

# 从手机短信到 3D 动画人物表情的自动生成<sup>①</sup>



赵檬檬

(北京工业大学 信息学部, 北京 100124)

**摘要:** 手机 3D 动画自动生成系统旨在通过输入一条手机短信得到与其相符的动画并发送给接收方. 动画中表情对于表达情绪主题以及强化动画效果等具有重要意义. 本文重点研究手机 3D 动画中人物表情的自动生成. 主要包括定性规划和定量计算两大部分. 定性规划部分主要利用语义网技术, 基于面部编码系统和情感轮模型构建表情本体库并建立相应公理, 然后根据短信的关键信息进行知识推理, 得到关于表情的定性描述. 在定量部分将定性描述信息转化为具体动画数据并处理表情平滑过渡等问题. 在实验的 270 条数据中除人为操作引起的异常外, 表情规划成功率为 71.2%, 其中多样性表情生成率为 86.89%. 实验表明该方法能够较好地生成表情动画.

**关键词:** 动画自动生成; 面部表情; 语义网; 本体; 定量计算

引用格式: 赵檬檬. 从手机短信到 3D 动画人物表情的自动生成. 计算机系统应用, 2019, 28(6): 13-21. <http://www.c-s-a.org.cn/1003-3254/6921.html>

## Automatic Generation of 3D Facial Animation from SMS

ZHAO Meng-Meng

(Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China)

**Abstract:** The 3D animation automatic generation system of mobile phone aims to obtain an animation which matches an input text message and send it to the recipient. Expressions in animation are important for expressing emotional themes and enhancing animation effects. This study focuses on the automatic generation of the character expression in the 3D animation system, mainly including the qualitative planning and the quantitative calculation. The qualitative planning part mainly uses the semantic Web technology to build an expression ontology library based on the facial coding system and the emotional wheel model and establish the appropriate axioms. Then knowledge reasoning is carried out according to the key information of short message to get the qualitative description of expressions. Followed is converting the qualitative description information into specific animation data and processing the expression smooth transition in the quantitative part, and finally a rich and diverse character expression animation is generated. Except for the abnormalities caused by human operation in the 270 data of the experiment, the success rate of expression planning is 71.2%, and the diversity expression generation rate is 83.76%. Experiments show that this proposed method can generate expression animations better.

**Key words:** animation auto-generation; facial expression; semantic web; ontology; quantitative calculate

## 1 引言

在网络高速发展的今天, 人们的通讯方式逐渐变成了短信、微信等, 为了使交流更具趣味性, 各种表情包、搞怪动画层出不穷. 出于相同目的, 中科院张松懋

研究员创造性地提出了将 3D 动画自动生成技术应用到手机短信上的想法, 通过一段生动的动画表达出短信内容. 4G、5G 通信技术的不断发展为该系统的实现提供了保障.

① 收稿时间: 2018-11-12; 修改时间: 2018-12-11, 2019-01-03; 采用时间: 2019-01-07; csa 在线出版时间: 2019-05-25

手机 3D 动画自动生成系统主要依靠全过程计算机辅助动画自动生成技术<sup>[1]</sup>,旨在通过对中文短信文本的处理生成一段能够表达短信内容的动画.系统以中文短信文本作为输入,利用哈尔滨工业大学自然语言处理系统对文本进行分词处理,进而进行命名实体识别、否定识别等操作,提取出短信中的主题、模板等信息,同时利用机器学习方法进行短信话题检测以及情感分析,并将其作为确定短信主题的参考值.将最终确定的主题、模板信息作为定性规划部分的输入,进行下一步处理.定性部分首先要完成情节规划,这也是整个动画开展的关键.情节规划需要根据主题、模板选择合适的场景、模型,并通过构建事件来丰富动画情节.其它子系统要分别实现布局规划、动作规划、表情规划等功能.定量计算是动画实现的真正部位,按照定性输出结果,通过三维动画软件 Maya 完成动画场景的搭建,并最终渲染生成动画.

目前,系统已经可以根据不同内容的短信生成动画,但是人物面部表情呆板,仅通过动作、环境烘托等来突出短信内容略显单薄,表达能力有限.相比于一个面无表情的人物在手舞足蹈,显然一个大笑的人物在手舞足蹈更能表现出短信内容.另外对于多个人物而言,即便表达的是同一个含义,表情也应该是不尽相同的.因此,在手机动画自动生成系统中研究人物表情规划是十分有必要的.

## 2 表情规划整体流程

整个动画自动生成系统主要包括信息抽取、定性规划、定量计算 3 个部分,其中各子系统相对独立,分别完成各自定性、定量部分的设计与实现.情节规划是定性规划部分中最先完成的,主要利用信息抽取得到的结果确定出动画选择哪个场景、添加哪些模型等,并通过规划人物多动作构建动画情节.在动作规划子系统中,需要对单动作、交互动作进行规划.表情规划子系统完成对人物表情类型、个数等的规划,其中由于动作与表情在一定程度上需要保持一致,故表情规划需要参考情节规划和动作规划后的结果.所有子系统定性规划结束后进而完成相应的定量计算:动作规划子系统要实现所有规划的动作类型的添加并计算人物模型运动轨迹等.表情规划子系统确定出表情操控参数以及不同表情间过渡的方式等,实现表情的添加.各子系统定量计算完成后即得到最终的动画.

