

# 基于 Kinect 骨骼信息与深度图像的指尖点检测

张登攀, 李国玄, 王黎阳

(河南理工大学 机械与动力工程学院, 河南 焦作 454000)

**摘要:** 针对普通摄像头手势识别系统易受复杂环境和光照条件等因素影响, 存在对指尖点的漏判、误判问题, 提出一种基于 Kinect 骨骼信息与深度图像的掌心点提取和指尖点检测的手势识别方法; 在 DRVI 平台上创建 Kinect 的接口控件, 对 Kinect 传感器获取人体骨骼信息和深度图像进行分析, 采用了坐标映射、图像分割、距离变换的关键技术和方法从深度图中分割出手势部分区域, 对手势区域形态学处理, 结合凸包和 K-曲率算法检测不同手势中指尖点的个数和位置, 计算不同手势凸包轮廓上的点集生成的 HOG (Histogram of Oriented Gradient) 特征描述子, 最后利用特征描述子对预定的 6 种数字手势进行识别; 经实验测试可以在复杂环境和不同光照情况下正确识别指尖点。

**关键词:** 手势识别; 骨骼信息; 深度图像; HOG 特征描述子; 指尖检测

## Fingertips Recognition Based on Kinect Skeleton Information and Depth Data

Zhang Dengpan, Li Guoxuan, Wang Liyang

(School of Mechanical and Power Engineering, Henan Polytechnic University, Jiaozuo 454000, China)

**Abstract:** The hand gesture recognition system is susceptible to the light conditions and complex environments by using ordinary camera, so the fingertips are always missed and misjudged. To solve these problems, a new method for fingertip detection and palm point extraction of hand gesture is proposed based on Kinect skeleton information and depth image. Creating a Kinect control on the DRVI platform and analyzing the information of the skeleton and depth image is acquired by Kinect sensor, the technology and method of coordinate mapping, image segmentation and distance conversion can be used to segment the hand area on the depth image, which need morphological processing. After the number of fingertips and the location for hand gesture being detection with combining convex hull and K-curvature algorithm, calculating the HOG feature descriptor generated by the point set on the contours of different gesture hulls. Finally, the HOG feature descriptor is applied to identify six scheduled number hand gesture. The experiment results show that the proposed method can identify fingertips in complex environments and different lighting conditions.

**Keywords:** hand gesture recognition; skeleton information; depth image; HOG feature descriptor; fingertip detection

## 0 引言

手势识别作为人机交互中的一项关键技术, 是当前计算机视觉领域的研究热点。其中的指尖检测是该技术的一个重要基础组成部分, 已在多个领域得到广泛应用, 如智能家居、人机协同、机器人智能控制等。早期的手势识别技术多采用普通摄像头获取图像信息, 该技术采集的图像易受到不同光照条件和复杂背景的影响, 采用光电传感器等设备采集图像成本高、约束条件多<sup>[1-2]</sup>。传统的基于数据手套的手势识别方法识别率虽然不受光照和复杂背景影响, 但是需要用户穿戴复杂设备影响交互的自然特性; 基于肤色模型在 YCrCb 或 HSV 颜色空间的聚类特性分割出手势区域易受光照的影响<sup>[3]</sup>; Simen Andresen, Stokkeland 等人利用 HSV 颜色阈值分割提取手掌部分, 循环遍历每一个像素点的深度值, 检测矩阵中突变的像素点, 但是受复杂环境影响较大。针对上述问题, 提出了一种基于 Kinect 骨骼

信息和深度图像的指尖点检测方法, 重点研究内容如下:

- 1) 在 DRVI 平台创建 Kinect 接口控件;
- 2) 深度图像跟踪手部区域;
- 3) 对不同手势凸包点进行 K-曲率计算, 生成相应的 HOG 特征描述子。

## 1 系统结构及原理

手势识别系统的结构包括图像的采集、手势跟踪与定位、手势分割与预处理、手势特征提取和指尖检测五个部分, 其中手势特征提取和指尖检测是识别系统的关键, 该部分的主要功能是分析设备采集的图像, 提供原始图像数据, 并根据不同用户对输入参数进行实时调整。

