

# 面向颜色深度图像手脸近距遮挡的手势识别

刘斌 赵兴 胡春海 万欣

燕山大学测试计量技术与仪器重点实验室, 河北 秦皇岛 066004

**摘要** 手脸近距遮挡属于深度传感器应用中具有代表性的难点问题,针对该问题提出了一种综合利用颜色与深度信息的手势识别方法。采用核模糊 C-均值聚类,对手脸遮挡图像进行粗分割和灰度增强,实现手脸分离。引入初始水平集函数,解决聚类方法导致的手势区域像元缺失问题。利用基于深度信息的梯度方向直方图(HOG)特征对手势进行分类识别。通过采集不同人体手脸近距遮挡情形下的多种手势图像建立了样本数据库,进行了对比实验,实验结果验证了该方法的可行性和有效性。本文方法能有效分离近距遮挡的手和脸,提取得到相对完整的手势信息,深度 HOG 特征能够对手势空间信息进行精确描述,具有比传统形状特征更准确的识别效果。

**关键词** 图像处理; 手势识别; 手脸近距遮挡; 多相水平集; 核模糊 C-均值聚类; 梯度方向直方图特征

中图分类号 TP391.4

文献标识码 A

doi: 10.3788/LOP53.061001

## Gesture Recognition Method of Hand over Face Occlusion in Color and Depth Images

Liu Bin Zhao Xing Hu Chunhai Wan Xin

Key Laboratory of Measurement Technology and Instrumentation, Yanshan University, Qinhuangdao, Hebei 066004, China

**Abstract** Hand over face occlusion is a typical difficulty during depth sensor application. Aiming at this point, a gesture recognition algorithm is proposed using color and depth information comprehensively. Kernel fuzzy C-means algorithm is used to get rough segments of hand over face occlusion image and gray enhancement. The hand face separation is achieved. The initial level set function is introduced to solve the clustering method of gesture area pixels missing problem. The classification of gestures is carried based on the histogram of oriented gradient (HOG) features of depth map. The sample database is established by collecting different human hand and face close covered case gesture images. Comparative result proves the feasibility and effectiveness of the proposed algorithm. The proposed method can separate hand from face efficiently and get relatively ideal result, and the HOG features can describe the spatial information more precisely and get higher recognition rate compared with traditional characteristics.

**Key words** image processing; gesture recognition; hand over face occlusion; multiphase level set; kernel fuzzy C-means clustering; histogram of oriented gradient features

**OCIS codes** 100.3008;100.5010;150.1135

## 1 引言

近年来,随着 RGB-D 数据传感器及其相关技术的快速发展,深度信息的应用受到广泛关注。在手语理解、人机交互、体感游戏等涉及手势识别的应用中,深度信息不受物体颜色、纹理和背景光线等因素影响,对解决遮挡问题有明显优势,但手脸交互过程中存在颜色相近且距离贴近的情形,此时手脸分离、手形提取与手势识别变得异常困难。

近距离遮挡情况下完整而准确地分割手形对识别手势具有重要意义。传统手势分割方法多针对彩色监

收稿日期: 2016-01-28; 收到修改稿日期: 2016-02-23; 网络出版日期: 2016-05-26

基金项目: 河北省自然科学基金(F2015203287)、河北省科技计划项目(15220324)

作者简介: 刘斌(1981—),男,博士,讲师,主要从事视觉检测、目标识别等方面的研究。E-mail: dqlb@ysu.edu.cn

控视频中的人体影像,利用颜色特征和肤色提取技术进行手势分割<sup>[1-3]</sup>。文献[4]提出融合边缘信息与亮度变化进行手脸遮挡情形下的手势提取,但该方法对环境要求苛刻,受光照、手形等因素影响严重,使得结果不完整且存在虚假边缘。陈晓雷等<sup>[5]</sup>提出了一种基于力场转换的手势检测算法,该方法分别计算纯脸区域与遮挡脸部区域的力场图像,然后通过分块直方图的方法检测手脸遮挡区域,这种方法只是对遮挡大致区域进行预测,不能用于手势提取。

Pham等<sup>[6]</sup>通过微软 Kinect 深度传感器捕获人体运动信息,利用经验阈值方法提取手势。徐鹏飞等<sup>[7]</sup>利用最大类间方差计算阈值,进而检测得到手势。上述方法对于手脸间距较远时有效,但是在手语识别、人机交互和异常行为识别等诸多应用中,存在大量手脸贴近且手形灵活多变的情形,此时,需要综合利用深度数据和颜色信息才能有效检测手势。

近距离手脸分离与手势完整分割是手势识别的关键环节,同时,深度信息对于手势特征的描述具有重要作用,深度信息能够精确描述手势的空间信息,能够弥补传统的形状特征或区域特征中缺少对手势内部信息描述问题。本文提出一种融合深度信息的模糊聚类协同偏移场矫正模型多相水平集分割方法。采用基于核模糊 C-均值聚类(KFCM)算法对手脸遮挡图像进行聚类,并对聚类结果进行灰度增强,将其引入初始化水平集函数,并利用基于多相偏移场矫正的水平集方法对图像进行精细分割,得到完整的手势图像,利用基于深度图像的梯度方向直方图(HOG)特征进行手势描述与识别。

