

引用格式: 李森, 谢翔, 李国林, 等. 一种基于颜色结构光的投影仪摄像头触摸板人机交互系统[J]. 微电子学与计算机, 2019, 36(1): 61-65. [LI S, XIE X, LI G L, et al. A Projector-Camera-Touchpad Human Computer Interaction System Based on Color Structured Light[J]. Microelectronics & Computer, 2019, 36(1): 61-65.]

一种基于颜色结构光的投影仪摄像头触摸板人机交互系统

李 森¹, 谢 翔¹, 李国林², 王志华¹

(1 清华大学 微电子与纳电子学系, 北京 100084; 2 清华大学 电子工程系, 北京 100084)

摘 要: 本文提出了一种只用投影仪和单摄像头搭建的投影触摸板人机交互系统. 本人机交互系统是利用投影仪摄像头系统对手指的运动进行捕捉, 可以使任何平面变成触摸板. 本文分析了系统测量三维坐标的误差, 通过优化标定提高了测量精度. 为了解决手指触碰的高度和位置的精准度问题, 基于建立的手指模型利用颜色结构光来实时获取手指三维信息和寻找触碰点, 并且利用卡尔曼滤波优化手指触碰后手指的运动轨迹, 提高了手指触碰的精度在 2 mm 以下 99% 准确率, 触碰后的位置精度达到 1 个像素点, 利用投影仪-摄像头系统实现了触摸板的功能和应用.

关键词: 投影仪-摄像头; 人机交互; 颜色结构光; 卡尔曼滤波

中图分类号: TP391

文献标识码: A

文章编号: 1000-7180(2019)01-0061-05

A Projector-Camera-Touchpad Human Computer Interaction System based on Color Structured Light

LI Sen¹, XIE Xiang¹, LI Guo-lin², WANG Zhi-hua¹

(1 Institute of Microelectronics, Tsinghua University, Beijing 100084, China;

2 Electronic Engineering Department, Tsinghua University, Beijing 100084, China)

Abstract: A method using only one projector and one camera is proposed to build a Projector-Camera-Touchpad (PCT). This HCI system uses projector-camera to catch the finger's motion to make any plane to a touchpad. This paper analyzes the error measuring the 3-dimensional coordinates of the system, and improves the accuracy by optimizing the calibration. To solve the problem of height when to touch happens and finger's position accuracy, based on a finger model the color structured light is adopted to get 3-dimension information and locate the finger touch point. At meanwhile, the trajectory of finger touch is optimized by Kalman filter. Finally, the touch accuracy is improved to 99% under 2mm and touch position precision are increased to 1 pixel. The function of the touchpad is realized by the projector-camera system.

Key words: projector-camera system; human-computer interaction; color structured light; Kalman filter

1 引言

为了满足人们对于更加便携和方便的人机交互方式的追求, 很多研究者将投影仪-摄像头系统安装到微型设备来提供更大的人机交互区域. 文献[1-2]

中的 PlayAnywhere 和文献[3]中 OmniTouch 利用微型投影仪和深度摄像头, 微软的 Kinect, 作为输入和输出设备, 但是问题在于深度摄像头因为部件太多, 其功耗和大小都不适合集成到越来越小的智能设备上, 而且深度摄像头的精度并不能满足触摸

板的精度要求. 文献[4]采用了单投影仪和单摄像头去把任意平面变成依靠触碰交互的平面. 我们的团队认为人机交互系统应该不仅仅是适应便携性的设备, 而且应该有更加便利的交互方式. 很多人们还是对于传统的人机交互方法更加熟悉, 比如键盘和鼠标, 所以我们利用投影仪-摄像头系统来实现类似传统交互方式来满足人们的需求. 文献[5]提出了一种新颖的手指触碰检测方法来利用投影仪-摄像头系统实现一个投影键盘的功能. 文献[6]提出了一种基于更加简单并且精确的手指模型的更加准确的指尖触碰检测方法, 并且将这个方用于文献[5]中的投影键盘中, 提高了触碰精确度. 文献[7]提出基于手指模型的投影触摸板的系统建立.

