“为你写诗”：面向中国古典诗歌的可视化交互创作系统

封颖超¹⁾, 周姿含¹⁾, 张玮¹⁾, 谭思危¹⁾, 邵瑞敏²⁾, 陈佳舟^{3)*}, 陈为¹⁾

¹⁾ (浙江大学 CAD&CG 国家重点实验室 杭州 310058)

²⁾ (浙江大学中国古代文学与文化研究所 杭州 310058)

³⁾ (浙江工业大学计算机科学与技术学院 杭州 310012)

(cjz@zjut.edu.cn)

摘 要：中国古典诗歌具有工整的句式和悠扬的韵律，这些特点对创作者提出了较高的要求。现有的诗歌生成系统大多着眼于生成模型的效果，用户的参与度和发挥空间有限。基于古诗文网、THUNLP-AIPoet 等开源数据，设计并实现了一个面向中国古典诗歌的可视化交互创作系统“为你写诗”。其自定义视图支持自定义句式情感，“挥毫泼墨”视图自动生成诗歌作为创作基础，并且可视化诗歌文本的韵律情感等多维度属性；修改辅助视图在创作过程中提供改进建议和候选推荐。通过实例分析和用户调研对系统视图的功能和创作结果的连贯性、情感和韵律进行评价，结果表明了它的有效性和实用性。

关键词：人机回环；文本可视化；诗歌创作；诗歌自动生成；文本评分

中图法分类号：TP391.41 **DOI:** 10.3724/SP.J.1089.2021.18980

WPFY: Visual Interactive Authoring System for Chinese Classical Poetry

Feng Yingchaojie¹⁾, Zhou Zihan¹⁾, Zhang Wei¹⁾, Tan Siwei¹⁾, Shao Ruimin²⁾, Chen Jiazhou^{3)*}, and Chen Wei¹⁾

¹⁾ (State Key Laboratory of CAD&CG, Zhejiang University, Hangzhou 310058)

²⁾ (Department of Chinese Language and Literature, Zhejiang University, Hangzhou 310058)

³⁾ (College of Computer Science and Technology, Zhejiang University of Technology, Hangzhou 310012)

Abstract: Chinese classical poetry exhibits neat sentence patterns and melodic rhythms, which pose high challenges for creators. Most existing systems focus on generation models without considering user participation. Based on the open-source data of the Gushiwen website, THUNLP-AIPoet, we design and implement WPFY, a visual authoring system for Chinese classical poetry: a configuration view supports specify sentence patterns and emotions; an authoring view automatically generates poetry as the basis, and visualizes the multi-dimensional attributes of poetry such as rhythms and emotions; an auxiliary view presents suggested recommendations. User studies concerning the fluency, emotions, and rhythms of the creations demonstrate the usefulness and effectiveness of WPFY.

Key words: man-machine loop; text visualization; poetry creation; automatic poetry generation; text evaluation

收稿日期: 2021-03-08; 修回日期: 2021-03-16. 基金项目: 中央高校基本科研业务费(2-2050205-21-688). 封颖超杰(1997—), 男, 博士研究生, 主要研究方向为可视化、数字人文; 周姿含(1999—), 女, 博士研究生, 主要研究方向为可视分析; 张玮(1988—), 女, 博士研究生, 主要研究方向为可视化; 谭思危(1997—), 男, 博士研究生, 主要研究方向为数据科学; 邵瑞敏(1997—), 女, 博士研究生, 主要研究方向为唐宋文学; 陈佳舟(1984—), 男, 博士, 副教授, 硕士生导师, CCF 会员, 论文通讯作者, 主要研究方向为计算机图形学、数字人文; 陈为(1976—), 男, 博士, 教授, 博士生导师, CCF 杰出会员, 主要研究方向为可视化、可视分析、大数据分析和人机混合智能。

中国古典诗歌是中华民族的文化瑰宝,古往今来,文人墨客通过这种优美的语言表达形式抒发情感,留下了无数精彩的诗歌。例如,在耳熟能详的《静夜思》中,李白借明月寄托对家乡的思念之情,很多抒情名句就是在这样的心境下诞生的。可见情感是诗歌创作的一个重要的出发点。对情感表达的忽视可能导致情感崩溃或情感偏差,损害语义和多样性^[1]。

