

文章编号: 1006-3080(2025)04-0538-14

DOI: 10.14135/j.cnki.1006-3080.20240831001

基于交互式进化算法的三维表情动画生成

张雯婷, 刘漫丹, 朱宝旭, 谢立志

(华东理工大学信息科学与工程学院, 上海 200237)

摘要: 三维面部表情动画已广泛融入人们的日常生活, 但目前大多数表情生成技术并未考虑到实际用户的偏好。这些方法大多依赖于预先定义的心理原型, 限制了其生成丰富且具有表现力的表情动画的潜力。针对这两个问题, 本文提出了一种全新的表情动画编码方案, 并采用交互式遗传算法 (Interactive Genetic Algorithm, IGA) 实现表情动画的自动进化, 解码时则采用 Dirichlet 自由变形算法 (Dirichlet Free-Form Deformation, DFFD) 控制人脸网格发生形变; 为了维持表情动画种群的多样性, 提出了两种变异算子: 带屏蔽区间的均匀变异和共享变异点的均匀变异。实验结果证实这两种算子有效降低了进化过程中种群内的相似度。实际用户参与的实验进一步验证了所提方法在生成符合用户偏好且丰富多样的表情动画方面的可行性。

关键词: 表情动画生成; 交互式遗传算法; 用户评价; 变异算子; 种群多样性

中图分类号: TP391

文献标志码: A

随着网络技术的不断发展, 人与人之间面对面的交流越来越少, 取而代之的是通过各种社交媒体, 使用文字或者是表情符号去交流。虽然表情符号可以辅助情感表达, 但其真实度十分有限。而且人类情感极其复杂且丰富, 即使传达的是同一个表情符号, 在不同的使用者看来, 也可能是不同的情绪, 所以用户在选用表情符号或动画时经常会遇到无法找到任何一个能准确表达自我的表情符号的情况。除情感表达之外, 三维表情动画也越来越多地应用在影视制作、虚拟现实、人机交互、医疗保健、心理研究等各种领域。因此, 提出一种支持用户引导的表情动画生成方法具有十分重要的意义。

现今大多数的表情生成技术都是生成指定心理原型下的表情, 如开心、惊讶等。例如深度神经网络在生成 6 种基本表情图像时表现突出, 但在生成难以描述或非典型心理原型相关的表情时, 却表现不佳, 导致可生成的表情类型受到限制。对于真实感更高的三维表情动画, 目前大多是采用数据采集设备或者专业的建模软件得到^[1-2], 过高的财力和人力成本

使得这些方法不具有普适性, 无法推广给用户个人使用。另外, 三维表情质量的高低很大程度上取决于用户个人的喜好, 一些用户认为搞怪夸张的表情更好, 另一些用户则可能偏好真实生动的表情, 这很难通过具体的数量指标来说明。若能在表情生成过程中直接将用户的个人想法考虑到其中, 定量评价指标缺失的问题将会得到有效的解决。

本文采用交互式进化算法 (Interactive Genetic Algorithm, IGA)^[3] 实现表情动画的自动进化, 基于由 FaceGen^[4] 导出的标准三维人脸模型, 结合 Dirichlet 自由变形算法 (Dirichlet Free-Form Deformations, DFFD)^[5], 提出了一套全新的表情动画个体编码。与传统的遗传算法 (Genetic Algorithm, GA)^[6] 相比, IGA 根据用户评价获取每个个体的适应度值, 这使得毫无专业知识背景的普通用户参与表情进化的过程中, 引导表情动画朝向个人偏好的方向进化。进化开始前初始种群随机生成, 无需事先给定相应心理原型下的目标表情编码, 进化过程由用户把控, 通过不断的进化迭代, 使进化得到的表情不再

收稿日期: 2024-08-31

作者简介: 张雯婷 (2001—), 女, 江西人, 硕士生, 主要从事交互式进化算法研究。E-mail: wtzhang2001@163.com

通信联系人: 刘漫丹, E-mail: liumandan@ecust.edu.cn

引用本文: 张雯婷, 刘漫丹, 朱宝旭, 等. 基于交互式进化算法的三维表情动画生成 [J]. 华东理工大学学报 (自然科学版), 2025, 51(4): 538-551.

Citation: ZHANG Wenting, LIU Mandan, ZHU Baoxu, et al. 3D Facial Expression Animation Generation Based on Interactive Evolutionary Algorithm[J]. Journal of East China University of Science and Technology, 2025, 51(4): 538-551.

局限于一些典型的心理原型,如“快乐”、“悲伤”等,相反基因解码后得到表情个体所对应的心理原型可能是用户所理解的“无语”、“假笑”等,甚至无法用简单的言语描述,全凭用户个人理解。但也正因为引入了用户评价过程,IGA通常会面临用户疲劳和种群多样性过早丧失的问题。所以本文选择采用多用户评价的模式,并针对性地提出了两种变异算子,旨在降低种群内表情个体之间的相似性,进化出丰富多样、符合大众审美的表情动画。

1 相关工作

1.1 表情生成

除情感表达之外,三维表情生成技术因其在游戏或电影角色制作^[7-8]、虚拟现实^[9-10]、人机交互^[11-12]、医疗美容^[13]、心理研究^[14-15]等领域的广泛应用而越来越多地受到学者们的关注。当前主流的表情生成技术主要分为两种:基于计算机图形学的表情合成方法和基于深度学习的表情合成方法。早期基于计算机图形学的方法通过变形三维模型得到表情,其中多需要专业人士参与调控^[16-17]。如, Liu等^[18]采用基于人工建模的图像插值和变形方法,通过对多帧不同表情的人脸纹理和细节进行线性插值以获取中间表情,从而得到连贯的表情动画。Zhang等^[19]提出的几何驱动的面部表情合成系统能够在面部表情特征点位置给定的前提下,自动合成具有相应纹理细节的表情图像。这类方法生成的表情精度很大程度上取决于使用的三维人脸模型的精度,故表情的精细度一直有待提升。基于生理结构的面部动作编码系统(Facial Action Coding System, FACS)^[20],定义了44个模拟面部肌肉运动的动作单元(Action Units, AUs)。崔韶^[21]基于柔性AUs识别和刚性姿态估计的手段,将用户实际的表情同步到虚拟角色上,但这在实际中并没有直接生成表情动画。Cuculo等^[22]提出的OpenFACS系统允许用户交互式地直接选择需要激活的AUs以及激活的强度,从而直接获取需要模拟的人脸表情。但由于AUs数量有限,采用FACS生成的表情通常都面临多样性不足的问题。基于深度学习的表情合成方法以生成对抗网络(GAN)^[23]为代表, GAN在生成二维表情图像方面表现十分突出,它的各种变体^[24-26]也被学者们不断提出。Pumarola等^[27]提出的GANimation网络结合FACS,允许控制每个动作单元的激活幅度,这使得合成的表情图像强度连续可变,克服了许多网络仅生成强度单一的表情的问题。林本旺^[28]提出的稠密分组卷积生成对

抗网络减轻头部姿态对于面部表情生成带来的影响,但应用场景也还是局限于6种基本表情图像的生成。Wang等^[29]提出了一种基于GAN的地标引导残差模块,用于捕捉微妙的表情运动,该方法可以很好地将视频序列中的任意表情转移到源图像上。然而,利用神经网络直接生成三维人脸表情的相关研究仍然较少,主要是由于三维人脸数据的敏感性使得获取合适的训练数据集较为困难^[30]。

以上这些方法都面临着两个问题:(1)缺少用户的主观评价,即没有考虑用户的个性化需求;(2)无法生成复杂多样的心理原型对应的表情动画。为此,本文设计表情动画的基因编码,采用交互式进化算法将用户的偏好引入进化过程,根据编码规则随机生成初始种群,解码后得到的表情动画对应的心理原型和适应度值完全由用户感知。

1.2 交互式遗传算法

IGA的概念最早是由日本学者Takagi^[3]提出,其最核心的思想是将人的主观评价引入到传统的遗传算法中,个体的适应度由用户决定,这使得IGA在解决隐式优化问题时表现突出。

目前IGA在产品设计^[31-34]、艺术创作^[35-37]、图像处理^[38]与检索^[39-40]等领域都有广泛的应用。但其在表情动画生成方面的应用还较少。2018年,Hailemariam等^[41]采用IGA进化表情动画,他们根据面部F曲线设计编码,在最初分别给定嘴部、眼部和头部的若干种动作曲线,进化过程中仅两父代相同的关键帧之间可以进行交叉,再将不同部位的动作曲线进行排列组合。这种情况下所有表情动画的帧数和帧率都相同,而且每个部位的动作都局限在最开始给出的动作库中,合成的表情不够多样。同年Reed等^[42]提出的基于用户引导的表情动画生成也具有相同的问题,所生成的表情都局限于FACS中所提供的44个运动单元的排列组合,很难进化出令人感到惊喜的表情。故本文的编码结合DFFD算法,旨在能够生成新颖、丰富多样的三维表情动画。

1.3 Dirichlet自由变形算法

自由变形(Free-Form Deformation, FFD)算法^[43]是一种与物体表示无关的变形方法,它能够被广泛地应用于计算机动画和几何建模中。本文使用的DFFD算法^[5]是FFD的一种传播较广的改进,其特点在于引入了Sibson局部坐标系^[44]之后,算法允许控制点集合在合理范围内自由设定。一旦控制点确定,每个顶点的Sibson坐标将保持不变,这避免了在多次变形过程中重复计算顶点坐标,从而提高了算法处理高维数据的效率。当控制点集 $P_c = \{p_c | i =$

