**使用眼球输入的人机互动**

THOMAS E. HUTCHINSON, K. PRESTON WHITE, JR., SENIOR MEMBER, IEEE, WORTHY N. MARTIN, KELLY C. REICHERT, AND LISA A. FREY

原文链接[Sci-Hub | Human-computer interaction using eye-gaze input | 10.1109/21.44068 (yncjkj.com)](https://sci-hub.yncjkj.com/10.1109/21.44068)

摘要 -Erica是一个具有独特用户界面的计算机工作站。该工作站配备了成像硬件和软件，能自动记录用户眼睛的数字肖像。根据当前肖像的特征，界面计算出用户的眼球在电脑屏幕上的大致位置。然后，计算机会执行与当前显示在这个屏幕位置的菜单选项相关的命令。通过这种方式，用户可以与计算机进行互动，运行应用软件，并管理外围设备，只需看着屏幕上显示的一系列适当的菜单选项。本文介绍了眼球界面技术，它在Erica中的实现，以及它作为一个假肢设备的应用。

1. 介绍

想象一下，你是一场严重的残废事故的受害者。你不能再移动或说话。你不能写字，不能指点，甚至不能点头。你只用眼睛来交流--上下移动表示 "是"；左右移动表示 "是"。向上和向下表示 "是"；向左和向右表示 "不是"。尽管你的身体有残疾，但你的智力和创造能力仍然没有减弱，然而你的思想却被困在你体内。

这种情况对我们大多数人来说是令人不安的，甚至想象不到。然而，对于成千上万的严重残疾人来说，这是一个活生生的现实，他们是受伤、中风或逐渐丧失能力的疾病的结果。由于没有声音或手势可供支配，这些人缺乏与他人交流的有效手段，无法控制他们的环境，也无法娱乐和丰富他们的思想。

在本文中，我们描述了一种独特的假肢装置，称为眼球反应界面计算机辅助装置（Erica）。Erica是一个独立的工作站，以标准的个人电脑为基础，专门配备了成像硬盘和软件。用户界面直接接受来自人眼的输入。菜单选项显示在计算机显示器的不同位置。只需通过观察一个特定的位置，相应的菜单选项就会被调用。通过这种方式，残疾的Erica用户可以与计算机互动，运行通信和其他应用软件，并管理外围设备。

虽然底层的眼球技术有许多潜在的应用，但Erica项目最初和持久的目标是帮助身体和声音残疾的人达到或恢复某种程度的独立通信和控制。这个群体包括美国大约23.8万名四肢瘫痪的人，他们需要援助来完成日常功能[1]。对于许多至少保留一定程度运动控制能力的四肢瘫痪者来说，Erica是现有假肢通信系统的一个有吸引力的替代方案[2]-[5]。眼睛注视的界面不那么麻烦，而且可能比其他的输入设备，如身体激活的开关和抿嘴的机制更快、更实用。最重要的是，对于大约15万名只能控制眼睛肌肉的严重肌肉功能障碍者来说[1]，眼球注视界面是无可替代的。

Erica是由弗吉尼亚大学工程和应用科学学院的教师和学生设计和开发的。

弗吉尼亚大学工程和应用科学学院的学生设计和开发的[6]。该项目开始于1984年。自开始以来，超过一百名工程专业的本科生和研究生贡献了他们的时间和才能，帮助Erica成为现实。 几年来，Erica一直在实验室和测试场中积极使用。该项目的进展得到了全国媒体的广泛报道，包括电视和印刷媒体[7]。弗吉尼亚州费尔法克斯的LC技术公司目前根据与该大学和弗吉尼亚创新技术中心（CIT）的专利许可协议制造和销售Erica I系统。第一批商用Erica装置于1988年底交付。

在下一节中，我们描述了标准的Erica硬件系统，解释了该系统的工作原理，并介绍了眼球位置检测算法。第三节介绍了系统运行和现有的软件应用。在第四节中，我们讨论了系统开发的一些历史，并描述了选定的测试经验。在最后几节中，我们考虑了当前系统的局限性，并讨论了为克服这些局限性而继续进行的研究和开发。



Figure 1Erica硬件系统的实验室配置

1. 系统描述
2. 标准硬件配置

Erica的硬件系统包括一台带硬盘的个人电脑、一台彩色显示器、一个近红外光源、一个带红外滤光片的光跟踪视频、监控摄像机和一个光强成像板。这个基本装置的实验室配置如图1所示。图中还显示了一台打印机和第二台较小的黑白视频监视器（位于彩色监视器的左边）。 在实验室中用于摄像机-图像反馈

