

自拍相机运动手势交互界面的设计与实现

褚少微¹, 陆东伟², 潘瑞芳¹

CHU Shaowei¹, LU Dongwei², PAN Ruifang¹

1. 浙江传媒学院 新媒体学院, 杭州 310018

2. 中国传媒大学 计算机学院, 北京 100024

1. School of New Media, Zhejiang University of Media and Communications, Hangzhou 310018, China

2. School of Computer, Communication University of China, Beijing 100024, China

CHU Shaowei, LU Dongwei, PAN Ruifang. Motion gesture interface for self-portrait camera. Computer Engineering and Applications, 2017, 53(10): 181-186.

Abstract: With taking self-portraits users can capture memories, create art, and advance their photography skills. However, most existing camera interfaces are limited in that they do not support device-less remote control, and real-time control over important camera functions, etc. In this paper, a new self-portrait camera system is described and a gesture interface for self-portraits is developed. This self-portrait camera system supports a rotatable front lens and a motion-based gesture interface to control camera functions including the shutter trigger, color balance, and exposure. The gesture-recognition accuracy is experimentally evaluated and the effectiveness of the system is examined in a user study. The results suggest that the proposed gesture recognition technique achieves 85% accuracy, and the gesture-based interface is effective for controlling self-portrait camera options and improves the user experience when taking self-portraits.

Key words: self-portrait; gesture interface; motion gesture; human computer interaction; user interface

摘 要: 随着智能移动终端的发展及摄像镜头的小型化, 自拍变得越来越流行。如何设计新型自拍交互方法使得用户在自拍过程中能够自由、实时地控制相机是自拍相机交互界面的关键问题。提出利用基于视觉的运动手势交互界面的新方法, 使自拍过程中用户只要挥一挥手臂就可以实现与自拍相机的交互功能。使用手势交互的方法, 用户可以把相机放在任意的平台上, 自由地摆出各种自拍姿态, 增加了自拍的丰富性, 提高了用户体验。主要提出挥手及画圈两种交互手势, 通过组合应用可以实现丰富高效的自拍交互控制功能, 如快门控制、白平衡、曝光度等。手势的识别利用相机摄像的实时图像进行处理, 采用稀疏光流算法来识别运动手势。用户评估实验表明, 所提出运动手势自拍交互界面具有较好的交互效率以及良好的用户满意度, 两种手势的识别效率约为 85%。

关键词: 自拍; 手势界面; 运动手势; 人机交互; 用户界面

文献标志码: A **中图分类号:** TP399 **doi:** 10.3778/j.issn.1002-8331.1512-0158

1 引言

自拍是随着智能移动终端, 特别是智能手机上摄像头的普遍引入以及社交网络的快速发展而流行起来的热门事物。智能手机的便捷性使得人们可以利用自拍随时随地记录自己的状态、生活中的点滴经历, 以及与亲友分享自己的状况。同时自拍还具有学习提高摄影技术、创作艺术作品等功能^[1-3]。

目前在智能手机上被普遍使用的自拍方法主要有: 按键控制、计时器、遥控等。利用按键控制的方法常常是用户用手拖住相机让镜头朝向自己, 然后按下快门键。此种方法直接快捷, 但局限是容易因为手抖而使所成图像模糊, 同时由于用户手臂长度的限制只能拍摄出大头照^[4]。利用计时器的方法用户可以把相机放在稳定的平台上, 自由地摆出各种姿势, 但往往需要用户反复

基金项目: 国家新闻出版广电总局科研项目 (No.Z412F14208); 国家自然科学基金青年项目 (No.61502415)。

作者简介: 褚少微 (1982—), 男, 博士, 讲师, CCF 会员, 研究领域为人机交互, E-mail: chu@zjicm.edu.cn; 陆东伟 (1993—), 男, 硕士研究生, 研究领域为人机交互; 潘瑞芳 (1959—), 女, 教授, 研究领域为数字媒体技术。

收稿日期: 2015-12-12 **修回日期:** 2016-03-01 **文章编号:** 1002-8331(2017)10-0181-06

CNKI 网络优先出版: 2016-03-25, <http://www.cnki.net/kcms/detail/11.2127.TP.20160325.2025.078.html>

