

# 基于机器视觉的新型鼠标

## A New Mouse Based on Machine Vision

卢向东 钱剑敏 (东华大学数字化纺织服装技术教育部工程研究中心 上海 201620)

**摘要:**设计了一种基于机器视觉的新型鼠标。本文提出了基于鼻子特征的实时跟踪算法来控制光标移动;同时介绍一种简易眼睛睁合判断算法来实现模拟击键。文中对整个设计思路和实现方法做了详细介绍。测试表明新型鼠标性能较好。

**关键词:**人脸跟踪 鼻子特征 眼睛睁合 鼠标

### 1 引言

普通鼠标是一种手动的传统交互方式,而手臂残疾的人和游戏爱好者期望采用多种交互方式来操作 PC。随着 CCD、CMOS 图像传感器制造工艺的成熟和图像、视觉核心算法的发展为基于机器视觉的鼠标的实现提供了条件。

### 2 系统的总体概括

新型鼠标的系统结构如图 1 所示。



图 1 系统结构图

系统使用一个 CCD 摄像头捕捉图像,此外还有一台 PC: Windows XP、P4/1.5 GHz。摄像头置于桌面上与屏幕同高度且上仰与水平成 20 度左右,图像分辨率为 320X240。该系统具有较高的识别速度和正确率,按照软件功能可以将系统划分为光标移动模块、模拟击键模块、图像采集驱动模块,各个模块设计的理论依据和具体实现在下面的章节中予以介绍。

### 3 光标移动模块

光标移动模块控制要求:光标移动有一定精度和鲁棒性;精度是要求光标能够点击桌面图标、关闭窗口等,而鲁棒性要允许用户方便、灵活地旋转头部(人脸);在保证精度和鲁棒性的同时不能忽略图像处理的

快速性。

#### 3.1 现有的光标移动方案分析

基于机器视觉的鼠标其光标移动方案是通过人脸跟踪确定人脸位置进而来控制光标。下面对常用人脸跟踪算法来控制光标的优劣性进行比较:

(1) 文献[5]利用彩色信息成功地实现了人脸跟踪,文献[4]对人脸建立了低价参数模型,对光照进行补偿,并可容忍部分遮挡。二者不受人脸姿态限制,但算法用于光标控制的精度不高。

(2) 文献[3]则利用器官的多级 Markov 模型,获得了准确的跟踪结果,但算法在获得较精确位置的要求下的跟踪速度不能为系统接受。

(3) 文献[1]采用鼻尖及附近像素点的亮度值来跟踪,该算法对受光线变化和光线照射角度影响。

(4) 文献[6]采用嘴巴与人脸的灰度差异来跟踪,但对人讲话时嘴型产生变化,难以精确确定人脸位置。

针对各个方案不足,本文设计一种基于鼻子特征的跟踪算法用于控制光标移动。

#### 3.2 算法的出发点

为了方便判断所选择的局部特征是否适合跟踪,定义两个性质作为判断依据:设  $S(V_i, V_j)$  表示任意两个特征  $V_i, V_j$  的相似度,用来衡量进行局部特征跟踪时产生误跟踪(误匹配)大小;被跟踪的局部特征为  $V_i$ ,其他特征为  $V_j$ 。

$$S(V_i, V_j) \rightarrow \min \quad (1)$$

$$S(V_i^0, V_i^1) \rightarrow \max \quad (2)$$

公式(1)表示被跟踪的局部特征  $V_i$  在人脸(甚至背景)中具有唯一性,保证决定人脸位置的局部特征是唯一的。公式(2)表示在跟踪过程中(不同姿态的人脸)被跟踪的局部特征变化很小,保证跟踪鲁棒性。

鼻子正好位于人脸中间,具有凸起、对称、唯一、不变性,比其他脸部器官更容易被跟踪和定位。

### 3.3 鼻子特征

根据上述的考虑,本文将鼻子特征定义如下:

定义:鼻子特征由鼻孔加若干边缘曲线构成。

鼻孔用来保证鼻子特征的唯一性,而若干边缘曲线是为了避免过分的依赖鼻孔,因为不同姿态的人脸中的鼻孔会产生一定的变形,导致匹配值降低,此时需要若干边缘曲线对匹配值进行补偿。