## 2 KFCM 协同多相水平集算法的手势分割

为了详细说明手脸交互行为中存在肤色相近且手脸贴近的问题,本文在手脸交互行为的视频序列中,分别提取手运动到脸部区域和离开脸部区域的 5 帧颜色和深度图像,颜色图像和深度图像分别如图 1(a)、(b)所示。其肤色区域深度直方图如图 1(c)所示,图 1(d)为采用文献[7]的方法进行图像分割的结果。从图中可以看出,当手没有到达脸部区域时,直方图呈现双峰特性,适合用阈值或文献[7]的方法进行手势分割;当手脸距离较近时,其直方图不再呈现双峰特性,很难利用传统方法进行分割。

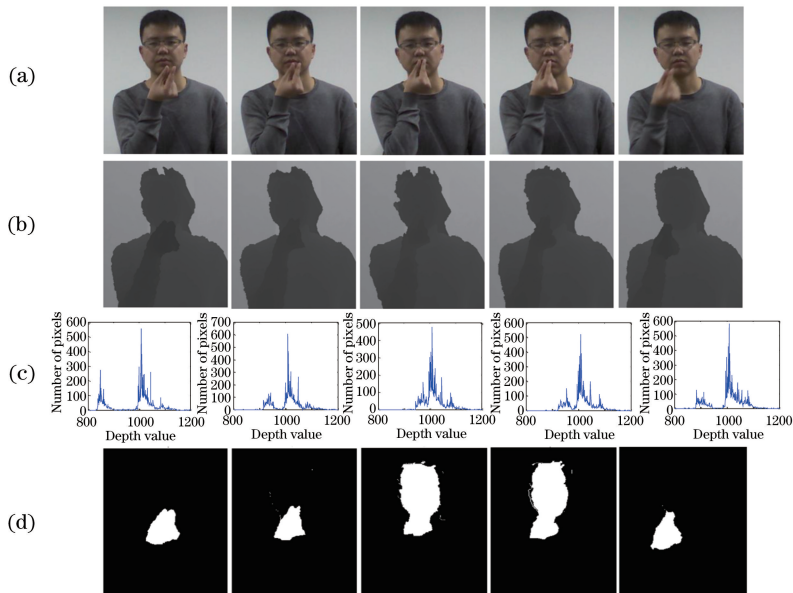


图 1 手脸交互行为对比图

Fig. 1 Comparison charts of hand face interactive behavior

### 2.1 基于 KFCM 的手脸分离

运用核模糊聚类算法<sup>[8-9]</sup>,引入高斯核函数,通过计算中心像元与高斯核窗口中其他像元的核函数关联度,使原始样本空间数据映射到高维特征空间,以实现更准确的模糊划分。

核模糊聚类算法目标函数定义为

$$J_m(\mathbf{U}, \mathbf{V}) = \sum_{i=1}^c \sum_{k=1}^n u_{ik}^m ||\Phi(x_k) - \Phi(v_i)||^2, \quad (1)$$

式中  $X = \{x_1, x_2, \dots, x_n\}$  为图像中  $n$  个像元的像元值所组成的数据集,  $x_k$  为数据集中第  $k$  个样本;  $V = \{v_1, v_2, \dots, v_c\}$  表示聚类中心,  $c$  为聚类的个数;  $\mathbf{U}$  为隶属度矩阵,  $u_{ik}^m$  是  $\mathbf{U}$  中的元素, 取值范围为  $[0, 1]$ , 表示第  $k$  个元素对聚类中心  $v_i$  的隶属程度, 且满足  $\sum_{i=1}^c u_{ik}^m = 1, \forall k = 1, 2, \dots, n$ ;  $m$  为加权指数一般取值为 2;  $||\Phi(x_k) - \Phi(v_i)||$  为距离函数,  $\Phi$  表示非线性映射, 以高斯核函数  $K(x, v) = \exp(-||x - v||^2 / 2\sigma^2)$  表示非线性映射,  $\sigma$  表示高斯核函数宽度。

此时距离函数表示为

$$||\Phi(x_k) - \Phi(v_i)|| = \sqrt{[\Phi(x_k) - \Phi(v_i)]^T [\Phi(x_k) - \Phi(v_i)]} = \sqrt{K(x_k, x_k) + K(v_i, v_i) - 2K(x_k, v_i)}, \quad (2)$$

由于  $K(x_k, x_k) = K(v_i, v_i) = 1$ , 则 KFCM 目标函数表达式为

$$J_m(\mathbf{U}, \mathbf{V}) = 2 \sum_{i=1}^c \sum_{k=1}^n u_{ik}^m [1 - K(x_k, v_i)]. \quad (3)$$

根据隶属度归一化约束, 通过迭代更新聚类中心和隶属度函数, 来达到最小化目标函数  $J$  的目的。

图 2(a)和(b)分别为采集的彩色图像和深度图像。深度图像为 16 位, 采用位图分割算法<sup>[10]</sup>提取低 8 位手势图像, 并进行亮度归一化增大图像对比度, 然后通过颜色特征去除非肤色区域的背景干扰, 结果如图 2(c)所示。图 2(d)~(j)为使用 KFCM 算法将手脸遮挡图像分为 7 类的结果。

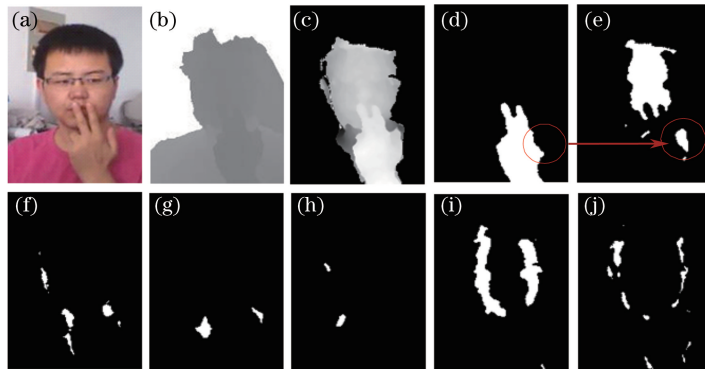


图 2 KFCM 聚类分割手脸遮挡图像。(a) 彩色图; (b) 深度图; (c) 增强图; (d)~(j) 聚类结果图

Fig. 2 KFCM cluster results of hand face occlusion image. (a) Color image; (b) depth image; (c) enhanced result; (d)~(j) cluster results