调整设置查看成像结果,交互效率不高。使用遥控器是一种较好的交互方法,可以解决距离和实时控制的问题,但物理遥控器占用用户的手臂,使得用户不能够自由地摆出各种自拍姿势^[5]。

近年来,出现了基于语音、视觉等技术来实现远距离交互并进行自拍的新方法。例如Oppo N1可以识别用户的口令,如“茄子”等,达到控制快门拍照的功能。以及如百度魔拍、笑脸检测等利用视觉的方法识别用户脸部的语义来达到控制相机快门拍照的功能等。但此类方法往往只提供单一的快门控制功能,无法提供更多的参数调节功能。然而,自拍过程中用户往往还需要调整必要的功能参数,如感光度等^[6]。

本文提出利用基于视觉的手势识别方法在智能手机上实现新型的自拍交互技术。其优点在于,(1)直接对自拍相机捕获的实时图像进行图像处理来识别手势,不需要对现有设备进行物理改造。(2)手势可以表达丰富的语义,从而实现更多的控制功能。(3)利用手势交互使得自拍过程更加自然,交互效率更高。

在基于视觉的手势识别方法中采用了光流跟踪算法在图像序列上分析手掌的运动,然后利用基于时间的有限状态机来识别指定的手势。该方法具有较快的运算处理速度,可以在目前主流的智能手机等移动终端上运行,取得了较好的执行效率及识别准确度。交互的手势方面,本文主要提出了挥手及画圈两种手势。挥手是弯曲手臂抬起手掌并朝向前,再左右挥动的手势;画圈是弯曲手臂抬起手掌并朝向前顺/逆时针旋转的手势。两种手势分别用在控制相机快门及控制感光参数的功能上,如调节照片曝光度、白平衡等。

通过用户的测试实验,分别对手势识别的识别正确率,以及手势自拍的交互界面进行了测试和评估。实验表明所提出的手势识别方法平均识别率达到85%。用户满意度达到平均4.7分(满分为5分)的评价,相比于传统触屏加计时器的方法用户体验更佳。

2 相关研究

2.1 自拍交互方法

手托式自拍方法,该方法将相机托在手上,伸长手臂将镜头面向自己进行拍照。该方法由于其方便快捷性而被普遍使用,特别是在旅途中^[7]。基于此种自拍方法,出现的配备镜头与屏幕同一方向来方便自拍的相机,例如Casio TR500、Sony 香水瓶相机,以及可翻转镜头的Canon 70D单反相机等。然而,使用手托相机自拍的方法难以防止手抖,而容易使成像产生对焦不准或模糊。因为手臂长度的限制往往只能拍摄出近距离的大头自拍照,缺乏自拍的自由性和多样性^[4,8-9]。自拍杆的出现部分克服了人与相机成像距离的问题,但仍然难以避免手抖的问题^[4,10]。

计时器方法,该方法可以把相机固定放在某一位置,如三脚架上,然后设定一个5秒或10秒的计时器时间,等待相机快门自动触发而完成自拍。然而该方法往往需要用户预先在指定位置放置一个物品,如花瓶等,作为用户的替身来确定相机的对焦点^[6]。当用户设定好计时器后需要快速移动到相机前,移走替身物品,然后准备一个自拍姿势完成自拍。此种方法效率较低容易让人产生挫败感,如要反复确定对焦准确性及成像的构图结构,往往需要多次尝试后才能够得到一种满意的照片。如果需要调整镜头缩放的功能,用户很难在设定计时器的同时获得自己成像预览的实时反馈^[11]。

遥控器的方法,使用手持式遥控器可以在摆出自拍姿势的同时按下控制键完成拍照。它不需要用户接触相机,也可以保证相机具有正确的对焦。如果使用具有触屏功能的移动终端作为遥控器,如在智能手机上的相机控制应用软件Canon EOS Remote,可以提供更多的相机控制功能,如白平衡、曝光度、缩放,实时预览等。然而,使用遥控器自拍时往往需要在快门触发前隐藏手上的遥控器^[12]。遥控器会占用用户的手臂导致自拍姿势不自然。

