



中国科学院软件研究所

杰出青年人才发展专项计划年度进展报告书

(2012 年度)

姓名	马翠霞	课题名称	基于草图的可视媒体交互技术
资助金额	180 万	课题起止时间	2012.1-2015.12
资助类别	应用基础研究类	所在部门	人机交互技术与智能信息处理实验室

研究工作主要进展和阶段性成果（含论文、研究生培养等）

一、研究工作主要进展

和谐自然的交互是进行可视媒体操作与构建的必要手段。本年度的研究工作围绕可视媒体交互的理论、方法及技术等方面展开研究，主要包括面向视频内容的草图认知模型、基于草图的视频语义分析与描述方法以及基于草图的视频摘要及其交互技术等内容。主要取得如下进展：

1. 草图认知模型分析

给出了描述草图表征认知过程的认知模型，探讨了认知模型中各认知单元的作用及相互关系，分析了草图表征与交互的认知过程。草图的信息加工过程包括选择注意、工作记忆、长时记忆和决策系统四大模块（图 1），选择注意阶段用户抽取草图蕴含的多种信息，对信息进行理解、学习和记忆，并可为进一步的认知加工提供基础；工作记忆中包括草图识别单元，它接受结构编码中的信息和长时记忆中反馈的草图特征信息，输出结果进入决策系统，工作记忆中的草图识别结果可以进入长时记忆；不同草图特征的整合以及对草图特征自上而下的加工离不开长时记忆。由于用户的认知资源有限，当用户面对连续的复杂多变的视频信息时，草图可以帮助用户快速理解视频的信息，而且草图具有的勾画、审查、修订等方面的优势，支持用户的连续思维，可促使新线索的发现。

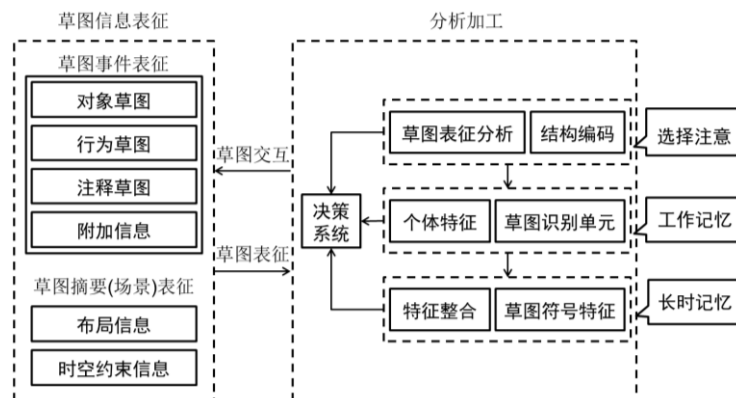


图 1 草图认知模型

2. 面向视频语义的草图描述与交互

(1) 草图注释与语义草图

提出了基于草图的视频注释方法和语义草图描述方法。利用草图描述视频的高层语义时将草图分为两类，一类是与人们的习惯认知普遍一致的、用于描述视频非有形类信息的特殊语义符号，如描述视频对象的声音、注释、运动轨迹、运动方式等以及运动的路径、方式、事件发生的环境；另一类是表现视频对象实体的形状类信息。这两类草图形式用来共同描述视频对象的多个信息侧面。为了直观有效地描述视频的动态或隐喻内容，草图语义符号的设计采用人们所普遍熟悉的具有共识的符号，用来描述时空、对象和行为等主要特征。用户可以通过添加草图注释来描述视频语义或者表达自己对视频内容的看法，也可以添加草图符号来标注视频中的场景或对象的某些特性，并进一步可以利用草图注释辅助用户对视频内容的索引和浏览（**错误!未找到引用源。&?`**）



图2 用户添加草图注释

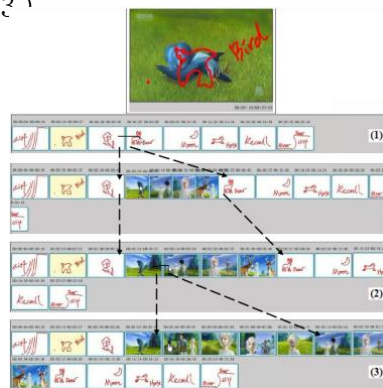


图3 基于注释和关键帧的视频浏览

(2) 基于场景结构图的可视媒体交互

从不同的层次上来给出对视频语义的描述，场景结构图主要包括两层：草图层和抽象图层。草图层以语义草图形式支持视频的编辑，使得用户从整体上把握视频的组织结构；抽象图以顶点和边构成的图结构来给出草图所代表的不同视频段之间的关联关系，并支持用户对超链接视频的构建。基于多视角的编辑环境提供给设计人员不同的辅助功能，支持对两类信息结构的创建、选择、删除等操作。基于当前操作状态，实现基于视频时空上下文和草图语义上下文的约束关系的主动捕捉和推荐机制（图4）。