对于很多人来说,结合情感的诗歌创作是困难的。由于经验不足,缺少诗歌的积累和韵律上的指导,创作的内容不尽如人意。此外,对诗歌的润色也是一个值得深究的过程。而现有的系统创作诗歌的效果大多仅依赖生成模型的表现,用户在其中的参与度和发挥空间都很有限。

因此,本文提出了一个中国古典诗歌的可视化创作系统“为你写诗”。它以用户作为创作主体,允许用户自定义句式和情感。本文构建了一个情感诗歌生成模型,能够生成出具有特定情感、符合句式韵律的诗歌,为用户提供创作基础。同时,从连贯性、情感表达和韵律3个角度对诗歌进行实时分析,并且根据诗歌的字符级别的评分为用户提供改进建议。在用户修改诗歌的过程中,系统将候选词的预期评分和生成概率进行可视化。

1 相关工作

本节从文本可视化、诗歌生成方法与系统、文本评分方法3个主要方面回顾相关工作。

1.1 文本可视化

文本可视化以高效展示文本信息为目的,主要包括文本挖掘、可视化呈现、交互等过程^[2-3]。ThemeRiver^[4-5]用河流隐喻文本主题随时间变化的过程;InfoSky^[6]将文档可视化为银河系中的恒星,展现文档的层次关系;Tag Clouds^[7]以及后来的Wordle^[8]是显示词频的经典方法,关键词的大小对应了其在文本中的使用频率高低;Word Tree^[9]则是针对特定词语呈现上下文语境;Westeros Sentinel^[10]使用SenticNet 3^[11]进行概念级的情感分析,用雷达图对多维度的情感进行可视化。

本文采用雷达图呈现诗歌中不同情感的表达

程度,采用类似于Word Tree的形式呈现后缀候选词的生成概率,并将这些可视化形式与交互结合起来,便于用户理解和使用。

1.2 诗歌生成方法与系统

传统的诗歌生成方法主要通过统计机器学习方法进行诗歌生成,如遗传算法^[12-13]和统计机器翻译方法^[14]。深度学习的引入将诗歌生成的效果提升至新的高度。Zhang等^[15]提出了一个基于递归神经网络的生成模型,在诗句生成的过程中考虑全部的历史内容;Yan^[16]提出了一种迭代抛光的生成模式,对诗歌进行多次生成以提升诗歌的质量;对于诗歌的风格,Yi等^[17]将其解释为影响因素的组合,通过控制影响因素的取值形成不同的风格,从而提升生成诗歌的多样性。

上述诗歌生成的成果中,很多被应用于诗歌生成系统。“九歌”^①能够支持多模态的输入和体裁风格的控制,并推荐修改选项和相似诗歌提供参考^[18];“乐府”^②提供了命题作诗和看图作诗2种方法;“微软绝句”^③允许用户在预设的关键词中选择一个作为诗歌主题,并生成一系列候选诗句。

不同于以上的这些工作,本文系统支持用户自定义诗歌的情感及其强弱程度,展示诗歌的多维度属性,尽可能地减少了交互上的约束,使用户能够流畅地输入想表达的内容。

1.3 文本评分方法

BLEU^[19]是一种经典的文本评分方法,常用于机器翻译和图像标注领域,也是诗歌自动评分的常用方法之一^[14-16],它通过计算文本的 n -gram重合程度衡量评价文本的质量;此后,很多工作对BLEU的一些缺陷提出了改进^[20-21]。困惑度^[22]是将待评估的语言模型经过训练后,在测试集中计算文本的生成概率,生成概率越高,困惑度越小。

本文从多个维度提出了诗歌的评分方法,并将评分结果用于提供改进建议。

2 数据

本文使用的数据主要来自古诗文网^④,THUNLP-AIPoet^⑤、平水韵和词林正韵。古诗文网整理了从先秦到清代的诗歌数据,包含诗歌的译

① <http://jiuge.thunlp.org/>

② <https://bbs.huaweicloud.com/blogs/118144>

③ <http://duilian.msra.cn/jueju/>

④ <http://www.gushiwen.org/>

⑤ <https://github.com/thunlp-aipoet>

文、注释、赏析和类型. THUNLP-AIPoet 开源了数个有标注诗歌数据集. 本文将平水韵和词林正韵作为韵律分析的标准.

