

基于视觉的投影 - 摄像机指尖触控系统

宋呈群 程 俊

(中国科学院深圳先进技术研究院 深圳 518055)

摘 要 文章提出了一种基于视觉的投影交互方法, 通过判断用户手指与其在屏幕上投射阴影的融合程度来检测投影屏幕上是否有触控事件发生。为提升系统的鲁棒性, 在检测阶段, 将手指与其阴影同时从相机图像中分割出来, 进而采用一种线性模型判断方法来判定手指与其阴影的融合程度, 如果判定触碰事件发生, 则通过指尖位置检测来判断触控的具体位置。该方法不需要用户使用任何辅助工具, 即可用手指在投影屏幕上直接与显示画面实现交互操作。

关键词 投影-摄像机系统; 人机交互; 前景提取; 手势识别

中图分类号 TP 751.1 **文献标志码** A

Vision-Based Fingertip Interactive Projector-Camera System

SONG Chengqun CHENG Jun

(Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China)

Abstract A vision-based fingertip interactive projector-camera system was presented in this paper. The proposed method can detect touch operations by judging the fusion level of fingertip and its shadow. To improve the robustness of the proposed method, in image detection stage, finger and its shadow are extracted and segmented from the image separately. And then, a linear detection model was introduced to judge the fusing level of fingertip and its shadow in the image. While a touching event was detected, the touching location can be calculated from the extracted fingertip image coordinate. With the proposed method, users can realize finger-based operation on the projector screen without any accessory tools.

Keywords projector-camera system; human-computer interaction; foreground extraction; hand gesture recognition

1 引 言

投影-摄像机触控系统^[1]投射计算机显示画面到任意平面并允许用户通过徒手来操控计算机。

它为用户提供了一种使用手指便可生动自然地与计算机进行交互的方式, 如点击虚拟键盘、拖动对象、打开文件、翻动网页等。投影-摄像机触控系统避免了在整个显示屏幕上铺设大量的电容或电阻阵列, 节省了触控大屏的成本, 并提高了

收稿日期: 2016-11-16 修回日期: 2016-12-11

基金项目: 国家自然科学基金项目(61403365); 中国博士后科学基金(2016M602543); 广东省科技发展专项资金(2016B010108010); 广东省自然科学基金项目(2015A030313744); 深圳科技计划(JCYJ20140901003939013); 深圳市技术攻关项目(JSJG20160331185256983)

作者简介: 宋呈群, 博士后, 助理研究员, 研究方向为人机交互、图像处理和光电信息技术等; 程俊(通讯作者), 研究员, 博士生导师, 研究方向为人机交互、模式识别、三维重建和机器视觉等, E-mail: jun.cheng@siat.ac.cn。

成品率。

韩国电子部品研究院的 Park 等^[2]提出了一种投影交互系统, 他们利用一台深度相机从周围场景中提取手部区域, 用于识别手势。但是深度相机昂贵, 并且由于周围照明的干扰, 其精度并不高。中国科学院深圳先进技术研究院的何明与程俊^[3]利用一种新型的视觉算法实现了投影显示交互触控检测, 该检测基于一种指尖局部投影自适应编码方法。然而, 其系统需要高速摄像机来实现, 增加了成本, 而且操控定位精度不是很精确。在许多交互系统中, 手指及其阴影均被使用。中国台湾交通大学的 Dung 等^[4, 5]提出了一种投影交互系统, 利用手指及其阴影的距离来判定触控动作。该触控系统不需要大量电子探测元件, 可以应用于交互式投影系统。然而, 额外附加红外光源进行接触探测并不方便, 而且红外光在投影屏幕上形成中间亮边缘暗的圆斑, 降低了触控探测精度。日本大阪大学的 Xu 等^[6]提出一种利用阴影的投影交互系统。中国台湾台北大学 Huang 等^[7]提出了一种基于阴影的指尖探测方法和一种快速的有限状态机 (Finite-State Machine) 模型来进行手势识别。但他们的系统需要提供外部理想的照明才可正常工作, 因此具有很大的局限性。

