

基于扩散距离的 SIFT 特征匹配算法^①

胡 刚, 刘侍刚, 吴清亮, 王 刚

(陕西师范大学 计算机科学学院, 西安 710062)

摘 要: SIFT (Scale Invariant Feature Transform)是目前最流行的局部特征提取及匹配算法. 但传统 SIFT 算法采用欧氏距离来度量特征之间的 SSD(Sum of Square Differences)并进行匹配, 而传统的欧氏距离不能使高维特征向量恢复到具有低维的几何结构, 导致错误匹配. 为了克服这缺点, 利用扩散距离代替欧氏距离进行匹配, 然后使用随机抽样一致从候选匹配中排除错误的匹配. 实验表明: 该方法在图像形变、光照变化和图像噪声方面优于原方法.

关键词: 计算机视觉; SIFT 特征描述符; 扩散距离; 图像匹配

SIFT Matching Algorithm Based on Diffusion Distance

HU Gang, LIU Shi-Gang, WU Qing-Liang, WANG Gang

(School of Computer Science, Shaanxi Normal University, Xi'an 710062, China)

Abstract: The SIFT(Scale Invariant Feature Transform) algorithm is now regarded as the best local feature extraction and matching algorithm. However, in the traditional SIFT algorithm, the Euclidean distance which could not change the high-dimensional feature vector into a low-dimensional geometry structure is used to measure the SSD(Sum of Square Differences) between two image features to match and results into mismatching. To overcome the shortcoming, an SIFT matching algorithm based on diffusion distance is proposed in this paper which replaces the Euclidean distance with the diffusion one. At the same time, RANSAC(Random Sample Consensus) is presented to exclude the mismatching points. Experimental results show that the proposed algorithm has more efficiency to deal with image deformation, illumination change and image noise than the traditional one.

Key words: computer vision; SIFT feature descriptor; diffusion distance; image matching

1 引言

图像的特征提取和匹配是计算机视觉的一个重要部分, 已经大量地应用于机器人导航、图像检索和目标识别等领域. 图像的特征可以分为特征点、特征边和特征区域三大类. 在实际的应用中, 特征点作为图像特征应用最为广泛, 其中比较有代表性的特征点提取算法有 Harris 方法、Shi-Tomasi 方法、FAST 方法等, 但这些方法都不具有尺度不变性, 因此, 使用这些方法进行图像匹配有一定的局限性.

2004 年 Lowe^[1]提出 SIFT 算法, 该算法采用一

种近似于尺度归一化的拉普拉斯高斯核(Laplacian of Gaussian kernel)的高斯差分 DoG(Difference of Gaussian)探测符, 使 SIFT 特征具有尺度不变性. 由于 SIFT 算法在尺度、旋转、亮度和视角的变化上有很好的鲁棒性并具有独特的图像特征, 该算法被用作特征提取的首选方法. 随后, 产生了一系列基于 SIFT 的改进算法, 如 Ke 和 Sukthankar 提出的基于 PCA 的 SIFT 算法^[2]; Mikolajczyk 和 Schmid 提出的 GLOH 算法^[3], 用对数-极坐标取代笛卡尔坐标; Bay 和 Ess 等人提出的 SURF 算法^[4], 用积分图像进行卷积. 虽然这些方

^① 基金项目: 国家自然科学基金项目(60805016); 高等学校博士学科点专项科研基金新教师基金课题(200807181007); 陕西省科技计划项目(2011JM8014); 中国博士后科学基金特别资助(200902594); 陕西师范大学中央高校基本科研业务费专项资金(GK201002016); 大学生创新性实验计划项目(1110718026)

收稿时间: 2011-12-25; 收到修改稿时间: 2012-02-16

法在某些性能方面改善了传统的 SIFT 算法的性能, 但这些方法都是采用欧氏距离来度量特征向量的差异, 然而使用欧氏距离仅仅得到的是像素灰度差异的累加, 一个较小的图像形变就可能使原来的欧氏距离变长, 从而导致匹配的误差较大。