本文选择了 6 种常见的情感类型, 分别为思念远方、离别不舍、深切爱情、军旅悲壮、忧国忧民和咏史怀古. 经过整理, 总共得到了绝句诗约 12 万首. 其中, 带有情感类型标注的约 3000 首, 每首诗包含至少一种上述情感类型.

3 数据整合与模型

本节主要介绍数据处理流程和建模, 包括诗歌的韵律和句式特征提取、诗歌生成模型的结构和优化方法以及多维度的诗歌评分方法.

3.1 韵律和位置提取

韵是诗歌格律的基本要素之一^[23]. 绝句中的律绝对诗句的平仄有严格的要求, 分为平起式和仄起式. 押韵要求诗中偶数句末尾的字具有相同的韵母. 本文参考 Yi 等^[24]的方法提取诗中押韵位置汉字所属的韵部, 对其他位置的汉字提取平仄属性. 位置信息包括汉字在诗中的行数和句子中的次序, 这些信息用于控制诗句长度.

3.2 诗歌生成方法

本文的诗歌语料中的诗歌大部分是无标注的, 因此, 参考了 Yi 等^[17]的方法, 采用半监督条件变分自编码器(variational auto-encoder, VAE)^[25]对有标注诗歌和无标注诗歌进行学习.

用 $x = \{x_i\}_{i=1}^n$ 表示一首诗. 其中, n 为诗句的数量; $x_i = \{x_{i,j}\}_{j=1}^l$; $x_{i,j}$ 为单个字; l 为 x_i 的长度. 用 y 表示诗歌的情感, z 表示隐变量, 包含了诗歌的情感和语义等信息.

本文模型的任务是学习诗歌生成的条件概率分布 $p(x|y)$, 加入隐变量 z 后得到

$$p(x|y) = \int p(x, z|y) dz \quad (1)$$

$$p(x, z|y) = p(z|y)p(x|y, z) \quad (2)$$

对于 $p(x|y, z)$, 本文参考了 Yan^[16]的方法, 以长短时记忆模型(long short-term memory, LSTM)作为生成器构建了一个层级结构. 先用一个全局 LSTM 生成句子级别的内容表示, 即

$$s_{i,0} = \text{LSTM}^{\text{global}}(s_{i-1,-1}, [z]) \quad (3)$$

其中, $\text{LSTM}^{\text{global}}$ 为全局 LSTM; $s_{i,0}$ 为隐状态. 再用一个局部 LSTM 根据内容表示生成诗句, 即

$$s_{i,j} = \text{LSTM}^{\text{local}}(s_{i,j-1}, [e(x_{i,j-1}), g_{i,j}]) \quad (4)$$

$$p(x_{i,j} | x_{i,1:j-1}, x_{i,i-1}, y) = \text{softmax}(Ws_{i,j}) \quad (5)$$

其中, $\text{LSTM}^{\text{local}}$ 为局部 LSTM; $e(\cdot)$ 为词嵌入向量; $g_{i,j}$ 为一个体裁嵌入向量, 包含了第 3.1 节中提到的韵律和位置信息; W 为线性投影参数.

式(3)中的 z 为隐变量, 本文假设 z 是服从高斯分布的, 其分布空间从训练过程或生成过程中得到, 如图 1 所示. 然后从分布空间中采样得到隐变量 z . 因此, 模型的学习目标 $p(x|y)$ 的一个变分下界可以表示为

$$\ln p_\theta(x|y) \geq E_{q_\phi(z|x,y)}[\ln p_\psi(x|y,z)] - \text{KL}[q_\phi(z|x,y) \| p_\theta(z|y)] = -\mathcal{L}(x,y) \quad (6)$$

其中, E 为期望; KL 为相对熵; q_ϕ 为后验网络; p_θ 为先验网络; ϕ 和 θ 为模型参数. 对于无标注数据, 将不可见的情感标签看做另一个隐变量, 构建了一个情感推断模型 q_ω , 其模型结构在第 3.3 节中介绍. 对于这部分数据, $p(x)$ 的变分下界为

$$\ln p(x) \geq E_{q_\omega(y|x)}[-\mathcal{L}(x,y)] + \mathcal{H}(q_\omega(y|x)) = -\mathcal{U}(x) \quad (7)$$

本文将这 2 个下界函数作为模型的优化目标, 通过超参数 α , β 和 γ 调节它们的权重, 即

$$\mathcal{L}_{\text{VAE}} = \alpha \cdot E_{p_1(x,y)}[\mathcal{L}(x,y) - \gamma \cdot \ln q_\omega(y|x)] + \beta \cdot E_{p_u(x)}[\mathcal{U}(x)] \quad (8)$$

其中, $p_1(x,y)$ 和 $p_u(x)$ 分别对应有标注数据和无标注数据.

