

基于角膜反射的非侵入式视线跟踪技术

高迪, 印桂生, 马春光

(哈尔滨工程大学 计算机科学与技术学院, 黑龙江 哈尔滨 150001)

摘要: 针对现有的视线跟踪技术设备复杂、对姿势限制过多、要头部佩戴专用设备等方面的不足, 提出一种实现视线跟踪技术的新方法, 采用 5 个近红外光灯作为角膜反射的光源。该方法不需要带任何头戴设备, 并且能适应头部的自然运动, 设计的算法计算出来的视线位置精度较高。提出一种新的瞳孔边缘拟合方案, 循环拟合剔除假点, 最终得到精确的瞳孔中心, 提高了映射的精度。利用图像的灰度进行人眼的初定位, 快速而准确。利用交比不变原理进行坐标映射, 精确计算出注视坐标。

关键词: 视线跟踪; 角膜反射; 瞳孔边缘拟合; 交比

中图分类号: TP391

文献标识码: B

文章编号: 1000-436X(2012)12-0133-07

Non-invasive eye tracking technology based on corneal reflex

GAO Di, YIN Gui-sheng, MA Chun-guang

(College of Computer Science and Technology, Harbin Engineering University, Harbin 150001, China)

Abstract: For the deficiencies that existing eye tracking devices are complex, have too many restrictions on the position, need to wear special equipment. A novel method of eye tracking technology using the five near-infrared light as the source of corneal reflex was put forward. This method did not need any wearing equipment, and could adapt to the natural head movement, the algorithm designed calculates the position of sight in high accuracy. A novel method of pupil edge fitting was put forward: fitted circularly and removed false points to get accurate position of pupil center, improved mapping accuracy. Used the GrayScale of image to initialize eye position quickly and accurately. Using the theory of cross-ratio invariance to make coordinate mapping, finally calculated the attention coordinate accurately.

Key words: eye tracking; corneal reflex; pupil edge fitting; cross-ratio

1 引言

人们获得的信息绝大部分是通过眼睛来获取的, 正因为通过视觉获取的信息对人们的生活有重要的意义, 对视觉方面的研究显得至关重要。近年来, 随着科技的飞速发展, 视觉跟踪日趋成为国内外专家学者研究的焦点, 然而, 由于起步较晚和技术支持的局限性, 这方面的研究仍处于萌芽状态, 没有一种理想的方案能将其广泛地应用到日常生活中。

视觉跟踪技术按其所借助的媒介分为以硬件

为基础和以软件为基础 2 种。视线跟踪装置有强迫式与非强迫式、穿戴式与非穿戴式、接触式与非接触式之分^[1]。

近年来, 人们为克服视线跟踪装置对人的干扰, 提出了用软件实现^[2]的对用户无干扰的视线跟踪方法。其基本工作原理是利用图像处理技术, 使用摄像机捕获人眼的图像, 对人眼的图像进行处理, 然后用软件实现图像中人眼的定位与跟踪, 能够精确检测和提取眼动位置信息, 通过分析计算出视线的注视坐标。

目前, 基于角膜反射原理的视线跟踪技术已经

有一些研究。但是对于瞳孔边缘的拟合、瞳孔中心定位、人眼注视点的计算是技术的核心, 这些问题没有精确的方法。在现有的一些研究(如刘瑞安等的研究^[3])中采取了随机椭圆拟合来进行瞳孔边缘的提取, 但是图像中会有噪点, 这样可能会把噪点也当作椭圆有效值来进行拟合了, 这样得到的结果不准确, 不能够精确定位瞳孔中心。采用卡尔曼滤波法^[4]来利用原有的数据估算将来的数据, 由于将卡尔曼滤波用于运动变化快的目标时, 滤波器所得的状态不能马上跟踪真实状态的变化, 从而产生较大的偏差。