本文主要在文献[7]的工作基础上, 利用投影仪-摄像头系统建立一个可以使任意平面成为触摸板的投影触摸板系统并改进整体系统的各个功能的精度. 通过对投影仪-摄像头系统的测量精度分析, 找到影响触碰判断精度的参数, 其中投影仪和摄像头的成像平面上的误差受到标定方法的影响, 所以通过改进标定方法中对于透视失真角点的亚像素坐标的寻找, 得到了更加精确的系统内外参数, 同时减小了三维测量误差, 得到更精确的触碰判断. 考虑到用户对于用户体验和便携性的需求, 系统对于实时性, 计算和存储都有一定的限制, 所以基于手指模型和颜色编码结构光, 改进了利用模型上少数关键点进行触碰判断的方法, 同样提高了触碰判断精度. 之后基于手指模型, 可以推测出的粗略触碰点, 利用卡尔曼滤波对手指进行了触碰后的运动追踪, 减小更多的误差带来的对于手指轨迹的影响, 得到更加精确触碰位置, 同时节省了计算和存储空间.

2 系统精度分析与标定方法改进

投影仪-摄像头系统整体上可以当做一个双目系统, 如图 1 所示, O_r 和 O_l 分别为投影仪和摄像头的光心, x_r 和 x_l 分别为投影仪和摄像头图像坐标系中成像的位置, P 和 P' 分别为真实物体位置和经过双目系统测量得到的位置, z 和 z' 分别为真实物体深度和经过双目系统测量得到的深度, f 为成像清晰时候投影仪和摄像头的像素焦距, 由焦距 f_p 像素点尺寸 d_p 得来, T 为基线长度.

由三角形的相似原理我们可以得到深度误差 $z - z' = \Delta z(1)$, 其中 δ 是物体在摄像头成像平面上的误差, 由测量系统标定不准确带来的误差造成.

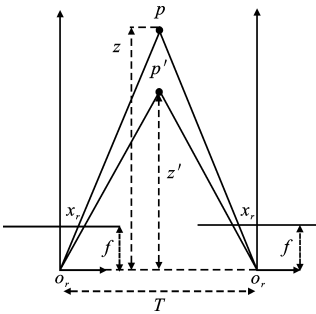


图 1 投影仪-摄像头系统模型

$$\Delta z = \frac{\delta}{f} \times \frac{z^2}{T} = \frac{\delta}{\frac{f_p}{d_p}} \times \frac{z^2}{T} \tag{1}$$

深度误差与像素点误差成正比, 与像素焦距成反比(与像素点大小成正比, 与焦距成反比), 与物距平方成正比, 与基线长度成反比. 在现实使用中, 我们会追求基线长度尽量小, 以适应便携设备的大小, 一般为 5~10 cm, 交互距离也是控制在 20~50 cm 之间, 方便大多数交互场景的使用. 像素焦距由硬件设备决定. 还要考虑标定方法的精度影响之后测量三维信息时产生的误差, 这个误差也会影响到深度测量的精度. 在本投影触摸板应用中, 深度测量会较大的影响手指触碰判断的功能, 所以深度测量的精度越高越好.

为了使触碰判断更加准确, 深度测量精度也要提高. 利用精度更高的标定方法能改善深度测量的精度. 本系统中之前使用的标定方法是基于张正友标定法的适用于投影仪-摄像头系统标定方法^[8]. 经过数据分析之后我们发现在此方法中, 寻找亚像素级角点误差较大的点都是靠近边缘的存在透视失真的角点, 原因是找角点的方法没有考虑透视失真, 找的角点的亚像素坐标就会存在误差, 最后导致了整体的测量系统上的误差. 基于文献[9]中的考虑了存在一定夹角的寻找亚像素级角点函数, 在投影仪-摄像头联合系统标定方法中利用此函数来寻找角点的亚像素坐标, 提高了存在透视失真的角点的亚像素坐标的精确度, 最后提高了系统的标定精度. 此标定方法的投影仪重投影误差达到了 0.140 1.

3 触碰判断方法改进

触碰判断方法基于文献[7]提出的手指模型, 如图 2 所示, 同时用其中的颜色编码结构光获取三维信息.

如果我们完全重建整个投影场景, 需要的计算量将是巨大的. 幸运的是, 我们引入了手指模型

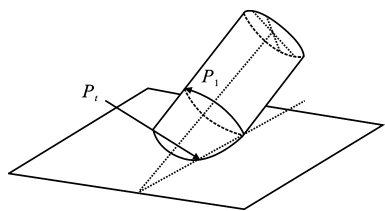


图 2 手指模型和关键点

辅助我们进行触碰的判断,即使获取的三维信息只有上表面,基于手指模型我们也可以获得手指下表面距离交互平面的距离,选取部分的关键点的三维信息来进行触碰的判断,减少计算量和存储需求^[10].

