



从故事到动画片—— 全过程计算机辅助动画自动生成¹⁾

陆汝钊 张松懋

(中国科学院数学与系统科学研究院数学研究所 北京 100080)

(E-mail: rqlu, smzhang@math08.math.ac.cn)

摘要 论述了作者提出并实现的全过程计算机辅助动画自动生成技术,即从用自然语言书写
的故事开始,直至最终生成动画,每一步都是在计算机辅助下完成的,这是一种新的基于人工智
能的动画生成技术.文中还分别从自然语言理解、故事理解、动画设计和生成、规划、动画知识
库、常识处理几个方面对该技术进行阐述.

关键词 计算机动画,基于知识的动画生成,自然语言理解,故事理解

中图分类号 TP391.41

FROM STORY TO ANIMATION—— FULL LIFE CYCLE COMPUTER AIDED ANIMATION GENERATION

LU Ru-Qian ZHANG Song-Mao

(Institute of Mathematics, AMSS, Academia Sinica, Beijing 100080)

(E-mail: rqlu, smzhang@math08.math.ac.cn)

Abstract This paper summarizes the technique of full life cycle computer aided
animation generation proposed by the authors. Starting from a children story written
in Chinese natural language, until the generation of the final animation, the whole
process is done with the aid of the computer. This new technique is based on artificial
intelligence, and the paper describes its various aspects, including among others
natural language understanding, story understanding, animation design and
generation, planning, animation knowledge base, and commonsense knowledge pro-
cessing.

1) 中国科学院院长特别基金、国家自然科学基金重点项目(69733020)、国家自然科学基金项目(69773017)、国家“八
六三”高技术项目(863-306-ZT02-01-2)、国家“九五”攻关项目(98-780-01-07-09)、高等学校重点实验室访问学者
基金(吉林大学“符号计算与知识工程教育部重点实验室”)、中国科学院数学与系统科学院创新基金、中国科学院
数学研究所创新基金、中国科学院智能信息处理开放实验室基金资助

收稿日期 2000-12-13 收修改稿日期 2001-08-31

Key words Computer animation, knowledge based animation generation, natural language understanding, story understanding

1 人工智能技术在动画中的应用

计算机动画技术在近十几年得到了突飞猛进的发展,被成功而又广泛地应用于可视化科学计算、特种技术培训、教育、军事、模拟、动画片、广告、电影、电视、游戏、虚拟现实等等领域.一般说来,计算机动画技术分为关键帧动画、变形物体的动画、过程动画、关节动画和人体动画及基于物理的动画等^[1].计算机动画的研究课题包括动画硬件和软件、计算机辅助动画、绘画系统、动作控制、关键帧动画、图形模拟、基于力学的动画、合成角色、动画中的图像 render、动画语言和系统、动画者界面、电影实例研究、路径规划、三维医学图像、分子图形学、飞行模拟器、语音合成、视频动画、应用于计算机动画的传统动画技术、科学可视化中的动画、工程中的动画、基于人工智能的动画、动画中的机器人学方法、特殊效果、实时动画、录像录音技术、场景制作、故事图板生成、数字化音乐和声音、等等,其中基于人工智能的动画研究如何将人工智能技术用于计算机动画,从而提高动画制作的自动化程度和智能性.本节我们对若干类基于人工智能的动画技术及其研究现状进行综述,从第 2 节开始介绍我们提出并实现的从故事开始直至动画片的全过程计算机辅助动画自动生成技术.

1.1 自然语言指令驱动的动画

自然语言指令驱动的动画技术研究如何由自然语言指令和其它高级任务规范说明来产生虚拟人类 Agent 完成任务的动画模拟.下面我们介绍若干个有关语言、动画和虚拟现实的系统.

1) Ulysse^[2]是一个使用户可以用自然语言在虚拟现实环境中漫游浏览的 Agent,它主要包括一个语法分析器、语义模块、语言到虚拟实体的映射器、几何推理器等,Ulysse 使用规划规则来构造动画,并能和语音识别系统交互. Ulysse 的研制者目前的工作是研究如何从语言文本来生成场景.

2) 微软的 Persona 项目^[3]致力于产生能够和用户进行自然语音对话的拟人格化的动画角色,该项目基于微软开发的一个独立于谈话者的连续语音识别系统和一个覆盖面广的英语理解系统,Persona 采用拟人化的对话机制,角色说话时还有动作相配合.

3) 美国海军研究室致力于研究面向浸入式虚拟现实环境的自然语言和语音理解界面 Nautilus 项目^[4],通过使用口头命令,用户可以在模拟的三维环境中漫游、移动或隐藏虚拟物体、控制模拟回放等,一个包含有关虚拟世界和物体信息的知识库支持了这些功能的实现.

4) 美国宾州大学对基于指令的动画生成技术已经进行了十多年的研究,研制成功了许多系统^[5~10],他们的 AnimNL(Animation from Natural Language Instructions)项目的目标是生成真实的动画来表现人执行自然语言指令所说明的任务的过程,它的功能流程一般说来包括自然语言语法分析、语义分析、规划推理、规划、模拟和人物动画;AnimNL 支持对指令的语义和语用理解,以及这种理解随着动作的演化;AnimNL 所生成的动画不仅包括人的一般肢体动作,还有人的面部表情、手势、和语音合成;AnimNL 提供了各种知识库以支持各部分的功能.

1.2 知识驱动的人体跑步动作的动画生成

文献[11~15]介绍了生成各种实时人体跑步动作的动画高层控制技术.所谓高层控制

是相对传统的低层关键帧技术而言的,是通过采用有关人类跑步的知识来实现的,该系统中的知识分为若干类:关于跑步动作参数之间关系的实验知识,用于计算人体运动轨迹的物理知识,关于跑步时肢干协调性的知识,等等.在这种基于知识的控制机制的支持下,用户可以交互式地调整高层参数以实时地得到满意的跑步动作效果.

1.3 模拟人类对话的动画形象

美国 MIT 媒体实验室的手势和叙事语言小组一直在研究和构造用于对话的逼真的动画形象^[16~18],基于对人类语言、认知和社交能力的深层理解,他们的动画形象能够模拟人类和用户对话,包括语音、语调、手势、面部表情、头部运动、眼神等.他们的系统的主要特点包括多模式输入和多模式输出、实时、命题性和交互性信息的理解与合成及具有诸如转折等对话能力.系统的静态知识库、动态对话知识库、规划模型和对话模型等支持了这些功能的实现.他们已经研制了若干能和人对话的动画角色,例如 Avatars, Gandalf, Rea 等.

1.4 一系列相关事件描述到角色语音和面部表情的自动生成

Byrne 系统^[19]以 RoboCup 模拟器的输出为输入,基于角色的个性、情感状态和比赛情形来产生角色富有感染力的言语和面部表情,Byrne 的输入是关于一场模拟足球比赛的当前状态和最近动作的一系列相关事件描述,每一个事件描述都被预赋一个相关值,事实上 Byrne 的动画和语音自动生成是基于比赛分析和事件相关性计算都已完成的假定的,这大大减少了它的处理复杂性. Byrne 的主要子系统包括情感生成、情感行为生成、文本和语音标记生成、表情标记生成等,它基于一组情感结构生成规则、评球员角色的个性、组合规则等. Byrne 的一个重要特点是它大量使用了有关语音合成、面部表情和自然语言语义语用处理和生成等的各种系统标准.此外还有 Maiocchi 关于“自动演员”的工作^[20].

1.5 交互式故事系统

一般交互式故事或游戏系统有故事图和模拟世界两种类型,前者的交互性太差,而后的行为时间相关性太差,难于成为故事.面向兼顾交互性和情节控制的目标^[21],描述了一个支持交互式情节的动态产生、管理和冲突解决的计算框架,它让用户充当故事的主角来确定当前的行为,同时根据角色的说明、关系、目标等来控制角色的行为.该框架的核心模块是一个情节管理器,它的输入是一系列初始情节条件,输出是角色动作序列,该系统目前已在网上发布运行.

1.6 自然语言故事到计算机动画的翻译

MULTRAN^[22,23]是一个从自然语言(日文)写的故事到计算机动画的翻译器,其主要步骤包括:从故事篇章中推导出脚本、角色动作生成、动画环境构造和虚拟摄像机定位. MULTRAN 的特点在于:中间表达由事件和事件间的约束组成;能够生成一个动画脚本,它是时间估算的一个最优解;调度执行的过程中角色可以估算自己动作的时间. MULTRAN 的研制者期望他们的工作可以支持自然语言处理中的对话理解和篇章的二义性排除.

1.7 交互式故事图板生成

故事图板(storyboard)是指电影、动画片、电视节目或商业广告等的情节串连图板,主要用于设计镜头效果和指导拍摄.现行的故事图板生成系统一般都是交互式的,由用户利用系统所提供的功能操作和库逐个图地进行画面设计,每个图可以附加标题和脚本文字说明.以 StoryBoard Quick^[24]为例,它是为导演、作家、制片人、摄影师、电视录像制作人所设

计的用于前期可视化的实用软件,在制作每幅画面时用户可进行的操作主要包括:从系统所提供的库中调入角色、选择角色的动作和方向、放大或缩小角色、添加背景图或图像、从系统所提供的库中调入道具和各种图示符号、添加文字说明等等. StoryBoard Quick 的目标是简单易学,便于用户对有关镜头的想法进行快速可视化和交流. 还有的交互式故事图板生成软件可以为画面附加声音效果、把画面输出到视频设备等^[25]. 相关的工作还有文献^[26,27].

1.8 动态实时摄像机自动控制

计算机动画中有关摄像机自动控制方面的研究比较少,虚拟摄影师系统^[28,29]面向捕捉虚拟三维环境中的事件而自动产生实时的、完整的摄像机规范说明,这是基于对电影摄影知识的总结和表达. 该系统中有两类知识表示结构:电影惯例和摄像机模块,每个电影惯例负责描述一种类型的场景,例如两个角色对话,电影惯例选择一组镜头并说明镜头和镜头转换的时间及镜头之间的排序;一组摄像机模块为电影惯例所共享,它们描述为拍摄每种镜头特定的摄像机的几何位置,还负责角色位置的微调以获得最佳拍摄效果.

1.9 基于角色情绪和动作的自动背景音乐生成

计算机动画中有关背景音乐和音响效果自动生成方面的研究比较少,文献^[30]论述了一个为已经制作好的动画自动配乐的原型系统,其输入是每个场景的音乐参数和动作参数. 音乐参数包括情绪类型和音乐主题,情绪类型是指场景中主要角色的诸如高兴、伤心、生气、困倦等的情绪及其程度. 该系统中背景音乐的节拍、旋律、伴奏等主要是根据角色情绪所产生的. 背景音乐的和弦级数的生成是基于协调规则的,这里的协调规则是根据 17 到 19 世纪古典音乐最常使用的和弦而总结出来的,而动作的音响效果是根据动作的特性和强烈程度来确定的.

