



(12) 发明专利

(10) 授权公告号 CN 103530892 B

(45) 授权公告日 2016.06.22

(21) 申请号 201310497334.X

(22) 申请日 2013.10.21

(73) 专利权人 清华大学深圳研究生院

地址 518055 广东省深圳市南山区西丽大学城清华校区

(72) 发明人 朱艳敏 袁博

(74) 专利代理机构 深圳新创友知识产权代理有限公司 44223

代理人 江耀纯

(51) Int. Cl.

G06T 7/20(2006.01)

G06K 9/66(2006.01)

(56) 对比文件

CN 103034851 A, 2013.04.10,

CN 101901052 A, 2010.12.01,

CN 103257713 A, 2013.08.21,

Yanmin Zhu et al. tracking the articulated motion of two strongly interacting hands. 《Service Sciences》. 2013, 260-265.

I. Oikonomidis. Tracking the articulated motion of two strongly interacting hands. 《Computer Vision and Pattern Recognition》. 2012, 1862-1869.

Yi Li. Hand Gesture Recognition Using Kinect. 《Software Engineering and Service Science》. 2012, 196-199.

审查员 王齐强

权利要求书2页 说明书7页 附图2页

(54) 发明名称

一种基于 Kinect 传感器的双手跟踪方法与装置

(57) 摘要

本发明提供了一种基于 Kinect 传感器的双手跟踪方法与装置,所述跟踪方法,包括:S1 视频采集步骤;S2 检测步骤一;S3 单手跟踪步骤;S4 检测步骤二;S5 双手跟踪步骤,对已经检测到的双手进行跟踪。由于检测第二只手时利用了第一只手的信息,所以本发明的跟踪方法能够快速、准确的跟踪用户双手的运动,且计算复杂度低。



CN 103530892 B

1. 一种基于Kinect传感器的双手跟踪方法,其特征在于包括:

S1视频采集步骤,从Kinect传感器获取分辨率和帧率都相同的彩色视频流和深度视频流;

S2检测步骤一,从所获取的彩色视频流和深度视频流的图像中检测第一只做起始手势的手;

S3单手跟踪步骤,利用前一帧或两帧以上图像中第一只手的位置信息和尺寸信息对第一只手进行定位和跟踪;

S4检测步骤二,对深度图做滤波,寻找面积和深度位置与第一只手都相似的目标,利用第一只手的位置信息和尺寸信息检测第二只手;

S5双手跟踪步骤,对已经检测到的双手进行跟踪。

2. 如权利要求1所述的基于Kinect传感器的双手跟踪方法,其特征在于所述步骤S2中进一步包括:

S2-1样本训练步骤;

S2-2模式选择步骤;

S2-3起始手势判定步骤。

3. 如权利要求2所述的基于Kinect传感器的双手跟踪方法,其特征在于:所述步骤S2-1中,选用SVM分类器(Support Vector Machine)学习和训练手部的形态学信息,选取几何不变矩作为训练特征。

4. 如权利要求2所述的基于Kinect传感器的双手跟踪方法,其特征在于:所述步骤S2-2中,当光线适当时,选用肤色模式,即用肤色滤波结合深度滤波的方法提取第一只手;当光线太暗或太亮时,选用形状模式,即用深度滤波结合形状滤波的方法提取第一只手。

5. 如权利要求2所述的基于Kinect传感器的双手跟踪方法,其特征在于:所述步骤S2-3中,起始手势定义为:手向前伸出且距离身体在阈值d以上。

6. 如权利要求1所述的基于Kinect传感器的双手跟踪方法,其特征在于,所述步骤S3中的定位是:利用已得到的前一帧或两帧以上图像中第一只手的位置及其外接矩形预测当前帧图像中第一只手的ROI,在此ROI内做深度滤波以定位当前帧图像中第一只手的位置。

7. 如权利要求1所述的基于Kinect传感器的双手跟踪方法,其特征在于,所述步骤S5中的双手跟踪包括:

1) 双手在相互遮挡之前的分离状态时,对两个目标进行分别跟踪,即根据前一帧或两帧以上图像中两只手的位置和尺寸信息分别预测当前帧图像两目标各自的ROI,并在这两个区域内分别检测;

2) 双手在相互遮挡状态时,检测到的两个目标轨迹重合,视为对同一目标进行跟踪;

3) 双手在相互遮挡后分开时,根据遮挡前后双手在深度方向上位置关系的不变性辨别双手,并分别跟踪。

8. 一种基于Kinect传感器的双手跟踪装置,其特征在于包括如下模块:

视频采集模块,用于从Kinect传感器获取分辨率和帧率都相同的彩色视频流和深度视频流;

检测模块一,用于从所获取的彩色图像和深度图像中检测第一只做起始手势的手;

单手跟踪模块,包括定位单元和跟踪单元,用于利用前一帧或者两帧以上图像中第一

只手的位置信息和尺寸信息对第一只手进行定位和跟踪；

检测模块二,用于利用第一只手的位置信息和尺寸信息检测第二只手；

双手跟踪模块,用于对已经检测到的双手进行跟踪。

9.如权利要求8所述的基于Kinect传感器的双手跟踪装置,其特征在于:所述检测模块一包括样本训练单元、模式选择单元和起始手势判定单元。

10.如权利要求9所述的基于Kinect传感器的双手跟踪装置,其特征在于:所述样本训练单元选用SVM分类器学习和训练手部的形态学信息,选取几何不变矩作为训练特征。

11.如权利要求9所述的基于Kinect传感器的双手跟踪装置,其特征在于,所述模式选择单元用于:

当光线适当时,选用肤色模式,即用肤色滤波结合深度滤波的方法提取第一只手;

当光线太暗或太亮时,选用形状模式,即用深度滤波结合形状滤波的方法提取第一只手。

12.如权利要求9所述的基于Kinect传感器的双手跟踪装置,其特征在于:所述起始手势判定单元中,起始手势定义为:手向前伸出且距离身体在阈值d以上。

13.如权利要求8所述的基于Kinect传感器的双手跟踪装置,其特征在于,所述定位单元用于:利用已得到的前一帧或者两帧以上图像中第一只手的位置及其外接矩形预测当前帧图像中第一只手的ROI,在此ROI内做深度滤波以定位当前帧图像中第一只手的位置。

14.如权利要求8所述的基于Kinect传感器的双手跟踪装置,其特征在于,所述双手跟踪模块用于:

1)双手在相互遮挡之前的分离状态时,对两个目标进行分别跟踪,即根据前一帧或者两帧以上图像中两只手的位置和尺寸信息分别预测当前帧图像两目标各自的ROI,并在这两个区域内分别检测;

2)双手在相互遮挡状态时,检测到的两个目标轨迹重合,视为对同一目标进行跟踪;

3)双手在相互遮挡后分开时,根据遮挡前后双手在深度方向上位置关系的不变性辨别双手,并分别跟踪。

## 一种基于Kinect传感器的双手跟踪方法与装置

### 技术领域

[0001] 本发明涉及人机交互领域中手势识别和双手跟踪,尤其涉一种基于Kinect传感器的双手跟踪方法与装置。

### 背景技术

[0002] 随着计算机技术的迅猛发展和人们科技观念的进步,用户对计算机的感知能力提出了越来越高的要求。传统的人机交互模式主要依赖于物理键盘和鼠标,其比较单一的以文本为主的输入方式已经不能满足人们的需求,而更注重“以人为本”的新型人机交互方式打破了传统模式的束缚,把输入方式转变成图像和声音等更加丰富自然的形式,大大改善了用户的体验。近年来人们在人脸识别、语音识别、人体姿态识别、手势识别等研究领域都取得了长足的进步。

[0003] 手势在人们的生活中扮演着极其重要的角色,基于计算机视觉的手势识别是实现新一代人机交互所不可或缺的一项关键技术。手势识别中的关键和大多数难点都存在于人手的跟踪部分。其面临的主要挑战有:一、复杂背景的干扰,如人脸和别的类肤色背景;二、手在运动过程中会发生形变;三、光照的变化对手的外观影响很大;四、系统运行实时性的要求。

[0004] 目前手势识别技术大多是针对单手情况,所用的定位跟踪方法有:Camshift (Continuously Adaptive Mean-SHIFT)算法,特征空间匹配法等。这些已有的方法只是在特定的条件下取得不错的效果。双手的跟踪和识别相比于单手又有新的特点和挑战,如在双手发生相互遮挡之后如何准确分辨和跟踪双手。现有技术中有的做法是限制两只手在遮挡前后形状不变,双手遮挡分开后,利用遮挡前的双手形状信息进行识别。但这种方法对人手限制较大,不够自然方便。微软公司的Kinect为手势跟踪和识别提供了很大的便利,它提供的深度信息大大简化了背景去除步骤。现有技术中有利用Kinect提供的深度图像和彩色图像建立双手的3D模型并进行跟踪的方法,可准确定位到双手关节等细节信息,但这种方法计算复杂度非常高,即便在使用GPU(Graphic Processing Unit)加速的情况下也不能达到实时跟踪。因Kinect本身带有提取和跟踪人体骨架的功能,故现有技术中也有用在骨架上寻找手部节点的方法提取手部位置从而跟踪人手,但这种方法要求用户的姿态是标准的坐姿或者站姿,对用户姿势限制太多,且识别和追踪效果不够理想。

