

## 基于改进最大类间方差法的手势分割方法研究

李擎<sup>1</sup> 唐欢<sup>1</sup> 迟健男<sup>1</sup> 邢永跃<sup>1</sup> 李华通<sup>1</sup>

**摘要** 针对手势图像中由于噪声和成像干扰造成的手势模糊和边界不清晰的问题, 提出了一种基于改进最大类间方差法的手势分割方法. 首先建立手势图像的二维灰度直方图, 在二维灰度直方图上确定噪声点位置, 在原图的相应区域滤除噪声. 然后重建二维灰度直方图将内点区的点集投影到 45 度线, 得到投影灰度直方图. 接下来在灰度投影直方图上采用全局 Otsu 确定局部 Otsu 的左边界, 用高斯函数拟合得到局部 Otsu 右边界, 最后采用局部 Otsu 分割手势. 该方法可以有效地对手势图像进行精确分割, 实验结果验证了本文算法的有效性.

**关键词** 手势分割, 改进最大类间方差法, 二维灰度直方图, 投影灰度直方图

**引用格式** 李擎, 唐欢, 迟健男, 邢永跃, 李华通. 基于改进最大类间方差法的手势分割方法研究. 自动化学报, 2017, 43(4): 528–537

**DOI** 10.16383/j.aas.2017.c150862

## Gesture Segmentation with Improved Maximum Between-cluster Variance Algorithm

LI Qing<sup>1</sup> TANG Huan<sup>1</sup> CHI Jian-Nan<sup>1</sup> XING Yong-Yue<sup>1</sup> LI Hua-Tong<sup>1</sup>

**Abstract** In this paper, in order to solve the problem of ambiguity or unclear boundary caused by noise and interference in gesture imaging, a gesture segmentation method based on the improved maximum between-cluster variance algorithm is proposed. Firstly, a two-dimensional gray histogram of gesture image is generated, and positions of noise points are determined on the two-dimensional gray histogram. After filtering noise in the corresponding region of the gesture image, a two-dimensional gray histogram is reconstructed. The point set of the inner point area are projected to the 45 degrees line to generate the gray projection histogram. Then, the global Otsu is used to determine the left boundary of the local Otsu and Gauss function is used to get the right boundary of the local Otsu in the projection gray histogram. Finally, the local Otsu is used to segment the gesture image. This method can effectively segment the gesture image accurately. Experimental results have verified the effectiveness of the proposed algorithm.

**Key words** Gesture segment, improved Otsu, two-dimensional gray histogram, projection gray histogram

**Citation** Li Qing, Tang Huan, Chi Jian-Nan, Xing Yong-Yue, Li Hua-Tong. Gesture segmentation with improved maximum between-cluster variance algorithm. *Acta Automatica Sinica*, 2017, 43(4): 528–537

据统计, 人类之间 70% 的信息是通过肢体语言或面部表情传递的, 肢体语言比自然语言更能表达人的意图. 因此, 手势是人们与外界交流沟通时最自然和直观的交流方式之一. 近年来, 基于机器视觉的手势识别技术作为一种重要的自然人机交互手段而受到众多研究人员的关注. 基于机器视觉的手势识别可以分为静态手势识别和动态手势识别, 静态手势识别通过人手图像提取手势特征去识别手势所具有的含义; 动态手势识别则通过人手运动的视频图像, 在序列图像中提取人手的动态特征; 在此基础上

进行聚类分析, 确定手势动作的含义. 无论是动态还是静态手势识别, 手势识别前需要首先对手势进行分割, 提取手势识别所需要的手部特征和运动信息, 因此, 手势分割是手势识别的前端和基础, 手势分割的效果直接影响手势识别的可靠性和精度.

目前的手势分割方法大致可以分为如下几类: 1) 基于肤色特征的分割方法. 肤色是人手的天然特征, 如文献 [1] 采用肤色检测与背景差分的方法, 结合降噪算法实现了在背景复杂的视频序列中的手势分割; 文献 [2] 提出用手势的肤色特性和光照条件, 分三个阶段 (颜色空间的转化、降噪、将手和阴影的像素平均值作为自动阈值) 实现了实时的手势检测与分割; 文献 [3] 提出了一种对手的外观和光照条件具有鲁棒性的手势分割算法, 通过在彩色空间中手像素分布的方法中设计一个距离度量的学习, 提高了分割的效果; 文献 [4] 针对传统肤色分割的不足, 提出先用改进的卡尔曼滤波先来预判手的位

收稿日期 2015-12-23 录用日期 2016-04-01  
Manuscript received December 23, 2015; accepted April 1, 2016  
北京市自然科学基金 (4122050) 资助  
Supported by Natural Science Foundation of Beijing (4122050)  
本文责任编辑 贺威  
Recommended by Associate Editor HE Wei  
1. 北京科技大学自动化学院 北京 100083  
1. School of Automation and Electrical Engineering, University of Science and Technology Beijing, Beijing 100083

置, 其次用 TSL (Tinct-saturation light) 肤色模型, 最后用形态学处理的方法进行手势分割; 文献 [5] 基于手势的肤色特征, 将 YCbCr 颜色空间中的 CbCr 分量映射到平面生成肤色聚集区, 提出了一种自由形式的肤色模型应用于复杂背景的手势分割. 2) 基于运动分析的分割方法. 运动分析一般用于动态手势的分割, 如文献 [6] 通过融合肤色的信息和手势的运动特性, 同时运用形态学的处理, 实现了单目视觉的手势分割; 文献 [7] 提出用正态分布中的  $3\sigma$  原则结合连续视频的背景差分减弱背景变化的影响, 然后用最大类间方差法选择阈值来进行手势分割; 文献 [8] 提出了基于模型的实时人机交互的手势分割, 分三个阶段 (运动分析、模型跟踪、模型匹配) 提取手势模型, 实现了复杂背景下的手势分割; 文献 [9] 提出了在基于隐藏条件随机场的分区矩阵的方法进行手势运动轨迹的评估, 实现了 2~4 个人的随机移动手势的选取. 3) 基于深度图像传感器的手势分割. Kinect 等三维体感传感器的出现为手势分割提供了新的方法和途径, 如文献 [10] 提出将遗传算法用于评价深度图像和 RGB 图像的关键点, 对提取的特征点的数量具有鲁棒性, 并将其用于手势分割; 文献 [11] 提出将改进的基于贝叶斯网络的 EM (Expectation maximization) 算法用于提取彩色和深度数据, 结合遗传算法来进行手势分割. 4) 除了上述几类, 还包括一些其他的手势分割方法. 如文献 [12] 用运动轮廓做椭圆拟合的方法从历史图像中筛选出手势区域, 能够更好地适应不同的光照环境; 文献 [13] 提出使用高斯多尺度聚集策略及其在时间上的优越特性去在移动设备上进行无背景约束情况下的手势分割.