表情规划整体流程如图 1 所示.主要以信息抽取得到的主题和模板以及情节规划和交互动作规划得到的动作类型为依据,通过知识推理得到相应的表情组合.将包含表情添加对象、类型、实现方式等信息的表情定性描述语句转化为 Maya 中控制模型面部变化的表情控制器的具体参数,并通过设置关键帧的方式生成表情动画.

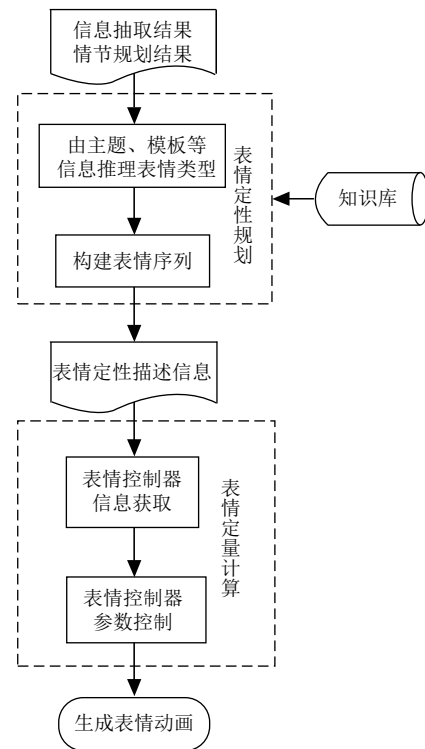


图 1 表情规划整体流程图

表情定性规划部分的主要任务是构建表情本体库并通过建立对象属性、构建公理实现从主题、模板等多种有效信息中推理得到最适合的表情种类.定量计算则是将对表情的客观描述在人物模型上加以实现,包括选择适当的关键帧位置及解决相邻表情切换时的平滑过渡等问题.

## 3 表情定性规划

表情定性规划是表情实现的关键,它决定了适合表达短信的表情类型、个数、展示时间等,使表情动画有理可依.定性规划主要包括构建知识库以及生成表情定性描述语句两部分.

### 3.1 表情本体库的构建

知识库是以描述型方法来存贮和管理知识的机构,

由知识和知识处理机构组成<sup>[2]</sup>. 与数据库系统不同, 知识库系统根据输入的数据信息进行推理后给出结果, 而不单是为用户提供可检索的信息. 目前系统已有主题库、模板库、事件库、动作库等. 主题库和模板库是系统定性部分的核心, 可按是否含情感分为情绪主题(模板)、非情绪主题(模板). 事件库依托于动作库, 按照一定规则规划出动作序列, 而动作库也同样包含情绪动作, 所以事件规划和动作规划结果都会对表情规划产生影响.

表情本体库主要由运动单元类和表情类两部分构成. 其中运动单元类建立的基础是面部基本运动单元(Action Unit, AU). Paul Ekman 在分析不同表情与不同面部肌肉动作的对应关系后, 提出了面部动作编码系统(Facial Action Coding System, FACS)<sup>[3]</sup>. 它以一种自下而上的方式将脸部分割成一系列基本运动单元, 通过 44 个独立运动的表情单元来描述面部动作, 绝大多数的表情都可以由这些单元组合而成. 根据动画人物表情的特点, 结合 FACS 中定义的 AU 类型, 构建运动单元类. 该类包括眉、眼、嘴三个子类, 每个子类分别包含相应部位的常见运动方式, 如图 2 中 mouth 子类包含嘴部向下(m\_downward)、嘴巴张大(m\_enlarge)等嘴部运动方式, 另外通过建立实例描述不同运动方式的具体实现形式, 如图 2 中 m\_downward 子类包含下嘴唇完全下拉(m\_alldown)、嘴角下拉(m\_down)等实例.

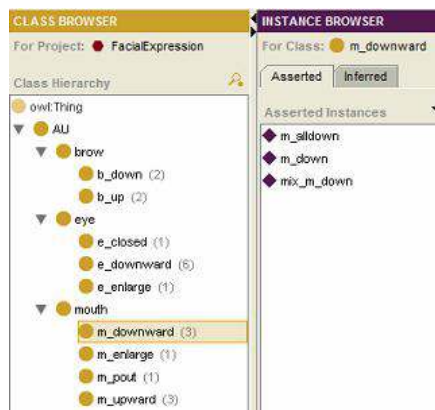


图 2 运动单元类

表情类以 Plutchik 提出的情感轮模型<sup>[4]</sup>为依据建立. Plutchik 将情感划分为八个维度, 包括信任、期待、喜悦、生气、恶心、害怕、伤心、惊讶, 在每个情感维度又按不同的情感强度划分为 3 组, 以这 24 个

情感特征来描述情绪. 利用情感轮模型分类结构, 结合目前系统能够处理的情绪类型, 构建出表情类(见图 3). 通过分析表情种类建立不同的表情实例, 如图 3 中狂喜(ecstasy)含 ecstasy\_120、ecstasy\_121 等 5 个表情实例, 表示该表情有 5 种不同的表达方式.