1, 2, \dots, n} 和待变形的点集 $P = \{p_i | i = 1, 2, \dots, q\}$ 确定后(其中 p_c 和 p_i 均表达了具有三维坐标的空间中的某个点), 依次找到待变形顶点 p_i 的 Sibson 邻居点集合 $P_s^{(i)} \{p_{s_j}^{(i)} | j = 1, 2, \dots, m, p_{s_j}^{(i)} \in P_c\}$, $P_s^{(i)} \subseteq P_c$, 便可计算 p_i 的 Sinson 坐标 $X^{(i)} = \{x_j^{(i)} | j = 1, 2, \dots, m\}$, 其中 m 为点 p_i 在控制点集 P_c 下的 Sibson 邻居数。点 p_i 变形后的坐标可以根据式 (1)^[45] 计算得出, 其中 $\Delta p_{s_j}^{(i)}$ 是点 p_i 的第 j 维个 Sibson 邻居点的坐标增量, 这由控制点集的坐标增量 $\Delta P_c = \{\Delta p_c | i = 1, 2, \dots, n\}$ 给出。

$$p_i' = p_i + \sum_{j=1}^m x_j^{(i)} \Delta p_{s_j}^{(i)} \quad (1)$$

考虑到面部表情的变化主要体现在眼睛、眉毛和嘴部, 本文的基因编码会主要针对这 3 个部位进行设计, 其中眉毛和嘴巴两部分的动作使用变形算法控制。本文首先选取了 10 个面部关键点作为第 1 级变形的控制点集合 $P_c^1 = \{p_c^1 | i = 1, 2, \dots, 10\}$, 如图 1 所示, 分别是左右眉的眉尖、眉峰和眉尾, 左右嘴角、上下嘴唇的中点。为优化变形效果, 第 1 级变形将控制点集 P_c^1 作为待变形的点, ΔP_c^1 可根据第 1 级变形的结果计算得出。第 2 级变形则通过控制点集 P_c^1 的移动去牵引整个面部网格其他顶点的移动, 最终表现为面部肌肉的运动。由于 P_c^1 数量较多且选取方式不唯一, 这里不再列出。

2 交互式进化算法的实现

2.1 IGA 进化流程

IGA 算法流程与 GA 基本一致, 差别在于个体适应度值不再由适应度函数计算, 而是根据用户的评价直接给出, 本文 IGA 的算法流程如图 2 所示。在开始进化前, 需要根据设计好的基因编码规则随机生成初始种群, 解码后呈现给用户的表情个体为 gif 动图文件。区别于常规的 IGA, 本文的进化流程有两点不同: (1) 引入精英替换策略, 对于当前代的种

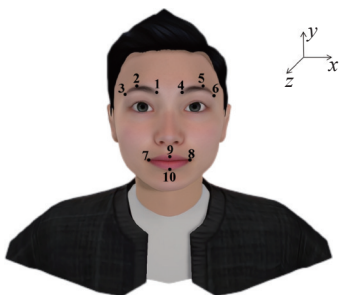


图 1 面部关键点

Fig. 1 Facial key points

群中适应度最低的两个个体, 用父代种群中适应度值最高的两个个体替换; (2) 在开始进化之前无需预设目标表情动画, 支持用户在任意一代保留自己偏好的个体而不终止进化, 故希望每一代种群中的个体尽可能多样化。

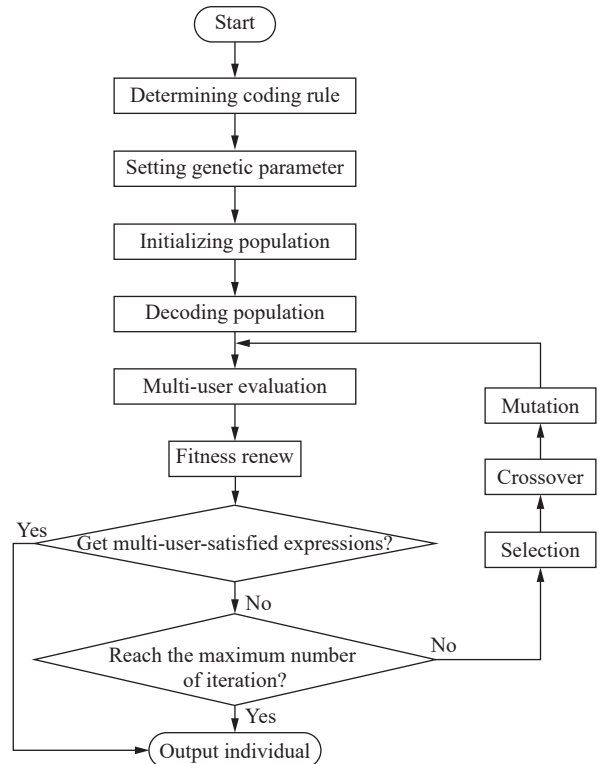


图 2 交互式进化算法流程图

Fig. 2 Flow chart of interactive genetic algorithm

2.2 编码方式

本文进化得到的三维面部表情动画最终以 gif 文件呈现, 一个表情动画的变化过程如图 3 所示。为了使个体的表现更加丰富, 每个表情动画个体包含两个静态目标表情 Expression 1 和 Expression 2, 初始的 Expression 0 为不带任何表情的标准人脸。在图 3 中, 从 Expression 0 变化到 Expression 1, 再变化到 Expression 1', 是一个完整的表情变化过程, 同样从 Expression 1' 到 Expression 2, 再到 Expression 2', 是第 2 个完整的变化过程, 两个过程的变化规则相同。

以第 1 个完整的表情变化过程为例, 如图 4 所示, 我们考虑 6 个部位的动作, 分别是左右眉、左右眼、嘴部和头部。每个部位单独运动, 这意味着不同部位动作的速度、起始帧和结束帧互不影响。同时每个部位都有 4 个动作过程, 依次是 P1-s, P1-m, P2-s 和 P2-m, 如图 5 所示, 图中矩形条的长度表示对应过程耗费的帧数。以嘴部的动作为例, 过程 P1 为朝向目标状态的变化过程, 过程 P2 为回复过程, 即当嘴部从初始状态 Mouth 0 变化到目标表情 Expression 1



图 3 表情动画的变化过程

Fig. 3 Changing process of an expression animation

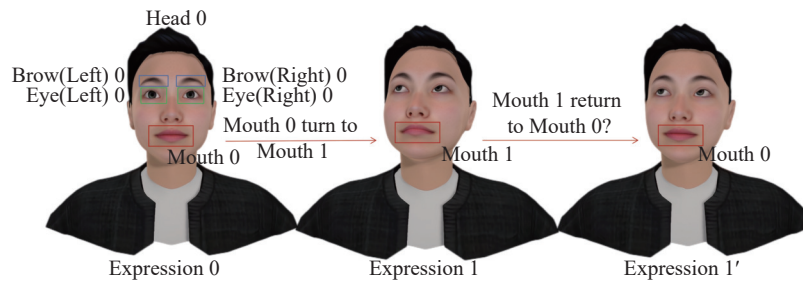


图 4 第 1 个完整的目标表情变化过程

Fig. 4 Whole changing process of the first target expression

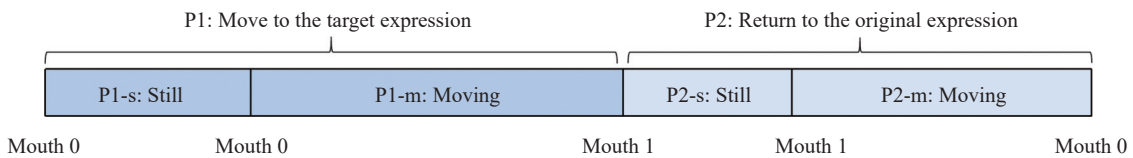


图 5 一个目标表情变化过程的 4 个动作过程(以嘴部动作为例)

Fig. 5 Four movement processes for one target expression (Taking mouth as an example)

的嘴部状态 Mouth 1 后,再回到初始状态 Mouth 0 的过程。但是此回复过程是否存在取决于嘴部的回复控制位,若该控制位基因的值为 0 则过程 P2 不存在。Px-s 和 Px-m (x 为 1 或 2) 决定着动作过程 Px 的起始帧与结束帧, Px-s 为动作开始前保持静止的帧数, Px-m 为持续动作的帧数,持续的帧数越少,动作变化越快。

综上,染色体编码需要包含两个目标表情的信息以及表情变化过程的帧数信息。图 6 示出了个体表情动画的染色体结构,1 条染色体由 4 段基因组成 (C1~C4),共有 134 个基因位。各基因位点的含义在单个目标表情的编码规则(表 1)和各部位 4 个动作过程的帧数信息编号规则(表 2)中也有说明。C1 和 C2 控制 Expression 1 的变化过程, C3 和 C4 控制 Expression 2 的变化过程。表情个体的染色体编码向量可以表示为 $I_a = \{c_{a1}, c_{a2}, \dots, c_{ai}, \dots, c_{a134}\}$, 其中 c_{ai} 表示第 i 个基因位。

本文采用实数编码的方式,首先两个静态目标表情的编码规则相同,表 1 示出了单个目标表情的编码规则,共 37 位基因,对应图 6 中的 C1 和 C3 基因段。就单个表情而言,左右眉和嘴部的动作形态通