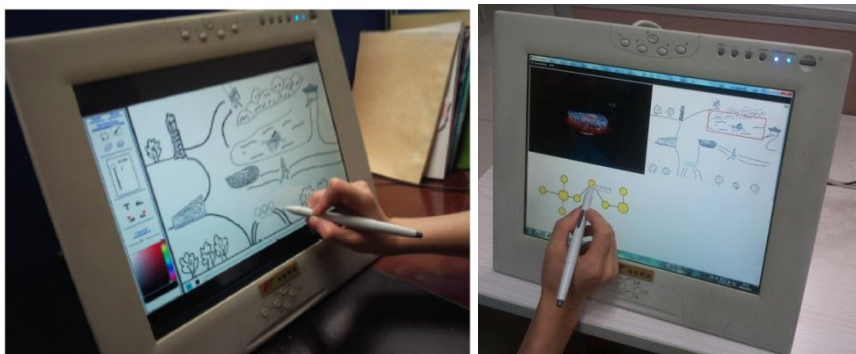


图4 基于场景结构图的视频编辑（左：草图层；右：基于草图层与抽象层的编辑）

(3) 多尺度视频表示与浏览

给出了一种多尺度的视频表示与浏览方法，包括增强时间轴、草图注释、

草图摘要以及紧凑的关键帧摘要等四种方式来提供不同尺度下视频的内容表达和浏览（图 5），有效克服了单一表达方式的局限，能够按照用户需求较好地表达从较粗尺度到较精细尺度的内容，帮助用户理解与定位视频内容。同时提供了与内容相关的聚类 and 分层方法，从而给用户不同层次内容的表达，更好地满足用户的不同浏览需要。

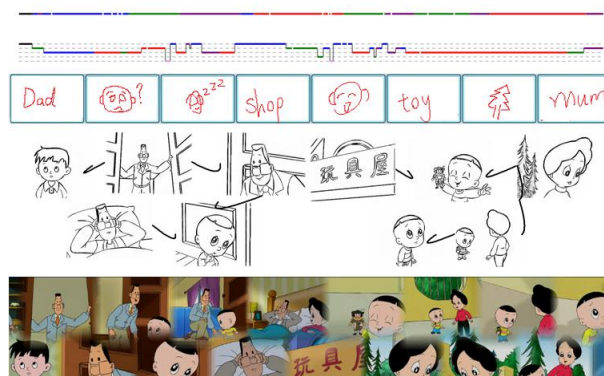


图 5 多尺度表示与浏览

3. 支持视频媒体高效展示和浏览的视频摘要生成与交互

(1) 螺旋摘要生成

提出了一种新的螺旋摘要形式，充分发挥了螺旋视图的连续性和空间优势，支持视频的高效浏览和定位，实现对视频的交互性编辑与操作。在抽取的关键帧以及相应的显著区域 (ROI) 基础上，根据不同的粒度将 ROIs 划分为不同的呈现层次，进而通过螺旋的平滑旋转实现根据用户的操作，意图呈现出不同层次不同粒度的视频信息展示。螺旋摘要这一新的视频内容呈现形式，也需引入与之相对应的高效交互方式。本课题与草图技术相结合，研究基于手势操作以及多指触摸交互模式的螺旋摘要高效交互方法（图 6），可以支持基于手势操作的整体缩放、螺旋的旋入与旋出、视频内以及视频间的超链接创建和直观的视频片断截取以及合并功能等。用户可以通过自由勾画的方式自然直观地表达交互意图，有利于建立足够独立性的交互资源，易被用户学习和记忆，降低了用户的认知负担和操作复杂度。

