

一种实时手势识别应用开发框架

张长青, 宋旭

(安阳师范学院 计算机与信息工程学院, 河南 安阳 455002)

摘要: 提出并设计了一种简化的、基于自然手势的实时识别的软件开发框架, 通过手势样本采集和消息映射机制, 开发人员可以自定义手势含义, 将手势含义封装成消息发送给具体应用。降低了开发实时手势识别软件的成本, 屏蔽了手势识别技术本身的复杂性。

关键词: 手势识别; 实时; 消息映射; 样本采集; 框架

中图分类号: TP311.1

文献标识码: A

文章编号: 1674-7720(2011)09-0009-03

A real-time gesture recognition based applications development framework

Zhang Changqing, Song Xu

(School of Computer & Information Engineering, Anyang Normal University, Anyang 455002, China)

Abstract: A simplified framework based on gesture recognition is proposed. Using sample collection and the gestures-message mapping mechanism, developers can define gesture meaning, and encapsulates the gesture meaning into message, then sent the message to the specific application. The development costs of real-time gesture recognition based application become slow, and the complexity of hand gesture recognition itself is shielded.

Key words: gesture recognition; real-time; message mapping; sample collection; framework

手势是一种自然直观的人机交流方式, 它符合人的生活习惯。手势识别研究的主要目的是将手势引入到人机接口 HCI(Human Computer Interface)中^[1]。

手势识别技术最初是基于数据手套, 人需要戴上有传感器的数据手套^[1], 通过传感器系统感知手的位置和关节的相对位置, 这种方法由于需要在使用者手上带上固定的物理装置, 影响了系统使用的自然性和友好性。另外, 物理装置的价格以及容易磨损的特性都严重影响了系统的整体使用, 尤其严重影响了用户体验。后来, 人们逐渐把手势识别研究的焦点转移到了自然状态的手上, 这种识别方式主要基于计算机视觉的识别。基于视觉的手势识别解除了物理装置的限制, 使人机交互更加自然、直接和简洁^[2]。

基于手势识别的应用方兴未艾, 目前手势识别的产品并不多见, 然而可以预见, 手势识别等更为丰富和自然的人机交互方式是未来研究和应用的一个趋势, 然而, 由于手势识别的专业性和复杂性, 阻碍了基于手势识别的系统的研究和开发。基于此, 本文提出并设计了一种通用的手势识别应用软件开发框架, 降低了软件开

发成本, 将主要精力集中于应用的业务本身, 而非对手势识别技术本身的研究。

1 应用框架

自然手势识别方法大致可以分为两大类: 基于三维手势建模的方法和基于表观建模的方法。基于三维手势建模的方法需要估算的参数值多, 比较复杂。而基于表观的方法直接从图像中提取相应的特征进行识别, 不需要恢复手势在三维空间中的信息^[3], 但可识别的手势种类较少。文献[4]的方法就是通过视频图像序列的前若干帧来建立静态手势框架, 文献[5]提出了一种基于2D的变形手势的实时跟踪方法。但是考虑到在给定软件的人机交互中, 有限的手势种类即可满足要求, 重点要考虑自然性和实时性, 因此本文采用基于表观的方法。本文将手势分为静态手势和动态手势, 定义如下:

定义1 静态手势: 单个手势, 静态手势是手势模式数据库的基本组成元素, 也是构成动态手势的基本要素。

定义2 动态手势: 由包含时间信息的一个或多个静态手势组成。在定义静态手势后, 可通过分析视频信

息定义动态手势。

本文提出的框架分为手势定义部分和手势识别部分,基本架构如图1所示。该框架又可分为几个具体部分:

(1)摄像设备:基于计算机视觉的手势识别,摄像设备是必要的硬件装置。通过摄像设备捕捉人动态信息,将视频文件传输到相应处理模块。

(2)视频分帧模块:负责初步处理视频信息。视频文件传输到此模块后,该模块将从连续的视频文件中提取图像序列。

(3)手势定义模块:负责分析该时间段内所形成的手势,如果该手势尚未定义,则提醒用户进行手势含义定义,并存储手势含义信息。

(4)手势识别模块:负责手势的识别和含义的确认,并将识别出的手势含义以消息的形式封装。

(5)手势模式数据库:数据库中存储了已定义手势模式及其含义。

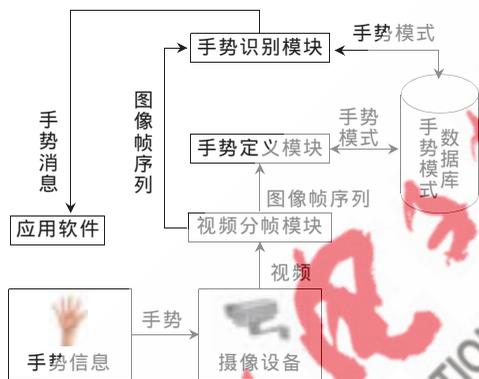


图1 实时手势识别框架

2 实时识别方法

2.1 静态手势定义方法

静态手势为单一手势,由一幅含手势信息的静态图像即可获得。关于静态手势的识别方法较多,文献[1]~[8]都提出了各自的识别方法。手势分析完成后,将会形成新的手势,对于每一种手势,用户根据应用需要,按照规定格式定义自己的手势消息。

2.2 动态手势定义方法

动态手势由静态手势序列组合而成。动态手势的定义步骤如下:

步骤1 摄像设备捕获用户的手形变化,形成一段视频信息;