语音控制的方法,如三星Galaxy S5相机等,利用语音识别的技术,用户通过喊出指定的口令,如“拍照”、“茄子”等来控制相机的功能。然而,由于语音识别技术的局限,往往需要在安静环境下或用户距离设备有较近的距离范围才可以达到较好的识别准确率。

基于视觉的方法,通过对相机捕获到即时图像进行图像处理,识别特定的信息标记、动作信息,甚至是手势,从而产生控制命令来与相机进行交互。此类技术有:如Sony Party Shot的笑脸快门、百度眨眼快门控制等,利用人脸识别技术识别用户的面部表情变化达到控制的功能^[9];还有如Casio的运动检测快门控制,当相机检测到图像中有大幅度的人体动作即可自动设定自拍计时器^[13]等。文献[14]提出了利用手势进行自拍相机镜头旋转控制的功能来完成图像构图的技术。

本文同样提出基于视觉的自拍控制方法。在智能手机上实现了挥手及画圈两种手势的交互界面,优化了手势识别效率,达到了实时控制相机快门、白平衡,曝光度等参数的功能。

2.2 手势识别方法

基于视觉的手势识别技术可分为静态的基于形态的识别、动态的基于运动的识别,以及混合静态和动态的识别方法等^[15]。

静态的手型识别有,如文[16-17]等,利用图像的颜色信息分隔出手的轮廓,然后与预先存储在数据库中的手轮廓模板进行匹配,识别出指定的手型。识别可实时还原手和手指的三维结构信息。被应用在增强现实,如HandyAR^[18]利用手掌与三维物体交互;以及空中手势三

维建模^[19]等交互界面中。此类方法的局限是需要较强的颜色信息源,通过控制环境光照或者穿戴指定颜色的手套等才能够使用,同时也需要较大计算量,限制了实际环境的应用。而在自拍的情况下需要考虑不同的环境光线变化,以及用户设定的镜头感光度改变图像的明暗等问题,静态的手型识别方法不是很适合。

动态的基于运动的识别有:如文[20-21]等,通过光流跟踪的方法,分析图像中局部或整体的运动信息来识别手势。被应用在游戏交互界面中,如文[22]。此类方法针对图像序列上相邻帧图像灰度图的差异进行计算分析,对光线的要求较低^[23],较适合于自拍的环境中进行手势的识别。然而,由于光流分析的运算效率较低,在移动平台上难以达到实时的识别效果。本文提出了限定识别区域的方法,利用有限的3×3排列的关键点利用光流跟踪方法识别。在移动平台设备上可达到22FPS的效率。

3 系统设计及实现

设计的自拍相机原型系统采用可旋转的具有1 300万像素摄像头的Oppo N1拍照手机,该手机可将摄像头旋转至屏幕前方方便自拍时用户观看屏幕上的预览画面。同时其5.9英寸的屏幕可使得用户远距离拍照时仍可以观看到屏幕中的画面。用户可以用支架等设备将设备摆放着桌子,或者挂在墙壁上,然后在相机前挥动手臂完成自拍。通过组合挥手及画圈手势可以完成调节照片曝光度、白平衡等参数控制的功能。系统的使用场景如图1所示。



图1 手势自拍系统

系统的实现过程中,主要考虑到两个方面的问题:(1)只利用两种手势来实现交互。提出利用日常生活中自然便捷的两种手势,如果提供更多的交互控制方法,但同时也会加大用户的学习负担,影响交互的目的。通过上下文组合两种手势的应用来达到丰富交互控制的功能。(2)考虑在移动平台上实现高效的基于视觉的手势识别方法。主要利用限定手势识别区域来减少图像处理负担,以及利用基于时间的有限状态机方法来识别手势的方法。

3.1 手势交互方法

首先,利用人脸检测确定人脸的区域及宽度 W ,然

后设定在人脸右侧水平距离一个人脸宽度的位置开始的一个宽度为 $1.1W$ 的矩形区域为手势交互与识别的区域,宽度为 $1.2W$,如图2所示。该区域经过设计人员与3位普通用户的实际体验估算得出,较为自然舒服的右手手势运动区域。手势识别过程中,在该区域中平均分布9个关键点作为光流跟踪的关键位置,从图像序列上每一帧所检测到的运动信息可以估算出手势的运动信息。图中紫色圆圈为手势识别区域,绿色点为光流跟踪关键点。