总之, 由于手势分割是手势识别的重要组成部分, 有着较长的研究历史, 也提出了很多分割方法, 其中基于肤色的分割方法应用于静态的手势分割较为广泛, 而基于运动分析的分割方法则主要应用于动态的手势分割. 本文的研究内容是多点触摸与手势识别人机交互系统中的一部分, 研究平台集成了光感应式多点触摸装置、平台式多点触摸装置、Kinect 传感器等设备, 构成了综合手势识别的研究平台, 对手势检测 (触点检测)、手势跟踪 (触点跟踪)、手势识别 (触摸手势识别) 等内容进行深入研究, 在各类传感器的手势分割检测中, 精确的手势分割都是不可避免的问题. 例如: 在运用 Kinect 的 RGB-Depth 图像传感器进行手势分割时, 手势目标的边界会由于噪声的影响和成像干扰而模糊、不清或缺 (如图 1 所示), 从而影响手势分割的精度. 而前文综述的各种手势检测方法, 多数手势分割方法一般都关注如何在复杂的背景下或者光照变化下的手势定位与检测或如何将运动手势分割出来, 或关

注三维手势分割等问题. 而对如何消除手势成像中的干扰, 以及手势目标在边界模糊等情况下的精确手势分割等问题则研究不多, 同时关注在全局图像中检测和定位手势, 而无法解决手势定位后的精确分割问题.

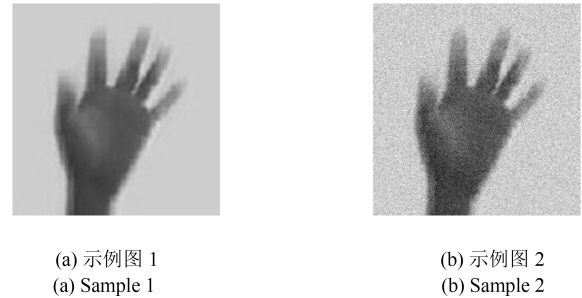


图 1 待精确分割的手势图像

Fig. 1 Gesture images to be segmented accurately

本文的研究工作在于提出了一种改进的最大类间方差法来解决手势定位后的手势精确分割问题. 该方法是针对 Kinect 的 RGB-Depth 图像在前景和背景分离过程中手势图像边界模糊而提出的, 但具有通用性, 也适用于其他类图像的相似问题的解决; 通常情况下, 阈值分割回归于本质还是调整阈值到最优, 而本文提出的手势分割方法优势在于: 1) 在目标手势初步背景分离后, 建立二维直方图定位噪声, 滤噪 (八邻域点的灰度均值代替噪声点, 越界则取八邻域的部分点), 减弱了噪声的影响; 2) 重建二维直方图后进行投影降维能够提高效率; 3) 全局 Otsu 结合高斯拟合的处理定位局部 Otsu 的阈值边界能够使阈值达到真正意义上的最优.

## 1 最大类间方差法<sup>[14]</sup>

对一幅灰度图像进行阈值分割时, 需要通过一个灰度级 (假定为  $k$ ) 去进行类别的区分. 假定这幅灰度图像的灰度级分为  $[0, 1, \dots, L-2, L-1]$ , 再令  $C_0$  代表前景,  $C_1$  代表背景, 则  $C_0$  表示灰度级为  $[0, 1, \dots, k]$  的像素点,  $C_1$  表示灰度级为  $[k+1, \dots, L-1]$  的像素点. 对于每一类出现的概率如式 (1) 所示:

$$\begin{cases} w_0 = P(C_0) = \sum_{i=0}^k p_i = w(k) \\ w_1 = P(C_1) = \sum_{i=k+1}^{L-1} p_i = 1 - w(k) \end{cases} \quad (1)$$

其中,  $p_i$  为灰度级  $i$  出现的概率, 而  $w_0 + w_1 = 1$ . 对于每一类的平均灰度如式 (2) 所示:

$$\left\{ \begin{array}{l} u_0 = \frac{\sum_{i=0}^k i \times P(i|C_0)}{\sum_{i=0}^k i \times p_i} = \frac{u(k)}{w_0} \\ u_1 = \frac{\sum_{i=k+1}^{L-1} i \times P(i|C_1)}{\sum_{i=k+1}^{L-1} i \times p_i} = \frac{u_T - u(k)}{1 - w(k)} \end{array} \right. \quad (2)$$

其中,  $u_T = \sum_{i=0}^{L-1} i \times p_i$ ,  $w_0 u_0 + w_1 u_1 = u_T$ . 这样, 类间方差、类内方差及总方差如式 (3) 所示:

$$\left\{ \begin{array}{l} \delta_B^2(k) = w_0(u_0 - u_T)^2 + w_1(u_1 - u_T)^2 \\ \delta_W^2(k) = w_0 \sum_{i=0}^k (i - u_0)^2 P(i|C_0) + \\ \quad w_1 \sum_{i=k+1}^{L-1} (i - u_1)^2 P(i|C_1) \\ \delta_T^2 = \delta_B^2(k) + \delta_W^2(k) = \text{const} \end{array} \right. \quad (3)$$

由于总方差为常量, 计算最大类间方差得到的阈值  $k$  即为最优, 然后简化类间方差如式 (4) 所示:

$$\delta_B^2(k) = \frac{[u_T w(k) - u(k)]^2}{w(k) \times (1 - w(k))} \quad (4)$$

选择适当的  $k$  使得  $\delta_B^2(k)$  最大即可, 这个  $k$  也就是要得到的阈值.

## 2 基于改进最大类间方差法的手势分割方法

### 2.1 方法概述

如图 2 所示, 首先获取待精确分割的手势图像, 其次建立手势图像的二维灰度直方图, 在二维灰度直方图上确定噪声点位置, 在原图的相应区域滤除噪声. 然后重建二维灰度直方图, 将内点区的点集投影到 45 度线, 进行降维处理, 得到投影灰度直方图. 接下来在投影灰度直方图上即用改进的 Otsu 处理: 先是对全局的投影灰度直方图进行 Otsu 处理得到一个阈值作为后续局部 Otsu 的左边界, 后用部分投影灰度直方图做镜像处理后拟合成高斯曲线, 用高斯曲线的  $3\sigma$  原则确定局部 Otsu 的左边界, 最后采用局部 Otsu 阈值分割手势. 该方法可以有效地对手势图像进行精确分割, 能够减弱噪声, 保证实时性和分割的准确性.