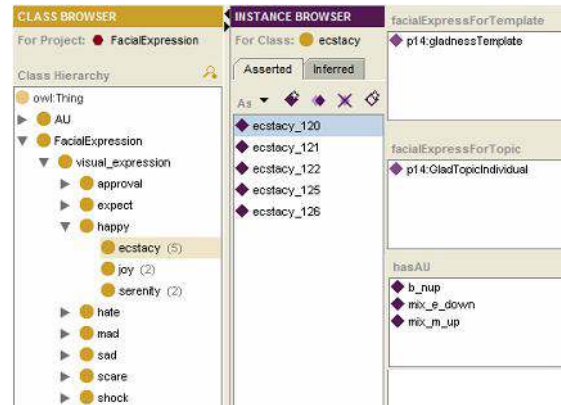


图 3 表情类及其对象属性

其中表情与 AU 组合的对应关系、主题与表情的对应关系等是通过创建对象属性并刻画相应公理建立起来的. 对象属性用以描述不同类之间的关系以及同一类下不同子类之间的关系, 其中类包含的实例均满足该关系. 本文创建了 hasAU、topicForFacialExpress、similarnext 及 templateForFacialExpress 四个对象属性, 其中 hasAU 属性将运动单元类和表情类联系在一起, 描述不同表情对应的 AU 组合情况. similarnext 属性在同一表情类下不同情感强度的子类之间建立关联, 用以描述诸如微笑-大笑的渐进式表情. 另外通过 topicForFacialExpress、templateForFacialExpress 属性描述表情类与主题类/模板类的关系, 表明该主题类/模板类对应的表情类型. 通过分析各类之间的相互关系建立公理, 图 4 展示了 anger 子类满足的公理, 表示生气的表情适合于争执模板(CollideTemplate)以及生气主题(AngryTopic), 存在渐进表情-烦恼(annoy), 能够通过眉毛向下(b\_down)、眼角低垂(e\_downward)、嘴巴撅起(m\_pout)或嘴部向下(m\_downward)的方式展示.

### 3.2 表情定性规划过程

本系统中, 确定表情类型的依据除主题、模板外, 还有动作类型及情感分类结果等. 其中情感分类结果只有喜、怒、哀、惧四类, 种类较少. 而基于事件的情节规划是以动作库为基础, 通过已有动作划分不同的事件类, 然后构建情节, 其结果以人物多动作来体现.



故通过短信主题、模板以及动作类型这三个参数即可确定表情. 表情定性规划流程如图 5 所示.

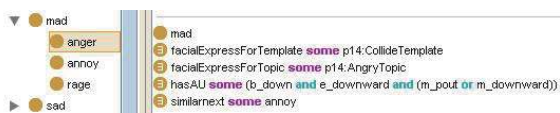


图 4 anger 子类相应公理

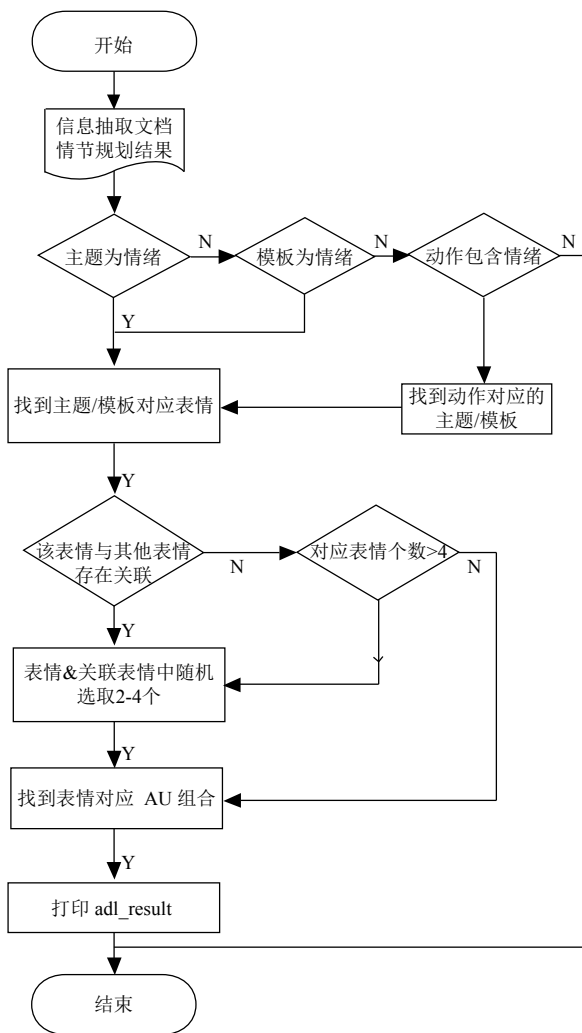


图 5 表情定性规划流程图

本模块以信息抽取和情节规划的结果作为输入. 信息抽取结果包括短信分词结果、短信主题以及短信包含的模板信息, 其中主题是能够表达短信核心思想的关键词, 目前本系统中存在包括节日、运动、情绪等 93 类不同的主题. 模板为短信中包含的有效信息, 包括地点、动作、情绪、生活用品等多个类型. 情节规划结果主要包括选定场景、主题、动画总长度、添

加模型的信息以及添加动作的类型等内容. 表情规划以信息抽取结果得到的短信主题、模板信息以及情节规划、动作规划得到的动作类型作为表情判断的依据. 其中以情绪主题为主, 情绪模板次之, 动作类型比重最低. 在确定出表情类型后, 根据相应的公理选择适当的 AU 组合方式作为表情定性规划的结果.