如图 2,我们最关心的点是手指和桌面的接触点 P_t ,但是这个点的三维坐标无法直接从摄像头中拍摄并获取的,所以我们要利用手指模型来辅助获取. 基于手指模型, P_t 是手指与桌面接触时手指模型前端球心 P_i 在桌面上的投影, P_i 到 P_t 的距离就是球的半径,且 $P_i P_t$ 垂直于交互平面,所以几何模型的参数可以得到 $P_i P_t$ 等于手指模型前端四分之一球半径 r ,同时也等于手指模型后端椭圆柱椭圆的半长轴 a .

由此我们触碰判断的准则变为如果 P_i 距离交互平面大于 r ,则没有触碰,如果小于 r ,则发生触碰. 所以接下来的工作是找到点 P_i 的位置和三维坐标. 因为手指模型在平面上的投影是一个半椭圆加一个矩形,通过提取的手指二维图像如图 3(a) 和手指模型的投影进行匹配可以寻找 P_i 的位置.



(a) 分割手指

(b) 手指主方向

图 3 指尖点 P_f 和球心点 P_i , 重心和中轴线与条纹交点

对比文献[7],本文提出一种更加鲁棒的先从二维图像上找到点 P_i ,再计算 P_i 距离交互平面高度的方法. 文献[7]在找手指主方向(手指模型中椭圆柱的中心轴指向指尖的方向)的时候,仅依靠了指腹部分近似直线的特点. 而我们提取的手指轮廓指腹部分并不是连续的,是由不间断的点组成的,大致主方向可以找对,其准确度却不高,稍微的方向偏移对于指尖点的影响却是极大的. 如图 3(b),在本文中我们加上轮廓点到重心的距离,加上曲率判断指尖,

再结合主方向寻找指尖点. 指腹轮廓的两条近似直线一个功能是确定手指模型的直径 $2r$,即两条线的距离,另一个功能是找到大致的指尖方向. 主方向指明手指大致方向,找到此方向上距离重心最远的两点,其分布在手指的指尖和末尾,再通过两点各自的曲率大小分辨出指尖点,曲率大的为指尖点 P_f .

接着通过指尖点 P_f 延主方向反方向寻找 r 距离的位置就可以找到 P_i 的位置. 由于手指和交互界面有一定的角度,需要找手指的俯仰角. 选择通过指尖点的主方向的直线,作为手指中轴线,选取中轴线与横条纹交叉的点,如图 3(b),中轴线和手指上虚线条纹交叉的点,而不是像文献[7]中选取等距离的点来计算俯仰角. 因为横条纹与极线垂直,利用极线约束可以更准确和方便地计算出中轴线上点的三维坐标,并且通过下面公式(2) 计算出俯仰角,公式(2) 中 x 代表在中轴线上的坐标, y 代表距离交互平面的高度.

$$\varphi_p = \arctan \frac{\sum_{k=1}^N (x_k - \bar{x})(y_k - \bar{y})}{\sum_{k=1}^N (x_k - \bar{x})} \quad (2)$$

从而在俯视图上 P_f 到 P_i 的距离 $D_p = r \times \cos \varphi_p$,如图 3(b). 如此找到了二维俯视图上 P_i 的位置,三角计算的方法得到 P_i 的高度 H_{pi} .

相比于文献[7]中基于单一阈值的触碰判断方法,由于触碰是一个连贯性的动作,本文利用前后连续多帧的 P_i 的高度来判断触碰. 我们认为在触碰动作之前手指至少距离交互界面有一定的高度 h 表示还没有进行交互. 初始 $H_{pi} > r + h$,之后的连续多帧中 H_{pi} 在逐渐减少,而且最后达到了 $H_{pi} < r$,那么发生了触碰. 初始 $H_{pi} < r$,在之后连续多帧中 H_{pi} 逐渐增加,最后 $H_{pi} > r + h$,则发生手指脱离交互界面,触碰结束.

基于手指模型的算法只需要计算几个关键点的三维坐标就可以精确判断出是否发生触碰,如此简化了整个计算过程,而且在计算出每张图中 P_i 的坐标之后这一帧图的原始图像数据无需再使用了,这样节约了很多存储空间,为整体系统的实时性和硬件实现提供了有力支持.

4 运动追踪

在触碰之后手指的移动轨迹变成了我们关心的重点,因为手指的移动控制着光标的移动. 但是手指是一个较大的物体,要将手指和光标对应起来才能

实现触摸板的交互. 基于手指模型,我们发现手指和交互界面接触点 P_i 可以代表手指对应光标的位置,因为 P_i 的三维坐标在触碰判断的时候已经计算出来了, P_i 在投影交互平面上移动时的投影的相对二维坐标变化代表光标的位置移动,以此来实现手指位置对于光标位置的控制.