鼻子特征作为局部特征进行跟踪不涉及与特征部分无关区域,它对鼻子附近区域变化不敏感,保证鼻子特征决定的人脸位置连贯变化。同时鼻子特征在跟踪过程依赖边界和相对位置,而不是像素灰度值,对照明条件的变化有很强的适应性。

不同姿态及带眼镜的人脸中鼻子特征变化,如图2所示。图2中可知鼻子特征在人脸范围或背景中具有唯一性且特征变化小,同时图2也表明不同姿态时,眼睛、嘴巴、眉毛的特征变化较大。实验也验证了鼻子特征在不同的照明条件变化很小。对500帧同一人脸不同姿态的图像,采用鼻子特征进行跟踪测试,正确识别达491帧,识别率达98.2%。



图2 不同姿态的人脸中鼻子特征

### 3.4 具体实现

首先明确几个概念:人脸位置——新一帧图像中鼻子特征的最佳匹配位置,此时模板覆盖下子图的重心;起始帧——提取鼻子特征模板;跟踪帧——进行人脸跟踪来得到人脸精确位置的帧;搜索区域——人脸

做较大运动时,当前帧的人脸位置为中心的圆区域内仍可以包含下一帧的鼻子特征,因此可作为跟踪鼻子特征的搜索区域。

#### 3.4.1 起始帧

起始帧中鼻子特征模板提取考虑两个方案:方案1:通过人脸检测得到人脸的大致位置后<sup>[7,8]</sup>,将图像二值化处理,利用水平、垂直积分投影法<sup>[7]</sup>得到鼻子位置,以鼻子位置为中心提取鼻子特征模板。

方案2:软件启动时视频窗口显示十字交叉线和矩形框;脸部尽量端正,鼻子对准视频窗口所显示十字交叉点,眼睛跟嘴巴恰好摆放在矩形框内,经过10秒调整后,系统捕获第一帧图像作为起始帧,裁剪以十字交叉点为中心大小  $h \times l$  的长方形作为鼻区,在鼻区中提取鼻子特征模板。

若采用方案1容易受光线干扰和背景环境的影响,所以方案2的可能烦琐一点,但是可靠性高。市场上的产品较多采用方案2。当然对于手臂正常用户可以直接用常规鼠标圈取鼻区。

#### 3.4.2 跟踪帧

采用传统的层次模板匹配方法<sup>[2]</sup>,对搜索区域进行模板匹配,鼻子特征模板如图3(a)。由于人脸旋转,在每个位置进行相关匹配时需要将鼻子特征模板进行旋转和比例尺度搜索,鼻子特征模板的旋转角度  $\theta$  可以限制在  $-\pi/4$  到  $\pi/4$  之间,比例尺度  $Scale$  可以限制在 0.75 到 1.15 之间。相关匹配结束得到当前帧最佳匹配位置时人脸位置、 $\theta$ 、 $Scale$ 、和最佳匹配系数  $Score$ 。为了显示跟踪,在当前帧图像中标记一个以为中心,边长4的正方形。如图3所示。其匹配参数为:  
 $(130.75, 155.42)$ 。  $Score = 0.855$ ,  $\theta = -0.0398$ ,  $Scale = 1.08$ 。



(a) 不同层次鼻子特征模板 (b) 鼻子位置

图3 人脸位置

### 3.5 Windows 操作系统下光标移动控制

设  $(x_c, y_c)$  表示光标在显示器屏幕上的位置。一般摄像头分辨率  $320 \times 240$  而屏幕分辨率  $1024 \times 768$ 。需采用 1:4 对应关系:  $if x_n \geq 32, x_c = 4(x_n - 32)$ ,  $else x_c$

$=0$ ; if  $y_N \geq 24, y_c = 4(y_N - 24)$ , else  $y_c = 0$  来保证光标在全屏幕范围内移动,也保证鼻子与图像边界有一定余量。这种模式下的光标分辨率可达  $256 \times 192$ 。若采用亚像素,可以获得更高分辨率,这样的分辨率满足将光标移动到所需桌面图标、窗口图标。如果用户需要更高的分辨率可以采用两级光标移动模式。最后调用 Windows 的 Mouse API 函数 SetCursorPos 实现光标移动。

#### 4 模拟击键模块

关于模拟击键模块参考文献较少,文献<sup>[8]</sup>采用手势识别的方案来实现,但对本系统不适用。文中考虑以下两个方案:

方案 1: 通过语音识别来产生单击右键、双击左键、单击左键的信号来配合光标移动模块的工作。

方案 2: 眼睛容易睁开,这可用于模拟鼠标击键。在很多场合,可以不使用鼠标右键,用户依然可以正常使用计算机:比如浏览网页、运行程序等,同时根据人们用眼的习惯:有人容易合右眼,有人容易合左眼,所选择容易合的那只眼睛模拟单双击左键。

方案 1 比较容易实现,且对模拟击键控制效果好,但需要语音表达。这里重点探讨方案 2。

##### 4.1 眼睛睁开判断算法

为了加快识别速度,眼睛睁开判断算法的设计不应太复杂。文中依据形态学运算和形状特征来实现。

###### 4.1.1 眼睛定位

利用已确定的人脸位置与眼睛之间距离来产生一个人脸区域,而不通过人脸肤色检测得到人脸的大致区域。对人脸区域进行二值化,如图 4(a),为了减少照明条件变化产生的影响,将  $(x_N, y_N)$  的亮度值减去一个常数作为 ROI 二值化的阈值;利用  $2 \times 2$  结构元进行膨胀填充二值人脸区域的空洞,以利于形状的验证。采用以下眼睛连通区域的面积和眼睛连通区域的中心位置两条形状特征进行眼睛的粗定位,得到眼睛候选连通区域如图 4(b)。

###### 4.1.2 眼睛睁开

考虑人眼睁开的宽高比不同且人脸旋转时眼睛也旋转,在实验反复测试表明,对眼睛候选连通区域  $w_i$  进行椭圆检测时,当  $3b \geq a$  可认为眼睛是睁状态。设  $n$  为连通区域  $w_i$  像素个数。计算连通区域的质心、惯

$$\bar{x} = \frac{1}{n} \sum_{(x,y) \in w_i} \sum x, \bar{y} = \frac{1}{n} \sum_{(x,y) \in w_i} \sum y \quad (3)$$

$$\mu_{pq} = \sum_{(x,y) \in w_i} \sum (x - \bar{x})^p (y - \bar{y})^q \quad (4)$$

由(3)(4)求连通区域的椭圆长短轴和方向:

$$a = \sqrt{8(\mu_{20} + \mu_{02} + \sqrt{(\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2})} / 2 \quad (5)$$

$$b = \sqrt{8(\mu_{20} + \mu_{02} - \sqrt{(\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2})} / 2 \quad (6)$$

$$\beta = -0.5 \arctan(2\mu_{11} / (\mu_{02} - \mu_{20})) \quad (7)$$

如图 4(c) 所示中左眼  $a = 12.24, b = 5.05$ , 满足  $3b \geq a$  认为眼睛睁开;右眼  $a = 13.00, b = 3.40$ , 不满足  $3b \geq a$  认为眼睛合上。

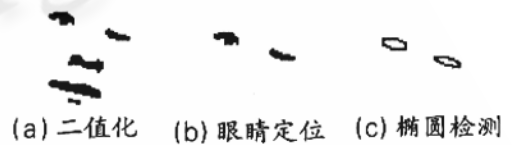


图 4 眼睛睁开判断

当然不同人脸姿态和带眼镜对眼睛睁开判断有影响,效果不佳,如表 1 所示:

表 1 眼睛睁开判断效果

测试图像内容(单眼)	测试数	正确识别	识别率
带眼镜、眼睛睁状态	245	208	84.9%
带眼镜、眼睛合状态	65	44	67.7%
不带眼镜、睁状态	185	169	91.4%
不带眼镜、合状态	55	45	81.8%

##### 4.2 Windows 操作系统下模拟击键实现

假设选择右眼来模拟单双击左键。当前帧图像中判断右眼是否合上,若右眼合上,则计数器 +1,否则置零,这样可排除人的正常眨眼;当计数器累加到 5,即连续 5 帧图像中的右眼合上,断定右眼连续合上约 1 秒,就模拟左键单击;当计数器累加到 10,断定右眼连续合上约 2 秒,则模拟左键双击,计数器清零。最后调用 Windows 的 Mouse API 函数 mouse\_event 实现光标移动。

