

基于 LBFSGS-EKF 算法的三维空间目标跟踪研究

雷安琪 冯玉田 唐子成

(上海大学通信与信息工程学院 上海 200444)

摘要: 人机交互中,人手与机器的交互是最通用的方式,因此手势交互研究是当今人机交互研究的重点之一。本文以 Leap Motion 为依据,对三维空间目标检测和跟踪进行研究,提出了一种限定内存扩展卡尔曼滤波算法(LBFSGS-EKF)。该算法在基于降低噪声所提出的 EKF 方法上,通过 LBFSGS 算法求最优解的方式来代替 EKF 算法在每次迭代中求逆 Hessian 矩阵而造成内存消耗,计算速率下降,导致实时性差的问题,从而形成一种新的目标跟踪算法。仿真结果表明,新算法用于手势识别时,可以降低误差、提高目标跟踪的精度。

关键词: Leap Motion ;扩展卡尔曼滤波 EKF; LBFSGS; 三维空间

中图分类号: TP391.9; TN911.4 **文献标识码:** A **国家标准学科分类代码:** 520.2040

Research on target tracking based on limited-memory BFGS and extended kalman filter in 3D-space

Lei Anqi Feng Yutian Tang Zicheng

(School of Communication and Information Engineering, Shanghai University, Shanghai 200444, China)

Abstract: In human-computer interaction, the interaction between human-hands and machine is the most common way, so gesture interaction research is one of the focuses of human-computer interaction research today. Based on Leap Motion, this paper studies the three-dimensional space object detection and tracking, and proposes a new algorithm-LBFSGS-EKF algorithm. The algorithm is based on the EKF method proposed by reducing noise, through the method of the LBFSGS algorithm is used to replace the Hessian matrix in each iteration instead of the EKF algorithm, which consumes memory and reduces the computational rate, resulting in poor real-time, thus forming a new target tracking algorithm. The simulation results show that by this new algorithm LBFSGS-EKF for gesture recognition, the error can be reduced and the real-time performance of target tracking can be improved.

Keywords: Leap Motion; extended kalman filter(EKF); limited-memory BFGS(L-BFGS); three-dimensional space

0 引言

近年来,随着计算机视觉和机器学习等相关学科的发展,人机交互技术(human computer interaction)已经成为人们日常生活中不可或缺的重要组成部分。其中,手势是人机交互中最常用的交互方式,基于手势的人机交互相比于传统交互方式具有更加简洁、自然、灵活的特点。因此,如何不借助中间媒介直接将人手作为计算机的输入设备来操作设备,已成为人机交互研究的重点^[1]。

人机交互中,计算机要能对人手的配置参数做到较准确的测量。设备的操作性能依赖于对手手位置的精准检测和跟踪。目前,手势检测和跟踪主要包括基于数据手套^[2]和基于视觉(如摄像机)的两类研究方向^[3]。

手部运动通常发生在三维空间。传统的二维空间只能

获取二维信息,忽略了Z轴方向的运动轨迹,这就导致实际手势检测和跟踪准确度不高。近年来,随着科技的快速发展,能直接获取三维空间信息的设备越来越多,而且识别的准确度、实时性都有了很大的提升。例如 TOF 相机与 Microsoft 公司推出的 Kinect 设备、Leap 公司推出的体感设备 Leap Motion。相应设备的研究也不断涌现。如基于 Kinect 设备,孟明等人提出人体运动检测算法^[4]或文献^[3]提出的基于卡尔曼滤波的三维手势跟踪。卡尔曼滤波主要用于线性环境中,然而实际操作环境下,目标运动过程中受到噪声的影响下往往是非线性情况,并且非线性因素是不能忽略的,所以单纯以卡尔曼滤波进行目标跟踪必然会因为周围噪声^[5]的干扰而导致目标跟踪精确度极差甚至发散;基于 Leap Motion 设备,潘佳佳等提出的三维自由手势操作^[6],因为相应的变量较少,计算量也会降低,对内存消