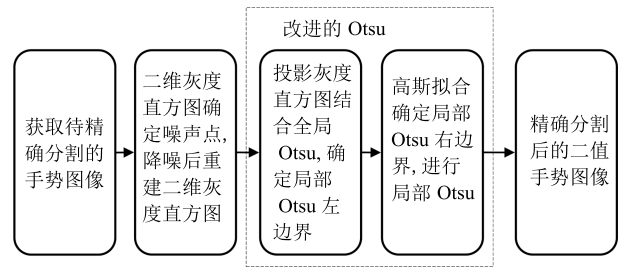


图 2 手势分割框架图

Fig. 2 Framework of gesture segmentation

### 2.2 基于改进最大类间方差法的手势分割方法

对于本文的手势图像, 用  $f(x, y)$  表示手势图像中点  $(x, y)$  的灰度值,  $g(x, y)$  为以点  $(x, y)$  为中心,  $k \times k$  邻域内的平均灰度值,  $g(x, y)$  定义为式 (5), 以  $f(x, y)$  和  $g(x, y)$  组成二元组  $(i, j)$  来定义二维灰度直方图<sup>[15-16]</sup>, 直方图中任意一点的联合概率密度定义为  $c_{ij}$ , 它表示二元组发生的频率, 则:  $p(i, j) = c_{ij}/(M \times N)$ , 其中  $c_{ij}$  是  $(i, j)$  出现的频数,  $0 \leq i, j \leq L-1$ ,  $M$  和  $N$  为手势图像的长和宽.

$$g(x, y) = \frac{1}{k^2} \sum_{m=-[k/2]}^{[k/2]} \sum_{n=-[k/2]}^{[k/2]} f(x+m, y+n) \quad (5)$$

对于本文手势图像二维灰度直方图 (取  $k=3$ ), 如图 3 所示, 首先通过位于主对角线两侧且与其平行的 4 条斜线  $L_1$ 、 $L_2$ 、 $L_3$ 、 $L_4$ , 将二维灰度直方图区域分成内点区、两个边界点区和两个噪声点区. 内点区:  $L_1$  和  $L_2$  之间的区域由于像素点灰度级和邻域平均灰度级相近, 认为是目标和背景; 边界点区:  $L_1$  和  $L_3$  之间、 $L_2$  和  $L_4$  之间的两个区域因像素点灰度级和邻域平均灰度级有一定差别, 视为目标和背景之间过渡; 噪声点区:  $L_3$  以外和  $L_4$  以外的两个区域由于像素点灰度级和邻域平均灰度级相差很大<sup>[17-19]</sup>. 这些像素点主要落在内点区, 为了提高分割效果的同时又不过量损失效率, 采用与横坐标夹 45 度的二维灰度直方图对角线投影内点区域的点集. 当内点区的点集在投影线上的时候, 将这些点集累加到对应灰度级之上, 这样就可以得到投影灰度直方图. 图 4 所示的投影原理图中可以看到点集中的点可以大致分为两类, 同时两类之间也存在较多的点.

通常情况下 (以黄种人为例) 手势作为一个特定的目标, 它的灰度级是很丰富的, 根据投影灰度直方图的分布特性可知, 手势目标和背景之间存在较长的过渡地带, 如果采用单一的最大类间方差法或者结合肤色分割去处理, 在目标边界较为模糊时候, 过渡带会出现类别误分的情况, 得到的分割结果

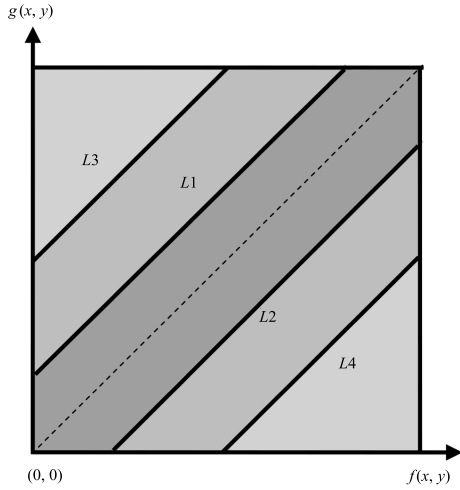


图3 二维灰度直方图区域划分图

Fig.3 Region division of two-dimensional gray histogram

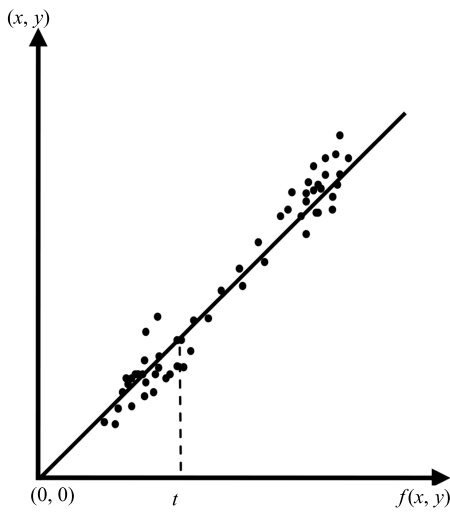


图4 投影原理图

Fig.4 Principle of projection

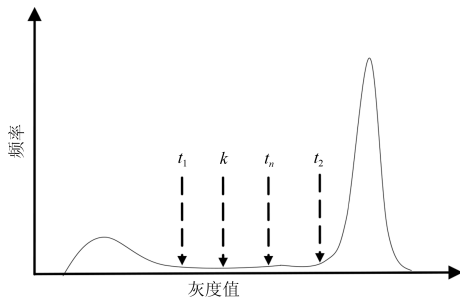


图5 投影灰度直方图

Fig.5 Projection gray histogram

将是很不理想的. 本文提出基于改进最大类间方差法来进行手势分割: 首先, 获取待精确分割的手势图像取二维灰度直方图, 并进行区域(内点区、边界点区、噪声点区)的划分, 滤噪; 其次, 重建二维灰度

直方图并将内点区的散点进行投影处理, 得到如图5所示. 鉴于本文的手势图像的边界较为模糊(存在背景覆盖目标的现象), 在全局进行最大类间方差法处理后, 在过渡带  $[t_1, t_2]$  区间内的  $[k, t_2 - 1]$  进行投影直方图局部的最大类间方差法处理, 此时, 将引出一个新的问题, 即寻找过渡带的右边界的方法. 本文先采用最小二乘法处理局部的投影灰度直方图得到高斯曲线将高斯函数描述(式(6)所示)转化为式(7)的形式.