我们以短信“一起去吃饭吧”为例, 详细介绍表情定性规划的整体流程. 根据信息抽取结果 (见图 6) 中的 topic 节点 (包含主题信息) 及 root 节点 (包含模板信息) 可知, 该短信以吃饭为主题, 并提取出一个“走:去”的动作模板, 均不包含情绪, 进一步判断动作类型. 情节规划结果 (见图 7) 中 maName 节点包含选定场景、主题、音乐及动画总长度的信息, rule 节点为规划规则, 包括添加模型规则 (addToMa) 及添加动作规则 (addActionToMa), 情节规划选择添加了一个人物模型 M\_girl 和一个物品模型 M\_teapot, 并为该人物模型添加了两个动作 ElatedWalk142\_06 和 eating80\_24, 构建出“开心地走+吃饭”的情节. 由于该动作包含情绪, 故以动作为依据进行表情规划. 在知识库中动作库与主题库存在对应关系, 根据动作可以找到相应的主题, 根据已构建好的公理可以通过主题找到相应的表情实例, 进而找到每个表情实例对应的 AU 实例, 得到最终规划结果 (见图 8).

```

<result negType="" negCont="">
<message value="一起去吃饭吧"/>
<nessage value="一起去吃饭吧"/>
<segmessage value="一起#n1 去#vb 吃饭#vb 吧#ax"/>
<topic name="吃饭" key="吃饭">
  <root name="动作" flag="" value="走:去" color="">
  </topic>
</result>

```

图 6 信息抽取结果

```

<maName name="coffeeshop.ma" topic="EatActionTopic" music="db03.mp3" maFrame="260">
<rule ruleType="addToMa" addModel="M_girl.ma" class="YoungGirl"
spaceName="SP_coffeeshop_F" type="people" isTarget="1"
degree="" number="1" addModelID="addModelID1"/>
<rule ruleType="addToMa" addModel="M_teapot.ma" class="Pot"
spaceName="SP_coffeeshop_D" type="model" isTarget="0"
degree="" number="1" addModelID="addModelID2"/>
<rule ruleType="addActionToMa" type="action" eventType="WalkEvent"
useModelID="addModelID1" useModelInMa="M_girl.ma" actionName="ElatedWalk142_06"
distance="Default" reference="" ifAddConstraint="false"/>
<rule ruleType="addActionToMa" type="action" eventType="EatingEvent"
useModelID="addModelID1" useModelInMa="M_girl.ma" actionName="eating80_24"
distance="Default" reference="" ifAddConstraint="true"/>

```

图 7 情节规划结果

```

<rule ruleType="addExpressionToMa" type="FacialExpression"
useModelInMa="M_girl.ma" useModelID="addModelID1" source="ac_topic"
expression="ecstasy" controller="b_wup mix_e_down mix_m_up"/>
<rule ruleType="addExpressionToMa" type="FacialExpression"
useModelInMa="M_girl.ma" useModelID="addModelID1" source="ac_topic"
expression="ecstasy" controller="b_wup mix_e_down mix_m_up"/>
<rule ruleType="addExpressionToMa" type="FacialExpression"
useModelInMa="M_girl.ma" useModelID="addModelID1" source="ac_topic"
expression="ecstasy" controller="b_wup mix_e_down mix_m_up"/>

```

图 8 表情规划结果

表情定性结果各参数的描述如表1所示。

表1 表情定性规划参数描述

参数	取值	说明
ruleType	addExpressionToMa	规则名
usedModelInMa	M_girl、M_boy..	表情添加对象
usedModelID	addModel1..	表情添加对象 ID
source	topic、template、ac_topic..	确定情绪的依据： 主题、模板或动作
expression	ecstasy、joy..	表情类型
controller	b_nup e_down m_up..	AU 组合的描述

其中 usedModelInMa 取值为情节规划确定出的需添加的人物模型, usedModelID 取值为该人物模型对应的 ID. source 参数表示得出表情规则的直接决定因素, 如上述例子中 source 取值为 ac\_topic, 表示该表情规则以动作类型为依据推出. 除此之外还可取值为主题或模板. expression 参数取值范围即为表情类中各子类的值, 每个表情实例对应的 AU 组合用 controller 表示, 每一个 controller 参数均包含眉、眼、嘴三部分的运动方式. 对同一人物模型的多条表情添加规则表示当前模型需实现多表情连续动画。