过图 1 中 10 个面部关键点的坐标变化量 ΔP_c^i 控制 (其中上角标 1 表示第 1 层变形,下角标 c 表示控制点集合)。从表中可以看出,对于左右眉,我们选取了眉尖、眉峰和眉尾的 x 、 y 坐标增量,角标 i 代表的是图 1 中关键点 p_c^i 的索引。另外,对于每个部位均设有对称约束位,用于控制左右两边的动作是否对称,故目标表情中左眉和右眉的形态共由 13 位基因 (C1 中 $c_{a1} \sim c_{a13}$ 、C3 中 $c_{a68} \sim c_{a80}$) 控制。嘴部共由 11 位基因 (C1 中 $c_{a21} \sim c_{a31}$ 、C3 中 $c_{a88} \sim c_{a98}$) 控制,本文选取了左右嘴角的 x 、 y 坐标,以及上下嘴部中点的 x 、 y 、 z 坐标增量,同样,对称约束位用于控制左右嘴角的动作是否一致。为了防止出现过分夸张的表情,我们事先给出了各个坐标值变化的上限。除了控制下嘴部中点 y 坐标的基因取值范围在 $[-1, 0.5]$ 外,其他基因取值都在 $[-1, 1]$ 。基因与相应的上限相乘即可得到图 1 中 10 个关键点的坐标增量 ΔP_c^i 。

眼部动作包括眼球的旋转和眼皮的闭合程度,基因取值均在 $[-1, 1]$,同样眼部也设一个对称约束位,用于控制左右眼的动作是否一致,故静态目标表情中控制眼部动作的基因共 7 位 (C1 中 $c_{a14} \sim c_{a20}$ 、C2 中 $c_{a81} \sim c_{a87}$)。头部涉及上下旋转、左右旋转和

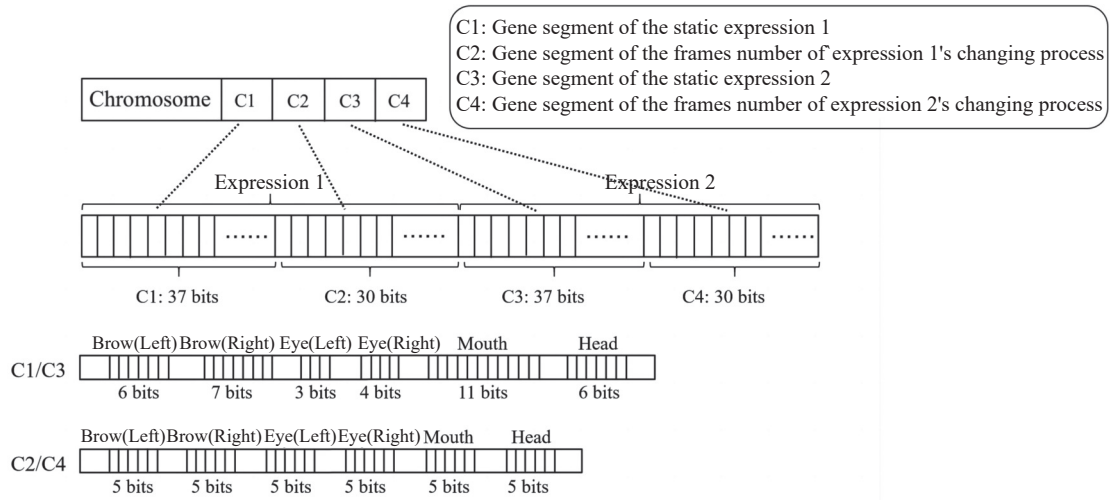


图 6 个体表情动画的染色体结构

Fig. 6 Chromosomal structure of expression animation individuals

歪头 3 个动作, 控制头部动作的基因共 6 位 (C1 中 $c_{a32} \sim c_{a37}$ 、C3 中 $c_{a99} \sim c_{a104}$), 其中 3 位用于控制这 3 个动作是否发生, 另外 3 位控制相应动作的幅度, 取值在 $[-1, 1]$ 。故控制一个静态目标表情的基因编码共 37 位, 即 C1 和 C3 基因段各 37 位, 分别负责控制两个静态目标表情。

表 1 中 6 个面部部位动作的过程主要由图 6 中的 C2 和 C4 基因段(各有 30 个基因位)控制。表 2 示出了 C2 和 C4 基因段中各基因位点的含义, 从表中可以看出, 对于每个部位都有相应的基因位点去控制 4 个动作过程(P1-s、P1-m、P2-s、P2-m)的帧数, 另外每个部位均设有回复控制位“Return or not”。为了控制生成的表情动图 gif 文件的大小, 解码时可以预设最大帧数的限制, 故控制 4 个动作过程的帧数的基因位取值范围在 0~1 之间, 同样与帧数上限相乘并

取整后得到相应过程的实际运动帧数。回复控制位“Return or not”取值为 0 或 1。

综上, 一个完整的表情变化过程的编码共有 67 位 (C1+C2 或 C3+C4 基因段), C1 和 C3、C2 和 C4 中的基因值不同, 但含义基本相同, 所以一个表情动画个体共包含两个相互独立的表情变化过程, 染色体共有 $67 \times 2 = 134$ 位基因。

2.3 遗传策略

针对本文表情动画的个体编码规则, 本文初步选取了表现较好的平坦交叉^[46]和均匀变异^[47]策略。如式(2)所示, 平坦交叉依次操作每一位基因, 子代的基因值随机落在两父代对应的基因值之间, 交叉一次只得到一个子代。其中两父代 I_a 与 I_b 的染色体编码为: $I_a = \{c_{a1}, c_{a2}, \dots, c_{ai}, \dots, c_{a134}\}$, $I_b = \{c_{b1}, c_{b2}, \dots, c_{bi}, \dots, c_{b134}\}$, u_i 为 0~1 之间的随机数, c'_i 为子

表 1 单个静态目标表情的编码规则

Table 1 Coding rules of a single static expression

Facial part	Meaning of gene locus	Number of gene locus
Brow (Left)	Coordinate increments: $\Delta x_i, \Delta y_i, i = 1, 2, 3; \Delta x_i, \Delta y_i \in [-1, 1]$	$c_{a1} \sim c_{a6} / c_{a68} \sim c_{a73}$
Brow (Right)	Coordinate increments: $\Delta x_i, \Delta y_i, i = 4, 5, 6; \Delta x_i, \Delta y_i \in [-1, 1]$; Symmetrical or not for brow	$c_{a7} \sim c_{a13} / c_{a74} \sim c_{a80}$
Eye (Left)	Amplitude of turn up-down; Amplitude of turn left-right; Degree of eyelid closure	$c_{a14} \sim c_{a16} / c_{a81} \sim c_{a83}$
Eye (Right)	Amplitude of turn up-down; Amplitude of turn left-right; Degree of eyelid closure; Symmetrical or not for eye	$c_{a17} \sim c_{a20} / c_{a84} \sim c_{a87}$
Mouth	Coordinate increments: $\Delta x_i, \Delta y_i, i = 7, 8, 9, 10; \Delta z_i, i = 9, 10;$ $\Delta x_i, \Delta y_i, \Delta z_i \in [-1, 1], y_{10} \in [-1, 0.5]$; Symmetrical or not for mouth	$c_{a21} \sim c_{a31} / c_{a88} \sim c_{a98}$
Head	Turn up-down or not; Turn left-right or not; Tilt or not; Amplitude of turn up-down; Amplitude of turn left-right; Amplitude of tilt	$c_{a32} \sim c_{a37} / c_{a99} \sim c_{a104}$

表 2 各部位 4 个动作过程的帧数信息编码规则

Table 2 Coding rules of frame number information of four movement processes in various parts

Facial part	P1-s	P1-m	P2-s	P2-m	Return or not
Brow (Left)	c_{a38}/c_{a105}	c_{a39}/c_{a106}	c_{a40}/c_{a107}	c_{a41}/c_{a108}	c_{a42}/c_{a109}
Brow (Right)	c_{a43}/c_{a110}	c_{a44}/c_{a111}	c_{a45}/c_{a112}	c_{a46}/c_{a113}	c_{a47}/c_{a114}
Eye (Left)	c_{a48}/c_{a115}	c_{a49}/c_{a116}	c_{a50}/c_{a117}	c_{a51}/c_{a118}	c_{a52}/c_{a119}
Eye (Right)	c_{a53}/c_{a120}	c_{a54}/c_{a121}	c_{a55}/c_{a122}	c_{a56}/c_{a123}	c_{a57}/c_{a124}
Mouth	c_{a58}/c_{a125}	c_{a59}/c_{a126}	c_{a60}/c_{a127}	c_{a61}/c_{a128}	c_{a62}/c_{a129}
Head	c_{a63}/c_{a130}	c_{a64}/c_{a131}	c_{a65}/c_{a132}	c_{a66}/c_{a133}	c_{a67}/c_{a134}

代第 i 位的基因值。

$$c'_i = \min(c_{ai}, c_{bi}) + u_i |c_{ai} - c_{bi}| \quad (2)$$

对于需要变异的基因位, 均匀变异直接在该基因位的取值范围内随机生成一个新值 c'_i 取代原有基因值, 如式 (3) 所示。 lo_i 和 up_i 分别是该基因位的上、下限值, r_i 为该基因位的变异点。

$$c'_i = lo_i + (up_i - lo_i) \times r_i, \quad r_i \sim U[0, 1] \quad (3)$$

为了解决 IGA 的小种群规模带来的种群多样性过早丧失的问题, 本文基于均匀变异提出了另外两种变异策略, 分别是带屏蔽区间的均匀变异 (Uniform mutation with exclusion zones) 以及共享变异点的均匀变异 (Uniform mutation with shared mutation points)。前者改进变异机制, 后者结合编码的实际意义改进基因变异点 r_i 的给定方式。本文从基因编码层面去量化个体间的相似性从而去衡量种群多样性, 种群内个体间平均相似性越高, 表明多样性越低。因此, 本文将比较不同变异算子在降低种群内平均相似度 (Intra-population similarity)^[15] 方面的效果。

