

EMS 人机界面及其交互技术研究

朱全胜, 刘 晓, 李卫东

(大连理工大学 电气工程与应用电子技术系, 辽宁 大连 116024)

摘要: 人机界面是能量管理系统(EMS)中的重要一环, 利用若干先进的人机交互技术, 对 EMS 输入端、输出端 2 方面进行了改进。输入端在原有的交互工具上加入语音识别, 从而和鼠标、键盘形成一个多模式的输入界面。该输入界面能够减少命令输入时间以及实现多命令并行输入等。输出端引入关联多屏显示技术, 并将 2-D 可视化和 3-D 可视化结合使用, 在听觉输出上使用语音合成技术。这些综合输出方式的结合使用能够将系统信息以更加合理、自然的方式呈现出来, 便于调度员更加快速地把控系统的运行情况。

关键词: 人机界面; EMS; 交互技术; 语音识别; 语音合成; 可视化

中图分类号: TM 76

文献标识码: A

文章编号: 1006-6047(2007)08-0077-04

0 引言

人机界面是能量管理系统(EMS)的一个关键环节, 好的人机界面能提升调度员分析、解决系统问题的速度, 提高调度员日常调度及处理事故时的效率, 为良好的经济调度和安全控制提供保障。相反, 不合理的人机界面会减低调度员的工作效率, 甚至会间接影响到系统的安全性和稳定性。

一个好的 EMS 人机界面应该做到: 调度员能够快速、准确地输入命令; 计算机能清晰、合理地展现调度员所关心的信息。

目前的人机交互技术给开发人员提供了多种选择。在输入端, 除了常用的“键盘+鼠标”组合模式外, 还有语音识别、视线跟踪、手势识别及面部识别等新的交互手段, 这些输入技术既可以单独使用, 也可以结合组成多通道界面。利用多通道技术, 可以实现文献[1]中提到的“动用人类各种感官, 实现人和计算机的全面沟通”。到目前为止, 除了语音识别之外, 以上提到的其他新型输入技术都未成熟, 并不能实际应用在 EMS 中。语音识别则可以通过和鼠标、键盘组合使用, 增加输入手段; 输出端的研究进展并没有输入端显著, 目前多媒体的输出模式能满足大多数的需要。在 EMS 中, 仍然要以视觉输出为主, 在努力提高视觉输出性能的同时研究其他感官输出的可能性。

1 EMS 人机界面的发展回顾

1.1 第一代人机界面——字符图形用户界面

字符图形用户界面, 又称有限图形用户界面, 是 EMS 的第一代人机界面。

字符图形用户界面除了可以显示数字和文字类型的信息, 还可以显示一些简单的字符图形(例如单线图)。早期的字符图形用户界面主要通过键盘控制, 随着新型指点设备的发明, 输入技术也有所提高, 广泛使用了轨迹球、光笔、鼠标等指点设备。

此外, 该界面的图形输出能力很有限, 只能显示简单的单线图。

1.2 第二代人机界面——全图形用户界面

EMS 的全图形用户界面逐渐取代了字符图形用户界面。全图形用户界面引入了 WIMP 的概念, 即窗口(Windows)、图标(Icons)、菜单(Menus)、指点设备(Pointing devices)。

全图形用户界面的特点有: 图形可以自由地被缩放、平移、旋转; 可以利用多窗口显示更多的观测信息; 窗口可以自由地缩放、打开、关闭、移动; 调度员主要利用对话框和计算机进行交互等^[2]。

与字符图形用户界面相比较, 使用全图形用户界面不需要记忆大量的指令, 从而减轻了调度员的认知负荷; 同时强大的图形输出能力以及更加人性化的界面减轻了调度员的操作负担。此外, 电力系统可视化技术的应用进一步提升了全图形用户界面的交互水平^[3]。