Ling 和 Okada 提出的扩散距离是一种非线性的距离判断方式^[5], 它通过模拟热扩散和记录随机扩散遍历的时间计算两点之间的距离. 本文利用扩散距离计算 SIFT 特征描述符向量间的距离, 实验表明该方法在图像形变、光照变化和图像噪声方面所得到的匹配比传统的 SIFT 算法具有更好的匹配效果。

2 SIFT 算法

2.1 SIFT 图像特征

传统的 SIFT 算法的 4 个主要步骤:

1) 尺度空间极值探测. 采用高斯差分(Difference of Gaussian)金字塔的方法探测候选关键点的位置. 尺度空间 $L(x, y, \delta)$ 定义为可变尺度高斯函数 $G(x, y, \delta)$ 和输入图像 $I(x, y)$ 的卷积:

$$L(x, y, \delta) = G(x, y, \delta) * I(x, y) \quad (1)$$

$$G(x, y, \delta) = 1 / (2\pi\delta^2) e^{-(x^2+y^2)/2\delta^2} \quad (2)$$

式中 (x, y) 为图像像素坐标, δ 为尺度空间因子。

为了在尺度和空间所在的三维上探测到稳定的关键点, Lowe 提出在高斯差分函数 $D(x, y, \delta)$ 中寻找尺度空间局部极值, 高斯差分函数 $D(x, y, \delta)$ 定义为相邻图像间的高斯函数之差与图像 $I(x, y)$ 的卷积:

$$\begin{aligned} D(x, y, \delta) &= (G(x, y, k\delta) - G(x, y, \delta)) * I(x, y) \\ &= L(x, y, k\delta) - L(x, y, \delta) \end{aligned} \quad (3)$$

式中 k 为一常数 ($k = 2^{1/s}$, s 为对每一组(Octave)高斯差分函数 $D(x, y, \delta)$ 的采样尺度个数。

由式(3)可知, 高斯差分函数 $D(x, y, \delta)$ 可由相邻尺度空间相减求得; Lowe 经过实验论证, 当 $s = 3$ 时, 关键点的重复率最高, 计算量相对较少, 故取 $k = 2^{1/3}$; 当 $s = 3$ 时, 高斯差分金字塔中每一组的层数为 $s + 2 = 5$, 高斯金字塔中每一组的层数为 $s + 3 = 6$ 。

构造高斯差分金字塔的过程如图 1 所示, 在高斯金字塔中(如图 1 左侧所示), 第一组的第一层至第六层, 由原始图像线性插值扩大一倍后依次与可变尺度

高斯函数卷积得到(第一层的尺度空间因子为 δ , 后一层的尺度空间因子依次为前一层的 $k = 2^{1/3}$ 倍); 从第二组开始, 每一组的第一层由前一组的第 $s + 1 = 4$ 层图像降采样得到(如图 1 左侧曲线箭头所示), 图像变为原来的 1/4 大小; 高斯差分金字塔的每一层由左边相邻的高斯图像相减得到(如图 1 右侧所示)。

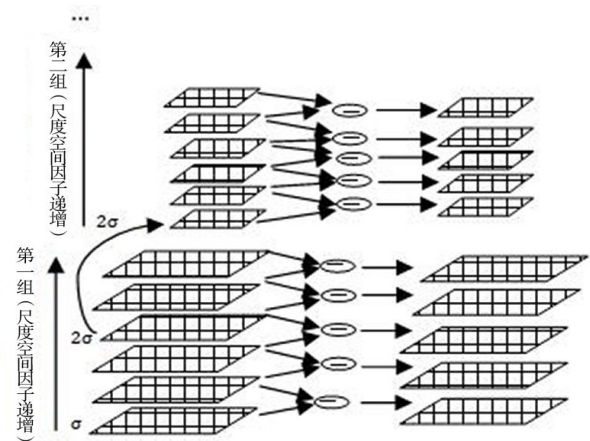


图 1 构造高斯差分金字塔

在高斯差分金字塔中, 每一组依次从第二层至第四层中的每一点与同层 8 个邻点和相邻上下层的各 9 个邻点(共 26 个点)比较, 寻找局部极值作为候选关键点, 如图 2 所示. 找到极值后输出该候选关键点所在的尺度空间因子和位置。

