

基于立体视觉的非穿戴指势识别

管业鹏

(上海大学 通信与信息工程学院, 上海 200072; 新型显示技术及应用集成教育部重点实验室, 上海 200072)

摘要: 基于彩色图像中红、绿、蓝3分量强度在阴影区域存在差异, 根据小波变换在时域和空域均具有优异的局部化特征, 结合背景差分, 进行小波多尺度变换, 提取视频指势对象, 所提方法不需场景学习与训练、手工校正及先验假设等信息, 可克服动态场景变化、阴影、噪声干扰等影响, 具有强的鲁棒性. 基于人类生物结构特征, 采用不易遮挡和不受人脸朝向、姿态、光照变化等影响的头顶特征代替人眼特征, 保证了人机交互活动的自由性和自然性, 且提高了人机交互的时效性. 融合手指尖特征和手臂中心轴线及其外极线的多几何约束策略, 采用求解反对应方法, 确保手指特征匹配对应的正确性. 通过实验验证, 证实了上述方法有效、可行, 可应用于实时、非穿戴的自然指势视觉3维人机交互中.

关键词: 指势识别; 人机交互; 非穿戴; 多尺度小波变换

中图分类号: TP391 **文献标识码:** A

Stereo vision-based recognition of nonwearable pointing gesture

GUAN Ye-peng

(School of Communication and Information Engineering, Shanghai University, Shanghai 200072, China;
Key Laboratory of Advanced Displays and System Application, Ministry of Education, Shanghai 200072, China)

Abstract: Based on the differences between the color intensities of R, G and B components on a color image in shadow regions, we develop a novel approach to the pointing object segmentation across a clutter background. Because the wavelet transformation is with outstanding local characteristics both in temporal and spatial fields, we suggest extracting the video pointing objects based on the combination of background subtraction with the wavelet multi-scale transformation. The proposed algorithm does not require the information which is necessary for existing methods in literature, such as scene learning and training, manual calibration and a priori hypothesis. It is also robust to dynamic scene variation, shadow and noise disturbance. Based on the biological structure characteristics, we employ the position of the human head-top instead of that of the human eyes in the pointing object segmentation. This is because that the human head-top is not easily occluded by other parts of the body, and is free from the effects of the facial orientation, posture and illumination variation. This provides the flexibility and casualness for the human-computer interaction(HCI), and ensures a high speed for the interaction. For the correct matching of finger tip characteristics, we present a stereo matching strategy based on the geometric constraints among the pointing arm finger tip, the central axis of the pointing arm and the corresponding epipolar line. A reverse matching criterion is employed to ensure the validity of the processed matching. Experiment results indicate that the developed approach is efficient for the recognition of the flexible and casual nonwearable pointing in the human-computer interaction.

Key words: pointing gesture recognition; HCI; nonwearable; multi-scale wavelet transformation

1 引言(Introduction)

随着计算机技术的迅猛发展, 研究符合人类自然交流习惯的新颖人机交互技术异常活跃, 且人机交互技术已从以计算机为中心逐步转移到以人为中心, 而多媒体用户界面则大大丰富了计算机信息的表现形式, 使用户可以交替或同时利用多个感觉通道. 然而, 多媒体用户界面的人机交互形式, 仍迫使用户使用常规的输入设备(如键盘、鼠标器和触摸屏

等)进行输入, 成为当今人机交互的瓶颈. 虚拟现实作为一种新型人机交互形式, 可实现和谐的、以人为中心的人机界面. 在虚拟现实中, 若以人手直接作为计算机输入设备, 则可充分利用人类的日常技能, 且不需特别训练或学习, 人机间的通信将不再需要中间媒介. 在以人手作为自然交互的工具领域中, 手势识别研究活跃^[1~3]. 但手势不仅受骨骼肌肉驱动, 且涉及人类思维活动, 导致它具有多样性、多义性. 相

对手势的多样性和多义性, 指势(pointing gesture)容易解释^[4]. 指势是人们日常生活中用手指对空间感兴趣目标的反映, 是人类语系发展和个体发生学的重要先驱, 可揭示人类社会智力^[5], 是一种理想的自然人机交互模式^[6].

