

基于 Mask R-CNN 的款式图衣领识别

黄振华^{1,2,3}, 李涛^{1,2,3}, 蒋玉萍^{1,2,3}, 杜磊^{*1,2,3}

(1. 浙江理工大学 丝绸文化传承与产品设计数字化技术文化和旅游部重点实验室, 浙江 杭州 310018; 2. 浙江理工大学 浙江省服装工程技术研究中心, 浙江 杭州 310018; 3. 浙江理工大学 服装数字化技术浙江省工程实验室, 浙江 杭州 310018)

摘要:为提升服装款式图领型识别精度,提出一种基于 Mask R-CNN 神经网络的服装款式图领型定位与识别方法。建立共 1 800 张包含无领、立领、翻领与驳领 4 种领型的款式图样本库,利用迁移学习与 Mask R-CNN 神经网络实现领型定位与识别。结果表明,4 种领型的平均识别精确度高于 98%,测试集平均精确度达到 99.2%,mAP 值达到 90%。该识别方法可以减少样板生成中的人工失误,为数字化样板生成提供参考。

关键词: Mask R-CNN 模型;款式图;服装领型;识别定位

中图分类号: TS 941.26 **文献标志码:** A **文章编号:** 2096-1928(2021)01-0036-06

Collar Flat Sketch Recognition Based on Mask R-CNN

HUANG Zhenhua^{1,2,3}, LI Tao^{1,2,3}, JIANG Yuping^{1,2,3}, DU Lei^{*1,2,3}

(1. Key Laboratory of Silk Culture Heritage and Products Design Digital Technology, Zhejiang Sci-Tech University, Hangzhou 310018, China; 2. Clothing Engineering Research Center of Zhejiang Province, Zhejiang Sci-Tech University, Hangzhou 310018, China; 3. Zhejiang Provincial Engineering Laboratory of Clothing Digital Technology, Zhejiang Sci-Tech University, Hangzhou 310018, China)

Abstract: In order to improve the recognition accuracy of collar flat, this paper proposed a fine-grained localization and recognition method of garment collar based on Mask R-CNN neural network. A database of 1800 images which contained collarless, lapel, turndown collar and stand collar, was created. Transfer learning and Mask R-CNN neural network were used to realize collar location and recognition. The results showed that the average recognition accuracy of four collar types were higher than 98%. The average accuracy of the test set and mAP were reached to 99.2% and 90% respectively. The recognition method can reduce the manual errors in the pattern generation, and provide a reference for the digital pattern generation.

Key words: Mask R-CNN model, garment flat, garment collar type, recognition and location

目前服装样板生成过程中,需要样板师人工解读款式图信息,容易造成识别错误或理解偏差,导致样板生成错误,降低生产效率^[1]。衣领是服装的重要组成部分,不仅展现服装的美感,而且很大程度上决定了服装风格^[2]。由于款式图对衣领的细

部工艺表达不完整,容易造成样板师人工识别错误^[3],因此,实现款式图衣领的自动识别具有重要现实意义。

为降低人工识别误差,国内外学者通过引入计算机视觉技术代替人工识别,其主要基于图像内容

收稿日期:2020-11-08; 修订日期:2020-12-16。
基金项目:文化和旅游部重点实验室开放基金项目(2020WLB09);国家级大学生创新创业训练计划项目(201910338011);浙江省教育厅一般科研项目(Y201840287);浙江省服装工程技术研究中心开放基金项目(2018FZKF13)。
作者简介:黄振华(1996—),男,硕士研究生。
*通信作者:杜磊(1989—),男,副教授,硕士生导师。主要研究方向为功能性纺织品与服装数字化技术。
Email:dulei@zstu.edu.cn

与深度学习两种方式^[4]。基于图像内容是通过人工设计方法提取图像特征,主要有几何特征^[5]、全局特征^[6]、局部特征^[7]及多特征融合^[8]等提取方式。其中,AN L X 等^[9]运用小波傅里叶描述子表征服装外部廓形特征,并利用主成分分析降维,结合极端学习机对款式图进行分类。由于小波傅里叶描述子不适合形状的实时分类,李东等^[6]提出一种通过预处理获取服装轮廓信息的算法,并训练多分类支持向量机实现实时分类。