$$f(x) = \frac{A}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (6)$$

$$F(x) = a + bx + cx^2 \quad (7)$$

式(7)中,  $F(x) = \ln f(x)$ ,  $a = \ln A - \frac{1}{2} \ln(2\pi) - \ln \sigma - \mu^2/(2\sigma^2)$ ,  $b = \mu/\sigma^2$ ,  $c = -1/(2\sigma^2)$ , 建立基于最小二乘拟合原理<sup>[20]</sup>的目标函数(式(8)所示)

$$G = \sum_{i=1}^n [F(x_i) - (a + bx_i + cx_i^2)]^2 \quad (8)$$

令  $\frac{\partial G}{\partial a} = 0$ ,  $\frac{\partial G}{\partial b} = 0$ ,  $\frac{\partial G}{\partial c} = 0$ , 整理各项并把  $a, b, c$  分离出来得式(9), 根据式(9)可求方程解  $[a, b, c]$  得式(10), 得到式(10)三个参数后即可得到高斯曲线, 然后采用  $3\sigma$  原则, 将拟合的高斯曲线的  $\mu - 3\sigma$  处对应的灰度级作为  $t_2$ .

$$\begin{bmatrix} n & \sum_{i=1}^n x_i & \sum_{i=1}^n x_i^2 \\ \sum_{i=1}^n x_i & \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i^3 \\ \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i^3 & \sum_{i=1}^n x_i^4 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n F(x_i) \\ \sum_{i=1}^n x_i F(x_i) \\ \sum_{i=1}^n x_i^2 F(x_i) \end{bmatrix} \quad (9)$$

$$\begin{cases} A = \sqrt{-\frac{\pi}{b}} e^{(a-\frac{b^2}{4c})} \\ \mu = -\frac{b}{2c} \\ \sigma = \sqrt{-\frac{1}{2c}} \end{cases} \quad (10)$$

找到右边界  $t_2$  后, 令  $D_0$  代表过渡带的前景,  $D_1$  代表过渡带的背景, 则  $D_0$  表示灰度级为  $[k, k + 1, \dots, t_n]$  的像素点,  $D_1$  表示灰度级为

$[t_n + 1, t_n + 2, \dots, t_2 - 1]$  的像素点. 对于每一类出现的概率如式 (11) 所示.

$$\begin{cases} \psi_0 = P(D_0) = \sum_{i=k}^{t_n} \rho_i = \psi(t_n) \\ \psi_1 = P(D_1) = \sum_{i=t_n+1}^{t_2-1} \rho_i = 1 - \psi(t_n) \\ \rho_i = \frac{p_i}{w(t_2) - w(k)} \\ \psi_0 + \psi_1 = 1 \end{cases} \quad (11)$$

其中  $\rho_i$  为在  $[k, k + 1, \dots, t_2 - 1]$  中灰度级  $i$  出现的概率,  $D_0$  和  $D_1$  的平均灰度如式 (12) 所示. 这样, 可以得到并化简在灰度级区间  $[k, k + 1, \dots, t_2 - 1]$  进行局部最大类间方差法处理的类间方差, 如式 (13) 所示. 接下来需要计算全局和局部最大类间方差法的方差和, 如式 (14) 所示, 然后选择适当的  $t_n$  ( $t_n \in [k, t_2 - 1]$ ) 使得  $\delta(k, t_n)$  最大, 这个  $t_n$  即为所求最优阈值.

$$\begin{cases} v_0 = \frac{\sum_{i=k}^{t_n} i \times P(i|D_0)}{\psi_0} = \frac{\sum_{i=k}^{t_n} i \times \rho_i}{\psi(t_n)} \\ v_1 = \frac{\sum_{i=t_n+1}^{t_2-1} i \times P(i|D_1)}{\psi_1} = \frac{\sum_{i=t_n+1}^{t_2-1} i \times \rho_i}{1 - \psi(t_n)} \\ v_T = \sum_{i=k}^{t_2-1} i \times \rho_i = \psi_0 v_0 + \psi_1 v_1 \\ \rho_i = \frac{p_i}{w(t_2) - w(k)} \end{cases} \quad (12)$$

$$[\delta(t_n)]^2 = \frac{[v_T \psi(t_n) - v(t_n)]^2}{\psi(t_n) \times (1 - \psi(t_n))} \quad (13)$$

$$\delta(k, t_n) = [\delta(t_n)]^2 + [\delta(k)]^2 \quad (14)$$

接下来需要计算全局和局部最大类间方差法的方差和, 如式 (14) 所示, 然后选择适当的  $t_n$  ( $t_n \in [k, t_2 - 1]$ ) 使得  $\delta(k, t_n)$  最大, 这个  $t_n$  即为所求最优阈值.

### 3 实验及分析

#### 3.1 样本的选择及实验设计

本文选用的手势分割的实验样本主要分为四个部分, 如表 1 所示. 实验 1 选用的是不加噪声且对比度高, 目标清晰的手势图像. 实验 2 选用的是不加噪声但手势目标模糊, 边界不清晰的图像. 实验 3 选用的是含有泊松噪声且手势目标模糊, 边界不清晰的图像. 实验 4 选用的是不同个体的手势, 且图像中手势目标模糊, 边界不清晰. 为了验证本文算法的有效性, 本文选取了两种有代表性的手势分割方法进行比较 (一种仅仅是阈值的调整, 一种是加上特定的处理之后选最优阈值进行分割), 分别运用普通的最大类间方差法进行手势图像的阈值分割, 常用的肤色检测结合最大类间方差法进行阈值分割, 进行不同的分割方法的比较; 同时针对最后一组样本不同的个体不同手势进行本文算法处理, 评估手势目标分割结果.

#### 3.2 实验过程及结果

根据第 3.1 节中选择的手势样本和设计的对比实验, 则本文的实验过程如下:

1) 获取待精确分割的手势图像, 部分加高斯噪声满足实验 3, 不加噪声满足实验 2 和实验 4, 而实验 1 图像来自网络图像.

2) 用第 1) 步中的图像做二维灰度直方图确定噪声点区、边界点区和内点区, 对噪声点和边界点进行滤噪处理, 得到滤除噪声的手势图像用于后续的手势分割处理.

3) 将滤噪后的图像的二维灰度直方图的内点区向 45 度线投影得到投影灰度直方图, 进行全局的最大类间方差法处理得到初始阈值  $k$  作为局部最大类间方差法的左边界.