### 发明内容

[0005] 本发明所要解决的技术问题是:克服前述现有技术的缺陷,提供一种可以处理复杂背景和不同光照情况且对人手姿态限制很少的基于Kinect传感器的双手跟踪方法,包括:

[0006] S1视频采集步骤,从Kinect传感器获取分辨率和帧率都相同的彩色视频流和深度视频流;

[0007] S2检测步骤一,从所获取的彩色视频流和深度视频流的图像中检测第一只做起始

手势的手；

[0008] S3单手跟踪步骤,利用前一帧或两帧以上图像中第一只手的位置信息和尺寸信息对第一只手进行定位和跟踪；

[0009] S4检测步骤二,利用第一只手的位置信息和尺寸信息检测第二只手；

[0010] S5双手跟踪步骤,对已经检测到的双手进行跟踪。

[0011] 根据实施例,本发明还可采用以下优选的技术方案：

[0012] 所述步骤S2中进一步包括：S2-1样本训练步骤；S2-2模式选择步骤；S2-3起始手势判定步骤。

[0013] 所述步骤S2-1中,选用SVM分类器(Support Vector Machine)学习和训练手部的形态学信息,选取几何不变矩作为训练特征。

[0014] 所述步骤S2-2中,当光线适当时,选用肤色模式,即用肤色滤波结合深度滤波的方法提取第一只手；当光线太暗或太亮时,选用形状模式,即用深度滤波结合形状滤波的方法提取第一只手。

[0015] 所述步骤S2-3中,起始手势定义为：手向前伸出且距离身体在阈值d以上。

[0016] 所述步骤S3中的定位是：利用已得到的前一帧或两帧以上图像中第一只手的位置及其外接矩形预测当前帧图像中第一只手的ROI(Region Of Interest),在此ROI内做深度滤波以定位当前帧图像中第一只手的位置。

[0017] 所述步骤S5中的双手跟踪包括：

[0018] 1)双手在相互遮挡之前的分离状态时,对两个目标进行分别跟踪,即根据前一帧或两帧以上图像中两只手的位置和尺寸信息分别预测当前帧图像两目标各自的ROI,并在这两个区域内分别检测；

[0019] 2)双手在相互遮挡状态时,检测到的两个目标轨迹重合,视为对同一目标进行跟踪；

[0020] 3)双手在相互遮挡后分开时,根据遮挡前后双手在深度方向上位置关系的不变性辨别双手,并分别跟踪。

[0021] 本发明还提供一种基于Kinect的双手跟踪装置,包括如下模块：

[0022] 视频采集模块,用于从Kinect传感器获取分辨率和帧率都相同的彩色视频流和深度视频流；

[0023] 检测模块一,用于从所获取的彩色图像和深度图像中检测第一只做起始手势的手；

[0024] 单手跟踪模块,包括定位单元和跟踪单元,用于利用前一帧或两帧以上图像中第一只手的位置信息和尺寸信息对第一只手进行定位和跟踪；

[0025] 检测模块二,用于利用第一只手的位置信息和尺寸信息检测第二只手；

[0026] 双手跟踪模块,用于对已经检测到的双手进行跟踪。

[0027] 根据实施例还可采用以下优选的技术方案：

[0028] 所述检测模块一包括样本训练单元、模式选择单元和起始手势判定单元。

[0029] 所述样本训练单元选用SVM分类器学习和训练手部的形态学信息,选取几何不变矩作为训练特征。

[0030] 所述模式选择单元用于：当光线适当时,选用肤色模式,即用肤色滤波结合深度滤

波的方法提取第一只手;当光线太暗或太亮时,选用形状模式,即用深度滤波结合形状滤波的方法提取第一只手。

[0031] 所述起始手势判定单元中,起始手势定义为:手向前伸出且距离身体在阈值d以上。

[0032] 所述定位单元用于:利用已得到的前一帧或两帧以上图像中第一只手的位置及其外接矩形预测当前帧图像中第一只手的ROI,在此ROI内做深度滤波以定位当前帧图像中第一只手的位置。

[0033] 所述双手跟踪模块用于:

[0034] 1)双手在相互遮挡之前的分离状态时,对两个目标进行分别跟踪,即根据前一帧或两帧以上图像中两只手的位置和尺寸信息分别预测当前帧图像两目标各自的ROI,并在这两个区域内分别检测;

[0035] 2)双手在相互遮挡状态时,检测到的两个目标轨迹重合,视为对同一目标进行跟踪;

[0036] 3)双手在相互遮挡后分开时,根据遮挡前后双手在深度方向上位置关系的不变性辨别双手,并分别跟踪。

[0037] 本发明与现有技术对比的有益效果是:

[0038] 由于检测第二只手时利用了第一只手的信息,所以本发明的跟踪方法能够快速、准确的跟踪用户双手的运动,且计算复杂度低。

[0039] 一个优选的技术方案中,由于采用了深度滤波的方法并在跟踪过程中设定了目标的感兴趣区域ROI,所以可以不受复杂背景的干扰,因为深度滤波可以去除目标后方干扰物如用户的脸部的影响;而设定ROI区域使得只在目标附近的区域检测下一帧图像中的目标,去除了ROI区域外的物体,如旁观者的手部和脸部等的干扰;以上还带来另一好处,就是对用户的姿态限制很少。

[0040] 另一个优选的技术方案中,用户可以根据光线情况选择合适的检测模式,使得此方法能够适应不同光线的情况。

[0041] 再一个优选的技术方案中,在跟踪过程中检测手时不使用其形状信息,因此手在运动过程中的形变不影响跟踪效果。

## 附图说明

[0042] 图1是本发明一个实施例的双手跟踪的流程图。

[0043] 图2是一个实施例的检测第一只手初始手势的流程图。

## 具体实施方式

[0044] 下面对照附图和结合优选具体实施方式对本发明进行详细的阐述。

[0045] 实施例1

[0046] 一种基于Kinect的双手跟踪方法,包括:

[0047] S1视频采集步骤,从Kinect传感器获取分辨率和帧率都相同的彩色视频流和深度视频流;

[0048] S2检测步骤一,从所获取的彩色视频流和深度视频流的图像中检测第一只做起始

手势的手；

[0049] S3单手跟踪步骤,利用前一帧或两帧以上中第一只手的位置信息和尺寸信息对第一只手进行定位和跟踪；

[0050] S4检测步骤二,利用第一只手的位置信息和尺寸信息检测第二只手；

[0051] S5双手跟踪步骤,对已经检测到的双手进行跟踪。

[0052] 所述步骤S2中进一步包括:S2-1样本训练步骤;S2-2模式选择步骤;S2-3起始手势判定步骤。

[0053] 所述步骤S2-1中,选用SVM分类器学习和训练手部的形态学信息,选取几何不变矩作为训练特征。

[0054] 所述步骤S2-2中,当光线适当时,选用肤色模式,即用肤色滤波结合深度滤波的方法提取第一只手;当光线太暗或太亮时,选用形状模式,即用深度滤波结合形状滤波的方法提取第一只手。

[0055] 所述步骤S2-3中,起始手势定义为:手向前伸出且距离身体在阈值d以上。

[0056] 所述步骤S3中的定位是:利用已得到的前一帧或两帧以上图像中第一只手的位置及其外接矩形预测当前帧图像中第一只手的ROI,在此ROI内做深度滤波以定位当前帧图像中第一只手的位置。

[0057] 所述步骤S5中的双手跟踪包括:

[0058] 1)双手在相互遮挡之前的分离状态时,对两个目标进行分别跟踪,即根据前一帧或两帧以上图像中两只手的位置和尺寸信息分别预测当前帧图像两目标各自的ROI,并在这两个区域内分别检测;

[0059] 2)双手在相互遮挡状态时,检测到的两个目标轨迹重合,视为对同一目标进行跟踪;

[0060] 3)双手在相互遮挡后分开时,根据遮挡前后双手在深度方向上位置关系的不变性辨别双手,并分别跟踪。

[0061] 实施例2

[0062] 如图1所示,是本实施例的双手跟踪的流程图,包括:

[0063] 步骤1)获取视频流。譬如使用Kinect传感器获取分辨率为640\*480,帧率为30fps的彩色视频流和深度视频流。

[0064] 步骤2)检测第一只手。即从获取到的彩色图和深度图中检测第一只手。具体是根据深度信息和肤色信息或者形状信息寻找最近的足够大的目标物体作为第一只手。

[0065] 步骤3)判断是否检测到初始手势。即判断找到的第一只手是否在做起始手势。当手向前伸出,距离身体在阈值d以上时,判定此手做出有效起始手势,跟踪开始。标记此手为Hand1,把其三维位置信息存入此手的轨迹traj1中。