光源是一个砷化镓、近红外、880纳米的发光二极管（LED），通常用于通信系统。图2显示了用于将LED同轴放置在相机镜头前的组件。LED被连接到一个短管的一端。一个球窝接头将这根管子的另一端连接到一个红外通滤波器的中心。滤光片被拧到相机镜头的前面。球窝接头提供了定位LED所需的两度自由，以便它能照亮用户的脸，而滤光片则减少了到达相机传感器的环境光量。此外，一个凝光透镜被安装在一个金属护套上，该护套可沿管子滑动。从镜头到LED的距离可以从0.5到2.0厘米不等，将红外光聚焦在用户的脸上。从图1中可以看出，光源/摄像机组件直接安装在彩色显示器下面。

1. 如何工作

在操作中，光源用无害的近红外光照亮用户的脸，如图3所示。用户注视的方向是由用户的一只眼睛反射出的这种光的摄像机图像来确定的。如图4所示，眼睛图像有两个重要特征，用于眼球位置检测。

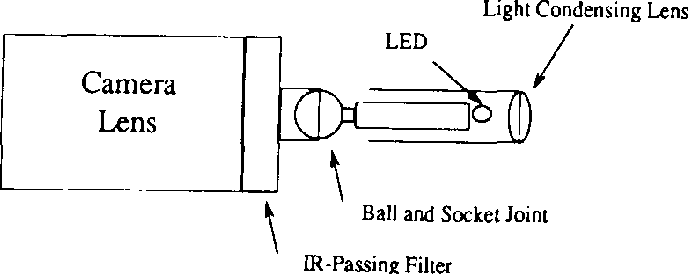


Figure 2Erica光源组件

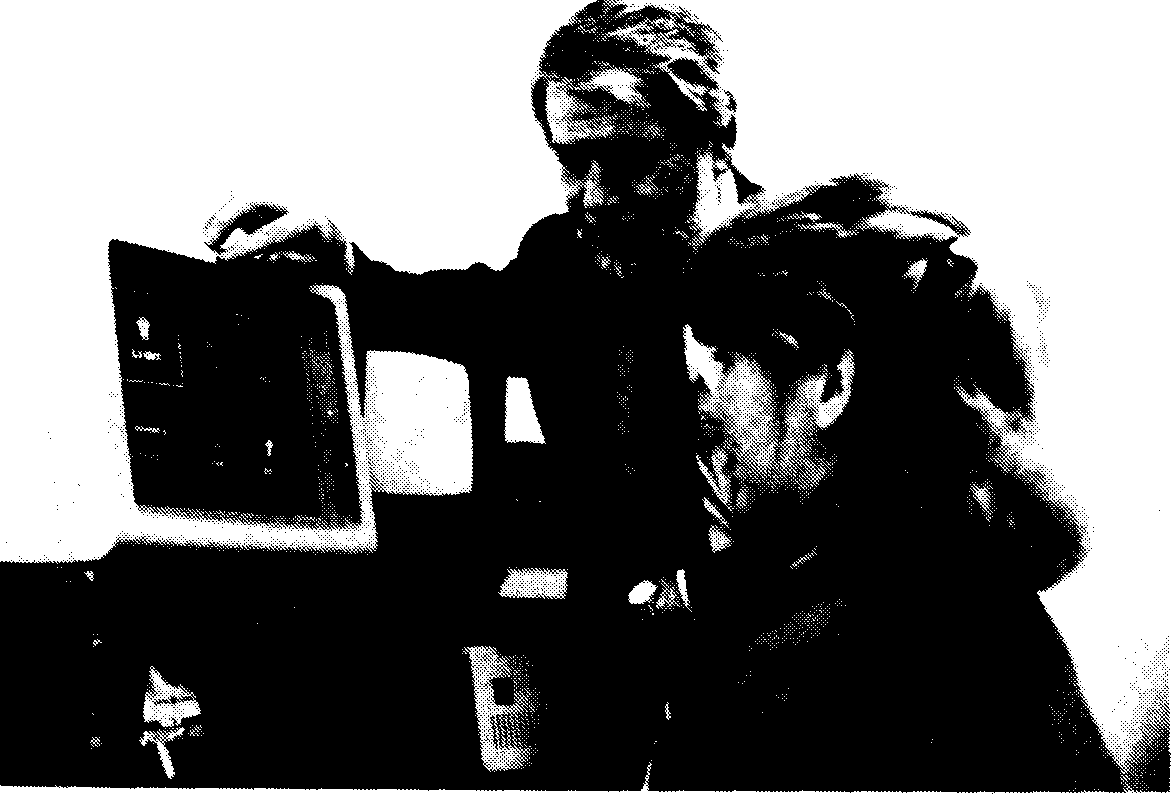
* 闪烁的光芒：一部分红外光被反射到角膜表面。这是LED的第一个Purkinje图像，在相机中显示为一小块红外光的密集区域，称为闪光。只要 只要用户的头部相对于摄像机保持静止，闪光的位置就在图像领域中保持固定。
* 明亮的眼睛：一部分红外光进入瞳孔并从视网膜上反射。这就是瞳孔的图像。称为亮眼（人类视网膜对红外光的反射，类似于夜间猫眼对可见光的反射）。亮眼在相机中显示为一个红外光区域，比闪光大，强度小，但比绕过虹膜的 "黑暗 "图像更强。亮眼的位置在相机的图像领域中移动，跟随眼睛的运动。