视线跟踪技术的应用范围极其广泛。视线跟踪技术的实现是为了向视觉控制的拓展, 只要涉及到人机交互的领域, 都可以用视觉控制来代替手工操作。而随着 IT 业的不断发展, 人机交互已经涉及到了人们生活的方方面面, 因此视觉跟踪技术能够渗透到人生活的方方面面, 视线跟踪技术的普及将会使社会的发展产生质的飞跃。如果视线跟踪技术在实际生活中得到广泛应用, 不仅能够提高人们的生活质量, 还能够为国防的现代化建设做出贡献。例如在虚拟现实、娱乐、军事、残疾人辅助、医疗、航天等方面的应用有着重要的意义。对视觉控制技术进行研究, 进而开发出各种个性化的视觉控制设备, 将各个领域进行拓展。视觉控制技术和数字人像技术、生物识别技术紧密相关, 在公共安全、经济和生活中具有广阔的应用前景。

本文在前人已有基础上提出了一种基于角膜反射的非侵入式视线跟踪技术, 不需要头戴任何设备, 而且能适应头部的自然运动; 提出了一种新的更精确的瞳孔边界的拟合方法, 进行循环椭圆拟合, 剔除假点, 最后拟合出精确的瞳孔边缘; 采用 5 个线外光源, 进行精确参数获取, 利用射影几何原理进行空间映射, 提高了视线定位的精度。

2 系统描述

在进行人眼运动捕捉的过程中, 系统需要在三维空间坐标系和平面坐标系之间进行转换和映射, 进而来获取准确的人眼注视坐标位置。首先, 调整和设定各项参数; 然后, 图像采集卡采集人眼的图像, 将图像送至 PC 进行处理, 计算出人眼的瞳孔中心位置; 之后, 利用单摄像机、显示屏幕、人眼及红外光源这几个空间坐标映射算出人眼注视屏

幕的坐标, 设定好的屏幕按钮根据注视坐标进行响应。实验系统如图 1 所示。

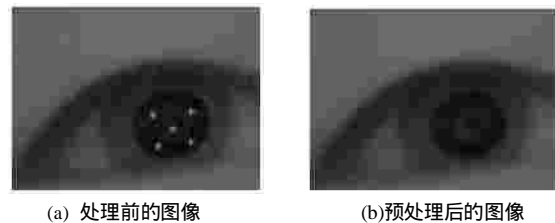


图 1 ROI 图像处理

3 视线跟踪方案

3.1 寻找眼睛感兴趣区域 ROI

眼睛特征的快速提取是进行下一步工作的基础。本研究在实验的基础上提出了一种新的快速检测方法。

眼睛不同部位对红外光的反射和折射率不同, 瞳孔对红外光的反射最弱, 因而瞳孔在红外光的照射下呈黑色, 而其他部分都不是黑色。利用这一点, 可以快速定位到人的眼睛, 这样可以节省时间和计算开销, 同时精度也很高。

采用经典的阈值技术处理灰度图像。利用所有小于阈值的点 $(x_i, y_i) (i=1, \dots, n)$ 在二值化图像粗定位瞳孔中心位置。

$$x_{\text{pupil}} = \frac{1}{N} \sum_{n=1}^N x_n \quad (1)$$

$$y_{\text{pupil}} = \frac{1}{N} \sum_{n=1}^N y_n \quad (2)$$

根据粗定的瞳孔中心, 把以它为中心的 60×60 pixel 的矩形区域设置为感兴趣区域 ROI(region of interest), 这样就缩小图像处理的范围, 仅对感兴趣区域进行操作, 提高了图像处理效率。

3.2 计算光斑坐标

要计算出精确的注视点位置, 必须首先获取红外灯在角膜中反射的 5 个光斑的坐标。因而快速准确地获取光斑坐标是至关重要的。由于图像噪声的存在和头部运动的影响, 光斑的强度可能不稳定, 有时会变得很暗淡, 这就增加了寻找光斑的难度, 为此本文提出了一种健壮快速的光斑提取方法。