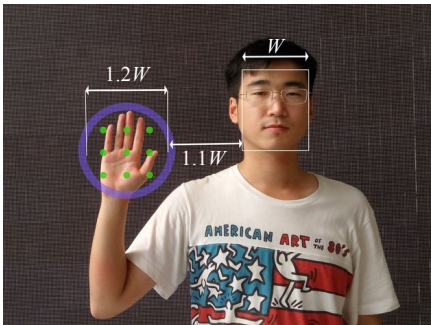


图2 人脸位置以及手势识别区域

3.1.1 运动识别方法

在识别区域确定后,将该区域的图像缩放一个固定大小的尺寸(本文为100像素×100像素)来进行图像处理。这样可以保证较好的运算效率,同时在该尺寸可以达到较好的运动检测效果来识别手势。

在运动检测的算法上采用了Lucas-Kanade^[24]稀疏光流算法,相比于Horn-Schunck^[25]等稠密光流算法,其优势在于具有更快的运算效率。Lucas-Kanade光流分析假设两幅连续的视频图像相同信息的偏移是细微的。对指定的一个像素点 P , 以及其四周的像素点为一个区域窗口,在连续的两帧图像上光流分析 P 点的位置即在这个区域窗口中。通过计算多个像素点的光流,以及利用多级金字塔图像可以精确地确定一个图像区域内的运动信息,如运动的方向和速度。

基于以上原理,在识别区域上设置了3行3列的关键点来覆盖识别区域并计算各点的光流,如图2所示。因为光流分析易受图像噪声等影响而产生错误,所以在每一帧上剔除了关键点中位移过大的计算结果。将所有指定关键点所检测到运动方向及速度信息进行加和平均作为当前帧上检测到手势运动信息。根据连续多帧上的手势运动信息可利用有限状态机的方法进行手势识别。目前在每一帧上光流计算效率约为45.5 ms,即22FPS。

3.1.2 挥手手势

挥手手势是用户将手掌放在识别区域中,左右来回挥动的手势。手掌在挥动的过程中,在识别区域中的每一个光流跟踪关键点所计算出的结果进行加和平均来估算出当前帧上手掌的运动方向。

在时间序列上利用有限状态机的方法,将手掌运动状态分为停止、左、右三个状态。当没有检测到运动时,

3.3 系统的实现

手机采用了可206°自由旋转、具有1 300万像素摄像头的Oppo N1拍照手机。可以提供1 920×1 080分辨率的30FPS摄像头图像实时预览,拥有5.9英寸的屏幕,以及1.7 GHz的高通四核处理器。

软件的实现使用了JNI(Java Native Interface)技术,可以让Java代码调用底层的C程序,从而加快图像处理的运算速度。光流跟踪算法使用了OpenCV中的Lucas-Kanade^[24]方法。为使图像处理的运算均衡,手势识别的算法与图像及UI界面的显示程序分别运行在两个不同的线程上,最终程序运行的评价效率达到22FPS。

4 用户实验

实验分别对手势识别的正确率,以及手势自拍的交互界面的用户满意度进行了测试和评估。

共邀请了12名大学生作为受验者,女生5名,男7名,平均年龄为22.7岁(±1.9)。

实验在白天室内环境下进行,受验者被邀请坐在自拍相机前约1米处。首先受验者接受约5分钟左右的系统功能介绍以及2分钟的使用体验。然后,受验者根据实验员的指示分别进行挥手及画圈手势各10次。同时,实验员记录手势识别的正确性,以及完成手势操作所需的时间。

实验共统计了两种手势各120次操作,其中挥手手势成功识别了106次,平均识别率为88%;画圈手势成功识别了101次,平均识别率为83%。受验者平均花费约1.8 s(±0.2)来完成一次挥手手势;花费约为2.1 s(±0.3)来完成一次画圈手势。

在手势识别实验之后,受验者被邀请自由尝试使用手势与自拍相机进行交互操作。受验者可以利用画圈手势调节白平衡、曝光度等参数,然后利用挥手手势完成自拍。同时受验者再使用相机原有的触屏界面加计时的方法完成参数调节及自拍操作。在受验者完全理解并掌握了手势自拍界面的操作方法以及传统的触屏加计时的方法后,被要求完成一份系统使用体验的满意度调查问卷。问卷采用5分制李克特量表(Likert scale)的方法进行量化评估。

