

2

S-11

IMMPS: 研究现状与发展趋势*

Intelligent MultiMedia Presentation System: The Research and Development

吴朝晖 陈瑜 潘云鹤

TP391

(浙江大学人工智能研究所 杭州 310027)

摘要 智能多媒体表现系统是一个通过基于知识技术处理多媒体信息来生成智能化表现的新兴研究领域。本文对智能多媒体表现系统的定义、任务、计算模型以及研究现状做了深入论述;分析了当前主要表现系统对表现机制的不同理解和技术实现;并总结、预测了智能多媒体表现系统研究与发展趋势。

关键词 多媒体,表现,内容规划,媒体生成,编排,协同

IMMPS 信息管理

当前新兴的智能多媒体表现系统(Intelligent MultiMedia Presentation System, IMMPS)正是一个试图通过基于知识的技术来生成智能多媒体表现以解决这些问题的新研究领域。所谓 IMMPS 的定义是:通过有效的表现技术和媒体的综合,以能达到给定通讯目标并支持用户完成任务的方法来设计一个表达给定目标的表现过程。

决过程。

一、系统分类

IMMPS 的主要特点在于不仅能处理符号化的数据,而且能处理图像、图表、文本、声音、动画等,能够完成各种表现技术和其相关的媒体综合。IMMPS 将承担原来由用户完成的表现过程设计,避免了由于用户专业知识,时间或设计兴趣的缺乏而带来的困难;即使用户已掌握设计表现的方法,IMMPS 也能协助用户以最有效的形式完成表现任务。因此,智能多媒体表现系统必须要能够处理新的信息类型以及它们间相互关系,辨识并管理扩增的系统部分,检查新增加的设计比较原则,智能化地完成规划、选择、综合、表现、编排等环节,并能有效地表现问题解

智能多媒体表现系统的应用极其广泛,主要涉及数据分析、商业报表、军事情报处理、模拟训练、设备维修、过程控制、虚拟现实、动画制作等领域。IMMPS 分类有几种方式,其中,最常用的分类是依据表现媒体的特性(如静态与动态)和类型(如 3D, 2D)来划分表现(表 1)。表 1 中总结了近来研究原型中各种类型的多媒体表现。在所有这些项目中,生成系统包括:桌面出版,超文本设计,多媒体界面设计及卡通动画表现。

表 1

	信息图象	三维图形
静态媒体	地图、表格、框图 例如: SAGE, FNN	二维图形显示 例如: WIP, COMET
动态媒体	超媒体表现 例如: Alfresco, IDAS	动画 例如: VITRA, SOCCER, ANIMNL

[4] George J. Klr. Principles of uncertainty: What are they? Why do we need them?, Fuzzy Sets and Systems, 74, 1995

[5] Simon Parsons 等. A rough set approach to reasoning under uncertainty. J. Expt. Theor. Artif. Intell. 7, 1995

[6] 王 珏、袁小红、石纯一等,关于知识表示的讨论,计算机学报,18(3)1995

[7] 赵南元,认知科学与广义进化论,北京:清华大学出版社,1994

[8] 李 凡,人工智能中的不确定性,北京:气象出版社,1992

[9] 魏长华,无权重神经网络的不确定知识表达,神经网络理论及应用'94 最新进展,武汉:华中理工大学出版社,1994,10

[10] 魏长华,人类语言的不确定性及一种表示方法,华中师范大学学报(自然科学版)1996,6

[11] R. R. Jordan. Academic Writing Course. Willian Collins Sons and Co. Ltd, 1980

*)受国家自然科学基金资助,课题编号 69503008,69233011。受教外司留【1994】660 号资助

第二种分类方式是依据表现信息属性来划分, IMMPS 又可以分为如表 2 几种类型:

表 2

信息类型	典型系统
空间信息	<ul style="list-style-type: none"> • 三维对象的智能表现,例如,Feiner and Seligmann^[21]在 COMET 工作表明:三维对象可由三维图像显示来描绘。 • 地理相对位置的智能表现,例如,Shapiro^[20]和 Aren^[8]的系统通过图标和映射技术已经实现了地理相对位置表示。
时间信息	<ul style="list-style-type: none"> • 动画模型的控制,例如,Badler 等^[4]的 AnimNL 系统对卡通元素的时序控制;以及 Feiner^[8]的动画设计系统。
过程信息	<ul style="list-style-type: none"> • 过程指导的智能表现,例如,可以通过自然语言和实际的图片来表示,Maybury^[17]的工作表明了这一点。 • 智能训练与操作仿真,例如,Goodman^[9]的智能训练系统和 Andre, Rist^[11]的电子设备模拟操作。
数字信息	<ul style="list-style-type: none"> • 商业表格通过表格的图标表达实现,例如, Wahlster^[26]。 • 气象数据的自动处理,例如, Kerpedjiev^[12]天气预报自动生成系统。

第三种分类方式则是根据对表现的不同认识,就表现的形式和功能进行划分。一种认为研究重点应放在表现系统通讯目标的建模上,将表现看成被传输的信息;而另一种认为研究重点应放在面向对象的信息处理上,使表现系统成为用户用以开发信息的界面。

表 3 列举了这两种认识方法的特点、系统、缺陷。

表 3

	任务分析方式	通讯行为方式
特点	考虑表现设计中认识任务的特征,表现系统的目的在于能使用户搜索精确的数据,完成任务分析。	考虑表现设计中独立于媒体表现目标的特征与实现,通讯被视为一个受精确度、知识、目标和规划影响的过程,系统的目标在于生成相应于不同目标和不同上下文的表现。
典型系统	SAGE 系统 ^[25] 分析了信息搜索目标来表征用户为了不同目的从表现中获取信息的方式。	WIP 系统 ^[23] 通过规划技术来表征特定表现目标的多模式表现。
缺陷	必须把所有感知任务构成序列化的操作,难于表达需要并行处理的任务。	缺乏对表现结果认知复杂度的考虑。

以上两种方法可以看成是表现功能的不同实现,从两者存在的缺点来看,一个新的综合方式是必

须的。

二、系统的研究状况

IMMPS 不仅仅表现信息,而且还必须通过表现技术和媒体的综合技术完成用户目标,并在任务执行上支持用户的交互。在设计表现时,不论是手工还是用自动表现生成器,所有表现系统应遵守的表现原则是 Norman 所提出的适宜性原则:人工制品所用的表层表达所传给用户的内容应该是任务所适合的信息——既不多也不少!

IMMPS 的计算结构包括如图 1 所示的一些关键计算过程。

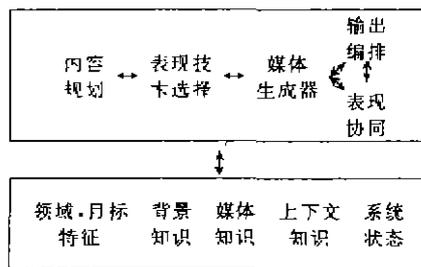


图 1 IMMPS 的计算结构

★领域内容规划:应用系统和多媒体表现系统间的界面,决定表现什么样的信息,选择支持用户任务所需的内容。

★表现技术选择:为实现选择合适的媒体或模式,协同不同信息选择的影响。

★媒体生成器:决定如何运用媒体来传送其所选的信息内容,本文将主要讨论文本生成器、图象生成器。

★表现协同:组合,组织,协调,解决,并通过反馈作用于规划过程,规划器,及文本、图象生成器等各部分,强调表现实现过程中各个部分间的相互依赖性,解决冲突,维护表现的一致性。