目前, 国内采集深度图像的设备主要有两种, 一种是因特尔的 RealSense, 具有体积小, 方便携带但深度图的质量较差, 噪声非常大; 另一种是采用 PrimeSense 公司技术的体感摄像头, 此摄像头芯片价格便宜, 能耗低, 而且其微小的体型适用于笔记本以及智能手机等终端。Kinect 最初作为 XBOX 360 主机的外设, 旨在让用户脱离控制器, 通过语音或手势来操作 XBOX 的系统界面,

是以 PrimeSense1080 系统级芯片技术为核心的图像采集装置, 由彩色摄像头、深度摄像头等部件构成, 还搭配

收稿日期: 2018-08-30; 修回日期: 2018-09-19。

基金项目: 河南省科技攻关项目(142102210051); 河南省教育厅科技攻关项目(13A460338)。

作者简介: 张登攀(1975-)男, 河南叶县人, 博士, 副教授, 主要从事精密测量技术的教学与研究工作。

追焦技术, 底座马达会随着对焦物体移动跟着转动, 内建阵列式麦克风, 由四个麦克风同时收音比后消除杂音, 并通过其采集声音进行语音识别和生源定位, 如图 1 所示。

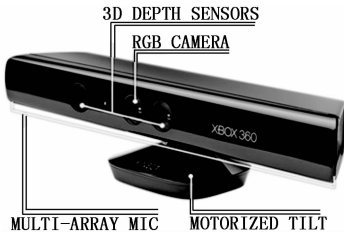
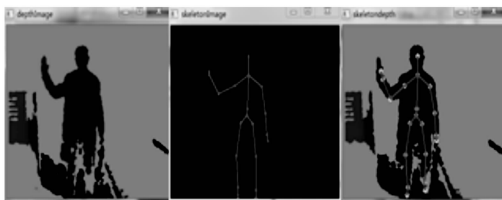


图 1 Kinect 结构

Kinect 传感器的主要功能就是可以产生三维数据, 获取深度图像的方法是将结构光投射至场景, 对获取的红外数据进行计算处理, 并由传感器捕获相应的带有结构光的图案, 采集的深度图像如图 2 (a) 所示; 微软将侦测到的 3D 深度图像转换到骨架追踪系统, 骨骼追踪技术是 Kinect 的核心技术, 它可以准确标定人体的 20 个关键点, 并能对这 20 个点的位置进行实时追踪。利用这项技术, 可以开发出各种基于体感人机交互的有趣应用如图 2 (b) 所示。Kinect 捕获的骨骼数据属于三维空间数据, 将坐标信息映射到所获取的深度图像上, 关节用圆形显示, 需要采用坐标映射的方法将节点从骨骼数据三维空间转换到深度数据的二维空间, 根据手掌节点的坐标确定手势的大致范围, 如图 2 (c) 所示。



(a) 深度图像 (b) 骨图像 (c) 映射图

图 2 Kinect 采集的图像

## 2 手势分割

手势分割作为手势识别的前提, 手势分割的好坏直接影响手势识别的识别率。一般的分割方法大致分为基于直方图的分割、基于局部区域信息的分割、基于颜色等一些物理特征的分割方法, 但这些分割技术大都需要对背景、用户以及视频采集加以约束, 受背景复杂度和光照变化的影响较大。传统的视觉的手势提取算法主要有基于肤色的手势提取、基于运动的手势提取及基于轮廓的手势提取等, 手部区域的提取通常包括手势分割、手势建模、特征匹配等过程<sup>[4]</sup>。为了更好地检测手势区域掌心点和指尖点, 首先需要把手掌从复杂的背景中分离出来, 在深度图中利用骨骼追踪系统确定手部、腕部坐标及灰度值, 然后搜索手腕节点的端点, 采用最近邻法则和阈值法则从深度图像中分离出手部区域<sup>[5]</sup>。Kinect 捕获到的腕坐标  $(X_w, Y_w, Z_w)$  和手部关节点坐标  $(X_h, Y_h, Z_h)$  是骨骼空间数据, 根据骨

骼空间数据计算手部、腕部在深度图的坐标  $(D_h, D_h)$  和  $(D_w, D_w)$ , 可以得到相应的深度值  $d_h, d_w$ , 对深度图 2 维数据可表示为  $addr(M_{ij}) = M.data + M.step[0] * i + M.step[1] * j; M.step[]$  表示元素地址  $(i_0, \dots, i_{M.dims-1})$ , 其中  $0 << i_k < M.size[k]$ , 式 (1) 分割出手部区域:

$$seg(x, y) = \begin{cases} 0, 1 > Y_w \text{ 或 } j > Y_h \text{ 或 } i < X_w \text{ 或 } i < X_h \\ maxv - 4, \text{ otherwise} \end{cases} \quad (1)$$