2.3.1 种群内平均相似度 由于文中种群多样性个体染色体编码涵盖不同的基因取值范围, 部分基因位点采用 0-1 编码的方式表示不同的属性, 如是否复位和是否转头等, 故而在进行相似度计算时需要将属性相似与空间数据相似进行融合; 并且由于编码长度较长, 维度灾难问题也需要得到妥善处理, 所以本文个体间相似度基于谢明霞等^[48] 的研究, 按式 (4) 进行计算。其中 $sim_{a,b}$ 为个体 $I_a = \{c_{a1}, c_{a2}, \dots, c_{a134}\}$ 与个体 $I_b = \{c_{b1}, c_{b2}, \dots, c_{b134}\}$ 之间的相似度; $l = 134$, 为个体基因编码的长度; w_i 为第 i 维的权重, $w_i > 1$, w_i 取值越小, 第 i 维对计算结果的影响越大。在本文中, 控制静态目标表情的基因位对应的 w_i 值相对较小。 $\delta(c_{ai}, c_{bi})$ 算子的引入减小了布尔位带来的偏差, 当第 i 维为布尔位, 且 $c_{ai} \neq c_{bi}$ 时, $\delta(c_{ai}, c_{bi}) = 0$; 其他时候, $\delta(c_{ai}, c_{bi}) = 1$ 。

$$\begin{cases} sim(I_a, I_b) = \frac{1}{1 + e^{-10(sim_{a,b} - 0.5)}} \\ sim_{a,b} = \sum_{i=1}^l \frac{\delta(c_{ai}, c_{bi})}{l} \cdot \frac{1}{1 + w_i |c_{ai} - c_{bi}|} \\ \delta(c_{ai}, c_{bi}) = \begin{cases} 0, & \text{if } c_{ai} \neq c_{bi} \text{ and } c_{ai}, c_{bi} \in \{0, 1\} \\ 1, & \text{others} \end{cases} \end{cases} \quad (4)$$

另外, 由于个体编码的数据维度过高, 个体之间的初始相似度 $sim_{a,b}$ 大多集中在 0.5 附近。本文为了放大不同相似度之间的差异, 再对 $sim_{a,b}$ 进行 sigmoid 放缩, 最终得到 $sim(I_a, I_b)$ 用于衡量个体 I_a 与 I_b 之间实际的相似度。所以, 若以 $I = \{I_1, I_2, \dots, I_N\}$ 表示当前种群集合, 则种群内平均相似度 $simI$ 可以由式 (5) 计算得出。

$$simI = \frac{2}{N(N-1)} \sum_{a=1}^{N-1} \sum_{b=a+1}^N sim(I_a, I_b) \quad (5)$$

2.3.2 带屏蔽区间的均匀变异 屏蔽区间的核心思想是通过避开原基因值附近的值以提高不同表情动画的染色体编码的差异性。屏蔽空间如图 7 所示, 在原基因值 c_i 的左右两边各有一段长度为 d 的屏蔽区间, 变异后得到的新值随机落在 $[lo_i, d_-] \cup [d_+, up_i]$ 中, 也就是在屏蔽区间之外。屏蔽区间长度 d 主要由 $simI$ 和当前代种群多样性决定。

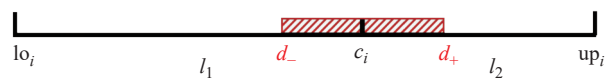


图 7 屏蔽区间

Fig. 7 Exclusion zones

种群多样性的计算主要与 $simI$ 和种群中个体分布相关, 式 (6)^[49] 为种群中个体分布的均匀度 $\gamma(t)$ 的计算方法, 其中 f_{avg}^t , f_{max}^t 和 f_{min}^t 分别为第 t 代种群适应度的平均值、最大值和最小值。 $\sigma(t)$ 是第 t 代种群的适应度方差, $\max\{\sigma(t)\}$ 和 $\min\{\sigma(t)\}$ 分别表示进化到第 t 代时适应度方差的最大值和最小值。 $\gamma(t)$ 的值越接近 1, 表示种群分布越均匀。因此, 种

群多样性指标定义如式 (7) 所示, 当 $\text{div}(t)$ 越接近 1 时, 种群多样性越高。

$$\gamma(t) = \frac{\sigma(t) - \min\{\sigma(t)\}}{\max\{\sigma(t)\}} \cdot \cos\left(\frac{\pi}{2} \cdot \left(1 - \frac{f_{\text{avg}}^t - f_{\text{min}}^t}{f_{\text{max}}^t - f_{\text{min}}^t}\right)\right) \quad (6)$$

$$\text{div}(t) = \gamma(t) \cdot [1 - \text{sim}I(t)] \quad (7)$$

我们希望当种群内相似度越高或种群多样性越低时, 屏蔽区间的长度更长。但为了避免局部最优解逃逸现象的发生, 仅当种群多样性低于一定范围时, 屏蔽区间才被引入。所以对于第 t 代的每个个体, 对每一位需要变异的基因, 屏蔽区间长度 d 根据式 (8) 计算。 $f[\text{div}(t)]$ 可以取 3 种形式, 分别是 \cos 、 \sin 和 linear , 用于控制 d 随 $\text{div}(t)$ 的变化趋势, 如式 (9) 所示。

$$\begin{cases} d = d_{\max} \cdot f[\text{div}(t)] \\ d_{\max} = [l_2 + (l_1 - l_2) \cdot \text{sim}I(t)^s] \\ l_1 = \max(\text{up}_i - c_i, c_i - \text{lo}_i) \\ l_2 = \min(\text{up}_i - c_i, c_i - \text{lo}_i) \end{cases} \quad (8)$$

$$f[\text{div}(t)] = \begin{cases} \cos\left(\frac{\pi}{2k} \cdot \text{div}(t)\right), f: \cos \\ 1 - \sin\left(\frac{\pi}{2k} \cdot \text{div}(t)\right), f: \sin \\ 1 - \frac{1}{k} \cdot \text{div}(t), f: \text{linear} \end{cases} \quad (9)$$

其中, s 用于控制屏蔽区间的最大长度 d_{\max} 随当前代的 $\text{sim}I$ 的变化趋势。 k 用于控制 $\text{div}(t)$ 值的阈值, 当 $\text{div}(t)$ 超过 k 时, 不再存在排除区间。

图 8(a) 示出了不同 s 值下 d_{\max} 随 $\text{sim}I$ 的变化情况, 从图中可以看出, $\text{sim}I$ 一定时, s 越小, d_{\max} 的值越大。图 8(b) 和图 8(c) 分别示出了不同的 k 值和 f 函数对 d 的影响。图 9 示出了 $l_1=0.8$ 、 $l_2=0.2$ 、 $s=2.0$ 、 $k=0.4$ 、 f 取 \cos 时, d 随 $\text{sim}I$ 和 $\text{div}(t)$ 的变化, 可以看出, 当 $\text{sim}I$ 越大, $\text{div}(t)$ 越小时, d 越大。

2.3.3 共享变异点的均匀变异 由 2.2 节可知, 一条表情动画的染色体主要可以分为静态目标表情的控制位 (C1 和 C3 基因段) 和运动过程的帧数控制位

(C2 和 C4 基因段) 两个部分, 其中对最终呈现效果影响更大的是静态目标表情的控制位。因此, 共享变异点的均匀变异主要是为静态目标表情的基因控制位 (C1 和 C3 基因段) 设计的, 也就是说对于 C2 和 C4 基因段还是采取原始的变异方式。此方法核心思想在于相同的面部部位倾向于有相同的变化, 对于控制同一个部位的基因位设置相同的变异点, 即式 (3) 中参数 r_i 的值被设置为相同的随机值, 但是变异方式还是采取均匀变异。面部部位根据表 3 划分, 以控制目标表情 Expression 1 左眼的 3 个基因位 $c_{a14} \sim c_{a16}$ 为例, 这 3 位基因的变异点 $r_{14} \sim r_{16}$ 共享相同的值 r_{14}^{pl} 。相较于表 1 中的面部区域划分, 这里再将嘴部动作的控制基因段细分为了左嘴角、右嘴角和嘴部中点 3 段。