[0066] 其中d一般设为15~25cm。

[0067] 步骤4)跟踪单只手。即跟踪检测到的所述第一只手Hand1。具体可以是:根据前两帧图像中手的位置和尺寸信息,预测当前帧图像中目标的可能位置区域,即感兴趣区域ROI。在深度图的ROI区域内寻找最可能的目标,把目标的三维位置信息存入Hand1的轨迹traj1中。所述前两帧是指跟踪过程中当前帧之前的两帧图像,如当前帧是 $f_t$ ,那么前两帧指 $f_{t-1}$ 和 $f_{t-2}$ 。当然,利用前两帧只是一个优选的做法,也不一定必须是前“两”帧,如当开始

跟踪时,前面只有一帧图像中(检测到第一只手的那帧)有手的信息,那就只能利用前一帧的手的位置尺寸信息。

[0068] 这个步骤主要作用是利用“前N帧”的信息对当前帧目标的位置做预测。N越大,可以参考的信息越多,但预测公式会更复杂。另外,N越大,前面第N帧图像的参考价值越小,所以N不是越大越好,取2或3可能比较合适。

[0069] 步骤5)检测第二只手。因双手手势的自然起始状态是两只手伸出相近的深度距离,摆出相似的手势,因此可以利用当前帧图像中第一只手Hand1的位置和尺寸信息在全局范围内检测第二只手。对深度图做滤波,寻找面积和深度位置与第一只手都相似的目标,这种方法相对于利用颜色信息或者形状信息重新搜索另外一只手的方法更加简单快速。

[0070] 步骤6)判断是否检测到目标。即判定步骤5是否找到第二只手,如果找到了,标为第二只手Hand2,将其三维位置信息存入Hand2的轨迹tra<sub>j</sub>2中;如果没找到,则继续跟踪单只手。

[0071] 步骤7)跟踪双手。即在找到第二只手之后对双手(Hand1和Hand2)进行跟踪。在两只手没有相互遮挡或者重叠的时候,相当于对两个单独的目标进行跟踪,根据前两帧图像中目标的位置和尺寸信息预测当前帧图像中两目标的可能位置区域,设定两个目标的各自的感兴趣区域ROI<sup>1</sup>和ROI<sup>2</sup>,并在深度图的ROI<sup>1</sup>和ROI<sup>2</sup>范围内做深度滤波分别寻找最有可能的目标。当两只手相互遮挡或者重叠时,在ROI<sup>1</sup>和ROI<sup>2</sup>内检测到的是同一目标,两只手的位置信息相同,轨迹重合。当两只手相互遮挡之后再分开时,根据遮挡前和遮挡后两手的深度值关系的不变性来分辨Hand1和Hand2,即如果在遮挡前Hand1在Hand2之前,那么遮挡分开后Hand1应该仍然在Hand2之前。

[0072] 步骤8)输出手的轨迹。即分辨出两手后,将其各自三维位置信息分别存入对应的轨迹中。

[0073] 如图2所示,是本实施例中检测第一只手初始手势的流程图,主要包括:

[0074] 步骤201,获取彩色图像和深度图像:从Kinect获取分辨率为640\*480,帧率为30fps的RGB彩色图像和灰度深度图像。

[0075] 步骤202判断光线是否合适。根据光照强弱选择检测模式,如果光照适中,选择肤色模式;如果光照太弱或者太亮,选择形状模式。

[0076] 步骤203肤色滤波。肤色模式首先把RGB图片转换到肤色更为集中的YCbCr颜色空间,其中Y表示明亮度(Luminance、Luma),Cb和Cr都指色彩,Cb指蓝色色度,Cr指红色色度。转换公式如下:

[0077]  $Y=0.299*R+0.587*G+0.114*B$

[0078]  $Cr=(R-Y)*0.713+128$

[0079]  $Cb=(B-Y)*0.564+128$

[0080] 然后进行肤色滤波,用椭圆肤色模型对YCbCr色彩空间的彩色图片进行滤波得到肤色的二值图像mask1。

[0081] 步骤204深度滤波:结合深度图对肤色二值图像mask1进行深度滤波。从mask1中找到所有面积大于手部最小面积A<sub>min</sub>的连通域,选出其中面积最大的3个,得到新的肤色二值图像mask2。在mask2中选取深度最小的连通域SR<sub>n</sub>,对其进行深度滤波,选取深度范围在[D<sub>min</sub>,D<sub>min</sub>+1]内的区域作为准目标区域R<sub>c</sub>。其中,D<sub>min</sub>为连通域SR<sub>n</sub>的最小深度值,1为用