为了提高触控探测在多变环境条件下的鲁棒性, 本文提出了一种投影-摄像机显示系统, 该系统采用一种利用手指及其阴影融合情况进行触控探测的新方法。这种方法避免了将手指与其阴影分离, 增强了投影-摄像机显示系统的鲁棒性。此外, 该系统只需要一台普通的摄像机, 不需要其他光源设备, 更加实用、低成本和人性化。

2 投影-摄像机触控系统

2.1 系统简介

该触控系统包含: 用户、电脑、投影机、摄像机和投影屏幕。其中, 用户的手部阴影通过投

影机提供的光源投射到投影屏幕上; 摄像机通过计算机视觉方法用来追踪用户的手及其阴影动作。当用户接触到投影屏幕, 系统首先从摄像机捕获的图像中同时提取出手部及其阴影, 接着探测指尖及其阴影的融合程度, 最后探测指尖位置并进一步通过阴影提供的线索判断指尖与屏幕是否有接触。

如图 1 所示, 系统的处理过程分为如下三个步骤: 第一步是通过比较预测图像和摄像机获取的图像, 将手及其阴影从背景中分割出来。第二步是检查手与其阴影是否完全融合。如果两者之间没有间隙, 就意味着一个触控动作可能即将发生。在这一步中, 因为手与其阴影图像的分离不是必要的, 所以该投影摄像机显示系统具有很强的鲁棒性。第三步是通过探测指尖位置来判定触控位置, 并结合进一步结合阴影信息来判断指尖是否接触到投影屏幕。与其他光学图像触控系统相比, 本系统仅需要一台低成本摄像机就能获取更加灵活、自然与直观的手指信息, 提供一种经济实惠的人机交互方式。

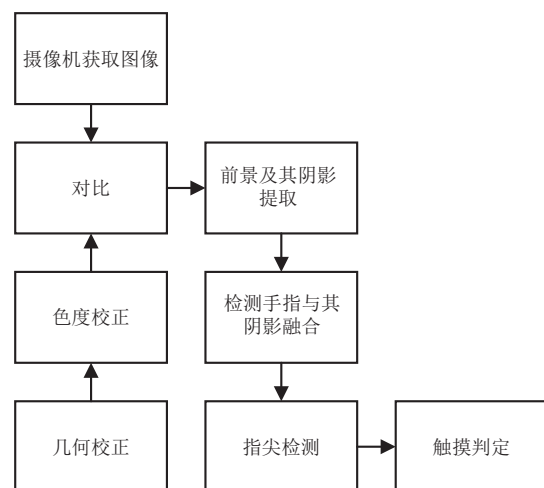


图 1 投影-摄像机系统接触判定流程图

Fig. 1 System block diagram and the process of touch detection

2.2 前景及其阴影提取

第一步是从摄像机获取的图像中提取前景及其阴影。摄像机获取的图像可以通过要投射的画

面预测到。

为了预测摄像机图像，我们需要知道相机视点和输入投影图像下对应像素点的位置关系。这些工作需要几何校正来做。几何校正的步骤如下：(1)从电脑中加载一幅棋盘格图像并传送到投影机；(2)投影机将棋盘格画面投射到投影屏幕；(3)利用一台固定的摄像机获取投影屏幕画面；(4)探测摄像机捕获画面上棋盘格的角点；(5)利用探测到的角点与对应的投射输入画面上棋盘格角点的位置来计算两者图像的 3×3 单应性矩阵 $\mathbf{H}^{[8]}$ ；(6)采用 3×3 单应性矩阵 \mathbf{H} 对摄像机捕获的图像进行几何校正。具体流程如图2所示。