式中,  $i$  表示深度图像手势区域的行数,  $j$  表示深度图像手势区域的列数,  $maxv$  表示深度图中轮廓像素的最大值。

本文在 DRVI 平台上做了 XBOX 的接口控件, DRVI 是采用软件总线开放结构和 COM/DCOM 组件的即插即用特性来设计的具有计算机硬件模块化组装特点的面向用户的可在线编程、调试和重组的新型虚拟仪器技术。从深度图分割出来的手势区域边缘存在很多图像缺陷影响手势识别的结果, 需要对手势区域进行高斯滤波处理, 如图 3 所示。

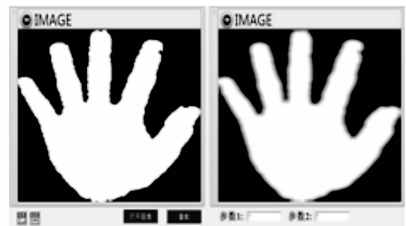


图 3 高斯滤波处理

图 4 (a) 中被黑色矩形框圈出来的深度图像为手部区域, 图 4 (b) 是用最近邻法则分离出来的深度手势图像。然后采用形态学处理中的开闭运算对其进行去除噪声处理, 能有效的消除和抑制噪声, 达到图像平滑的效果<sup>[6]</sup>, 经过滤波和形态学处理后的手势图像边缘锯齿状明显减少, 如图 4 (c) 所示。



(a) 手势跟踪 (b) 深度图 (c) 高斯处理图

图 4 手部区域图

## 3 手势识别

在计算机科学和人机交互领域中, 手势识别是通过数学算法来识别人类手势的一个议题。手势识别可以来自人的身体各部位的运动, 但是一般是指手的运动或者手形识别, 根据交互功能的要求有动态的识别, 比如左右挥手, 还有静态的手形识别, 比如识别伸出手指数, 用户可以使用简单的手势来控制或与设备交互, 让计算机理解人的意图, 此手势识别系统主要研究静态手形的识别。

在连续情况下, 矩特征主要表现了图像区域的几何特

性,具有旋转、平移、尺度等特性的不变特征,可以据此此特征来对图像进行分类。图像函数为  $f(x, y)$ , 那么图像的  $p+q$ 阶中心矩和  $p+q$ 阶几何矩分别定义为: 如式 (2) 所示:

$$\begin{aligned} \mu_{pq} &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x-\bar{x})^p (y-\bar{y})^q f(x, y) dx dy \\ m_{pq} &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x^p y^q f(x, y) dx dy \end{aligned} \quad (2)$$

其中:  $\bar{x}$  和  $\bar{y}$  代表图像的重心, 如式 (3) 所示:

$$\begin{aligned} \bar{x} &= m_{10}/m_{00} \\ \bar{y} &= m_{01}/m_{00} \end{aligned} \quad (3)$$

利用二阶和三阶归一化中心矩构造了 7 个不变矩  $M_1 \sim M_7$ : 如式 (4) 所示:

$$\begin{aligned} M_1 &= \delta_{20} + \delta_{02} \\ M_2 &= (\delta_{20} - \delta_{02})^2 + 4\delta_{11}^2 \\ M_3 &= (\delta_{30} - 3\delta_{12})^2 + (3\delta_{21} - \delta_{03})^2 \\ M_4 &= (\delta_{30} + \delta_{12})^2 + (\delta_{21} + \delta_{03})^2 \\ M_5 &= (\delta_{30} + 3\delta_{12})(\delta_{30} + \delta_{12})[(\delta_{30} + \delta_{12})^2 - 3(\delta_{21} + \delta_{03})^2] \\ &+ (3\delta_{21} + \delta_{03})(\delta_{21} + \delta_{03})[3(\delta_{30} + \delta_{12})^2 - (\delta_{21} + \delta_{03})^2] \\ M_6 &= (\delta_{20} + \delta_{02})[(\delta_{30} + \delta_{12})^2 - (\delta_{21} + \delta_{03})^2] + 4\delta_{11}(\delta_{30} + \delta_{12})(\delta_{21} + \delta_{03}) \\ M_7 &= (3\delta_{21} - \delta_{03})(\delta_{30} + \delta_{12})[(\delta_{30} + \delta_{12})^2 - 3(\delta_{21} + \delta_{03})^2] \\ &- (\delta_{30} - 3\delta_{12})(\delta_{21} + \delta_{03})[3(\delta_{30} + \delta_{12})^2 - (\delta_{21} + \delta_{03})^2] \end{aligned} \quad (4)$$