表 1 实验样本表

Table 1 Experimental samples

样本名称	样本特点	样本选择目的
实验 1	不加噪声, 对比度较高, 边界清晰	验证本文算法对成像质量高的手势分割效果
实验 2	不加噪声, 对比度一般, 边界模糊, 存在背景覆盖目标的情况	验证本文算法对边界模糊的手势分割效果
实验 3	存在噪声, 对比度一般, 边界模糊, 存在背景覆盖目标的情况	验证本文算法对存在噪声且边界模糊的手势分割效果
实验 4	不同个体的不同手势图像, 边界模糊, 存在背景覆盖目标的情况	验证本文算法对不同个体手势分割效果

4) 高斯函数拟合局部投影灰度直方图 (灰度级在  $k$  和投影灰度直方图右峰值之间, 由于  $t_2$  在右峰值点的左侧, 则以右峰值的左侧附近点对应的灰度级做镜像后拟合), 采用  $3\sigma$  原则确定局部最大类间方差法的右边界  $t_2$ , 在  $[k, t_2 - 1]$  内局部最大类间方差法处理得到最终阈值 (等于  $\mu - 3\sigma + k$ ), 而后进

行手势分割得到最终结果.

5) 将第 1) 步得到的图像分别做最大类间方差法和肤色模型结合最大类间方差法的处理用于比较.

6) 用前述 2) ~ 5) 步分别做完前 3 个实验, 对实验 4 用本文算法处理.

实验结果如图 6 ~ 9 所示.

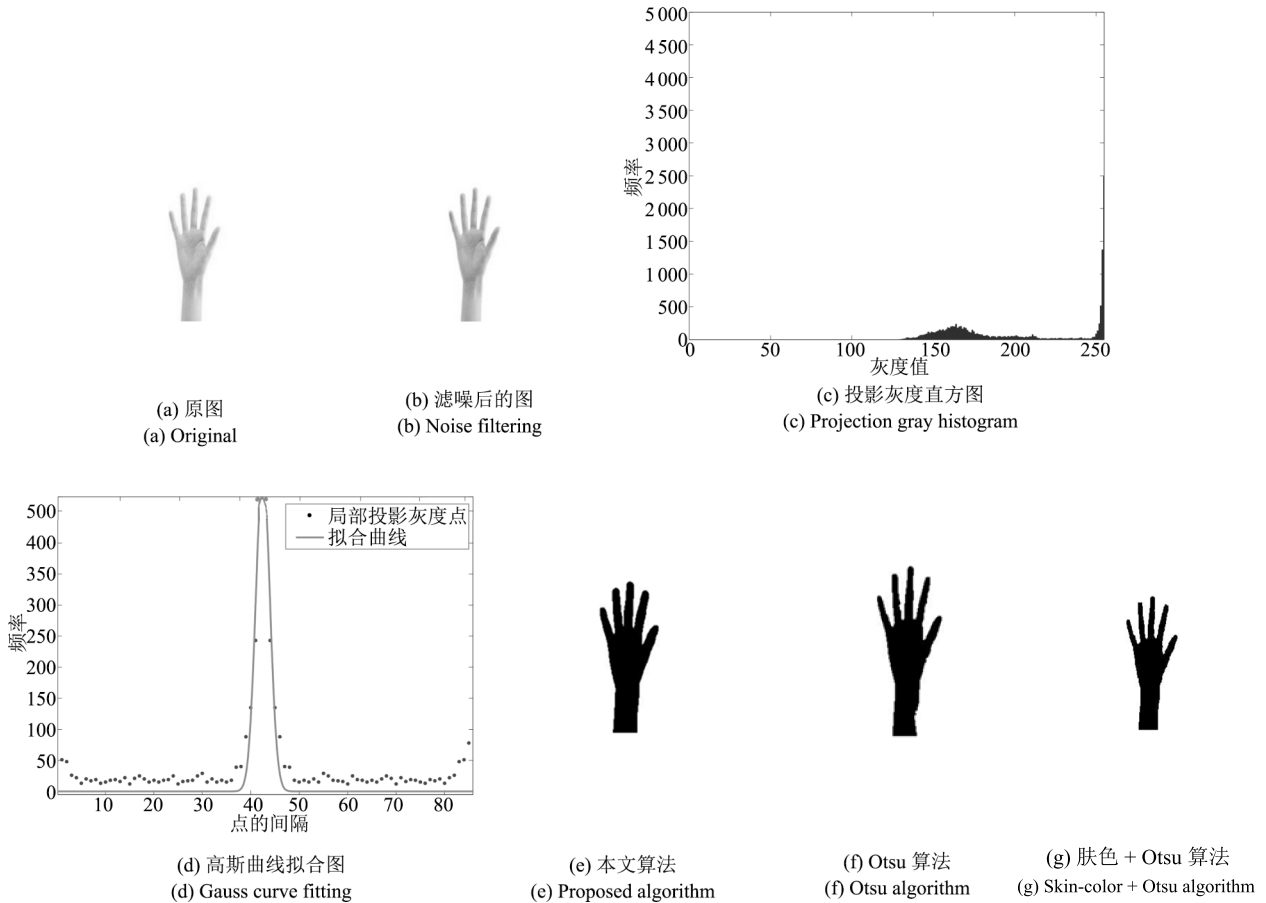
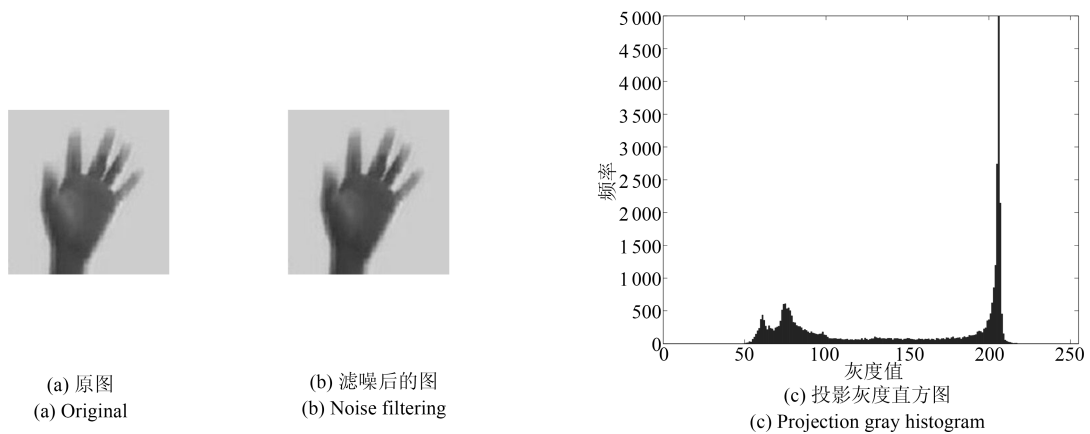