由此便带来了手指和光标对应点是否准确的问题,其精度会影响到投影触摸板的用户体验,所以对于手指在交互平面上的相对移动位置的检测结果我们期望是精度越高越好. 但是对于 P_i 的相对坐标获取问题,整个流程中存在很多噪声,其中包括图像拍摄的噪声,图像处理的噪声,手指模型匹配的噪声等. 这么多的噪声会影响 P_i 的坐标获取,导致在实现拖动等需要获取相对运动位置的操作的时候存在误差.

我们认为获取到的影响 P_i 位置的所有噪声都是服从高斯分布. 问题转化为已知 T 时刻及之前所有的 P_i 的运动位置,要求更加精确的 $T+1$ 时刻的 P_i 的位置. 基于 $T+1$ 时刻之前的连续帧中已获得的 P_i 的位置可以计算出 P_i 在 $T+1$ 时刻之前每一时刻的运动状态,依赖这些运动状态和 T 时刻 P_i 位置预测 $T+1$ 时刻 P_i 的运动状态和位置,这就得到了 $T+1$ 时刻 P_i 的位置预测值. 将此预测值结合 $T+1$ 时刻对 P_i 通过第 4 部分中算法获得的真实观测值来获得 $T+1$ 时刻误差较小的 P_i 的位置,以此作为 P_i 在 $T+1$ 时刻的最终位置. $T+1$ 时刻的预测值和观测值结合方法是两者加权平均,权重由预测值的方差和观测值的方差确定,其方差越大说明此值受到的噪声影响越大,则权重越小,由此来获得更加准确的 $T+1$ 时刻的位置,同时减小噪声的影响. 这个方法就是利用卡尔曼滤波^[11] 来进行 P_i 运动轨迹的优化,减小误差的影响. 本方法中直接对 P_i 进行卡尔曼滤波,以获得更加精确的 P_i 的运动位置和轨迹,从而提升触摸板系统在对手指运动过程中位置精度.

5 实验结果及分析

在本实验中我们使用了 Optoma PK301 DLP 投影仪,分辨率为 848×480 , 和 Logitech Pro9000 摄像头,分辨率为 960×720 . 投影仪和摄像头固定在一起,保证其角度和基线长度的固定,然后利用塑料手指模型固定在支架上来进行模拟操作实验. 系统工作的高度范围距离交互平面 $20 \sim 50$ cm. 本实验中系统距离交互平面约 35 cm.

为了检测触碰判断在高度的精度和运动跟踪的精度,检测实验分为两部分:第一部分检测触碰判断在不同高度上的精度,第二部分检测手指运动追踪的精确度.

实验一是检测触碰判断在不同高度上的准确度的实验. 通过硬币来控制触碰的高度,即将硬币垫在手指要触碰的位置,手指碰到硬币的高度为最低高度,一个硬币厚度约为 1.7 mm,在一定高度(垫一定数量的硬币)分别进行 100 次的触碰,得到的触碰判断结果为在 2 mm 以下(1 个硬币厚度),触碰判断可以达到 99% 的识别率,这样的识别率可以满足大部分人在日常生活中对于触摸板触碰高度的需求. 本方法与其他文章中方法触碰准确率对比如下表 1,可见本方法在高度为 2 mm 的情况下,触碰准确率是比较高的,可以满足人们在日常生活中对于触碰功能的使用.

表 1 实验结果比较

	本文	文献[9]	OmniTouch ^[3]	smartUI ^[5]
高度	2 mm	2 mm	—	2mm
准确率	99%	98%	95%	91%

在触碰判断过程中由于只需要存储少量关键点的三维数据,可以将每一帧图像的信息减少到我们所需要的关键点的三维信息的大小,从而节约了很大的存储空间. 在同一时刻只需要保存当前帧图像和初始图像进行来获取当前时刻 P_i 的三维坐标,同时需要保存手指参数 r ,前 30 帧的连续的 P_i 的三维坐标,交互界面的方程 3 个参数来进行触碰判断. 除了必须存储的当前帧和背景图像两张图以外,其他所有需要的点和参数的数据存储都用双精度浮点数,总和为 70 kB 左右. 而保存一帧无损图片就需要 2 MB 左右,如果对场景进行重建,图像中每一个像素点还需再加一维信息,需要占用 3 MB 左右. 利用本算法,大大节约了每帧所需要存储的信息,同时实现了高精度的触碰判断. 触碰判断后当前帧图像无需再保存,直接将 P_i 点的坐标保存,利用 P_i 在交互平面上的投影点计算出 P_i 的坐标,直接传递到下一个步骤手指的运动追踪中去,同时利用 P_i 的高度来进行触碰判断,两步并行处理,进一步提高整体算法的运行效率.