1) 首先根据经验阈值分割图像, 由于光斑的灰度值在图像中是最高的, 因而用适当的阈值处理可以获取光斑的信息。然后找出面积最大的 5 个光斑。

2) 利用质心法计算出它们的中心坐标, 同时利

用光斑的相对位置进行检测，剔除假的光斑信息。

3) 如果存在没有找到的光斑点, 则用找到的光斑坐标预测没有找到的光斑位置, 同时减小阈值, 在预测区域搜索没有找到的光斑。重复进行直到找到光斑为止。

3.3 瞳孔中心精确定位

3.3.1 图像预处理

由于噪声和光斑的影响以及上眼睑和眉毛的遮挡, 瞳孔边缘的精确提取变得十分困难, 因此减少噪声对图像的干扰就显得十分重要。

首先根据找到的光斑的坐标, 对其进行涂黑, 这样就是减少了光斑对瞳孔边缘的影响, 然后用中值滤波对图像进行平滑, 进一步减少光斑和噪声的影响。由于中值滤波^[5]运算量大, 所以笔者只对 ROI 进行处理, 以减少运算量, 达到实时性的要求。处理后的图像如图 1(b) 所示。

3.3.2 瞳孔阈值提取

阈值的选取直接影响瞳孔定位的精度, 由于眼睛的成像图像对红外光线比较敏感, 当光强较强时, 图像的整体灰度值增加, 反之则减小。如果设置固定阈值, 那么当图像的明暗程度随光线条件变化时, 就不能很好地适应自然条件并使瞳孔的定位产生很大的误差, 所以为保证瞳孔定位的准确性, 要使用自适应阈值。通过分析眼睛的直方图来获取瞳孔的阈值, 这样可以保证每幅图像中的阈值都合适地自适应阈值。

图像直方图如图 2 所示, 对于灰度图像, 可以返回每个灰度上的像素点数目, 由于图像中存在噪声, 需要用式 (3) 对直方图数据进行相邻 3 点平均低通滤波, 这样才能得到比较平滑的直方图以供分析。

$$y_s = (y(n-1) + y(n) + y(n+1))/3 \quad (3)$$

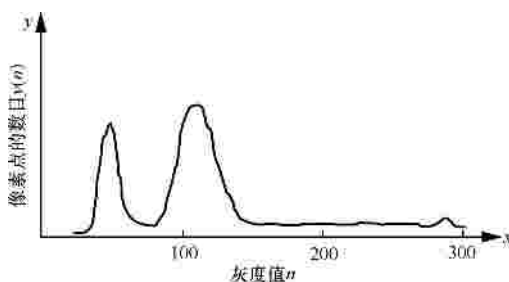


图 2 眼睛图像直方示意

左边波峰对应瞳孔, 往右依次对应虹膜及光斑, 通过寻找波谷提取自适应阈值。

3.3.3 瞳孔边界点提取

用简单的阈值分割技术不能很好地获取瞳孔的边界点, 因为精确的分割对阈值十分敏感, 并且角膜上光斑的影响增加了处理的难度。由于这些因素的影响, 用阈值分割出来的瞳孔轮廓会和真实的大小有所差距。

现在利用卷积算子^[6]的边界检测算法(如 Sobel^[7]和 Canny^[8]算子)并不适合用来进行瞳孔边界点的提取。原因是实际瞳孔图像中存在噪点, 利用卷积算子的算法会把瞳孔实际边缘内圈的噪点当做边界点进行计算, 这样就不能够正确地提取瞳孔边界点, 自然也就不能进行有效的椭圆拟合。

本文中瞳孔的边界点的提取采用阈值技术, 通过上面得到的阈值对图像进行处理。首先利用阈值简单计算出瞳孔的大致中心位置, 然后对瞳孔的图像利用一维边缘检测算子^[11000]^[9]进行扫描, 进而获取瞳孔边缘的候选点。