2) 关键点的定位. (a) 排除具有低对比度的关键点. 对高斯差分函数 $D(x, y, \delta)$ 进行三元二次拟合, 使极值点达到亚像素精度。

$$D(X) = D + \frac{\partial D}{\partial X} X + \frac{1}{2} X^T \frac{\partial^2 D}{\partial X^2} X \quad (4)$$

式中 $X = (x, y, \delta)$ 。

对高斯差分函数 $D(X)$ 求导, 得到高斯差分函数 $D(X)$ 的极值点:

$$\hat{X} = - \left(\frac{\partial^2 D}{\partial X^2} \right)^{-1} \frac{\partial D}{\partial X} \quad (5)$$

为了使极值点与其周围像素点之间具有足够大的差异性, 过滤掉极值点 $\hat{X} = (\hat{x}, \hat{y}, \hat{\delta})$ 中任何项大于 0.5 的点, 确保得到的关键点具有独特性。

将式(5)代入式(4), 得到高斯差分函数 $D(x, y, \delta)$ 的极值:

$$D(\hat{X}) = D + \frac{1}{2} \frac{\partial D^T}{\partial X} \hat{X} \quad (6)$$

(b)消除边缘响应. 由于高斯差分函数(DoG)具有很强的边缘响应. 利用高斯差分函数沿着边缘具有很大的曲率, 垂直边缘具有很小曲率的性质, 使用 Hessian 矩阵的特征值与高斯差分函数的曲率成正比的关系, 过滤掉靠近边缘的不稳定关键点.

Hessian 矩阵定义为:

$$H = \begin{pmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{pmatrix} \quad (7)$$

假设 Hessian 矩阵的两个特征值为 $\alpha, \beta (\alpha > \beta)$, 常数 γ 为 α, β 的比值, 即 $\alpha = \beta \gamma$;

$$\text{Tr}(H) = D_{xx} + D_{yy} = \alpha + \beta \quad (8)$$

$$\text{Det}(H) = D_{xx}D_{yy} - (D_{xy})^2 = \alpha\beta \quad (9)$$

$$\text{Tr}(H)^2 / \text{Det}(H) = (\alpha + \beta)^2 / \alpha\beta = (\gamma + 1)^2 / \gamma \quad (10)$$

过滤掉 γ 大于 10, 即式(10)大于 10 的边缘关键点.

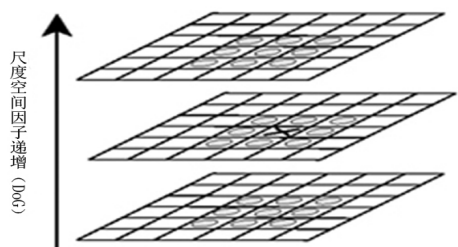


图 2 极值点的探测

3)方向的分配. 在尺度空间因子为最接近关键点所在尺度空间因子的高斯金字塔图像中(图1左侧), 计算特征点周围邻域点的梯度和方向. 每一点的梯度模值 $m(x, y)$ 和梯度方向 $\theta(x, y)$ 的计算公式分别如式(11), (12):

$$m(x, y) = \sqrt{(L(x, y) - L(x, y-1))^2 + (L(x+1, y) - L(x-1, y))^2} \quad (11)$$

$$\theta(x, y) = \tan^{-1}((L(x, y) - L(x, y-1)) / (L(x+1, y) - L(x-1, y))) \quad (12)$$

在特征点邻域 16×16 的采样点中, 每个采样点分别由自身梯度模值和标准差为该特征点所在尺度空间因子 1.5 倍的高斯权重圆形窗口加权, 累加采样点的梯度方向建立一个局部梯度方向直方图, 360度划分为

36 柱, 每一柱 10 度. 选取局部梯度方向直方图中最高峰值所在方向为特征点的主方向, 大于最高峰值 0.8 倍的峰值所在方向作为特征点的辅方向, Lowe 统计了大概有 15% 的特征点拥有多个方向, 即产生方向不同而具有相同位置和尺度的特征点.