指势识别分为穿戴式和非穿戴式两类^[4,6]. 前者基于数据头盔、数据手套和身体标记等, 这类方法可获得较准确的用户位置和用户行动, 但该方法为侵入式, 用户需专门训练, 操作不便; 后者基于非接触传感器(如摄像机). 在日常生活中, 人类习惯并大量使用非精确的信息交流, 允许使用模糊的表达手段, 以避免不必要的认识负荷, 提高交互活动的自然性和高效性. 且在人类信息的获取中, 80%以上来自于视觉, 因此, 非穿戴的指势视觉识别方法倍受青睐, 日益成为研究热点^[5,6].

2 目前非穿戴指势视觉识别方法(Non-wearable pointing vision recognition methods at present)

Colombo等人^[6]提出非侵入式指势视觉识别方法, 该法基于肤色检测手和头部区域, 采用匹配手指尖与头部重心, 确定用户指向的空间目标, 为避免人眼定位困难, 采用头部重心作为人眼位置. 该法除受光照条件影响外, 还受指势对象肤色和发色对手指与头部区域提取以及投影模式选择等因素影响. Cipolla等人^[7]基于用户指势为“手枪”模式, 基于无校正立体视觉, 采用约束性透视变换方法, 确定立体图像对中的食指, 并以立体图像对中的食指延长方向, 确定所指目标位置. 所提方法要求设置标记, 以求解透视变换参数, 且标记点与用户手臂应位于摄像机的视野中. 该方法受用户手势、标记点设置以及立体匹配等因素影响. Nickel等人^[8]基于隐马尔柯夫模型, 采用立体匹配策略结合磁力镜, 进行指势识别. Malerczyk^[9]基于检测、匹配双目摄像机获取的手指尖, 以手指尖的交线作为指势方位, 实现博物馆人机交互. 除要求两摄像机正交或近正交放置外, 还要求用户站立于指定地点(如设置的鞋印位置), 且用户为成人. Starner等人^[10]实现了基于指势识别的人机交互感知工作台, 为便于目标检测, 要求用于人机交互的感知工作台光照条件可控, 用户手臂不能超出工作台面或同时存在多个用户交互或遮挡等. Vira等人^[11]基于指势视觉识别方法进行外科医生手术时, 要求医生手持具有明显色彩标志的指示器或磁力棒, 通过检测、立体匹配指示器或磁力棒上的色彩标记, 确定指向目标的空间位置. Yamamoto等人^[12]通过安装于天花板四角的摄像机, 捕获人体和人脸, 通过视频对象分割和人体朝向与视线之间夹

角的计算, 确定所指空间目标. 所提方法除要求用户指势方向一定范围内无遮挡或障碍外, 还要求用户站立于某范围内, 且用户为成人. Sugi等人^[13]提出了基于指势识别的留意工作台定位所指目标方法. 在所提的方法中, 除要求用户佩戴颜色手套外, 还要求光源可控、稳定. Cernekova等人^[14]采用单目摄像机, 基于SVM分类识别的方法确定所指空间目标, 在所提方法中, 要求在地板上设置明显标记以规定用户活动范围.

经上述分析知: 上述方法在获取指势对象信息时, 均采用了增加限制方法, 如基于肤色、光源条件可控、用户对象单一或其它条件; 同时, 上述方法在确定指势空间目标时, 条件苛刻, 如要求指势对象为成人, 用户须站立于指定位置或限制摄像机位置. 上述条件与要求, 极大地限制了人机交互活动的普适性和自然性.