表1呈现了最终的评估结果。所设计的手势自拍系统得到了大多数用户的认可,满意度达4.7(±0.4),较好地触屏加计时的方法4.1(±0.6)。方差分析结果 $F(1, 22)=8.0$, $p=0.01$,表明二者具有显著差异。所

提出的两种手势被认为是自然的交互方法,其中挥手手势用户满意度达4.6(±0.1),画圈手势用户满意度达4.2(±0.3)。

5 讨论

挥手手势得达到了88%的识别率及较高的用户满意度。该手势用于控制相机快门拍照的功能较为合适。本文提出的预设识别区域的方法有效的避免了由于背景或身体的移动所可能产生的误识别,同时也提高了识别的运算效率,相比于文献[26]具有更好的实际应用价值。

画圈手势达到了83%的识别率,其出错主要是用户在手势操作过程中将手掌移出识别区域所导致。用户在刚开始操作时往往偏向于较大幅度的画圈动作而使得手掌移出识别区域,导致识别不能持续进行,在实验人员的引导下可熟悉正确操作的模式。该手势可以识别手掌移动的圆周角度以及画圈的顺/逆时针方向,设计中可以将这些数值映射到参数调整的控制界面中。

手势自拍交互最主要是解放了用户的双手,有用户评论到:“手势自拍自由度增加了,可以方便尝试不同的自拍姿势”。也有用户评价说:“手掌运动幅度有些大,增加了操作负担,如果能够识别手指的运动,如手掌张开到抓取的动作就更好了”。

系统的图形界面有效地提醒了用户操作及系统的状态,在用户接受了实验人员的介绍后可以快速理解界面的意义及操作的方法。由于手势识别难以达到完美,有效的即时图形反馈很重要。系统界面利用不同的颜色区分不同的系统状态给予了用户较好的直观感受。“界面中有颜色的区域很吸引人伸手去触控”有用户评价道。

在手势的识别方面,需要用户运用合适的手势运动速度,如果运动太快,图像会产生运动模糊使得光流跟踪错误率较高。当前摄像头提供30 frame/s的帧速,但识别效率为22FPS,今后具有更高执行效率的设备将改善识别的精度。

系统的主要局限是只有两种手势进行操作,对于菜单项的选择必须要用画圈手势一个一个的切换来选择,交互效率较低。将来考虑加入更多的手势,以及研究面部运动,如点头、摇头等,来辅助交互的方法。

6 结束语

本文提出了基于视觉的运动手势自拍交互界面,并在智能手机上加以实现,使得自拍过程用户只要挥一挥手臂就可以完成。主要提出挥手及画圈两种手势交互方法,通过组合两种手势可以实现相机的白平衡、曝光度等参数的控制。图形界面的设计中考虑了利用颜色实时给予用户反馈,提高了用户体验。所采用的区域运

表1 手势自拍用户满意度

	用户满意度
挥手	4.7(±0.4)
画圈	4.2(±0.3)
手势界面	4.7(±0.4)
触屏+计时器	4.1(±0.6)

动手势识别方法取得了85%的识别率。在对手势自拍交互界面的用户实验中,用户给予了手势界面4.7分评价,好于传统的触屏加计时器的方法。在今后的研究中将对多种不同的自拍交互方法,如遥控、语音、表情识别等,开展用户可用性评估实验,发现不同技术的优劣以及不同应用情景的可用性问题等。

参考文献:

- [1] Saltz J. Art at arm's length: A history of the selfie[J]. New York Magazine, 2014, 47(2): 71-75.
- [2] Barry C T, Doucette H, Loflin D C, et al. Let me take a selfie: associations between self-photography, narcissism, and self-Esteem[J]. Psychology of Popular Media Culture, 2015, 1(1): 1-13.
- [3] Chua T H H, Chang L. Follow me and like my beautiful selfies: Singapore teenage girls' engagement in self-presentation and peer comparison on social media[J]. Computers in Human Behavior, 2016, 55: 190-197.
- [4] Lee T P, Kosasih W. Exploring the distance, location, and style of "selfies" with the self-pano mobile application[C]// New Contributions in Information Systems and Technologies. [S.l.]: Springer, 2015: 875-884.
- [5] Chu S, Tanaka J. Development of a head gesture interface for a self-portrait camera[J]. The Transactions of Human Interface Society, 2013, 15(3): 247-259.
- [6] Chu S, Tanaka J. Interacting with a self-portrait camera using motion-based hand gestures[C]// Proceedings of the 11th Asia Pacific Conference on Computer Human Interaction ACM, India, 2013: 93-101.
- [7] Lee T P, Kabbany N, Jun C K. Self-portrait images for mobile application[C]// Advances in Information Systems and Technologies. [S.l.]: Springer, 2013: 827-838.
- [8] Samani H. Take a selfie with brand; effective approach or not?[C]// 6th International Conference of Cognitive Science, 2015.
- [9] Yeh M C, Lin H W. Taking good selfies on your phone[C]// Proceedings of the ACM International Conference on Multimedia. ACM, 2014: 751-752.
- [10] Lo C C, Huang S P, Ren Y, et al. Design of a self-portrait application with sensor-assisted guiding for smart devices[C]// 2013 First International Black Sea Conference on Communications and Networking (BlackSeaCom). IEEE, 2013: 99-101.
- [11] Chu S, Tanaka J. Head nod and shake gesture interface for a self-portrait camera[C]// ACHI 2012, The Fifth International Conference on Advances in Computer-Human Interactions, Spain, 2012: 112-117.
- [12] Chu S, Tanaka J. Design of a motion-based gestural menu-selection interface for a self-portrait camera[J]. Personal and Ubiquitous Computing, 2015, 19(2): 415-424.
- [13] Garcia C, Bazin J C, Lancelle M, et al. Automatic jumping photos on smartphones[C]// 2014 IEEE International Conference on Image Processing (ICIP). IEEE, 2014: 2522-2526.
- [14] Chu S, Tanaka J. Hand gesture for taking self portrait[C]// 14th International Conference on Human-Computer Interaction, Part II. [S.l.]: Springer, 2011: 238-247.
- [15] Wachs J P, Kölsch M, Stern H, et al. Vision-based hand-gesture applications[J]. Communications of the ACM, 2011, 54(2): 60-71.
- [16] Wang R Y, Popović J. Real-time hand-tracking with a color glove[J]. ACM Transactions on Graphics (TOG), 2009, 28(3).
- [17] Hoshino K, Tomida M. Quick and accurate estimation of human 3D hand posture[J]. Intelligent Service Robotics, 2010, 3(1): 11-19.
- [18] Lee T, Höllerer T. Handy AR: Markerless inspection of augmented reality objects using fingertip tracking[C]// 2007 11th IEEE International Symposium on Wearable Computers. IEEE, 2007: 83-90.
- [19] Wang R, Paris S, Popović J, et al. 6D hands: markerless hand-tracking for computer aided design[C]// Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology, Santa Barbara, California, USA 2011. ACM, 2011: 549-558.
- [20] Bayazit M, Couture-beil A, Mori G. Real-time motion-based gesture recognition using the GPU[C]// MVA, 2009: 9-12.
- [21] Zivkovic Z. Optical-flow-driven gadgets for gaming user interface[C]// Entertainment Computing-ICEC. [S.l.]: Springer, 2004: 90-100.
- [22] Eisenstein J, Mackay W E. Interacting with communication appliances: an evaluation of two computer vision-based selection techniques[C]// Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM, 2006: 1111-1114.
- [23] Molnár J, Chetverikov D, Fazekas S. Illumination-robust variational optical flow using cross-correlation[J]. Computer Vision and Image Understanding, 2010, 114(10): 1104-1114.
- [24] Viola P, Jones M. Rapid object detection using a boosted cascade of simple features[J]. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2001, 1: 511-518.
- [25] Horn B, Schunck B. Determining optical flow[J]. Artificial Intelligence, 1981, 17(1/3): 185-203.
- [26] Irie K, Umeda K. Detection of waving hands from images using time series of intensity values[J]. Journal-robotics Society of Japan, 2003, 21(8): 95-103.