

基于手绘草图的中保真原型自动生成

梁旭东, 黎 鸣, 赵慧玲, 范 益, 林 涛

(四川大学计算机学院, 成都 610065)

摘 要: 从手绘草图生成数字化的中保真原型是界面设计的必要步骤, 它需要花费设计师大量的时间和精力. 针对这一问题, 本文提出了基于目标检测和启发式规则的中保真原型自动生成方法, 并基于该方法实现了一个自动化工具. 首先, 本文提出基于置换注意力机制(SA)的目标检测模型 SA-FRCNN 对手绘草图中的组件进行检测; 然后, 提出启发式规则对检测结果进行布局优化. 实验表明, 本文方法能够提升对手绘草图的检测精度, 较基线模型提高了 2.3%; 此外, 本文工具可以有效提高原型设计效率.

关键词: 手绘草图; 中保真原型; 目标检测; 启发式规则

中图分类号: TP391.7; TP183 **文献标识码:** A **DOI:** 10.19907/j.0490-6756.2022.022005

Automatic generation of the mid-fidelity prototype based on hand-drawn sketch

LIANG Xu-Dong, LI Ming, ZHAO Hui-Ling, FAN Yi, LIN Tao

(College of Computer Science, Sichuan University, Chengdu 610065, China)

Abstract: Generating digital mid-fidelity prototypes from hand-drawn sketches is a necessary step in User Interface design, which takes designers a lot of time and effort. In order to solve the problem, this paper proposes an automatic method for generating mid-fidelity prototypes based on object detection and heuristic rules, and an automatic tool is implemented based on this method. Firstly, an object detection model SA-FRCNN based on the Shuffle Attention (SA) mechanism is proposed to detect components in hand-drawn sketches; and then, heuristic rules are proposed to optimize the layout of the detected results. Experiments show that our method can improve the detection accuracy of hand-drawn sketches by 2.3% compared with the baseline model; in addition, our tool can effectively improve the efficiency of prototyping.

Keywords: Hand-drawn sketch; Mid-fidelity prototype; Object detection; Heuristic rules

1 引 言

界面原型通常分为低保真和高保真两种保真度^[1]. 低保真原型通常是手绘草图, 其创建快速, 便于比较多个设计方案, 但需要较高的沟通成本^[1-3]; 高保真原型与最终界面外观相似, 便于沟通, 但不易制作和修改^[1,3]. 在实际的原型设计过

程中, 还存在着介于这两者之间的中保真原型, 它是对低保真原型的数字化, 具有美观的布局, 但不具备视觉设计细节^[1]. 中保真原型为后续的高保真原型设计奠定了必要基础.

在原型设计过程中, 根据手绘草图制作中保真原型是必要和关键的一步, 该步骤完成的好坏将直接影响后续高保真原型的设计. 同时, 快速将手绘

收稿日期: 2021-10-26

基金项目: 四川大学-泸州战略合作项目(2017CDLZ-G30)

作者简介: 梁旭东(1995-), 男, 硕士研究生, 研究方向为人机交互. E-mail: 753533360@qq.com

通讯作者: 林涛. E-mail: lintao@scu.edu.cn

草图转换为可以修改或重新排列组件的数字原型,对加快设计过程很有价值^[4]。此外,一项对设计师工作负荷进行分析的研究表明,相较于低保真原型和高保真原型,制作中保真原型需要花费设计师更多的体力和更少的脑力^[5],这说明中保真原型设计是一项高重复性、低创造性的工作,该过程具有很大的自动化潜力。

近年来,人工智能技术特别是目标检测算法已取得了很大的进展^[6,7],该技术的引入对加快界面设计很有意义。典型的研究思路是先对草图进行检测,然后生成代码^[8,9],虽然取得了不错的效果,但不便于设计师进行修改,他们更习惯于在设计工具的图形界面上创作^[3]。因此,研究者们提出了面向设计师的原型界面自动生成方法^[5,10],然而,该方面的研究还有待完善,界面布局仍然需要设计师花费大量时间手动调整。界面的美观度对用户感知的可用性和满意度有积极影响,美观是界面设计的重要目标^[11]。手绘草图由于创建随意的特点,缺乏美观的布局,这导致其不便于沟通。然而,目前还缺乏基于手绘草图进行布局优化的研究。