用户的注视方向可以从照相机图像中的亮眼和闪光的相对位置来确定。当视线方向与摄像机的光轴重合时（即，当用户直接看向光源时），亮眼的中心与闪光的中心重合。由于光源/相机组件紧挨着计算机显示器的下方，每当用户看向显示屏时，亮眼中心就会相对于闪光点升高。图5显示了一个被照亮的眼睛的例子，图6显示了在显示屏幕上各种眼睛注视位置的亮眼和闪光中心之间的空间关系。

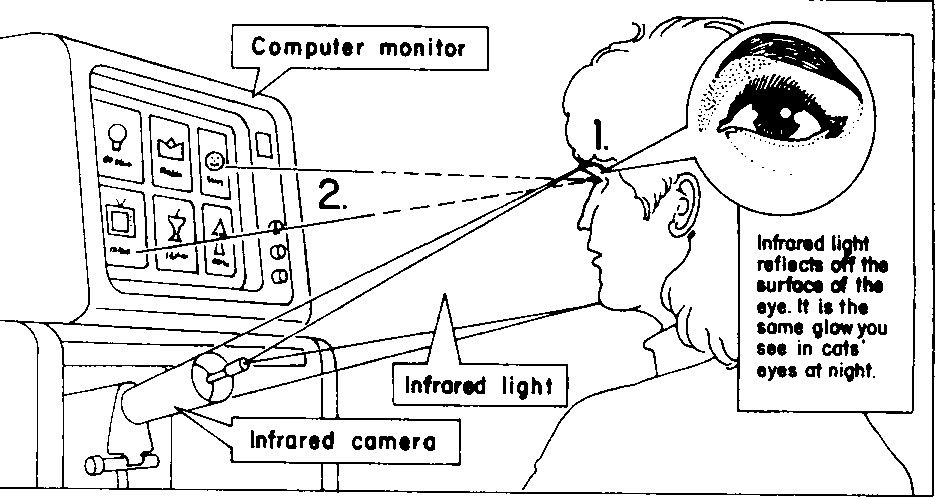
为了测量亮眼中心和闪光中心之间的矢量关系，将眼睛的虚拟图像传给计算机.目前的图像处理板以每秒30帧的速度在512 X 480位的帧中记录摄像机图像。该帧包含眼睛图像中每个像素的数字化光级强度。该帧存储在成像板的物理存储器中，主机可以一次访问该帧的四分之一。图7显示了经成像板和主机软件处理后的数字化图像。

1. 眼球位置检测算法

在目前的发展状态下，Erica可以准确区分多达9个参考区域，或菜单框，在显示器上以3 X 3的矩阵排列。屏幕上以3 X 3矩阵排列的九个参考区域或菜单框。这是通过眼球位置检测算法实现的,该算法1）从代表当前眼球的框架中提取闪光中心和亮眼中心的水平和垂直笛卡尔坐标位置；2）将这些坐标位置映射到与九个菜单框之一相对应的存储参考范围中。一个典型的屏幕布局显示在图8中。



(a)



(b)