针对上述不足, 论文提出了一种新颖的非穿戴指势视觉识别方法. 与现有指势视觉识别方法相比, 论文主要贡献体现在: 1) 基于RGB色彩信息与多尺度小波时/空域融合, 提取视频指势对象, 克服环境条件约束, 适应复杂背景下人机交互的普适性和自然性. 2) 采用安置于头顶上方的立体摄像机取代需严格定位的摄像机, 捕获视频指势对象, 降低指势对象互遮挡, 扩大用户交互活动范围, 满足用户交互活动的自由性和自然性要求. 3) 采用不易遮挡的头顶特征取代人眼特征, 降低指势特征提取难度, 提高指势人机交互的时效性. 4) 基于指势手臂及其外极线的多几何约束关系及其反对应求解方法, 提高指势手指特征立体匹配的正确性和快捷性. 通过实验对比, 验证了文中所提方法有效、可行.

3 前景指势对象提取(Foreground pointing objects extraction)

背景差法是日前运动分割中常用方法. 一种简单的背景模型是时间平均图像, 但该法对于动态场景的变化较敏感. 利用图像序列中2个或3个相邻帧, 基于像素的时间差分法, 是提取运动目标区域的另一种常用方法^[15]. 该法对于动态环境具有较强的自适应性, 但一般不能完整提取出运动目标相关特征, 在运动实体内部也容易产生空洞现象. 虽然光流法能在背景运动存在的情况下, 也能检测出独立的运动目标, 但要求物体边缘与背景之间有明显的速度差, 且需要一定的后处理工作, 才能完整地检测出运动物体; 另外, 基于光流法检测运动目标时, 计算复杂、抗干扰能力差. 而快速准确的运动目标分割却又十分重要, 因在后处理过程中可仅考虑图像中对应于运动区域的像素, 大大简化工作计算量. 但由于

动态环境变化, 如光照变化、运动目标阴影等, 给准确有效的运动目标分割带来了困难. 其中, 阴影可能与被检测的运动目标相连或分离、扭曲目标形状或被误认为场景中一个完全错误的假目标^[16,17].

针对上述问题, 根据多尺度小波变换在时域和空域均具有优异的局部化特征, 结合背景差分, 利用小波多尺度特性, 对差分图像进行小波多尺度变换, 提取视频指势对象.

2维图像 $f(x, y)$ 在尺度 2^j 及方向 k 下的小波变换为^[18,19]:

$$W_{2^j}^k f(x, y) = f * \psi_{2^j}^k(x, y), k = 1, 2. \quad (1)$$

则在 x, y 方向上的小波 $\psi_{2^j}^k$, 可通过如下微分求取

$$\psi_{2^j}^1(x, y) = \frac{\partial \theta(x, y)}{\partial x}, \psi_{2^j}^2(x, y) = \frac{\partial \theta(x, y)}{\partial y}. \quad (2)$$

其中 $\theta(x, y)$ 为滤波尺度函数.

在不同尺度经 $\theta(x, y)$ 滤波后的2维小波为:

$$\nabla_{2^j} f(x, y) = \frac{1}{2^{2j}} \nabla f * \theta_{2^j}(x, y). \quad (3)$$

则不同尺度下的局部梯度幅值为:

$$M_{2^j} f(x, y) = \sqrt{(W_{2^j}^1 f(x, y))^2 + (W_{2^j}^2 f(x, y))^2}. \quad (4)$$

若某点 (x, y) 在尺度 2^j 上的梯度幅值 $M_{2^j} f$ 位于梯度相位 $A_{2^j} f$ 上, 则该点为多尺度边缘点, 其中,

$$A_{2^j} f(x, y) \equiv \arctan\left[\frac{W_{2^j}^2 f(x, y)}{W_{2^j}^1 f(x, y)}\right]. \quad (5)$$

通过不同尺度下的边缘点求取, 即可获取不同尺度 2^j 下的局部梯度集:

$$P_{2^j}(f) = \{p_{2^j,i} = (x_i, y_i); \nabla_{2^j} f(x_i, y_i)\}. \quad (6)$$

对于 J 层2维动态小波变换, 下列集合称之为图像 $f(x, y)$ 的多尺度边缘

$$\rho(I) = \{S_{2^j} f(x, y), [P_{2^j}(f)]_{1 \leq j \leq J}\}. \quad (7)$$

其中 $S_{2^j} f(x, y)$ 为 $f(x, y)$ 在 2^j 尺度下的低通滤波算子.