耗也少,所以能够很好的对静态手势做出识别,但对动态手势跟踪上准确度欠佳,因为在三维空间为了更加准确获取运动目标的信息,有一个较高的识别率,通常采取多维参数来提取手势特征。但参数维度越高对处理器的负载就越重,导致计算速率下降,最终影响手势检测和跟踪的实时性^[5]。

本文中提出的三维空间目标跟踪采用的是 Leap Motion 三维数据跟踪体感设备。考虑目标跟踪环境的非线性问题采用扩展卡尔曼滤波 EKF 来处理噪声提高目标检测的准确度。同时,为了降低运动目标在三维空间运算减小消耗内存方面,采用限定内存的 BFGS 方法能够有效的降低内存,提高运算速度。因此,本文提出一种结合扩展卡尔曼滤波和限定内存 BFGS 算法的新方法,新算法继承了上述两个方法的优点,可以有效地消减噪声和减少内存消耗保证实时性。

1 扩展卡尔曼滤波目标跟踪

1.1 Leap Motion 介绍

Leap Motion 利用双目红外成像的原理重建出手的三维空间运动信息,能检测并跟踪手、手指和类似手指的工具,并获取它们的位置、方向和部分手势信息^[7]。Leap Motion 的坐标系统采用右手笛卡尔坐标系。坐标原点在 Leap Motion 控制器中心, x 轴和 z 轴在器件的水平面上, x 轴与设备的长边平行, y 轴与设备垂直,其正方向朝上(与计算机图形学中的坐标系相反), z 轴垂直于屏幕,距离计算机屏幕越远,其值正向增加,所有坐标均以 mm 为单位,坐标系如图 1 所示,下面内容将基于该坐标系建立手部模型进行算法仿真分析。

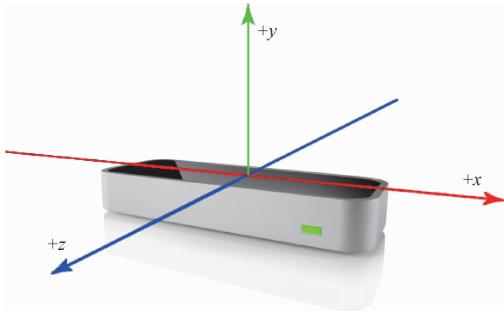


图 1 Leap Motion 的右手坐标系

1.2 卡尔曼滤波

在许多工程实践中,往往不能直接得到所需要的状态变量的真实值。例如,在本文中基于 Leap Motion 设备在三维空间去探测跟踪目标的时候,根据 Leap Motion 内置的红外立体摄像头的反射波等信息能获取目标跟踪数据,但是摄像头在探测过程中会因为复杂的背景、光线的强度等产生随机干扰的问题,这就导致在观测的得到的信号中往往夹杂有随机噪声。要从夹杂有随机噪声的观测信号中

分离出手势的运动状态量,要准确地得到所需的状态变量是不可能的,只能根据观测信号来估计或预测这些状态变量。卡尔曼滤波(kalman filter, KF)是一种能有效降低噪声影响的利器。它可以根据系统的历史测量值,建立最大化后验概率的系统状态模型,对下一状态做最优估计,有限地解决状态空间模型的估计与预测。其基本思想是以系统的状态空间模型为基础,利用前一时刻的状态估计值和现在时刻的状态测量值来更新对状态变量的估计,求解出现在时刻的状态估计值^[8]。