★输出编排:协调表现的最终物理输出,实现语义关系到视觉安排的一致性映射。

从图 1 所描绘的表现智能化设计中关键过程间的相互关系,我们可以初步地理解智能多媒体表现的概念结构问题。许多人把 IMMPS 的结构看成是包括了以上计算过程的一个几乎连续的流或者一个导致多媒体表现的完全范例化、具有表达力的分层规划过程。但是,一个有效表现设计过程必须包括平行交互的过程(Wahlster^[26], Seligmann&Feiner^[21]提出)。这个观点支持有关媒体和模式决定模块之间的外部反馈,而不是严格地分层化,序列化抉择过程。

从计算机发展来看,人们对表现的认识通常分为三个阶段:1)显示过程;2)编排过程;3)规划过程。在图1所示的结构中,这些过程的确是相互关联的,并都由数据和任务特征,所处上下文以及过程状态所驱动。

以上的计算模型是普遍适用的,为理解一些高层次设计问题提供了一个应用框架。实际上,绝大多数过程都存在于多种层次中。这些问题包括:分配,协同,从后续阶段回馈到内容选择,任务的考虑,在技术选择中的用户和数据特征以及相关的设计选择。我们提出这一统一计算模型在于:可以用于比较研究目前各多媒体表现系统。

在 SAGE^[25]中,首先选择相关信息,然后经文本和图形生成来组织信息,最后将形成的通用结构转化成文本和图形表现。这种方法的缺点是文本和图形的生成是相互独立的。

在 COMET^[7]中,系统首先建立一个映射表现组织形式的树状结构,然后由特定媒体生成器以单调的方式来扩充这个树状结构。COMET 系统的缺陷是不允许由媒体选择而导致的反馈。

Aren 和其同事在 Cicero^[2]的研究工作中,将规划和媒体选择严格分离开。在规划过程中,由表现通讯目标和内容决定的对话结构已由系统完全说明;此后,通过特定规则来实现媒体的选择;然后,对话结构从底到顶反向转化成面向表现的结构。这种方法的一个问题很明显是表现结构无法影响对话结构。媒体的选择受对话结构的影响,但内容的选择却是与媒体特性无关。

在 AIMI^[6]系统中,内容选择和内容组织相并行。虽然 AIMI 在内容选择过程中执行媒体选择,但是内容选择和媒体选择的相互依赖性被限制在一定范围内。例如,当 AIMI 的某个通讯行为既可由文本表现,也可由图形表现时,AIMI 并不将媒体约束和行为联系起来,因此系统难以表现在特定上下文中对行为实现媒体约束。

与上面各系统不同,WIP 系统通过规划策略将内容选择,内容组织和媒体选择集成起来,其优点在于方便不同模块间的协同。例如,在系统的表现策略中,首部包括媒体槽,在其中定义媒体的各种约束,直到策略的主行为得以实现后,媒体的具体形式才确定。

三、多媒体表现机制

3.1 内容选择

为了简化应用,维护独立性,理论上表现系统应将内容选择与系统划分开。然而,严格地将其区分也有欠妥,因为内容规划必须受到如下影响:(1)表现的规模和复杂度;(2)信息量;(3)硬件设备;(4)表现完备性和一致性需求的限制。例如:表现内容的选择必然受到屏幕显示空间和硬件功能的限制,必须考虑在试图传送不同内容规划结果时可能导致的辨识复杂度,必须借鉴先前传输信息的设计手法等等。这些方面的协调工作也只有通过内容规划与表现单元间的反馈(或其他机制)来完成。

目前,大多数表现系统处理内容规划的方法大致为:应用系统或多或少地决定内容,而表现系统只是负责表现内容的选取与组织,普通做法是由表现系统通过与领域内容知识库的通讯来获取并修正相关信息。这种基于反馈思想的内容规划方法最初由 Roth 等^[24]在 1991 年提出,并在自然语言生成模块中使用修辞模型将内容规划与表现设计集为一体。修辞模型用来定义某一将被表现的概念类型与组织结构的一系列通讯行为,例如,设备维修过程的指导。修辞模型为对话规划器(Cicero 系统^[2])提供了从较大内容库中选择相关内容,序列化相关知识。

