

基于非线性扩散滤波的双目视觉跟踪与测量研究实现

于 菲, 战兴群, 刘 铭

(上海交通大学 航空航天学院, 上海 200240)

摘要: 针对双目视觉跟踪与测量技术的研究主要集中在缩短时间, 精确性问题考虑不够充分; 结合 KAZE 特征提取算法, 利用卡尔曼预测对目标进行跟踪, 提出一种基于图像匹配的双目视觉跟踪与测量算法; 将 KAZE 特征提取算法和 SIFT、SURF 等算法进行仿真对比, 结果表明 KAZE 算法具有更加出色的匹配率, 而跟踪速度相当; 运用 KAZE 算法与卡尔曼预测跟踪技术实现了基于 Visual Studio 环境下的双目视觉测量与跟踪系统, 测试结果表明, 该系统可进行有效的高精度测量。

关键词: 双目视觉; 图像匹配; 非线性扩散滤波; 视觉跟踪; 卡尔曼预测

Stereo Vision Tracking and Measurement System based on KAZE feature

Yu Fei, Zhan Xingqun, Liu Ming

(Shanghai Jiaotong University, Shanghai 200240, China)

Abstract: The intention of this research is achieve an accurate stereo visual tracking and measurement system. The emphasis of this application lies in the KAZE feature, a novel image matching algorithm, Kalman prediction for target tracking. The result of KAZE comparison experiment indicates that comparing the traditional feature extraction algorithm matching algorithm KAZE has higher match ratio and better robustness. The application of Kalman prediction algorithm with KAZE in stereo visual tracking and measurement system under Visual Studio environment shows that the prediction accuracy of binocular tracking has been improved, and achieve better accuracy in the medium-long range.

Keywords: stereo vision; image matching; KAZE feature; visual tracking; Kalman prediction

0 引言

视觉信息是人类识别理解周围环境状况的重要信息来源, 自计算机技术诞生以来, 人们一直在寻找如人眼一样灵活高效的计算机视觉系统。在众多的视觉系统中, 双目视觉以其在深度信息上的优势, 成为近年来计算机视觉与图像处理领域的研究热点。双目视觉技术在机器人的定位导航、避障、地图构建和测距等方面均得到了广泛应用。其中立体视觉技术中一个核心的技术是图像匹配算法, 它是视觉跟踪与测量的关键步骤, 它的匹配精度会直接影响到系统最终跟踪与测量的精度。传统的 SIFT、SURF 等特征检测算法通过线性高斯金字塔进行多尺度分解^[1], 这样容易造成细节丢失和边界模糊, 且传统方法通过基于正向欧拉法求解非线性扩散方程时, 迭代收敛步长短, 计算复杂度高, 耗时长。

本文利用非线性扩散滤波(KAZE)算法, 采用加性算子分裂算法(additive operator splitting, AOS)^[2]来进行非线性扩散滤波, 采用任意步长来构造稳定的非线性尺度空间, 对图像进行非线性扩散处理, 融合卡尔曼^[3]预测模型对目标中心点位置进行预测。进而运用局部搜索, 降低搜索区域, 提高跟踪速度。同时提出一种融合图像分割和图像匹配以及自学习的新算法, 提高了跟踪稳健性, 最后并将其应用到双目测量跟踪实验中, 以实现精确的跟踪与测量效果。

1 KAZE 特征检测匹配算法

KAZE 特征检测匹配算法是一种非线性尺度空间的多尺度 2D 特征提取和描述算法^[4]。KAZE 在日文代表“风”, 其命名是为了纪念尺度空间分析的开创者——日本学者 Iijima^[5]。在自然界中, 风被定义为受非线性约束的大规模空气流动。KAZE 与风类似, 也是通过这种方式在图像域中进行非线性扩散处理。该算法的最大特点是使用非线性扩散滤波方法来构造一个非线性尺度空间。