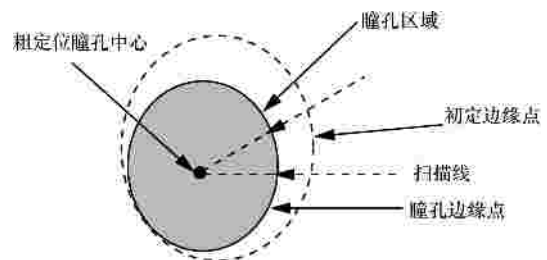


图 3 瞳孔像素级边缘点提取示意

3.3.4 椭圆拟合瞳孔边缘

本文提出了一种新的边界点提取方案。对边界点提取中的所有候选点进行多次的椭圆拟合, 采用直接最小二乘椭圆拟合法^[10]来进行椭圆拟合, 并剔除与椭圆轮廓较远的点, 重复进行直到获取的椭圆中心稳定, 这样求出的椭圆中心就是最终用到的瞳孔中心。

3.4 坐标映射

根据计算出的屏幕上 4 个光斑和计算出的瞳孔中心坐标, 利用射影几何原理进行空间坐标的映射。应用交比不变原理^[11]进行计算。计算模型如图 4 所示。

摄像机图像平面上的点由于和眼睛上的光斑存在空间映射关系, 这些点不能直接通过在图像上建立坐标系来求得。经眼球中心将眼角膜上的光斑通过晶状体折射到与摄像机光轴垂直的切平面上, 如图 5 所示, r_1 和 r_2 分别对应折射后的投射点 v_1 和 v_2 。

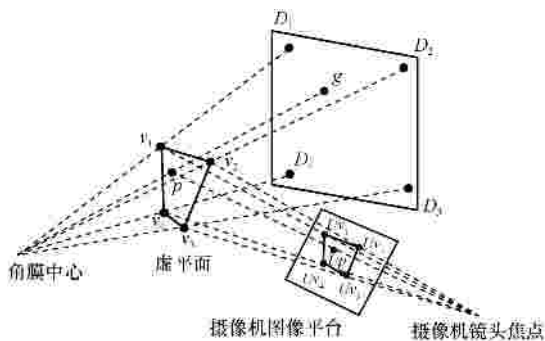


图 4 空间映射关系

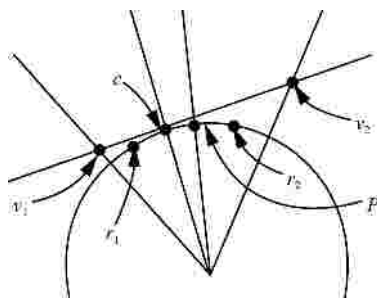


图 5 实际光斑与虚点的关系

这些折射后的共面点与实际的光斑位置是有差别的，实际计算中的空间坐标投影是基于四点共面来进行的。将这个投射出来的切平面称为虚平面，图 4 中 v_1 、 v_2 、 v_3 、 v_4 就是屏幕上 4 个光斑投影到虚平面上的虚点。

虚平面上的点投影到摄像机图像平面也同样是虚点，图 4 中的 Uv_1 、 Uv_2 、 Uv_3 、 Uv_4 就是虚点。这些点是不能在图像中直接找到的。实际的人眼角膜反射光斑在捕获到图像中的光斑是实点，摄像机上方光源在人眼角膜上产生的光斑是 r_0 ，屏幕上 4 个光源在人眼的光斑是 r_1 、 r_2 、 r_3 、 r_4 。这 5 个光斑在图像中对应的实点是 I_0 、 I_1 、 I_2 、 I_3 、 I_4 ， I_0 、 I_1 、 I_2 、 I_3 、 I_4 可以通过计算得出， I_0 、 I_1 、 I_2 、 I_3 、 I_4 和虚点 Uv_1 、 Uv_2 、 Uv_3 、 Uv_4 之间存在一个转换关系^[12]，如图 6 所示，图 6 中，白色的是实点，灰色的为虚点。