2 《天鹅》——全过程计算机辅助动画自动生成

我们将人工智能技术和基于知识的方法^[31,32]引进动画生成的全过程,提出了一种新的计算机动画技术——全过程计算机辅助动画制作. 这实质上是一条动画片生产自动化的技术路线,它的目标是只要有了一个适当的故事,以受限自然语言的形式把它输入计算机里,从此时开始,直到最终生成动画,每一步都是在计算机辅助下完成的. 粗略说来,这些步骤是

第 1 步. 对故事文本作自然语言理解;

第 2 步. 故事情节理解;

第 3 步. 把故事改编为分场景剧本;

第 4 步. 根据分场景剧本作动画设计,包括角色、背景、动作等设计,其中体现了时间、空间规划;

第 5 步. 根据上述设计和规划,利用事先构造好的动画素材库和声音库,生成完整的动画.

经过多年的努力,我们已经在计算机上实现了这条技术路线,并研制了相应的软件系统《天鹅》^[33]. 在一定的限制条件下,用户可以用中文自然语言编写一个童话故事,《天鹅》就可以把这个用中文写的故事自动转换成一个动画片. 转换过程原则上不需要人的干预,但《天鹅》也在若干个不同的转换层次上提供了人工干预的手段,以使用人的智慧补充系统艺术创

作能力的不足.

《天鹅》系统的功能主要包括:中文自然语言理解,上下文无关常识性处理,上下文有关常识性处理,故事情节的可表达性处理,故事情节的展开和生成,故事情节的完整化处理,故事中心思想分析,分场景剧本生成,场景规划(包括角色、道具、布景等在内的空间规划),路径规划(包括二维和三维路径规划),角色关节动作规划,摄像机规划,动画定量描述剧本生成,动画生成,动画元素库及库管理,语言库及库管理,设计库及库管理,声音元素库及库管理,动画连续演示,总控及用户界面,等等. 下面我们以一个非常简单的故事为例,分别从自然语言理解、故事理解、动画设计和生成、规划、动画知识库、常识处理几个方面来对《天鹅》进行论述.

3 中文自然语言单句理解

3.1 受限中文语法语义分析

显然,用一个计算机系统来完整地表达和理解中文自然语言是不现实的,因此我们根据儿童童话故事的语言特点总结出一组语法规则,称之为受限自然语言,它是中文自然语言的一个子集. 我们的自然语言单句理解系统能对这种语言进行识别处理,其主要功能为分词、语法分析和语义分析. 对分词,我们采用了通常的计算机分词方法——最大匹配法,并建立了一个较大的汉语词汇底表;对语法分析,我们定义了一种改进的适用于中文的 ATN,称为 CATN,其中句子的结构主要为“主语+谓语+宾语”的形式;由于 ATN 涉及细节太多,结构化不强(CATN 亦不例外),我们又以定子句的形式用 Prolog 语言重写了句法分析程序;对语义分析,我们以格框架树的形式表示分析结果. 1968 年 Fillmore 提出的格框架模型在分析句子语义时仅以动词为中心,用它表达中文自然语言不是十分方便. 根据理解童话故事的需求,我们扩充了格框架根节点的类型,定义了自己的格框架树结构 GF_2 ,并以一种文本性的表达输出,以使用户修改.

假设输入的自然语言故事是“黑炭王子过生日,白雪公主是黑炭王子的朋友,她很高兴,她向他祝贺生日”. 经过语法和语义分析,生成的格框架文本表达如表 1 所示.

3.2 基于常识的自然语言理解

常识在《天鹅》的自然语言理解中起着很大的作用,这里介绍我们两个方面的工作. 故事理解要用到更多的常识,将在下节介绍.

表 1 自然语言单句理解输出实例

Role-key: R-1
Role-name: Coaldark
Role-identity: prince
End-of-role
Role-key: R-2
Role-name: Snowwhite
Role-identity: princess
End-of-role
Role-key: R-3
Role-pronoun: she
End-of-role
Role-key: R-4
Role-pronoun: he
End-of-role
Action-key: A-1
Action: felicitate birthday
Action-subject: R-1
End-of-action
Action-key: A-2
Action: being friends
Action-subject-1: R-1
Action-subject-1: R-2
End-of-action
Action-key: A-3
Action: feel happy
Action-subject-1: R-2
End-of-action
Action-key: A-4
Action: congratulate birthday
Action-subject-1: R-2
Action-indirect-object-1: R-1
End-of-action

3.2.1 基于常识的指代处理

在上面的例子中,指代问题是利用常识解决的.两条常识起了作用:1)王子是男的,公主是女的;2)一个人不能向自己祝贺.于是故事成为“黑炭王子过生日,白雪公主是黑炭王子的朋友,白雪公主很高兴,白雪公主向黑炭王子祝贺生日”.但如果把白雪公主改成白雪王子,最后两句话就会成为“他很高兴.他向他表示祝贺”.此时就需进行指代的二义性处理.

二义性是自然语言中普遍存在的问题,它实质上是意义与形式之间的矛盾问题,通常有以下表现形式:词的用法二义性,词义二义性,语法结构二义性,范围的二义性,省略的二义性,指代的二义性,关系的不确定性,分词引起的二义性,格语义的二义性,等等.其中指代问题可以分为三个处理层次:语法语义分析(利用语法语义学知识来找到指代对象),句内知识分析(利用简单常识性知识在句间分析出指代对象),句群知识分析(利用常识性知识在全文范围内基于上下文来找到指代对象).目前的研究工作大部分集中在语法语义处理层上,相继提出了一些正确率颇高的算法,也有的研究是基于规划和意图的话语理解.我们在研究用常识处理二义性时,提出了一种利用角色情感来处理指代问题的方法,这是一种基于知识的二义性排除法,基本的处理过程是:对故事逐句分析,记录每个动作引起的其角色情感的动态变化;当遇到“代词+情感形容词”时,比较各个角色的情感变化历史,选出代词指代对象.为此,我们进行了人类情感分类、描述情感的形容词分类、基于动作对角色情感影响的动词分类、动词与情感形容词之间的对应关系分析等研究,相应地建立了基于情感公理的知识库,并给出了代词指代的处理算法.

从表1的格框架可以看出,该故事一共有4个角色.其中的R-3和R-4是虚角色,它们是代词,需要在推理过程中代真.推理所用的情感公理包括

$\text{felicitate}(x, z) = > \text{feel}(x, \text{happy}),$

$\text{being friends}(x, y), \text{feel}(x, \text{happy}) = > \text{feel}(y, \text{happy}).$

这两条公理当然仍不能解决上面的二义性.但是如果再加上语言学上的排二义性原理——按代词和被指代对象之间的距离的倒数加权,即可判定高兴的是白雪王子.

不难看出,这种基于情感的推理是纯语言学方法所不能代替的.

3.2.2 自然语言故事的常识检查

常识检查主要分上下文无关和上下文有关两类.

1) 上下文无关常识检查

上下文无关常识检查是针对故事中单个语句的语义语用而言的,主要检查分析单句描述的情节是否符合常识,例如“苹果吃皇宫”、“国王和王子结婚”就不符合常识.我们建立了一个常识限制库,规定了所有词的分类和属性,词属性可以为生物型、非生物型、地点型、动词型、静态描述型、动态描述型等,其中每一类型又可继续细分,特别地我们规定了动词的属性,规定了它的施动格和受动格属性、地点格属性等.包括:

① 对动作主体格的限制

有些动作只能由生物化的角色才能做,例如“皇宫”不能“吃饭”;

② 对动作对象格的限制

包括对外型、尺寸和递归等的限制,例如“公主吃了公主”是不允许的;

③ 对动作地点格的限制

包括对位置、介质、尺寸、距离、障碍等的限制,例如“王子在桌子上游泳”是错误的;

④ 对动作数量格的限制

例如,一个人同时只能与一个人结婚,一个句子中的地点个数应为 1,因为动画表现中一个动作同时只能发生在一个背景中,我们专门建立了一个成分数量库,用来规定对某些动作成分的数量要求;

⑤ 对动作本身的限制

动作在进行过程中由于受外界影响而必须满足一定的条件,例如“天刮风下雨”时雨线应随风飘动而不是垂直下落.

我们为动作的每个格关系设立各种属性,规定属性的匹配原则,当原则被违背时或判错或进行某种处理.需要指出的是,考虑到动画所特有的夸张性和非现实性,我们常识库中的属性规定并不是固定一成不变的,允许放开一些限制以表现特殊情节.

2) 上下文有关常识检查

上下文有关常识检查是针对故事中多个语句描述而言的,如“公主死了,公主拍球”不符合常识,但“公主死了,公主复活,公主拍球”可以;又如人的一只手一般最多能被砍五次手指头,而头只能被砍一次.为此我们定义角色的一些特殊属性,在情节发展过程中动态地记录这些属性值的变化,并以此为依据进行常识检查.但是这些功能在《天鹅》中还是非常初等的.

4 故事理解

经过自然语言单句理解后,《天鹅》开始进行故事理解.故事理解的内容主要包括该故事的主题、主要角色、人物特征、故事发展线索等,这都是一般的自然语言理解所不涉及的.在故事理解技术和故事理解深度之间有一种交换关系:采用的技术简单,则实现方便,但理解的程度也浅;反之,如果要加深故事理解的程度,则技术就要复杂,实现也困难.下面分别论述我们提出的两种故事理解方法:浅方法和深方法.

4.1 浅方法

这主要是角色图方法.角色图是一个无向图,其节点为故事中的角色,每个节点含一个整数集,称为出现集,表示该角色所参与的动作(语句)编号.每两个节点的连线也含一个整数集,称为联络集,是这两个角色共同参与的所有动作(语句)的编号.利用角色图可以对故事做多方位的分析.例如,角色图的最大连通分量可能是故事的主要情节,最大连通分量中连线最多的节点可能是故事的主角,等等.我们还可以定义事件图,它是一个有向图,表示故事中事件的走向.利用事件图可以分析故事的情节发展线索和任务发展线索,等等.这两种图统称为故事图.基于故事图的分析虽然不一定准,但有相当的参考价值.

上面的故事可以生成有两个节点:黑炭王子 $=\{1,2,4\}$,白雪公主 $=\{2,3,4\}$;三条边:(黑炭王子,黑炭王子) $=\{1\}$, (白雪公主,白雪公主) $=\{3\}$, (白雪公主,黑炭王子) $=\{2,4\}$ 的角色图.在这里整数都表示故事中句子的编号.其中前两条边都由一个节点通向自己.因此一般来说角色图是一个超图.该图只有一个连通分量.在这个简单的故事中,两个角色都是主角.