1.1 非线性扩散滤波

KAZE 算法中使用 Perona 和 Malik 提出的各向异性扩散(Anisotropic diffusion)公式^[6-7], 其主要思路是构造一个各向异性的扩散方程, 使得灰度的扩散在平缓的图像区域比较快, 在边缘处(即梯度较大的地方)扩散比较慢, 即非线性化。具体而言, 非线性扩散滤波方法是将图像亮度 $[L]$ 在不同尺度上的变化视为某种形式的流动函数的散度, 可以通过非线性偏微分方程来描述:

$$\frac{\partial L}{\partial t} = \text{div}(c(x, y, t) \cdot \Delta L) \quad (1)$$

在 KAZE 算法中使用 Perona 和 Malik 给出的传导函数定义为:

$$c(x, y, t) = g(|\Delta L_\sigma(x, y, t)|) \quad (2)$$

其中: ΔL_σ 是高斯平滑后的图像 L_σ 的梯度[gradient]。在 KAZE 算法中, 默认:

$$g = \frac{1}{1 + \frac{|\Delta L_\sigma|^2}{k^2}} \quad (3)$$

KAZE 算法中, 参数 k 是对比度因子, 它的值是梯度图像 ΔL_σ 的直方图 70% 上的值。

1.2 AOS 算法

KAZE 算法使用 AOS 算法来构造尺度空间, AOS 算法采

收稿日期:2015-08-07; 修回日期:2015-09-17。

基金项目:国家高技术研究发展计划(863 计划)(2014AA123103)。

作者简介:于 菲(1991-),女,河南人,硕士研究生,主要从事视觉导航方向的研究。

战兴群(1970-),男,教授,博士生导师,主要从事组合导航技术方向的研究。

用隐式差分格式来改善显示差分格式求解非线性偏微分方程时收敛慢的缺点:

$$\frac{L^{i+1}-L^i}{\tau}=\sum_{i=1}^m\mathbf{A}_i(L^i)L^{i+1}\tag{4}$$

其中: \mathbf{A}_i 是表示图像在各维度 (l) 上传导性矩阵。该方程的解如下:

$$L^{i+1}=\left(I-\tau\sum_{i=1}^m\mathbf{A}_i(L^i)\right)^{-1}L^i\tag{5}$$

这种求解方法对任意时间步长 τ 都有效。

1.3 构造非线性尺度空间

KAZE算法的尺度空间呈一个金字塔型。尺度级别按对数递增,共有 O 个层级,每个层级有 S 个子层级。但是 KAZE 的各个层级均采用与原始图像相同的分辨率。不同的层级和子层级分别通过序号 o 和 s 来标记,尺度参数可以用以下来表示:

$$\sigma_i(o,s)=\sigma_o2^{o+s},\alpha[0,\cdots,O-1],\times[0,\cdots,S-1],i_k[0,\cdots,N]\tag{6}$$

其中: σ_o 是尺度参数的初始基准值, $N=O\times S$ 是整个尺度空间包含的图像总数。非线性扩散滤波模型以时间为单位,因此需要将像素为单位的尺度参数 σ_i 转换至时间单位。

在 KAZE 算法中,根据一组进化时间,利用 AOS 的方程就可以构造出 KAZE 算法的非线性尺度空间:

$$L^{i+1}=\left(I-(t_{i+1}-t_i)\sum_{i=1}^m\mathbf{A}_i(L^i)\right)^{-1}L^i\tag{7}$$

1.4 特征点检测

与 SIFT 算法类似 KAZE 算法也是通过寻找局部极大值点来检测特征点,在不同尺度寻找像素点进行归一化的 Hessian 局部极大值点。在寻找极值点时,同样也是把每一个像素点和它同尺度及相邻 2 个尺度上的 26 个相邻点进行比较,当其大于或小于所有的相邻点时,即为极值点。