如图1为基于小波多尺度变换提取的某视频指势对象结果.



原始图像 初始分割结果 最终分割结果

图 1 频指势对象分割结果

Fig. 1 Video pointing object segmentation results

由于动态环境变化、噪声等因素影响, 造成阴影也被视为前景, 以及指势对象分割存在孔洞等(见图1中图).

为克服上述不利影响, 由阴影特性分析知: 在RGB彩色空间中, 由于 R, G, B 3色波长不同, 导致它们在阴影区域内强度变化存在差异, 为消除阴影影响, 将满足下式的各点视为阴影, 并从提取的前景区域中剔除.

$$\Delta I_R > \Delta I_G > \Delta I_B, \quad (8)$$

$$\Delta I_i = I_{fi} - I_{bi}, i = R, G, B. \quad (9)$$

其中: I_{fi} 代表 R, G, B 前景光强度, I_{bi} 代表背景中 R, G, B 光强度.

同时, 结合数学形态学对随机噪声进行压制. 图1右图为经上述处理后所得结果.

4 指势视觉识别(Pointing vision recognition)

基于指势进行人机交互时, 指势空间目标由指势手指尖与人眼视线的连线同目标所在平面的交点确定^[6,20]. 可见: 基于指势视觉进行指势目标识别, 需确定指势人眼及其手指尖的3维位置.

人眼特征的提取, 受到人脸朝向、姿态及光照变化、遮挡等因素影响, 造成人眼特征提取困难^[21,22]. 但基于人类生物结构特征知: 人类站立或坐立时, 人眼位于头顶下某范围, 且头顶特征不易遮挡, 不受人脸朝向、姿态及光照变化等因素影响, 易于提取. 因此, 可基于头顶特征代替人眼特征^[20]. 实验证实: 基于头顶特征代替人眼特征是可行的. 采用头顶特征代替人眼特征, 不仅保证了人机交互活动的自由性和自然性, 且大大提高了人机交互的时效性.

手指尖由于在前景对象分割图像中特征明显, 因此, 易于提取. 但由于存在场景变化、噪声干扰等因素影响, 造成指尖特征提取错误, 因此, 直接基于立体图像对中的手指尖特征进行3维重建将存在歧义性. 为有效确定手指尖的3维坐标, 基于指势手指特征点、指势手臂中心轴线及其外极线等多几何约束关系, 确定立体图像对上的手指尖对应匹配.

4.1 手指区域分割(Segmentation of pointing arm region)

基于前景指势对象分割结果, 提取手指区域, 根据手指在图像中的空间位置关系, 采用自适应且大小为 $1 \times N/2$ 的矩形结构元素对目标区域水平分割, 其中, N 为提取的指势目标区域水平方向最大连续域宽度. 采用上述方法, 对图1右图进行手指区域分割, 所得结果如图2.



图2 指势手指区域分割

Fig. 2 Pointing arm region segmentation

4.2 手指特征点与手臂轴线(Fingertip feature points and pointing arm axis line)

基于手指区域分割结果和手指在图像空间中的位置关系,手指尖2维图像位置为:

$$x_{\text{top}} = \arg \min(f(x, \min y)), \quad (10)$$

$$y_{\text{top}} = \arg(f(x_{\text{top}}, y)). \quad (11)$$

式中 $f(x, y)$ 为手指区域分割图像.