2 现有的交互技术

2.1 输入端交互技术

从字符图形界面后期开始, 调度员一直使用鼠标和键盘作为输入工具。文献[4]的例子验证了使用这种输入方法造成的低效与不便: 在字符图形用户界面或者全图形用户界面中采用传统的输入方式(鼠标+键盘), 调度员需要完成 17 个动作实现一个很小的任务(断开断路器), 并且双手需要在鼠标和键盘中不停地移动。

2.2 输出端交互技术

多屏显示作为电力系统中较为常见的显示技术, 在系统的监控工作中起着重要的作用。目前的

多屏显示系统中,各个屏幕是相互孤立的,这种孤立的屏幕并不利于信息显示和调度员操作。

目前广泛使用的全图形用户界面已经完全可以结合 2-D 可视化技术进行使用,然而 2-D 可视化技术虽然可以同时显示海量数据,但并不能显示太多的数据种类^[5]。

除了视觉输出之外,EMS 还经常使用到听觉输出,即利用计算机发出一些简单的报警声来提醒调度员注意。报警形式较为单一,通常除了通知调度员系统发生了事故之外,并不能有效传达一些事故的相关情况。

3 输入端交互技术的改进

3.1 新型的输入技术以及多通道界面

除了鼠标和键盘等比较常见的输入手段之外,几种新型的输入技术主要包括:语音识别^[6-8]、手势识别^[9-10]、视线跟踪^[11-12]等。与鼠标和键盘不同的是,这些输入方式都近似于人与人日常的交互方式,直接通过人的效应通道向计算机发布指令,交互方式更加自然。

以上的交互技术除了可以单独使用之外,也可以互相整合在一起组成多通道界面。多通道界面通过多个通道的协同工作以及相互配合,从而产生较之单一通道更大的效率。目前,比较成功的整合案例有“语音+手势”^[13-14]、“语音识别+视线跟踪”^[15-16]、“语音+笔输入”^[17-18]等。其中,“语音+手势”的组合方式最为流行,多通道交互的第一个具体例子应用的正是“语音+手势”的结合^[13]。

3.2 语音识别技术

语音识别就是让机器通过识别和理解过程把语音信号转变为相应的文本或命令的技术^[7]。

在 20 世纪 50 年代,贝尔实验室就开始研究语音识别技术。从 20 世纪 80 年代末至今,计算机的语音识别技术逐渐有了很大突破。到目前为止,在实验室条件下,语音识别系统的精度已经达到了相当的高度^[7]。

尽管如此,语音识别技术还存在着诸多问题:在噪音环境下识别率较低,口语化的语言识别率低以及方言的识别率低等。但上述问题不会对语音识别技术在 EMS 的应用有所影响:控制中心的环境很好,不存在噪音干扰的问题;尽量在语音输入过程中使用专业术语和正规的书面语言,从而避免口语化;使用标准语言(例如汉语的普通话)进行语音输入。

因此,尽管目前语音识别技术的使用范围仍有局限性,用户也不能随意地和计算机进行流利的自然语言(natural speech)交流,但其完全可以作为一种交互技术应用在 EMS 中。

文献[19]设计了 2 种针对电力系统的语音输入(日语)界面:控制台操作界面以及错误操作预防界面。通过实验得出:控制台操作语音输入界面能够减少操作时间、减少操作的次数,从而能够提高系统

的操作效率;错误操作预防界面则利用语音识别和指点动作的联合使用减少了操作的失误次数。该实验充分验证了语音识别的可用性。

文献[20]分析了语音技术(语音识别+语音合成)在控制中心中能够起到的作用。认为在控制中心使用语音技术能够减少成本、提高效率,以及能够更快地与关键的应用程序进行交互。

3.2.1 语音识别+鼠标+键盘

语音识别既能代替鼠标的点击和选择功能,也能代替键盘的输入功能,但在 EMS 中完全用语音识别代替鼠标和键盘是不切实际的,因为很多时候,语音识别并不是最有效的办法。例如,在地理信息图中,鼠标能很快地选择某一个模糊区域(没有确切命名),语音识别则因不知该区域的名称而无法进行定位。此外,在文字输入方面,由于语音识别在识别率以及错误更正等多方面存在的问题,单纯地使用语音识别要比“鼠标+键盘”组合输入慢很多^[21-22]。