图 6 实验 1 处理过程及结果图

Fig. 6 Processing and results of Experiment 1



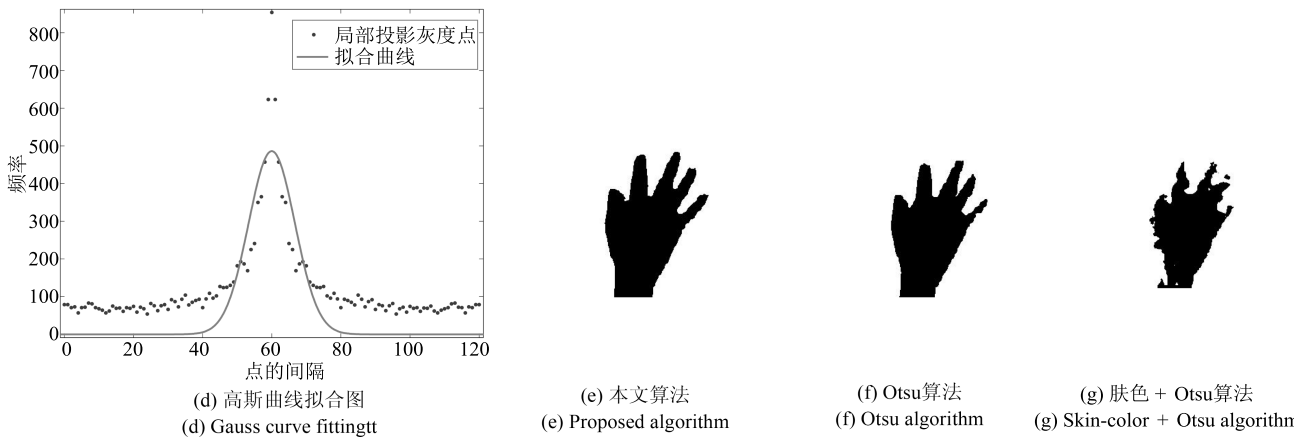


图 7 实验 2 处理过程及结果图  
Fig. 7 Processing and results of Experiment 2

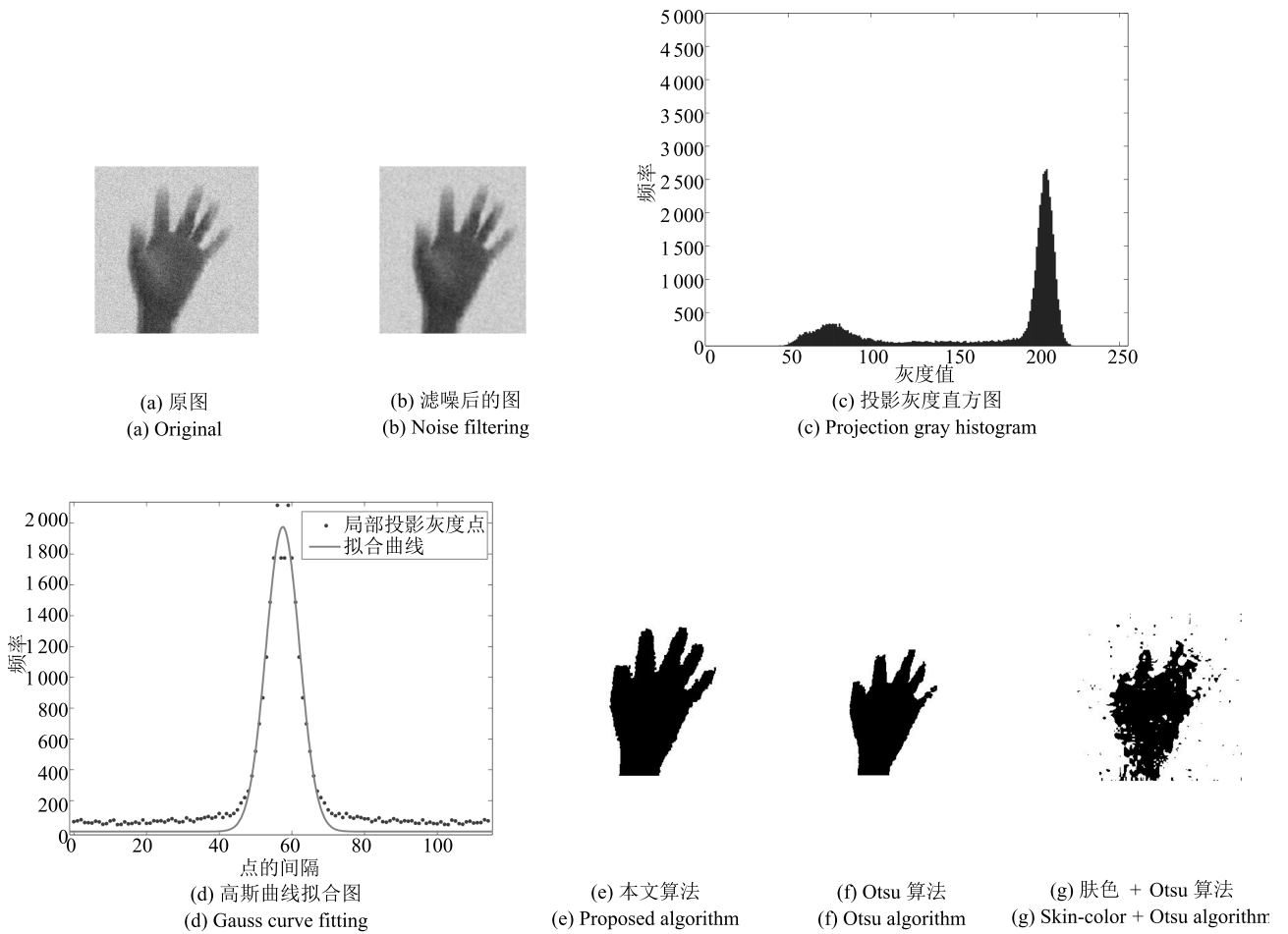


图 8 实验 3 处理过程及结果图  
Fig. 8 Processing and results of Experiment 3



图 9 实验 4 处理结果图 (本文算法)

Fig. 9 Processing and results of Experiment 4 (Proposed algorithm)

表 2 拟合曲线参数表 (本文算法)  
Table 2 Parameters of fitting curve  
(Proposed algorithm)

实验	$\mu$	$\sigma$
实验 1	42	1.4487
实验 2	60	6.7082
实验 3	57	4.6043

表 3 边界灰度表 (本文算法)  
Table 3 Edge gray (Proposed algorithm)

实验	$k$	$t_2$
实验 1	210	252
实验 2	142	203
实验 3	145	202

表 4 阈值表 (三种算法)  
Table 4 Threshold (Three algorithms)

实验	Otsu	肤色 + Otsu	本文算法
实验 1	210	143	248
实验 2	142	136	183
实验 3	145	136	188

### 3.3 实验效果评价

实验 1 的处理结果如图 6(e), (f), (g) 所示, 可以看到最大类间方差法分割后手势轮廓的右下部分有缺陷, 肤色结合最大类间方差法能够较好地呈现轮廓, 本文算法在滤噪时对该图像作用不明显, 投影灰度直方图呈现双峰且过渡带距离较长, 高斯曲线拟合较好 (寻找本文算法的右边界), 分割后能够较好地呈现轮廓; 与此同时, 三种方法都能够较好地得到手势的指尖信息. 说明本文算法在应用于不加噪声且边界清晰的手势图像能够取得较好的效果, 性能优于常用的最大类间方差法和肤色结合最大类间方差法.