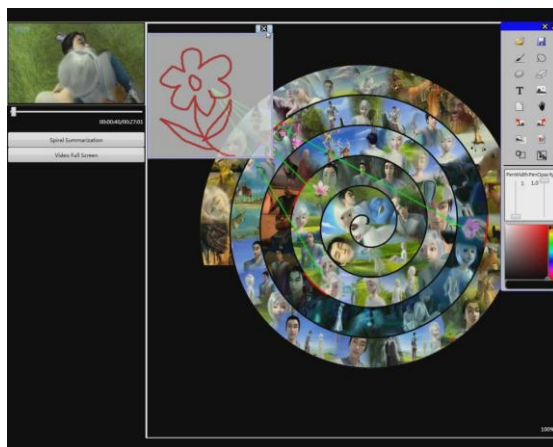


图 6 螺旋摘要

(2) 移动设备上自适应视频摘要的呈现与交互

移动设备体积小、易携带且交互性高，小屏幕带来的显示界面和交互

界面的冲突使得用户在使用功能元件对视频进行交互时，会降低视频观看效果。为了降低功能元件对视频播放的影响，使用触摸手势代替功能元件实现对视频的交互功能。另外，考虑到利用传统视频线性时间轴的定位方式，用户难以通过内容定位视频帧，在设计视频浏览方式的时候引入了视频摘要的概念，将基于视频摘要的非线性时间轴代替线性时间轴。并根据移动操作的具体状态自动调节其呈现方式，用户利用新的时间轴能够通过视频内容定位视频位置，交互直观高效（图7）。











图7 摘要时间轴

4. 基于草图手势的自然交互

草图是用户间进行交流的最具表现力和最直接的表达方式之一，符合人们长久以来的书写习惯，并且可以快速、准确表达用户意图，建立起人脑中的概念模型与计算机的可计算模型之间差异性的桥梁。用草图手势来支持对视频的操作以及对摘要的编辑高效自然，我们把草图手势按照其表现方式主要分为单笔手势、多笔手势以及多点触控手势。

单笔手势是指用户绘制完一笔手势后抬笔时完成手势命令。多笔手势是把多个单笔手势进行分组，组与组之间通过时间间隔长度来确定，两个单笔手势相隔较短就被分到一个手势组中，相隔较长则代表一个多笔手势结束。

表1 草图手势示例

单笔手势	手势名称	手势含义	单笔手势	手势名称	手势含义
	Clear-all	清空草图面板		V-move	上下移动
	Select	选定操作对象		H-move	左右移动
	Clear	删除当前对象		insert	插入
	Confirm	确认输入		Undo/redo	撤销/重做

二、阶段性成果

方法与技术

(1) 基于场景结构图的视频创作(video authoring)

我们结合空间、时间属性以及笔迹操作的实时属性等特征描述,用草图来进行视频注释以及描述视频内及视频间的语义和各类关系,采用草图形式的描述方法可以保留所表达内容的主要信息而去掉冗余信息,辅助用户高效获取视频的语义,通过场景结构图的方式实现面向视频媒体语义的两层视图的同步编辑技术。相关成果发表在国际期刊 *IEEE Transactions on Multimedia*。

(2) 面向视频内容的多尺度表示与浏览技术

利用增强时间轴、草图注释、草图摘要以及紧凑的关键帧摘要等四种方式来提供不同尺度下视频的内容表达和浏览;提供了与内容相关的聚类 and 分层方法,有效体现镜头之间的相关性,帮助用户更好理解视频内容,而且能够根据重要性的不同对关键帧分层,从而给用户不同层次内容的表达。相关成果将发表在中国科学 (*Science China Information Sciences*)。

发表/录用文章

1. **Ma Cui-Xia**, Liu Yong-Jin, Wang Hong-An, Teng Dong-Xing, Dai Guo-Zhong, Sketch-Based Annotation and Visualization in Video Authoring, *IEEE Transactions On Multimedia*, 14(4), pp 1153-1165,2012
2. Hongan Wang, **Cuixia Ma**. Interactive Multi-Scale Structures for Summarizing Video Content. Accepted, *SCIENCE CHINA*,2012
3. Yongjin Liu, Xi Luo, A. Joneja, **Cuixia Ma**, xiaolan Fu, User-adaptive sketch-based 3D CAD model retrieval, *IEEE Transactions On Automation Science and Engineering*, DOI: 10.1109/TASE.2012.2228481, Accepted, 2012
4. 詹启, 马翠霞, 倪美娟, 张艳秋, 王宏安, 戴国忠.一种基于草图注释的视频浏览技术.计算机辅助设计和图形学学报,Accepted,2012
5. Dongxing Teng, Haiyan Yang, Cuixia Ma, Hongan Wang.VDQAM: A Toolkit for Database Quality Evaluation based on Visual Morphology , *IEEE Conference on Visual Analytics Science and Technology*, 2012, pp 145-146, Seattle
6. 杨海燕, 滕东兴, 马翠霞, 戴国忠, 王宏安.故事地图: 一种基于草图界面的高效视频浏览方法.*International Journal of Software and Informatics*, 6(1), pp 89-105, 2012
7. 杜一,田丰,马翠霞,戴国忠,王宏安.基于多尺度描述方法的移动用户界面生成框架.计算机学报, Accepted,2012