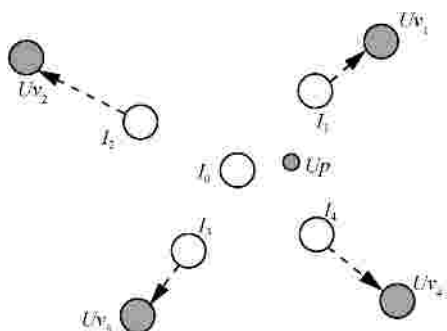


图 6 摄像机平面上的虚点与实点的计算关系

$$Uv_1 = I_0 + a_1(I_1 - I_0) \quad (4)$$

其中， a_1 是固定的， I_0 和 I_1 可以通过计算得出，只要 Uv_1 已知， a_1 就可求。 g 为人眼注视屏幕的点， P 点为瞳孔中心在虚平面上的点， U_p 为 P 投影到摄像机图像平面的点。当注视 D_1 时， $Uv_1 = U_p = I_0 + a_1(I_1 - I_0)$ 。 U_p 就是图像中计算出来的图像上的瞳孔中心。这样就可求出 a_1 ，同理可计算出其他 3 个方向上的 a_2 、 a_3 、 a_4 。虚点 Uv_1 、 Uv_2 、 Uv_3 、 Uv_4 的坐标就可计算出来。

根据交比不变原理，以屏幕上 D_1D_4 的交比为例， D_1D_4 的交比和 V_1V_4 的交比相等，并且 V_1V_4 的交比和 Uv_1Uv_4 的交比相等，则 D_1D_4 的交比和 Uv_1Uv_4 的交比相等。摄像机图像平面上各点的坐标可以计算出来，等价于 D_1D_4 的交比可以计算出来了。只要在计算过程中把未知点 g 用上，就可以通过公式变换把 g 求出来。对应关系模型如图 7 所示。

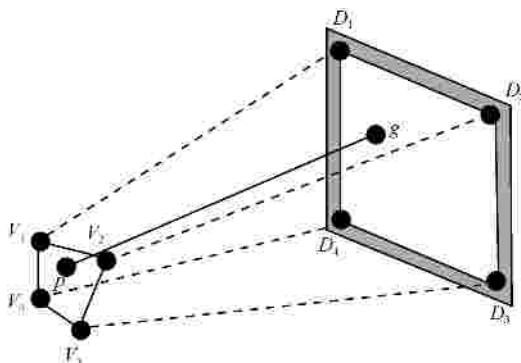


图 7 角膜切面上虚点与屏幕光源及注视点间的关系示意

用交比的方法对眼睛注视点进行估算。图 8 和图 9 显示了怎么计算交比， E 是 Uv_1Uv_3 和 Uv_2Uv_4 的交点。先计算出 Uv_1Uv_2 和 Uv_3Uv_4 延长线的交点 N_2 ，从 N_2 出发过 U_p 和 E 的直线交 Uv_1Uv_2 于 M_1 和 M_2 。

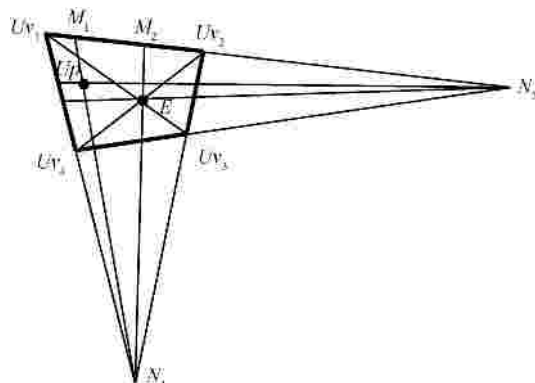


图 8 摄像机图像平面虚点的交比计算

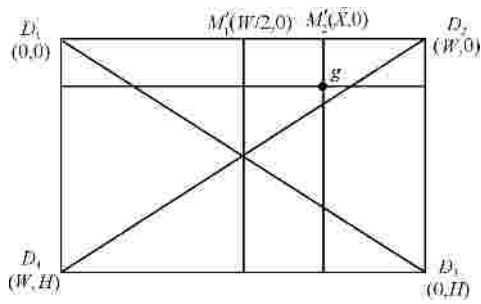


图 9 屏幕交比计算

直线 Uv_1Uv_2 上 4 点 $Uv_1(x_1,y_1)$ 、 $M_1(x_2,y_2)$ 、 $M_2(x_3,y_3)$ 、 $Uv_2(x_4,y_4)$ 4 个点的交比 CR_x 满足下面等式：

$$CR_x = \frac{(x_1y_2 - x_2y_1)(x_3y_4 - x_4y_3)}{(x_1y_3 - x_3y_1)(x_2y_4 - x_4y_2)} \quad (5)$$