随着深度学习在图像分类上的不断发展,学者们利用深度学习方法进行服装图像识别。吴欢等^[10]针对服装相似廓形分类效果不理想的问题,引入卷积神经网络 CaffeNet 模型进行女裤廓形分类,分类准确率达到 95%。针对衣领部位分类效果不理想的问题,尹光灿等^[11]通过人工截取领型区域,运用 AlexNet 卷积神经网络实现对服装领型的识别与分类。由于衣领部件相对服装整体目标较小,其他服装部件会形成噪声,导致分类准确率低,通过人工截取目标区域虽然可以减少噪声,但被分类图像需进行裁剪以保证部件位置统一,因此无法广泛适用于服装图像的分类。

针对小目标识别困难的问题,以服装款式图中衣领作为研究对象,选用神经网络模型 (mask region convolutional neural networks, Mask R-CNN) 进行试验^[12],该模型对数据要求低,不需要对数据进行裁剪、灰度化等预处理,对小目标物体识别性能高。文中根据款式图特点调整模型参数,结合迁移学习方法实现对款式图领型的识别与定位。

1 实验方法

1.1 Mask R-CNN 模型

实验采用 Mask R-CNN 深度学习算法,该方法在 Faster R-CNN 基础上进行改进,解决了 Faster R-CNN 在识别小目标物体上出现丢失的问题,并提升了识别准确度^[13]。Mask R-CNN 结构如图 1 所示,主要由 3 个模块组成,分别是特征提取网络 (CNN)、区域建议网络 (region proposal network, RPN) 及全连接层回归网络 (full connect)。最终网络输出包含 3 个分支,分别是目标的分类 (class)、目标包围框的坐标 (box) 及目标物体的二值掩码 (mask)。

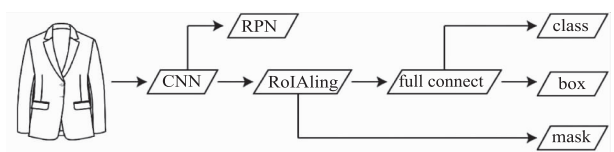


图 1 Mask R-CNN 结构
Fig. 1 Structure of Mask R-CNN

1.1.1 特征提取网络 Mask R-CNN 在特征提取网络中采用残差特征金字塔结构,该网络将 ResNet 与特征金字塔网络 (feature pyramid networks, FPN) 进行结合,实现仅增加极小的计算量,同时处理好物体检测的多尺度变化问题,能够使小目标在下采样过程中不丢失。ResNet 还可以根据不同任务目标选取不同的骨干网络,例如 ResNet-50 与 ResNet-101。其中 ResNet-50 是由 4 个模块组成,每个模块中含有若干个残差块,第一个残差块将上一个模块通道数量翻倍,而高与宽减半,共包含了 49 层卷积层与 1 层全连接层。ResNet-101 是在 ResNet-50 的基础上增加第 4 层残差块。在 Mask R-CNN 中,ResNet 仅作为特征提取器对图像进行特征提取,因此将 ResNet 的全连接层去除。

1.1.2 区域建议网络 Mask R-CNN 利用 ResNet-FPN 结构将不同阶段的特征图输入 RPN 生成多个候选信息区域 (region of interest, RoI),通过 RPN 对 RoI 进行训练,得到标定具有完整标定正样本与含有部分标定区域负样本,将剩余的 RoI 进行 RoIAlign 操作,将图像上生成的 RoI 与 Resnet-FPN 提取的特征层进行对齐,最后计算 RPN 网络的损失。RoIAlign 通过双线性插值得到固定的像素值,减少返回原图时的误差。其中 RoIAlign 的对齐公式为

$$K = K_0 + \text{lb}(\sqrt{wh}/244)。$$
 (1)

式中: w, h 分别表示 RoI 宽度和高度; K 为这个 RoI 的特征层; K_0 是 $w = 224, h = 224$ 时的特征层。

1.1.3 全连接层回归

将获取的 RoI 运用全卷积网络 (fully convolution network, FCN) 进行类别分类、标定框回归与 Mask 掩码回归。该模型将训练与预测分离,在训练中同时进行分类与掩码回归,而在预测中先得到分类结果,再将结果导入掩码回归中得到对应掩码。FCN 类别分类和回归与 RPN 网络中分类和回归相同,损失函数也与 Softmax 层交叉熵损失函数相同。其中训练采用多任务损失函数,通过学习不断减小损失函数的值,直至获得全局最优解。损失函数的公式为

$$L = L_c + L_b + L_m。$$
 (2)