2.4 用户评价和适应度更新

本文的用户评价环节允许多位用户同时根据自己的喜好选择自己喜欢的表情动画。为减轻用户的任务量, 我们不需要用户对每个表情个体评分, 取而代之的是“收藏 (Collect)”和“保存 (Save)”两个操作, 如图 10 所示 (图中进化代数 23)。用户对于满意的表情可以点击“收藏”, 而对于想留下的表情可以进一步“保存”到自己的表情库中。这样简单的选择方式同时也减小了因为主观因素引起的评分偏差。个体 I_a 的适应度值 $F(I_a)$ 根据其收获的收藏数 $c(U, I_a)$ 和保存数 $v(U, I_a)$ 按式 (10) 进行更新, 其中 U 为用户集合, $U = \{U_1, U_2, \dots, U_M\}$, M 是参与评价的用户的总数。

$$F(I_a) = c(U, I_a) + 2v(U, I_a) \quad (10)$$

3 实验部分

为了克服无法获取足够多的实际用户数据用于多次对比实验的问题, 本文引入了一个简易的多用

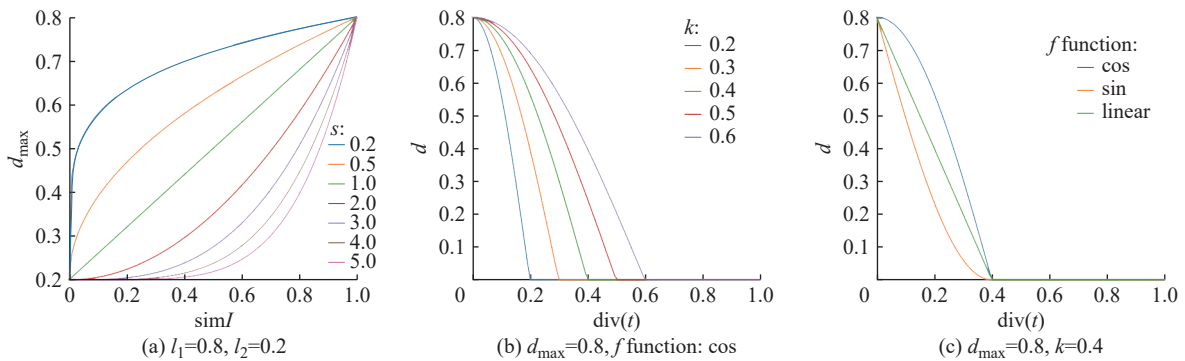


图 8 不同的 k 、 s 、 f 对屏蔽区间长度的影响 ($l_1=0.8$, $l_2=0.2$)

Fig. 8 Impact of different k , s values and f function forms on the length of exclusion zones ($l_1=0.8$, $l_2=0.2$)

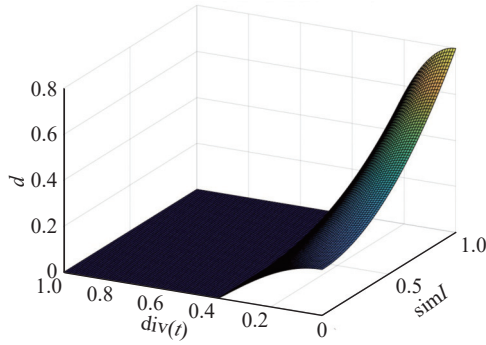


图9 屏蔽区间长度随 simI 和 div(t) 的变化趋势

Fig. 9 Changing trends of the length of exclusion zones with simI and div(t)

户行为代理模型用以辅助实验,在参数确定以及算子有效性验证时使用用户伪代理模型进行实验,验证本文所提方法有效性时则邀请了实际用户参与实验。

3.1 多用户伪代理模型

用户伪代理模型用来模拟用户在选择偏好表情

$$\begin{cases} c(U, I_a) = \sum_{j=1}^M c(U_j, I_a), c(U_j, I_a) = \begin{cases} 1, & \text{if } \text{sim}(U_j, I_a) > \text{thc} \\ 0, & \text{if } \text{sim}(U_j, I_a) \leq \text{thc} \end{cases} \\ v(U, I_a) = \sum_{j=1}^M v(U_j, I_a), v(U_j, I_a) = \begin{cases} 1, & \text{if } \text{sim}(U_j, I_a) > \text{thv} \\ 0, & \text{if } \text{sim}(U_j, I_a) \leq \text{thv} \end{cases} \end{cases} \quad (11)$$

鉴于提到的第(2)点,每一代收藏的阈值 thc 和保存的阈值 thv 根据当代的 simI 进行调整,如式(12)所示。保存是在收藏的基础上进行,所以 thv 要适当高于 thc。在同一代种群中,用户往往不会选择两个高度相似的个体,保存两个相似个体的可能性更小。所以,在本文中,当 simI 超过一定值时, thc 和 thv 都随 simI 的增大而增大,且 thv 增大的趋势被设置得更快。图 11 所示为实际用户引导进化时, simI 对种群所获的收藏数和保存数的影响,从图中可以看出,当 simI 达到 0.50 时,收藏数和保存数的上升趋势明显停滞,保存数甚至开始下降。同时考虑用户的直观感知,当 simI 超过 0.60 时,就能明显感觉到种群中同代表情动画个体的表现越来越相似,此时种群多样性已丧失,当 simI 维持在 0.50 以下时,表情动画的多样性较高。故 thc 和 thv 的起始值分别设置为 0.45 和 0.55,当 simI 超过 0.45 之后, thc 和 thv 开始随 simI 增大而增大。

时的行为,通过收集多位用户在选择表情动画时的心理,发现了两个最明显的规律:(1) 每位用户都有自己独特的偏好,有些用户更倾向保存夸张搞怪的表情动画,他也愿意收藏普通但真实的表情;然而有些用户只看重表情的真实性,他们只收藏或保存能明显看出心理原型的表情,如高兴、悲伤等。(2) 当进入到进化后期,种群中表情的相似度高,对于同一种群中类似的表情动画,用户可能只会保存其中一个,而不会每一个都保存。对于第(1)点,本文记录了由多个实际用户选出的 100 个偏好表情动画的编码,作为 100 个用户的伪代理模型,用来表示不同用户的不同偏好,则表情个体 I_a 收获到的来自用户伪代理模型的收藏数和保存数可根据式(11)计算,其中 $U = \{U_1, U_2, \dots, U_j, \dots, U_M\}, M = 100$, 为用户伪代理模型的集合, $\text{sim}(U_j, I_a)$ 为个体 I_a 与伪用户模型 U_j 的相似度, thc 和 thv 分别是达到收藏和保存条件的相似度阈值。

$$\begin{cases} \text{thc} = \begin{cases} 0.45 + (\text{simI} - 0.45)^3, & \text{simI} > 0.45 \\ 0.45, & \text{simI} \leq 0.45 \end{cases} \\ \text{thv} = \begin{cases} 0.55 + (\text{simI} - 0.45)^2, & \text{simI} > 0.45 \\ 0.55, & \text{simI} \leq 0.45 \end{cases} \end{cases} \quad (12)$$

3.2 基于用户伪代理模型的实验

实验所用交叉策略为 2.3 节中的平坦交叉,选择采用轮盘赌^[50]的方式,通过对比不同变异策略和参数下的种群平均适应度值 ($\bar{F}(I)$)、最大适应度值 ($F_{\max}(I)$) 和 simI, 确定式(8)和式(9)中的最佳参数以及最佳变异策略。实验中种群大小 $N = 10$, 最大迭代次数 $G = 50$, 交叉率 (pc) 和变异率 (pm) 分别被设置为 1 和 0.2。为了减少单次实验带来的偶然性,每种变异策略下的评价指标都是 100 次重复实验之后的均值。

3.2.1 带屏蔽区间的均匀变异 第 1 组实验选择式(9)中的 f 为 cos 函数、并将 k 值设为 0.4, 对比不同 s 取值下 $\bar{F}(I)$ 、 $F_{\max}(I)$ 以及 simI 随进化过程的变化,结果如图 12 所示。从图 12 中可观察到,当

表3 带屏蔽区间的均匀变异中不同面部部位的变异点

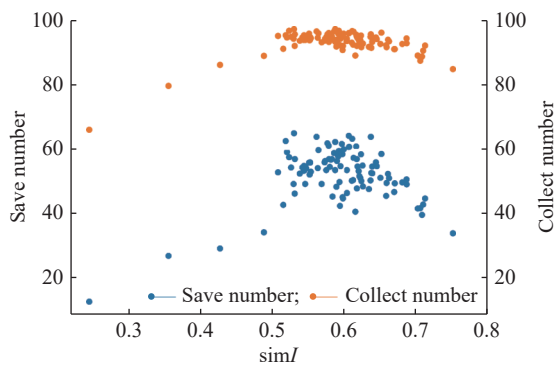
Table 3 Mutation points for different facial parts in uniform mutation with exclusion zones

Expression	Brow (Left)	Brow (Right)	Eye (Left)	Eye (Right)	Mouth_corner (Left)	Mouth_corner (Right)	Mouth_mid	Head
Expression 1	$r_1 \sim r_6: r_1^{\text{BL}}$	$r_7 \sim r_{12}: r_1^{\text{BR}}$	$r_{14} \sim r_{16}: r_1^{\text{EL}}$	$r_{17} \sim r_{19}: r_1^{\text{ER}}$	$r_{21} \sim r_{22}: r_1^{\text{ML}}$	$r_{23} \sim r_{24}: r_1^{\text{MR}}$	$r_{25} \sim r_{30}: r_1^{\text{MM}}$	$r_{35} \sim r_{37}: r_1^{\text{H}}$
Expression 2	$r_{68} \sim r_{73}: r_2^{\text{BL}}$	$r_{74} \sim r_{79}: r_2^{\text{BR}}$	$r_{81} \sim r_{83}: r_2^{\text{EL}}$	$r_{84} \sim r_{86}: r_2^{\text{ER}}$	$r_{88} \sim r_{89}: r_2^{\text{ML}}$	$r_{90} \sim r_{91}: r_2^{\text{MR}}$	$r_{92} \sim r_{97}: r_2^{\text{MM}}$	$r_{102} \sim r_{104}: r_2^{\text{H}}$



图 10 用户评价界面

Fig. 10 User evaluation interface

图 11 收藏数和保存数随 $simI$ 的变化趋势Fig. 11 Changing trends of save number and collect number with $simI$