1.5 特征值主方向和描述向量

KAZE 算法描述特征点的方法与 SURF 类似,在梯度图像上以一个尺度参数为的特征点为中心,取周围的正方形区域,并将这个区域划分为个子区域,让相邻的子区域有宽度为的相交部分,每个子区域大小为。然后使用高斯核对每个子区域进行高斯加权,每个子区域计算出一个 4 维的描述向量:

$$d_v=(\sum L_x,\sum L_y,\sum |L_x|,\sum |L_y|)\tag{8}$$

最后进行归一化处理后,可得到 16 个区域,每个 4 维的 64 维描述向量。

1.6 KAZE 性能

本文使用了 KAZE 作者在论文中提到的测试图像库^[3],通过 Visual Studio 2010 和 OpenCV 对该图库的仿真,再与 matlab 结合,从高斯模糊度、旋转角度、缩放程度、亮度变化程度等方面对 KAZE、SIFT、SURF、ORB 进行性能仿真测试。其中性能测试包括正确匹配点百分比、匹配点百分比、匹配率等。仿真结果如图 1、图 2、图 3,其横坐标范围见表 1。

仿真结果表明 KAZE 与 SIFT、SURF、ORB 相比,从各方面而言,正确匹配点百分比,匹配点百分比以及匹配率都表现最好,精确度最高。

2 双目视觉跟踪与测量

2.1 常见跟踪算法

目前常见的视觉跟踪算法,按照跟踪原理主要分为:基于区域的跟踪算法、基于特征的跟踪算法、基于变形模板的跟

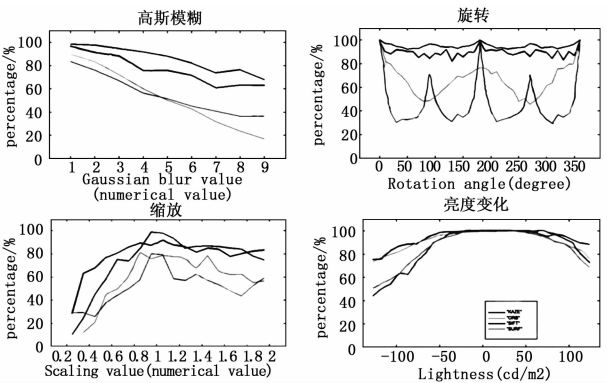


图 1 正确匹配点百分比

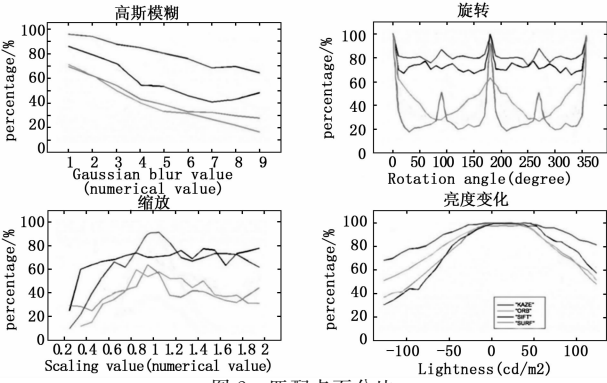


图 2 匹配点百分比

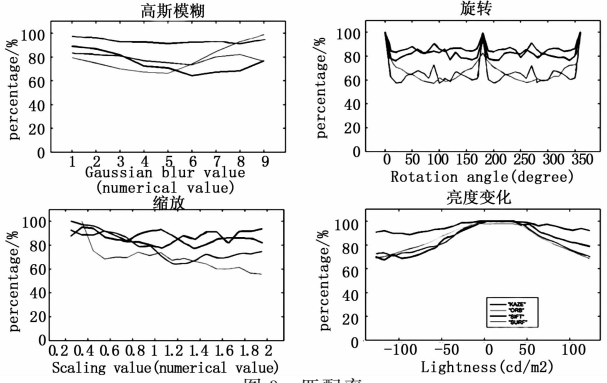


图 3 匹配率

表 1 4 种特征匹配算法对比图横坐标范围

性能	最小值	最大值	步长
高斯模糊度	1	9	1
旋转角度	0	360	10
缩放程度	0.25	2	0.1
亮度变化	-127	127	10

踪、基于模型的跟踪、基于运动估计的跟踪^[8-9]。本文选取的跟踪算法是基于特征点匹配的算法。

基于特征的跟踪算法和基于区域的跟踪算法相比,都需要计算目标图像与搜索窗中图像的相关性。但是它们选取的特征类型不同,基于区域的算法易受到周围环境的影响,如在纹理、颜色类似的背景中,该算法易丢失目标。而基于特征的算法,用的是目标的局部特征信息,如角点、边缘、轮廓等,能够抑制一定程度的遮挡,只要部分区域中仍存在目标中的特征信息,就可以跟踪到目标。基于特征的跟踪模型与卡尔曼滤波