式中: L_c 为分类误差, L_b 为标定框误差, L_m 为掩码误差。

1.2 结果评价指标

在目标检测中常用精确度与平均 AP 值 (mean average precision, mAP) 作为评价指标。精确度是衡量分类准确性的指标,以分类结果为正类的概率进

行表征。mAP 是多个类别 AP 的平均值,AP 是单个类别的平均精确度,通过计算精确度、召回率以及预测框和真实框的交并比(intersection over union, IoU)判断预测框是否准确预测到了位置信息,具体如图 2 所示,其中并集 U 如图 2(a)所示,阴影部分为交集 I 如图 2(b)所示。 S_{IoU} 值越高表示领型目标定位越精确。IoU 公式为

$$S_{IoU} = \frac{I}{U}。$$

(3)

式中: I 为预测边框与真实边框交集, U 为预测边框与真实边框并集。

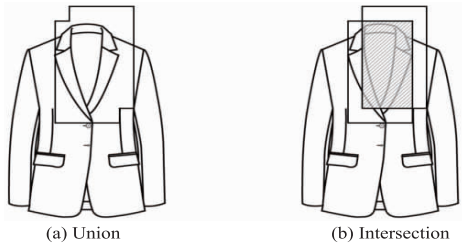


图 2 IoU 示意
Fig. 2 Schematic diagram of IoU

2 识别实验

2.1 样本库构建





根据领型结构进行分类,将领型分为无领、立领、翻领、驳领 4 种领型^[11],其特点与款式图见表 1^[14],每种领型下含有不同细分领型,但实验仅针对经典领型进行分类。在网络上收集包含 4 种领型的款式图,创建一个款式图样本库。为减少样本不均匀造成的误差,每种领型各收集 200 张,共得到 800 张服装款式图,通过数据增强^[13]的方式,将样本库扩充为 1 800 张,其中训练集为 1 600 张,测试集为 200 张,4 种领型样本数量均匀分布。款式图源于网络设计素材,主要为白底背景,不含有噪音。由于 Mask R-CNN 对图像质量要求较低,样本库中图像分辨率范围为 437×459 像素到 2 480×3 508 像素。

2.2 数据集制作

将样本库中样本运用深度学习图像标注工具 VIA (VGG Image Annotation) 进行人工 Mask 框选以

表 1 领型特点与款式图

Tab. 1 The features of collar type and garment flat

衣领类别	款式特点	款式图	标 签
无领	只有领窝而无领身,用领窝线造型表示领型		L_w
立领	只有领座没有翻领部分		L_L
翻领	领座部分和翻领部分通过缝制连接成一体或领座与翻领连成一体		L_F
驳领	由翻领和驳领两部分组成,翻领与衣片领口缝合,驳领由衣片的挂面翻出形成		L_B

及类别标注,使用多边形线段描出所有领型区域与领型轮廓,根据领型赋予标签“L_L”“L_F”“L_W”和“L_B”。当一张图像中含有2个款式时,对2个款式的衣领处进行对应标注,款式图标注格式如图3所示,阴影区域为该款式图领型的mask,并在属性中选取对应class标签。将标记好的图像保存为json格式,作为神经网络读取数据。



图3 款式图标注格式

Fig. 3 Labeling format of garment flat

2.3 Mask R-CNN 实验

实验采用 Python 语言、TensorFlow 框架编程实现,基于台式电脑 Windows 10 操作系统[Intel (R) Core (TM) i7-8700 CPU,3.20 GHz,32 GB],GPU 为 MSI RTX 2060。实验针对款式图特点与样本库数量进行参数修改。

由于人工采集的数据集较少,为了防止算法在训练过程中发生过拟合现象,采用迁移学习方法,利用 MS COCO 公用数据集进行网络预训练。该数据集适用于图像分类与语义分割,保存训练后网络的权重。

网络主要结构参数见表 2。由于款式图中领型特征信息较少,运用多层数隐含层骨架网络容易产生过拟合现象,因此将 Mask R-CNN 网络默认的卷积神经网络模型 ResNet-101 修改为 ResNet-50,迭代次数(epoch)为 500,每次迭代训练图片张数(steps-per-epoch)为 50,学习率为 0.001,权重下降为 0.000 1。Mask R-CNN 分类过程如图 4 所示。