图 1 诗歌生成模型结构

3.3 诗歌评分方法

本节介绍连贯性、情感表达、韵律准确性 3 个维度诗歌评分方法.

3.3.1 连贯性评分方法

无论是自动评分^[26]还是人工评分, 连贯性都是诗歌评分的重要指标. 本文认为连贯性是局部的, 即文本中某处的连贯性不会受到远处的影响.

同时, 本文认为连贯性与情感是无关的, 因此任何情感的诗歌都可以作为评分方法的参考. 基于以上的观点, 本文提出了一个连贯性的评分算法, 即

$$score^f(d) = \tanh(c_d \cdot w).$$

其中, d 为被评分的文本序列; c_d 表示该序列在诗歌语料中出现的次数; w 为缩放系数, 可以根据语料库大小调整. 本文将 $x_{i,j}$ 与周围的字组成不同长度的序列, 对这些序列的连贯性评分加权求和, 作为 $x_{i,j}$ 的连贯性评分 $S_{i,j}^f = \max(S_{i,j}^{\text{pre}}, S_{i,j}^{\text{post}})$. 其中,

$$S_{i,j}^{\text{pre}} = \sum_{k=2}^4 score^f(x_{i,j-k:j}) \cdot \alpha_k;$$

$$S_{i,j}^{\text{post}} = \sum_{k=2}^4 score^f(x_{i,j:j+k}) \cdot \alpha_k;$$

α_k 为权重. 在此基础上, 本文用所有 $S_{i,j}^f$ 的平均值作为诗歌 x 的整体连贯性评分 $S^f = \sum S_{i,j}^f / (n \cdot l)$. $S_{i,j}^f$ 的计算过程需要多次遍历庞大的参照语料库, 因此, 本文构建并预训练了一个神经网络 q_f , 对 $S_{i,j}^f$ 进行估计.

3.3.2 情感评分方法

情感推断模型 q_ω 也是一个神经网络. 类似于 Kim^[27] 的工作, q_ω 以 x 的独热向量作为输入, 通过 $y_\rho = \text{sigmoid}(Wf + b)$ 推断诗中包含的情感. 其中, W 和 b 为线性投影参数; $f = \text{pool}(\text{conv}(e(x)))$. conv 为卷积层; pool 为池化层, 本文采用最大池化层. 定义用户输入的诗歌情感为 y_ε , 则 x 的情感表达评分为 $S^e = \frac{\sum (1 - |y_\varepsilon - y_\rho|) \cdot y_\varepsilon}{\sum y_\varepsilon}$.

3.3.3 韵律评分方法

本文使用第 3.1 节方法提取 x 的韵律 r_ρ , 与诗歌预设的韵律 r_ε 进行比较, 得到 x 的韵律评分为

$$S^r = \frac{\sum S_{i,j}^r}{n \cdot l};$$

其中,

$$S_{i,j}^r = \begin{cases} 1, & r_{i,j}^\varepsilon = r_{i,j}^\rho \\ 0, & r_{i,j}^\varepsilon \neq r_{i,j}^\rho \end{cases}.$$

综上, 本文实现了从连贯性、情感和韵律 3 个维度对创作诗歌评分的方法, 其中连贯性和韵律的评分是字符级别的. 此外, 在这 3 项评分中, 连贯性仍然是“诗人无关”的, 即没有对 VAE 模型的训练过程产生影响. 因此, 本文在式(8)的基础上增加了强化损失^[28], 得到

$$\mathcal{L}_{\text{DRL}} = -\sum_{p(x_{i,j})} E_{p(x_{i,j})} (S_{i,j}^f - S_{\text{mean}}^f),$$

$$\mathcal{L} = \mathcal{L}_{\text{VAE}} + \eta \cdot \mathcal{L}_{\text{DRL}}.$$

其中, S_{mean}^f 为训练集的连贯性平均分; η 为权重.