在手指区域分割过程中,由于存在手指区域分割不完整或前景/背景复杂,造成手指区域定位“过头”或缺损.为保证手臂中心轴线正确,基于手指分割区域,确定手指区域重心,结合手指尖位置,确定手臂中心轴线.

4.3 3维重建与指势识别(3D reconstruction and pointing recognition)

基于指尖特征点、指势手臂中心轴线及其外极线融合等多几何约束,重建3维指势手指尖.

在立体视觉系统中,设由双目摄像机获取的左、右图像分别为 P_1, P_2 ,则在图像 P_1 中的指尖点 p_1 必位于 P_2 图像中由 p_1 与两摄像机相对几何位置决定的极线上,其数学描述为:

$$u_2^T F u_1 = 0. \quad (12)$$

式中: F 为基础矩阵,由左、右摄像机对应的投影矩阵确定; u_1, u_2 为同名对应点的图像齐次坐标.

设双目左、右摄像机已标定,且在视频序列左、右图像中的手指尖、手臂区域重心位置已确定,分别为 $(x_{\text{top}}^i, y_{\text{top}}^i), (x_g^i, y_g^i) (i = 1, 2)$,则可得手臂中心轴线方程为:

$$y = (x - x_{\text{top}}^i)(y_g^i - y_{\text{top}}^i)/(x_g^i - x_{\text{top}}^i) + y_{\text{top}}^i. \quad (13)$$

联立方程(11), (12), 可由已知左(或右)图像指尖特征,确定其在另一图像中的对应点.

为消除上述求解过程中出现的歧义性,基于反对求解方法,即将求解对应点与已知参考点之间的欧氏距离小于某阈值 T ,作为求解的对应点为正确匹配对应点判据.基于确定的立体图像对的对应手指尖点、对应头顶特征点和摄像机投影矩阵,采用最小二乘法确定它们的3维坐标.

基于上述3维坐标,确定指势视线,并进而确定指势空间目标,实现指势视觉识别.

5 实验结果与分析(Experimental results and analysis)

采用如下实验装置(图3),验证文中所提方法的有效性.

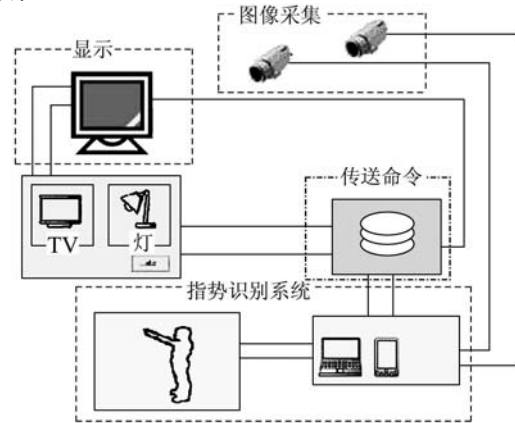
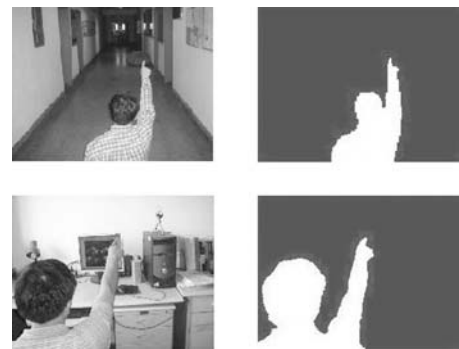


图3 实验装置示意图

Fig. 3 Experimental setup map

在上述实验装置中,利用安置于头顶上方的双目摄像机捕获指势视觉对象,通过视频对象分割方法提取指势对象及其相关特征,基于所提方法进行指势识别,确定所指空间目标.通过计算机发送触发命令,显示反馈所指目标结果.

基于上述前景视频对象分割方法,提取不同复杂背景条件下的指势对象的部分实验结果,如图4.



原始图像 指势对象分割二值图

图4 不同背景条件下指势对象分割结果.

Fig.4 Pointing object segmentation results in different background circumstances.