4.2 深方法之一——基于高维和上下文有关文法的故事理解法

我们的研究是基于这样一种思想:用形式语言中的文法来表达关于故事理解的知识,使用这些知识来理解故事的过程就是用文法的语法分析器对故事进行识别的过程.基于这种

文法描述故事的思想,我们将形式语言和编译理论与故事理解相结合,为表达语义和语用,形式语言中的上下文无关串文法显然不够用了,必须改造使之具有高维结构和上下文有关性,从而表达故事的深层含义和故事情节之间的内在联系.用这样的新型文法对故事进行理解的过程实质上就是它的语法分析算法对故事进行语法、语义和语用的综合分析过程,我们称这种理解为故事分析,称这样的文法为故事分析文法(story parsing grammars).有关工作的细节见文献[34~44].这里只是介绍一下故事理解的主要思想.

1) 故事分析文法

故事分析时故事已表达为代表各个单句语义的格框架树,这些树并置在一起称为格框架森林,所以故事分析文法不是简单的串文法,它的产生式应具有高维结构,同时它也不应是上下文无关的.为此,我们引入一类高维文法——森林文法,它的产生式左部和右部都是树结构的并置,即森林.故事分析文法是一种特殊的上下文有关的森林文法的子类文法,故事分析文法实质上显式地表示语法、语义和语用信息,其中对语法和语义的表达主要体现在产生式中各树上节点的标记及其之间的关系上,而对语用的表达则主要体现在产生式左部和右部的重写关系上.实际上,故事分析文法的推导过程是对故事内容从抽象到具体的逐步生成过程,即故事分析文法表示了对故事内容的不同深度和抽象程度从高到低的转换.抽象的最高层次是某种故事分析目标,如获取中心思想,抽象的最低层次就是故事本身.故事分析文法所产生的语言是格框架森林的集合.

2) 故事分析过程

根据编译理论的思想和方法,针对某类故事和某种故事分析目标,构造一个故事分析文法及其语法分析器.任给一个故事,如果经过该分析器证明这个故事属于该故事分析文法所产生的语言,则表示该故事可以被抽象成这个故事分析文法中的分析目标,由此完成了一个故事分析过程.

3) 故事内容的多层次分析

一般说来,故事理解的最高抽象层次是获取故事的中心思想,与故事的中心思想的总结有关的理解内容包括对故事主要情节的分析、对故事进行分段以及分析每个段落的段落大意等等.一个故事的中心思想代表了整个故事所要表达的主要的深层含义.故事的一个情节是指表达了一定含义具有一定语义联系的故事中的某些动作或事件等;故事的主要情节是与故事的中心思想的总结密切相关的那些情节,而非主要情节常常是对故事中心思想来说可有可无的一些情节;故事的分段是根据中心思想把故事按其叙述次序分成若干个部分,每个部分,即段落,相对独立地表达了某种深层含义,即段落大意.故事内容的多层次分析的结果应该尽可能地接近现实世界中人的这种常识性的理解.故事分析文法的构造亦应支持这种分析.

我们对故事的中心思想、主要情节、段落和段落大意进行了形式化定义,并较为深入地研究了对这些故事分析内容的各种启发式的分析方法并给出了有关的算法.我们进行这些研究的意义首先在于加深了故事分析的深度,扩大了故事分析的广度,使得能够对故事进行多层次的理解.这些研究的意义还在于尽可能地避免故事分析文法的异常庞大和由此而引起的处理的低效和灵活性差.对故事内容的各种层次的分析都是启发式的过程,我们研究的各种启发式的方法、规则和函数,避免了穷举方法的指数复杂度.

4.3 深方法之二——角色动态特征分析法

角色动态特征分析法是一种以故事结构信息为导引,以语句的语用含义分析为核心的故事理解方法,是面向角色的分析方法.面向角色的方法比面向事件的方法更易于理解故事中非常规的情节(即出人意料的情节).遗憾的是,迄今为止的面向角色的研究往往只着重故事角色某一动态特征的分析,却没有一个模型能用统一的机制分析角色的多种特征,而多种角色动态特征的综合分析无疑对于深入理解故事及故事知识获取有着重要的意义.因此这就构成了我们探讨故事角色动态特征分析的原因.

显然,角色的动态特征不同于其静态特征,静态特征可以是一些在故事中不随时间变化而变化的特征,而动态特征则是在故事中随着时间变化和情节发展而变化的特征.我们选取了五个动态特征作为分析对象——意图、信念、情感、人际关系、品格.角色的这些动态特征便构成了模型中读者信念(在故事理解中这读者即模型本身)的主要内容.每分析一个故事语句,读者信念中的这些特征便动态地发生变化,反映了模型对语句语用含义的理解,而全部故事语句分析完毕之后得到的一系列读者信念便反映了模型对整个故事的理解.我们的角色动态特征分析以指向分析、进展分析和结局分析为基础,用故事结构信息来引导分析的进行.

1) 指向分析

以往的故事理解模型均没有考虑指向分析而只研究故事的进展和结局部分,事实上大部分的故事均有指向部分,对故事发生的背景做出交代,它是故事理解中不可缺少的内容.指向部分由描述一切静物和一切行动状态的描写句构成,它描述了故事发生的时间、地点、角色静态特征等等.指向分析的目的就是建立起初始的读者信念,更具体地说就是加入故事的背景信息,并且根据角色静态特征的描写,在理解故事进展部分之前对角色有一种先入为主的印象,如建立起初始的角色品格信息等等.

2) 进展分析

故事的进展部分由描述人物行动和事物变化的叙事句和对话构成,进展分析是故事理解的核心内容.我们给出的进展分析算法的功能包括分析角色动作和话语的主、被动性、角色信念集合的增量构造、角色情感反应集合的增量构造、角色品格信息集合的增量构造、语句候选意图集的推导、全局意图的筛选、对方当前意图的筛选、角色全局意图结构的再构造、对方意图实现与否的评价、角色关系值的再确定等等.

3) 结局分析

结局部分描述了故事里角色的命运,目的实现或失败等等.故事中一般有三类矛盾,即角色间意图冲突、角色信念与现实矛盾及角色目标与环境的冲突,结局分析就是对矛盾冲突的结果进行分析.相应地三种结局分析的内容分别是谁胜谁负的分析、就错误的信念是否已得到更正的分析和角色目标是否实现的分析.

4) 评议分析

评议部分实际上可看成局部的指向或结局,也就是对故事某一个情节的背景作出交代或对一个情节的结局作出评价.因此,评议分析算法就是指向分析算法和结局分析算法的结合.

5) 故事创作意图及创作技巧分析

在角色动态特征分析的基础上,我们进一步探讨了故事整体含义的理解,其中重要一部分是研究故事创作意图及创作技巧的分析.我们提出了一种用于描述故事创作信息(包括主

配角的信息、创作类型信息、情节发展过程和结局信息)的故事创作框架,并且对故事主角的识别、主配角之间关系的分析、主要情节分析等等进行了详细的讨论,为以角色动态特征信息为基础建立起故事创作框架提供了一系列算法.

应该说明,目前故事分析的深方法只有一个实验系统,尚未应用于《天鹅》中.

5 动画定性规划

如前所述,《天鹅》的工作流程分为(故事)分析和(动画)规划两大部分.而动画规划又分为定性规划和定量规划两大阶段,规划的结果分别通过动画定性描述语言和动画定量描述语言表示.本节只讨论定性规划部分.

5.1 情节规划

在用户给出的原始童话故事中可能会有大量的情节是无法用图形语言直接表达的,因为它们不够具体.例如,黑炭王子过生日,怎么个过法?是出门去喝酒?去跳舞?还是邀请亲友共同庆祝?白雪公主向黑炭王子祝贺生日,如何祝贺?是送鲜花?打电话?寄贺卡?还是送蛋糕?至于白雪公主很高兴这类精神活动,则更难直接表示了.我们统称这类活动为抽象活动.动画自动生成中不可缺少的一步就是要把所有的抽象活动转换为具体活动.在《天鹅》中我们称之为情节规划.情节规划要基于知识,特别是常识.为此我们建立了一个情节常识库,针对各种生活情节建立了一系列的常识单元.它们分成两类:一类是剧本agent,另一类是情节agent.每一类都有自己的继承体系.

定义 1(剧本 agent 和情节 agent 的简化语法).

〈剧本 Agent〉::=

DRAMA AGENT(〈Agent 名〉):

Father: 〈父剧本 Agent〉

Child: 〈带条件的子剧本 Agent 序列〉

Drama Description Begin:

[〈剧本描述部分〉]

Drama Description End

[T-Specification: 〈时态条件序列〉]

[Candidate: [Object: 〈对象序列〉]

[Role: 〈角色序列〉]

[Action: 〈动作序列〉]]

END (〈Agent 名〉)

| PLOT AGENT (〈Agent 名〉):

Father: 〈父情节 Agent〉

Child: 〈带条件的子情节 Agent 序列〉

Plot Description Begin:

[〈情节描述部分〉]

Plot Description End

[T-Specification: 〈时态条件序列〉]

[Candidate: [Background: 〈背景序列〉]]

[Object: <对象序列>]
 [Role: <角色序列>]
 [Action: <动作序列>]]

END (<Agent 名>)

在上述语法中,<剧本描述部分>由剧本说明部分和一组不含 Action 语句的广义格框架语句(简称 GCF)组成,其核心部分是对其他 剧本/情节 agent 的调用(Call 语句).<情节描述部分>由情节说明部分和一组不含 Call 语句的 GCF 组成,其核心部分是 Action 语句. GCF 可以是格框架语句(CF)或格框架变量(FR). 剧本说明部分主要含角色(即人物)和对象(三维空间物体)说明,情节说明部分除含角色和对象说明外,还可以含背景和人物布局说明(即基本场景,见 5.3 节).

下面我们给出格框架语句的简化和近似语法,共有 8 种格框架语句.

定义 2(格框架语句).

1) <Action 语句> ::= <Action 名> (<一些参数>)

Action 语句表示一个具体动作,其中的参数可以是角色、对象等具体的参数值;

2) <Enrich 语句> ::= Enrich (<父辈 agent 名> [<一些参数>])

Enrich 语句利用它带的参数把父 agent 中的某些情节具体化(丰富化);

3) <Call 语句> ::= Call (<Agent 名> [<一些参数>])

Call 语句通过它带的参数调用指定的 agent;

4) <Conditional 语句> ::= If <岗哨> Then <无标号 GCF> [Else <GCF>] Fi

Conditional 语句是一种带哨语句,根据岗哨的值确定选择哪个 GCF;

5) <Case 语句> ::= Case <Conditional 语句序列> Esac

Case 语句是一种分情形语句;

6) <Bloc 语句> ::= Bloc <GCF 序列> Colb

Bloc 语句把一组 GCF 编在一起,成为一个语句;

7) <Loop 语句> ::= Until <事件> Do <Action 语句序列> Od

| While <状态> Do <Action 语句序列> Od

| Repeat <Action 语句序列> For <自然数> Times

Loop 语句指定动画播映时应该执行的循环,其效果在静态时不可确定;

8) <Replace 语句> ::= Replace <父辈 Agent> (<标号 1>) With (<标号 2>)

Replace 语句用具体的 GCF 例化父辈 Agent 中的格框架变量.