考虑离散的线性随机差分方程的耦合系统由式(1)、(2)给出:

$$x_k = M_k x_{k-1} + W_k \quad (1)$$

$$y_k = K_k x_k + V_k \quad (2)$$

在式(1)中, x_k 表示系统在 k 时刻 $n \times 1$ 状态; M_k 是 $n \times n$ 的状态转移矩阵; W_k 是被称作预测误差的 $n \times 1$ 的随机向量。在等式(2)中, y_k 表示为 $m \times 1$ 的观测数据; K_k 表示为 $m \times n$ 的线性观测矩阵; V_k 是被称作观测误差的 $m \times 1$ 的随机向量。其中 W_k 和 V_k 被认为是均值为 0、方差阵各为 Q 和 R 的不相关白噪声,并且对称正定义协方差矩阵各为 C_{W_k} 和 C_{V_k} 。由此,便可以计算出时刻 k 状态的先验估计 \hat{x}_k^- 和 \hat{x}_k^- 的协方差矩阵 C_k^- 。接着就可以更新 Kalman 滤波器增益、 x_k 和 C_k ^[9]。下面给出卡尔曼滤波过程。

卡尔曼滤波(kalman filter, KF)过程^[8]:

1)选择初始猜测值 x_0 和协方差矩阵 C_0 ,并设置 $k=0$

2)计算状态一步预测 \hat{x}_k^- 和一步预测协方差阵 C_k^-

a. 计算 $\hat{x}_k^- = M_k \hat{x}_{k-1}$

b. 计算 $C_k^- = M_k C_{k-1} M_k^T + C_{W_k}$

3)计算卡尔曼滤波增益矩阵 G_k ,卡尔曼滤波估计值 x_k 和其协方差值 C_k

a. 计算卡尔曼增益矩阵 $G_k = C_k^- K_k^T (K_k C_k^- K_k^T + C_{V_k})^{-1}$

b. 计算卡尔曼估计 $x_k = \hat{x}_k^- + G_k (y_k - K_k \hat{x}_k^-)$

c. 计算协方差估计 $C_k = (I - G_k K_k) C_k^-$

1.3 扩展卡尔曼滤波

由于 KF 主要用于线性系统中,而实际系统总是存在不同程度的非线性问题。如飞机的飞行状态、导弹的制导系统、卫星导航系统等,其中的非线性因素都是不可忽略的,必须建立适用于非线性系统的滤波算法。对于非线性系统滤波问题,常用的方法是利用线性化技巧将其转为一个近似的线性滤波问题,应用最广泛的就是扩展卡尔曼滤波(extended kalman filter, EKF)^[10]。

EKF 的离散非线性系统的动态方程如式(3)、(4)所示,用于取代 KF 的式(1)和(2)。

$$x_k = M(x_{k-1}) + W_k \quad (3)$$

$$y_k = K(x_k) + V_k \quad (4)$$

在式(3)和(4)中 M 和 K 是非线性函数。EKF 的获取可以通过修改上面 KF 过程:在 2) a 中,用 $X_k^- = M(X_{k-1})$,并定义

$$M_k = \frac{\partial M(x_{k-1})}{\partial x} K_k = \frac{\partial K(x_k)}{\partial x} \quad (5)$$

通过在目标跟踪中引用扩展卡尔曼滤波 EKF,既可以在最大程度上滤出随机噪声对目标跟踪的干扰,又可以对目标下一步目标位置进行预测。

2 限定内存扩展卡尔曼滤波

本文在对三维空间手势跟踪检测过程中,不仅要考虑手在三维空间的位置,还要考虑手势在运动过程中的速度、加速度以及噪声对其产生的影响,这就导致在实际的运算中,不是一个三维参数而是高维参数的运算。而我们使用 EKF 作为目标跟踪的算法中,其构建的状态方程和测量方程中必然涉及高维矩阵存储,并且在求逆过程中造成内存不足,这将是制约状态空间算法可行性的一大关键因素^[11]。所以,本文采用拟牛顿法^[12]中限定内存 BFGS 来解决存储和内存不足的问题,这对于 Leap Motion 这样小巧的设备是非常有帮助的。该方法保持了 Hessian 矩阵简单且紧凑的近似:代替存储完全密集的 $n \times n$ 近似值,仅仅保存了隐式表示近似值的长度为 n 的几个向量。而且, LBFGS^[9,13] (limited-memory broyden fletcher Ggoldfarb shanno)优化方法也被用来获取状态估计协方差矩阵和预测误差协方差矩阵的低存储近似。这进一步证明了 LBFGS 方法在与高维 EKF 运算结合中降低运算复杂度、减小占据内存的优势所在。