图4中的两幅原始图像,背景均复杂多变,其中,左上图中的背景主要受到镜面反射光、走廊前方透射光、白炽灯光等光照变化影响,而左下图中的背景则受到工作中的台灯、日光灯及计算机荧光屏等光源变化影响,并在不同区域间存在阴影.

由图4的指势对象分割结果知,文中所提方法,有效地从上述复杂的背景图像中提取出了指势对象.

为进一步验证文中所提方法在视频对象分割的

有效性, 采用具有复杂背景且存在较严重阴影、来自MPEG-4测试系列的Hall Monitor进行视频对象提取, 部分实验结果如图5.



图 5 基于Hall Monitor视频对象分割结果

Fig. 5 Results segmented of Hall Monitor video sequence

其中, 图5左图为原始图像, 右图为对象提取结果. 为反映分割结果与实际视频对象是否一致, 右图采用高亮灰度象素附加于当前图像方法, 突现分割结果. 由右图可见, 采用文中所提方法有效地提取了视频对象.

上述结果定性地表明文中方法分割视频对象有效、可行, 为进一步定量评估所提方法有效, 论文基于ROC曲线分析方法^[23~25]与文献[26,27]所提方法, 就上述视频测试系列进行了视频对象分割对比, 结果如图6.

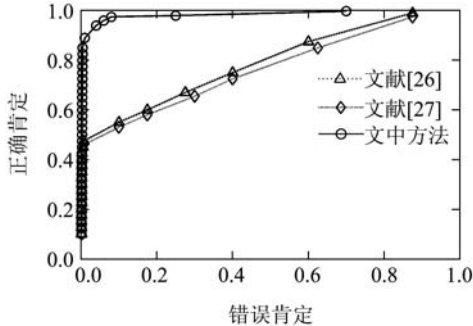


图 6 不同分割方法ROC曲线图

Fig. 6 ROC curve of different methods

经视频分割实验对比, 由图6知: 文中所提方法明显优于其它两方法, 从而进一步验证了文中方法在视频对象分割时具有强的鲁棒性.

为了验证指势视觉识别的有效性, 由年龄为20岁~65岁之间的36名学生和教师, 手指放置于摄像机前方3.5 m至12 m的12块标志牌(大小: 120 mm×200 mm), 其中, 36名志愿者使用各自指势习惯, 采用左、右手分别指向前方12块目标, 共得到864个指势集(12块目标×36用户×2手), 分别采用文中所提方法与文献[6,14]中方法对比, 其指势识别结果如表1.

由表1见, 不同对象手指前方不同目标时, 文中所提方法的识别结果较文献[6,14]方法理想, 表明文中所提指势识别方法可行、有效.

在运行时间上, 使用PIV 1.5G 512M RAM计算机,

在VC++6.0环境下, 采用DH-CG400图像采集卡进行图像采集(每帧图像大小: 320×240 pixels), 系统处理速度平均为350 ms, 基本满足人机交互实时要求.

由于立体匹配存在歧义且其本身问题的复杂性, 至今该问题尚未得到很好解决, 导致指势3维信息不准确, 从而影响指势识别, 因此, 为提高指势识别结果, 需融合多个摄像机并从中选择具有最佳视角摄像机, 利用多几何约束关系并基于最佳视角单目摄像机重建指势3维信息.

表 1 不同方法指势视觉识别结果

Table 1 Pointing recognition results of different methods

目标	位置/mm			识别率/(%)		
	X	Y	Z	文献[6]	文献[14]	文中方法
1	-1250	125	3500	97.4	94.5	97.1
2	0	125	3500	97.3	95.6	96.9
3	1250	125	3500	97.4	94.6	96.8
4	-1250	125	6000	94.2	92.4	96.0
5	0	125	6000	94.6	92.7	95.9
6	1250	125	6000	93.8	91.8	96.0
7	-1250	125	8500	92.6	90.2	95.3
8	0	125	8500	92.4	90.6	94.8
9	1250	125	8500	92.7	90.7	95.7
10	-1250	125	12000	90.0	89.6	94.0
11	0	125	12000	90.6	89.2	93.4
12	1250	125	12000	90.2	89.6	92.9