在图 2(d)~(j)的聚类结果中, 图 2(d)为部分手势区域, 图 2(e)中手指与面部区域能够分离, 图 2(f)~(j)为面部区域或颈部区域, 说明 KFCM 能很好地将手与面部分离, 但是由于手型灵活多变, 灰度不均匀, 聚类时不能将完整的手势分割出来, 例如图 2(d)~(e)中红色圈出的拇指区域。而完整手势轮廓对于量化手势特征, 定义手势类别, 进而有效区分各种手势具有重要作用。

由于手部区域灰度与面部区域灰度值相近, 为了提取完整的手势, 把面积最大的区域与深度均值低或区域边界平滑的邻接区域进行合并, 并对合并结果进行灰度增强, 进一步对增强后图像中的手势进行精细提取。其中, 灰度增强可表示为

$$I_{\text{new}}(i, j) = I_{\text{mean}} + \{1 + [I(i, j) - I_{\text{mean}}]\}^\gamma, \quad (4)$$

式中  $I_{\text{new}}(i, j)$  为区域中灰度变换后的灰度值,  $I(i, j)$  为图像原来的灰度值,  $I_{\text{mean}}$  为非面部、非手势区域灰度的平均值,  $\gamma = -10$ , 为了计算简便, 直接用  $I_{\text{mean}}$  代替  $I_{\text{new}}(i, j)$ 。

## 2.2 基于偏移场矫正多项水平集的手势提取

偏移场可以用来表示灰度不均匀图像<sup>[11]</sup>, 基于偏移场的灰度不均匀图像可以表示为

$$\mathbf{I} = \mathbf{bJ} + \mathbf{n}, \quad (5)$$

式中  $I$  为采集到的图像,  $n$  为附加高斯噪声,  $J$  为真实图像, 一般假设符合分段常数的特性,  $b$  为偏移场, 一般假设具有灰度缓慢变化和平滑的特点。

灰度矫正的目地就是通过观测图像  $I$ , 估算真实原始图像  $J$  和估算偏移场  $b$ 。对于只有目标和背景的理想图像  $\Omega$ , 可以将其划分为 2 个区域  $\Omega_1$  和  $\Omega_2$ ,  $\Omega_1$  和  $\Omega_2$  可以通过一个水平集函数  $\varphi(x)$  的符号表示为

$$\Omega_1 = \{x : |\varphi(x)| > 0\}, \Omega_2 = \{x : |\varphi(x)| < 0\}。 \quad (6)$$

为了更直观表示  $\Omega_1$  和  $\Omega_2$ , 定义区域函数  $M_i(\varphi)$  为  $M_1(\varphi) = u[\varphi(x)]$ , 其中  $u(x)$  为阶跃函数。对于区域  $\Omega_1$  内的元素有  $M_1(\varphi) = u[\varphi(x)]$ , 对于区域  $\Omega_2$  内的元素有  $M_2(\varphi) = 1 - u[\varphi(x)]$ 。

图片中的目标多于两个时, 如本文手脸遮挡情况下, 包含手、脸和背景至少三个区域, 定义两个或多个水平集函数  $\varphi_1$  和  $\varphi_2$ 。设图像中包含三个待分割目标时区域函数可以表示为

$$\begin{cases} M_1(\varphi) = u(\varphi_1)u(\varphi_2) \\ M_2(\varphi) = u(\varphi_1)[1 - u(\varphi_2)], \\ M_3(\varphi) = 1 - u(\varphi_1) \end{cases} \quad (7)$$

根据偏移场理论的假设, 将真实图像中每个区域  $\Omega_i$  的灰度值近似表示为一个常数  $c_i$ , 那么真实图像表示为

$$J(x) = \sum_{i=1}^N c_i M_i(\varphi), \quad (8)$$

在适当邻域  $O_y = \{x : ||x - y| \leq \rho\}$  内, 有  $b(y) \approx b(x)$ , 这个邻域用高斯窗口函数表示, 则观测图像  $I$  表示为

$$I(x) = b(y)c_i + n(x), x \in O_y \cap \Omega_i, \quad (9)$$

利用真实图像与观测图像之间的拟合误差构造能量泛函为

$$\epsilon = \int \left[ \sum_{i=1}^N \int K(y-x) |I(x) - b(y)c_i|^2 M_i(\varphi) dx \right] dy, \quad (10)$$

为了提高水平集函数演化的稳定性, 在上述能量函数的基础上, 加入计算演化曲线长度的规则项, 根据梯度下降流理论求解并迭代更新偏移场、平均灰度值和水平集  $\varphi(x)$ 。