表情类中喜怒哀惊等子类是最常见的表情类型, 相应的表情方式也多样, 而诸如期待、信任等子类对应的表情特点不突出, 种类也就相对较少. 为实现表情动画的多样性, 定性规划确定出表情后, 若该情绪子类对应实例较多, 则每次随机选取若干实例进行下一步规划, 若该情绪子类对应实例较少, 则将该子类与相近情绪子类的实例构成渐进式表情集合, 从该集合中选择若干实例进行下一步规划. 图9为上述短信的动画片段。



图9 动画片段

#### 4 表情定量计算

表情定量计算是表情动画实现的部分. 定性输出是对表情的客观描述, 真正让模型表情动起来还需在 Maya 中进行相应操作. Maya 中表情动画主要分为面部表情制作和关键帧动画制作两部分。

#### 4.1 面部表情制作

Maya 中制作表情主要通过三种方式实现: 表情装配、Maya 内置语言语句 (Maya Embedded Language, MEL)、表情捕捉<sup>[5]</sup>. 表情装配主要通过建立骨骼系统控制头部、下颌和口部的运动关系, 依靠变形工具和融合变形方法调节五官变化, 并利用非线性动画编辑器实现各表情之间的融合效果等方式在面部建立表情控制系统. MEL 语言动画是通过 Maya 内置语言编写动作程序从而控制角色动作生成, 主要用于集群及动力学动画制作中. 表情捕捉则是通过动态采集真人表情数据并映射到三维角色中从而生成极具真实感的表情动画。

本文主要利用表情装配的方法制作表情. 其中骨骼方向约束及建立父子关系的装配手段主要用于受骨骼控制的部位, 动画制作中常用于控制眼球、牙齿的运动. 簇及融合变形方法则主要用于对蒙皮的修改。

首先建立头部骨骼、下颌骨骼及其与面部、眼睛、牙齿等蒙皮之间的关系, 使得骨骼可以正确控制整个头部各部位的运动. 其中下颌控制主要通过建立下颌骨骼并与头部骨骼建立方向约束实现, 牙齿实质是蒙皮, 将其与下颌骨骼建立父子约束并调节权重即可通过骨骼控制牙齿运动 (见图10). 进而通过变形工具和融合变形方法调节五官变化, 分别参照 FACS 中定义的眉、眼、嘴三部分的常见运动方式做出各自的融合形状, 此时融合图形窗口中每个参数对应一种融合形状, 如图11所示。

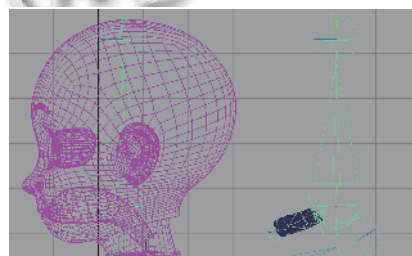


图10 头部骨骼系统

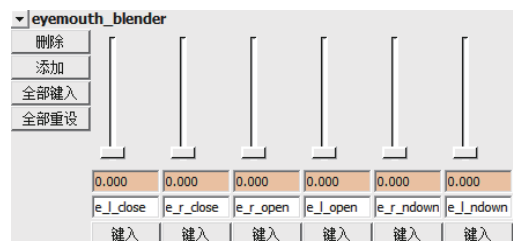


图11 融合图形窗口部分参数

由于融合图形窗口内的参数较多且互相独立,导致创建表情过程繁琐,故还需建立面部控制器来驱动融合形状.面部控制器实质上是与头部骨骼绑定的一组曲线,通过驱动关键帧方法建立各曲线的属性值如Y轴位移(translateY值)与融合图形窗口中各参数之间的关系,实现控制器对融合形状的直接控制.利用面部控制器能够直观地对面部进行操控,极大地简化了表情制作流程<sup>[6]</sup>.图12展示了面部控制器与生气表情的对应关系,图中加粗的部分为相应调节的局部控制器(若未加说明,下文中提到的控制器均表示局部控制器).

为精细控制面部表情动作,单一功能型控制器是必不可缺的.单一功能型控制器是指该控制器控制同一部位的融和形状,如图12中内侧眉毛控制器可以控

制内侧眉毛抬升、低垂.单一功能型控制器越多,角色表情设定就越自由,但其数量过多会降低动画制作效率.因此还需复合功能型控制器,该控制器可以同时控制多个部位的融和形状,把应用频率较高的表情动作设置为复合功能型控制器可以提高工作效率<sup>[7]</sup>.部分表情控制器设计见表2.