4)特征点描述符. 继续在尺度空间因子为最接近关键点所在的尺度空间因子的高斯金字塔图像中, 使用标准差为描述符窗口宽度一半的高斯权重函数对采样点进行加权, 计算特征点周围采样点的梯度模值和方向, 将描述符坐标系和梯度方向坐标系分别旋转至特征点方向, 如图 3 所示.

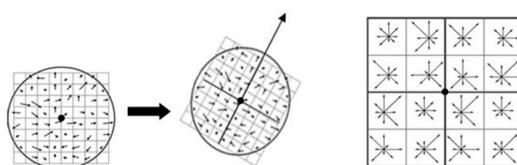


图 3 8×8 采样点生成 2×2 种子点 图 4 SIFT 特征描述符

如图 4 所示, 每一个关键点拥有 $4 \times 4 = 16$ 个种子点, 生成 16 个种子点 $\times 8$ 个方向 = 128 维非负元素组成的特征描述符向量. 将特征描述符向量归一化, 从而使特征描述符具有光照不变性; 进一步抑制非线性的光照改变, 再将上一步归一化后的特征描述符向量中的元素大于 0.2 的元素设为 0.2, 再一次归一化; 最终得到 SIFT 特征描述符向量.

3 SIFT特征匹配

由于传统的欧氏距离对图像形变比较敏感及其缺少考虑像素间的空间关系, 为了克服该缺点, 本文引入扩散距离.

3.1 扩散距离

一维分布 $h_1(x)$ 和 $h_2(x)$ 的差异可以表示成:

$$d(x) = h_1(x) - h_2(x) \quad (13)$$

假设在一个独立的温度场 $T(x, t)$ 中, 在时间 $t = 0$ 时, 有 $T(x, 0) = d(x)$. 温度场 $T(x, t)$ 的计算公式:

$$T(x, t) = T_0(x) * \phi(x, t) \quad (14)$$

式中 $T_0(x) = T(x, 0) = d(x)$, $\phi(x, t) = \frac{1}{(2\pi)^{1/2} t} e^{-\frac{x^2}{2t^2}}$.

一维分布 $h_1(x)$ 和 $h_2(x)$ 的扩散距离定义为:

$$\hat{K}(h_1, h_2) = \int_0^{\bar{t}} k(|T(x, t)|) dt \quad (15)$$

式中 \bar{t} 为一个正常数, $k(|T(x, t)|) = \int_{-\infty}^{\infty} |T(x, t)| dx$.

推广到高维, 假设有两个 m 维的直方图 $h_1(\mathbf{X})$ 和 $h_2(\mathbf{X})$, 其中 \mathbf{x} 为 m 维的实矩阵. 由于高斯金字塔中图像的平滑子采样不会产生混淆现象, 能有效地描述温度场 $T(\mathbf{X}, t)$ 的连续扩散过程, 所以使用基于高斯金字塔的交替距离函数定义高维扩散距离:

$$K(h_1, h_2) = \sum_{l=0}^L k(|d_l(\mathbf{x})|) \quad (16)$$

其中,

$$d_0(\mathbf{X}) = h_1(\mathbf{X}) - h_2(\mathbf{X}) \quad (17)$$

$$d_l(\mathbf{X}) = [d_{l-1}(\mathbf{X}) * \phi(\mathbf{X}, \sigma)] \downarrow_2 \quad (l=1, \dots, L) \quad (18)$$

$$\phi(\mathbf{X}, t) = \frac{1}{(2\pi)^{m/2} t} e^{-\frac{\mathbf{x}^T \mathbf{x}}{2t^2}} \quad (19)$$

式(18)中 \downarrow_2 表示降采样的一半大小, L 是高斯金字塔的层数, σ 是高斯函数 $\phi(\mathbf{X}, \sigma)$ 的标准方差, 式(19)中 m 为直方图的维数.