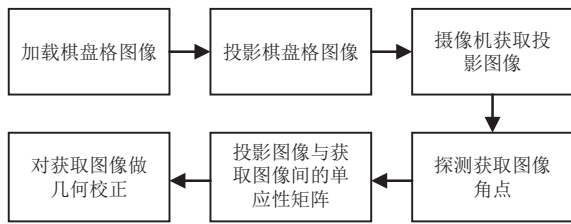


图2 几何校正流程图

Fig. 2 Block diagram of geometric calibration

在上述几何校正的第(4)步中，棋盘格角点探测方法如下：①将RGB图像转换为灰度图像；②找到灰度图像的边缘；③利用Hough变换探测边缘图像的直线；④利用上述直线拟合出四边形；⑤找到四边形的角点。

为预测摄像机图像，还要对投影图像做色度校正^[9]。对每个像素点采用的色度校正模型如下：

$$\mathbf{P} = \mathbf{A}(\mathbf{VI} + \mathbf{F}) \quad (1)$$

其中，

$$\mathbf{P} = \begin{bmatrix} P_R \\ P_G \\ P_B \end{bmatrix}, \mathbf{A} = \begin{bmatrix} A_R & 0 & 0 \\ 0 & A_G & 0 \\ 0 & 0 & A_B \end{bmatrix} \quad (2)$$

$$\mathbf{F} = \begin{bmatrix} F_R \\ F_G \\ F_B \end{bmatrix}, \mathbf{V} = \begin{bmatrix} V_{RR} & V_{GR} & V_{BR} \\ V_{RG} & V_{GG} & V_{BG} \\ V_{RB} & V_{GB} & V_{BB} \end{bmatrix}, \mathbf{I} = \begin{bmatrix} I_R \\ I_G \\ I_B \end{bmatrix} \quad (3)$$

在公式(1)中，矩阵 \mathbf{P} 是预测图像的像素；

矩阵 \mathbf{I} 是投影图像的像素；矩阵 \mathbf{A} 是投影屏幕反射率；向量 \mathbf{F} 是环境光的影响因素； 3×3 矩阵 \mathbf{V} 是颜色混合矩阵，它描述系统中颜色通道间的相互影响。计算 \mathbf{AV} 和 \mathbf{AF} 的主要步骤如下：

(1)投射纯黑色的画面来计算 \mathbf{AF} ；

(2)投射纯红、纯绿和纯蓝画面来计算 \mathbf{AV} 。

获取预测图像以后，前景及其阴影可以通过预测图像及其摄像机图像间的差别来提取出来。如果图像被直接投射在屏幕上，摄像机捕获图像 C 应当与预测图像 P 相同。如果在投影屏幕前有一只手，则投影屏幕的投射画面反射率会发生改变。像素 $[x, y]$ 的反射率变化可以通过反射比 $\alpha[x, y]$ 计算得到：

$$\alpha[x, y] = \frac{C_{g[x, y]}}{P_{g[x, y]}} \quad (4)$$

其中， C_g 是图像 C 的灰度图像； $C_{g[x, y]}$ 是像素 $[x, y]$ 在图像 C_g 中的灰度值； P_g 是图像 P 的灰度图像； $P_{g[x, y]}$ 是像素 $[x, y]$ 在图像 P_g 中的灰度值。

如果没有前置物体，投影屏幕的反射率 $\alpha[x, y]$ 应当接近于1。我们采用以下规则把前景及其阴影从摄像机捕获图像中分离出来。

如果像素 $[x, y]$ 满足以下条件，则该像素点属于前景及其阴影：

$$\alpha[x, y] < 1 - s \text{ 或 } \alpha[x, y] > 1 + s \quad (5)$$

其中， s 是反射率容错率，通常取值 $0.5 \sim 0.8$ 。

如果像素 $[x, y]$ 不属于前景及其阴影，则设置像素 $[x, y]$ 的灰度等级为255。

图3所示为不同光照条件下前景提取实验结果。图3(a)和(b)是在自然光条件下亮背景和暗背景投影图像中的前景提取结果。图3(c)和(d)是在LED光照条件下亮背景和暗背景投影图像中的前景提取结果。图3(e)是暗环境下暗背景投影图像的前景提取结果。从图3可以看出，本文方法在前景提取方面表现出很强的鲁棒性。