$s \leq 1.0$ 时, $\bar{F}(I)$ 的值明显比 s 取其他值时要低; 当 $s \geq 2.0$ 时, 不同 s 值下 $\bar{F}(I)$ 之间的差异不明显。同时, 不同 s 值下的最大适应度 $F_{\max}(I)$ 几乎相等, 这表明不同的 s 取值对每一代种群中最优个体的搜索并没有显著影响。值得注意的是, s 的值越小, 在进化过程中 $simI$ 也越低。在保持种群质量, 即种群的平均适应度值 $\bar{F}(I)$ 没有明显下降的前提下, 当 $s = 2.0$ 时 $simI$ 最低。

第 2 组实验同样选择 \cos 函数, 并固定 s 取值为 2.0, 对比不同 k 值下带屏蔽区间的变异算子的表现, 如图 13 所示。从图中可以明显看出 $\bar{F}(I)$ 和 $simI$ 都随 k 值的增大而减小, 然而当 $k > 0.4$ 时, $F_{\max}(I)$ 也开始有明显的下降。当 $k = 0.4$ 时, $\bar{F}(I)$ 相比 $k \leq 0.3$ 时略微有下降, 但此时可以保证最大适应度不下降, 并同时保持 $simI$ 保持在 0.50 以下。综合考虑不同的 s 和 k 取值对适应度值和种群相似度的影响, 最终选定 $s = 2.0$ 、 $k = 0.4$ 。

最后, 本文还对比了不同的 f 函数对带屏蔽区间的均匀变异算子性能的影响, 如图 14 所示, 从图中可以看出, 不同的函数形式对该变异算子的表现并

没有很大的影响, 但当 f 采用 \cos 形式时, $simI$ 在进化过程中最低。所以最终确定屏蔽区间长度 d 的计算公式如式(13)所示:

$$\begin{cases} d = [l_2 + (l_1 - l_2) \cdot \cos(\frac{\pi}{2 \times 0.4} \cdot \text{div}(t))] \cdot \cos(\frac{\pi}{2 \times 0.4} \cdot \text{div}(t)) \\ l_1 = \max(\text{up}_i - c_i, c_i - \text{lo}_i) \\ l_2 = \min(\text{up}_i - c_i, c_i - \text{lo}_i) \end{cases} \quad (13)$$

3.2.2 不同变异算子的对比实验 将原始的均匀变异作为对照组, 对比本文提出的两种变异算子, 本文考察了 $\bar{F}(I)$ 、 $F_{\max}(I)$ 、 $simI$ 的变化趋势, 如图 15 所示。从图中可以看出, 在表情进化过程中, 采用均匀变异算子时, 大约在第 10 代时 $simI$ 值超过了 0.60, 而本文所提出的两种算子可以在不影响种群适应度的前提下, 很好地将 $simI$ 保持在 0.60 甚至是 0.50 以下, 从而在提高了种群多样性的同时保持了整个种群的质量, 尤其是带屏蔽区间的均匀变异能够将 $simI$ 保持在 0.50 以下。

图 16 和图 17 分别示出了在指定进化代数下、3 种变异策略各自实验 100 次, $simI$ 超过 0.50 或 0.60 的实验次数。结果表明, 在 100 次实验中, 采用带屏蔽区间的变异策略时, $simI$ 超过 0.50 和 0.60 的次数都是最少的, 尤其是 $simI$ 超过 0.60 的次数基本都在 5 次以下。和原始的均匀变异相比, 采用共享变异点的均匀变异时, $simI$ 超过 0.50 的次数略少, 但超过 0.60 的次数相比均匀变异有明显下降。表 4 示出了在指定进化代数下整个种群的总保存数, 即每个个体被保存的次数的总和, 总保存数越高, 说明当前种群越受用户所喜爱。从表中的数据可以看出, 在整个进化过程中, 采用均匀变异策略时种群收到的总保存次数总是最少的。

3.3 真实用户评估

在实际用户的评估阶段, 本文邀请了 20 位用户

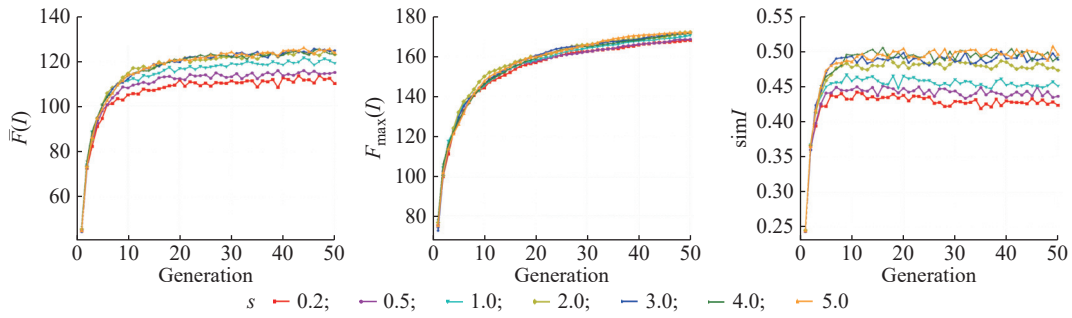


图 12 带屏蔽区间的均匀变异中, 不同 s 值下 $\bar{F}(I)$ 、 $F_{\max}(I)$ 和 $simI$ 的变化趋势

Fig. 12 Trends of $\bar{F}(I)$, $F_{\max}(I)$ and $simI$ under different s values of uniform mutation with exclusion zones

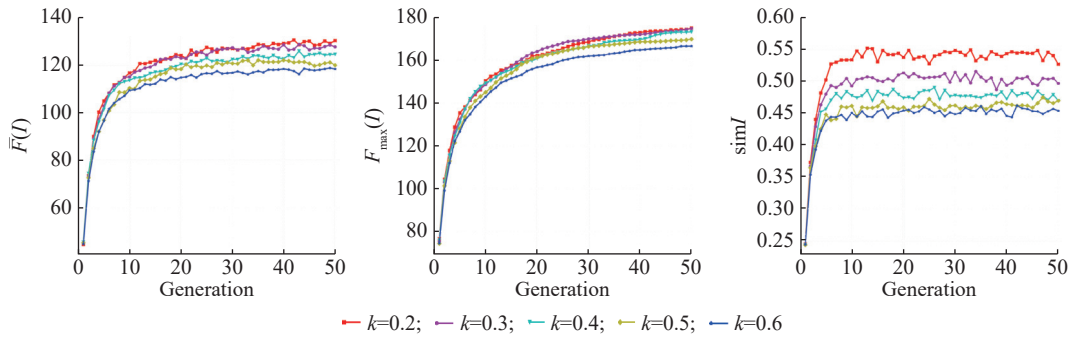


图 13 带屏蔽区间的均匀变异中, 不同 k 值下 $\bar{F}(I)$ 、 $F_{\max}(I)$ 和 $simI$ 的变化趋势

Fig. 13 Trends of $\bar{F}(I)$, $F_{\max}(I)$ and $simI$ under different k values of uniform mutation with exclusion zones

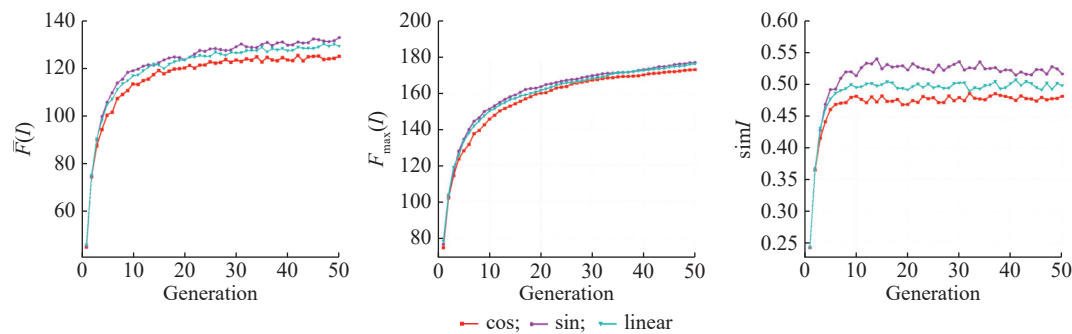


图 14 带屏蔽区间的均匀变异中, 不同 f 函数下 $\bar{F}(I)$ 、 $F_{\max}(I)$ 和 $simI$ 的变化趋势

Fig. 14 Trends of $\bar{F}(I)$, $F_{\max}(I)$ and $simI$ under different f functions of uniform mutation with exclusion zones

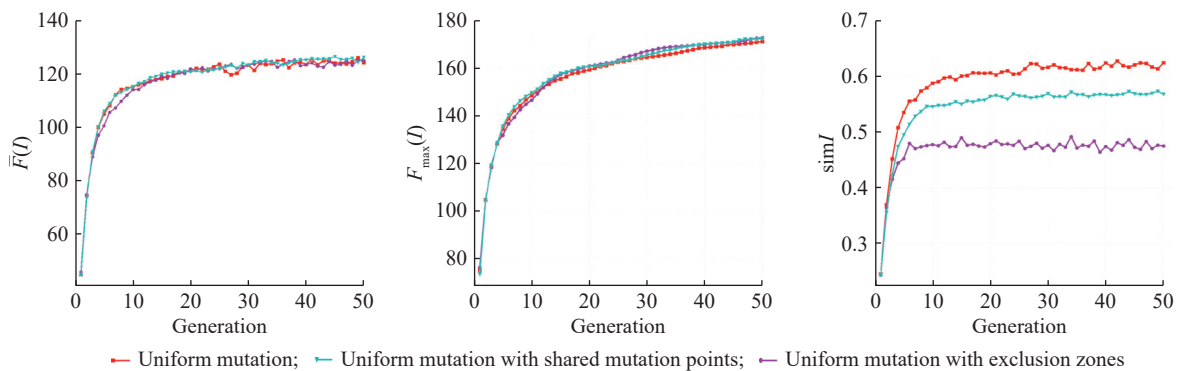


图 15 3 种变异策略下, $\bar{F}(I)$ 、 $F_{\max}(I)$ 和 $simI$ 的变化趋势

Fig. 15 Trends of $\bar{F}(I)$, $F_{\max}(I)$ and $simI$ under three mutation operators