结合能进一步提升跟踪性能。

本文在其基础上又加入学习机制, 将其应用到双目环境中, 并融入卡尔曼预测模型预估搜索区, 减少匹配搜索面积, 从而提高系统实时性。

2.2 卡尔曼预测模型

卡尔曼预测就是根据当前的跟踪目标的位置信息, 预测目标下一时刻的位置信息, 只在预测目标的邻域内搜索, 从而减少搜索区域, 提高匹配跟踪速度^[10]。基于卡尔曼的双目视觉跟踪与测量系统如图 4 所示。

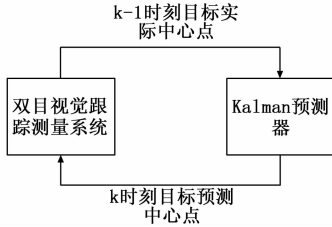


图 4 基于卡尔曼双目视觉跟踪系统流程图

本文将跟踪目标的特征点集中区域, 左图像中目标中心点 $p_l(u_l, v_l)$ 和右图像中目标中心点 $p_r(ur, vr)$ 作为观测点, 以此作为观测向量, 预测下一帧图像中该点的位置。跟踪预测总体框架如下:

- 1) 获取 $k-1$ 时刻时, 左右图像中目标的中心点 $p_l(k-1)$ 和 $p_r(k-1)$;
- 2) 根据 $p_l(k-1)$ 和 $p_r(k-1)$, 用卡尔曼预测器估计出 k 时的 $\hat{p}_l(k)$ 和 $\hat{p}_r(k)$;
- 3) 根据 k 时刻的中心点, 在其邻域设置 ROI 进行搜索, 重复 1)、2) 步。

卡尔曼预测模型的状态向量为 $\mathbf{X}(k) = (u(k), v(k), \dot{u}(k), \dot{v}(k))^T$, 观测向量为 $\mathbf{Y}(k) = (u(k), v(k))^T$, 设该图像跟踪离散系统的状态方程为:

$$\mathbf{X}(k) = \mathbf{A}(k, k-1)\mathbf{X}(k-1) + \mathbf{W}(k-1) \quad (9)$$

观测方程为:

$$\mathbf{Y}(k) = \mathbf{H}(k)\mathbf{X}(k) + \mathbf{V}(k) \quad (10)$$

其中: $\mathbf{A}(k, k-1)$ 为状态转移矩阵, $\mathbf{W}(k-1)$ 为系统噪声。 $\mathbf{H}(k)$ 观测矩阵, $\mathbf{V}(k)$ 为观测噪声, 并设 $\mathbf{W}(k)$ 和 $\mathbf{V}(k)$ 为相互独立, 且为零均值的高斯白噪声。它们满足高斯分布 $p(w) \sim N(0, \mathbf{Q}(k))$, $p(v) \sim N(0, \mathbf{R}(k))$, $\mathbf{Q}(k)$ 、 $\mathbf{R}(k)$ 分别是 $\mathbf{W}(k)$ 、 $\mathbf{V}(k)$ 的协方差矩阵。

考虑图像坐标系的 u 轴和 v 轴, 状态方程为:

$$\mathbf{X}(k) = \begin{bmatrix} u(k) \\ v(k) \\ \dot{u}(k) \\ \dot{v}(k) \end{bmatrix} = \begin{bmatrix} 1 & 0 & T & 0 \\ 0 & 1 & 0 & T \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u(k-1) \\ v(k-1) \\ \dot{u}(k-1) \\ \dot{v}(k-1) \end{bmatrix} + \begin{bmatrix} T^2/2 \\ T^2/2 \\ T \\ T \end{bmatrix} \mathbf{W}(k-1) \quad (11)$$