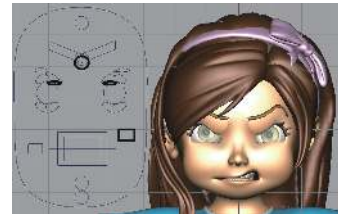


图12 控制器与表情对应关系

表2 部分表情控制器设计

控制器类别	命名	动作	装配手段	备注
骨骼控制器	jaw_con	下颌开、合	骨骼、方向约束	单一功能型控制器,配合口部动作
	tooth_con	牙齿开合	父子关系	
口部控制器	mouth_con	嘴角抬升、下拉		
	mouth_open	嘴巴张大、掀起		
眉毛控制器	l_brow	左眉内侧抬升、低垂	簇、融合变形驱动关键帧	单一功能型控制器
	l_browl	左眉外侧抬升、低垂		
	r_brow	右眉内侧抬升、低垂		
	r_browl	右眉外侧抬升、低垂		
眼部控制器	l_blink	左眼张大、紧闭		
	r_blink	右眼张大、紧闭		
	l_twist	左眼弯曲		
	r_twist	右眼弯曲		
	brow_mad	双眼紧皱		复合功能型控制器

#### 4.2 表情动画的定量实现

利用非线性动画编辑器可以实现多个表情之间的融合效果也可以实现不同表情之间的平滑过渡,但较强的依赖于手动调节.为此本文主要以关键帧动画为基础,针对相邻表情过渡时可能出现的各种情况作出处理,自动化生成表情动画.

关键帧动画是通过给出若干时间点(关键帧)处的状态并由电脑完成中间状态的变化和衔接的一种动画制作方法<sup>[8]</sup>.表情动画中关键帧处的状态即为不同的表情,而表情受面部控制器控制,故主要通过设置关键帧处的控制器参数(下文简称为设置关键帧)完成表情动画制作.本文根据动画总时长( $T$ )和表情个数( $n$ )初步确定每个表情的关键帧位置为 $Tk/n(0 < k \leq n)$ .由于各表情之间相互独立,除相邻表情有相同部位运动方式一

致的情况外,前一表情的面部各部位需在下一表情开始后回归初始位置,但各部位运动由面部控制器控制,某些控制器可控制两种运动方式,若这两种运动方式相邻,会导致控制器参数调节矛盾,表情规划异常.针对相邻控制器的不同情况需要采取不同的策略.

若相邻两个运动方式分别对应不同控制器,则在时间轴 $Tk/n(0 < k \leq n)$ 位置处置第一个控制器参数为最大值、 $T(k+1)/2n$ 位置处将该值置零设置关键帧即可.若相邻两个运动方式完全一致,控制器不做归零处理.若相邻两个运动方式不同但对应同一控制器,为解决切换过程中的归零问题,选择在一个较小时间范围内对参数值取反设置关键帧.

通过调节控制器参数大小可以得到不同幅度的表情,但效果不明显.实验发现通过同时调节使眼睛/嘴巴



张大的控制器(统称为 open 控制器)可以得到情绪强度不同的表情,如嘴角抬升的动作叠加嘴巴张大的动作可以得到咧嘴笑的表情效果等.该混合控制器以 mix 为前缀.其中 open 控制器参数以 0.5 为宜.由于此类控制器同时操纵 open 控制器,在与单独的 open 控制器相邻时,同样会遇到控制器参数设置矛盾的问题.对于 open 控制器在前,混合控制器在后的情况,可在时间轴  $T(k+1)/2n$  位置处将 open 控制器参数调至 0.5 并设置关键帧.对于 open 控制器在后,混合控制器在前的情况,由于不涉及控制器参数归零的问题,故只需在  $T(k+1)/2n$  将参数调至最大即可.以眼部运动方式 e\_l\_up-e\_l\_down-e\_l\_down-e\_open-mix\_e\_up-e\_open-e\_down 这组序列为例对控制器冲突问题的解决方法进行说明(见图 13).

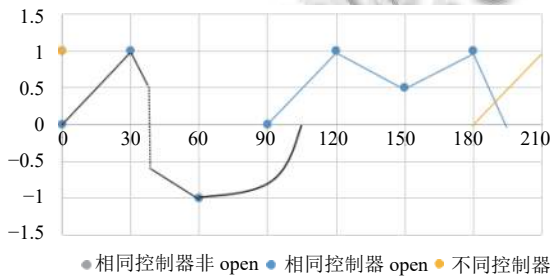


图 13 控制器冲突解决示意图

假定动画总长为 210 帧.眼部表情个数为 7 个,故每个表情展示时长为 30 帧.基本处理方式:30 帧内控制器参数达到最大(小),在下一表情开始后该表情参数逐渐归零.

对于 e\_l\_up-e\_l\_down-e\_l\_down,前两个关系为同一控制器不同参数,为解决切换过程中的归零问题,选择在 40 帧处将此时的参数值取反并将该值设置为 41 帧处控制器的参数值.而后两个是完全相同的控制器,则该时间处无需设置关键帧.对于 e\_open-mix\_e\_up-e\_open,关键在于 open 控制器的设置,据上述分析,open 控制器整体变化为从最大值减少到 0.5 最后升至最大值.具体见上图蓝色部分.对于 e\_open-e\_down,即为两个完全不同控制器相邻的情况,只需在最后一控制器调节过程中将前一控制器归零即可.

## 5 实验结果及分析

为了验证表情规划的结果是否符合预期,现分别编辑 3 条以生气、害怕、伤心、喜悦、惊喜等八种情

感为主题的短信以及不含情感因素的短信,每条短信测试 10 遍,最后统计这 270 条短信中不同表情规划的结果(见图 14).其中有 13 条未成功输出规划结果的短信,主要是人为启动程序不当造成.由于情节规划多样性的需求,并不是所有短信都会在定性规划过程中添加人物模型,故而导致 59 条短信因为没有表情规划对象而无法进行规划.另外还存在 15 条由于没有明显的情绪主题、模板、动作数据导致表情规划失败的短信.