同理，假设 g 的坐标为 $g(\hat{x}, \hat{y})$ ，则屏幕上的 4 个点 $D_1(0,0)$ 、 $M_1'(w/2,0)$ 、 $M_2'(\hat{x},0)$ 、 $D_2(w,0)$ 的交比满足式(6)，即

$$CR_x = \frac{(w - \frac{w}{2})\hat{x}}{(w - \hat{x})\frac{w}{2}} = \frac{\hat{x}}{w - \hat{x}} \quad (6)$$

由于式(5)中的值都可计算出来，这样就可以求出人眼注视屏幕点的 \hat{x} 坐标，同理可以求得 \hat{y} 的坐标。这样屏幕注视点的坐标 $g(\hat{x}, \hat{y})$ 可被计算出来。

4 实验与分析

本文实现了一个视觉控制鼠标的系统，所用操作系统为 Windows XP 系统，采用 VC++ 进行实现。系统实现后，在一般的实验室办公环境下进行测试，人以正常操作电脑的习惯坐在电脑面前，屏幕尺寸为 17 英寸，分辨率为 1 280×1 024pixel，在屏幕上确定 49 个十字叉标定的注视位置，屏幕可显示测试者视线定位的实际结果，而且当在以十字叉中心为中心的 200×120pixel 范围内，原地点击鼠标，该范围会以按钮形式响应，产生一次闪烁。对不同的测试者首先提取其特征参数，并进行校验，然后进行视线定位实验。

4.1 校验实验

对 5 名实验者进行实验。结果如图 10 所示。

对每一位实验者，首先对图像进行处理，得出实验者的 5 个光斑坐标，实验过程中将计算出的光斑位置显示地标定出来，以保证光斑获取无误，运用交比原理对其进行处理，计算实验者的 a 值并输出，实验结果如表 1 所示。

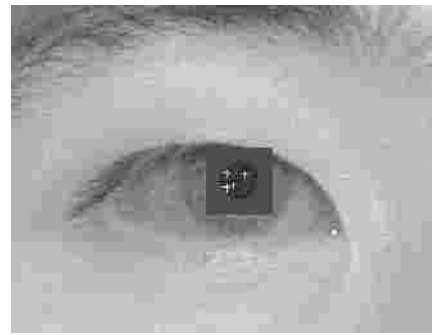


图 10 光斑计算

表 1 校验实验结果

对象	a_1	a_2	a_3	a_4	屏幕宽度	高度
1	1.288 1	1.338 2	1.289 6	0.915 9	1 250	1 400
2	1.277 8	1.414 2	1.399 2	1.073 5	1 324	1 400
3	1.186 2	1.190 6	1.180 4	1.116 9	1 464	1 280
4	1.344 4	1.220 0	1.033 4	1.279 8	1 350	1 350
5	1.231 0	1.240 0	0.980 0	1.157 0	1 280	1 024

人和人之间的 a 值差别没有特定的规律可循，微小的变化在映射过程中都会有较大的影响。在采集过程中，对每个光源要注视 20s 左右才能采集足够多的数据来计算出比较精确的 a 值。对不同的人计算用的屏幕宽度和高度也是有差别的。4 个 a 值和 2 个屏幕计算参数共同影响计算结果。

4.2 视线定位实验

实验 1 定点注视实验。在屏幕上确定 49 个十字叉标定的注视位置，屏幕可显示测试者视线定位的实际结果，实验注视的实际结果由鼠标指针显示地反应出来，通过实际注视点与测试界面上十字叉中心坐标的差别来检验结果的精确度。