发明专利

获授权：

一种基于草图的超视频构建方法 专利号：ZL201010119438.3

受理：

一种面向视频内容的螺旋摘要生成方法 专利号：201210218870.7

一种多尺度表示和浏览方法 专利号：201210557460.5

学生培养

目前指导在读学生3名；

2012年毕业硕士生一名；

在读一硕士生参加十七届全国计算机辅助设计与图形学学术会议（CAD/CG2012），论文被评为大会优秀学生论文。

开设春季博士生课程《sketch based User Interface》

国内外合作与学术交流：

- (1) 参加全国第十七届计算机辅助设计与图形学学术会议；
- (2) 参加Computational Visual Media conference(CVM 2012)国际会议；
- (3) 参加中法虚拟现实研讨会并作交流报告；
- (4) 参加2012杰出人机交互学者讲座；
- (5) 与英国Brunel大学的Dr. Shengfeng Qin (Senior Lecturer in Product Design, Head of Computer Aided Design Research Group)就草图用户界面方面开展合作研究；
- (6) 参与多个国际会议审稿工作；担任国际会议“International Conference on Pervasive Computing and Applications”和“国内会议和谐人机环境联合学术会议（HHME）”的程序委员会委员；
- (7) 中国计算机学会高级会员；担任ACM人机交互学会中国分会秘书长；负责中国自动化学会计算机图形和人机交互专委会的年报总结工作。

下一年度工作计划，包括国内外合作与交流计划

本课题下一步研究将紧密围绕项目计划进行，深入研究面向视频媒体的草图技术，主要包括：

1. 研究并改进草图认知模型，加强同心理学方面的合作，基于分布式认知理论深入分析草图表示和交互的认知机理。基于事件边界效应和视频数据的连续性特点，利用时间、空间以及对象运动等表征事件活动连续性的信息线索扩展传统的视频内容描述方法，获得更多的描述各个帧之间的关系和约束的信息，从而有助于用户更快更好地理解视频内容。

2. 研究并改进语义草图的描述方法，提高关键帧图像的草图化效果；研究和改进面向视频内容的草图摘要布局算法，加入审美约束、轨迹约束等信息，面向用户需求为主，生成个性化草图摘要并研究基于草图摘要的交互技术。

3. 研究移动设备上基于草图交互的视频共享技术，支持手机之间、手

机与电脑之间的面向视频内容的协同交互；研究多指触摸手势的设计与识别，实现视频注释、视频多尺度浏览以及视频搜索、视频编辑等功能。

国内外合作与交流计划

- (1) 加强同中科院心理所的合作，结合认知理论来研究草图表征与交互的认知过程，定期举行研讨。
- (2) 推进与清华的交叉基金的合作研究的进展，定期举行研讨。
- (3) 邀请英国 Brunel 大学 Senior lecturer Dr. Shengfeng Qin 来所交流。
- (4) 参加 ACM Multimedia 国际会议。

年度经费使用情况概要

该项目于 2012 年 4 月到款 46 万元。本年度共支出经费 43.35 万元。其中设备费支出 2.39 万元，用于科研过程中设备的购买；材料费支出 1.7 万元，用于研发过程中材料的购买；燃料动力费支出 8.33 万元，用于科研设备使用过程中所消耗的水、电等燃料动力费；差旅会议费支出 8.43 万元，用于项目过程中科研及参加会议所产生的差旅及会议费；出版文献等费用支出 6.57 万元，用于论文发表、资料购买、网络资源等；劳务费支出 15.93 万元，用于学生及临时聘用人员劳务费的发放。

存在的问题、建议及其他需要说明的情况

无

- 备注：1.杰出青年人才入选者应实事求是地填写此报告，禁止弄虚作假；
2.此报告作为专项计划跟踪、管理的主要依据，每年报送人力资源处；
3.此报告将在所网站对外发布。