步骤2 将视频(或连续的图像帧)作为输入,若为视频文件,系统将对 T_0-T_n 时间段内的视频转化为连续的图像帧作为输入;

步骤3 用户设置相应的视频待分析时间范围,系统将根据视频和时间信息进行分析,尝试识别出手势信息;

步骤4 如果识别成功,即形成一个完整手势模式,然后将识别完成的手势展示给用户,用户确认后,即形成新的动态手势模式。

2.3 实时手势识别策略

手势模式完成即手势模式数据库初始化完毕,之后,可进行手势的识别。从分割出的连续视频帧中识别出手势信息存在的问题,即手势的时间效用范围。采用滑动窗口策略,假设视频分帧率为 f ,手势有效时间长度为 t ,滑动窗口大小即一个手势所包含的连续图像数目为 n 。则有如下关系:

$$n = f \times t$$

滑动窗口策略识别步骤如下:

步骤1 初始化滑动窗口状态,起始窗口状态如图2(a)所示,包含第1幅视频帧至第 n 幅(此处假设 $n=3$)。

步骤2 此时手势分析模块尝试根据已定义手势模式数据库分析手势。若存在匹配手势,则转到步骤4,否则转入步骤3;

步骤3 窗口向后滑动1步(帧),如图2(b),转到步骤2;

步骤4 窗口向后滑动 n 步(帧),如图2(c),转到步骤2。



图2 滑动窗口

滑动窗口的大小对系统将会产生重要影响,滑动窗口越大,分析的图像序列越多,则信息相对越完整,识别准确程度越高。但同时,分析时间越长,实时效果越差。

3 消息机制

基于本框架的应用软件分为框架层和应用层,如图3所示。

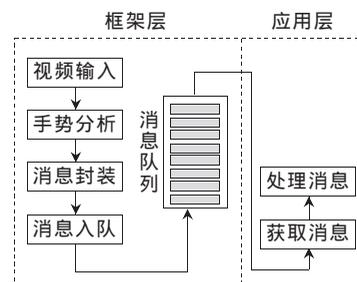


图3 消息机制框架

框架层接收视频输入后进行手势分析,将分析所得手势转化为手势消息,并将消息放入框架所维护的消息

队列;应用层负责从消息队列中取出消息,并进行相应处理。

系统实现示例代码主要包括三个部分:消息结构体定义、消息封装函数、消息处理函数。

(1) 手势消息结构体

```
typedef struct GESTRUE_STRUCT
{
    unsigned int iMessageType;    //手势消息类型
    unsigned long iTime;         //消息发生时间
    unsigned long lParam;        //消息保留参数
    unsigned long wParam;        //消息保留参数
}
```

(2) 消息封装函数

如用户定义了手势 GM_MOUSEMOVE,该手势表示鼠标指针的移动,对应消息封装函数如下:

```
void OnCreateMOUSEMOVEMsg()
{
    GM_MESSAGE msg;
    msg.iMessageType=GM_MOUSEMOVE;
    msg.iTime=GetSystemTime();
    msg.lParam=GetCurPos();
```

```
        //获取当前光标位置,由框架提供
    msg.wParam=0;        //该参数在此消息中暂不使用
```

(3) 应用层消息处理代码

```
switch(iMessageType)
case : GM_MOUSEMOVE :
    iMouseX=HIWORD(lParam);
        //从消息保留参数中获取手势位置 X 坐标
    iMouseY=LOWORD(lParam);
        //从消息保留参数中获取手势位置 Y 坐标
    mouse_event(MOUSEEVENTF_MOVE,
        //调用 windows API 函数,使鼠标移动
        iMouseX,
        iMouseY,
        NULL,
        NULL);
break;
```

.....

本文在介绍手势识别技术的基础上,提出并设计了了手势识别应用的一种通用框架,使基于手势识别的软件开发框架化,将手势识别软件的开发重点转移到了应用业务本身,降低了手势识别软件开发的难度。

由于手势变化的复杂性,手势识别技术的准确性和实时性还是制约手势识别应用发展一个瓶颈。因此,今后将进一步研究和关注手势图像识别算法,继续改进手势识别的实时性和准确性,提高框架的适应性和鲁棒性。

参考文献

- [1] 朱继玉,王西颖,王威信,等.基于结构分析的手势识别[J].计算机学报,2006,29(12):2130.
- [2] 肖志勇,秦华标.基于视线跟踪和手势识别的人机交互[J].计算机工程,2009,35(15):198.
- [3] 任海兵,祝远新,徐光祐,等.连续动态手势的时空表观建模及识别[J].计算机学报,2000,23(8):824.
- [4] NING H Z, TAN T N, WANG L, et al. People tracking based on motion model and motion constrains with automatic initialization[J]. Pattern Recognition, 2004, 37(7): 1423.
- [5] 王西颖,张习文,戴国忠.一种面向实时交互的变形手势跟踪方法[J].软件学报,2007,18(10):2423.
- [6] 章毓晋.基于内容的视觉信息检索[M].北京:科学出版社,2003.
- [7] 汤勇,顾宏斌,周来.交互系统中实时手势分割及指尖检测方法[J].光电工程,2010,37(7):145.
- [8] 刘军,田国会,李荣宽,等.智能空间下基于手势识别的人机交互[J].北京联合大学学报(自然科学版),2010,26(6):14.

(收稿日期:2010-12-13)

作者简介:

张长青,男,1982年生,硕士,讲师,主要研究方向:软件工程,图像处理。

宋旭,男,1980年生,硕士,讲师,主要研究方向:Web应用,图像处理。