实验 2 的处理结果如图 7(e), (f), (g) 所示, 可以看到最大类间方差法分割后手势轮廓的小拇指明显过细, 影响了整体的轮廓, 指尖能够反映该手势的特征信息: 肤色结合最大类间方差法分割后轮廓和指尖的特征信息完全体现不出来, 误分严重: 本文算法在滤噪时对该图像作用不明显, 投影灰度直方图基本呈现双峰且过渡带距离较长, 高斯曲线拟合较



好, 分割整体效果较好, 只是大拇指存在些许误分, 但不影响特征信息. 说明本文算法在应用于不加噪声、边界模糊的手势图像能够取得较好的效果, 性能远远优于常用的最大类间方差法和肤色结合最大类间方差法.

实验 3 的处理结果如图 8(e), (f), (g) 所示, 可以看到最大类间方差法分割后手势轮廓的小拇指完全被误分, 中指的指尖分割结果也较差, 对于噪声的抑制作用也一般; 肤色结合最大类间方差法分割后手势的轮廓和指尖的特征信息完全体现不出来, 同时抑噪能力较差; 而本文算法在滤噪时对该图像作用明显, 投影灰度直方图呈现双峰且过渡带距离较长; 高斯曲线拟合很好, 分割后手势轮廓较好, 食指和无名指虽然分割不是很精确但不影响该手势特征信息的体现, 抑噪能力较好. 说明本文算法在应用于加噪声、边界模糊的手势图像能够取得较好的效果, 性能远远优于常用的最大类间方差法和肤色结合最大类间方差法.

实验 4 的处理结果部分如图 9 所示, 结果评价如表 5 所示, 可以看到不同个体的不同手势图像在应用本文算法进行手势精确分割的时候都能够取得较好的效果, 指尖和轮廓分割能够很好地反映出手势的特征信息. 说明本文算法的适用性较好.

表 5 不同方法处理的结果评价表

Table 5 Results evaluation of different algorithms

实验	评价指标	Otsu	肤色 + Otsu	本文算法
实验 1	指尖	较好	较好	较好
实验 1	轮廓	一般	一般	较好
实验 1	时间 (ms)	2.75	3.21	31.25
实验 2	指尖	较好	较差	较好
实验 2	轮廓	一般	较差	较好
实验 2	时间 (ms)	2.85	3.82	37.63
实验 3	指尖	较差	较差	一般
实验 3	轮廓	较差	较差	较好
实验 3	抑噪	一般	较差	较好
实验 3	时间 (ms)	2.99	3.67	39.50

同时, 就三类算法而言, 本文算法在实验 1~3 中耗时相对较长, 在实验 4 中的耗时大概在 33 ms 到 40 ms 的区间内, 但是本文算法分割效果明显优于另外两类, 耗时就手势分割或者后续的手势识别的实时性来说, 是可以接受的.

### 3.4 实验结论

在多点触摸与手势识别人机交互系统中, 应用 Kinect 初步分离背景的手势图像采用本文算法有如下特点: 1) 本文算法对手势的阈值分割中的噪声具有抑制作用, 对存在边界模糊的手势目标更具针对性, 能够较好地分离背景得到手势目标; 2) 本文算法在阈值分割方面明显优于普通的最大类

间方差法, 肤色结合最大类间方差法的分割; 3) 上文实验 4 中对多个手势样本进行阈值分割, 仍能取得不错的效果, 同时耗时控制较好, 说明本文算法在多点触摸与手势识别人机交互系统中的适用性较好.

## 4 结语

本文针对手势图像中由于噪声和成像干扰造成的手势模糊和边界不清晰的问题, 提出了一种基于改进最大类间方差法的手势分割方法. 首先通过二维灰度直方图确定噪声点后滤除噪声. 其次重建二维灰度直方图做投影处理得到投影灰度直方图. 然后在灰度投影直方图上采用全局 Otsu 确定局部 Otsu 的左边界, 用高斯函数拟合得到局部 Otsu 右边界, 最后采用局部 Otsu 分割手势. 本文算法能够很好地精确分割手势图像, 能够尽可能地保留手势图像中的特征信息, 为后续的手势识别提供了良好的保证. 本文的算法还有很大的研究空间, 例如使用不同的滤噪算法进行滤噪, 用更多的方法 (例如背景建模处理) 初步处理后结合本文的算法提高分割精度, 进一步提高效率等, 作者将继续进行相关的研究.

## References

- Bao H, Zhao X G. Study on hand gesture segmentation. In: Proceedings of the 2010 International Conference on Multimedia Technology (ICMT). Ningbo, China: IEEE, 2010. 1-4
- Rahmat R W, Al-Tairi Z H, Saripan M I, Sulaiman P S. Removing shadow for hand segmentation based on background subtraction. In: Proceedings of the 2012 International Conference on Advanced Computer Science Applications and Technologies (ACSAT). Kuala Lumpur, Malaysia: IEEE, 2012. 481-485
- Hou G D, Chang D Q, Zhang C S. Hand segmentation with metric learning superpixels. In: Proceedings of the 2014 IEEE China Summit & International Conference on Signal and Information Processing (ChinaSIP). Xi'an, China: IEEE, 2014. 455-459
- Mo S, Cheng S H, Xing X F. Hand gesture segmentation based on improved Kalman filter and TSL skin color model. In: Proceedings of the 2011 International Conference on Multimedia Technology (ICMT). Hangzhou, China: IEEE, 2011. 3543-3546
- Dawod A Y, Abdullah J, Alam M J. A new method for hand segmentation using free-form skin color model. In: Proceedings of the 3rd International Conference on Advanced Computer Theory and Engineering (ICACTE). Chengdu, China: IEEE, 2010. V2-562-V2-566
- Cao Xin-Yan, Zhao Ji-Yin, Li Min. Monocular vision gesture segmentation based on skin color and motion detection. *Journal of Hu'nan University (Natural Sciences)*, 2011, **38**(1): 78-83  
(曹昕燕, 赵继印, 李敏. 基于肤色和运动检测技术的单目视觉手势分割. *湖南大学学报 (自然科学版)*, 2011, **38**(1): 78-83)
- Zhang Q Y, Chen F, Liu X W. Hand gesture detection and segmentation based on difference background image with complex background. In: Proceedings of the 2008 International Conference on Embedded Software and Systems (ICCESS2008). Sichuan, China: IEEE, 2008. 338-343

