56-61,67.

[5] 孙刚,任尊松,辛欣,等. 高速动车组齿轮传动系统振动特性[J]. 机械工程学报,2019,55(18):104-111.

[6] SANCHEZ M B,PLEGUEZUELOS M,PEDRERO J I. Influence of profile modifications on meshing stiffness,load sharing, and trasnsmission error of involute spur gears[J]. Mechanism and Machine Theory,2019,139:506-525.

[7] ZHOU J,YI F,XU X,et al. Effect of tooth profile modification on dynamic tooth load of planetary gear train[J]. Shock and Vibration,2019,2019:1-20.

[8] XIAO Z,CAO J,YU Y. Mathematical modeling and dynamic analysis of planetary gears system with time-varying parameters[J]. Mathematical Problems in Engineering,2020,2020:1-9.

[9] 邢彬,侯明曦,孔祥锋. 齿向修形齿轮啮合刚度及齿向载荷分布系数的计算研究[J]. 机械传动,2017,41(10):96-103.

[10] YUE G,NIU W,ZHAO J,et al. Gear whine resolution by tooth modification and multi-body dynamics analysis[R]. SAE Technical Paper,2016.

[11] 甘纯,何泽银,张涛. 考虑啮入冲击激励的跨座式单轨牵引齿轮箱振动噪声预估与试验研究[J]. 噪声与振动控制,2022,42(2):23-28.

[12] 范孝良,石鹏飞,刘建民. 风电齿轮箱微观修形对振动与声振粗糙度性能的影响[J]. 中国工程机械学报,2019,17(1):43-48.

[13] 彭俊祥. 齿轮传动系统动态特性对传递误差的影响分析[D]. 重庆:重庆理工大学,2019.

[14] 杨红波,史文库,陈志勇等. 基于某二级减速齿轮系统的齿面修形优化[J]. 吉林大学学报(工学版),2022,52(07):1541-1551.

(责任编辑:陈雯)