2.1 限定内存 BFGS 算法

首先,给出一个 LBFGS 求解最小化问题的一个通用表达式

$$f(x) = \frac{1}{2} \langle Ax, x \rangle - \langle b, x \rangle \quad (6)$$

$$Ax^* = b \Leftrightarrow f(x^*) \Leftrightarrow \min f(x) \quad (7)$$

其中 A 是一个 $n \times n$ 对称正定矩阵, b 是一个 $n \times 1$ 向量。应用关于 x 的最小化 $f(x)$ 问题的 LBFGS 方法能够被用来获取 $f(x)$ 的最小化估计,以及 A 和 A^{-1} 低存储近似,这两者都将用于我们的方法之中。在详细介绍 LBFGS 算法之前,必须对拟牛顿中的 BFGS 算法有一定的了解,详细过程可以参考^[14],其中 BFGS 方法中每一步都有如下形式:

$$x_{k+1} = x_k - \alpha_k H_k \nabla f_k \quad (8)$$

其中, α_k 是步长, H_k 通过式(8)在每次迭代中更新

$$H_{k+1} = V_k^T H_k V_k + \rho_k s_k s_k^T \quad (9)$$

其中,

$$\rho_k = \frac{1}{y_k^T s_k}, V_k = I - \rho_k y_k s_k^T \quad (10)$$

$$s_k = x_{k+1} - x_k, y_k = \nabla f_{k+1} - \nabla f_k \quad (11)$$

由于逆 Hessian 近似通常是密集型的,在高维参数下,存储操作它都是耗时和占内存的过程,在实际工程中是不允许的。为了解决这个问题,通过存储式(8)、(9)中使用一定数量的矢量对 $\{s_i, y_i\}$ (例如, m) 来隐式地存储 H_k 修改

版本。由此 $H_k \nabla f_k$ 的获得便可通过矢量对 $\{s_i, y_i\}$ 、 ∇f_k 以及内部产物序列求解出来。在每一次循环迭代中,最老矢量对 $\{s_i, y_i\}$ 被从式(11)获得的新的矢量对 $\{s_k, y_k\}$ 所取代。矢量对的集合来自最近的曲率迭代信息。实践表明 m 取值在(3~20)在为理想,参考下表1。由此可知,实际工程中, m 的取值越少越好。

表1 不同存储器在LBFGS下CPU速率比较

Problem n	LBFG-S	LBFG-S	LBFG-S	LBFG-S
	$m=3$	$m=5$	$m=17$	$m=29$
	nfg time	nfg time	nfg time	nfg time
DIXMAA	146	13.4	120	125
NL 1500	16.5	17.4	28.2	44.4
EIGENAL	821	569	363	168
H 110	21.5	15.7	16.2	12.5
FREUROT	>999	>999	69	38
H 1000	—	—	8.1	6.3
TRIDIA	876	611	531	462
1000	46.6	41.4	84.6	127.1

在第 k 次迭代中,当前迭代值是 x_k ,同时矢量对 $\{s_i, y_i\}$ 由 $i = k - m, \dots, k - 1$ 给出。必须选择一些初始 Hessian 近似 H_k^0 重复利用式(8)发现 LBFGS 满足下式:

$$H_{k+1} = (V_k^T \dots V_{k-m+1}^T) H_k^0 (V_{k-m+1} \dots V_{k-1} V_k) + (V_k^T \dots V_{k-m+2}^T) \rho_0 s_0 s_0^T (V_{k-m+1} \dots V_{k-1} V_k) + (V_k^T \dots V_{k-m+3}^T) \rho_1 s_1 s_1^T (V_{k-m+3} \dots V_{k-1} V_k) + \dots + V_k^T \rho_{k-1} s_{k-1} s_{k-1}^T V_k + \rho_k s_k s_k^T \quad (12)$$