2.3 融合判定

我们通过手指与其阴影的融合信息来判断手

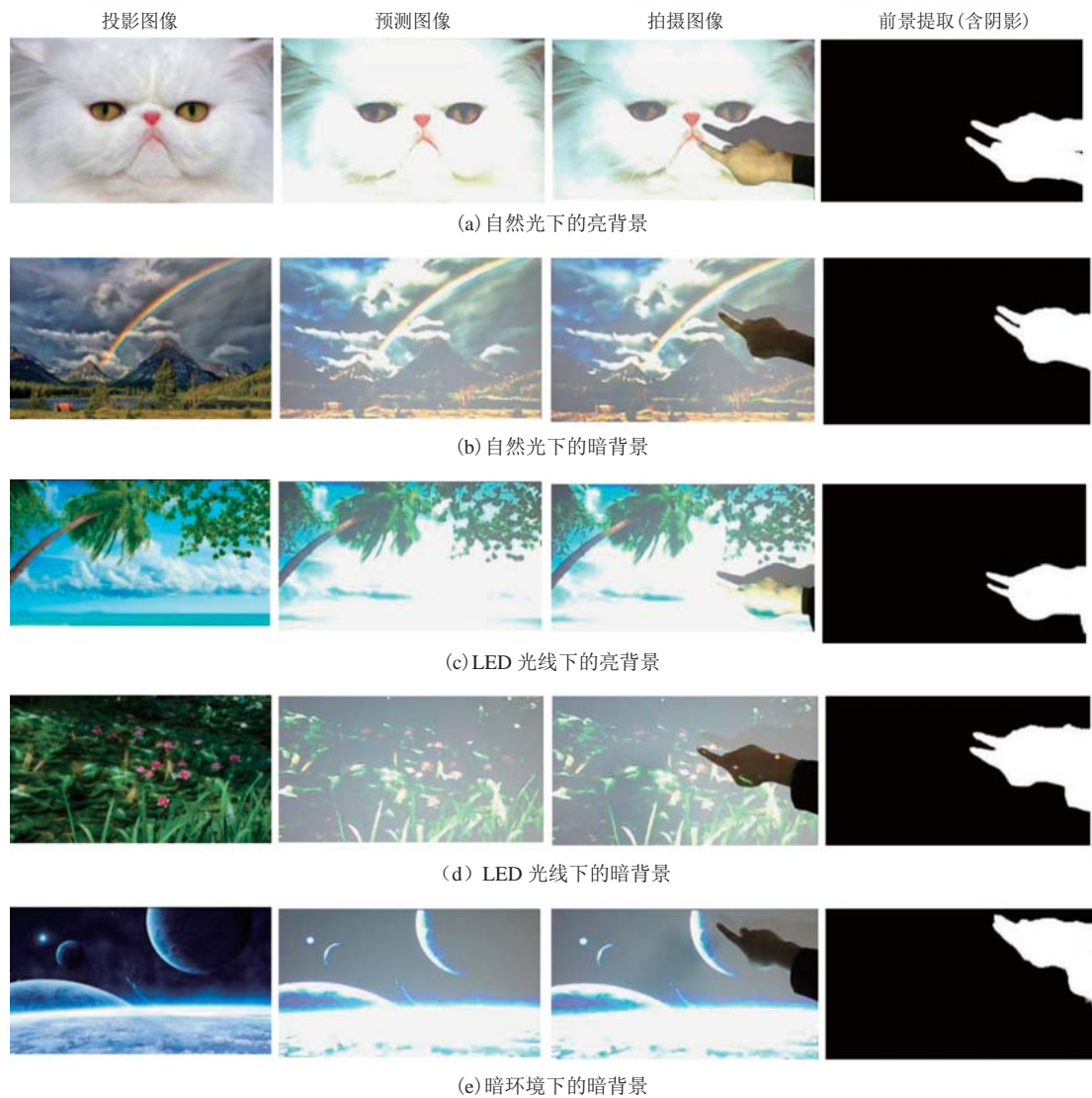


图 3 不同照明环境下前景提取结果

Fig. 3 User region detection results under different lighting conditions

指与投影屏幕的触碰情况。通常情况下, 需要分别探测手指与其阴影的顶点, 并根据顶点之间的距离来判断手指与其阴影的顶点是否已经融合^[4]。然而, 在某些光照条件下, 手指与其阴影像素相同, 这样就很难将两者清晰地分离。为此, 我们提出了一种简单稳固的方法, 不需要解决手指与其阴影的分离问题, 从而增强了投影-摄像机显示系统的鲁棒性: 首先判断手指与其阴影的融合情况; 然后探测指尖位置并进一步寻找投影屏幕

上的触控位置。

我们建立了一个简易竖直线模型来判断手指与其阴影是否融合。融合信息可以被用来判断用户的手指与投影屏幕接触情况。在第 2.2 节, 手指与其阴影被从摄像机捕获的图像中分割出来。竖直线性模型被用来扫描这些分离的图像。如果手指和其阴影没有完全融合, 竖直线性模型上会始终存在一个中部有限的 0 像素部分, 该部分两端的像素灰度值都是 255。如果手指与其阴影完

全融合，一部分竖直线性模型中心部分的像素灰度值是 255；对于其他竖直线性模型，从中心到末端至少有一半的竖直线性模型上一半的像素灰度值为 0。

2.4 触控检测

利用 2.3 节所描述的方法，当探测到指尖与其阴影没有完全融合时，即手指与其阴影之间还存在一定的间隙。在这种情况下，可以判定手指与投影屏幕之间还存在一定的距离，即手指与投影屏幕间没有触碰交互动作。当探测到手指与其阴影融合以后，如上面 2.3 节所描述，指尖与其阴影即使在顶端也不会存在向内凹陷的状况。即指尖与其阴影在指尖顶端融合成一个连续的整体。这种情况是手指与投影屏幕相接触的一个必要条件。当指尖与其阴影融合以后，下一步是找到指尖的位置用来判定触控。我们提出了一种精准快速探测指尖的方法，如图 4 所示，该方法描述如下：