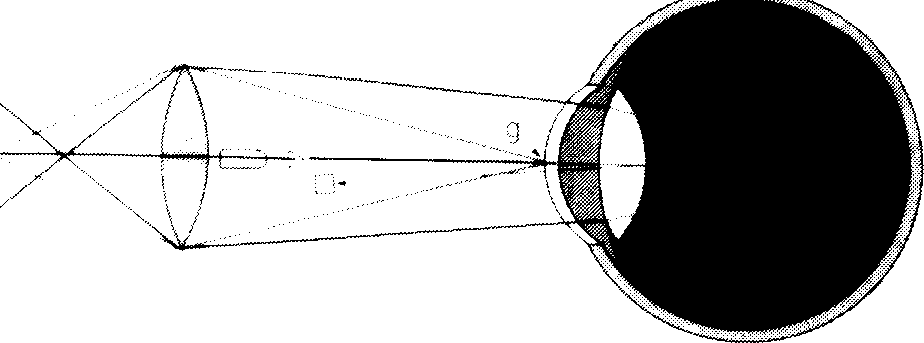
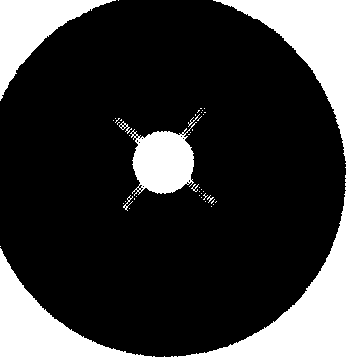
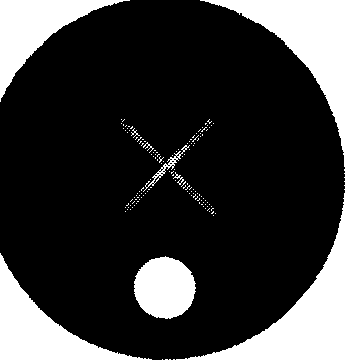
Figure 3 (a) 使用中的Erica系统。(b) 显示眼睛-眼睛界面如何运作的示意图。(1) 光束射向眼睛。反射到相机中的光提醒计算机注意眼睛的位置 (2) 盯着计算机屏幕上显示的一个命令0.5秒或更长时间，自动触发系统。

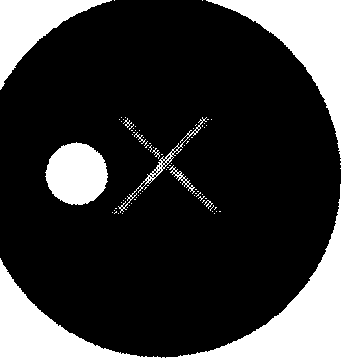
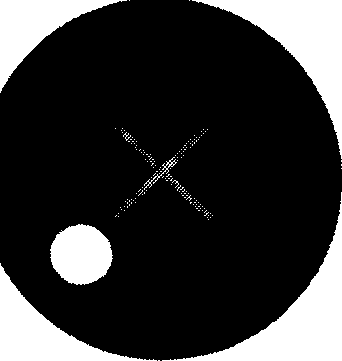
Figure 4显示闪光反射和亮眼效应的示意图。感兴趣的图像特征是。1）眼睛前角膜表面的闪光反射，2）瞳孔的 "亮眼 "图像来自于后角膜表面反射的光。



Figure 5发光的眼睛，显示出闪烁的光芒和明亮的眼睛

(a) (b)

(c) (d)