本文图像中包括手部区域  $\Omega_1$ , 面部区域  $\Omega_2$ , 背景区域  $\Omega_3$  和未分类区域  $\Omega_4$ , 将未分类区域融入面部区域, 以便进行初始化和进一步精细分割。此时将水平集函数初始化为

$$\begin{cases} \varphi_1(x, y) = \begin{cases} 1, (x, y) \in \Omega_1 \cup \Omega_2 \\ -1, (x, y) \notin \Omega_1 \cup \Omega_2 \end{cases} \\ \varphi_2(x, y) = \begin{cases} 1, (x, y) \in \Omega_2 \cup \Omega_3 \\ -1, (x, y) \notin \Omega_2 \cup \Omega_3 \end{cases} \end{cases}, \quad (11)$$

并将水平集函数代入(7)式完成区域函数的初始化。

## 3 基于深度信息的手势特征提取与识别

### 3.1 基于深度信息的手势特征表示

手势分割、特征提取和分类器设计是手势识别中三个关键步骤, 采用基于深度信息的 HOG 特征, 详细描述不同手势之间的差异。

HOG 特征通过计算局部区域上的梯度方向直方图来对目标的形状信息进行描述, 提取的特征向量对局部区域的描述具有极强的辨识能力<sup>[12-13]</sup>。传统的 HOG 特征一般对 RGB 图像进行处理, 但由于手势本身的颜色和纹理极为相近, 受环境影响严重, 单纯使用颜色信息的 HOG 特征不能很好地对手势局部区域进行描述。深度图像能精确地检测到手势内部的细微形变, 利用深度信息的 HOG 特征对手势进行描述具有更大的优势。

传统的 HOG 特征对尺度和方向敏感, 因此本文利用主轴方向角进行旋转变换, 并利用双线性插值方法对图像进行缩放来降低特征受图像尺度和方向影响较大的特点。具体步骤如下所示。

1) 对分割图像进行预处理: 首先对于分割的二值图像去除手臂干扰, 每  $3^\circ$  顺时针逆旋转一次, 并通过比较每次旋转之后图像每列像素总数的最大值来获取图像的主轴角度, 并根据主轴角度对图像进行旋转, 然后利

用双线性插值理论将图像缩放为  $64 \text{ pixel} \times 64 \text{ pixel}$ 。

2) 梯度计算:对于深度图像中的像元点  $(x, y)$  的梯度计算方法为

$$\begin{cases} G_x(x, y) = D(x+1, y) - D(x-1, y) \\ G_y(x, y) = D(x, y+1) - D(x, y-1) \end{cases}, \quad (12)$$

式中  $G_x(x, y)$  和  $G_y(x, y)$  分别为像元的水平和垂直方向的梯度值。 $D(x, y)$  为图像的深度值。像元点的梯度大小  $G(x, y)$  和方向  $\alpha(x, y)$  为

$$\begin{cases} G(x, y) = \sqrt{G_x(x, y)^2 + G_y(x, y)^2} \\ \alpha(x, y) = 1 / \left\{ \arctan \left[ \frac{G_y(x, y)}{G_x(x, y)} \right] \right\} \end{cases} \quad (13)$$

3) 构建梯度方向直方图:定义每个 block 大小为  $m \times m$ , 将缩放后的图像每隔  $m$  个像元进行划分得到  $(64/m) \times (64/m)$  个 block, 每个 block 包含 4 个不重叠的胞元, 大小为  $(m/2) \times (m/2)$  像元。将梯度方向量化为 9 个通道, 对胞元内每个像元的梯度方向对应的通道进行累加, 即得到梯度方向直方图。

4) 归一化:为了增加特征的稳健性,对提取的直方图进行归一化,归一化方法为

$$\mathbf{V} = \mathbf{v} / \sqrt{\|\mathbf{v}\|_2^2 + \epsilon^2}, \quad (14)$$

式中  $\|\mathbf{v}\|_k$  为向量  $\mathbf{v}$  的  $k$  范数,  $\epsilon$  为一个很小的值。深度 HOG 算法的流程如图 3 所示。图 3(a) 为原始深度图像, 图 3(b)、(c) 为经过旋转和缩放后求取的梯度幅值和方向图像, 再对梯度方向进行量化和归一化, 组合成最终的特征向量。

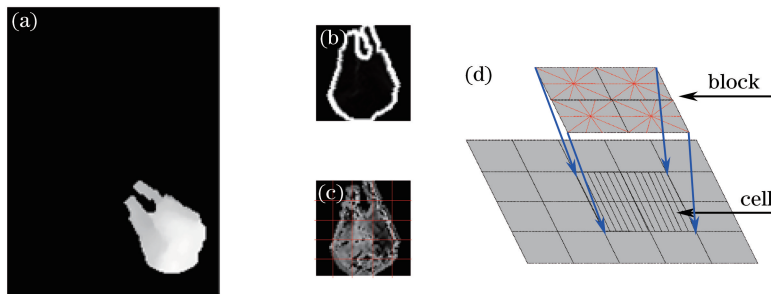


图 3 深度 HOG 计算示意图。(a) 深度手势图; (b) 幅值图; (c) 方向图; (d) 特征示意图

Fig. 3 Sketch map of HOG calculation. (a) Hand image; (b) amplitude image; (c) direction image; (d) sketch of feature

### 3.2 基于支持向量机的手势识别

支持向量机(SVM)的机理是寻找一个满足分类要求的最优分类超平面,使得该超平面在保证分类精度的同时,能够使超平面到两侧支持向量平面的距离最大。理论上,支持向量机能够对线性数据和非线性数据进行分类<sup>[14-15]</sup>。