为了更好地进行输入,可以将语音识别和其他输入工具整合使用,目前比较可行的办法是结合传统输入工具——鼠标和键盘,语音识别作为其中的一种输入手段供调度员有选择性地使用。这也是一种类似于多通道界面的组合。文献[23]中的实验也证明,语音识别和鼠标、键盘的结合使用更有价值,它们可以互相弥补各自的不足,通过相互的融合提高了界面的性能以及被接受度^[24]。

3.2.2 语音识别+鼠标+键盘的优点

与传统的“鼠标+键盘”的输入模式相比较,加入语音识别的人机界面有诸多优点。

a. 语音输入通常可以减少输入命令的步骤和时间。完成图形用户界面的一个任务,调度员通常需要点击很多按钮,并且经常需要在鼠标和键盘之间切换^[4],而语音输入只需要很少的步骤就可以实现。此外,语音识别在进行导航时的优势也很明显,调度员不需要花时间寻找所要观测的区域或者对象(有确定命名的区域或对象),利用语音识别输入可以迅速地定位目标^[19],特别是当目标比较隐蔽,难以寻找时。

b. 语音识别可以和键盘或鼠标同时进行输入操作,从而实现并行的多命令输入。在同一时间内,语音、鼠标、键盘三者可以单独使用,同时也可以相互组合成“语音+键盘”或者“语音+鼠标”的方式。例如调度员在用鼠标点击命令按钮的同时,可以用语音输入另一个指令,从而实现多命令的并行输入。

c. 在特定的时候调度员可以一边注视着屏幕的某个地方,一边用语音输入指令(监控和输入同步进行)。调度员在鼠标或者键盘输入命令时,眼睛不能一直注视感兴趣的地方:移动鼠标时,视线和指针的位置是一致的;使用键盘输入时,即使能够熟练操作计算机的用户仍然需要视线结合键盘进行输入。使用语音识别,则可以在认真观察屏幕的同时输入指令。

d. 用语音识别代替鼠标和键盘去输入指令,可

以减轻手的负担。文献[4]的例子中,采用传统的输入方式(鼠标+键盘),调度员需要13个步骤去完成一个小任务,双手在鼠标和键盘中不停的移动,整个输入过程过于依赖双手(特别是调度员的惯用手),手和手臂因此负担了很大压力。长时间如此,容易引起肌肉损伤^[25],特别是容易引发腕管综合症。加入语音识别的人机界面,可以在一定程度上减轻手的工作负担,避免疾病的产生。

4 输出端交互技术的改进

4.1 关联多屏显示系统

目前的多屏显示在应用过程中,各个屏幕间是互相孤立的,没有真正的内在联系,只是简单地扩充了监控的有效面积,这样做虽然能够提升调度员的交互效率,但没有最大程度地发挥多屏的作用。此外,很多情况下调度员所执行的任务有着很大的区别,传统多屏显示中各个屏幕的显示内容和显示方式是固定不变的,不能随着任务类型的变更而进行调整。

文献[5]通过对多屏显示深入的剖析,提出了关联多屏显示这一全新的概念。文章以电力系统为研究对象,对关联多屏显示中最重要的问题——关联多屏之间的关联关系进行了分析,并且依据作者设定的分类依据,归纳出对象相关、信息相关、完全相关和完全不相关等4种关联关系。

关联多屏显示较之传统的多屏显示主要有以下的优点:符合调度员操作习惯,快速实现任务,并联的多任务显示等。

4.2 3-D 可视化

3-D 可视化在其他很多领域的应用都很有成效,特别是利用科学可视化技术显示一些物理现象时。3-D 可视化可以生动地模拟实际的场景(例如导弹发射过程),也可以很精准地把微观的东西(例如DNA结构的显示)展现在用户的面前。这是因为这些被展示的事物大都有物理实体,所以用3-D的可视化更能很好地将他们呈现出来。然而电力系统所显示的很多信息没有相对应的实体表现形式(例如功率、电压等)^[26],所以3-D可视化的实用性并不是特别明显。