Figure 6闪光和亮眼中心之间的矢量关系，用于确定目视位置。(a)对准相机。(b)正对相机的上方。(c) 照相机的右边。(d)向上和向右

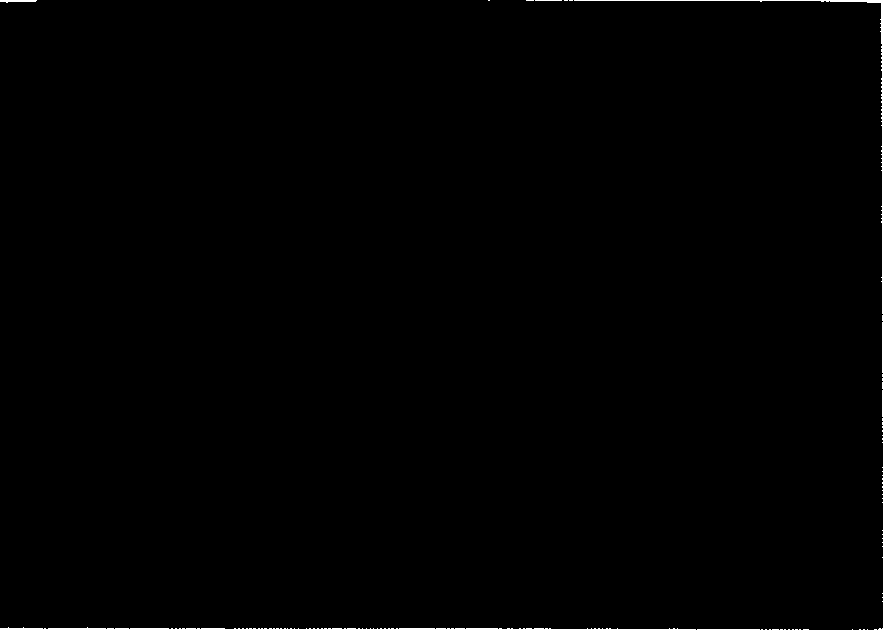


Figure 7在屏幕上显示数字化的眼球图像

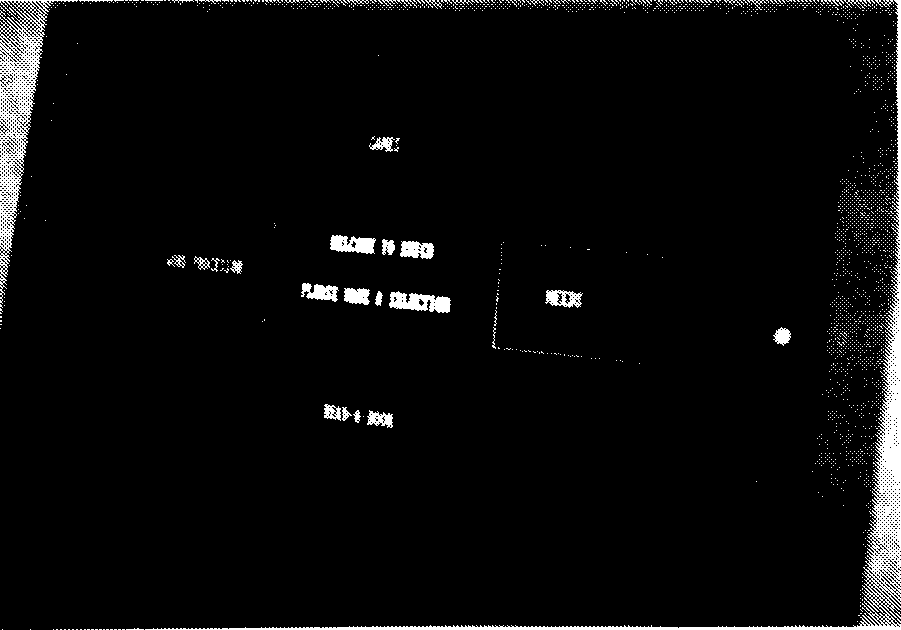


Figure 8 典型的屏幕布局显示四个活动菜单框

在每个Erica会话开始时执行一个校准程序，以确定算法所需的参考值。在校准过程中，一个图标被连续放置在显示屏幕的三个位置，并提示用户用他或她的眼睛注视该图标。第一个图标被放置在屏幕的中心。获得的眼睛图像被用来建立亮眼和闪光的强度阈值。 然后，该图标被移到屏幕的左上角和右下角。在每个角落，记录从闪光点到亮眼中心的矢量距离的X和Y分量，并对若干图像进行平均。对这些平均值的线性回归

对这些平均值进行线性回归，以确定横跨整个显示屏幕的80×25矩阵的凝视位置的参考坐标。然后，这些注视位置与相应的菜单框相关联。整个校准过程通常需要大约5秒钟。

在校准过程中使用的成像算法和随后的操作都是一样的。从画面的第一行开始，大约每20个像素就被检查一次（这个分辨率只比最小的预期瞳孔大小略小），直到找到一个强度大于亮眼阈值的像素。这个像素被用来在画面中定义一个较小的区域，包含亮眼和闪光。然后，这个较小的区域被逐个像素地重新切割，以确定：1）亮眼圆周上的像素（即在有低强度邻居的边缘上）和2）代表闪光的帧内像素（通常是5至8个非常高强度的像素）。如果不能检测到闪光点，就抓取一个新的画面，重复这一过程。反复检测不到闪光点会自动导致重新校准，在指定的时间延迟后进行。

1. 操作和当前应用

Erica目前使用树状结构的菜单层次进行操作。菜单选项出现在一到九个菜单框中。用户通过盯着所需的选项看一小段时间来做出选择。 这个时间通常预设为两到三秒，但可以根据用户的经验、技能和应用情况来改变。

当用户的视线固定在这段时间内，就会发出提示音，并在与视线一致的菜单框中出现一个图标（光标）。如果用户继续盯着这个启用的选项，就会发出第二声提示音，这个选项就会执行。听觉和视觉反馈的目的是让用户有时间通过相应地改变他或她的视线来改变或放弃已启用的选项。大多数菜单包含一个 "备份 "选项，允许用户在需要时返回到前一个菜单。