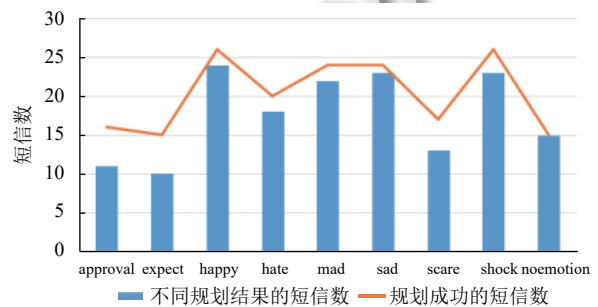


图 14 表情规划多样性测试结果

从图中可以看出,喜、怒、哀、惊四种最常见表情的多样性结果较好,基本达到 92.00%.赞成、期待、憎恶等表情以及非情绪表情的多样性结果稍差,在 80.72% 左右.这是由于这些情绪之前粗分子上述四种情绪中,所以系统在单独处理这些情绪时会有很大概率以不添加人物模型的简单方式处理,导致无法进行表情规划.另外,这些情绪对应的表情数目也没有以上四种情绪对应的表情多,故而会对表情的多样性展示造成一定局限性.

通过对以上信息的分析可知,本文提出的方法对大部分情绪可以做出较好的规划,但仍存在一些无法规划表情的情况,包括短信不含情感、情节规划没有添加人物模型等.这也使我们充分认识到目前的表情库并不完备,与其他各库之间的关系亟待进一步挖掘.我们将在以后的工作中继续做出改进.

图 15 是根据短信“一群人在开心地玩耍”对不同人物模型做出的不同表情规划的展示.

## 6 相关工作

在人物表情动画研究中的常用方法有:关键帧插值方法、参数化方法、基于肌肉模型的方法和行为驱动的方法.其中关键帧插值方法<sup>[9,10]</sup>在早期动画制作中应用较普遍,主要通过关键帧处插入面部数据,

进而采用插值算法在相邻表情间插入若干经过修改的面部网格点数据从而生成动画。关键帧插值方法能够快速产生基本人脸动画,但大量依赖于样本且产生动画种类有限,有一定的局限性。Parke<sup>[11]</sup>最早提出用表情参数和形状参数来描述表情变化,通过调整这些参数获得不同的人脸表情,在之后的30多年不断有研究人员加入到该方法的研究中,其中最成功的是1998年提出的MPEG-4标准<sup>[12]</sup>。MPEG-4表中定义的脸定义参数(Face Definition Parameters, FDPs)和人脸动画参数(Face Animation Parameters, FAPs)分别用于代表人脸几何特征和人脸基本运动,同时定义了特定人无关的人脸动画参数单元(Facial Animation Parameter Unit, FAPU)用于在不同模型面部展现相同表情,目前表情动画的研究多依据MPEG-4标准实现目标模型的重定向<sup>[13,14]</sup>。基于肌肉模型的方法是采用肌肉仿真技术来描述人脸上不同部位的肌肉运动。上文中提到的Ekman P所提出的面部动作编码系统就是基于该方法。Ahn S<sup>[15]</sup>等人进一步研究了通过估计肌肉收缩参数生成面部表情的方法。该方法能够实现逼真自然的人脸动画,但模型制作过程中需要一定的解剖学知识,不利于推广。而行为驱动方法是近几年兴起的技术,主要通过动态采集人脸面部数据,利用特征提取方法得到人脸三维运动轨迹,再通过变形传输技术驱动虚拟人脸产生同样的表情动画<sup>[16,17]</sup>。基于该方法主要有两大研究方向:实时动画生成、音频驱动口型动画的实现<sup>[18]</sup>。其中实时动画的研究包括利用深度摄像装置采集三维面部数据实现面部的三维重构(如真人CG电影)、利用单目摄像装置或根据视频动态采集面部数据驱动表情生成<sup>[19]</sup>及通过静态图片实现相应表情到三维模型的映射<sup>[20]</sup>。



图15 表情规划多样性展示

本文研究的重点是如何根据有限的信息生成多样化的表情动画,由于系统中的人物模型拓扑结构复杂,将模型参数化比较困难且运动边界很难划分,故参数化方法不适宜采用。而行为驱动的方法需要采集大量人脸数据,解决特征提取、模型映射、表情解耦等问题,且表情种类受限于采集的人脸数据。相较之下基于FACS并利用Maya融合变形技术能够更简单、快速地生成表情动画。

## 7 结语

手机3D动画自动生成系统旨在自动化地生成能够表达短信内容的动画。在传统动画制作过程中,人物表情都是需要动画师根据情境手绘出来,耗时费力。在数字动画盛行的今天,表情动画制作除关键帧动画技术之外,广泛采用了动作捕捉技术,通过动态采集真人的面部数据,驱动模型人物做出相应的面部表情。目前对表情动画的大部分研究都是基于该技术并利用机器学习方法开展的,包括人物表情仿真、声音驱动的面部表情自动生成、表情识别等。但本系统不追求表情的逼真程度,也无需添加口型动画,重点在于生成与短信内容相符的多种多样的流畅的表情动画,故选择通过基于知识库的方法加以实现。