下列算法概括了利用情节知识库展开一个抽象动作的关键步骤.

算法 1(展开抽象动作).

- 1) 给定一个 GF2 格框架 c, 设该格框架代表一个抽象动作. 从剧本库中寻找一个剧本 Agent (或情节 Agent) ag, ag 的名字正好等于格框架 c 的根. 如果找不到这样的 Agent, 则算法失败.
- 2) 若 ag 有子 Agent, 找出那些满足 c 的需求(由该格框架中的诸格的值确定)以及 CON 中的上下文条件(CON 是一个数据结构, 其中存放着该抽象动作展开过程中积累的所有信息)的子 Agent. 任选其中的一个, 仍称之为 ag.
- 3) 重复执行步骤 2) 以选择辈分更小的子 Agent, 直至再无满足条件的子 Agent 可选. 把选子 Agent 过程中产生的所有新信息存入 CON 中.

- 4) 把原来的格框架语句 c 看成是一个虚拟的剧本调用语句 $\text{call}(\text{ag}, \dots)$. 应用算法 3 于 ag , 并利用 c 中所含的信息来例化 ag 中的变量.
- 5) 运用算法 2 使 ag 继承它的父辈的所有信息.
- 6) 如果 ag 是一个剧本 agent, 则
 - Until ag 不再含 call 语句,
 - DO 根据算法 3 执行一个 call 语句,
 - OD.
- 7) While ag 中有一个语句 s , s 中的变量可用 CON 及 ag 的说明部分所含信息例化
 - DO 例化这些变量,
 - OD.
- 8) 如果仍有语句含未被例化的变量, 则算法失败.
- 9) 否则, 算法成功结束.

算法 2 (继承).

- 1) 设 Agent ah 准备继承它的所有父辈节点的所有属性及其值. 构造 ah 的一个拷贝, 称为 ag .
2. 假设 ag 共有 n 个父辈. 把它们依次排列为 $\text{ag}_1, \text{ag}_2, \dots, \text{ag}_n$, 其中 ag_1 辈分最高, 是顶 Agent. 对每个 i , ag_i 是 ag_{i+1} 的父节点.
- 3) 对 $i=2, \dots, n$, 依次令 ag_i 继承 ag_{i-1} , 并令 ag 继承 ag_n . 每一次继承都包括下面的第 4) 步至第 11) 步. 其中统一以 ag 称呼继承者.
- 4) 若 ag 有一个背景, 则它的父节点的背景不被继承.
- 5) 否则, 若它的父节点有背景, 则 ag 继承此背景.
- 6) 令 ag 继承它的父节点的所有角色和对象.
- 7) 构造父节点的剧情部分的一个拷贝 ca . 对于 ag 的每个 replace 语句 $\text{Replace} \langle \text{父辈 Agent} \rangle (\langle \text{标号 1} \rangle) \text{ With } (\langle \text{标号 2} \rangle)$, 执行下列动作: 把 ca 中以 $\langle \text{标号 1} \rangle$ 为标号的框架变量或框架语句换成 ag 中以 $\langle \text{标号 2} \rangle$ 为标号的框架语句.
- 8) 从 ag 中除去所有的 replace 语句即用于置换的语句.
- 9) 对 ag 的每个 Enrich 语句, 用它提供的信息填入 ca 中有待充实的语句.
- 10) 从 ag 中除去所有 Enrich 语句.
- 11) 令 ag 继承 ca 的所有语句, 继承时应让所有被继承的语句之前加上父 Agent 的名字以示区别.

算法 3 (Call 语句的例化).

- 1) 以 s 表示 call 语句 $\text{call}(\text{hg}; x_1, x_2, \dots, x_n)$, 其中 hg 是另一个 agent 的名字, $x_1, x_2, \dots, x_n$ 是一些用于情节具体化的参数. 此外假定 s 属于 agent cag .
- 2) 如果 hg 有满足故事情节及 CON (见算法 1) 中信息的子 agent, 则根据算法 1 从中选择一个, 仍然称之为 hg , 并把所做的选择决策存放在 CON 中.
- 3) 对于新得的 agent 再执行上面的第 2) 步来选择新的子 agent, 在选择时注意和 CON 中的信息保持一致并及时把新的选择信息存入 CON 中. 如此反复执行, 直至无新的子 agent 可被选择.
- 4) 如果 s 不被其他 call 语句所调用, 则建立一个空的上下文关系集 CTX, 把 s 及 s 的所有父辈节点, hg 及 hg 的所有父辈节点均存入 CTX 中. 如果 s 被其他某个调用语句所调

用,则检查 hg 是否已在 CTX 中. 如果是,则表示出现了递归调用,算法失败;否则,把 hg 及 hg 的所有父辈节点放入 CTX 中.

5) 构造 hg 的一个拷贝,并称之为 chg1.

6) 反复执行算法 2 的第 1)至第 6)步,使 chg1 继承它的父辈的所有角色、对象和背景.

7) 利用诸参数 x_i , 使 chg1 中的情节具体化.

8) 我们称上述步骤 6)和 7)为值的例化. 执行下列步骤以传播在值例化过程中获得的信息:

8.1) 把它传播到 chg1 和它的父辈 agent 的所有语句中去;

8.2) 把它传播到 cag 的说明部分和语句中去.

9) 展开 chg1 中所有的 conditional 语句、case 语句和 bloc 语句.

10) 令 chg1 根据算法 2 的第 7)至 11)步继承它的父辈们的所有语句,得到 chg1 的一个新版本,并仍称之为 chg1.

11) 对 chg1 的每个尚未例化的属性名赋予一个缺省值. 称结果为 chg2.

12) chg2 取代 cag 中的调用语句 s(它现在已不含变量).

13) 对本算法开始时 s 和 cag-s 共享的所有变量 v_i , 令它们在 cag-s 中的所有出现都取它们在 s(现在是 chg2)中的相应出现(经过上述步骤后)所具有的值.

为了简化问题和不过多地陷入细节,我们只选上述故事中的最后一个语句展开.

用“白雪公主向黑炭王子祝贺生日”调用剧本 Agent “birthday congratulation”, 执行上述算法后,得如下结果

Role: Woman (Snowwhite),

Man (Coaldark),

Clerk (Any clerk)

Object: Car(Any car),

Gift ({Book, Whiskey})

Presentation: 1) Background: street_1

Solo. moving (Snowwhite)

Enter (Snowwhite, house, Type(house)=shop)

2) Background: shop_intern

Double. facing(Snowwhite, Any clerk)

Give (Any clerk, Snowwhite, {Book, Whiskey})

Give (Snowwhite, Any clerk, Money)

3) Background: street_3

Solo. moving (Snowwhite)

Drive(Snowwhite, Any car, building_1, building_2)

4) Background: street_2

Solo. moving (Snowwhite)

Enter (Snowwhite, house, Type (house)=home)

5) Background: none

Double. facing (Snowwhite, Coaldark)

Give (Snowwhite, Coaldark, {Whiskey, Book})

6) Background: none

Double.facing (Snowwhite, Coaldark)

Speak(Snowwhite, Coaldark, smiling, "happy birthday to you")

Speak(Coaldark, Snowwhite, smiling, "thank you")

由于故事情节中出现模糊描述等原因,定性规划过程中也有部分定量规划问题.有关的工作见文献[45,46].

5.2 场戏分割

上面得到的场戏序列是一个部分分割了的连续动作流.分割的边界是各剧本 agent 和情节 agent 调用展开的开始和结束处,其中有些情节 agent 是带背景的.但是这个分割不一定是完备的,某些情节可能还没有背景,某些分割可能还要细化,也有些分割可能需要合并.为了生成合理的分场戏剧本,需要执行场戏分割,得到一幕一幕的场戏.

把动作流分割成场戏要做两件事:一是确定哪些动作应该组合在一起;二是对组合后的动作集团(场戏)确定其局部的角色、物体和背景.分割动作通过两个阶段完成,在情节规划时作初步分割,到导演规划时还要作进一步分割.

我们把分割的结果称为场戏,而不是分镜头.因为同一个场戏中可能会有多种不同的镜头类型.原则上,一个场戏中发生的事件一定是在同一个地点(运动镜头除外)且在时间上是连续的,但反过来不一定(《天鹅》在动画编辑中也采用蒙太奇手法).划分场戏是根据一系列启发式的原则进行的,其中包括背景原则(不同背景的情节不在同一幕中)、时间原则(不同时间的情节不在同一幕中)、身份原则(不同身份的人物有不同的典型环境)等.凡是框架语句中确定的动作、物体和角色都作为分割时的参考信息,称为关键词.通过场戏分割获得的最主要信息是故事的场面(背景).我们奉行的是“极大外延”原则,这有点像物理中的结晶过程,一般来说,对应于原来抽象动作的那些具体动作起结晶中心的作用,其他动作各自朝结晶中心靠拢、凝聚.

算法 4.

1) 利用剧本 Agent 和情节 Agent 把所有的抽象动作变成具体动作.

2) 对所有从抽象动作转变而来的具体动作做如下操作:

2.1) 每个情节 Agent 生成的所有具体动作构成一个动作组,采用该情节 Agent 规定的背景和基本场景;

2.2) 把使用同一背景,同一基本场景,角色相同,又无明显时间差别的相邻动作组合并成更大的动作组.

3) 对所有不是从抽象动作转变而来的具体动作作如下操作:

3.1) 如果动作语句中出现关键词 K,则根据 K 对该具体动作确定一组可能的背景;

3.2) 如果动作语句中不出现关键词,则该动作背景暂缺;

3.3) 每个动作自成一个动作组.

4) 根据如下原则合并动作组:

4.1) 无背景的动作组如果和它的上一动作组或下一动作组有相同角色,又无明显时间差别,则可以并入;

4.2) 相同背景、相同基本场景,且角色相同的动作组可以合并;

4.3) 背景集之交非空且角色和基本场景相同的动作组可以合并,并以此交集为合并后动作组的背景;

4.4) 不同背景的动作组不能合并;

4.5) 虽在同一背景,但时间上不连续的动作组不能合并;

4.6) 如果动作组 A 和 C 之间有动作组 B,且 B 的背景与 A 或 C 的背景不同,则 A 和 C 视为时间上不连续.

上面例子中动作流的初步分割中,第 5)和第 6)个场戏没有背景.利用上述算法第 3)步,把“王子”和“公主”看成关键词,则知识库的启发式规则提示用 palace_room 作为缺省背景.此时第 5)和第 6)的场景的背景、角色和人物布局完全一样,可以把它们合并,得到新的场戏:

5) Background: palace_room

Double. facing (Snowwhite, Coaldark)

Give (Snowwhite, Coaldark, {Whiskey, Book})

Speak (Snowwhite, Coaldark, smiling, “happy birthday to you”)

Speak (Coaldark, Snowwhite, smiling, “thank you”)

至此,我们的小故事的场戏分割任务已经完成.全剧共分 5 个场戏,场戏也称为幕.