表2 Mask R-CNN 主要参数

Tab.2 Main parameters of Mask R-CNN

配 置	属 性
骨干网络	Resnet-50
图片尺寸	[1 024,1 024,3]
学习动量	0.9
学习率	0.001
类别数	4
RPN 非极大值抑制閾值	0.7
单次迭代步长	50
迭代次数	500
权重衰减	0.000 1

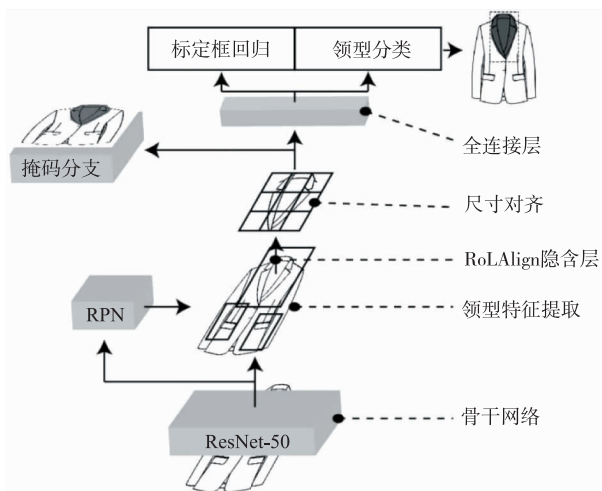


图 4 Mask R-CNN 分类过程

Fig. 4 Classification process of Mask R-CNN

3 结果与分析

3.1 Mask R-CNN 识别结果

3.1.1 训练结果分析 实验采用 Mask R-CNN 网络是为了将服装款式图中衣领位置进行定位,并根据标定框中提取的领型特征信息进行类型识别。网络的各损失函数如图 5 所示,从网络损失图可知,网络训练结果良好,没有出现过拟合现象。

由图 5 可知,总损失、class 损失、bbox 损失和 mask 损失都在稳定收敛。其中总体损失函数在迭代次数达到 400 后,总损失趋近 0.1;class 损失在迭代次数达到 400 时,损失低于 0.01;bbox 损失在迭代次数达到 400 后已经稳定波动,mask 在迭代次数达到 450 次后也趋于稳定。

3.1.2 识别结果 随机抽取测试集中的款式图进行识别测试, 针对领型识别精度而言, 单个领子识别精确度为 99.9%; 包含 2 个领型的款式图的识别精确度为 99.8%, 识别结果如图 6 所示。其中掩码 mask 能够完整覆盖领型区域, 标定框 box 准确检测出领型所在区域。

表3为Mask R-CNN的实验结果。由表3可知,所有领型的平均识别精确度达到98%以上,无领、立领与翻领的平均识别精确度达到99%以上,并且mAP值为90%,表明文中实验方法取得良好结果,对款式图领型分类与定位具有较高的精确度。其中,无领与驳领的AP不足90%,这主要是由于4种领型中无领和驳领的细部领型形态和形状差异较大,box定位与mask标注不能完全与测试集标注重合,导致在计算AP时IoU值下降。