线性可分数据超平面方程可以表示为

$$(\omega \cdot x) + b = 0, \quad (15)$$

为使分类超平面能够对样本正确分类并且具备最大的分类间隔,约束如下:

$$y_i [(\omega \cdot x_i) + b] - 1 \geq 0, i = 1, 2, \dots, m, \quad (16)$$

时分类间隔为  $2 / \|\omega\|$ , 最优超平面需要使  $2 / \|\omega\|$  最大化, 因此求解最优超平面的问题就转化为在约束式下求最小值问题:

$$\min \varphi(\omega) = 1 / (2 \|\omega\|^2). \quad (17)$$

上述问题可以利用凸二次规划完成求解。对于非线性可分数据,采用核函数的方法将低维特征转换为高维特征空间,在高维空间中利用线性可分理论对目标进行分类。本文手势识别框架如图 4 所示。

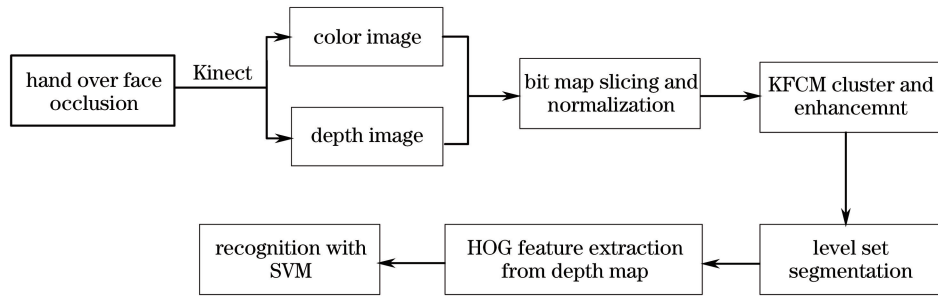


图 4 手势识别框架

Fig. 4 Framework of gesture recognition

## 4 实验结果与分析

实验利用 Matlab 2014a 软件和微软 Kinect for windows 传感器搭建数据采集平台,并采集实验样本。数据采集平台如图 5 所示,采集系统主要包括 kinect 传感器、支架和计算机。计算机的硬件配置如下:处理器为 Intel(R)coreTM(i5-3230 M),CPU 频率为 2.6 GHz,内存 4 GB。

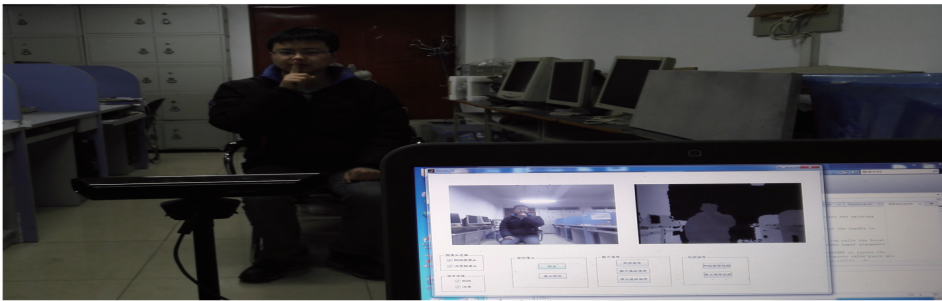


图 5 数据采集系统示意图

Fig. 5 Sketch map of data acquisition platform

为了验证本文方法的有效性与可行性,在室内环境下,采集不同性别、不同体型 10 个人的 6 种常见手脸交互手势样本,图 6 为手势提取实验结果,分别为 6 种不同手势的处理过程,图 6(a)为原始彩色图像、经位图分割算法和肤色提取算法提取的低 8 位面部深度图像,图 6(b)~(f)分别为进行归一化的预处理结果、面部区域灰度直方图、采用 KFCM 算法的分割结果、采用本文算法提取的轮廓结果、本文算法最终检测得到的手势图像。

从图 6 中可以看出,手脸交互图像的灰度直方图并不呈现明显的双峰特性,不宜采用最大类间方差算法进行阈值分割,而且由于手脸相距较近,手势灵活性较大,经验阈值也无法选取。采用 KFCM 算法能够很好地将手势从面部区域分离出来,再通过灰度变换和水平集精细分割过程,能够获得相对满意的结果。

为了定量验证分割效果,采集手脸距离较远的样本 840 个,手脸贴近的样本 840 个,其中,每种手势 140 幅图像。对手脸距离较远的深度图像采用文献[7]中的方法进行手势提取,手脸贴近的图像分别采用本文方法和文献[7]方法提取手势,然后分别将每种手势中的 40 幅,共 240 幅远距离手势和 240 张手脸贴近手势样本作为训练集,提取特征,并进行训练。对其余 600 幅手脸贴近手势和 600 幅远距离手势作为测试集,用于识别测试。其中部分实验手势样本深度图像如图 7 所示。

利用提取的手势库分别对手脸距离较远手势测试集、文献[7]手势分割方法测试集和本文分割方法测试集进行特征提取与识别,其中深度 HOG 特征的 block 为 16 pixel×16 pixel 像元,识别结果如图 8 所示。图 8(a)为远距离手势采用文献[7]方法进行识别的结果图,图 8(b)为手脸贴近图像采用文献[7]方法识别的结果,图 8(c)为对手脸贴近图像采用本文方法识别的结果。从图中可以看出,文献[7]方法对手脸距离较远图像识别效果较好,但是对于手脸贴近图像的手势识别结果不理想,说明文献[7]方法对手脸贴近图像手势提取结果不理想。采用本文方法对手脸贴近手势识别与手脸距离较远图像识别结果相近,平均准确率达到 97%,验证了近距离手脸贴近图像手势分割算法的有效性和可行性。由于深度传感器成像误差,所以有少数