于分割手部的阈值,一般设为5~8cm(此处是指从手的最前端如指尖到手腕的深度距离,后文所述“分割手部的阈值”的设定含义与此处相同)。

[0082] 步骤205样本训练:形状模式要求用户和摄像头之间没有物体干扰,用户处于最前方,且做起始手势时手掌要张开。首先采集张开手掌的正负样本数据库,再选取特征和分类器学习训练手形。由于几何不变矩——Hu矩具有平移、旋转和尺度不变性,选择它作为训练分类器的特征。Hu矩定义如下:

[0083] 其中, $\eta_{pq}$ 是(p+q)阶归一化中心矩。

[0084]  $I_1 = \eta_{20} + \eta_{02}$

[0085]  $I_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2$

[0086]  $I_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2$

[0087]  $I_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2$

[0088]  $I_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] +$

[0089]  $(3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]$

[0090]  $I_6 = (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03})$

[0091]  $I_7 = (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] -$

[0092]  $(\eta_{30} - 3\eta_{12})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]$

[0093] SVM(Support Vector Machine,支持向量机)在解决小样本、非线性及高维度分类问题时有许多特有的优势,因此选用SVM对手部的Hu矩特征进行学习和训练,生成手形分类器。

[0094] 步骤206深度滤波:对深度图进行深度滤波,定位深度范围在 $[D_{min}, D_{min}+1]$ 内的最大物体轮廓,如果其面积大于手型最小面积 $A_{min}$ ,则此轮廓包含的区域为可能目标区域 $R_n$ 。其中, $D_{min}$ 为深度图最小深度值,1为用于分割手部的阈值,一般设为5~8cm。

[0095] 步骤207形状辨别:用训练好的SVM手形分类器对可能目标区域 $R_n$ 做分类,判别是否是张开的手部。如果是,则此可能目标区域为准目标区域 $R_c$ 。

[0096] 步骤208判断是否找到目标手:判定准目标区域 $R_c$ 是否是要做手势的人手,即判定手是否做出起始手势。首先对深度图像进行滤波,取深度范围在 $[D_{min}, D_{min}+d]$ 内的区域,选取其中包含 $R_c$ 的区域 $R_b$ ,计算两个区域的面积比值。

[0097]  $ratio = area(R_c) / area(R_b)$

[0098] 根据ratio的范围判断目标是否是伸出的手。

[0099]  $ratio \in \begin{cases} [0.5, 1], & \text{有效起始手势} \\ [0, 0.5), & \text{无效起始手势} \end{cases}$

[0100] 如果ratio在 $[0.5, 1]$ 范围内,则判定 $R_b$ 不包括身体,只包括手臂部分,此时判定准目标 $R_c$ 是伸出的手,是有效的起始手势,否则不是有效的起始手势。

[0101] 其中, $D_{min}$ 为准目标区域 $R_c$ 的最小深度,d为用于判定人手是否伸出的阈值,一般设为15~25cm。

[0102] 步骤209锁定目标:如果找到了有效的起始手势,标记找到的目标手为Hand1,把其三维位置信息存入Hand1的轨迹

[0103] 找到起始手势之后,后面步骤的详细实施过程描述如下:

[0104] 单手跟踪,找到第一只手后进入了单手跟踪模块,利用前两帧 $f_{t-2}$ 和 $f_{t-1}$ 中Hand1的位置和尺寸信息确定当前帧 $f_t$ 的矩形感兴趣区域 $ROI^1$ 。 $ROI^1$ 的确定方法为:

$$[0105] \quad x=2x_{t-1}-x_{t-2}$$

$$[0106] \quad y=2y_{t-1}-y_{t-2}$$

$$[0107] \quad width=1.5*width(boundRect)$$

$$[0108] \quad height=1.5*height(boundRect)$$

[0109] 其中 $(x, y)$ 为 $ROI^1$ 的中心点, $width$ 和 $height$ 分别为其宽和高, $(x_{t-2}, y_{t-2})$ 和 $(x_{t-1}, y_{t-1})$ 为Hand1中心点在前两帧图像 $f_{t-2}$ 和 $f_{t-1}$ 中 $xy$ 平面的坐标, $boundRect$ 为前一帧 $f_{t-1}$ 中手形区域的外接矩形。如果 $t < 2$ ,即当前帧为检测到起始手势后的第一帧,此时仅有前一帧的手的信息,则 $ROI^1$ 的中心点为 $x=x_{t-1}, y=y_{t-1}$ ,宽和高与上面相同。在深度图的 $ROI^1$ 范围内寻找深度在 $[Dmin, Dmin+1]$ 内的面积大于 $Amin$ 的最大连通区域,判定为跟踪目标,将其三维位置信息存入Hand1的轨迹 $traj1$ 中。其中, $Dmin$ 为 $ROI^1$ 范围内的最小深度值, $l$ 为用于分割手部的阈值,一般设为5~8cm。

[0110] 检测第二只手,在当前帧中寻找深度值和面积都与Hand1相似的目标,即深度在 $[z_t-2cm, z_t+1+2cm]$ 范围内,且面积在 $[area_t*0.8, area_t*1.2]$ 范围内的连通区域,标记为第二只手Hand2,将其三维位置信息存入其对应的轨迹 $traj2$ 中。其中, $z_t$ 为Hand1的最小深度值, $area_t$ 为Hand1的面积, $l$ 为分割手部的阈值,一般设为5~8cm。

[0111] 双手跟踪,当两手处于遮挡前分离的状态时,根据前两帧 $f_{t-2}$ 和 $f_{t-1}$ 中Hand1和Hand2的位置和大小信息设定当前帧 $f_t$ 中各自的感兴趣区域 $ROI^1$ 和 $ROI^2$ 。在 $ROI^1$ 中寻找目标Hand1,在 $ROI^2$ 中寻找目标Hand2,方法与单手跟踪相同。当两只手发生相互遮挡时, $ROI^1$ 与 $ROI^2$ 重合,检测到的两个目标实际为同一目标,此时两只手的轨迹重合。当两只手遮挡之后分开时,根据两只手遮挡前后在其深度方向上位置关系的不变性来分辨双手。

$$[0112] \quad z_s^1 < z_s^2 \Rightarrow z_t^1 < z_t^2$$

[0113] 其中 $s$ 为遮挡前的时刻, $t$ 为遮挡分开后的时刻, $z_s^1$ 和 $z_s^2$ 分别为 $s$ 时刻Hand1和Hand2的深度值, $z_t^1$ 和 $z_t^2$ 分别为 $t$ 时刻Hand1和Hand2的深度值。即如果在相互遮挡前的时刻Hand1处于Hand2之前,那么遮挡结束后的时刻,Hand1也在Hand2之前。用这种方法分辨出双手后,对其做标示并把它们的三维位置信息分别存入各自的轨迹中。

[0114] 应当理解,本发明的装置实施例中的各个模块和单元的具体操作过程可与方法实施例中的描述相同,此处不再详细描述。

[0115] 以上内容是结合具体的优选实施方式对本发明所作的进一步详细说明,不能认定本发明的具体实施只局限于这些说明。对于本发明所属技术领域的普通技术人员来说,在不脱离本发明构思的前提下,做出若干等同替代或明显变型,而且性能或用途相同,都应当视为属于本发明的保护范围。