针对上述问题,本文提出了基于目标检测和启发式规则的中保真原型自动生成方法,旨在帮助设计师从低保真手绘草图快速实现数字化的中保真原型。首先,本文使用置换注意力(Shuffle Attention, SA)^[12]改进 Faster-RCNN^[13](SA-FRCNN)以提高对手绘草图的检测精度;然后,提出启发式规则基于检测结果进行布局优化,生成对应的中保真原型。此外,本文基于该方法开发了一个中保真原型自动生成工具,该工具实现为界面设计软件 Adobe XD 的一个插件,接受拍照上传的手绘草图,能帮助设计师快速从手绘草图自动生成数字化的中保真原型。

2 中保真原型自动生成

中保真原型自动生成分为界面组件检测和界面布局优化两部分,其总体框架如图 1 所示。首先,本文使用目标检测模型 SA-FRCNN 检测手绘草图中的组件,得到组件的类型、位置和大小等信息;接着采用启发式规则,基于检测结果划分组件间的包含关系,并调整组件的大小和位置,自动生成可编辑、数字化的中保真原型。

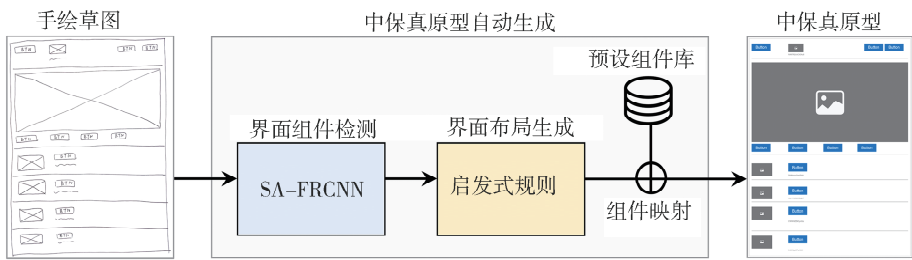


图 1 中保真原型自动生成框架
Fig. 1 Framework of the mid-fidelity prototype automatic genetation

2.1 基于 SA-FRCNN 的界面组件检测

在组件检测阶段,本文模型采用了引入置换注意力的两阶段目标检测网络 SA-FRCNN,其网络架构如图 2 所示。该模型由以下几部分组成:(1)基于 ResNet50^[14]改进的特征提取网络 SA-Net50,用于提取输入草图的特征;(2)区域建议网络(Region Proposal Network, RPN)^[13]将特征图划分为多个子区域,判断每个子区域是前景还是背景,并对前景进行边框回归,得到感兴趣区域(Region of Interest, ROI)^[13]; (3)ROI 池化结合特征图和 ROI 得到 ROI 特征,并将其变换为固定长度,用于分类并再次精修位置。

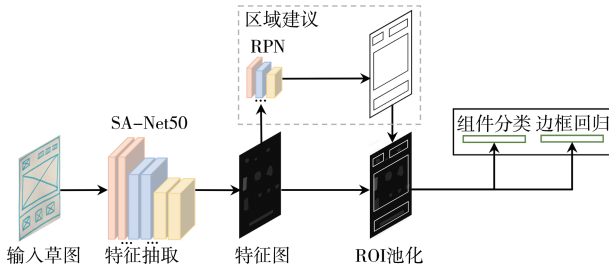


图 2 SA-FRCNN 网络架构
Fig. 2 Network architecture of SA-FRCNN

置换注意力采用置换单元高效组合通道注意力和空域注意力,其整体结构如图 3 所示,分为 4 个步骤:(1)沿通道维度将特征 X 分为多个子特征 $[X_1, \dots, X_g]$,并对每个子特征 X_a 再次沿着通道维度等分为 X_{a1} 和 X_{a2} , X_{a1} 用于学习通道注意力特

征, X_{a2} 用于学习空域注意力特征; (2) 对 X_{a1} 采用全局平均池化 (Global Average Pooling, GAP)^[15] F_{gp} 嵌入全局信息以生成通道统计, 并通过带有 Sigmoid 激活的门控机制 σ 紧凑特征, 得到 X_{a1} 经通道变换后的 X_{a1}^* ; (3) 在 X_{a2} 上采用组归一化 (Group Normalization, GN)^[16] 来获得空间统计,

然后采用 F_c 来增强 X_{a2} 的表示, 得到空域变换后的 X_{a2}^* ; (4) 合并 X_{a1}^* 和 X_{a2}^* , 并用通道置换操作进行组间通信, 得到变换后的 X^* . 置换注意力的最终输出具有与输入相同的尺寸, 这使得置换注意力可以轻易的嵌入到现有网络架构中.