在完成初始校准程序后，屏幕上会出现一个主菜单。使用这个菜单，用户可以选择一个应用领域。目前的Erica软件组合包括四个应用套件--控制，包括环境控制和个人需求的非发声通信；通信，包括文字处理和合成语音；娱乐，包括计算机游戏、数字化音乐和教育项目；以及文本阅读，包括一个小型书籍和其他文本图书馆。这些应用将在以下几个小节中描述。

1. 控制

控制应用允许用户通过一个X-10动力室系统来操作周围的电气设备。这个商业系统由一个编码器单元和一套电器模块组成。编码器和电器模块都插在同一交流电路的标准交流插座上。 每个需要控制的电气设备都被插入一个电器模块，该模块已经为该设备预设了一个独特的数字代码。当用户从Erica菜单中选择控制设备的命令时，应用软件通过计算机串行端口将命令信号传递给X-10编码器。编码器反过来通过交流电路发送适当的代码。然后，设备模块改变所连接设备的状态。通过这种方式，电器可以被打开和关闭，灯光可以被提高和调暗。

标准的X-10系统使用双态逻辑信号来控制电器的电源。不能以这种方式控制的电器功能，则使用专门适应的遥控外围设备。例如，为了选择电视频道，设计了一块PC扩展板，它具有向可编程遥控装置发送指令所需的寻址逻辑。该单元将指令转化为用于录像机（VCR）频道选择的红外代码，从而改变电视频道。这种控制策略也被成功地适用于装有遥控器的立体声音响。

控制应用还包括对频繁的个人需求进行紧急、非语言交流的能力。例如，在紧急情况下，一个标有 "呼叫护士 "的菜单选项会激活一个响亮的蜂鸣器。类似的选项使用户能够交流口渴、疼痛或瘙痒，并指定与疼痛或瘙痒有关的身体区域。

1. 交流

通信应用包含了几个功能，使残疾人能够通过语言进行交流。这套软件的核心程序是一个功能齐全的文字处理器，旨在模拟标准文字处理软件。它包括大多数常见的打字、编辑、文件创建和打印功能。此外，文本文件可以通过语音合成器转换为口头交流。

Erica文字处理器的主要区别是，触摸打字被眼睛打字所取代，即键盘输入被眼球输入所取代。显然，眼球位置检测的有限分辨率使得文字处理器无法使用标准键盘的全屏图像。与标准AT键盘上的84个键相比，目前的文字处理器只有6个 "键"（菜单框）可用于字符输入和编辑功能（最下面一排的菜单框已被禁用，以便在输入和编辑时在屏幕上提供显示文本的空间）。字符输入是通过一个树形结构的菜单层次来完成的。在第一个菜单上，可用的字符集被分割成大小大致相同的子集。连续的菜单将最初选择的子集扩展开来，直到每个菜单框最终与一个独特的按键相关。输入一个字符需要两个、三个、甚至多达四个菜单选择。

直接用眼睛打字是一件相对缓慢的事情，一个有经验的用户使用静态的菜单层次结构，一次输入一个字符，可能需要将近85分钟的时间来完成整个页面的文字。为了加快文本创作的进程，文字处理器中加入了几个独特的功能。首先，有一套常用的短语，可以使用文字处理机的菜单层次选择和输入短语。用户可以编辑和定制短语文件，并且可以根据交流的内容使用备用的短语文件。第二，开发并实施了一种字符预测算法，该算法根据前面两个字符条目，动态地改变根文本输入菜单上的字符选项。这个方案利用了英语文本字符串的条件概率结构（马尔可夫链），将文本输入时间平均减少了30%[8]。目前正在研究更高级别的预测方案、人工智能和其他概念，以进一步提高眼球打字的便利性、准确性和速度。

1. 娱乐

娱乐套件包括一些标准游戏，如21点、万事通和西蒙，这些游戏经过重写以允许眼球操作。许多娱乐程序是以教育为导向的，有一个应用程序甚至允许用户创作音乐，并通过计算机扬声器播放作品。这些程序本身作为娱乐就很有价值。此外，我们还发现，游戏提供了一种简单、无畏和令人信服的手段，向新手介绍Erica操作。

1. 文字阅读

文本文件可以用文字处理器，或一个专门设计的阅读应用程序来阅读。这个应用程序的根菜单提供了一个存储在磁盘上的文本文件的主题索引。选择一个主题类别就会出现该区域的文件名列表。该列表由一个光标来索引。菜单选项滚动光标，或选择当前文件进行阅读。