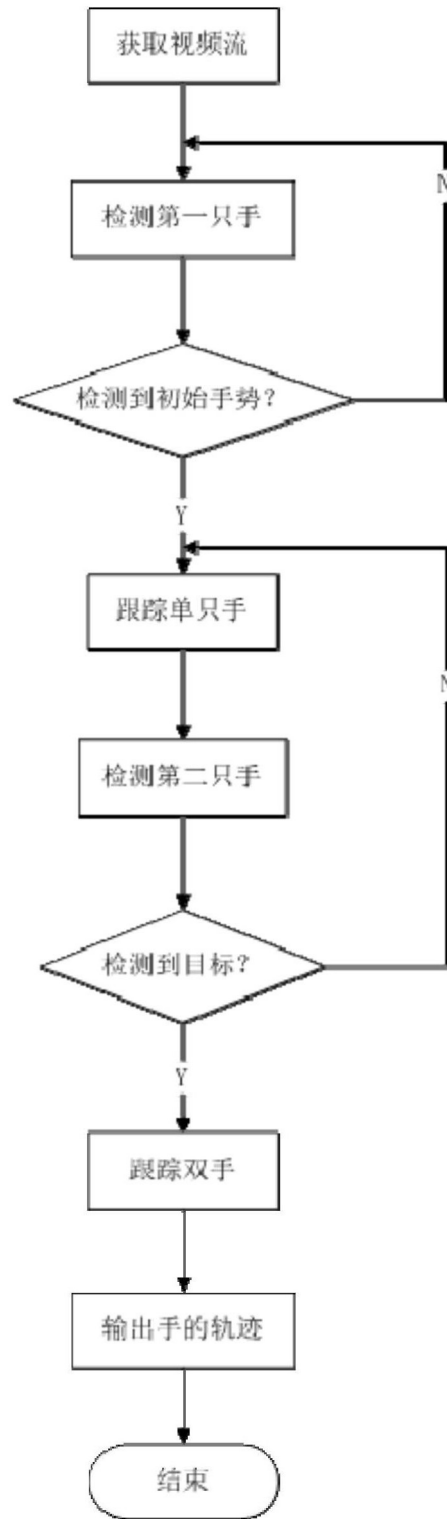


图1

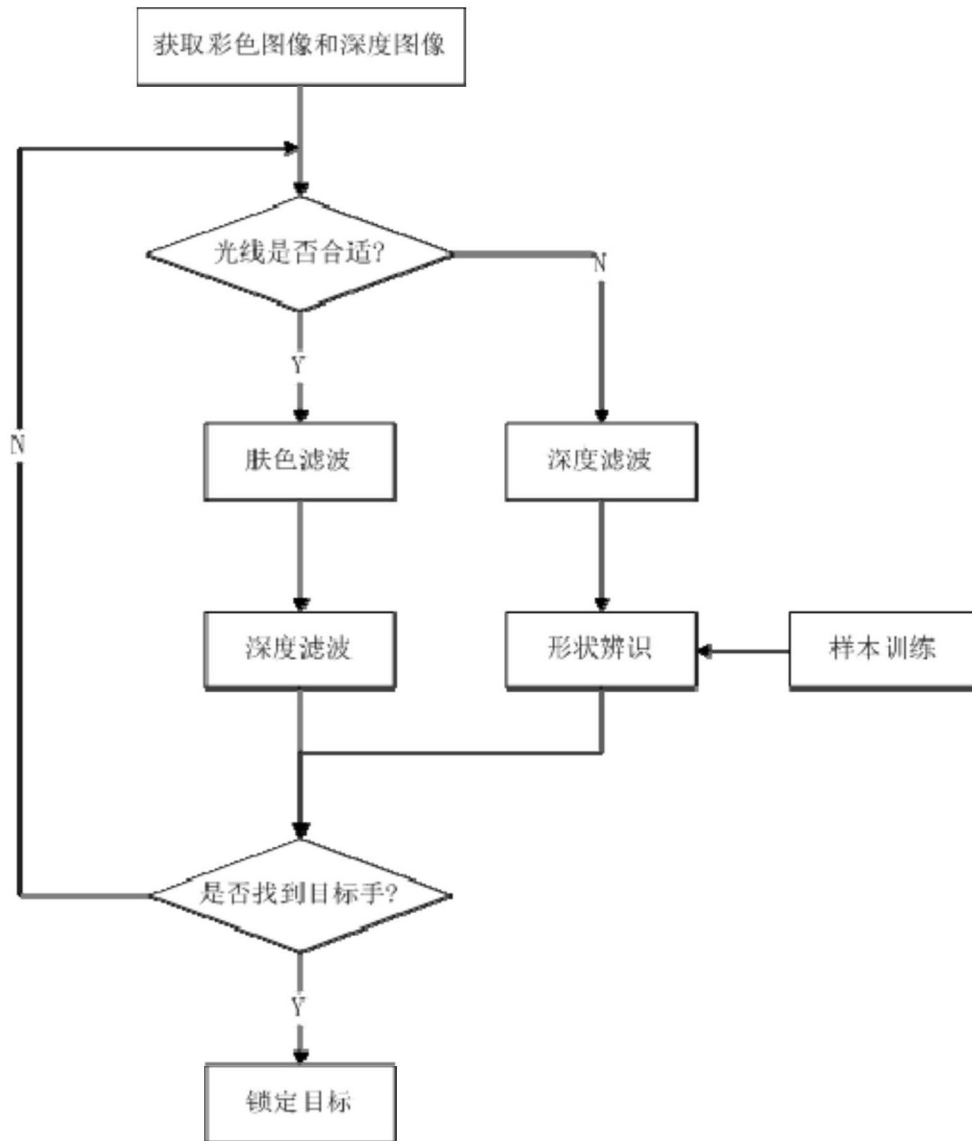


图2