在电力系统中,3-D 还是一个比较新的课题。目前仍停留在比较浅的层次,其应用潜力还没有得到充分的发挥^[3]。

文献[26-29]采用透视投影(perspective projection)的方法创建电力系统的3-D 可视化景象,该方法是在原有2-D图形的基础上稍作改进而成:2-D 可视化图形作为X-Y平面(水平平面),新增加了一个Z轴(垂直方向)用来显示一些其他信息。上述的3-D 可视化技术可以同时显示很多的信息种类^[28],例如调度员对电压进行安全分析时,需要同时显示节点电压、发电机的无功情况(包括无功储备)等,这时候用3-D 可视化图形就可以很轻易地显示出全部内容^[26]。

但同时,受限于3-D 的特点,距离视线近的对象往往显示得比较大,容易被观测到,远处的比较小而难以分辨,甚至可能被遮盖,不利于调度员掌握该图形所显示内容的全局。而2-D 虽然不能在一张图例上显示过多的信息种类,但由于图像呈现的是一个平面,调度员更容易把握全局。

从实用性的角度分析,2-D 和3-D 各有特色,用3-D 完全取代2-D 是不合适的,研究人员更应该考虑如何将二者联系在一起应用,以弥补各自的不足。例如在观测全局时使用2-D 的地理信息图,在观测局部时使用3-D 可视化,以便于在细节上观察到更多的信息种类。

4.3 语音合成技术

语音合成技术(text-to-speech)又称文语转换,是指计算机把文本或其他形式的信息以语音的方式输出。

较之语音识别技术,语音合成技术更为成熟。到目前为止,合成的语音输出清晰度和可懂度很高,但是自然程度和流畅程度不够,与自然语言有较大的差距。尽管如此,但不会影响此项技术在EMS中的应用。

在EMS中,语音合成技术主要有2个作用。

a. 计算机用语音回答或提出问题,和调度员进行简单的对话交互。例如,在语音输入的过程中,当计算机不能完全识别调度员的指令时,用语音的方式指引调度员输入能够被正确识别的指令。

b. 代替传统的报警方式。传统的计算机报警通常是通过发出一些特殊的声音来提醒调度员,报警形式过于单一。利用语音合成技术,可以将一些必要的信息用精炼的语言告诉调度员。例如通知调度员何时何地以及发生了怎样的问题,便于调度员在最短的时间内初步了解事故的简要信息。

5 结语

人机界面是EMS中至关重要的一个环节,改进其交互技术势必会提高调度员的工作效率,减少调度员的操作负担。通过引入多种先进的人机交互技术,从输入端和输出端2个方面提出了改进的办法。

在输入端,语音识别是比较成熟的输入技术,可以通过和鼠标、键盘整合使用,发挥各自的特长,从而实现减少输入命令的步骤和时间、多命令并行输入、监控和输入同步进行以及减轻双手负担等功能。

在输出端,利用关联多屏显示代替传统的多屏显示技术,该技术较之传统的多屏显示主要有以下的优点:符合调度员操作习惯,快速实现任务,并联的多任务显示等;3-D 可视化已经在电力系统被初步研究,与2-D 可视化相比较,其更适合用于显示局部区域的多种类信息,在实际应用中应该与2-D 可视化配合使用;通过语音合成技术,计算机能够将一些简单的文本信息以语言形式输出,从而完成与调度员进行简单对话以及事故报警等任务。

参考文献:

- [1] 张伯明. 现代能量控制中心概念的扩展与前景展望[J]. 电力系统自动化, 2003, 27(15): 1-6.
ZHANG Bo-ming. Concept extension and prospects for modern energy control centers[J]. Automation of Electric Power Systems, 2003, 27(15): 1-6.
- [2] D'AMOUR P R, BLOCK W R. Modern user interface revolutionizes supervisory systems[J]. Computer Application in Power, 1994, 7(1): 34-39.
- [3] 刘尧, 李卫东, 吕阳. 电力系统运行状态可视化技术综述[J]. 电力系统自动化, 2004, 28(8): 92-98.
LIU Rao, LI Wei-dong, LÜ Yang. Surveys of power system operating state visualization research[J]. Automation of Electric Power Systems, 2004, 28(8): 92-98.
- [4] CHRISTIE R D. Towards a higher level of user interaction in the energy management task[C]//IEEE International Conference on Systems, Man, and Cybernetics. Piscataway, NJ, USA: [s.n.], 1994: 1086-1091.
- [5] 朱全胜, 李卫东. 电力系统多屏显示技术的深入研究[C]//中国高等学校电力系统及其自动化专业第二十一届全国学术年会论文. 南宁: 广西大学, 2005: 1629-1633.
ZHU Quan-sheng, LI Wei-dong. Correlated multi-screen display in power system[C]//Proceeding of 21th CUS-EPSA. Nanning: Guangxi University, 2005: 1629-1633.
- [6] ZADEH J. Technology of speech for a computer system[J]. Potentials, IEEE, 2004, 22(5): 35-38.
- [7] 刘加, 刘润生. 语音识别技术及应用(上)[J]. 世界电子元器件, 2001(11): 21-23.
LIU Jia, LIU Run-sheng. Voice recognition: technology and application[J]. Global Electronics China, 2001(11): 21-23.
- [8] 刘加, 刘润生. 语音识别技术及应用(下)[J]. 世界电子元器件, 2001(12): 23-24.
LIU Jia, LIU Run-sheng. Voice recognition: technology and application[J]. Global Electronics China, 2001(12): 23-24.
- [9] 方志刚. 计算机手势输入及其在人机交互技术中的应用[J]. 小型微型计算机, 1999, 20(6): 418-421.
FANG Zhi-gang. Computer gesture input and its application in human computer interaction[J]. Mini-micro Systems, 1999, 20(6): 418-421.
- [10] 李清水, 方志刚, 沈模卫, 等. 手势识别技术及其在人机交互中的应用[J]. 人类工效学, 2002, 8(1): 27-33.
LI Qing-shui, FANG Zhi-gang, SHEN Mo-wei, et al. Gesture recognition technology and its application in human computer interaction[J]. Chinese Journal of Ergonomics, 2002, 8(1): 27-33.
- [11] JACOB R J K. What you look at is what you get: eye movement-based interaction techniques[C]//Proc ACM CHI'90 Human Factors in Computing Systems Conference. Seattle, Washington, USA: [s.n.], 1990: 11-18.
- [12] 方志刚. 人机交互中的视线跟踪技术[J]. 人类工效学, 2000, 6(1): 42-45.
FANG Zhi-gang. Eye-gaze technology in human computer interaction[J]. Chinese Journal of Ergonomics, 2000, 6(1): 42-45.
- [13] BOLT R. Put that there: voice and gesture at the graphic interface[J]. Computer Graphics, 1990, 14(3): 262-270.
- [14] SHARMA R, YEASIN M, KRAHNSTÖVER N, et al. Speech-gesture driven multimodal interfaces for crisis management[J]. Proceedings of IEEE (Special Issue on Multimodal Human-Computer Interface), 2003, 91(9): 1327-1354.
- [15] ZHANG Qiao-hui, IMAMIYA A, GO K, et al. A gaze and speech multimodal interface[C]//Proceedings of the 24th International Conference on Distributed Computing Systems Workshops (ICDCSW'04). Tokyo, Japan: [s.n.], 2004: 208-213.
- [16] TAN Y K, SHERKAT N, ALLEN T. Eye gaze and speech for data entry: a comparison of different data entry methods[C]//International Conference of Multimedia and Expo. Baltimore, USA: [s.n.], 2003: 41-44.
- [17] OVIATT S. Multimodal interfaces for dynamic interactive maps[C]//Proceedings of CHI'96 Human Factors in Computing Systems. NY, USA: ACM Press, 1996: 95-102.
- [18] OVIATT S, COHEN P R, WU L, et al. Designing the user interface for multimodal speech and gesture applications: state-of-the-art systems and research directions[C]//Human Computer Interaction in the New Millennium. Boston: Addison-Wesley, 2000: 263-322.
- [19] SEGAWA O, TAKAHASHI M, NISHIMATA N, et al. A study on speech-input interfaces for power control and information system[C]//People in Control: an International Conference on Human Interfaces in Control Rooms, Cockpits and Command Centres, University of Bath. Bath, UK: [s.n.], 1999: 21-23.
- [20] HUBER K. Does speech technology have a place in the control room[C]//Power Engineering Society General Meeting. San Francisco, USA: [s.n.], 2005: 2702-2703.
- [21] KARAT C M, HALVERSON C, HORN D, et al. Patterns of entry and correction in large vocabulary continuous speech recognition systems[C]//Proceedings of the SIGCHI conference on Human Factors in Computing Systems: the CHI is the Limit. Pittsburgh, Pennsylvania, USA: [s.n.], 1999: 568-575.
- [22] FENG J, SEARS A. Are we speaking slower than we type? Exploring the gap between natural speech, typing, and speech-based dictation[J]. Accessibility and Computing, 2004(79): 6-9.
- [23] SCHMANDT C, ACKERMAN M S, HINDUS D. Augmenting a window system with speech input[J]. IEEE Computer Society, 1990, 23(8): 50-56.
- [24] GRASSO M A, EBERT D S, FININ T W. The integrality of speech in multimodal interfaces[J]. ACM Transactions on Computer-Human Interaction, 1998, 5(4): 303-325.
- [25] HEDBERG S R. Dictating this article to my computer: automatic speech recognition is coming of age[J]. IEEE Expert, 1997, 12(6): 9-11.
- [26] OVERBYE T J, WEBER J D. Visualizing the electric grid[J]. IEEE Spectrum, 2001, 38(2): 52-58.
- [27] OVERBYE T J, WEBER J D. New methods for the visualization of electric power system information[C]//IEEE Symposium on Information Visualization. Salt Lake City, UT, USA: [s.n.], 2000: 131-136.
- [28] SUN Y, OVERBYE T J. Visualizations for power system contingency analysis data[J]. IEEE Trans on Power Systems, 2004, 19(4): 1859-1866.
- [29] OVERBYE T J, WEBER J D. Visualization of power system data[C]//Proceedings of the 33rd Hawaii International Conference on System Sciences. Hawaii, USA: [s.n.], 2000: 1228-1234.