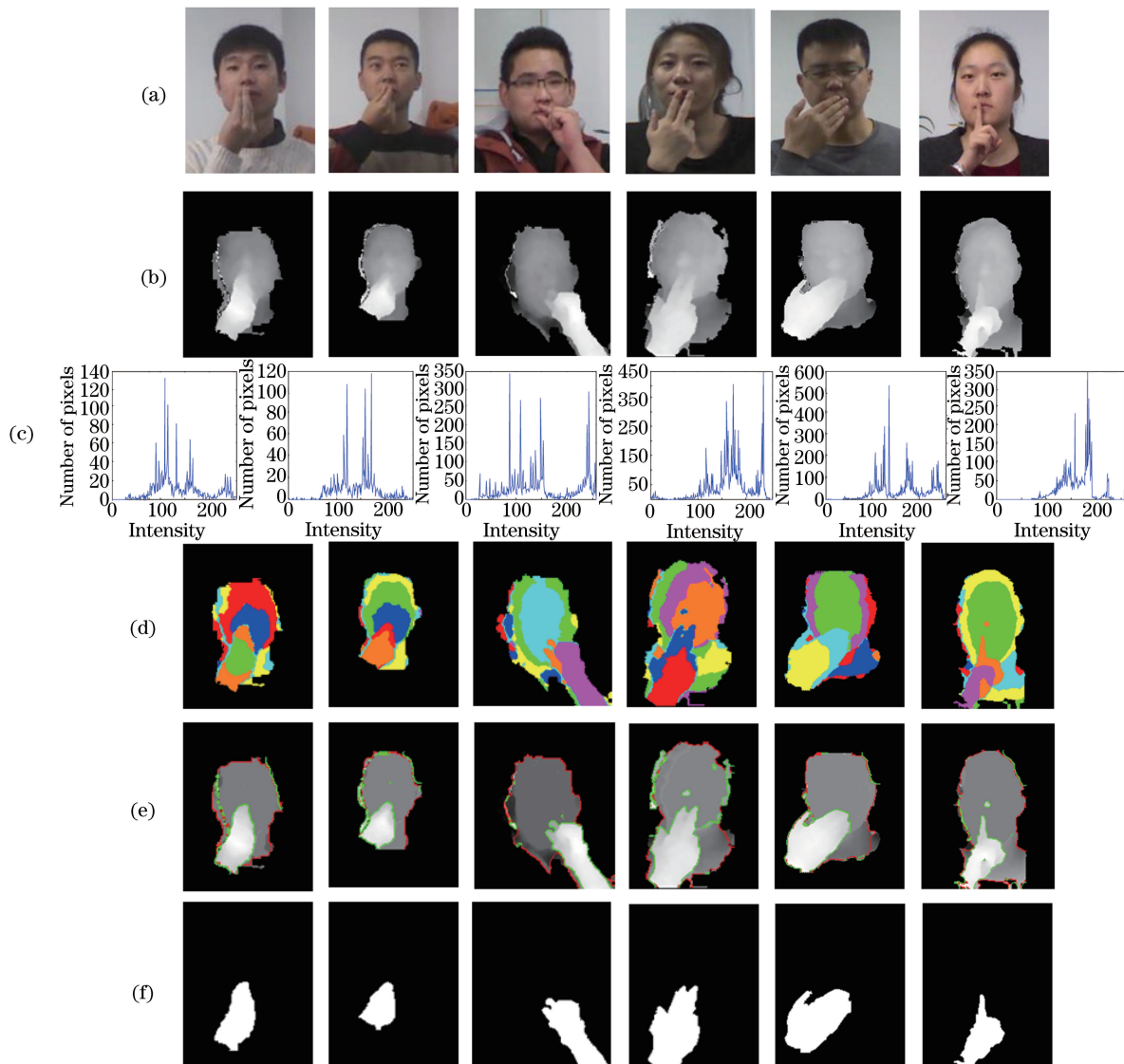


图 6 不同手势分割效果图

Fig. 6 Segmentation results of different gestures

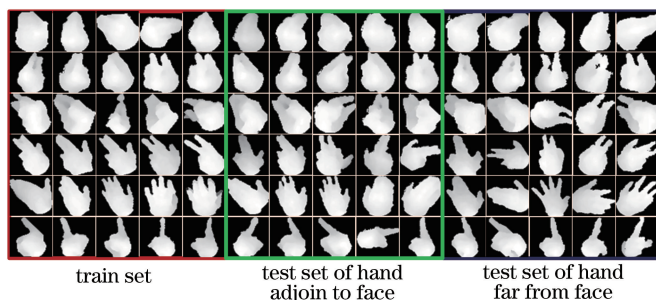


图 7 部分样本集手势深度图像

Fig. 7 Partial depth gesture images of sample set

手势样本未被正确识别。

深度图像的 HOG 特征能够对物体的空间信息进行有效描述,但是图像选取的 block 大小能够对目标描述产生的特征向量维数和运算时间产生较大影响。本文对不同 block 像元大小的识别结果进行对比。对手脸近距离遮挡手势测试集,分别提取 block 大小为  $8 \text{ pixel} \times 8 \text{ pixel}$ 、 $16 \text{ pixel} \times 16 \text{ pixel}$  和  $32 \text{ pixel} \times 32 \text{ pixel}$ ,总特征维数分别为 2304、576 和 144,对应的识别结果如图 9 所示。从识别结果中可以看