HOG 通过计算和统计图像局部区域的梯度方向直方图, 如图 5 所示。

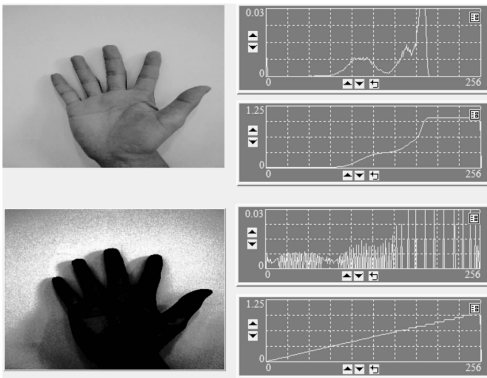


图 5 标准图像直方图

与 Hu 矩的 7 个特征构成一组特征向量, 每一种手形对应一个 XML 文件, 生成对应的标签。根据 SVM 分类思想训练出一个分割超平面, 即分类的决策边界, 当 Kinect 场景内出现与预定义手形最相近的特征, 即被贴上对应的标签, 如图 6 所示。

## 4 指尖检测

### 4.1 手部区域边缘检测

稳定的指尖检测算法可以应用在人机交互系统中, 提升交互技术的灵活性, 用 Graham Scan Algorithm 扫描算法寻找边缘轮廓点中横坐标 (纵坐标) 最小的点作为第一



图 6 识别效果图

个点;

1) 其他的点按照极角大小顺时针排列, 如果有共线的取最近的那一个, 依次将  $P_0, P_1, P_2$  三个坐标压入栈;

2) 比较  $\vec{P_0P_1}$  和  $\vec{P_0P_2}$  与极坐标轴夹角比较大小较小的  $P_1$  是凸包点, 对于剩余的点,  $i$  依次从 3 到  $P_i$  ( $i=0, 1, 2 \dots$ ) 比较相邻点序的极角, 较小的压入栈中, 如图 7 所示。

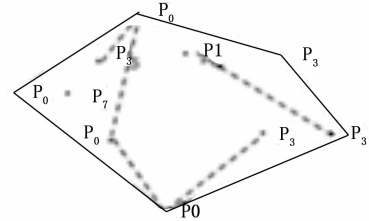


图 7 点集 P 的凸包

边缘是指图像局部强度变化最显著的部分, 边缘主要存在目标与目标、目标与背景、区域与区域之间。多次实验得阈值为 200 效果最好, 如果  $T_1 < 200$  则判定为伪轮廓, 本文采用 Canny 边缘检测算法是对信噪比和定位之乘积的最优化逼近算子, 用于手势边缘的双阈值算法检测和连接边缘<sup>[8-9]</sup>, 如图 8 (a) 是手部区域的轮廓图。为去除手腕信息的干扰, 首先用 ApproxpolyDP 逼近多边形算法只找到轮廓上的关键点, 目的就是缩小指尖点所在轮廓上的范围, 只是此轮廓粗犷一些<sup>[10-11]</sup>; 可以通过扫描图像法找到手势图像最上面的点 top, 以 top 为基点沿逼近多边形轮廓点顺时针逆时针各区 30 个像素点, 以  $P_{top}, P_{top+30}, P_{top-30}$  三点作圆, 如图 8 (b) 所示。从实验结果可以看出, 这两种约束条件可以很好的去除冗余的手腕信息。