4 可视化创作系统

根据设计目标, 本文设计并实现了一个面向中国古典诗歌的可视化交互创作系统. 系统主界面包含了自定义视图、“挥毫泼墨”视图和修改辅助视图. 如图 2 所示, 用户创作从选择诗歌的句式和情感开始, 由诗歌生成方法生成一首诗歌. 用户可以任意修改诗歌内容, 系统针对选择的字词推荐候选词, 对候选词的预期评分和生成概率进行可视化, 修改更新后提供实时的反馈, 形成交互的闭环, 最终帮助用户创作出一首满意的诗歌.

图2 本文创作过程

4.1 自定义视图

自定义视图提供了句式和情感的选择. 如图 3 中 A 区域所示, 句式包括五言绝句和七言绝句. 为

了降低新用户的学习成本, 系统默认随机选择一种平仄的规则和句尾的韵部, 并以此作为诗歌韵律评分的参照 r_ε .

图3 本文系统主界面

图3中B区域是情感选择视图,用户可以选择一种或多种情感,选中后会出现C区域的情感雷达图中,每条轴上的圆点坐标对应了情感的强弱程度。用户选择和诗歌分析的情感分别用红蓝曲线表示,反映诗歌中不同情感的差异。

4.2 挥毫泼墨视图

用户完成自定义后,就可以在挥毫泼墨视图中开始创作。图3中的E区域是主要的创作区域,每个田字格容纳一个汉字。田字格的下方是韵律的可视化,图4a所示为五度标记法,展示了不同声调类型的相对音高走势。高水平的用户可以用声调为参考,检查诗句的平仄是否遵循规范。

图4 韵律可视化编码

图3中F区域展示了诗歌在连贯性、情感和韵律3个维度的评分,这些评分通过第3.3节中的算法得到,然后换算成10分制的数值。用户可以选择性地参考,并以此确定改进的方向。在创作和改进的过程中,系统实时分析诗歌内容,提供即时的反馈和交互。单击“分享诗歌”按钮可以将诗歌内容复制到系统剪切板,方便用户向他人分享。

图3中D区域是工具栏,包含了选择文本、修改建议、撤销和恢复功能。诗歌的连贯性和韵律评

分是字符级别的,本文在系统中设置了一个低分阈值,将得分低于该阈值的字符以醒目的背景样式进行呈现,为用户提供修改建议,如E区域的“霜”和“乡”所示。选择文本功能支持单选、多选和框选等多种方式,被选中的文字变成红色,与其他文字进行区分,如E区域的“十”“二”“乡”所示。选中内容后就可以在修改辅助视图中查看和挑选候选词。撤销和恢复功能允许用户在诗歌内容的历史版本之间自由切换。

4.3 修改辅助视图

修改辅助视图中提供了2种选择候选词的方法。一种方法是用户选择诗中内容后,系统针对选中的位置集束搜索,生成一组相同长度的候选词,并对替换后的诗歌重新评分,作为它们的预期得分,帮助用户评估候选词的改进效果。如图3中的G区域所示,系统将评分绘制在3条横轴上,横轴的数值范围根据评分的区间动态调整,更充分地利用横轴空间体现评分差异。候选词的3个评分通过贝塞尔曲线依次相连,初始词和候选词通过线条的颜色和宽度进行区分。用户单击候选词就可以完成修改。

另一种方法是逐字选择。如图3中H区域所示,将同一位置的不同选项绘制成一个圆环,背景图案的透明度编码了选项的生成概率。如果选中内容的长度大于1,则绘制2个圆环,外部圆环是内圈选项的后缀,单击内圈的选项查看它的后缀字。将满意的选项拖到中心区域,圆环更新为后续字符位置的高概率选项。按照这种方式逐字选择,直到候选词的长度和选中内容相同。单击“提交”按钮替

换选中的原始内容,单击“撤销”按钮重新选择。

5 实验与分析

本节通过实例分析和用户研究来介绍系统的功能和交互。图3所示的系统视图就是一个实际的案例,本节通过另一个实例分析实际创作过程。用户研究介绍参与者使用系统的过程和反馈。

5.1 实例分析

诗歌的创作从自定义视图开始,用户选择诗歌的句式和情感。如图5a所示,本例选择创作五

言绝句,表达思念远方、离别不舍和深切爱情3种情感,在情感雷达图中将3种情感的强弱程度分别设置为1.0,0.6和0.6。写诗前,用户预先设想了诗歌的情景:一位多愁善感的妻子送别了远行的丈夫,在家中写下这首诗表达对丈夫的思念和眷恋。