(1) 二值化提取出的前景手部图像；

(2) 提取二值图像的轮廓；

(3) 以轮廓图像上的像素点为中心，构建正方形搜索块；

(4) 遍历轮廓点，计算搜索块内包含手部区域像素个数；

(5) 如果搜索块内手部像素个数比前一个小，则当下的搜索块取代上一个作为优值，遍历轮廓点，找到最优值的点作为指尖。

探测到融合了阴影的指尖位置以后，并不能确保在此时此位置手指尖接触到投影屏幕。在该投影-摄像机系统中，为了进一步判断指尖与投影屏幕是否相接触，在探测到的指尖位置，获取搜索块中融合了指尖与其阴影像素的总数 $N_{tip-shadow}$ ，并采用以下公式在我们的投影摄像机系统中进行触控检测：

$$DT_{touch} = \begin{cases} 1, & \text{if } N_{tip-shadow} \leq N_{threshold} \\ 0, & \text{if } N_{tip-shadow} > N_{threshold} \end{cases} \quad (6)$$

其中， DT_{touch} 为触控探测因子。当 $DT_{touch}=1$ ，表

示手指触碰到投影屏幕。当 $DT_{touch}=0$ ，表示投影屏幕上没有触控动作。 $N_{threshold}$ 是一个阈值。如公式 (6) 所示，如果搜索块中的手部及其阴影像素总数 $N_{tip-shadow}$ 小于或等于阈值，表示手指与其阴影完全融合，此时屏幕上有触控动作发生。如果搜索块中的手部及其阴影像素总数 $N_{tip-shadow}$ 大于阈值，表示手指与其阴影没有完全融合。此时，指尖与投影屏幕之间存在一定距离，即没有触碰动作发生。