### 4.2 计算手势区域掌心点位置

指尖检测的一个关键问题就是掌心点提取, 掌心点包含手势许多重要的信息, 比如指尖点的位置、方向以及手势的种类等。目前, 在图像处理中寻找掌心点有图像矩和距离变换等两种常用算法。在图像处理、计算机视觉和相关领域中, 通过图像矩可以轻易得到图像质心和关于其方向的信息, 图像矩是沿  $x, y$  方向对轮廓边界像素的积分, 利用零阶矩和一阶矩求重心, 深度图中的手部区域可以用手部节点粗略表示手势位置但不能准确代表手掌掌心, 手势骨骼节点根据人体姿态不同发生抖动<sup>[12-13]</sup>; 距离变换的基本思想就是计算图像中像素点到最近零像素点的距离,



(a) 手部区域轮廓图 (b) 手部区域轮廓图

图 8 手势轮廓图

象棋格距离 (Chessboard Distance), 公式 (5) 如下:

$$D = \max(|x_2 - x_1|, |y_2 - y_1|) \quad (5)$$

此距离变换减少了计算量, 将二值化手掌图像内的手掌部分设为白色, 外部区域设为黑色, 该算法可以精确找到手掌掌心, 如图 9 所示。



图 9 距离变换掌心图

计算一个图像中非零像素点到最近的零像素点距离, 通过合理的阈值限制可以得到去除手指的图像, 手部质心即为该图像的几何中心, 如图 10 所示手势的质心及坐标。



图 10 手势质心及坐标

### 4.3 指尖点检测

指尖检测的效果影响交互的丰富度、灵活性。Lee 等人通过计算掌心与轮廓的边缘距离检测单个指尖的位置实现虚拟交互<sup>[14]</sup>。本文采用手势轮廓凸包的缺陷和轮廓点曲率结合的方法检测指尖点, 计算不同手势指尖点的 HOG 特征描述子, 并识别 0~5 的数字手势。利用计算机视觉库函数 convexityDefects 检测出手势轮廓与逼近多边形围成的凹缺陷和凸包, 在此创建一个迭代器对顺序或关联容器类型中的每个元素进行连续访问, 如图 11 所示 convexity hull 与手掌之间的部分为 convexity defects。每个 convexity defect 区域有四个特征量: 起始点 (startPoint), 结束点 (endPoint), 距离 convexity hull 最远点 (farPoint), 最远点到 convexity hull 的距离 (depth), A~F 区域为检测出的缺陷, 则 startPoint、endPoint 为指尖候选点。基于曲率的指尖检测法主要依靠计算手部区域中类指尖点的近似 K 曲率,

蓝色折线是手指细化的结果, 目的就是为了让图像信息的特征更加突出, 细化算法一定要满足收敛性、保持原图的基本形状及减少笔画相交处的畸变等, 式 (6) 所示:

$$y = p_0 * 2^0 + p_1 * 2^1 + p_2 * 2^2 + p_3 * 2^3 + p_4 * 2^4 + \dots, \quad (6)$$

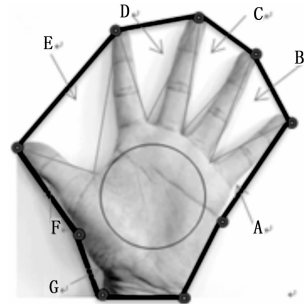


图 11 手势缺陷图

$p_0, p_1, p_2, p_3, p_4$  是细化后的像素点, 如图 12 所示, 定义  $P(i)$  为手势轮廓上一点, 按照逆时针和顺时针遍历  $k$  个像素的点定义为  $P(i+k)$  和  $P(i-k)$ 。利用  $P(i+k)$  与  $P(i)$  形成向量  $\vec{P}_1$ ,  $P(i-k)$  与  $P(i)$  形成向量  $\vec{P}_2$ , 那么点  $P(i)$  处的曲率  $K$  即是这两个向量夹角的余弦值, 如式 (7) 所示:

$$K = \frac{\vec{P}_1 \cdot \vec{P}_2}{|\vec{P}_1| |\vec{P}_2|} \quad (7)$$

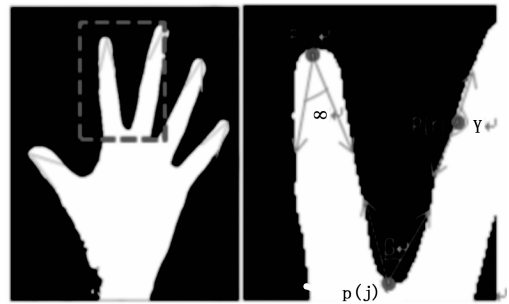


图 12 K 曲率算法

一般在指尖点位置的曲率达到最大, 曲率  $K$  的取值范围  $(-1, 1)$ , 在其范围中设定阈值  $T$ ,  $K \gg T$  此点才可能是类指尖点<sup>[15-16]</sup>, 需要满足式 (8):