#### 5 测试平台实现和性能分析

测试平台的开发环境:VB6.0、halcon7.0,使用 ez-VidC60.ocx 简化图像采集模块编程,也保证编写的程

序支持大部分摄像头。测试平台如图 5 所示,人脸可以在摄像头镜头内任自由摆动。用黑色的小正方形标记人脸位置,具体数值可见图中的测试参数,耗时约 100 ~ 150ms/frame。



图 5 新型鼠标测试平台

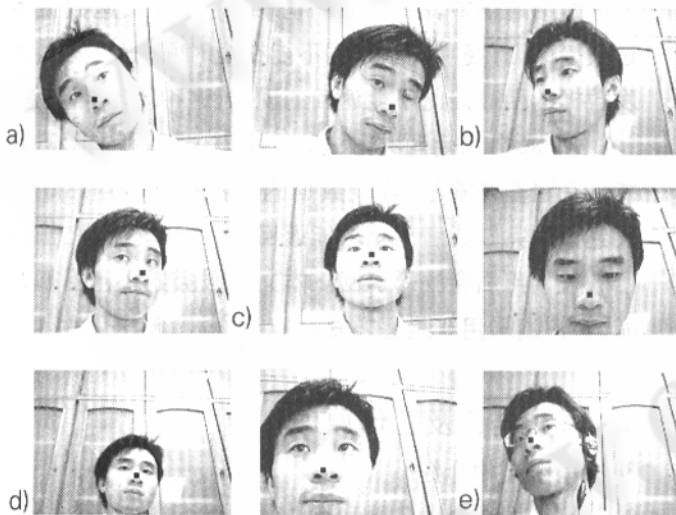


图 6 极限性能评价

由于照明条件、实验条件及人脸在镜头中运动很难重构,这给新型鼠标评估带来一定难度,文中主要从极限性能、抗干扰与其他产品比较来评价。

### 5.1 极限性能评价

对人脸平面内平面外左右旋转如图 6-a 6-b,上仰下俯如图 6-c,远离靠近摄像头如图 6-d 等测试。图 6 表明人脸可以在镜头内自由摆动而不影响跟踪效果,但模拟击键模块受平面外左右旋转影响较大。

### 5.2 抗干扰评价

人为加入干扰因素如手干扰、改变照明条件、人为离开镜头范围等都能较好跟踪人脸位置。

### 5.3 与 D. Gorodnichy 的 NOUSE 软件比较

本文开发的软件对光标移动控制的鲁棒性不如 NOUSE 软件,但是对光标的控制精度超过了 NOUSE,且对误识别情况能自行恢复以及添加了模拟击键模块,这是 NOUSE 所没有的。

## 6 总结展望

新型鼠标初步实现了普通鼠标的功能,光标移动模块具有较高精度和鲁棒性,但是对模拟击键模块的控制效果不佳,将进一步的改进。随着研究的深入,新型鼠标将可以让手臂残疾的人能够比较轻易地用上计算机,同时它也很可能受到游戏爱好者们青睐。

### 参考文献

- 1 D. Gorodnichy. On importance of nose for face tracking [J]. In Proc. Intern. Conf. on Automatic Face and Gesture Recognition, 2002.
- 2 Halcon/Hdevelop Reference Manual [B/OL]. Mvtec.com, 2005.
- 3 C. Kervrann, F. Heitz. A hierarchical Markov modeling approach for the segment and tracking of deformable shapes [J]. Graphical Models And Image Processing, 1998, 60(3): 173 ~ 195.
- 4 G. Hager, P. Belhumeur. Efficient region tracking with parametric models of geometry and illumination [J]. IEEE Trans. PAMI1998, 20(10): 1025 ~ 1039.
- 5 刘明宝、姚鸿勋、高文,彩色图像的实时人脸跟踪方法[J],计算机学报,1998,21(6):527 ~ 531.
- 6 邢昕、汪孔桥、沈兰荪,基于器官跟踪的人脸实时跟踪方法[J],电子学报,2000,28(6):29 ~ 31.
- 7 钟威、刘智明、周激流,人脸检测中眼睛精确定位的研究[J],计算机工程与应用,2004,40(36):73 ~ 76.
- 8 杜威、李华,基于单目视觉的手势识别系统[J],工程图学学报,2000,21(3):133 ~ 138.