5.3 场景规划

场景规划的任务是确定人物的出场分布.由于不存在确定的规则,只能按基于知识的启发式原理来设计.这要分两步进行:第 1 步确定人物分布的基本场景;第 2 步根据基本场景确定人物的定性空间分布.基本场景是从那些出现在电影或计算机动画中的基本场景总结出来的,它的核心内容包括主要角色的数量、这些角色的空间分布以及他们彼此之间的行为动作等.我们在上面引进情节 Agent 时已经提到过基本场景,但未定义.基本场景分为两层:定量层(主要角色和次要角色数量,按模糊及相对原则处理)和定性层(角色的空间分布和彼此之间的基本行为).定量层确定(一个或多个)注意中心;定性层指出各注意中心的相对位置及其变动.

基本场景的定量层有下列八种模式

1) Solo: 只有一个角色;

2) Double: 两个角色并发操作;

3) Few: 多于两个角色,但不是太多(比如说少于 10 个);

4) Few_Times_Few: 几小组人,每组为 Solo, Double 或 Few;

5) Many: 一大群人;

6) Few_Plus_Many: 一小组人加一大组人;

7) Few_Times_Many: (不多的)几个大组;

8) Many_Times_Few: 许多小组.

对于每一种定量层,都有几种定性层来描述人物的静态分布和动态行为.例如,与定量层 solo 相对应的定性层有三种: no_action, action 和 moving, 分别表示该角色是无动作、原地动作还是移动.而与定量层 double 相对应的定性层则有 facing (两人面对面)、siding (两人并排)、following (一先一后)、collecting (两人相聚)、separating (两人分开) 等等.例如, double. facing (张三, 李四) 是两人面对面的基本场景(可以代表谈话、售票、打架等两人面对面的各种动作).

场戏的人物分布一旦被确定为上述八种模式之一,就可以根据一些启发式的原则确定人物的定性空间分布.考虑到客观情况的多样性和电影艺术的特点,我们没有试图去穷尽这些原则,而是采取系统建议,用户调整的方法.下面是人物出场规划算法的要点.

1) 确定基本场景:

(i) 凡是已有基本场景的场戏均采用此基本场景;

(ii) 对其他场戏根据其中人物情况用启发式规则推导其基本场景.

2) 区分有障碍(其他三维空间物体)场景和无障碍场景.

3) 对于无障碍场景,区分正面、侧(斜)面和横面三种基本的人物出场模式.

4) 对于每一种基本场景和每一种人物出场模式,给出一种或多种人物定性分布规划.

5) 人物定性分布说明有两类:相对于画面(左、右、远、近)的说明和相对于其他人物或物体(在...之前、之后、之左、之右)的说明.

6) 系统建议出一种定性分布规划后,用户如不满意可做交互式调整.

7) 有障碍场景的人物出场规划主要采取 5) 中相对于障碍物体的规划.

用这个人物出场规划算法对上述故事的第五幕作规划. 由于在前面的步骤中已经从情节知识库中获取了该幕的基本场景 Double. facing, 因此可以直接从此基本场景推导人物空间定性布局. 根据启发式算法,如果是按照无障碍场景来作规划,且采取正面的人物出场模式,则推导结果为一个人物(选定黑炭王子)在场景中间偏左,另一个人物(白雪公主)在场景中间偏右. 两个人面对面站着. 人物出场分布说明通过 ADL 语言(见下面 5.5 节)的 Appear 类语句给出. 在我们的例子中是

```
appear (Coaldark, middle_left, facing_right, t1),
```

```
appear (Snowwhite, in_front_of (Coaldark), facing (Coaldark), t1).
```

这两个语句要作为情节语句的一部分加到 ADL 剧本中去.

5.4 人物外观规划

人物规划是导演创作中十分重要的任务,它的重要性有时甚至超过了情节规划,并且情节规划往往还是为人物规划服务的,因此在《天鹅》的总体设计中,人物规划也包含部分情节规划,目的是以人物的典型行为来刻画人物的典型性格. 但由于这个问题难度很大,在目前的版本中尚未实现.《天鹅》现在的人物规划只做了人物外观规划. 对于故事中的每个角色,以及通过情节规划生成的每个角色,《天鹅》通过如下步骤进行人物塑造:

1) 给定故事中的角色 A,从角色库中选出一个极大原型 R,使得从身体部件的意义上说 A 完全包含 R;极大的含义是库中不存在这样的原型 T,使得(从身体部件的意义上)T 包含 R, A 包含 T, T 不等于 R;

2) 如果 A 的身体部件多于 R 的身体部件,则增添 R 的身体部件,直至它们有相同的身体部件,称新的原型为 R';

3) 如果 A 的某个身体部件与 R 的对应身体部件有很大差别,不是通过简单的调整或变换可以使它们相同的,则更换 R 的对应身体部件,称新的原型为 R'';

4) 如果根据剧情 A 有(不属于身体部件的)额外的装备,则把同样的装备加在 R'' 的身上,我们称这样的额外装备为物理部件,所得的新原型称为 R''';

5) 如果 R''' 的(某个或某些)身体部件与/或物理部件的绝对尺寸或相对尺寸与所要刻画的角色 A 不符,则放大或缩小比例尺以调整之,称所得的新原型为 R'''';

6) 如果 R'''' 的(某个或某些)服装穿着或颜色与所要刻画的角色 A 不符,则使用着色法或粘贴服装法来修改 R'''' 的外表,称所得的新原型为 R'''''.

人物塑造的过程是在动画程序加工过程中自动进行的. 但是各种人物的塑造规划是在建立动画角色库时就设计好了的. 角色库中存放的原型是比例正常且未加修饰的原始角色.

原始角色可以不完整,例如角色的各部分都可以缺少,不仅是手、腿等可以缺少,连头都可以缺少(用于表示神话人物).另一方面,故事人物的身体部件也可以多于库中角色的身体部件,就像许多人格化的动物在动画片中被表示为人的形状后拖一条尾巴那样.

对应身体部件差别很大的例子在童话故事中极为常见,例如人的脑袋和狮子的脑袋.没有简单的变换可以把人脑袋换成狮子脑袋,只能更换.额外的装备是指帽子、手杖,以及手里拿的枪等等在当前幕的剧情演变过程中不会与角色的身体分开的物品.通常把它们做成角色关节的一部分.

对于角色外形的修改包括塑造高、矮、胖、瘦等外型特点,这里通常是指偏离标准尺寸的调整,还包括对穿着衣服的颜色和色样的调整.

在这个故事的第五幕中需要塑造王子和公主两个形象.《天鹅》从库中取出一个原型 boy. 增加 boy 的高度,并使他头戴金冠,身穿高贵的服装,成为王子的形象.又取出另一个原型 girl. 瘦小 girl 的身材,使她穿纯白服装,成为公主的形象.此外,两个人的脸都应该是带微笑的.至此,人物外观规划便完成了.规划成果体现在下节将要描述的 ADL 语言的角色说明部分.

5.5 情节定性描述语言

下一步的任务是用一种定性的故事描述语言把上述情节规范化地描述一遍,并在描述中体现场景规划和人物规划的结果.《天鹅》的动画定性描述语言称为 ADL. ADL 以一个场戏为一个段落(程序单位).它的设计原则主要是使用定性的、相对的方式来描述动画角色和它们在各个背景中的动作及摄像机的拍摄,我们在空间上引入定性的相对位置和方向,在时间上引入只代表先后次序的定性时间点序列,角色及其动作分别采用宏结构和宏动作来描述,角色的具体关节结构和组成元件,以及动作的具体关节运动细节等都不必出现. ADL 既可以作为计算机自动生成定性动画剧本的目标语言,又可以作为编程语言供动画导演编写分场景剧本之用.

一个完整的 ADL 剧本的语法结构如下:

```

<ADL 程序> ::= # ADL(<程序名>)
                # Environment: <背景说明>;
                # Role: <角色说明>;
                # Object: <对象说明>;
                # <场戏说明序列>
                # End of ADL
<场戏说明> ::= # Scene (<场戏名>)
                <背景选择>,
                <角色选择>,
                <对象选择>,
                <摄像机规划>,
                <情节说明>
                # End of 场戏名
<摄像机规划> ::= # Camera
                <基本摄像原语序列>
                # End of Camera
<情节说明> ::= # Plot

```


〈ADL 语句序列〉

End of Plot

〈基本摄像原语〉 ::= 〈基本摄像原语类型〉 (〈基本摄像原语参数序列〉),

〈ADL 语句〉 ::= 〈宏动作名〉 (〈动作主体〉, 〈动作对象〉, 〈格参数序列〉, 〈时间参数序列〉).

在这一节中我们主要谈情节说明. 摄像机规划此时还没有生成. 上述小故事的 ADL 情节描述见表 2. ADL 语句中的第 1 个 t 表示动作开始时间, 第 2 个 t 表示动作结束时间. t 的下标表示时间的先后. 时间描述在这个阶段还是定性的.

在 ADL 中, 一个背景由两部分组成: 固定部分和可变部分. 前者包含一个典型环境应该有的一般物体, 例如花园中的树木; 后者包括故事情节要求, 另加的物体, 例如情节“公主在花园里玩球, 球掉到井里去了”, 要求在花园背景里加一口井. 我们在这里只考虑第 5 幕的背景, 它对王宫内室没有特殊要求, 利用库中的典型背景就可以了.

至此, 我们已完成了上述小故事中第 5 幕的除摄像机规划之外的所有定性规划, 所得的结果综合在表 2 中.

表 2 动画定性描述语言 ADL 剧本示例

script(Snowwhite)
character:crown
type; crown
end-of-char
character:Coaldark
type; boy-1
property: figure(tall);
color(yellow,skirt-art);
face (smiling);
decorate: crown(on-top-of(head-art));
end-of-char
character:Snowwhite
type; girl-1
property: figure(small);
color(white,skirt-art);
face (smiling);
end-of-char
scene;1
background: palace-room
color;blue
static;bedroom
object:big-mirror
type;mirror
motion;fixed(on(back-wall(bedroom)));
end-of-obj
end-of-back
camera:
missing for the moment #
end-of-camera
plot :
appear(Coaldark, middle-left, facing-right,t ₁),
appear(Snowwhite, in-front-of (Coaldark), facing (Coaldark), t ₁),
give (Snowwhite, Coaldark, {Whiskey, Book}, t ₂ , t ₃),
Speak (Snowwhite, Coaldark,smiling, “happy birthday to you”,t ₄ ,t ₅),
Speak (Coaldark, Snowwhite,smiling, “thank you”,t ₆ ,t ₇)
end-of-plot
end-of-scene
end-of-script

6 导演规划

一部电影片子的成功在很大程度上决定于导演. 动画片也不例外. 在初步情节规划成功以后,《天鹅》还要经历一个导演规划的阶段. 这个阶段的任务主要是

- 1) 塑造主要和次要角色;
- 2) 规划摄像机的动作;
- 3) 照明设计;
- 4) 色彩设计;
- 5) 音乐设计;

在目前的《天鹅》版本中,只实现了上述导演任务的前两项. 并且人物规划还是简化了的. 在未来的新版本中,我们将要设计和实现一个导演规划语言,用以描述导演规划需求说明,并在导演规划需求的指导下推导出上述五项具体规划. 在本文中,我们只简释人物外观规划,并直接从摄影需求说明开始介绍摄影规划.