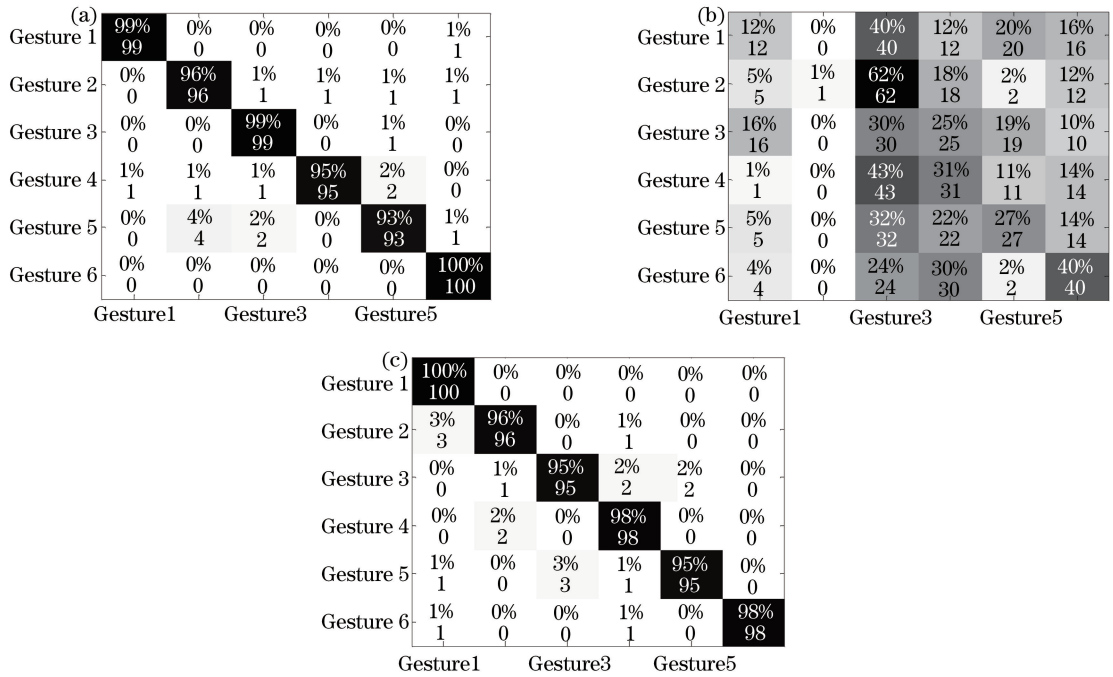


图 8 不同分割方法手势识别效果图。(a) 远距离识别结果；(b) 文献[7]方法识别结果；(c) 本文方法识别结果

Fig. 8 Recognition results of different segmentation methods. (a) Result of hand away from face;

(b) result of Ref. [7]; (b) result of proposed method

出, block 采用  $8 \text{ pixel} \times 8 \text{ pixel}$  和  $16 \text{ pixel} \times 16 \text{ pixel}$  的识别率相差不大, 而  $32 \text{ pixel} \times 32 \text{ pixel}$  窗口大小的识别效果较差, 这是因为大窗口引入大量的冗余信息, 将手势中的有用信息淹没。进一步说明  $16 \text{ pixel} \times 16 \text{ pixel}$  窗口大小既保持了对手势特征进行精确描述的能力, 又降低了特征维数。

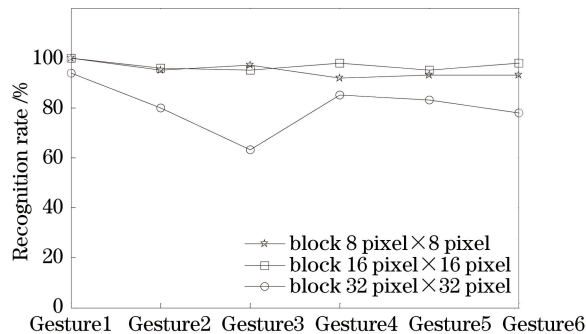


图 9 不同 block 大小识别结果对比图

Fig. 9 Recognition results of different block sizes

为了验证不同特征对识别结果的影响, 本文选取经验阈值 25, 分别对训练集和测试集的深度手势图像进行二值化, 然后对二值化结果提取 Hu 矩特征和边界傅里叶描述子特征, 并将识别结果与深度 HOG 特征识别结果进行比较, 如表 1 所示。结果表明, 深度 HOG 明显优于传统形状特征, 且远距离手势与近距离手势识别结果相比没有明显差别。这是因为传统的形状特征更加关注手势的外形特征, 而深度 HOG 特征在提取局部轮廓边缘特征的同时还对手势内部的空间信息进行描述, 增强了对手势空间信息的描述能力。

计算效率是评价手势识别算法的重要指标, 为了验证深度 HOG 特征的计算效率, 本文对测试集中的各种样本提取窗口大小为  $16 \text{ pixel} \times 16 \text{ pixel}$  的深度 HOG 特征, 然后对提取的特征进行识别, 并计算运行时间。结果表明, 识别一个手势的平均时间为 48.1 ms, 说明提取的特征可以快速地对手势进行描述, 具有较高的效率。同时, 继续对分割算法和特征描述算法进行优化, 以实现实时手势识别, 并将成为下一步主要研究工作。



表 1 不同特征识别结果对比

Table 1 Contrast of different feature recognition results

	Hu moment		Fourier descriptor		Depth HOG	
	Distant occlusion	Close occlusion	Distant occlusion	Close occlusion	Distant occlusion	Close occlusion
Gesture 1	98%	56%	88%	83%	99%	100%
Gesture 2	74%	64%	68%	60%	96%	96%
Gesture 3	48%	71%	68%	39%	99%	95%
Gesture 4	61%	84%	88%	89%	95%	98%
Gesture 5	87%	75%	74%	46%	93%	95%
Gesture 6	100%	88%	97%	87%	100%	98%
Average	78.0%	73.0%	80.5%	67.3%	96.0%	97.0%