McKeown 等^[7]在 COMET 系统中开发了支持表现设计单元中为陌生条目获取附加信息的方法。在系统文本生成器中,若词汇单元发现一陌生新条目时,就反馈给内容规划单元加以辨认或提供表现所需的信息。

WIP 系统^[3]列举了几种单靠严格的序列化方法(即输出结果只从内容规划流到表现设计)难以完成的任务:媒体交互参照(Crossmedia referring expressions),首语参照表达(Anaphoric referring expressions)等。WIP 系统虽然认为内容规划不应是表现系统的任务,但为了反馈的需要,用 RAT(Representation of actions using terminological logics)将领域知识形式化,(RAT 模式可被用作有关领域初步知识的文本生成及图象生成)。表现规划器能利用 RAT 提供的一些推理机制,WIP 使用了一个非正规的规划技术将内容规划与媒体选择交错起来。

目前,有几个实验正试图将内容规划集成到系统中,但实验结果基本上是表现限于应用特殊性中,这样表现就成为事先预制的过程,而不是依靠数据特征和表现知识的综合动态结果:很明显,这种方法只是适用于有限的特定场合。

选择了要传输的信息类型在一定程度上也就确

定了以何种方法来表现它,因此,内容规划必须是一个综合的、整体的过程。规划的选择依据应建立在以下知识之上:表现的(所有)通讯目标;表现数据的特征;制作者、用户、使用环境的特征;多媒体信息推理知识。

无论用怎样的方法规划内容,用怎样的机制实现反馈,系统都应遵照 Grice^[11]提出的最大原则来决定表现内容。一方面,确保所有相关信息都被编码;另一方面,不必告诉用户他已了解的信息。

在 IMMPS 中,为了支持交互功能,内容选择与表现单元之间的双向通讯是必然的,其中最自然的方法是通过类似数据库漫游或搜索的机制。双向通讯的优化是将来研究方向之一。同时,为了实现并扩展这种协同反馈关系,IMMPS 必须要能处理大信息集的表现任务,然而这一点当前大多数 IMMPS 都无法胜任。无疑这也正是 IMMPS 未来的研究方向之一。

3.2 信息在媒体间协同分配

一旦系统决定了表现什么信息,就须开始决定以哪些媒体来转载这些信息。

关于媒体选择知识,有三大类:媒体选择与协同的知识;表现技术选择知识;信息到表现间映射知识。其中媒体选择中所需的信息特征包括:内容、途径、时序、抽象度、与其他信息的关系等等。

目前,在绝大多数系统,选择知识表达形式采用规则。规则用来定义不同媒体的表达能力以及各媒体的不同实现技术(例如,在信息图形中,各对象的水平位置,形状,大小等的安排)。基于规则的媒体选择表达研究起源于信息图形自动设计系统 APT^[16]。该系统中开发了“表达规则”用以根据信息特征决定图象显示的编码技术。同样地,The Integrated Interfaces System^[3]使用简单“表现规则”来设计美国海军太平洋舰队每日报告的文本、图象、表格和地图的混合显示。与 APT 不同,该系统是基于 KL-TWO 知识表现机制来处理更大规模的数据。

然而,媒体不同表现能力的理论基础却难以建立,所以大多数系统目前只能采用相应的领域启发知识。例如 SAGE,以混合自然语言和信息图形来解释数值模型,其中媒体选择的启发知识,如:若内容中包括偶发率,过程,结论及其他抽象概念的表达或小规模数据的关系属性,则选用自然语言表达形式。

在 WIP 中,进一步将这些启发知识形式化(元级规划)并且成为独立知识库的子库。WIP 的元级规划知识具体分为:Design Rules, Style Rules, Lay-

out Rules, Expressionness & Effectiveness Constraints。其标准形式采用产生式规则。例如,以下规则表示:建议在表现空间信息时使用图象而不是文本:

```
IF (IS-A current-attribute-value Spatial-Concept)
THEN (DoBefore "Graphics-Strategies" "Text-Strategies")
```