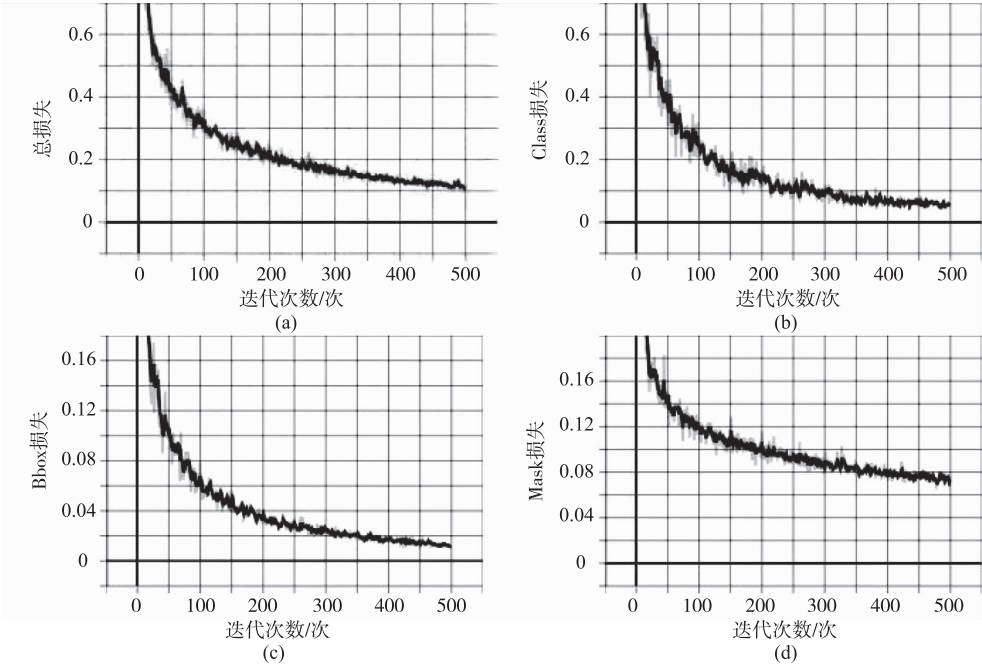


图 5 损失函数曲线
Fig.5 Loss function curve

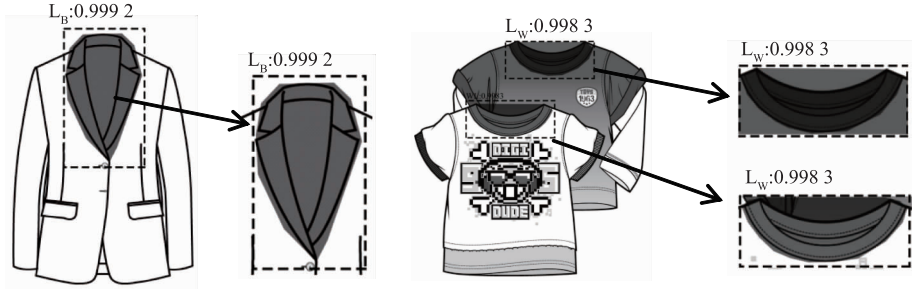


图 6 识别结果
Fig.6 Recognition result

表 3 Mask R-CNN 实验结果

Tab.3 Experimental results of Mask R-CNN

领型种类	训练集/张	测试集/张	平均精确度/%	AP/%	mAP/%
无领	400	50	99.41	87.5	90
立领	400	50	99.71	90.0	
翻领	400	50	99.33	97.5	
驳领	400	50	98.39	85.0	

3.2 实验对比分析

为验证改进后的 Mask R-CNN 网络对服装款式图领型分类的优势,将基于 Mask R-CNN 识别分类方法与尹光灿等^[11]在领型分类中提出的 Canny 边缘检测算法结合 AlexNet 卷积神经网络模型的方法进行对比。采用数据增强后的数据库进行实验,研究发现,相对于 AlexNet 的衣领识别,Mask R-CNN 方法在提高服装领型分类准确率方面具有优势,且不需要人工截取衣领信息,可有效减少人为干预因素。由于 mAP 是目标检测评价指标,所以仅用精确度进行对比,实验对比结果见表 4。

表 4 识别精度实验对比

Tab.4 The contrast of the experimental precision

实验方案	平均精确度/%
Mask R-CNN	99.8
AlexNet	20.3
Canny + AlexNet	29.5

由表 4 可知,AlexNet 识别精确度为 20.3%,识别结果较差,主要因为 AlexNet 卷积神经网络对图像质量要求较高,增加随机噪声的数据增强方式对 AlexNet 卷积神经网络造成较大干扰。利用 Canny 边缘检测算法对图片进行处理,虽然可以降低网络

提取领型特征的难度,减少色彩对领型识别的影响,进而提高领型分类的准确率,但调整后的Mask R-CNN网络识别效果更为理想。Mask R-CNN与AlexNet卷积神经网络方法相比,精确度提高了79.5%;与Canny边缘检测结合AlexNet卷积神经网络方法相比,精确度提高了70.3%,并且相较于文献[11]的方法,Mask R-CNN不需要对样本进行裁剪、灰度化和边缘检测等预处理,适用于各类服装图像。