由式(12)就可以计算出 $H_k \nabla f_k$ 获取搜索方向,但是并不是直接通过式(12)来计算获取搜索方向的,而是通过一种快速算法求解 $H_k \nabla f_k$ ^[15]。

由此得出 LBFGS 算法的求解过程如下:

1) 设置初始猜测值 x_0 和,以及 $m > 0$ 的值,并置 $k=0$

2) 选择初始 H_0 ^[16]

3) 计算 $f(x)$ 的一阶梯度: $g_k = \nabla f(x_k) = Ax_k - b$

4) 计算搜索方向 $d_k = -H_k g_k$

5) 计算步长 α_k : 根据 wolfe 准则或者 strong wolfe 准则确定,一般首先 $\alpha_k = 1$ 作为实验步长

6) 去下一个迭代点: $x_{k+1} = x_k + \alpha_k d_k$

7) 如果 $k > m$

摒弃内存中的矢量对 $\{s_{k-m}, y_{k-m}\}$; 计算和保存最新的矢量对 $\{s_k, y_k\}$

$s_k = x_{k+1} - x_k, y_k = \nabla f_{k+1} - \nabla f_k$

8) $K=k+1$; 跳转至步骤 2)

LBFGS 算法是一种用于无约束优化线性搜索的方法,隐舍地将近似 H_k 更新为逆 Hessian。通过式(12)可知,计算 H_{k+1} 只需要存储 $m+1$ 组 $\{s_i, y_i\}$ 矢量对即可^[15],这和直接求逆矩阵需要存储 $n \times n$ 维数据相比,在减少内存和提高

计算速度方面都有明显的优势。因此这就激发用 LBFGS 算法来替换高阶矩阵来重新构造新的 EKF。

2.2 LBFGS-EKF 算法

基于所描述的 EKF 概率模型,利用拟牛顿法中能够避免 Hessian 矩阵计算的优势对增益矩阵计算过程设计矩阵求逆的问题进行修正建立 LBFGS-EKF 新的算法模型。

LBFGS-EKF 算法迭代方程组如下:

1) 选择初始猜测值 x_0 和协方差矩阵 $B_{\#} = C_0$, 并设置 $k=0$

2) 计算状态一步预测 x_k^- 和一步预测协方差阵 C_k^-

a. 计算 $x_k^- = M_k x_{k-1}$

b. 计算 $C_k^- = M_k B_{\#} M_k^T + C_{w_k}$

3) 计算卡尔曼滤波估计值 x_k 和其协方差值 C_k

a. 定义 $A = (K_k C_k^- K_k^T + C_{v_k})$, $b = y_k - K_k x_k^-$ 用于式(6)

中;

利用 LBFGS 法求解满足 $\min f(x)$ 的最优解 $x^* = A^{-1}b$, 以及最优逆矩阵 $B^* = A^{-1}$

b. 计算 LBFGS-EKF 估计 $x_k = x_k^- + C_k^- K_k^T x^*$

c. 定义 $A = C_k^- - C_k^- K_k^T B^* K_k C_k^- (\approx C_k)$, $b = 0$ 用于式(6)中; 利用 L-BFGS 法即可求解出估计协方差矩阵 C_k

4) $k = k + 1$; 跳转至 2)

3 仿真结果与分析

接下来对本文所提及的算法进行仿真验证,并与 LBFGS-EKF 算法进行对比。基于 Leap Motion 识别范围的有限性,考虑目标跟踪实在匀速运动。定义目标的状态向量 $X = [x; \dot{x}; y; \dot{y}; z; \dot{z}]$, 分别表示目标在 x, y, z 3 个方向的位置和速度。在 MATLAB 仿真环境下进行蒙特卡洛仿真 10 次。实验环境是一台 HP Pavilion 14-AL028TX CPU 2.3 GHz 计算机上用 MATLAB R2014a 进行软件仿真。仿真结果如图 2~6 所示。