其他规则还包括:

图形优于文本若表现具体信息(如:形状,颜色,纹理等);

图形优于文本若表现空间信息(如:位置,方向,结构等)以及物理行为和事件(除非要精确表示速度值);

文本优于图形若表现时序重叠,频率(如:通常或偶尔),时间变动(如,六天前)等;

文本优于图形若表现语义关系(如:因果,条件等),以避免仅用图序列表示易造成的二义现象;

文本优于图形若表现数量概念(特别是如:绝大多数或一些或确定的数值)。

WIP 利用了几种子类型和十个通讯功能单元(吸引,比较,阐述,使能,解释,标示,激发,显示,背景,总结)识别七种信息类型(具体的,抽象的,空间的,协方差的,暂时的,量化的,否定的),并检查哪一种模式或模式的混合能最佳地传送它们。

与 WIP 的规划方法相反,COMET 采用了管线方式在文本和三维图形间进行信息分配。其启发知识包括:仅用图形来表现位置和物理属性;仅用文本来表现抽象行为和行为间的联系;并用文图来表现简单和复杂的行为。

以上方法的特点是,系统都是将信息特征映射到媒体类型(如:文本、图形)或媒体对象(如:表,框,饼图)上。与此相反,Hovy & Aren^[13]分析并提取了一套媒体特征,(如:MEDIUM, EXHIBIT, SUBSTRATE, INFORMATION CARRIER, CARRIER ITEM, INTERNAL SEMANTIC 等),提出从信息特征到媒体特征映射的媒体选择规则。他们将媒体的选择特征化为多步骤过程:

(1)表现数据元组,如位置时,使用平面媒体,如图形,报表,地图等;

(2)表现有特别注释的数据时,如空间的注释,使用有相同注释的媒体,如地图上的方位;

(3)如果有多种媒体可用,并且有一个已存在的表现手法,则使用当前表现已使用的媒体;

(4)媒体的选择必须能包容最大量的信息。

进一步,他们在系统化的框架中定义了一套规则来描述信息与媒体相互依赖所造成的复杂性。例如,若要传送说明:“危急”,规则之一是:

如果信息已经被显示了,则使用当前媒体,但将一种或几种以前固定的部分变为相应的时变状态。

若系统只是将信息特征映射到媒体类型如:文本、图形或媒体对象如:表,框,饼图上,这在一定程度上限制了系统选择媒体的自由度,也在一定程度上影响了表现输出的多样性。而在信息特征与媒体对象间加入一个中间层:媒体特征,将大大扩展系统的自由度,方便系统日后的媒体类型对象的扩充。

一旦媒体类型选定之后,媒体表现目标的实现在很大程度上依赖于其技术的选择。技术选择的依据为:感性操作,信息搜索的目标,通讯行为或意图,实用指示。

另外,需要强调的是:媒体选择过程并不能独立地完成,必须受到内容规划的反馈信息影响,并受到表现协同的进一步约束。

一个成功的多媒体表现系统并不是在表现信息时能使用不同的媒体,而是能在媒体替换上做出周密考虑,知道如何以一种一致的形式来集成不同媒体,还须顾及不同媒体中信息表现完备性和冗余性的适宜程度。一个不包括任何冗余信息的表现是不可能一致的,但如果太多冗余信息,又会影响用户的注意力。

当前的媒体选择工作在于形式化这些规则,如何进行评估,已涉及大量相关心理学方面的研究。媒体选择机制的完善将取决于认知心理学,设计心理学及美学的深入研究。

3.3 媒体生成器

当完成媒体选择,信息协调分配之后,系统必须能开始处理特定媒体的信息编码工作,一种直接的方法就是依靠指定的生成单元,例如,文本生成器,图形生成器,动画生成器等。这些单元集成了专家设计知识,提供了媒体技术自动选择,生成与混合的机制。