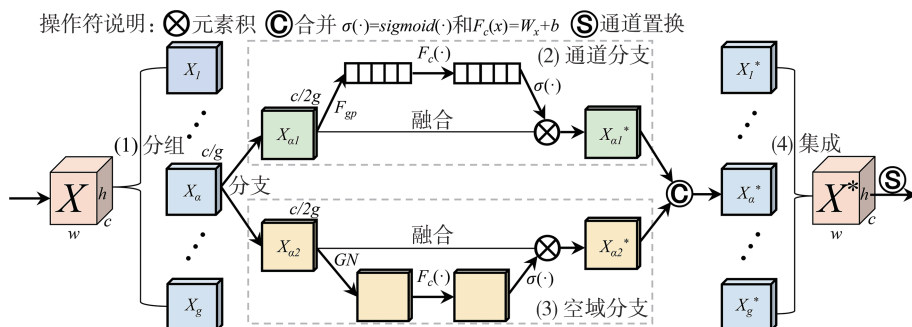


图 3 置换注意力结构
Fig. 3 Architecture of shuffle attention

2.2 基于启发式规则的界面布局优化

上一阶段的输出是一组对应于被检测组件的包围框, 由于手绘草图的随意性, 这些包围框在布局上可能存在排列错乱和重叠等问题, 本文基于启发式规则进行布局优化, 算法描述如算法 1 所示.

算法 1 基于启发式规则的布局优化算法

输入: 经检测所得的组件列表 L

输出: 布局优化后的组件列表 L^*

Begin

- (1) 对于任意两组件 C_1 和 C_2 :
- (2) $\alpha := \min(C_1.area, C_2.area)$
- (3) 如果 C_1 和 C_2 产生重叠区域 A 且 $A.area > 0.8\alpha$:
- (4) 如果 $C_1.area > C_2.area$, C_1 为 C_2 的父节点, 裁去 C_2 超出 C_1 的区域
- (5) 否则, C_2 为 C_1 的父节点, 裁去 C_1 超出 C_2 的区域
- (6) $\omega := \min(C_1.w, C_2.w)$
- (7) 如果 $|C_1.x - C_2.x| < \omega/4$, 进行左对齐
- (8) 如果 $|C_1.w - C_2.w| < \omega/4$, 使其宽相等
- (9) $\eta := \min(C_1.h, C_2.h)$
- (10) 如果 $|C_1.y - C_2.y| < \eta/4$, 进行上边对齐
- (11) 如果 $|C_1.h - C_2.h| < \eta/4$, 使其高相等
- (12) 对任意父节点相同的 C_1 和 C_2 :
- (13) 当 C_1 和 C_2 产生重叠区域 A 时, 记 A 的

宽和高分别为 Δw 和 Δh

- (14) 如果 $\Delta w > \Delta h$, 所有 $C.y \geq A.y$ 的组件 C 向下平移 Δh 个单位
- (15) 否则, 所有 $C.x \geq A.x$ 的组件 C 向右平移 Δw 个单位
- (16) 调整父节点大小以完全包含平移后的所有子节点

End

算法 1 中, 首先, 组件间有时存在包含关系, 本文通过计算组件间的重叠面积占比来生成它们的容器包含关系, 便于在各容器内分别进行布局优化. 对于重叠时超出边界的部分, 本文将其裁剪删去, 如步骤(1)~(5)所示. 然后, 调整组件的位置和大小, 使其在视觉上保持对齐和一致. 通常而言, 水平坐标或垂直坐标大小接近的组件在该方向上应保持对齐; 大小相近的组件应该保持相同的尺寸, 如步骤(6)~(11)所示.

最后, 采取整体向下或向右平移的方式解决重叠问题, 该做法可以有效避免二次重叠. 同时, 容器边界也将更新以保证完全包含平移后的组件, 如步骤(12)~(16)所示.

3 实验与评估

本文从准确率与可用性两方面对本文方法进行了评估.

3.1 数据集及实现细节

本文在 DrawnUI^[17]数据集上进行实验,该数据集由 4291 张手绘草图组成,包含 21 类组件,涵盖了大多数用例. 所有草图均为移动应用程序截图和网页截图的手绘版本.