而观测方程为:

$$\mathbf{Y}(k) = \begin{bmatrix} u(k) \\ v(k) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} u(k) \\ v(k) \\ \dot{u}(k) \\ \dot{v}(k) \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} \mathbf{V}(k) \quad (12)$$

设 $\mathbf{P}(k)$ 为 $\hat{\mathbf{X}}(k)$ 的协方差矩阵, $\mathbf{P}'(k)$ 为 $\mathbf{X}(k)$ 和 $\hat{\mathbf{X}}(k)$ 的误差向量的协方差矩阵。可得到卡尔曼预测方程:

$$\hat{\mathbf{X}}(k) = \mathbf{A}(k, k-1)\hat{\mathbf{X}}(k-1) + \mathbf{K}(k)(\mathbf{Y}(k) - \mathbf{H}(k)\mathbf{A}(k, k-1)\hat{\mathbf{X}}(k-1)) \quad (13)$$

其中:

$$\mathbf{P}'(k) = \mathbf{A}(k, k-1)\mathbf{P}(k-1)\mathbf{A}(k, k-1)^T + \mathbf{Q}(k)\mathbf{K}(k) = \mathbf{P}'(k)\mathbf{H}(k)^T(\mathbf{H}(k)\mathbf{P}'(k)\mathbf{H}(k)^T + \mathbf{R}(k))^{-1}\mathbf{P}(k) = (1 - \mathbf{K}(k)\mathbf{H}(k))\mathbf{P}'(k)$$

由此我们使用第一帧图像获得的目标图像中心点 $\mathbf{X}(0)$, 在跟踪系统中, 一般系统噪声较大, 测量观测噪声较小, 可以选取, $\mathbf{P}(0) = \mathbf{R}(0) = \mathbf{I}$, $\mathbf{Q}(0) = 2\mathbf{I}$, \mathbf{I} 为单位矩阵^[11]。本文的双目视觉模型, 需要将状态向量和观测向量进行扩展^[12], 极线约束的加入使预测更加准确, 减少了搜索区域, 提高了双目视觉跟踪稳定性。

2.3 双目视觉跟踪

2.3.1 获取跟踪目标信息

跟踪目标通过人工方式选取, 通过提取目标的图像大小, 作为 ROI 窗口的尺寸大小。再将目标的图像导入到跟踪系统的目标库, 进行初始化。提取目标上的特征信息, 并做备份, 保存一份原始目标图像。

2.3.2 跟踪初始化

1) 目标库初始化: 首先对相机获取的视频序列进行全帧搜索, 提取第 1 帧中的特征点, 并与跟踪目标上的特征点进行匹配, 获得目标在第 1 帧中的匹配点集, 对这个匹配点集的坐标求均值, 获取该匹配点集的中心点。然后以这个中心点为中心绘制 ROI 窗口, 标出目标的图像坐标。若连续 2 帧都能检测到目标的位置。对 ROI 窗口内的目标进行截取并保存到目标图片库中。

2) 卡尔曼预测: 为了减少搜索区域, 在算法中引入了卡尔曼预测器, 根据卡尔曼预测算法。对目标中心点下一时刻的位置进行预测。

2.3.3 目标丢失时应对机制

1) 局部搜索目标丢失: 当在卡尔曼预测的范围内搜索不到目标时, 算法转为全帧搜索。找到目标后, 再进入预测邻域内的局部搜索。

2) 全局搜索丢失目标: 首先倒序从目标图像库中进行搜索匹配。跟踪到目标则继续目标学习并跟踪。若目标图像库无目标, 则用 1) 中备份的原始目标图像的特征点去搜索全帧图像中的目标。