$$\begin{cases} K \gg K(P_{i-r}) \\ K \gg K(P_{i+r}) \end{cases} \quad (8)$$

目前检测出来的点称为“类指尖点”可能是指尖点、手指之间的凹槽点或者手掌末端与手臂连接处的点, 需要添加一个过滤条件排除凹槽点:

1) 由图 12 所知, 以轮廓上任意一点  $P(i)$  前后相隔  $m$  个像素点形成的夹角, 如果夹角  $25^\circ < \alpha < 35^\circ$ , 则此点为指尖点, 如果夹角  $\gamma$  接近  $180^\circ$  则此点在手指轮廓上; 如果夹角  $\beta$  大于  $50^\circ$  则判断为指尖凹槽点;

2) 如图 13 所示可以比较单指尖点到掌心点的距离, 如果指尖点的到掌心的距离大于 1.5 倍相邻凹槽点到掌心

点距离, 则此点为指尖点, 否则没有手指伸出, 此方法既可以检测单指尖又可以检测多指尖;

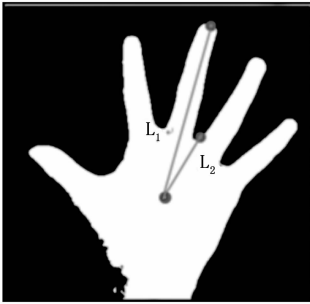


图 13 重心距离法

3) 由图 11 知凹缺陷深度应该大于掌心圆半径, 但小于手部轮廓最小外接圆半径, 经过五指伸开近似正方形这一条件过滤掉手腕干扰, 如果类指尖点符合公式, 则此点为凹槽点, 如图 14 所示。

$$\begin{cases} angle < 70^\circ \\ length > 0.5l_w \end{cases} \quad (9)$$

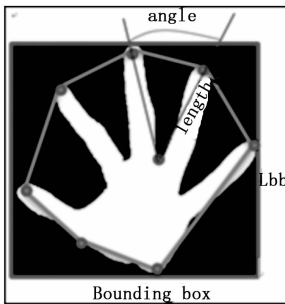


图 14 凹槽点检测法

把检测到的凸轮廓的凸包点储存在 hullpoint 向量容器中, 遍历窗口中所有轮廓, 算法中定义 ofstream 类型, 把 .txt 里面的坐标点导入 origin 数据处理软件中生成散点图, 如图 15 (a) 所示, 在一幅图像中, HOG 利用梯度或边缘的方向密度分布能够很好地描述局部目标区域的特征的思想, 对凸包点集的梯度信息作出统计, 生成 HOG 特征描述, 图像的梯度针对的是每一个像素计算得到, 水平边缘梯度算子:  $[-1, 0, 1]$ , 垂直边缘算子  $[-1, 0, 1]^T$ , 图像中像素点  $(x, y)$  的梯度为:

$$\begin{cases} g_x(x, y) = h(x+1, y) - h(x-1, y) \\ g_y(x, y) = h(x, y+1) - h(x, y-1) \end{cases} \quad (10)$$

式中,  $g_x(x, y)$ ,  $g_y(x, y)$ ,  $h(x, y)$  分别表示输入图像像素点  $(x, y)$  处的水平方向梯度、垂直方向梯度和像素值。

由式 (10) 得, 类指尖点  $(x, y)$  的幅值和梯度方向分别为:

$$\begin{aligned} g(x, y) &= \sqrt{g_x(x, y)^2 + g_y(x, y)^2} \\ \varphi(x, y) &= \tan^{-1} \left( \frac{g_y(x, y)}{g_x(x, y)} \right) \end{aligned} \quad (11)$$

然后在 cell 中进行方向梯度直方图的构建, 统计局部图

像信息并进行编码, 如图 15 (b) 所示:

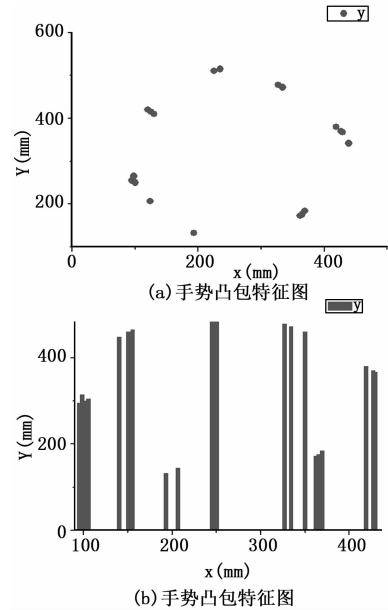


图 15 HOG 手势特征图

每种不同的手势帧生成对应的 HOG 特征描述向量, 其本质是扩大了量化间隔, 形成了一片相同的区域, 各区域之间有明显的轮廓<sup>[17-18]</sup>, 形成最终的 feature vector, 所以此特征就描述了检测窗口的图像内容。对不同的手势进行归一化处理, 引入  $\vec{r}$  表示一个没有被归一化的向量, 此时包含检测窗口的所有信息  $\|\vec{r}\|_n$  表示  $\vec{r}$  的  $n$  阶范数, 用  $\epsilon$  表示一个很小的常数  $w = \frac{\vec{r}}{\sqrt{\|\vec{r}\|_2^2 + \epsilon^2}}$ ,  $w$  表示归一化因子, 该标准形式图像对平移、旋转、缩放等仿射变换具有不变特性, 此算法把图像特征值缩小到一个很小的范围, 增强其稳定性。

### 5 实验结果与分析

为了验证本文指尖检测和手势提取算法的可靠性和准确性, 实验采用微软 Kinect XBOX 360 一代摄像头采集深度图像和骨骼数据, 深度图像分辨率为  $320 \times 240$ 。基于 Microsoft Visual Studio 2010 环境, 通过配置计算机开源视觉库 OpenCV2.4.9、Kinect SDK1.8, DRV14.5 控件接口, 用 C++ 进行编程。本文对指尖数目 0~5 的 6 种情况进行测试, 识别效果如图 16 所示。

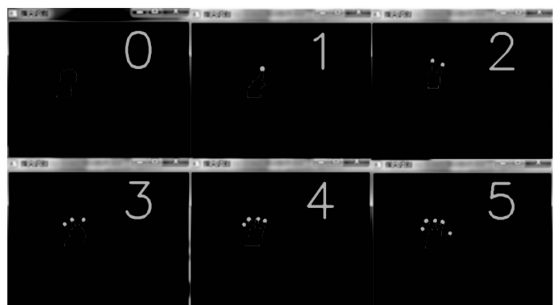


图 16 不同数目指尖识别效果

针对图 16 所示的 6 种数字手势进行识别,识别邀请 5 位实验者进行试验测试,每位实验者在不同光照下对每个手势分别做 10 次,总共采集 300 张样本图像,根据测试的结果每种手势的识别率如表 1 所示:

表 1 6 种数字手势检测结果

指尖数目	提取帧数	误检帧数	手势识别率
0	50	0	100%
1	50	4	92%
2	50	1	98%
3	50	3	94%
4	50	3	94%
5	50	0	100%

由图 16 和表 1 可以看出,数字手势 0 和数字手势 5 识别率最高,对于数字 5 五指伸开时各手指之间间距较大相互不影响,数字手势 0 无指尖特征;数字手势 2 只有两个缺陷起始点和一个凹槽点识别率也比较高;数字手势 3 和 4 容易识别混淆手势发生抖动时检测不到指尖点导致错检漏检;由于数字手势 1 无轮廓缺陷点无法使用曲率算法根据中心距离法求得,所以漏检率较高。图 17 (a) 和图 17 (b) 是在不同光照和复杂环境下检测的结果。