本文基于 PyTorch 框架实现了本文的组件检测模型,随机选择了 75% 的 DrawnUI 数据集训练模型. 在训练模型时,设置锚框的尺度分别为 32×32, 64×64, 128×128, 256×256, 512×512,长宽比分别为 1:1, 1:2, 2:1, 以便生成尺度更多样化的锚框. 初始学习率为 0.0025,每 10 个 epoch 进行衰减,批量大小设置为 2. 在输入网络训练之前已对草图进行了归一化处理并调整了尺寸,使其短边为 800,长边不超过 1333.

3.2 评价指标

本文使用平均精度均值(Mean Average Precision, mAP)作为评价指标来评估本文模型在手绘草图组件检测任务上的性能. mAP 需要由精确率计算得出. 精确率计算公式如下.

C_precision = C_TP / (C_TP + C_FP) (1)

式中,C_{TP}为真正例(True Positive, TP),即正确分类的组件个数;C_{FP}为假正例(False Positive, FP),即错误分类的组件个数.

平均精确率计算公式如下.

C_AP = (Σ C_precision) / N(Totalsketches)_C (2)

式中,C_{precision}为 C 类别组件的精确率;N(Total-Sketches)_C为包含 C 类别组件的手绘草图数量.

平均精度均值计算公式如下.

mAP = (Σ C_AP) / N(classes) (3)

式中,C_{AP}为每种组件的检测精确率;N(Classes)为数据集中所有组件的类别数.

3.3 准确率评估

本文使用随机选择的 25% 的 DrawnUI 数据集作为识别准确率评估的测试数据. 模型训练数据中没有出现任何测试数据. 本文统计了模型加入置换注意力前后的两项指标并进行分析,即交并比(Intersection over Union, IoU)为 0.5 下的 mAP(mAP@50)和模型的参数量,结果如表 1 所示. 由表 1 可知,将置换注意力加入 FRCNN 的特征提取网络后,模型的组件检测效果有明显的提升(mAP@50 提高了 2.3%),同时几乎没有增加额外的计算量.

表 1 SA-FRCNN 的识别准确率

Tab. 1 Recognition accuracy of SA-FRCNN

Detector	Backbone	mAP@50/%	Param/M
FRCNN	ResNet50	86.2	41.53
	+SA	88.5	41.53

图 4 展示了两个典型的例子,使用红框和绿框分别标出了识别错误和布局错误. 如图 4 所示,本文提出的组件识别模型总体展现了较好的性能. 然而,对复杂组件的识别还不够理想(如图 4b 中将“时间选择器”识别为“按钮”,未检测出“列表”). 分析有两点原因:(1) 这些组件的局部特征和其他组件较为相似;(2) 训练数据相对缺乏.

此外,本文所提的布局优化方法在对齐和重叠问题上表现出较好的性能. 然而,仍然可能存在局部区域未能对齐的情况,分析原因有两点:(1) 受限于人工设定的阈值;(2) 对组件平移可能导致本应对齐的组件不再对齐(如图 4a 中标出的图像与文本未能对齐). 此外,尽管整体平移能有效避免二次重叠问题,但也可能改变界面原本的布局(如图 4b 中为处理“按钮”与“容器”的重叠问题而使界面下方布局改变).

3.4 可用性评估

系统可用性是用户做选择时重要考虑的因素^[18]. 本文选取了 10 个被试,通过设计两个任务,对本文工具进行了可用性评估.

3.4.1 实验任务及过程 本文采用非概率抽样在四川大学学生中选取了 10 名被试者(5 名女性,5 名男性),所有被试者从未使用过本文工具. 实验前,先向被试者介绍草图组件的含义及工具的功能,确保被试者能够正确理解界面和使用工具.

本文指定了界面的功能为“登录”,要求被试者自由手绘草图界面,然后完成以下两项任务:(1) 使用本文工具直接生成先前手绘草图对应的中保真原型,然后手动调整自动生成的中保真原型中的错误;(2) 采用完全手工的方式在空白画布上制作先前手绘草图对应的中保真原型. 两个任务均完成后,被试者需填写系统可用性量表(the System Usability Scale, SUS)^[19]对本文工具进行可用性评估. SUS 的原始打分区间为 1 分到 5 分,1 分表示强烈反对,5 分表示非常赞成. 为了计算系统的综合得分,将各项目得分转换为 1 分到 4 分区间内.