跟进 20 代的表情动画进化过程, 每位用户同时对 3 种变异策略下所得的种群进行选择, 并根据自己的偏好去收藏或保存表情个体。为确保结果的全面性

和客观性, 选取的 20 位用户性别比例均衡(男女各半), 并涵盖了多样化的专业背景。其中, 4 名用户为本实验室成员, 具备与本研究相关的专业背景, 另外

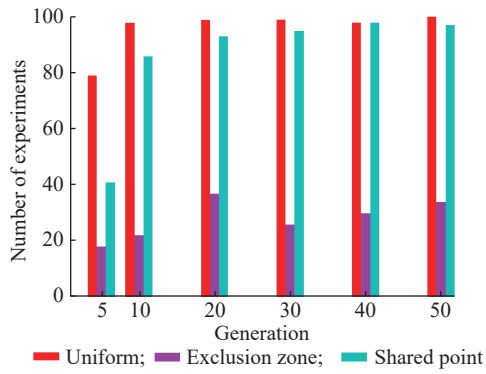


图 16 $simI > 0.50$ 的实验次数

Fig. 16 Number of experiments of $simI > 0.50$

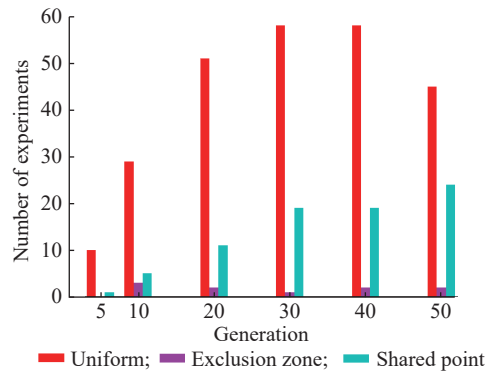


图 17 $simI > 0.60$ 的实验次数

Fig. 17 Number of experiments of $simI > 0.60$

6 名用户来自校内其他专业, 余下的 10 名用户为社会人士, 他们年龄分布在 20~50 岁, 学历从高中至硕士, 职业包括医生、销售代表、软硬件工程师等, 代表了广泛的社会群体。

实验过程中, 为了降低用户多次重复评价的不稳定性, 首先开始实验前, 每位用户都有机会熟悉评分规则和系统; 正式开始实验时, 3 种变异算子下的进化同步进行, 即用户每次对 3 种算子下的种群同时进行选择, 但是用户并不知晓当前评价的种群背后使用的变异算子类型。实验过程中, 用户“收藏/保存”的操作从评价方式上减轻了用户疲劳, 一定程度上降低了用户疲劳带来的评价偏差。图 18 示出了实际用户参与时、3 种变异策略下, $\bar{F}(I)$ 、 $F_{max}(I)$ 和 $simI$ 的变化趋势。从图中可以看出, 我们提出的变异策略在实际用户参与评估时也起到了良好的作用, 3 种变异方法下种群的 $\bar{F}(I)$ 都稳步提升, 但本文提出的两种变异策略在降低 $simI$ 方面有明显效果, 同时还一定程度上提升了种群中最优个体的适应度值。

图 19 示出了实际用户选择下第 20 代的表情动画个体。由于我们的初始种群是随机生成的, 并且不同用户的选择偏好也不尽相同, 所以每次进化过程最终都会产生出各种各样, 甚至意想不到的表

表 4 100 次实验中总保存数均值

Table 4 Average total save number of 100 experiments

Generation	Average total save number		
	Uniform mutation	Exclusion zone	Shared point
5	158.02	160.66	169.11
10	188.50	196.31	201.11
20	209.94	217.71	210.49
30	211.12	231.72	217.70
40	209.31	237.13	229.24
50	233.65	239.24	232.43

情。这也说明了在不预先给定心理原型的条件下, 通过本文的方法自动进化出复杂多样的表情动画供给用户挑选是切实可行的。

4 结束语

针对目前大部分的表情生成方法所面临的多样性不足和缺少实际用户评价的问题, 本文采用 IGA 引入多用户评价, 并结合 DFFD 算法提出了一套全新的三维表情动画编码。区别于一般的表情生成

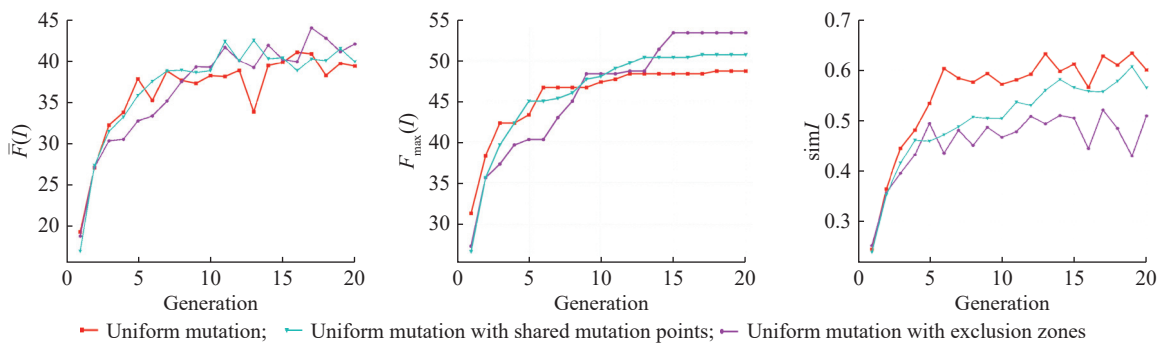


图 18 实际用户参与时, 3 种变异策略下, $\bar{F}(I)$ 、 $F_{max}(I)$ 和 $simI$ 的变化趋势

Fig. 18 Trends of $\bar{F}(I)$, $F_{max}(I)$ and $simI$ under three mutation operators, evaluated by actual users



图 19 实际用户选择下第 20 代的表情个体示例

Fig. 19 Several expression animations of the 20th generation under the actual user's choices

方法,本文方法不需要预设目标表情的心理原型,这使得进化得到更加复杂多样的表情动画成为现实。将用户评价嵌入到整个表情生成过程的方式,降低了表情生成的门槛,使得普通用户也可以轻松生成自己喜欢的表情动画,很好地解决了生成表情时缺少用户主观评价指标的问题。另外,考虑到采用 IGA 进化表情动画时所面临的种群多样性过早丧失的问题,本文提出了带屏蔽区间的均匀变异和共享变异点的均匀变异算子,实验结果表明,这两种变异算子在降低种群内表情个体之间的相似度方面具有明显的优势。两种算子的思想也可以用于其他使用 IGA 的场景,防止种群在进化过程中过早收敛。

本文后续将设计更加有趣的用户评价方式并进一步完善用户伪代理模型,旨在减轻用户评价负担的同时进一步提高表情动画的多样性。此外,目前的表情动画仅在标准人脸模型上呈现,后续将尝试迁移到各种不同的人脸模型上,进一步提高用户在表情生成过程的可操作空间。

参考文献:

- [1] MIN R, KOSE N, DUGELAY J L. KinectFaceDB: A kinect database for face recognition[J]. *IEEE Transactions on Systems Man & Cybernetics Systems*, 2017, 44(11): 1534-1548.
- [2] CHEN H, WANG L, LIU W, *et al.* Combined X-ray and facial videos for phoneme-level articulator dynamics[J]. *Visual Computer*, 2010, 26(6/8): 477-486.
- [3] TAKAGI H. Interactive evolutionary computation: Fusion of the capabilities of EC optimization and human evaluation[J]. *Proceedings of the IEEE*, 2001, 89(9): 1275-1295.
- [4] 卞芳明. 巧做人面面部表情动画——浅谈 FaceGen Modeller 与 Maya 的结合制作 [J]. *电视字幕 (特技与动画)*, 2007(9): 39-41.
- [5] MOCCOZET L, THALMANN N M. Dirichlet free-form deformations and their application to hand simulation[C]//Computer Animation and Simulation 97: Proceedings of the Eurographics Workshop. Geneva, Switzerland: IEEE, 1997: 93-102.
- [6] EIBEN A E, SMITH J E. Introduction to evolutionary computing[J]. *Assembly Automation*, 2004, 24(3): 324.
- [7] AMINI R, LISETTI C, RUIZ G. HapFACS 3.0: FACS-based facial expression generator for 3D speaking virtual characters[J]. *IEEE Transactions on Affective Computing*, 2015, 6(4): 348-360.
- [8] MOLANO J S V, DÍAZ G M, SARMIENTO W J. Parametric facial animation for affective interaction workflow for avatar retargeting[J]. *Electronic Notes in Theoretical Computer Science*, 2019, 343: 73-88.
- [9] CAO C, BRADLEY D, ZHOU K, *et al.* Real-time high-fidelity facial performance capture[J]. *ACM Transactions on Graphics (TOG)*, 2015, 34(6): 225.
- [10] THIES J, ZOLLHOFER M, STAMMINGER M, *et al.* Face2Face: Real-time face capture and reenactment of RGB videos[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA: IEEE, 2016: 2387-2395.
- [11] HYUNG H J, YOON H U, CHOI D, *et al.* Optimizing android facial expressions using genetic algorithms[J]. *Applied Sciences*, 2019, 9(16): 3379.
- [12] NI H, LIU J. 3D Face dynamic expression synthesis system based on DFFD[C]//2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC). Chengdu, China: IEEE, 2019: 1125-1128.
- [13] CHARLINE G, LAURENCE C, HUN STÉPHANIE, *et al.* Children facial expression production: Influence of age, gender, emotion subtype, elicitation condition and culture[J]. *Frontiers in Psychology*, 2018, 9: 446.
- [14] CARLISI C O, REED K, HELMINK F G L, *et al.* Using genetic algorithms to uncover individual differences in how humans represent facial emotion[J]. *The Royal Society*, 2021(10): 202251.
- [15] YAN S, SOLADIÉ C, SEGUIER R. Exploring mental prototypes by an efficient interdisciplinary approach: Interactive microbial genetic algorithm[C]//2023 IEEE 17th International Conference on Automatic Face and Gesture Recognition (FG). Waikoloa Beach, HI, USA: IEEE, 2023: 1-8.
- [16] ZHANG Y, JI Q, ZHU Z, *et al.* Dynamic facial expression analysis and synthesis with MPEG-4 facial animation parameters[J]. *IEEE Transactions on Circuits & Systems for*