## 5 结 论

针对深度传感器应用中具有代表性的前景与背景颜色相似、距离相近的手脸近距遮挡问题,提出了一种融合颜色和深度信息的手势识别方法。综合实验分析结果,采用的 KFCM 算法能够有效地将手形从面部区域分离,并通过灰度变换和水平集提取完整手形,为后续手势特征描述提供有效支撑。在此基础上,采用深度 HOG 特征能够对手势空间信息进行细致描述,具有比传统形状特征更为准确的识别效果。所提方法对于解决手语理解、人机交互、体感游戏等应用中手脸近距遮挡导致无法识别手势的问题提供了有效途径,后续工作将研究如何将该方法运用于深度传感器实时监控视频序列。

## 参 考 文 献

- Jalilian B, Chalechale A. Face and hand shape segmentation using statistical skin detection for sign language recognition[J]. Computer Science and Information Technology, 2013, 1(3): 196-201.
- Wang W, Pan J. Hand segmentation using skin color and background information [C]. International Conference on Machine Learning and Cybernetics (ICMLC), 2012, 4: 1487-1492.
- Yaming W, Jiansheng C, Guangda S U. Efficient hand segmentation and fingertip detection using color features of skin and fingernail[J]. Ieice Transactions on Information and Systems, 2013, 96(8): 1894-1897.
- Gonzalez M, Collet C, Rémi D. Head tracking and hand segmentation during hand over face occlusion in sign language [C]. Proceedings of the 11<sup>th</sup> European Conference on Trends and Topics in Computer Vision, 2012:234-243.
- Chen Xiaolei, Zhang Aihua, Lin Dongmei, *et al.*. Hand detection algorithm during hand over face occlusion in sign language video[J]. Journal of Sichuan University: Engineering Science Deition, 2014, 46(2): 66-71.  
陈晓雷, 张爱华, 林冬梅, 等. 适应手脸遮挡手语视频的手势检测算法[J]. 四川大学学报: 工程科学版, 2014, 46(2): 66-71.
- Pham N Q, Le H S, Nguyen D D, *et al.*. A study of feature combination in gesture recognition with kinect[M]. Berlin: Springer, 2015, 326: 459-471.
- Xu Pengfei, Zhang Hongying. Hand gesture segmentation and fingertip detection based on depth image of the Kinect[J]. Journal of Southwest University of Science & Technology, 2014, 29(1): 49-54.  
徐鹏飞, 张红英. 基于 Kinect 深度图像信息的手势分割和指尖检测算法[J]. 西南科技大学学报, 2014, 29(1): 49-54.
- Yang Xiukun, Zhong Mingliang, Jing Xiaojun, *et al.*. Near-infrared microscopic image segmentation based on W2DPCA-FCM[J]. Acta Optica Sinica, 2013, 33(8): 0811002.  
杨秀坤, 钟明亮, 景晓军, 等. 基于 W2DPCA-FCM 的近红外显微图像分割[J]. 光学学报, 2013, 33(8): 0811002.
- Jeph A, Rhee F C H. Interval type-2 fuzzy C-means using multiple kernels[C]. IEEE International Conference on Fuzzy Systems, 2013, 446: 1-8.
- Karthika M, James A. A proposed method for document image binarization based on bit plane slicing[C]. International Conference on Advances in Engineering and Technology Research, 2014: 1-5.
- Li C M, Huang R, Ding Z H, *et al.*. A Level set method for image segmentation in the presence of intensity inhomogeneities with application to MRI[J]. IEEE Transactions on Image Process, 2011, 20(7): 2007-2016.

- 12 Liu Shuping, Liu Yu, Yu Jun, *et al.*. Hierarchical static hand gesture recognition by combining finger detection and HOG features[J]. Journal of Image and Graphics, 2015, 20(6):781-788.  
刘淑萍, 刘 羽, 於 俊, 等. 结合手指检测和 HOG 特征的分层静态手势识别[J]. 中国图象图形学报, 2015, 20(6): 781-788.
- 13 Guo Pengyu, Su Ang, Zhang Hongliang, *et al.*. Online mixture of random naive bayes tracker combined texture with shape feature[J]. Acta Optica Sinica, 2015, 35(3): 0315002.  
郭鹏宇, 苏 昂, 张红良, 等. 结合纹理和形状特征的在线混合随机朴素贝叶斯视觉跟踪器[J]. 光学学报, 2015, 35(3): 0315002.
- 14 Wu Yiquan, Zhou Yang, Long Yunlin. Small target detection in Hyperspectral remote sensing image based on adaptive parameter SVM[J]. Acta Optica Sinica, 2015, 35(9): 0928001.  
吴一全, 周 杨, 龙云淋. 基于自适应参数支持向量机的高光谱遥感图像小目标检测[J]. 光学学报, 2015, 35(9): 0928001.
- 15 Sun Tao, Wang Canjin, Wang Rui, *et al.*. Contour bag of features applied in laser active lighting recognition system[J]. Chinese J Lasers, 2015, 42(1): 0109002.  
孙 涛, 王灿进, 王 锐, 等. 轮廓特征包在激光主动照明识别系统中的应用[J]. 中国激光, 2015, 42(1): 0109002.