4 结 语

文中以衣领款式图作为研究对象,通过改进Mask R-CNN目标检测方法,结合迁移学习对模型进行预训练,修改Mask R-CNN中的骨干网络,实现了对服装款式图中的领型识别,平均精确度达到99.2%,mAP值达到90%。相较于卷积神经网络方法,该方法对图像质量要求低,并且不需要对图像进行裁剪预处理,可以识别服装款式图部件,准确率提高79.5%,为二维数字化样板生成提供参考与借鉴。

参考文献:

- [1] 李涛,杜磊,黄振华,等.服装款式图识别与样板转换技术研究进展[J].纺织学报,2020,41(8):145-151.
LI Tao, DU Lei, HUANG Zhenhua, et al. Review on pattern conversion technology based on garment flat recognition[J]. Journal of Textile Research, 2020, 41(8): 145-151. (in Chinese)
- [2] 尹晓芳,蒋晓文.基于感性工学的女衬衫领型评价[J].服装学报,2020,5(2):111-118.
YIN Xiaofang, JIANG Xiaowen. Evaluation of blouse collar type based on kansei engineering[J]. Journal of Clothing Research, 2020, 5(2): 111-118. (in Chinese)
- [3] 李楠.基于行业需求的服装款式图绘制方法[J].服装学报,2016,1(1):53-58.
LI Nan. Research on the fashion style drawing method according to industry needs[J]. Journal of Clothing Research, 2016, 1(1): 53-58. (in Chinese)
- [4] 陈巧红,陈翊,孙麒,等.服装图像分类技术综述[J].浙江理工大学学报(自然科学版),2019,41(5):631-643.
CHEN Qiaohong, CHEN Yi, SUN Qi, et al. Overview of clothing image classification technology[J]. Journal of Zhejiang Sci-Tech University(Natural Sciences Edition), 2019, 41(5): 631-643. (in Chinese)
- [5] HUANG P, YAO J, ZHAO H. Automatic realistic 3D garment generation based on two images[C]//China VR. 2016 International Conference on Virtual Reality and Visualization. New York:IEEE, 2016: 250-257.
- [6] 李东,万贤福,汪军.采用傅里叶描述子和支持向量机的服装款式识别方法[J].纺织学报,2017,38(5):122-127.
LI Dong, WANG Xianfu, WAN Jun. Clothing style recognition approach using fourier descriptors and support vector machines[J]. Journal of Textile Research, 2017, 38(5): 122-127. (in Chinese)
- [7] DING X J, ZOU C H, CHEN K J, et al. Extraction and classification of She nationality clothing via visual features[J]. Textile Research Journal, 2016, 86(12): 1259-1269.
- [8] SURAKARIN W, CHONGSTIVATANA P. Classification of clothing with weighted SURF and local binary patterns[C]//International Computer Science and Engineering Conference. 2015 International Computer Science and Engineering Conference (ICSEC). New York:IEEE, 2015: 1-4.
- [9] AN L X, LI W. An integrated approach to fashion flat sketches classification[J]. International Journal of Clothing Science and Technology, 2014, 26(5): 346-366.
- [10] 吴欢,丁笑君,李秦曼,等.采用卷积神经网络CaffeNet模型的女裤廓形分类[J].纺织学报,2019,40(4):117-121.
WU Huan, DING Xiaojun, LI Qinman, et al. Classification of women's trousers silhouette using convolution neural network CaffeNet model[J]. Journal of Textile Research, 2019, 40(4): 117-121. (in Chinese)
- [11] 尹光灿,罗戎蕾.基于卷积神经网络的服装领型识别与分类研究[J].现代纺织技术,2020(3):48-53.
YIN Guangcan, LUO Ronglei. Research on recognition and classification of garment collar type based on convolutional neural network[J]. Advanced Textile Technology, 2020(3): 48-53. (in Chinese)
- [12] HE K M, GKIOXARI G, DOLLAR P, et al. Mask R-CNN[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017(99):1.
- [13] 刘舒康,唐鹏,金炜东.基于智能数据增强和改进YOLOv3算法的接触网吊弦及支架检测研究[J].计算机科学,2020,47(Sup.2):178-182.
LIU Shukang, TANG Peng, JIN Weidong. Study on catenary dropper and support detection based on intelligent data augmentation and improved YOLOv3[J]. Computer Science, 2020, 47(Sup.2): 178-182. (in Chinese)
- [14] 张文斌.服装结构设计[M].3版.北京:中国纺织出版社,2007:88-89.