- Video Technology, 2008, 18(10): 1383-1396.
- [17] 王俊培. 真实感三维人脸表情动画生成 [D]. 北京: 北京工业大学, 2011.
- [18] LIU Z, SHAN Y, ZHANG Z. Expressive expression mapping with ratio images[C]//Computer Graphics and Interactive Techniques. USA: ACM, 2001: 271-276.
- [19] ZHANG Q, LIU Z, GUO B, *et al.* Geometry-driven photorealistic facial expression synthesis[J]. IEEE Transactions on Visualization and Computer Graphics, 2006; 12(1): 48-60.
- [20] EKMAN P, FRIESEN W V. Facial Action coding system (FACS): A technique for the measurement of facial actions[J]. Rivista Di Psichiatria, 1978, 47(2): 126-138.
- [21] 崔韶. 人脸面部特征识别与虚拟角色表情生成算法研究 [D]. 河北: 河北科技大学, 2023.
- [22] CUCULO V, D'AMELIO A. OpenFACS: An open source FACS-Based 3D face animation system[C]// International Conference on Image and Graphics. Beijing, China: Springer International Publishing, 2019: 232-242.
- [23] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, *et al.* Generative adversarial nets[C]//Advances in Neural Information Processing Systems. USA: MIT Press, 2014: 2672-2680.
- [24] CHOI Y, CHOI M, KIM M, *et al.* Stargan: Unified generative adversarial networks for multi-domain image-to-image translation[C]//IEEE Conference on Computer Vision and Pattern Recognition. Saltlake City, UT, USA: IEEE, 2018: 8789-8797.
- [25] KARRAS T, LAINE S, AILA T. A style-based generator architecture for generative adversarial networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 43(12): 4217-4228.
- [26] EKMAN P. Constants across cultures in the face and emotion[J]. Journal of Personality and Social Psychology, 1971, 17(2): 124-129.
- [27] PUMAROLA A, AGUDO A, MARTINEZ A M, *et al.* Ganimation: Anatomically-aware facial animation from a single image[C]//Proceedings of the European Conference on Computer Vision (ECCV). Springer International Publishing, 2018, 818-833.
- [28] 林本旺. 基于生成对抗网络的人脸表情生成方法研究 [D]. 北京: 北京建筑大学, 2023.
- [29] WANG X, WANG Y, LI W, *et al.* Facial expression animation by landmark guided residual module[J]. IEEE Transactions on Affective Computing, 2023, 14(2): 878-894.
- [30] POTAMIAS R A, ZHENG J, PLOUMPIS S, *et al.* Learning to generate customized dynamic 3D facial expressions[C]//Computer Vision-ECCV 2020: 16th European Conference. Glasgow, UK: Springer International Publishing, 2020: 278-294.
- [31] ZHU X, LI X, CHEN Y, *et al.* Interactive genetic algorithm based on typical style for clothing customization[J]. Journal of Engineered Fibers and Fabrics, 2020, 15(10): 1-9.
- [32] NUTTHANON L, SUNISA R. Generating Kranok patterns with an interactive evolutionary algorithm[J]. Applied Soft Computing, 2020, 89: 106121.
- [33] 谢元媛, 王磊. 基于改进型交互式遗传算法的医疗器械优化设计 [J]. 微型电脑应用, 2022, 38(11): 129-133.
- [34] 张幂, 陈庆, 刘肖健. 基于交互式遗传算法的布艺沙发 CMF 情感化设计研究 [J]. 包装工程, 2023, 44(16): 79-88.
- [35] GUO Q. Computer-assisted music composition algorithm design dependent on interactive genetic algorithm with interval fitness[J]. Journal of Physics: Conference Series, 2021, 2066(1): 1-7.
- [36] FUKUMOTO M, NOMURA K. Music melodies suited to multiple users' feelings composed by asynchronous distributed interactive genetic algorithm[J]. International Journal of Software Innovation, 2018, 6(2): 26-36.
- [37] 苏胜, 顾森, 宋志强, 等. 基于深度表征学习和遗传算法的军用座舱色彩设计方法 [J]. 兵工学报, 2024, 45(4): 1060-1069.
- [38] YUNIARTI A, ANGGARA S, AMALIAH B. Resize My Image: A mobile app for interactive image resizing using multi operator and interactive genetic algorithm[C]//2016 International Conference on Information & Communication Technology and Systems(ICTS). Surabaya, Indonesia: IEEE, 2016: 58-62.
- [39] LAI C C, CHEN Y C. A user-oriented image retrieval system based on interactive genetic algorithm[J]. IEEE Transactions on Instrumentation and Measurement, 2011, 60(10): 3318-3325.
- [40] MIGUEL A, FERRI F J, MORENO-PICOT S. Distance-based relevance feedback using a hybrid interactive genetic algorithm for image retrieval[J]. Applied Soft Computing, 2011, 11(2): 1782-1791.
- [41] HAILEMARIAM M, GOERTZEL B, YOHANNES T. Evolving 3D facial expressions using interactive genetic algorithms[C]//International Conference on Advances of Science and Technology. Cham: Springer International Publishing, 2018: 492-502.
- [42] REED K, COSKER D. User-guided facial animation through an evolutionary interface[J]. Computer Graphics Forum, 2019, 38(6): 165-176.
- [43] SEDERBERG T W, PARRY S R. Free-form deformation of solid geometric models[C]//Proceedings of the 13th Annual Conference on Computer Graphics and Interactive techniques. USA: ACM, 1986: 151-160.
- [44] SIBSON R. A brief description of natural neighbour interpolation[J]. Interpreting Multivariate Data, 1981, 27(1): 21-36.
- [45] 倪虎. 基于 Dirichlet 自由变形算法的人脸表情动画技术研究 [D]. 武汉: 武汉理工大学, 2019.
- [46] RADCLIFFE N J. Equivalence class analysis of genetic algorithms[J]. Complex Systems, 1991, 5(2): 183-205.
- [47] MAITI A K, MAITI M. Discounted multi-item inventory

- model via genetic algorithm with Roulette wheel selection, arithmetic crossover and uniform mutation in constraints bounded domains[J]. *International Journal of Computer Mathematics*, 2008, 85(9): 1341-1353.
- [48] 谢明霞, 郭建忠, 张海波, 等. 高维数据相似性度量方法研究 [J]. *计算机工程与科学*, 2010, 32(5): 92-96.
- [49] SUN N, LU Y. A self-adaptive genetic algorithm with improved mutation mode based on measurement of population diversity[J]. *Neural Computing and Applications*, 2019, 31(5): 1435-1443.
- [50] QIAN W, CHAI J, XU Z, *et al.* Differential evolution algorithm with multiple mutation strategies based on roulette wheel selection[J]. *Applied Intelligence*, 2018, 48: 3612-3629.

3D Facial Expression Animation Generation Based on Interactive Evolutionary Algorithm

ZHANG Wenting, LIU Mandan, ZHU Baoxu, XIE Lizhi

(School of Information Science and Engineering, East China University of Science and Technology, Shanghai 200237, China)

Abstract: Three-dimensional facial expression animation has been widely integrated into people's daily lives, but most existing expression generation techniques have not taken into account the end-user preferences. These methods mostly rely on pre-defined psychological archetypes, limiting their potential to generate rich and expressive facial expressions. Aiming at the above issues, this paper proposes a novel encoding scheme for individual facial expression animations, and uses Interactive Genetic Algorithm (IGA) to achieve automatic evolution of facial expression animations. During decoding, the Dirichlet Free-Form Deformation (DFFD) algorithm is used to control the deformation of the entire facial mesh. By introducing human evaluation and generating the initial population randomly, the aforementioned issues are effectively resolved. Additionally, to maintain the diversity of the facial expression animation population, two mutation operators are designed: Uniform Mutation with Exclusion Zones introduces the concept of exclusion zones, and Uniform Mutation with Shared Mutation Points considers the actual meaning of chromosomes and employs the same mutation points for gene segments that control the same parts. Finally, the experimental results show that these two operators effectively reduce the similarity within the population during the evolution process. The actual user participation experiments further validate the feasibility of the proposed approach in generating various facial expression animations that meet user preferences.

Key words: facial expression animation; IGA; user evaluation; mutation operators; population diversity

(责任编辑: 李娟)