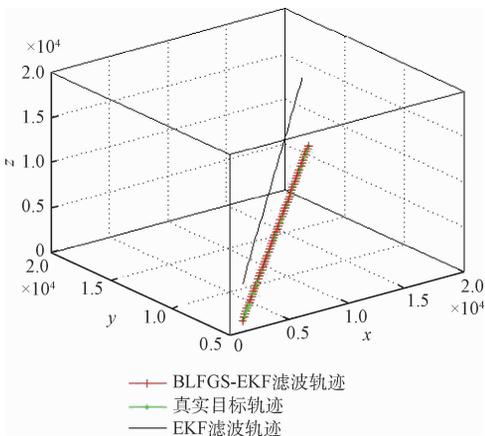


图 2 跟踪轨迹

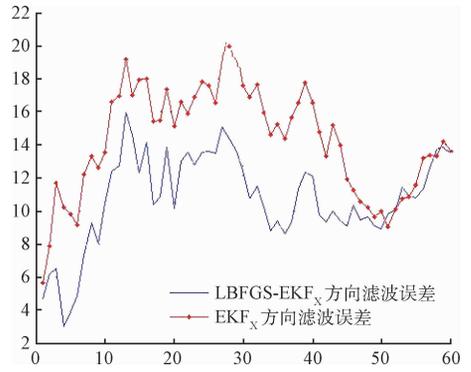


图 3 X 方向偏差

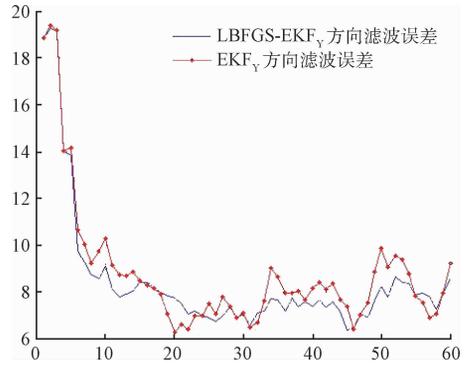


图 4 Y 方向偏差

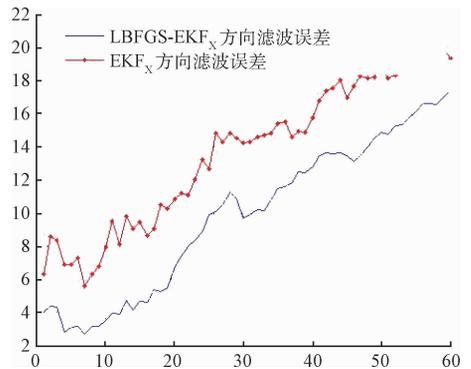


图 5 Z 方向偏差

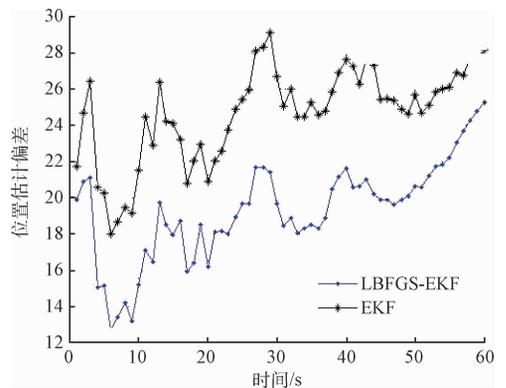


图 6 位置估计偏差

从图2可以看出,目标在整个运动期间,本文所提出的改进算法能够很好地对目标进行跟踪。并且从图3~6可以更加清晰地看到,本文所提出的算法整体误差相对于EKF算法较小,尤其是在大规模高维运算中,EKF算法在目标跟踪时,其算法中的求逆矩阵过程将会占据内存空间,提高处理器负载,导致计算速度下降;而LBFGS-EKF只选取前 m 个状态量就能很好的实现目标跟踪,这明显对减少计算机内存消耗,提高计算速率很有益的。