采用随机椭圆拟合瞳孔边缘、卡尔曼滤波等传统方法进行视线定位结果如图 11 所示。

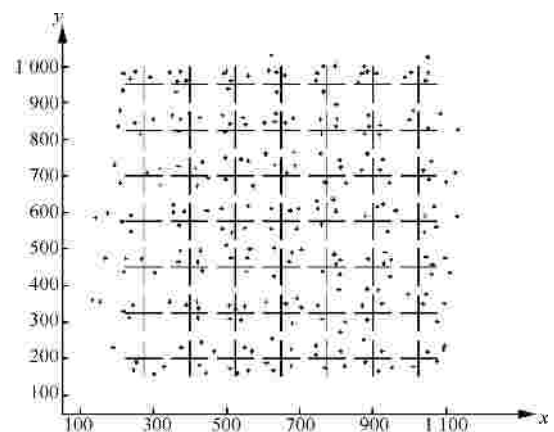


图 11 采用传统方法进行视线定位的结果

图 11 中的大十字叉表示测试者实际注视的位置，周围的小十字叉表示当 5 名测试者各自注视大十字叉时，视线定位实际得到的结果。

传统方法的实验结果表明，该方法虽然可以进行视线定位，但是结果与人眼实际注视的位置之间的误差较大。

本文视线定位实验结果如图 12 所示。

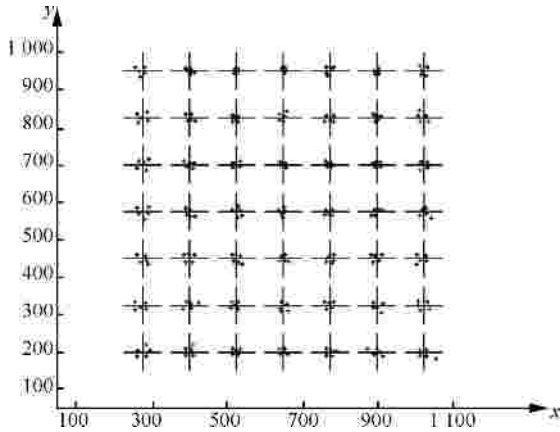


图 12 本文视线定位结果

由定点注视实验结果可看出，本文提出的方法较传统的方法在精度上有明显的提升，本研究的结果更接近于实际注视点，虽然实际计算得的结果不能与注视点完全重合，但是在精度上已经较传统方法有了很大的提高。

实验 2 按钮响应实验。实现有 20 个 200×120pixel 大小的测试界面，测试者用眼睛注视按钮中心，来实现鼠标指针位置的定位，原地点击鼠标实现人眼定位后鼠标事件的响应，如果系统实际计算出的人眼注视坐标在该按钮范围内，则该按钮产生一次闪烁，否则无响应或响应成其他按钮。

5 名测试者测试数据统计情况如表 2 和表 3 所示。

表 2 传统方法 5 名测试者测试数据统计

测试者	点击次数	响应次数	正确次数	有效率
1	100	43	37	37%
2	100	51	45	45%
3	100	47	40	40%
4	100	38	30	30%
5	100	61	50	50%

表 3 本文所提方法 5 名测试者测试数据统计

测试者	点击次数	响应次数	正确次数	有效率
1	100	89	85	85%
2	100	97	93	93%
3	100	98	95	95%
4	100	92	90	90%
5	100	95	88	88%

由按钮响应实验统计数据可看出，本文的方法在定位的准确度上有了很大的提升，5 名测试者的有效数据较传统方法提高了 40% 以上，一方面说明精度提高，另一方面也说明注视点的稳定程度也有所提高。

实验过程中，人头部能够在 CCD 摄像机可捕获图像的范围内自由移动，头部正常转动对实验精度的影响较小，在头部正对屏幕且距离屏幕 50~80cm 时，效果最好，而且注视点越靠近屏幕中心，坐标点定位越准。视线对屏幕时各方块按钮的定位精度都很高，屏幕最左侧、最右侧的视线定位精度略低于中间部分。