- 8 Tsai T H, Lin C Y. Visual hand gesture segmentation using signer model for real-time human-computer interaction application. In: Proceedings of the 2007 IEEE Workshop on Signal Processing Systems. Shanghai, China: IEEE, 2007. 567–572
- 9 Yao Y, Li C T. Hand gesture segmentation in uncontrolled environments with partition matrix and a spotting scheme based on hidden conditional random fields. In: Proceedings of the 2nd IAPR Asian Conference on Pattern Recognition (ACPR). Naha, Japan: IEEE, 2013. 842–846
- 10 Ju Z J, Wang Y H, Zeng W, Chen S Y, Liu H H. Depth and RGB image alignment for hand gesture segmentation using Kinect. In: Proceedings of the 2013 International Conference on Machine Learning and Cybernetics (ICMLC). Tianjin, China: IEEE, 2013. 913–919
- 11 Ju Z J, Wang Y H, Zeng W, Cai H B, Liu H H. A modified EM algorithm for hand gesture segmentation in RGB-D data. In: Proceedings of the 2014 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE). Beijing, China: IEEE, 2014. 1736–1742
- 12 de Santos Sierra A, Ávila C S, del Pozo G B, Casanova J G. Gaussian multiscale aggregation oriented to hand biometric segmentation in mobile devices. In: Proceedings of the 3rd World Congress on Nature and Biologically Inspired Computing (NaBIC). Salamanca, Spain: IEEE, 2011. 237–242
- 13 Shi Jiu-Gen, Chen Zhi-Hui. Hand gesture segmentation based on MHI and ellipse fitting. *Computer Engineering and Applications*, 2014, **50**(22): 199–202  
(史久根, 陈志辉. 基于运动历史图像和椭圆拟合的手势分割. *计算机工程与应用*, 2014, **50**(22): 199–202)
- 14 Otsu N. A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 1979, **9**(1): 62–66
- 15 Yang Hui-Xian, Li Miao, Tan Zheng-Hua, Zhai Yun-Long, Zhang Jian-Bo. Maximum scatter difference image thresholding segmentation algorithm based on two-dimensional histogram oblique. *Laser and Infrared*, 2014, **44**(4): 463–468  
(杨恢先, 李淼, 谭正华, 翟云龙, 张建波. 二维直方图斜分最大散度差阈值分割算法. *激光与红外*, 2014, **44**(4): 463–468)
- 16 Wu Yi-Quan, Pan Zhe, Wu Wen-Yi. Image thresholding based on two-dimensional histogram oblique segmentation and its fast recurring algorithm. *Journal on Communications*, 2008, **29**(4): 77–83  
(吴一全, 潘喆, 吴文怡. 二维直方图区域斜分阈值分割及快速递推算法. *通信学报*, 2008, **29**(4): 77–83)
- 17 Chen Q, Zhao L, Lu J, Kuang G, Wang N, Jiang Y. Modified two-dimensional Otsu image segmentation algorithm and fast realisation. *IET Image Processing*, 2012, **6**(4): 426–433
- 18 Xu Guo-Hua, Zhang Bao-Ming, Li Xu. Change detection for remote sensing images based on improved Otsu algorithm. *Science of Surveying and Mapping*, 2012, **37**(1): 80–82  
(徐国华, 张保明, 李旭. 基于改进的最大类间方差法的遥感影像变化检测. *测绘科学*, 2012, **37**(1): 80–82)
- 19 Jing Xiao-Jun, Cai An-Ni, Sun Jing-Ao. Image segmentation based on 2D maximum between-cluster variance. *Journal of China Institute of Communications*, 2001, **22**(4): 71–76  
(景晓军, 蔡安妮, 孙景鳌. 一种基于二维最大类间方差的图像分割算法. *通信学报*, 2001, **22**(4): 71–76)
- 20 Gao Shi-Yi, Yang Kai-Zhen. Research on central position extraction of laser strip based on varied-boundary Gaussian fitting. *Chinese Journal of Scientific Instrument*, 2011, **32**(5): 1132–1137  
(高世一, 杨凯珍. 变边界高斯拟合提取激光条纹中心线方法的研究. *仪器仪表学报*, 2011, **32**(5): 1132–1137)



**李 擎** 北京科技大学自动化学院教授。主要研究方向为智能优化理论及其在路径规划、基于数据驱动建模中的应用。

E-mail: liqing@ies.ustb.edu.cn

(**LI Qing** Professor at the School of Automation and Electrical Engineering, University of Science and Technology Beijing. His research interest covers

intelligent optimization theory and its application in path planning and data driven modeling.)



**唐 欢** 北京科技大学自动化学院硕士研究生。主要研究方向为图像处理与分析, 手势识别, 视线追踪。

E-mail: tanghuanyl@126.com

(**TANG Huan** Master student at the School of Automation and Electrical Engineering, University of Science and Technology Beijing. His re-

search interest covers image processing and analysis, gesture recognition and gaze tracking.)



**迟健男** 北京科技大学自动化学院副教授。主要研究方向为视线追踪, 人机交互, 车辆辅助驾驶, 嵌入式系统。本文通信作者。E-mail: sy\_jnchi@126.com

(**CHI Jian-Nan** Associate professor at the School of Automation and Electrical Engineering, University of Science and Technology Beijing. His re-

search interest covers gaze tracking, human computer interaction, driver assistance system and embedded system. Corresponding author of this paper.)



**邢永跃** 北京科技大学自动化学院硕士研究生。主要研究方向为机器学习, 图像处理与分析, 视线追踪, 计算机视觉。

E-mail: xingyongyue@outlook.com

(**XING Yong-Yue** Master student at the School of Automation and Electrical Engineering, University of Science and Technology Beijing. His re-

search interest covers machine learning, image processing and analysis, gaze tracking and computer vision.)



**李华通** 北京科技大学自动化学院硕士研究生。主要研究方向为图像处理与分析, 视觉伺服, 计算机视觉。

E-mail: lhuatong@126.com

(**LI Hua-Tong** Master student at the School of Automation and Electrical Engineering, University of Science and Technology Beijing. His re-

search interest covers image processing and analysis, vision serving and computer vision.)