6 结论(Conclusions)

基于彩色图像中红、绿、蓝3分量光强度在阴影区域存在差异, 根据小波变换在时域和空域均具有优异的局部化特征, 结合背景差分, 利用小波多尺度特性进行小波多尺度变换, 提取视频指势对象, 可有效提取不同复杂背景条件下的视频指势对象.

基于人类生物结构特征, 采用不受人脸朝向、姿态、光照变化等影响, 以及不易遮挡的头顶特征代替人眼特征, 不仅保证了人机交互活动的自由性和自然性, 且大大提高了人机交互的时效性.

为克服指示手指特征匹配歧义性, 融合手指指尖特征点、指示手指手臂中心轴线及其外极线的多几何约束的立体匹配策略, 采用求解反对应方法, 确保了手指特征匹配对应的正确性.

通过实验验证, 证实了上述方法的有效性, 可应用于实时、非穿戴的自然指势视觉3维人机交互中.

参考文献(References):

[1] 徐光祐, 陶霖密, 史元春, 等. 普适计算模式下的人机交互[J]. 计算机学报, 2007, 30(7): 1041 - 1053.
(XU Guangyou, TAO Linmi, SHI Yuanchun, et al. Human computer interaction for ubiquitous/pervasive computing mode[J]. Chi-

- nese Journal of Computers, 2007, 30(7): 1041 – 1053.)
- [2] HONGMO J, DAIJIN K. Hand gesture recognition to understand musical conducting action[C] // *Proceedings of 16th IEEE International Symposium on Robot and Human Interactive Communication*. New York: IEEE Press, 2007: 163 – 168.
 - [3] VAFADAR M, BEHRAD A. Human hand gesture recognition using spatio-temporal volumes for human-computer interaction[C] // *Proceedings of International Symposium on Telecommunications*. New York: IEEE Press, 2008: 713–718.
 - [4] JOJIC N, BRUMITT B, MEYERS B, et al. Detection and estimation of pointing gestures in dense disparity maps[C] // *Proceedings of 4th IEEE International Conference on Automatic Face and Gesture Recognition*. Washington DC: IEEE Computer Society Press, 2000: 468 – 475.
 - [5] ENFIELD N J, KITA S, DE RUITER J P. Primary and secondary pragmatic functions of pointing gestures[J]. *Journal of Pragmatics*, 2007, 39(10): 1722 – 1741.
 - [6] COLOMBO C, BIMBO A D, VALLI A. Visual capture and understanding of hand pointing actions in 3-D environment[J]. *IEEE Transactions on Systems, Man and Cybernetics-Part B: Cybernetics*, 2003, 33(4): 677 – 686.
 - [7] CIPOLLA R, HOLLINGHURST N J. Human-robot interface by pointing with uncalibrated stereo vision[J]. *Image and Vision Computing*, 1996, 14(3): 171 – 178.
 - [8] NICKEL K, STIEFELHAGEN R. Visual recognition of pointing gestures for human-robot interaction[J]. *Image and Vision Computing*, 2007, 25(12): 1875 – 1884.
 - [9] MALERCZYK C. Interactive museum exhibit using pointing gesture recognition[C] // *Proceedings of 12th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision*. Plzen, Czech Republic: UNION Agency-Science Press, 2004: 165 – 172.
 - [10] STARNER T, LEIBE B, MINNEN D. The perceptive workbench: computer-vision-based gesture tracking, object tracking and 3D reconstruction for augmented desks[J]. *Machine Vision and Applications*, 2003, 14(1): 59 – 71.
 - [11] VIRA N, VIRA S. Surgeon's magic wand: a screen pointing interactive method[C] // *Proceedings of IEEE International Conference on Bioinformatics and Computational Biology*. New York: IEEE Press, 2006: 222 – 228.
 - [12] YAMAMOTO Y, YODA I, SAKAUE K. Arm-pointing gesture interface using surrounded stereo cameras system[C] // *Proceedings of 17th International Conference on Pattern Recognition*. Washington DC: IEEE Computer Society Press, 2004, 4: 965 – 970.