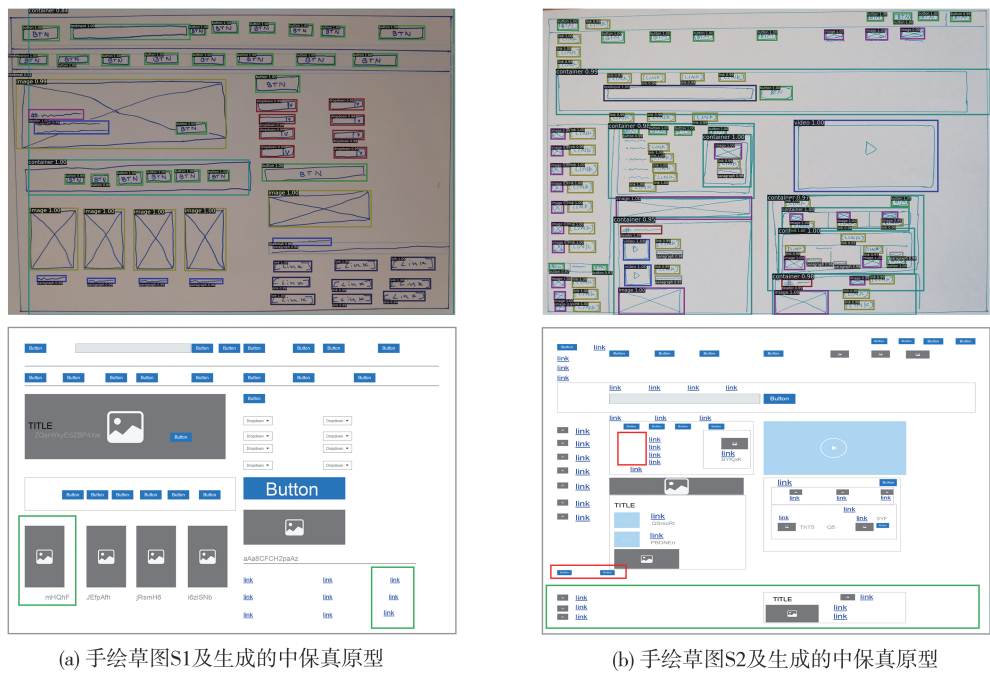


图 4 组件检测及中保真原型自动生成的结果
Fig. 4 Results of components detection and mid-fidelity prototypes automated generation

3. 4. 2 实验结果分析 本文采用威尔科克森符号秩检验对被试者完成上述两个任务所花费的时间进行了分析,以验证本文工具能否有效节约中保真原型设计的时间成本,结果如表 2 所示. 表 2 中, time_A、time_M 分别表示使用自动生成工具和手工完成中保真原型的时间.

表 2 自动化和手工设计所花费的时间			
Tab. 2 Time of automated and manual design			
time_A/min	time_M/min	time_M~time_A/min	
		Z 值	p 值
3.90±0.64	8.99±1.85	2.752	9.766×10 ⁻⁴

实验结果表明,与使用本文工具相比,手工进行中保真原型设计显著增加了时间成本($P < 0.05$),这说明使用本文工具可以有效节约时间开销,提高原型设计效率.

表 3 显示了本文工具在 10 个方面的平均得分. 由表 3 可知,被试者认为本文工具简洁、易用且操作简单,同时不需要过多技术干预且具有较好的一致性;然而,被试者认为本文工具可能不具有良好的常用性和完整性.

SUS 的综合得分代表了对所研究系统的整体可用性的综合度量. 经过计算,本文工具的综合得分为 83 分,评级为 B,表明总体上具有良好的可用性.

表 3 处理后的 SUS 得分	
Tab. 3 Processed SUS score	
项目	得分(均值±标准差)
常用性	2.80±0.79
复杂性	3.70±0.48
易用性	3.70±0.48
需要技术支持	3.60±0.70
完整性	2.40±0.70
不一致性	3.60±0.52
可学习性	3.20±0.79
操作繁琐	3.90±0.32
鼓舞人心	3.00±0.82
需要先验知识	3.33±0.67

4 结 论

本文提出了基于目标检测和启发式规则的中保真原型自动生成方法,通过基于手绘草图的组件识别和布局优化生成对应的中保真原型. 实验证明,本文方法能够有效提高原型设计效率. 未来工作中可以从数据集层面着手增强复杂组件的训练数据,并尝试使用运筹学知识解决界面布局优化问题.