实验二是检测手指移动固定距离之后获得鼠标移动的像素点数目. 在触碰的状态下移动塑料手指模型不同的距离,同时获取鼠标移动的像素点数,每个距离进行 10 次试验,结果如表 2.

表 2 手指移动距离对应光标移动像素点数

手指移动距离	5 cm	7 cm	10 cm
平均像素点数	163.3	225.9	323.5
每厘米对应像素点数	32.66	32.27	32.35

在移动不同距离情况下每厘米对应的像素点数目误差在 1 像素点内,这不会影响用户对于触摸板的日常使用.

6 结束语

本文采用投影仪-摄像头系统,利用彩色结构光建立了投影触摸板人机交互系统.主要分析了系统的误差并针对标定存在的问题对标定方法进行了改进,提高了三维测量的精度.接着利用手指模型实现了触碰检测和运动跟踪等功能,解决了触碰判断的深度信息获取问题,同时提高了手指触碰判断的正确率和光标对应点的运动精度.同时,我们简化了投影触摸板的实现过程,只需要获取手指模型上部分关键点,而且利用卡尔曼滤波优化了运动轨迹,提高了运动过程中的手指位置精度.最后,实验结果表明,我们的投影触摸板的触碰检测能够准确地判断手指在距投影平面 2mm 以下的触碰,运动跟踪精度能达到低于每厘米 1 像素,整个投影触摸板可以像一个真正的触摸板来使用.在未来,我们将添加更多的功能到这个触摸板,如由多指和绘图应用程序,并且提高投影触摸板硬件系统的便携性.

参考文献:

[1] Wilson A D. Using a depth camera as a touch sensor [C]// proc. of ACM int. conf. interactive tabletops and surfaces, Saarbrcken, Germany, 2010: 69-72.

[2] Wilson A D. Play Anywhere: a compact interactive tabletop projection-vision system[C]// Proc. of the 18th annual ACM symp. User interface software and technology. New York, 2005: 83-92.

[3] Harrison C, Benko H, Wilson A D. OmniTouch: wearable multi touch interaction everywhere [C]// Proc. of the 24th annual ACM symp. User interface software and technology. Berlin, 2011: 441-450.

[4] Dai J, Chung R. Making any planar surface into a tou-

ch sensitive display by a mere projector and camera [C]// Proc. of IEEE Computer Vision and Pattern Recognition Workshops (CVPRW). CVPRW). Providence, Rhode Island, USA,2012: 35-42.

[5] Hu J, Li G, Xie X, et al. Bare-fingers touch detection by the button's distortion in projector-camera system [J]. IEEE Trans. On Circuits and Systems for Video Technology, 2014, 24(4): 566-575.

[6] Cai X, Xie X, Li G, et al. A new method of detecting fingertip touch for the projector-camera HCI system [C]// Proc. of IEEE Int. Symp. on Circuits and Systems. Melbourne, Australia, 2014:526-529.

[7] Li S, Xie X, Li G, et al. Bare-finger projector-camera-touchpad (PCT) HCI system using color structured light[C]// IEEE International Symposium on Circuits and Systems. Baltimore, MD, USA, IEEE, 2017: 1-4.

[8] 宋玮,郑毅,谢翔,等. 一种投影仪-摄像机系统的简易高精度标定方法[J]. 微电子学与计算机, 2015, 32 (2):28-32.

[9] Zhang Y, Li G, Xie X, et al. A new algorithm for accurate and automatic chessboard corner detection[C]// IEEE International Symposium on Circuits and Systems. Baltimore, MD, USA, IEEE, 2017:1-4.

[10] Lanman D, Taubin G. Build your own 3D scanner: 3D photography for beginners [EB/OL]. (2016-09-26). <http://mesh.brown.edu/byo3d/>

[11] Faragher R. Understanding the basis of the kalman filter via a simple and intuitive derivation [Lecture Notes][J]. IEEE Signal Processing Magazine, 2012, 29(5):128-132.

作者简介:

李 森 男,(1993-),硕士研究生.研究方向为三维测量、人机交互. E-mail:lslzy712@163.com.

谢 翔 男,(1971-),博士,副教授.研究方向包括集成电路设计、图像处理、无线通信、生物医学电子学等.

李国林 男,(1970-),博士,副教授.研究方向包括互连分析以及 RFIC 设计等.

王志华 男,(1960-),博士,教授.研究方向包括集成电路设计、图像处理、生物医学电子学等.