 - [13] SUGI M, NIKAIDO M, TAMURA Y. Development of gesture-based interface for deskwork support system[C] // *Proceedings of IEEE/RJS International Conference on Intelligent Robots and Systems*. New York: IEEE Press, 2006: 5171 – 5176.
 - [14] CERNEKOVA Z, NIKOLAIDIS N, PITAS I. Single camera pointing gesture recognition using spatial features and support vector machines[C] // *Proceedings of 15th European Signal Processing Conference*. Washington DC: IEEE Computer Society Press, 2007: 130 – 134.
 - [15] WREN C, AZABAYEJANI A, DARRELL T, et al. Pfunder: real-time tracking of the human body[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1997, 19(7): 780 – 785.
 - [16] PRATI A, MIKIC I, TRIVEDI M M, et al. Detecting moving shadows: formulation, algorithms and evaluation[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2003, 25(7): 918 – 924.
 - [17] 管业鹏, 顾伟康. 二维场景阴影区域的自动鲁棒分割[J]. *电子学报*, 2006, 34(4): 624 – 627.
(GUAN Yepeng, GU Weikang. Automatic and robust shadow segmentation from two-dimensional scenes[J]. *Acta Electronica Sinica*, 2006, 34(4): 624 – 627.)
 - [18] GUAN Y P. Wavelet multi-scale transform based foreground segmentation and shadow elimination[J]. *The Open Signal Processing Journal*, 2008, 1(6): 1 – 6.
 - [19] MALLAT S, ZHONG S. Characterization of signals from multiscale edges[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1992, 14(7): 710 – 732.
 - [20] GUAN Y P, ZHENG M. Real-time 3D pointing gesture recognition for natural HCI[C] // *Proceedings of 7th World Congress on Intelligent Control and Automation*. New York: IEEE Press, 2008: 2433 – 2436.
 - [21] GUAN Y P. Robust eye detection from facial image based on multi-cue facial information[C] // *Proceedings of IEEE International Conference on Control and Automation*. Washington DC: IEEE Computer Society Press, 2007: 1775 – 1778.
 - [22] WANG P, JI Q. Multi-view face and eye detection using discriminant features[J]. *Computer Vision and Image Understanding*, 2007, 105(2): 99 – 111.
 - [23] BAKER S G. The central role of receiver operating characteristic(ROC)curves in evaluating tests for the early detection of cancer[J]. *Journal of the National Cancer Institute*, 2003, 95(7): 511 – 515.
 - [24] FAWCETT T. An introduction to ROC analysis[J]. *Pattern Recognition Letters*, 2006, 27(8): 861 – 874.
 - [25] KEREKES J. Receiver operating characteristic curve confidence intervals and regions[J]. *IEEE Geosciences Remote Sensing Letters*, 2008, 5(2): 251 – 255.
 - [26] HOWE N, DESCHAMPS A. Better foreground segmentation through graph cuts[C] // *Proceedings of International Conference on Computational Intelligence and Multimedia Applications*. New York: IEEE Press, 2007: 267 – 271.
 - [27] ZIVKOVIC Z, VAN DER HEIJDEN F. Efficient adaptive density estimation per image pixel for the task of background subtraction[J]. *Pattern Recognition Letters*, 2006, 27(7): 773 – 780.

作者简介:

管业鹏 (1967—), 男, 博士, 教授, 上海市曙光学者, 主要研究方向为计算机视觉、模式识别等, E-mail: ypguan@shu.edu.cn.