3) 仍未搜索到目标: 目标已不在相机的视野范围内或者目标跟踪丢失。

2.4 双目视觉测量

双目视觉测量, 是在完成双目相机的建模, 标定, 图像采集, 图像匹配过程后, 运用匹配的特征点对左右图像坐标系下的坐标, 根据双目视觉数学模型, 还原出该空间点的三维坐标的过程。

2.4.1 匹配特征点三维坐标测量原理

对于一个标定过的双目视觉测量系统, 在获得左右相机图像中的匹配特征点对 $p_{left}(u_l, v_l)$ 和 $p_{right}(u_r, v_r)$ 后, 根据双目视觉测量模型^[13], 就可以通过反投影变换解算出该特征点对应的空间点 $P(x, y, z)$ 的摄像机坐标系下的三维坐标值。

$$\begin{cases} x = zX_1/f_l \\ y = zY_1/f_l \\ z = \frac{f_l(f_r t_x - X_r t_z)}{X_r(r_7 X_l + r_8 Y_l + f_l r_9) - f_r(r_1 X_l + r_2 Y_l + f_l r_3)} = \frac{f_l(f_r t_x - Y_r t_z)}{Y_r(r_7 X_l + r_8 Y_l + f_l r_9) - f_r(r_4 X_l + r_5 Y_l + f_l r_6)} \end{cases} \quad (14)$$

2.4.2 极线约束补偿测量

在双目跟踪与测量中, 需要考虑极线约束补偿^[14-15]的情况。

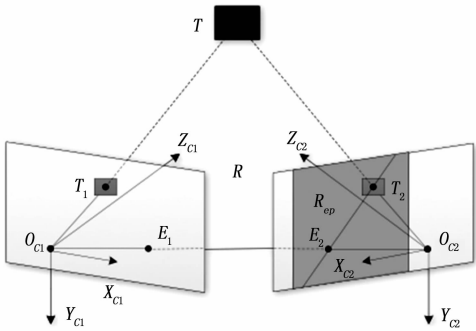


图 5 极线约束补偿解算空间坐标示意图

如果只在左相机搜索到目标 ROI 区域, 其中心点为 T_1 , 这时需要对右相机内的目标进行极线约束预测。 T_1 对应的极线为 E_2 T_2 , 取以该极线为对角线的区域加上左右各扩展 0.5a 的长度的区域为局部搜索区域 R_{ep} , 如图 5 右图像中深色区域, 如果仍搜索不到目标再进行全帧搜索。

基于图像匹配的双目跟踪首先要获取跟踪目标的图像, 然后通过两个相机同时获取跟踪目标, 再根据双目视觉测量模型求出目标的中心点, 假设这两个点为双目视觉中实际目标中心点的投影点对。通过这两个点, 进行双目视觉跟踪与测量。

2.5 实验结果与分析

首先使用张正友标定法^[16]对双目相机进行标定。利用 Matlab 提供的标定工具箱完成双目视觉系统的标定: 实验使用视觉看 seeker 型号 USB 网络摄像头, 分辨率为, 焦距 4.3 mm。图 6 是用十二张标定图片标定的双目相机的三维立体标定结果。相机参数见表 2。

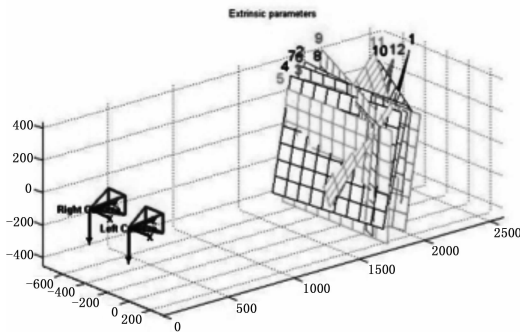


图 6 立体标定结果

表 2 双目相机标定参数

	左相机	右相机
畸变系数 k	831.82432	816.35316
焦距 f	0.21914	0.21095
透视中心	[321.64761 252.78004]	[340.44573 241.70518]
摄像机旋转矩阵 R		
摄像机平移向量 T		