确定诗歌的句式和情感之后,单击下方的“系统生成”按钮,系统为用户生成一首诗作为基础,生成的诗歌如图5b所示。诗歌内容下方的实时评分中,韵律评分为满分,连贯性和情感评分都取得了8分以上的成绩,同时,诗歌的情感分析结果也实时地呈现在自定义视图的情感雷达图中。

图5 本文系统实例图

由于诗歌在连贯性和韵律2个方面的较高评分,系统不再提供修改建议。但经过理解和体会,用户认为第3句“更是无情思”中的“无”不太恰当,于是通过工具栏的选择功能选中这个字,如图5c区域所示。在系统生成的候选字中,用户认为“多”字更能体现诗歌的情感,右侧的预期得分也印证了用户的想法,所以用户单击“多”字完成替换,完成了一次修改过程。至此,本例创作的诗歌内容为“花开春已晚,昨夜梦魂惊。更是多情思,何如怨别声。”用户还可以通过相似的操作不断地改进诗句。

5.2 用户研究

为了验证为你写诗系统的可用性,本文进行了一项用户研究。该研究包含系统介绍、自由创作、问卷评估3个环节。

本文召集了7位用户(4位男性,3位女性)参与,包括2位专家用户和5位普通用户。2位专家用户在诗歌领域有3年以上的研究经历,普通用户中有一位具有2年左右的数据可视化经验,其他几位没有相关研究经历或非常有限。

在系统介绍环节,按照诗歌创作的流程介绍系统界面和功能,具体过程与第5.1节中的实例分析类似。该环节大约持续20 min,在介绍完成以后,所有参与者都表示能够理解系统功能的使用方式和设计意图。

在自由创作环节,参与者需要独立完成一次诗歌创作和改进的完整过程,创作内容不限。对参与者的操作进行了记录。该过程持续5~15 min,所有参与者都在没有提示的情况下完成了这一过程,图6所示为用户研究中创作的诗歌。

黄河通汉阙 边城远接水连淮 乡山望断楚云初
 白月上天台 一夜烽烟万里来 万里长城一夜无
 但见龙沙客 行尽西风天上月 估客归来无尽说
 谁知戍马来 胡笳吹入戍楼台 城前不见白头书

图6 本文用户创作结果

在问卷评估环节,本文在问卷中设置了一系列半结构化的问题.每个问题中包含一个评分项,评分采用 5 分制,1 分表示非常不喜欢,5 分表示非常喜欢.用户在评分之后,还需要提供一些主观说明,说明的内容是开放式的.在问卷评估结束之后,对数据进行了整理,统计的结果如下.

(1) 在雷达图对于不同情感的呈现效果上,4/7 非常喜欢,3/7 比较喜欢.参与者认为“情感差异用雷达图来表示清晰明了”“可视化做得非常棒,不同情感的差异一目了然”“我能够按自己的喜好来调节不同情感的程度”.

(2) 关于挥毫泼墨视图中提供的修改建议是否易用和实用的问题,3/7 非常喜欢,4/7 比较喜欢.所有的参与者都给予了肯定的评价,如“非常人性化”“采用高亮背景突出,效果很好,易懂有效”“大大增强了用户后续修改的便利性”.

(3) 在修改辅助视图的满意度上,3/7 非常喜欢,4/7 比较喜欢.其中,4/7 更喜欢“雕章琢句”视图中的方法,因为它“简洁明了”“帮助我更好地判断和填词”;3/7 更喜欢逐字选择的方式,因为它“提供的自主选择性比较大”“水墨风格的呈现方式,更加符合为你写诗的整体风格”.

(4) 最后是对系统功能与交互的整体评价,3/7 非常喜欢,4/7 比较喜欢.大部分参与者都给予了较高的评价,如“带给我的视觉体验极佳,这就是我想象中的可视化佳作”“功能上十分健全,类似宣纸的底色让人心旷神怡,很有写诗的欲望”“不仅实现了‘为你写诗’,还通过修改功能激发了用户自己写诗的主动性和积极性”.