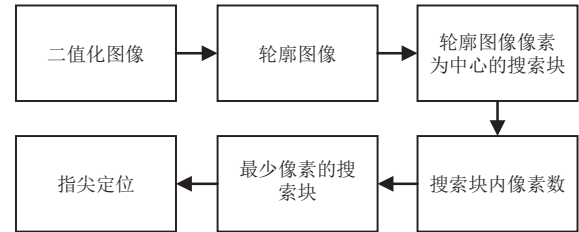


图 4 指尖定位方法

Fig. 4 Flowchart for fingertip detection

3 实验结果与讨论

我们通过实验来评估所提出的投影-摄像机系统的性能。如图 5 所示，实验平台包括：一台 3.6 Ghz CPU 和 4.0 G RAM 的电脑，一台分辨率为 1440×1050 的索尼投影仪，一台摄像机和一个投影屏幕。投影机与摄像机位于同一水平线，

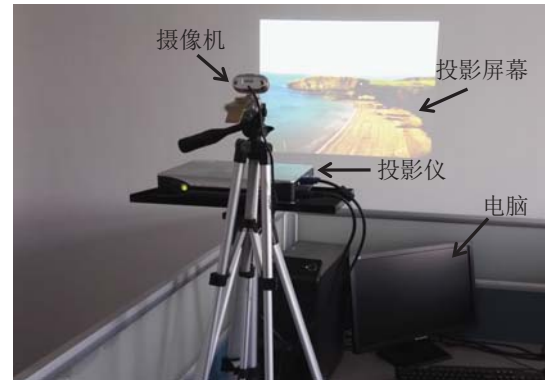


图 5 投影-摄像机系统

Fig. 5 The interactive projector-camera system

相隔 200 mm。投影机与屏幕的距离是 1 500 mm。投影屏幕尺寸为 50 英寸。

图 6 所示为本文提出的投影-摄像机系统在工作中不同阶段的状态。在图 6(a)中, 手指与其阴影有明显的距离, 说明手指与投影屏幕还很远, 计算机鼠标没有被触发。在图 6(b)中, 手指与其阴影已经融合, 如 2.3 节所述, 手指与其阴影之间不存在缝隙。但此时手指尖与投影屏幕之间仍存在一定的距离, 计算机的鼠标并未被触发。在图 6(c)中, 手指与其阴影已融合, 如 2.4 节所述, 探测到指尖位置, 并且以指尖为中心的搜索块内手指与其阴影像素总数低于设定的阈值, 此时, 手指碰触到投影屏幕。本系统中, 实验表明当指尖与投影屏幕之间的距离小于 ~ 0.5 cm 时会检测到触控。

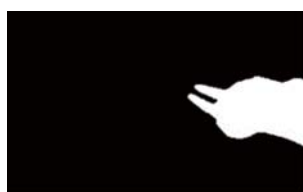
表 1 显示了 Hardenberg 和 Bérard^[10]及 Song^[11]的方法与本文方法在触控探测准确率方面的对比。在 200 帧图像中, 分四组对自然光和 LED 光下的亮背景和暗背景分别进行了比较实验。结果显示, 本文方法的触控探测准确率(90.0%)高于前两者的方法(55.5% 和 88.0%)。我们使用了预测方法来比较摄像机捕获图像与投影图像, 进而较为精确地提取前景图像。

表 2 显示了该系统关键模块的执行时间。对 200 帧图像进行实验, 处理一帧所需要的平均时间约为 30 毫秒, 因此我们可以以每秒 33 帧的速

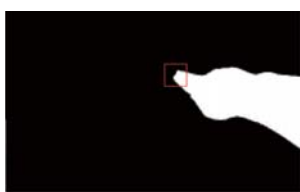
度来探测触控位置。该系统总的执行时间方差为 0.7, 总的执行时间的极差为 5。该系统对视频流图像, 可以满足 25 帧/秒的正常播放速度下对视频流的实时交互, 不会产生明显的延迟。也就是说, 本文方法能够在普通的计算机系统中实现实时计算。