参考文献:

[1] Engelberg D, Seffah A. A framework for rapid mid-

- fidelity prototyping of web sites [C]//Proceeding of the IFIP 17th World Computer Congress-TC13 Stream on Usability: Gaining A Competitive Edge. Deventer, Netherlands: Kluwer B V, 2002.
- [2] Huang F, Canny J, Nichols J. Swire: sketch-based user interface retrieval [C]//Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. New York: ACM, 2019.
- [3] Suleri S, Pandian V P. BlackBox Toolkit: intelligent assistance to UI design [EB/OL]. [2021-10-15]. <https://arxiv.org/pdf/2004.01949.pdf>.
- [4] Manandhar D, Jin H, Collomosse J. Magic layouts: structural prior for component detection in user interface designs [C]//Proceedings of the 2021 IEEE Conference on Computer Vision and Pattern Recognition. Nashville, USA: IEEE, 2021.
- [5] Suleri S, Pandian V P, Shishkovets S, *et al.* Eve: a sketch-based software prototyping workbench [C]//Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems. New York: ACM, 2019.
- [6] 郁强, 王宽, 王海. 一种多尺度 YOLOv3 的道路场景目标检测算法[J]. 江苏大学学报: 自然科学版, 2021, 42: 628.
- [7] 彭大芹, 刘恒, 许国良. 使用候选框进行全卷积网络修正的目标分割算法[J]. 重庆邮电大学学报: 自然科学版, 2021, 33: 135.
- [8] 周子洪, 周志斌, 张于扬, 等. 人工智能赋能数字创意设计: 进展与趋势[J]. 计算机集成制造系统, 2020, 26: 2603.
- [9] 王诗宇, 刘鸿星, 范家佳. 移动用户界面概念模型到代码的转换方法研究[J]. 计算机工程与应用, 2020, 56: 240.
- [10] Pandian V P, Suleri S, Beecks C, *et al.* Meta Morph: AI assistance to transform Lo-Fi sketches to higher fidelities [C]//The 32nd Australian Conference on Human-Computer Interaction. New York: ACM, 2020.
- [11] 陈锐, 谢应涛, 林涛, 等. 自动生成美观的信息墙[J]. 四川大学学报: 自然科学版, 2016, 53: 61.
- [12] Zhang Q L, Yang Y B. SA-NET: shuffle attention for deep convolutional neural networks [C]//Proceedings of the 2021 IEEE International Conference on Acoustics, Speech and Signal Processing. Toronto: IEEE, 2021.
- [13] Ren S, He K, Girshick R, *et al.* Faster R-CNN: towards real-time object detection with region proposal networks[J]. IEEE T Pattern Anal, 2017, 39: 1137.
- [14] He K, Zhang X, Ren S, *et al.* Deep residual learning for image recognition [C]//Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016.
- [15] Lin M, Chen Q, Yan S. Network in network [EB/OL]. [2021-10-15]. <https://arxiv.org/pdf/1312.4400.pdf>.
- [16] Wu Y, He K. Group normalization[J]. Int J Comput Vision, 2020, 128: 742.
- [17] Berari R, Tauteanu A, Fichou D, *et al.* Overview of ImageCLEFdrawnUI 2021: the detection and recognition of hand drawn and digital website UIs task [C]//CEUR Workshop Proceedings. Bucharest: CEUR-WS, 2021.
- [18] 陈伟鹏, 柳戌昊, 唐宁九, 等. 基于灰色关联度和 AHP 的用户满意度综合评价模型[J]. 四川大学学报: 自然科学版, 2017, 54: 713.
- [19] Brooke J. Usability evaluation in industry [M]. London: Taylor & Francis, 1996.

引用本文格式:

中文: 梁旭东, 黎鸣, 赵慧玲, 等. 基于手绘草图的中保真原型自动生成[J]. 四川大学学报: 自然科学版, 2022, 59: 022005.

英文: Liang X D, Li M, Zhao H L, *et al.* Automatic generation of the mid-fidelity prototype based on hand-drawn sketch [J]. J Sichuan Univ: Nat Sci Ed, 2022, 59: 022005.