除了上述评价,参与者也提出了很多有价值的建议,如在“修改建议”中提供判断依据,对字词提供用法、典故等注释信息,构建一个分享社区.这些建议充分反映出用户创作过程中的需求,对于提高系统的用户体验有很大帮助.

本文也请参与者对自己创作的诗歌进行满意度评分,评分结果的平均分如表 1 所示.总体来看,不论是专家用户还是普通用户,对于自己创作的诗歌都是比较满意的,特别是在诗歌的韵律方面,诗歌的平仄押韵让他们感受到音律上的美感.在评分相对较低的情感表达上,专家指出诗歌的情

感表达比较平面化,在情感的递进升华等方面还有改进的空间.

6 结 语

本文提出了一种情感诗歌的生成方法,能够生成具有特定情感的诗歌,并设计了一套诗歌自动评分的方法,从连贯性、情感表达、韵律 3 个角度对诗歌进行评分.在此基础上,本文设计并实现了一个面向中国古典诗歌的可视化交互创作系统,将诗歌文本的多维属性可视化,提供多种交互工具帮助用户创作和改进.用户研究的结果表明,本文系统提供了良好的用户体验,能够帮助用户创作出满意的作品.

本文也从用户研究中挖掘出了更多实际创作的功能需求,如诗中字词用法和典故信息的可视化、诗歌作品的社区分享、情感表达更好的诗歌生成方法,这将是未来工作的方向.

本文的工作以用户为主体,以中国古典诗歌和人工智能为支撑,通过可视化将二者连接在一起,是中国古代文学、可视化和人工智能相结合的一次新的尝试.如何更好地借助以人工智能为代表的现代信息技术探索文学中的规律,如何利用可视化技术将这些信息传递给用户,更好地为大众服务,是未来持之以恒探索的目标.

参考文献(References):

- [1] Chen H M, Yi X Y, Sun M S, *et al.* Sentiment-controllable Chinese poetry generation[C] //Proceedings of the 28th International Joint Conference on Artificial Intelligence. San Francisco: Morgan Kaufmann, 2019: 4925-4931
- [2] Wu Yingcai, Cui Weiwei, Song Yangqiu, *et al.* A survey on topic-based visual text analytics[J]. Journal of Computer-Aided Design & Computer Graphics, 2012, 24(10): 1266-1272(in Chinese)
(巫英才, 崔为炜, 宋阳秋, 等. 基于主题的文本可视分析研究[J]. 计算机辅助设计与图形学学报, 2012, 24(10): 1266-1272)
- [3] Tang Jiayu, Liu Zhiyuan, Sun Maosong. A survey of text visualization[J]. Journal of Computer-Aided Design & Computer Graphics, 2013, 15(3): 273-285(in Chinese)
(唐家渝, 刘知远, 孙茂松. 文本可视化研究综述[J]. 计算机辅助设计与图形学学报, 2013, 15(3): 273-285)
- [4] Havre S, Hetzler B, Nowell L. ThemeRiver: visualizing theme changes over time[C] //Proceedings of the IEEE Symposium on Information Visualization. Los Alamitos: IEEE Computer Society Press, 2000: 115-123
- [5] Havre S, Hetzler E, Whitney P, *et al.* ThemeRiver: visualizing