WIP 实现了文图的自动集成,其媒体生成由文本生成器和图形生成器共同完成。由于受到渐进过程要求的强烈影响,各生成器中的基本过程单元,内部数据流的形式和大小以及各单元间相互作用都体现了渐进生成原则。WIP 的文本生成器分为两大部分:文本设计和文本实现,实现部分又有两个层次:功能层和位置层。文本设计部分接收输入的表现内容,将其组成语义完整的句子,并送至文本实现部

分,实现部分根据约束来确定文本输出的适当形式。为此,设计部分与实现部分之间存在反馈。WIP 的图形生成器区分了几种基本的图形类型,将表现任务和一套约束联系起来,这样图形表现任务能以约束满足方式完成。WIP 开发了一套图形设计策略来逐步将表现任务细化为生成器能够实现的元行为。目前,WIP 归纳了以约 50 条规则为基础的注释技术。

COMET 的文本生成器也分为两个部分:词汇选择和句子生成。与 WIP 的自然语言处理机制不同,COMET 只是根据不同的基本约束来选择适当词汇,使得同一概念能用不同词汇来表达。这些约束来自基础知识库,先前对话,用户模式以及文法。此外,COMET 还能保证不使用生疏词汇。COMET 的图形生成器 IBIS(Intent-based Illustration System)采取生成-测试方式,对输入的通讯目标至少有一个设计规则与之相应,每个设计规则又激发一套类型策略以说明高层视觉效果,而这一切都被交给更低层的规划去具体实现。同时 IBIS 有一套评估任务完成质量的规则,若测试结果失败,则返回,试用另一个表现策略(这一点与 WIP 相似)。

作为智能多媒体表现系统的媒体生成器,它必须能相互协调,相互了解表现情况,因此各个生成器必须提供其编码的详细表达形式;必须能够与表现协同单元进行通讯;必须具有根据新的约束条件对生成结果进行修改的功能;必须能够与其他媒体生成器协同生成交互参照表现。目前,关于媒体生成器的研究除了继续在各生成器间的协同,参照表达机制以及生成结果的多样化方面进行拓展以外,大量的研究已涉及到了新类型媒体生成器(诸如:声音,动画,视频等)的设计。

3.4 协同表现过程

协同是指表现元素包括了两种或两种以上媒体编码的信息,因此在一定程度上可以看成是媒体的混合。智能多媒体表现的协同支持机制研究主要包括:(1)媒体间和媒体内部协同编码技术;(2)显示编排协同技术。本小节主要讨论媒体编码技术协同,输出编排协同将在下一小节讨论。

成功的表现不仅要为信息选择适当的媒体,而且要保证多媒体间及媒体内部的协同。因此表现系统必须作到,首先必须保证媒体间内容是一致的;其次,必须保证结果的输出形式,表达形式是一致的。

在早期的表现系统中,如:SAGE 只是简单地根据一些规则完成内容到表现的映射。而 COMET 系

统统使用一个逻辑形式系统来描述各生成器中的信息,为各生成器提供更多的信息,以及生成所需上下文内容。系统通过集成语句片段和图形片段,来实现从文到图的交叉参照功能,并且允许文本生成器和图象生成器之间相互作用来修正生成内容。COMET^[7]进一步考虑了基于时间推理机制并通过管理通讯行为的次序和持续时间来控制时间媒体的表现。

与 COMET 系统在内容规划之后再实现媒体选择不同,WIP 和 TEXPLAN^[14]使用规划操作元来同时完成内容选择和媒体选择,这使得媒体的显式编码与内容约束具有同样的形式,以生成最终分层多媒体表现规划。媒体生成器能够与规划操作元进行交互,也可与各生成器间进行交互,向设计模块提供反馈。例如,在 WIP 中,文本生成器与图形生成器可以相互生成对象明确的语言和视觉参照。这种交互使得文本生成器可以生成诸如:“on/off 开关位于图片的左上角”的视觉参照说明。