我们动画系统的一个重要特点就是多样性,对相同的内容,生成的动画应该有多样化的展示效果。由于同一种情绪可以由很多种表情表现,我们不可能将所有的表现形式一一列举,只能在有限的范围内产生多种表情组合去表达该情绪,这也意味着本系统有一定的局限性。如何通过有限的信息产生出更丰富、准确的表情动画将是以后研究的重点。

## 参考文献

- 1 陆汝钤,张松懋.从故事到动画片——全过程计算机辅助动画自动生成.自动化学报,2002,28(3):321-348.
- 2 吴丹,易辉.知识库系统中语义网知识的表示.电脑与信息技术,2004,(1):9-11.
- 3 Cohn JF, Ambadar Z, Ekman P. Observer-based measurement of facial expression with the facial action coding system. Neuroscience Letters, 2005, 10: 203-219.
- 4 饶元,吴连伟,王一鸣,等.基于语义分析的情感计算技术研究进展.软件学报,2018,29(8):2397-2426.
- 5 吕新欣.Maya高级角色动画教程.北京:北京大学出版社,2007.



- 6 张天翔, 刘超. 三维动画创作-动画制作. 北京: 中国传媒大学出版社, 2014. 5-23.
- 7 柴兴祝. Maya 软件中面部表情动画控制器的设计与实现. 计算机教育, 2015, (24): 81-85. [doi: [10.3969/j.issn.1672-5913.2015.24.021](https://doi.org/10.3969/j.issn.1672-5913.2015.24.021)]
- 8 伊慧娟. 三维动画中的关键帧动画. 科技经济导刊, 2017, (11): 163.
- 9 Pighin F, Szeliski R, Salesin DH. Modeling and animating realistic faces from images. International Journal of Computer Vision, 2002, 50(2): 143-169. [doi: [10.1023/A:1020393915769](https://doi.org/10.1023/A:1020393915769)]
- 10 Lewis J P, Mooser J, Deng ZG, *et al*. Reducing blendshape interference by selected motion attenuation. Proceedings of the 2005 Symposium on Interactive 3D Graphics and Games. Washington, District of Columbia, USA. 2005. 25-29.
- 11 Parke FI. Parameterized models for facial animation. IEEE Computer Graphics and Applications, 1982, 2(9): 61-68. [doi: [10.1109/MCG.1982.1674492](https://doi.org/10.1109/MCG.1982.1674492)]
- 12 Abrantes GA, Pereira F. MPEG-4 facial animation technology: Survey, implementation, and results. IEEE Transactions on Circuits and Systems for Video Technology, 1999, 9(2): 290-305. [doi: [10.1109/76.752096](https://doi.org/10.1109/76.752096)]
- 13 李俊龙, 章登义, 黄珺. Kinect 驱动的人脸动画合成技术研究. 计算机工程, 2015, 41(3): 237-241. [doi: [10.3969/j.issn.1000-3428.2015.03.044](https://doi.org/10.3969/j.issn.1000-3428.2015.03.044)]
- 14 Obaid M, Mukundan R, Billingham M, *et al*. Expressive MPEG-4 facial animation using quadratic deformation models. Proceedings of 2010 Seventh International Conference on Computer Graphics, Imaging and Visualization. Sydney, NSW, Australia. 2010. 8.
- 15 Ahn S, Ozawa S. Facial animation based on muscular contraction parameters. Proceedings of 2015 IEEE International Conference on Systems, Man and Cybernetics. Waikoloa, HI, USA. 2005. 10.
- 16 Li BB, Zhang Q, Zhou DS, *et al*. Facial animation based on feature points. Indonesian Journal of Electrical Engineering and Computer Science, 2013, 11(3). [doi: [10.11591/telkomnika.v11i3.2039](https://doi.org/10.11591/telkomnika.v11i3.2039)]
- 17 Ruhland K, Prasad M, McDonnell R. Data-driven approach to synthesizing facial animation using motion capture. IEEE Computer Graphics and Applications, 2017, 37(4): 30-41. [doi: [10.1109/MCG.2017.3271467](https://doi.org/10.1109/MCG.2017.3271467)]
- 18 Schabus D, Pucher M, Hofer G. Joint audiovisual hidden semi-Markov model-based speech synthesis. IEEE Journal of Selected Topics in Signal Processing, 2014, 8(2): 337-338.
- 19 黄明阳. 基于实时表情驱动的三维人脸模型控制研究[硕士学位论文]. 哈尔滨: 哈尔滨工程大学, 2018.
- 20 Moeini A, Moeini H, Faez K. Pose-invariant facial expression recognition based on 3D face reconstruction and synthesis from a single 2D image. Proceedings of 2014 22nd International Conference on Pattern Recognition. Stockholm, Sweden. 2014. 1746-1751.