表 1 参与者的满意度评分

用户类型	连贯性	情感表达	韵律
专家用户	4.0	3.0	4.5
普通用户	4.4	4.2	4.8

- thematic changes in large document collections[J]. IEEE Transactions on Visualization and Computer Graphics, 2002, 8(1): 9-20
- [6] Andrews K, Kienreich W, Sabol V, *et al.* The InfoSky visual explorer: exploiting hierarchical structure and document similarities[J]. Information Visualization, 2002, 1(3/4): 166-181
- [7] Viégas F B, Wattenberg M. TIMELINE: tag clouds and the case for vernacular visualization[J]. Interactions, 2008, 15(4): 49-52
- [8] Viégas F B, Wattenberg M, Feinberg J. Participatory visualization with Wordle[J]. IEEE Transactions on Visualization and Computer Graphics, 2009, 15(6): 1137-1144
- [9] Wattenberg M, Viégas F B. The word tree, an interactive visual concordance[J]. IEEE Transactions on Visualization and Computer Graphics, 2008, 14(6): 1221-1228
- [10] Scharl A, Hubmann-Haidvogel A, Jones A, *et al.* Analyzing the public discourse on works of fiction - detection and visualization of emotion in online coverage about HBO's game of thrones[J]. Information Processing & Management, 2016, 52(1): 129-138
- [11] Cambria E, Olsher D, Rajagopal D. SenticNet 3: a common and common-sense knowledge base for cognition-driven sentiment analysis[C] //Proceedings of the 28th AAAI Conference on Artificial Intelligence. Palo Alto: AAAI Press, 2014: 1515-1521
- [12] Manurung H N. An evolutionary algorithm approach to poetry generation[D]. Edinburgh: University of Edinburgh, 2003
- [13] Zhou Changle, You Wei, Ding Xiaojun. Genetic algorithm and its implementation of automatic generation of Chinese SONGCI [J]. Journal of Software, 2010, 21(3): 427-437(in Chinese)
(周昌乐, 游维, 丁晓君. 一种宋词自动生成的遗传算法及其机器实现[J]. 软件学报, 2010, 21(3): 427-437)
- [14] He J, Zhou M, Jiang L. Generating Chinese classical poems with statistical machine translation models[C] //Proceedings of the 26th AAAI Conference on Artificial Intelligence. Palo Alto: AAAI Press, 2012: 1650-1656
- [15] Zhang X X, Lapata M. Chinese poetry generation with recurrent neural networks[C] //Proceedings of the Conference on Empirical Methods in Natural Language Processing. Stroudsburg: ACL, 2014: 670-680
- [16] Yan R. i, Poet: automatic poetry composition through recurrent neural networks with iterative polishing schema[C] //Proceedings of the 25th International Joint Conference on Artificial Intelligence. San Francisco: Morgan Kaufmann, 2016: 2238-2244
- [17] Yi X Y, Li R Y, Yang C, *et al.* MixPoet: diverse poetry generation via learning controllable mixed latent space[C] //Proceedings of the AAAI Conference on Artificial Intelligence. Palo Alto: AAAI Press, 2020, 34(5): 9450-9457
- [18] Guo Z P, Yi X Y, Sun M S, *et al.* Jiuge: a human-machine collaborative Chinese classical poetry generation system[C] //Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics. Stroudsburg: ACL, 2019: 25-30
- [19] Papineni K, Roukos S, Ward T, *et al.* BLEU: a method for automatic evaluation of machine translation[C] //Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics. Stroudsburg: ACL, 2002: 311-318
- [20] Banerjee S, Lavie A. METEOR: an automatic metric for MT evaluation with improved correlation with human judgments[C] //Proceedings of Workshop on Intrinsic and Extrinsic Evaluation Measures for Machine Translation and/or Summarization. Stroudsburg: ACL, 2005: 65-72
- [21] Lin C Y. ROUGE: a package for automatic evaluation of summaries[C] //Proceedings of Workshop on Text Summarization Branches Out. Stroudsburg: ACL, 2004: 74-81
- [22] Manurung R, Ritchie G, Thompson H. Using genetic algorithms to create meaningful poetic text[J]. Journal of Experimental & Theoretical Artificial Intelligence, 2012, 24(1): 43-64
- [23] Wang Li. Rhyme of poetry[M]. Beijing: Zhonghua Book Company, 2009(in Chinese)
(王力. 诗词格律[M]. 北京: 中华书局, 2009)
- [24] Yi X Y, Sun M S, Li R Y, *et al.* Chinese poetry generation with a working memory model[C] //Proceedings of the 27th International Joint Conference on Artificial Intelligence. San Francisco: Morgan Kaufmann, 2018: 4553-4559
- [25] Kingma D P, Rezende D J, Mohamed S, *et al.* Semi-supervised learning with deep generative models[C] //Proceedings of the 27th Neural Information Processing Systems. Cambridge: MIT Press, 2014: 3581-3589
- [26] Yi X Y, Sun M S, Li R Y, *et al.* Automatic poetry generation with mutual reinforcement learning[C] //Proceedings of the Conference on Empirical Methods in Natural Language Processing. Stroudsburg: ACL, 2018: 3143-3153
- [27] Kim Y. Convolutional neural networks for sentence classification[OL]. [2021-03-08]. <https://arxiv.org/abs/1408.5882>
- [28] Williams R J. Simple statistical gradient-following algorithms for connectionist reinforcement learning[J]. Machine Learning, 1992, 8: 229-256