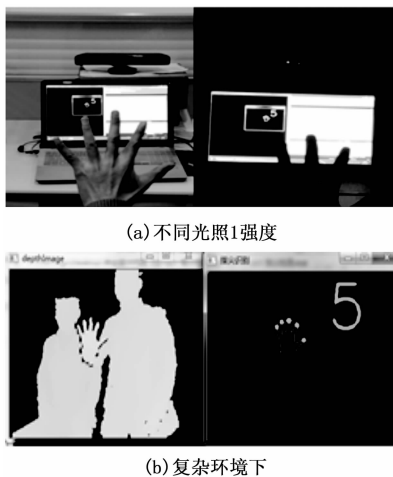


图 17 不同光照和复杂环境识别结果

## 6 结语

本文首先分析了手势识别研究的背景和意义,重点对手势分割和指尖检测的方法和关键技术做了深入研究,主要工作为:在静态手势识别方面利用 Kinect 骨骼数据快速定位手势在深度图的位置;在 DRVI 平台创建 XBOX 接口控件快速处理图像,减小运算量,利用设定阈值去除为轮廓,距离变换处理寻找掌心点;利用凸包和 K-曲率结合方法寻找指尖点,生成相应的特征描述算子。本文的研究内容可以解决不同光照和复杂背景下手势识别的难题,避免采用肤色转换算法中的类肤色区域的干扰。当人距离 Kinect 较远或者坐着的时候手掌节点发生抖动可能会导致跟

踪失败,而且对预定义的手势种类有限,所以该算法还需进一步深入研究。研究成果可以将基于视觉的指尖检测这一新型技术融入到机器人控制之中,简化机器人控制的复杂度,提高人与机器人之间的交互性。

## 参考文献:

- [1] 高晨,张亚军. 基于 Kinect 深度图像的指尖检测与手势识别 [J]. 计算机系统应用, 2017, 23 (6): 193-197.
- [2] 杜京义,胡益民,刘宇程. 基于区域分块的 SIFT 图像匹配技术研究及实现 [J]. 光电工程, 2013, 40 (8): 52-58.
- [3] 刘小建,张元. 基于 Kinect 的手势识别及其在场景驱动中的应用 [D]. 太原: 中北大学, 2017.
- [4] Simen Andresen, Martin Stokkeland, Vegar qsthus. Hand Detection Using Color Recognition [J]. Object Tracking and Gesture Recognition - Shortened Version. 2013, 29 (6).
- [5] 方华,刘诗雄,田敬北. 基于 Kinect 骨骼系统的手势识别研究 [J]. 2015, 22 (5): 65-67.
- [6] 罗元,杨明珠,张毅. 基于改进去伪匹配 SURF 算法的静态手语字母识别 [J]. 激志, 2014, 35 (7): 1-4.
- [7] Cui Y T, Weng J J. Viewased hand segmentation and hand-sequence recognition with complex backgrounds [A]. In Proceedings of the IEEE International Conference on Pattern Recognition [C]. Osaka, Japan, 1997: 617-621.
- [8] 宋海生,刘平和,王全州,等. 基于人体骨骼和深度图像信息的指尖检测方法 [J]. 计算机工程与科学, 2014, 9.
- [9] 黄山,罗琳. 基于视觉的手势识别系统关键技术研究 [D]. 南京: 东南大学, 2015.
- [10] 张毅,张烁,罗元,等. 基于 Kinect 深度图像信息的手势轨迹识别及应用 [J]. 计算机应用研究, 2012, 9.
- [11] 李长龙,董秀成. 基于 Kinect 深度图像的手势识别研究 [D]. 成都: 西华大学, 2014.
- [12] 毛雁明,张立亮. 基于 Kinect 深度信息的手势分割与识别 [J]. 系统仿真学报, 2015, 4.
- [13] Zhang Jinyu, Yan Chen, Huang Xianxiang. Edge detection of images based on improved sobel operator and genetic algorithms [A]. Proceedings of 2009 International Conference on Image Analysis and Signal Processing [C]. 2009: 32-35.
- [14] Shotton, J. Fitzgibbon A. Cook M. et al. Real-time human pose recognition in parts from single depth images [A]. 2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) [C]. 2011, 1297-1304.
- [15] 周凯,马行. 基于肤色和 SVM 的手势识别及其应用研究 [D]. 银川: 北方民族大学, 2016.
- [16] 郑斌珏,赵疗英. 基于 Kinect 深度信息的手势识别 [D]. 杭州: 杭州电子科技大学, 2014.
- [17] 陈祖雪,马苗. 基于深度卷积神经网络的手势识别研究 [D]. 西安: 陕西师范大学, 2016.
- [18] 董立峰,阮军. 基于 Hu 矩和支持向量机的静态手势识别及应用 [D]. 武汉: 武汉理工大学, 2012.