4 结 论

本文结合扩展卡尔曼滤波EKF和LBFGS算法各自的优势,提出了基于三维空间目标跟踪的LBFGS-EKF算法。仿真结果表明,该算法跟踪精度明显要高于原始的EKF算法。但由于我们是在匀速运动的环境下进行测试,考虑的变量参数相对有限,无法实质性体现出高维运算对运算速率的影响,所以当维度量较少时,实验结果的实时性并不是如期望中的那么理想。由于手势在运动过程中不可能只做匀速运动,可能出现突然加速、减速、转弯等比较复杂多变的情况,如何在这种复杂情况下提高跟踪的精度将是今后的研究方向。同时,这种多变环境必然考虑加入更多参数参与运算,如何体现高维参数下减少内存消耗、提高运算速率也将是今后的研究内容。

参考文献

- [1] 易靖国,程江华,库锡树. 视觉手势识别综述[J]. 计算机科学, 2016, 43(6A):103-108.
- [2] SUN Z S, BAO G, LI X N. Experimental research of force feedback dataglove based on pneumatic artificial muscle [J]. CHINESE JOURNAL OF MECHANICAL ENGINEERING, 2009, 22 (4): 489-495.
- [3] SANGHEON P, SUNJIN Y, JOONGROCK K, et al. 3D hand tracking using Kalman filter in depth space[J]. On Advances in Signal Processing, 2012, 1: 36-55.
- [4] 孟明,杨方波,余青山,等. 基于Kinect深度图像信息的人体运动检测[J]. 仪器仪表学报, 2015, 36(2): 386-393.
- [5] 郭雷. 动态手势识别技术综述[J]. 软件导刊, 2015, 14(7): 8-10.

- [6] 潘佳佳,徐昆. 基于Leap Motion的三维自由手势操作[J]. 中国科技论文, 2015,(2): 207-212.
- [7] 胡弘,晁建刚,杨进,等. Leap Motion关键点模型手姿态估计方法[J]. 计算机辅助设计与图形学学报, 2015, 27(7): 1211-1216.
- [8] 朱明茗,景红. 基于体感技术的手势追踪与识别[J]. 计算机系统应用, 2014, 23(8): 228-232.
- [9] AUVINEN H, BARDSLEY J M, HAARIO H, et al. Large-scale kalman filtering using the limited memory BFGS method [J]. Electronic Transactions on Numerical Analysis, 2009, 35: 217-233.
- [10] 李莉. 衰减记忆扩展卡尔曼滤波在目标跟踪中的应用[J]. 电子测量技术, 2011, 34(2): 36-65.
- [11] 张俊超. 结构约束下的动态PET图像重建研究[D]. 杭州:浙江大学, 2012.
- [12] 何炼坚. 拟牛顿矩阵的紧凑表示及其在有限存储方法中的应用[J]. 成都大学学报:自然科学版, 2009, 28(1): 29-32.
- [13] 卢慧芳,杨月婷. 一种新的修正有限内存拟牛顿法[J]. 华东师范大学学报:自然科学版, 2010(1): 35-37.
- [14] DAI Y H. A perfect example for the BFGS method[J]. Springer and Mathematical Optimization Society, 2013, 138: 501-530.
- [15] QIAN X Y, SHI Q S, LIU H, et al. A class of LBFGS-Type algorithms for large-scale unconstrained optimization[J]. 运筹学学报, 2011, 15(3): 9-20.

作者简介

雷安琪, 硕士研究生, 主要研究方向为增强现实、目标跟踪。

E-mail: leianqi3816@163.com

冯玉田, 工学博士, 副教授, 主要研究方向信号检测与处理、物联网等。

E-mail: fengyt@shu.edu.cn

唐子成, 硕士研究生, 主要研究方向为嵌入式开发、图像处理。

E-mail: 504651015@qq.com