7 定性摄影规划

7.1 规划流程概述

动画也是电影. 电影不同于舞台剧的特点之一是它把观众看剧的三维空间(三维物理空间中发生的故事)变成了四维空间(调动摄影技术让观众在预设的时间、观察点、距离、角度、镜头变换、变换速度等环境下观察事件和人物).《天鹅》中完成这个任务的模块叫摄影规划. 摄影规划也是定性规划的一部分,因其技术复杂而在此单列一节. 一般摄影技术的要点可以概括为光线处理、色彩处理、构图处理和运动处理四大部分.《天鹅》的摄影规划初步实现了运动处理和构图处理这两部分,具体步骤如下:

- 1) 生成一系列摄影规划要求原语;
- 2) 将摄影规划要求原语转换成群组摄影原语;
- 3) 将群组摄影原语翻译成高级摄影原语;
- 4) 将高级摄影原语翻译成基本摄影原语;
- 5) 将基本摄影原语编译成动画定量描述中量化的摄影语句.

摄影规划要求原语、群组摄影原语、高级摄影原语、基本摄影原语和量化的摄影语句的集合构成了摄影规划表示的层次模型. 摄影规划是一个非常复杂的过程. 下面我们把它和软件工程作一对比:

摄影规划要求原语→软件需求说明

群组摄影原语→软件系统设计

高级摄影原语→软件模块设计

基本摄影原语→高级语言编程

量化的摄影语句→汇编语言编程.

所不同的是,包括摄影规划要求原语在内的每一层都是自动生成的. 其中基本摄影原语编入 ADL 这一级.

7.2 摄影规划要求原语¹⁾

摄影规划要求原语建立在前面定义过的基本场景的基础上. 它的语法是:

〈摄影规划要求原语〉 ::= 〈典型场景〉 | 〈摄影要求原语〉 ; 〈典型场景〉

〈典型场景〉 ::= 〈基本场景〉 | 〈典型场景〉 , 〈附加因素〉

由此可见, 现在所缺的只是〈附加因素〉的定义. 由于基本场景不能完全概括现实生活中各种可能的情况, 不能只依靠它来制定导演规划, 因此必须有一些附加因素. 附加因素给出定量层和定性层都不包含的信息, 而包括环境气氛、人物特征和情感, 以及导演意图等, 用来结合基本场景以组成典型场景. 下面是〈附加因素〉的简化语法:

〈附加因素〉 ::= 〈附加因素类型〉 = 〈附加因素值〉

〈附加因素类型〉 ::= 〈环境附加因素〉 | 〈角色附加因素〉 | 〈角色关系附加因素〉 |

〈导演态度附加因素〉 | 〈导演意图附加因素〉

这里〈环境附加因素〉指的是与人无关的自然因素, 如天气、时间、地点、背景、气氛, 等等; 〈角色附加因素〉指的是诸角色的各种特征, 包括年龄、性别、外形、性格、情感、社会地位, 等等; 〈角色关系附加因素〉指的是角色之间的相互关系, 包括亲戚、朋友、仇敌、主仆, 等等; 〈导演态度附加因素〉是指导演对动画中的人物的态度, 包括同情、批判、赞扬、讥讽, 等等; 〈导演意图附加因素〉是指导演要用什么样的艺术手法来创造这部电影, 包括动画片的节奏、风格、镜头运用、色彩运用、采光, 等等.

摄影要求原语的确定由以下步骤组成: 对每一个故事情节, 首先确定基本场景, 再确定典型场景, 最后才确定附加因素. 一般来说, 组成典型场景的附加成分与基本场景有关, 这就是说不是每个附加因素都可以加到任何典型场景中去. 但不同的基本场景可能有相似的附加因素.

附加因素的规划也是自动进行的. 经过规划, 得到上述故事第 5 幕的摄影要求原语:

Double. facing (Coaldark, Snowwhite) [t_1, t_7],

place = inside room,

role's sentiment (Coaldark, Snowwhite) = happy,

role's attitude (Coaldark, Snowwhite) = friendly,

spatial focus = gift,

atmosphere = warm.

7.3 群组摄影原语

摄影要求原语与群组摄影原语的区别在于: 前者既独立于摄像机 (即不含任何与摄像机有关的术语) 又独立于观众, 它只指明故事自身显示的场景特征, 而不包含任何观众欣赏要求; 后者则独立于摄像机而不独立于观众, 它指明应该让观众看到什么, 但不指明应采用何种摄影技术. 群组摄影原语的简要语法如下:

〈群组摄影原语〉 := 〈群组摄影原语名〉 [(〈群组摄影原语参数序列〉)]

〈群组摄影原语参数〉 := 〈群体角色〉 | 〈个人角色〉

〈群体角色〉 ::= group[〈数字〉]

〈个人角色〉 ::= target[〈数字〉]

1) 由于摄影规划要求原语中所含的信息对光线、音乐、色彩等规划均有指导作用, 在我们设计的新版本中, 已经改称导演要求原语.

每个群组摄影原语表示一种观赏要求. 其内容包括介绍人物、突出人物、突出动作、突出气氛、突出环境、突出多个视觉中心, 等等. 以介绍人物为例, 又分介绍单人、介绍双人、介绍多人、介绍群体等. 下面只列出介绍静态双人的几个群组摄影原语:

```
side_by_side (target1, target2)      # 观众看到两个目标并排 #
face_to_face (target1, target2)      # 观众看到两个目标面对面 #
one_after_another(target1, target2)  # 观众看到两个目标一前一后 #
```

在我们的故事的第 5 幕中也有两个人(黑炭王子和白雪公主), 它们也是面对面. 但不能简单地引用上面列出的第 2 个群组摄影原语, 因为中间还有一个送礼的动作. 上面的摄影规划要求原语经自动翻译转换成相应的群组摄影原语:

```
object_passing(Snowwhite, Coaldark, {whiskey, book}) # 观众看到白雪公主送礼 #
static_focus_of_attention(whiskey, book)             # 观众看到礼物 #
face_to_face(Snowwhite, Coaldark)                   # 观众看到他们面对面说话 #
one_expression(Snowwhite)                            # 观众看到白雪公主的笑容 #
one_expression(Coaldark)                             # 观众看到黑炭王子的笑容 #
```

7.4 高级摄影原语

高级摄影原语指明应采用何种摄影技术才能实现群组摄影原语的规定, 是可以独立实现某种观众欣赏要求的摄影基本套路, 主要通过摄影专家历年积累的经验总结出来. 例如, 外反拍和内反拍可以用来拍摄二人对话; 直角拍一般用来拍摄站成直角的两个人; 平行拍能介绍多个注意中心; 同轴拍可以强调某种渐进效果等等. 基本语法为

〈高级摄影原语〉 ::= 〈高级摄影原语名〉(〈高级摄影原语参数序列〉)

〈高级摄影原语名〉 ::= right_angle(直角照) | 4_parallel(四角色平行照) | 3_coaxial(三角色同轴照) | in_contra(内反拍) | out_contra(外反拍) | over_shoulder_shot(过肩照) | own_fix(主观镜头) | ……

〈高级摄影原语参数〉 ::= 〈目标〉 | 〈镜头〉 | 〈方向〉 | 〈位置〉 | 〈角度〉 | 〈速度〉 | 〈时间〉

〈目标〉 ::= 〈目标名〉 | {〈目标名序列〉}

〈镜头〉 ::= BCU | CU | MCU | MS | MLS | LS | VLS | ELS

〈方向〉 ::= left | right | front | back | beneath | above | ……

〈位置〉 ::= left | right | middle

〈角度〉 ::= up | down | a little〈角度〉 | moderate〈角度〉 | largely〈角度〉

〈速度〉 ::= quickly | slowly | a little〈速度〉 | quite〈速度〉 | very〈速度〉

〈时间〉 ::= long | short | a little〈时间〉 | quite〈时间〉 | very〈时间〉

其中 BCU, CU, MCU, MS, MLS, LS, VLS, ELS 分别表示大特写、特写、近景、中近景、中景、全景、远景、大远景.

上述群组摄影原语可转换成下列高级摄影原语:

```
three_coaxial ({Snowwhite, Coaldark}, MLS, MS, MCU, front, slowly) # 三机同轴慢推进 #
down_normal ({whiskey, book}, MLS, slowly) # 中景慢俯拍 #
two_parallel ({Snowwhite, Coaldark}, MLS, front) # 两机平行照 #
```


in_contra (Coaldark, Snowwhite,MCU) #近景内反拍#

in_contra (Snowwhite, Coaldark,MCU) #近景内反拍#

7.5 基本摄影原语

基本摄影原语是在动画定性描述剧本中描述摄像机动作的基本单位,每个基本摄影原语对应摄像机的一次动作,如平移、推进、拉出、摇摄和固定镜头等.一般基本摄影原语的参数有被摄目标、景别、方向、俯仰角度、摇摄角度、速度参数、持续时间等.

经转换后,上述高级摄影原语中的多镜头操作可以分解为单镜头操作.例如,高级摄影原语 down_normal 将被转换为如下的基本摄影原语:

```
tilt(MLS(front({Sonwwhite,Coaldark})),MLS({whiskey,book}),down,slowly)
    #从白雪公主和黑炭王子中景缓慢下旋镜头至俯视威士忌和书中景#
track(MLS({whiskey,book}),CU({whiskey,book}),in,slowly)
    #从威士忌和书中景缓慢推进至威士忌和书特写#
```

基本摄影原语是定性动画描述语言 ADL 的组成部分.把这些语句插入表 2 中尚属空白的摄影规划部分,即得到一个完整的 ADL 程序.

8 动画定量规划

8.1 位置规划

ADL 程序中的人物和物体的定性位置需要转换为定量位置,其中包括出场位置和运动终结位置.规划的依据是人物、物体和环境的几何特征.位置规划的难点在于如何把静态位置计算和动态位置变化结合起来.例如,假定白雪公主出场时站在皇后面前,过若干时间后她要在皇后面前跪下来.如果只按站的要求计算白雪公主的位置,则她跪下时由于裙子铺开,有可能使裙子与皇后的裙子或脚相交,造成动画失真.

8.2 路径规划

8.2.1 二维空间路径规划

最理想的应该是首先进行路径的定性规划,然后再使之定量化.但是我们现在只做了定量的路径规划,并且只做了二维的定量路径规划.这是因为二维空间中的移动点运动规划问题的研究比较成熟,出现了许多多项式复杂度的算法.而三维问题比较困难,迄今没有令人满意的算法.经验表明,采用一些启发式的方法,二维算法可以解决大部分的三维问题.