这种只读应用程序的主要优点是，计算机屏幕上有更多的空间可用于显示文本。当选择一个文件进行阅读时，只有最下面一排的菜单框被启用。有两个方框用于翻页，向前和向后翻页，第三个方框用于调用一个子菜单。子菜单的选项允许用户在当前页面上放置一个书签，选择一个备用文本，并退出应用程序。当前图书馆中的文本资料的数量是有限的，既受可用磁盘空间的限制，也受在磁盘上获取文本的费用限制。

1. 一些测试经验

Erica的整体设计和开发遵循了众所周知的系统工程方法--确定问题、评估需求和目标、从现有的备选方案中选择最合适的设计，以及实施。 该项目借鉴了广泛领域的知识--光学、图像处理、数字逻辑、设计、软件工程、人因工程、康复工程、机械设计、项目管理和市场营销。然而，与大多数系统工程项目一样，方法的应用和这些知识领域的综合是通过一个反复的、进化的和启发性的发展过程完成的。最初的目标反映了设计者对健全人的预想，而对残疾人用户的困境没有什么背景知识或经验。在系统开发和测试的过程中，这些目标经常被改变，并发现新的目标。这个过程还在继续，因为它必须继续。在这一节中，我们认为应该提供一个设计过程的简单例子，与一些开发历史和一个具体的测试经验有关。用户界面的设计开始于对文献的回顾以及对特殊教育和医学专家和从业人员的访谈。这项研究的目的是建立一个残疾人和非语言使用者的档案，从中我们可以了解严重残疾的人的需求、愿望、能力和限制。有了这些资料，我们开始设计一个具有环境控制和文本阅读能力的系统原型，在实验室里，在眼球输入的限制下工作。在本科生中，身体健康的测试对象很多，而且很容易找到。这些测试对象对原型设计的性能提供了反馈。这些反馈是一致的。

然而，向从事残疾人工作的专家展示原型，以及随后与几个残疾人一起工作，得到的评价却不那么积极。对专家们来说，明显的是对一个能够进行交流的系统的迫切需求，但对项目组或有能力的对象来说却不是。主动的环境控制功能是令人印象深刻的技术成就，无疑是有用的，但严重的残疾人周围通常都有保健专业人员，他们经常以极大的能力为残疾人执行这些功能。这种关系中最困难和乏味的方面是对话，在对话中，残疾人可以交流需求、愿望以及对当前状况的满意或不满意的感觉。由于原型缺乏足够的沟通能力，我们亲身经历了这一事实，作为我们自己与残疾人志愿者合作的困难的一部分。这使我们回到了实验室，立即设计了一个简单的文字处理器。

第一个测试在1986年秋季开始。测试对象是史蒂文-麦克唐纳警官，一位受过大学教育的纽约市警察，他在前年夏天执勤时被枪击。他的C-2椎体所受的伤使他完全瘫痪，而且由于他依赖呼吸器，几乎完全不能说话。经过几次由是和非问题组成的访谈，麦克唐纳能说出答案，我们能够评估他的具体需求并为他提供一个原型系统。正如我们所希望和预料的那样，麦克唐纳警官的评估、建议和特殊见解是我们的主要来源。

正如我们所希望和预期的那样，麦克唐纳警官的评估、建议和特殊见解是应用软件设计改进的一个主要来源--这是测试的主要目的。

然而，在第一次测试中遇到的许多意想不到的困难中，包括对硬件重新设计的迫切需求。具体来说，计算机显示器和光源/摄像机组件的标准安装系统，在实验室和轮椅上的病人身上使用很成功，但对于下肢瘫痪的病人来说却不适应。最初为麦克唐纳开发的安装系统也无法使用，因为它干扰了医生和护士对病人的护理。由于病人床铺的特殊设计（可以减少褥疮的可能性）；由于需要保护病人、他的护理人员和他的访客（以及Erica原型），防止在经常拥挤的医院房间里发生意外的破坏，以及由于我们的 "首要指令 "是生产一个负担得起的系统，所以功能支架的机械设计更加复杂。解决这个机械设计问题使软件测试推迟了几个月。