COMET 和 WIP 两个系统在文本和图形的协同上都费了很多的精力,由于在系统结构、处理策略、表达语言上的不同,两者的协同表现机制有很大的差异,也都存在各自的不足之处;COMET 首先建立表现组织结构,导致系统只允许媒体生成器有限的反馈,造成系统不允许由媒体选择导致的修改。WIP 使用基于规划的方法较好地解决了各部分间的交流、反馈,但因为 WIP 采用渐进生成方法,信息被逐步地传送给生成器,在生成器获得足够信息以生成句子或图示之前,规划器必须一直评估实现的成功率,这增加了系统的消耗。

3.5 表现自动编排

表现物理形式的编排通常传达了信息的结构,意图和重要性;使表现呈现出一致性,绝大多数对编排的研究都集中在单一媒体上,例如,Hovy&Aren^[13]开发的文本生成和修辞结构,使用文本形式化程序(Tex)来指导最终文本的形式,例如,当文本规划模块用 SEQUENCE 关系来构造文本时,可以通过 Tex 命令/bullet 来实现。类似地,在 ANDD 系统中,Mark^[15]研究了点-弧图示编排和编码。ANDD 将有相同图形属性(如:形状,颜色,大小)的节点集合起来以强调某种类型图形。

与上述单一媒体的研究不同,Feiner^[6]的图形界面设计系统 GRID 研究了文本,图形以及虚拟输入设备。编排在一个类似 OPS5 生成系统中由一个图形设计栅格指导完成。与这种基于规划的方法相反,

Graf^[10]指出满足美学要求的输出编排设计可以被特征化为约束满足问题,WIP 中使用的就是 Graf 的 LayLab system 基于约束的编排管理器。通过考虑自动生成的文本片段和图形片段的混合体的视觉安排中的语义和实用关系以实现一致、有效的输出。LayLab 包括了文档模板知识(例如:滑条,显示环境),设计启发知识(例:垂直还是水平安置)以及图形约束知识。编排包括了通过特殊可视化技术实现从语义和实用关系(例如:“序列”、“对比”关系)到几何拓扑/时间约束的映射。例如:两个相同大小的图形可以并列在一起相互比较或叠加在一起相互比较。因此,编排被视为含义的一个重要载体。编排不仅可以通过静态文-图混合实现,而且可以通过动态渐增的表现手法和更多的媒体形式(如:动画,视频)来完成。Graf 准备继续开发一个约束获取模块,能够从人类专家的草稿中获取设计约束。

另外,在 CUBRICON 系统中,Neal&shapiro^[20]开发了智能窗口管理器 IWM 来支持多显示器/窗口的交互。IWM 通过计算窗口重要性的权值来控制生成媒体在各显示器/窗口的分配。

3.6 交互能力

当前表现系统主要缺少交互能力,交互能力作为当前研究的主要拓展方向正日益得到重视。系统交互能力应该完成的工作是,允许用户说明目标或任务;允许用户选择技术和协同方法;允许用户选择渲染方法等。但是,由于自然语言处理的困难,交互机制的复杂性和用户认知过程的不确定性,导致当前多媒体表现系统只是部分地实现了这一重要能力。

WIP 系统目前最显著的缺陷是在表现过程中它并不支持用户的交互性。一个交互用户在对表现的当前类型不满意或对目前为止生成的表现有疑问时,他会要求与系统进行交互作用。这就要求系统的表现规划器能进行扩充以允许必要的响应规划,并允许用户在表现期间改变表现参数。与 WIP 相似,COMET 对人机交互的实现只是允许用户选定目标(任务),并不支持用户的动态干涉。

Cicero 系统的对话结构通过人机交互以获取最新信息,为了避免自然语言理解的困难,它采用了高层结构化语言——脚本,这样也就要求每个领域的典型用户任务必须认同这个脚本。