在《天鹅》中,基本上采用了 Perez-Lozano 的可视图方法,并予以实现.实践证明,该方法已能满足现阶段动画自动生成的基本需要.Perez-Lozano 方法假定障碍物都取多边形的形式,否则可以用多边形把障碍物包起来.

定义 3(可见图).可见图是一个无向图 $VG(N,L)$,其中 $N=V\cup\{S,T\}$ 是节点集合, V 是所有障碍多边形的顶点集合, S 是出发点, T 是目标点. L 是 VG 的所有边 (a,b) 的集合, a,b 属于 N .所有 (a,b) 边与所有的多边形均不相交(不进入任何多边形之内).

Perez-Lozano 算法要点

1) 设要考虑半径为 r 的圆的移动问题.开始时,该圆的圆心位于 a 点上.目标是要把圆心移到 b 点去.

2) 使每个多边形的边都往外长出 r 距离(多边形变胖).把因此而相接的多边形看成一

个多边形.

- 3) 把新的多边形的所有顶点加上 a 和 b 成为一个新的顶点集. 连成一个可见图.
- 4) 把每条边的长度定义为该边的权.
- 5) 运用图搜索算法寻找 a 到 b 的最短路径.

8.2.2 三维空间路径规划

三维空间中的移动点运动规划问题非常困难,直到最近才有人证明了在包含一组相互分离凸多面体的三维空间中为任意两点寻找最短路径是个 NP 问题. 迄今为止,对于多面体三维空间中的规划问题,指数时间复杂度的算法是最好结果,不过当把问题进行简化时还是可以得到一些有效算法. 施海虎在研究这个问题的过程中,将障碍物包装成凸多面体,并假设工作空间中的障碍物的数目固定. 在研究两个凸多面体表面上移动点的最短路径问题时,得到了复杂度为 $O(n^2)$ 的结果.

在实际应用中,我们采取了三种办法来解决三维空间路径规划问题.

办法 1. 当障碍物形状离多边形柱体不太远时,用多边形柱体把它包裹起来. 这样,三维规划问题就化为二维规划问题了. 迄今为止在我们用《天鹅》生成的动画片中,用此方法已经够了.

办法 2. 当某些障碍物形状离多边形柱体较远时,可以把障碍物按水平距离分成段. 使得在每一段上可以大体把障碍物看成是一个多边形柱体. 此方法的一个典型应用是人在树林中行走的处理. 把树分成两段:树冠和树干. 一般树冠都高于人. 所以人的行走可以和树干一起规划.

办法 3. 当障碍物的分布不是很均匀时,可以把角色的活动范围分成若干区,每区单独规划. 例如,院内、院外,或楼上、楼下的规划均可按此办理.

8.3 定量摄影规划

在第 7 节中生成的定性摄影规划需要定量化. 由于对摄像机位置和动作的任何说明都是以具体的背景、人物和物体为参考坐标系的,因此只有当这些参考坐标系的定量位置都确定以后,才能够把摄像机的位置和动作计算出来. 又由于参考坐标系的可变性很大,有时计算的位置不是很精确,造成摄影效果不好. 为解决此类问题,《天鹅》有一个对摄影规划作模糊调整模块. 一旦发现摄影规划的效果不好,例如某个角色被其他角色挡住了,该模块即对当前的摄影规划作微调,以便找一个更好的位置;如微调不成,则重新规划.

8.4 动画定量描述语言

8.4.1 语言设计

动画定量规划的所有结果都通过动画定量描述语言 CAL 表示成程序形式. CAL 的简化语法如下(因该语言的程序含细节和数据太多,在本文中不再给出具体的 CAL 程序):

〈CAL 程序〉 ::= CAL (〈程序名〉) {〈程序体〉}_{1-n} End of CAL

〈程序体〉 ::= {〈图像说明〉; 〈物体说明〉; 〈场景说明〉;
 〈摄像机说明〉; 〈光照说明〉; 〈动作说明〉}

各说明部分语义如下:

图像说明——该动画的起始帧数号、结束帧数号、步长和总帧数;

物体说明——物体的初始位置、大小、颜色、反光性和接触关系等,物体本身用关节结构树描述,同时指出树中各叶节点对应的元件;

场景说明——是诸说明中最重要的,它指明运动物体的轨迹和速度,背景说明(二维背景)也在其内;

摄像机说明——摄像机的坐标位置、拍摄中心的坐标位置和广角度的变化,包括变化方向和速度,还有焦距;

光照说明——光源的位置、光照方向、光源运动速度及颜色等;

动作说明——角色说明部分引用到的关节动作所对应的各关节的转动、平移和比例变换的数值.

上述的说明都是定量的,既可以为常量也可以用关于时间的参数方程来表示.

8.4.2 语言实现

《天鹅》通过计算将动画定性描述转换为定量描述,ADL 定性描述剧本中的每个场景说明生成一个 CAL 定量描述剧本.该转换系统主要包括剧本分析、场景计算和定量剧本生成三部分,其中场景计算包括角色动作计算、摄像机动作计算、灯光变化计算、调度程序等.

我们选择了 TDITM公司的动画软件 EXPLORE 作为《天鹅》的底层动画支持系统,SCRIPT 语言是 EXPLORE 的文本式可读动画语言.

《天鹅》对 CAL 剧本进行编译,生成的目标程序(包括一个 SCRIPT 剧本和一系列运动轨迹文件)可以在 EXPLORE 中运行,可进行实时动画演示或经 EXPLORE 中的 RENDER 计算生成一帧一帧的画面后播放.

动画定性描述到定量描述转换,以及定量描述最后表现成动画中必不可少的一个重要功能就是各种动画素材库的支持,我们将在下节进行论述.

从中文自然语言、表达语义的可读性文本、动画定性描述、动画定量描述、直至底层动画支持系统剧本,《天鹅》逐层地进行自动转换,生成动画;同时这些层次的语言都是可读的,用户均可以进行人工修改.

9 动画知识库

全过程计算机辅助动画自动生成是基于知识的,我们建立了各种类型的知识库以支持这个全过程的每一步,例如用于自然语言单句理解的词汇库、语法库、语义库,用于导演功能的场景设计库、角色设计库、背景设计库,用于摄影规划的摄影要求原语库,等等.我们实现了知识库管理系统,用于对各个库的内容进行查询、增加、删除、修改等交互式操作.

这里我们主要介绍为动画定性描述到定量描述转换所建立的知识库,主要有元件库、结构库、角色库、动作库、背景库等,用于存放角色原型、静态背景和关节动作原型,它们的内容均为定量描述.

1) 元件库:元件为构成角色的基本不可分的三维物体,如人的脖子、门的把手等.

2) 结构库:结构表示角色的各个关节之间的关系,为一树结构,子关节随着父关节的运动而运动,系统通过用元件填充结构中的叶节点后得到角色.在这里,角色统一表示有生命对象和无生命对象.

3) 动作库:每个动作记录了角色关节的标准动作,每一个动作表示成一个关键状态序列,每个关键状态是指在某一时刻各个关节相对于父坐标系的平移、转动和比例变换.

4) 背景库:每个背景由一组三维物体按一定的位置和方向排列而成,背景库包含了童

话故事中常用到的一些典型背景,如皇宫、花园、森林等。

在动画设计和生成过程中,系统根据剧情需要从这些库中调出原型并对其进行加工和修饰,生成所需的角色、背景、动作等。

10 结束语

全过程计算机辅助动画自动生成是一个崭新的课题,目前国内外计算机辅助动画制作方面的工作很多,但就我们所知,像《天鹅》这样,从自然语言书写的故事开始,由计算机完成全过程动画创作的研究却还没有过,这项研究难度极大。《天鹅》生产的动画片《三兄弟》已经在中央电视台大风车节目中播放。使用《天鹅》,我们只花了两周的时间就完成了这部十分钟动画片的制作,如果使用传统的方法则至少要半年以上。

《天鹅》的技术内涵是人工智能加图形学加电影艺术,其中人工智能技术和电影艺术是这个系统的上层,图形学技术是这个系统的基础。目前,以人工智能技术为核心的《天鹅》原型系统已经实现,但存在的问题还很多。

从自然语言形式的故事自动生成动画,这毕竟是一个十分复杂的过程。我们是先有了一个初步的想法,然后在实践过程中逐步细化、深化和修正,其中有一些比较大的修正,有些局部模块的实现方法至今不很理想。本文描述的内容并非全部都是实现了的,其中有相当一部分是准备重新实现的新方案。

是否能够从自然语言故事自动生成动画,以及这种研究是否有意义,是一个有争议的问题。在此,我们谈谈自己的观点。我们觉得它的意义在于:第一,动画自动生成是计算机科学与技术,特别是人工智能技术的重要试验场,对它的研究有助于推动相关学科分支的发展;第二,计算机艺术的研究本身就具有它的重要性,特别是,我们认为在各种计算机艺术的研究中,计算机动画是最能带来实际效益的一支;第三,计算机动画自动生成完全有希望发展成为一种产业,并在很大程度上改变动画制造业的面貌,这种过程很可能是渐进的,采取在各个环节上逐步用计算机代替人工的办法。

有一种看法认为,动画是一种艺术,是创造性思维的产物,计算机不可能代替人来创造。我们也认为计算机不可能完全代替人来创造。但是,我们同时认为有艺术家和艺术匠人之分。事实上,我们在市场上买来的很多艺术品都是一些艺术匠人制造的。许多进口动画的质量很粗糙,无非也出自艺术匠人之手。但在孩子们那里依然很受欢迎。我们不奢求计算机成为艺术家,而只想把它培养成艺术匠人。这一点应该可以做到,至少是逐步做到。

计算机缺乏创造性思维的重要表现之一,是由计算机生成的动画很可能陷于千篇一律,这种危险确实是存在的。对此,我们采取的对策主要有两条。一条是提供动画制作者以介入动画生成过程的机会,用交互方式来丰富动画产品的多样化和体现制作者的创造性思维。为此,我们为动画自动生成的每个关键环节设计了高级编程语言,总数将近十种,它们是各中间产品的表示语言。制作者对中间产品不满意时,可以通过修改这些语言的程序来干预动画生成过程。另一条对策是建立丰富的知识库和灵活运用知识库的工具,使得有限的知识能够发挥更大的作用。

客观情况多变复杂,难于考虑周全,这也是认为不能用程序来实现动画生成自动化的重要理由。事实上,这种情况在人工智能中是正常的。例如,我们都知道没有一个机器翻译程序

能把任何文章都正确翻译出来. 解决的办法就是用启发式算法. 本文给出的启发式算法可能不会在任何情况下都给出满意的解, 遇到这种情况需要人工干预. 因此, 对于我们来说, “全自动生成”只是研究的目标, 是我们不断向之前进的目标. 至于在这个过程中推出的任何原型或产品, 都只是半自动加人工干预的. 目前的《天鹅》还不能一点不靠人工干预就自动生成艺术上达到电视台播放水平的动画片. 上面提到制造十分钟动画《三兄弟》用了两周时间, 就是因为反复用人工干预手段改进《天鹅》生成的动画的艺术质量.