1. 系统的局限性和未来的改进

尽管第一代Erica系统的数量不多，但正在进行的研究和开发为第二代设计的重大功能改进提供了希望。也许目前技术最重要的限制是关于亮眼效应。这种效应的强度因受试者而异，并不是所有申请该系统的人都有足够强烈的亮眼，以允许对他们的眼睛注视方向进行一致和可靠的检测。实验室研究和测试表明，这可能会在不同程度上抑制5%至10%的人使用目前的系统，至少在目前采用的近红外光频率下是如此。 这种变异性的来源还不是很清楚，可能与遗传或环境因素有关，或者是这些因素的某种组合。 为了提高我们对亮眼效应的认识，并开发克服这一基本问题的设计方案，已经进行了大量的研究。

为了操作Erica，使用者必须将他或她的头保持在一个几乎固定的位置。头部在任何一个方向上的横向移动超过两英寸，就会使眼睛的图像离开摄像机的视野；朝向或远离摄像机的移动超过几英寸，就会使眼睛的图像失去焦点。可悲的是，这对许多目标人群来说不是一个问题。然而，患有脑瘫和类似疾病的病人，其头部运动不受控制，目前抑制了他们对该系统的使用。正在评估能使埃里卡能够为这些人所用的硬件替代品，如头部追踪系统和自动对焦镜头。

菜单系统的操作可以通过增加菜单层次结构中任何给定级别的可用选项数量来加快。目前正在进行研究，以提高检测眼球位置的总体准确性和精确度。这将增加屏幕上允许的菜单框的密度。正在研究的还有更多利用整个视野的策略，超出当前显示器的边界，以支持额外的静态和/或动态菜单选项。目前还在研究菜单框选择机制的各种替代方案，以及为冗长和常用的选项序列开发动态和上下文敏感的菜单宏。

交流是眼球界面最重要的假体应用，研究导致文字处理器的进一步增强是该项目的主要兴趣。除了目前正在运行的字母预测方案外，还有几个设计概念似乎很有前途。这些概念利用了我们对单词语境和高层次英语语言结构的了解。正在调查的替代方案包括使用单词和短语预测、非连续字符输入、速记符号和象形文字、双显示器硬件配置、替代的挑选机制和菜单结构、有限和特殊用途的通信世界以及学习系统。正在调查的还有将文字处理器作为键盘模拟器使用，以允许使用不是专门为眼球计算设计的商业软件。

我们在本文中的主要重点是为残疾人社区的利益应用眼球技术。在这个应用领域中，正在开发的其他改进措施包括为残疾学龄儿童提供的教育工具、允许在家中重新配置控制应用套件的设置程序，以及用于移动机器人的目视控制系统。显然，眼球界面并不限于这里描述的假肢应用。在许多正在积极考虑的领域中，包括在医疗、教育、商业、工业和军事环境中使用眼球位置检测，如测试、训练、瞄准、数据输入和控制等应用。鉴于过去几年眼球界面的快速发展，我们乐观地期待着这项技术在未来十年内推广到通用的计算环境中。

鸣谢

特别感谢Janine M. Carley, E. Marshall Newton, IV, Kevin S. Spetz, 和Lisa A. Onufrak对本文初稿提出的许多有益意见。我们要感谢这些学生，以及弗吉尼亚大学过去和现在的许多研究生和本科生，他们为Erica的实现贡献了自己的时间和才能。我们还要感谢许多残疾人志愿者，特别是史蒂文-麦克唐纳警官和他的家人，以及杰米-米切尔，他们的力量、勇气和观点激发并加强了Erica的研究和开发。最后，我们要感谢大学内外的许多顾问，他们在不同时期为该项目捐赠了他们的专业知识和/或物质支持。

参考文献

[1] B.T.Pyle.*An Analysis of the High-Quadriplegics in the United States for use in Erica Marketing Slrategies*. Undergraduate Thesis.Univ Virginia.Charlottesville,VA.1987

[2] Aesir Software Engineering.AA RON—*A Communication Aid*, Pinedale. CA, 1984.

[3] L. Cory, P.H. Viall. and R. Walder. *A Versatile Communications System for High Quadriplegics*.Southern Massuchusetts Univ .1985

[4]Sentient Systems Technology, *EyeTyper—Model 300*,Pittsburgh PA. 1986.

[5] Word+，Inc.. *Equalizer and Equalizer II*, Sunnyvale. CA. 1986

[6] K. C. Reichert. *ERICA Hardware and Software Design*, M.S. Thesis, Univ. Virginia. Charlottesville. VA. 1987

[7]"*With Thomas Hutchinson's marvelous ERICA, a flick of an eye brings help to the helpless*," People, July 20, 1987, p. 87

[8]L. A. Frey. *Preliminary Design of an Intelligent Word Processor for Use in ERICA*. M.S. Thesis, Univ. Virginia, Charlottesville, VA, 1988.