(责任编辑: 汪仪珍)

作者简介:

朱全胜(1982-), 男, 河南博爱人, 博士研究生, 研究方向为电力系统人机交互及运行状态可视化(E-mail: zhuquansheng@student.dlut.edu.cn);

刘尧(1967-), 女, 辽宁大连人, 副教授, 主要研究领域为电力市场理论与应用和电力系统运行状态可视化;

李卫东(1964-), 男, 辽宁铁岭人, 教授, 博士, 博士生导师, 从事电力系统分析方面的教学与科研工作, 目前研究领域为互联电网运行控制性能评价、电力系统运行状态可视化及电力市场理论与应用。

HCI and interaction technologies of EMS

ZHU Quan-sheng, LIU Rao, LI Wei-dong

(Department of Electrical and Electronics Engineering, Dalian University
of Technology, Dalian 116024, China)

Abstract: HCI(Human Computer Interface) is an important part of EMS(Energy Management System). Advanced HCI technologies are adopted to improve the input and output of EMS. Speech recognition added at the input side form a multi-mode input interface together with mouse and keyboard, which can shorten the order input time and realize multi order inputs in parallel. Correlated multi-screen display technology introduced at the output side combines 2-D visualization with 3-D visualization. The text-to-speech technology combined with other output modes can present the system information more reasonably and naturally to help operators to control the system operating conditions quickly.

This project is supported by the National Natural Science Foundation of China(50577002).

Key words: human computer interface; EMS; interaction technology; speech recognition; text-to-speech; visualization