最后, 我们还要说一下关于常识的应用问题. 一般说来, 全过程动画自动生成中的不同功能涉及不同领域的知识, 例如自然语言处理需用到语言学知识、动画情节设计需用到艺术创作知识. 但是, 我们认为常识是动画自动生成所需的最本质的知识, 常识知识处理是动画自动生成过程中每一步都无法回避的难题. 前面我们已经初步介绍了《天鹅》在自然语言处理、故事理解、故事情节展开和分场景剧本生成方面的应用常识的机理. 这还只是常识在动画生成中可能的应用的一小部分, 需要研究的问题还很多.

进一步的工作包括以下各方面.

研究以各种数学模型为基础的计算机画面模拟, 包括自然景观、社会生活、人体动作、人物表情、动植物世界及特种动画技术等, 并把这些技术集成起来, 建立一个功能强的综合计算机动画平台, 其核心是一个内容丰富的、包含上述计算机模拟功能的库.

进行角色、动作和场景等的快速大容量建库研究.

通过计算机科学工作者和影视艺术家的合作, 总结和归纳有关的影视艺术规律, 用以指导《天鹅》系统的改造.

扩充故事分析技术, 使之实用化, 并引入导演规划, 使故事分析技术与导演规划相衔接, 形成较为有效的故事-剧本改编技术.

改进和完善空间行动规划和关节规划, 引入时间、空间、关节三要素综合规划.

深入研究从自然语言到动画转换过程中的常识推理技术和理论, 改进并丰富相应的常识知识库; 分析现有各种逻辑系统和推理方法对这种常识知识库的可用性, 研究新的、有效的常识推理方法, 并把这些方法与故事情节理解系统和动画自动生成系统连接起来, 使常识知识库和常识推理投入实用, 提高计算机生成动画的自动化程度.

将《天鹅》技术应用到因特网上, 研究个人网上实时动画定做.

将《天鹅》技术应用到计算机游戏上, 研究游戏自动生成和个人游戏实时设计实现.

可以看出, 全过程计算机辅助动画自动生成技术不仅提出了一系列具有研究意义的理论问题, 而且在动画片、网络、游戏等领域均有很大的应用潜力.

致谢 感谢曾经参加或仍在我们课题组的耿清、陈红捷(他们参加了定量描述语言工作), 朱文虹、李怡真(他们参加了定性描述语言工作), 李小滨、徐越、周月皎、李媛、高波、刘冷宁、杨帆(他们参加了自然语言理解工作), 杨德杰、罗翊(他们参加了故事分析工作), 施海虎、金正皓、李璐、靳小龙、孟祥亮(他们参加了情节规划和路径规划工作), 杜海霞、孙成敏(他们参加了摄像机规划工作), 余力新、赵莹、陈照兵、万荣林、姬广峰、张雪生、杨萍、范璐、马莹昊、王宇杰、蔡经纬(他们参加了动画库和动画生成工作)等. 他们的工作是富有探索性和创造性的. 感谢韦梓楚教授在项目的设计和实现过程中参与的大量讨论和在讨论中提出的许多宝贵意见, 以及在参与课题组管理时付出的许多精力.

本文在有关段落列出了一些参考文献,但仍有不到之处.特别要指出的是,许多问题的解决方案都是经过集体讨论的.许多学生的毕业论文未能在此一一列出.

参 考 文 献

- 1 金小刚,鲍虎军,彭群生. 计算机动画技术综述. 软件学报,1997, 8(4):241~251
- 2 Ulysse. <http://www.info.unicaen.fr/~nugues/research.html>
- 3 Persona. Chapter on Persona, Bradshaw J(eds.). Software Agents, MIT Press, 1997
- 4 Nautilus. <http://www.aic.nrl.navy.mil/~severett/vr.html>
- 5 Webber B. Instructing Animated Agents: Viewing Language in Behavioral Terms, LNCS 1374, Berlin: Springer-Verlag, 1998
- 6 Cassell J, Pelachaud C, Badler N I, Steedman M. Animated conversation: rule-based generation for facial expression, gesture and spoken intonation for multiple conversational agents. In: Proc. of Siggraph'94, Orlando, USA, 1994
- 7 Levison L, Badler N I. How animated agents perform tasks: connecting planning and manipulation through object-specific reasoning. In: Proc. of AAAI Spring Symposium: Toward Physical Interaction and Manipulation, March, 1994
- 8 Pelachaud C, Prevost S. Sight and sound: generating facial expressions and spoken intonation from context. In: Proc. of the 2nd ESCA/AAAI/IEEE Workshop on Speech Synthesis, New York, USA, September, 1994
- 9 Badler N I *et al.* Simulating Humans: Computer Graphics, Animation and Control. London: Oxford University Press, 1993
- 10 Badler N I, Webber B, Kalita J, Esakov J. Animation from Instructions, in Badler N I, Barsky B A, Zeltzer D (eds.). Make Them Move, San Mateo, CA: Morgan-Kaufmann, 1991, 51~93
- 11 Amaya K, Bruderlin A, Calvert T. Emotion from motion. In: Proc. of Graphic Interface'96, Toronto, Canada, May, 1996
- 12 Bruderlin A, Soraku-gun S, Calvert T. Knowledge-driven, interactive animation of human running. In: Proc. of Graphic Interface'96, Toronto, Canada, May, 1996
- 13 Calvert T, Bruderlin A, Dill J, Schiphorst T, Welman C. Desktop animation of multiple human figures. *IEEE Computer Graphics and Applications*, 1993, 13(3):18~26
- 14 Philips C *et al.* Jack: a toolkit for manipulating articulated figures. In: Proc. of ACM/SIGGRAPH Symposium on User Interface Software, Banff, Canada, 1988, 221~229
- 15 Raibert M H *et al.* Animation of dynamic legged locomotion. *Computer Graphics*, 1991, 25(4):349~358
- 16 Vilhjalmsen H, Cassell J. BodyChat: autonomous communicative behaviors in avatars. In: Proc. of the 2nd Annual ACM International Conference on Autonomous Agents, Minneapolis, USA, 1998
- 17 Cassell J, Bickmore T, Campbell L, Chang K, Vilhjalmsen H, Yan H. Requirements for an Architecture for Embodied Conversational Characters. to appear in Computer Animation and Simulation'99, Vienna, Austria: Springer-Verlag
- 18 Cassell J. A Framework for Gesture Generation and Interpretation, Cipolla R, Pentland A(eds.). Computer Vision in Human-Machine Interaction, Cambridge University Press, 2000
- 19 Binsted K. Character Design for Soccer Commentary. arXiv:cmp-lg/9807012, July, 1998
- 20 Maiocchi R *et al.* Directing an animation scene with autonomous actors. *The Visual Computer*, 1990, 6(6):359~371
- 21 Sgouros N M. Dynamic generation, management and resolution of interactive plots. *Artificial Intelligence*, 1999, 107: 29~62
- 22 Noma T, Kai K, Nakamura J, Okada N. Translating from natural language story to computer animation. In: Proc. of SPICIS'92, 1992, 475~480
- 23 Takashima Y *et al.* Story driven animation. In: Proc. CHI+GI'87, 1987, 149~153

- 24 StoryBoard Quick Demo. <http://www.powerproduction.com/quick/example/example.html>
- 25 BoardMaster StoryBoard and Timing Software. <http://www.boardmastersoftware.com/>
- 26 Sturmman D. Interactive keyframe animation of 3-D articulated models. In: Proc. of Graphics Interface'86, Tutrial on Computer Animation, 1986
- 27 Girard M. Interactive design of 3-D computer animated legged animal motion. *IEEE Computer Graphics and Applications*, 1987, 7(6):39~51
- 28 Christianson D, Anderson S, He L, Weld D. Declarative camera control for automatic cinematography. In: Proc. of AAAI-96, 1996
- 29 He L, Cohen M, Salesin D. The virtual cinematographer; a paradigm for automatic real-time camera control and directing. In: Technical Report, August 1996
- 30 Nakamura J, Kaku T, Hyun K, Noma T, Yoshida S. Automatic background music generation based on actors' mood and motions. *The Journal of Visualization and Computer Animation*, 1994, 5:247~264
- 31 陆汝铃. 人工智能. 北京:科学出版社,上册 1988,下册 1996
- 32 LU Ruqian. New Approaches to Knowledge Acquisition. Singapore: World Scientific, 1994
- 33 Lu Ruqian, Zhang Songmao, Wei Zichu. Generate computer animation from natural language stories. In: Proc. of Pacific Asian Conference on Expert Systems, Los Angeles, USA, 1999
- 34 Zhang Songmao. Story parsing grammar with attributes. In: Proc. 10th Int'l Conf. of FLAIRS, Florida, USA, March, 1997
- 35 Zhang Songmao. Weak precedence story parsing grammar. *J. of Compute Science and Technology*, 1995, 10(1):53~64
- 36 Zhang Songmao. A heuristic approach for acquiring the sketch of story. In: Proc. of Pacific-Asian Conf. on Expert Systems, Huangshan, China, May, 1995
- 37 Zhang Songmao. Forest grammar(I). *Scientia Sinica, Series A*. 1994, 37(6):761~768
- 38 Zhang Songmao. Forest grammar(II). *Scientia Sinica, Series A*. 1994, 37(8):998~1008
- 39 Zhang Songmao. Story parsing grammar. *J. of Computer Science and Technology*, 1994, 9(3):215~228
- 40 Zhang Songmao. Story grammar—a failed approach or a promising technique? In: Proc. 14th Int'l Conf. on AI, KBS, Expert Systems and Natural Language, Paris, France, June, 1994
- 41 张松懋. 相关的弱优先故事分析文法. 计算机学报, 1994, 17(6):477~480
- 42 张松懋. 含无序产生式的故事分析文法的研究. 软件学报, 1994, 5(1):11~18
- 43 张松懋. 故事分析实验系统 SPARS. 见:中国人工智能学术会议 93 论文集, 1993, 254~259
- 44 张松懋. 故事情节表达与高维文法. 见:中国人工智能学术会议 92 论文集, 1992, 139~146
- 45 Lu Ruqian, Zhang Songmao, Shi Haihu, Yu Lixin. The quantification problem in animation generation. In: Proc. of ANZIS'95 (The 3rd Australian and New Zealand Conference on Intelligent Information Systems), 1995
- 46 Lu Ruqian, Zhang Songmao, Shi Haihu, Yu Lixin. Solving the quantification problem in animation generation. *Australian J. of Intelligent Information Processing Systems*, 1995, 2(4):35~45

陆汝铃 研究员、中科院院士. 研究兴趣包括知识科学、基于知识的计算机动画生成技术、基于知识的软件工程等.

张松懋 副研究员、博士. 研究兴趣包括自然语言故事理解、基于知识的计算机动画生成技术、智能数据挖掘等.