很明显,以上的系统在多媒体信息的充分交互上存在着缺陷,有许多基础问题还有待解决,其中包括:支持智能多媒体交互的系统结构问题、知识问

题、媒体集成技术问题以及评估方法问题。

- 系统结构问题：对于多媒体交互，知识源和知识处理过程哪一个是关键问题？它们又分别需要什么样的功能来支持？各自适合的控制过程应是怎样的？（序列化，交错，还是根据共同约束）。

- 表现知识的获取、表达、维护、开发。系统所需的知识包括：信息特征、制作目标、媒体特征、用户特征、交互过程的记录、任务完成情况的记录等等。

- 媒体集成技术：包括处理结构不良的和不完全的输入，处理不同抽象层次上的输出，例如：语法上、语义上、实用主义上；使用独立于媒体的抽象行为概念，例如：分析、生成、比较等；开发新的媒体类型，例如：面部表情、肢体语言等，和新的模式，例如：接触，气味等。

- 评估方式：某些情况下，我们必须根据表现的结构优劣、一致性、协同性等来决定各种表现的好坏；某些情况下，系统必须根据教学效果，任务完成的效果来决定是否增加还是减少输入、输出的投入。所有这些都需要标准的评价准则、标准的使用方法。

尽管目前存在这么多的困难，但是 IMMPS 从结构上，技术上都是适合开发交互能力的。这需要开发设计过程中交互信息的获取，理解，集成，抽象，分配方法，需要相关领域（如：认知科学）的进一步研究。

四、研究趋势

本文阐述了智能多媒体表现系统的由来，概念结构，研究方法以及当前主要的研究方向。从总体上

来说，智能多媒体表现系统将在以下三个方面使得系统和用户能最充分地利用媒体信息：第一、能增加人机间的信息流量，例如：通过使用最适宜的媒体或各种媒体的混合形式；第二、能协助用户理解信息，例如：突出表现最重要的特征；第三、目前对媒体精确模型的研究将有助于将来人机界面的设计，例如：将来的界面可以加入人类交流的方式：语音语调的变化，面部表情，手势表达等。总之，我们认为当前的研究重点和将来完善的目标主要包括：

- ★构造选择适宜通讯的信息的机制；

- ★表征与表现生成有关的信息的关键特征；

- ★表征通讯内容或目标以及表现设计任务；

- ★总结归纳媒体选择知识；

- ★总结归纳表现媒体内部技术选择知识；

- ★总结归纳各媒体间和媒体内部各模式间的协同知识；

- ★开发表现系统的交互能力。

所有当前的研究都是将表现系统视为以计算机为中心的连接人、机和任务的过程。但我们认为，由于智能表现系统的最终对象是各个特定环境下的特定用户，因此随着认知科学等相关研究领域的发展，智能表现系统将重新以用户为中心，充分考虑上下文来扩展/构造上述各个模块。另外 IMMPS 的进一步发展将不仅仅得益于实用主义和用户任务的完善，而且将得益于 IMMPS 更加广泛的应用；自动交互表现设计，作为辅助专家评判工具，作为信息开发和模式化工具。（参考文献共 28 篇略）

全国第一家科技半月刊

中国科技信息杂志 权威新颖、开卷有益

每月一期综合版：将诠释这样一个命题：科学家的头脑+企业家经营=财富

每月一期生活版：将关注这样一个话题：日新月异的科技将如何改变我们的生活方式。

中科信时空（刊中刊）：技术、投资、产品、人才、市场、会展、图书等高密度分类。

中国科协主管，中国科技新闻学会主办，国际大 16 开 56 页，每月 1 日、15 日出版，每期定价 8 元，全年定价 192 元，全国各地邮局均可订阅。邮发代号：82-415

本刊发行部地址：北京复外真武庙二条财政部招待所五层

邮编：100045 电话：(010)68529305 68521028