4 结 论

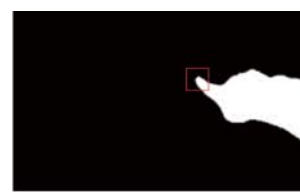
本文提出一种由单台摄像机实现的鲁棒性交互式投影显示系统。在该系统中, 用户能够将任何平面作为虚拟触控屏幕, 并直接通过手指与该屏幕进行交互。该系统主要有三大优势。首先, 基于几何校正与颜色校正的前景手势提取方法受到外界环境状况的影响较小, 使得用于交互的手势提取更加精确和具有普适性。其次, 手指与其阴影不需要做分离处理, 避免了由于手指与其阴影分离不彻底而造成的错误, 由此增加了系统的鲁棒性和准确性。最后, 通过基于搜索块方法的手指指尖定位与指尖像素阈值操作提高了指尖位置检测与触控检测的精度。本文方法存在的局限性在于提取的复杂前景影响了对指尖触控的判断, 降低了触控检测正确率。后期解决的方法是根据用户使用投影交互的特点, 针对手部所在的大致区域进行触控判定。这样可以进一步提升触控检测的准确率和系统的鲁棒性。



(a) 手指与其阴影未融合, 此时鼠标未被触发



(b) 手指与其阴影已融合, 但未完全融合, 此时鼠标未被触发



(c) 手指与其阴影完全融合, 指尖部位搜索块内的手指与其阴影的像素总数满足阈值条件, 此时触发鼠标动作

图 6 投影-摄像机系统在不同阶段的工作状态

Fig. 6 Working condition of projection-camera systems at different stages

表1 触控探测准确率

Table 1 Touch detection accuracy rates

方法	光线类型	检测结果	帧数	各类型准确率 (%)	平均准确率 (%)
Hardenberg 和 Bérard ^[10] 的方法	自然光下亮背景	正确	35	70	55.5
		不正确	15		
	自然光下暗背景	正确	23	46	
		不正确	27		
	LED 光线下亮背景	正确	32	64	
		不正确	18		
	LED 光线下暗背景	正确	21	42	
		不正确	29		
Song 等 ^[11] 的 方法	自然光下亮背景	正确	49	98	88
		不正确	1		
	自然光下暗背景	正确	44	88	
		不正确	6		
	LED 光线下亮背景	正确	44	88	
		不正确	6		
	LED 光线下暗背景	正确	39	78	
		不正确	11		
本文方法	自然光下亮背景	正确	49	98	90
		不正确	1		
	自然光下暗背景	正确	43	86	
		不正确	7		
	LED 光线下亮背景	正确	47	94	
		不正确	3		
	LED 光线下暗背景	正确	41	82	
		不正确	9		

表2 执行时间

Table 2 Execution time measured

模块	平均时间(毫秒)	方差	极差
前景提取	24	0.8	5
初次融合判定	3	0	0
指尖识别及触控判定	3	0	0
合计	30	0.7	5

参考文献

- [1] Cheng J, Wang Q, Song R, et al. Fingertip-based interactive projector-camera system [J]. Signal Processing, 2015, 110: 54-66.
- [2] Park YS, Park SH, Kim TG, et al. Implementation of gesture interface for projected surfaces [J]. KSII Transactions on Internet and Information Systems, 2015, 9(1): 378-390.
- [3] He M, Cheng J. Self-adaptive coding-based touch detection for interactive projector system [C] // 2014 4th IEEE International Conference on Information Science and Technology (ICIST), 2014: 656-659.
- [4] Dung LR, Lai GY, Wu YY. Shadow touching for interactive projectors [C] // 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, 2013: 1798-1802.
- [5] Dung LR, Huang RY. A shadow touching technique for interactive projector devices [C] // International