再将双目跟踪与双目静态测量相结合, 在 Visual Studio 2010, OpenCV2.4.11 环境下, 用 MFC 建立了一套双目视觉测量与跟踪系统。将得到的相机标定结果输入系统中, 其解算得到的目标中心点坐标为 (24.2047, 68.000, 629.027), 单

位 mm。通过普通卡尺测量得到目标实际距离约是 618 mm。相对误差为 1.78%。考虑到相机本身的工艺问题和实际测量带来的测量误差, 该方法的测量误差很小, 测量精度很高, 可以满足机器人视觉中对跟踪测量的精确要求。

3 结论

本文针对双目视觉跟踪与测量问题进行了研究, 着重解决之前图像匹配中存在的精度差的问题。研究了基于 KAZE 图像的匹配算法, 通过非线性滤波来任意步长的构造稳定的非线性尺度空间来消除边界模糊和细节丢失, 并创新性的将其应用到目标跟踪上, 获得了良好的效果。通过改变亮度, 改变高斯模糊程度, 旋转角度, 缩放变化的 matlab 性能仿真实验验证了 KAZE 的优良性。相比之前的跟踪算法只能根据特定的颜色或形状实现双目实时测距, 本文的算法可以跟踪任何特征点较丰富的目标, 并能够通过自学习, 完善目标图像库, 具有很出色的精确性, 较好地满足移动机器人对三维环境的感知需求, 具有较高的实用价值。

参考文献:

[1] 黄 超, 齐英剑. SIFT 算法研究和应用 [J]. 中国传媒大学学报: 自然科学版, 2012, 19 (1): 68-72.

[2] Weickert J, Romeny B M T H, Viergever M A. Efficient and reliable schemes for nonlinear diffusion filtering [J]. IEEE Transactions on Image Processing. 1998; 398-410

[3] Grewal M S, Andrews A P. Kalman filtering; theory and practice with MATLAB [M]. John Wiley & Sons, 2014.

[4] Alcantarilla P F, Bartoli A, Davison A J. KAZE features [A]. Computer Vision - ECCV 2012 [C]. Springer Berlin Heidelberg, 2012; 214-227.

[5] Weickert J, Ishikawa S, Imiya A. Linear scale-space has first been proposed in Japan [J]. Journal of Mathematical Imaging and Vision, 1999, 10 (3): 237-252.

[6] Perona P, Malik J. Scale-space and edge detection using anisotropic diffusion [J]. Pattern Analysis & Machine Intelligence, 1990, 12 (7): 629-639.

[7] 许冠军. 数字图像去噪算法研究 [D]. 杭州: 浙江大学, 2006.

[8] 杨 戈, 刘 宏. 视觉跟踪算法综述 [J]. 智能系统学报, 2010, 5 (2): 95-105.

[9] 王 亮, 胡卫明, 谭铁牛. 人运动的视觉分析综述 [J]. 计算机学报, 2002, 25 (3): 225-237.

[10] 王 江, 付 强, 全 权, 等. 基于 Kalman 滤波和直方图匹配的双目视觉跟踪 [J]. 新型工业化, 2013, 3 (2): 23-33.

[11] 任少盟, 魏振忠, 张广军. 基于扩展 Kalman 预测模型的双目立体视觉图像特征点跟踪方法 [J]. 电子测量技术, 2012, 35 (1): 71-75.

[12] 曲云腾, 李平康, 杜秀霞. 基于 Kalman 预测的人体运动目标跟踪 [J]. 计算机系统应用, 2011, 01: 137-140.

[13] 马颂德, 张正友. 计算机视觉—计算理论与算法基础 [M]. 北京: 科学出版社, 1998.

[14] 邱茂林, 马颂德. 计算机视觉中摄像机定标综述 [J]. 自动化学报, 2000, 26 (1): 43-55.

[15] 段 峰, 王耀南. 机器视觉技术及其应用综述 [J]. 自动化博览, 2002, 19 (3): 59-61.

[16] Zhang Z. A flexible new technique for camera calibration [J]. Pattern Analysis and Machine Intelligence, 2000, 22 (11): 1330-1334.