实验结果表明，采用本文提出的循环椭圆拟合方法进行瞳孔边界点的提取时，在剔除了假点后，可以更精确地计算出瞳孔中心位置。较之传统的随机椭圆拟合方法有了很大的改进。采用交比的方法来进行映射，也比传统的卡尔曼滤波法估算得出的结果更为精确。

5 结束语

本文提出了一种能够适应人头部自然运动的视线跟踪新方法。本文提出的方法不需要配戴其他设备，大大地降低了对使用者的限制；提出了一种新的瞳孔边缘拟合方案，循环拟合剔除假点，最终得到精确的瞳孔中心，提高了映射的精度。本研究与现有的相关研究相比，在精度上有了较大的提升。本研究可应用于涉及人机交互的各个方面，具有广阔的应用前景。

参考文献：

[1] LI X K; WEE W G. An efficient method for eye tracking and eye-gazed FOV estimation[A]. Image Processing (ICIP), 2009 16th IEEE International Conference on Digital Object[C]. Jilin, China, 2009. 2597-2600.
 [2] PANG L T, Eye-gaze tracking research based on image

- processing[A]. Image and Signal Processing, 2008. CISP 08[C]. Cherbourg-Octeville, France, 2008. 176-180.
- [3] LIU R, ZHOU X, WANG N L. Adaptive regulation of CCD camera in eye gaze tracking system image and signal[A]. 2009 CISP '09 2nd International Congress[C]. Tianjin, China, 2009.1982-1987.
- [4] RAN C, DENG Z L. Two correlated measurement fusion Kalman filtering algorithms based on orthogonal transformation and their functional equivalence[A]. Decision and Control, 2009 held jointly with the 2009 28th Chinese Control Conference. CDC/CCC 2009. Proceedings of the 48th IEEE Conference[C]. Irish, 2009. 2351-2356.
- [5] HONG Y, KWONG S, WANG H L. Decision-based median filter using k-nearest noise-free pixels[A]. Acoustics, Speech and Signal Processing, ICASSP 2009, IEEE International Conference[C]. 2009. 1193-1196.
- [6] MAKILA P M. Intrinsic difficulties in stochastic control of unstable convolution operators on Z[J]. Automatic Control IEEE Transactions on Volume, 2003, 48(11):2015-2018.
- [7] ZHANG J Y, CHEN Y, HUANG X X. Edge detection of images based on improved sobel operator and genetic algorithms[A]. Image Analysis and Signal Processing, 2009. IASP 2009. International Conference[C]. Triest, Italy, 2009. 31-35
- [8] HUO Y K, WEI G, ZHANG Y D, *et al.* An adaptive threshold for the canny operator of edge detection[A]. Image Analysis and Signal Processing (IASP), 2010 International Conference[C]. 2010. 371-374.
- [9] KERR D, COLEMAN S, SCOTNEY B. Near-circular corner and edge detection operators[A]. Machine Vision and Image Processing Conference[C]. Irish, 2007. 7-14.
- [10] FLIZGIBBON A, PILU M, FISHER B. Direct least square fitting of ellipses[J]. IEEE Trans Pattern Anal Mach Intell, 1999,21(5):476-480.
- [11] BATHE J, MOUADDIB E, SALVI J. Recent progress in coded struc-

tured light as a technique to solve the correspondence problem: a survey[J]. Pattern Recognition, 1998, 31(7):963-982

- [12] YOO D H, CHUNG M J. A novel non-intrusive eye gaze estimation using cross-ratio under large head motion[J]. Computer Vision and Image Understanding, 2005, 98(12):25-51.

作者简介：



高迪 (1981-), 女, 黑龙江哈尔滨人, 哈尔滨工程大学博士生, 主要研究方向为计算机视觉、信息安全、图像处理等。



印桂生 (1964-), 男, 江苏泰兴人, 哈尔滨工程大学计算机科学与技术学院院长, 主要研究方向为数据库、虚拟现实、信息安全等。



马春光 (1974-), 男, 黑龙江齐齐哈尔人, 哈尔滨工程大学教授、博士生导师, 主要研究方向为计算机网络与信息安全。