使用 L1 范数标准, 式(16)可以简化为:

$$K(h_1, h_2) = \sum_{l=0}^L |d_l(\mathbf{X})| \quad (20)$$

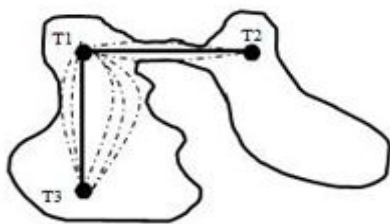


图 5 实线为欧氏距离, 虚线为扩散距离

如图 5 所示, 路径越多, 热扩散的越快, 相似度就越高; 点 T1 和点 T3 的欧氏距离与点 T1 和点 T2 的欧氏距离相等, 如图 5 实线所示; 而点 T1 和点 T3 之间的扩散距离比点 T1 和点 T2 之间的扩散路径多, 可知点 T1 和点 T3 的相似度比点 T1 和点 T2 高.

3.2 特征点的匹配

首先, 在生成的 128 维 SIFT 特征描述符中, 使用

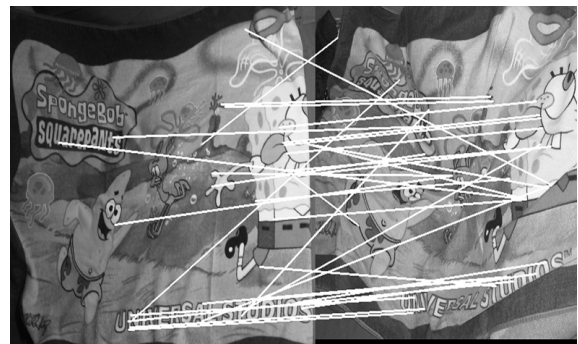
式(20)计算两幅图像关键点特征向量的扩散距离作为关键点相似性判定度量, 并用来判断特征向量间的 SSD.

然后, 使用近似最邻近查找算法 BBF(Best Bin First)查找关键点特征向量间的最近扩散距离与次最近扩散距离, 用最近扩散距离与次最近扩散距离的比值 (SSD_1/SSD_2)作为阈值(一般为 0.3—0.8)确定关键点的匹配, 去除相似特征点, 即排除所有比值大于某一阈值的匹配.

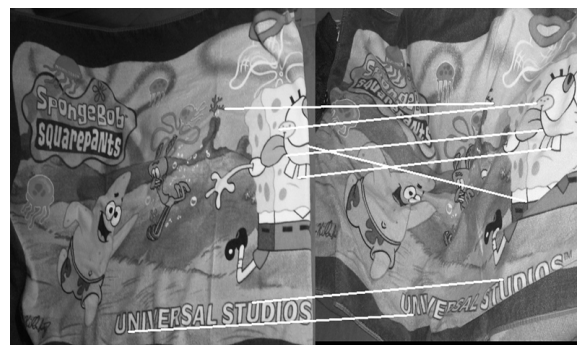
最后, 使用随机抽样一致算法迭代地从候选匹配特征中计算出最佳的图像变换模型, 找出所有满足该图像变换模型的候选匹配特征作为匹配特征, 排除不符合该图像变换模型的候选匹配特征, 即排除掉异常匹配(Outliers).

4 实验及分析

为了比较本文方法和传统的基于欧式距离的 SIFT 算法在图像存在形变及光照的变化情况下的匹配精度, 从 Haibin Ling 的图像库^[8]中选取十幅图像对进行匹配实验, 部分实验结果如图 6 和图 7 所示.

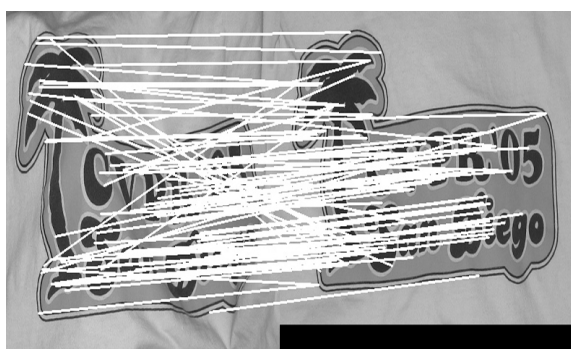


(a) 传统 SIFT 方法



(b) 本文方法

图 6 本文方法与传统 SIFT 方法对比实验图 1



(a)传统 SIFT 方法



(b)本文方法

图 7 本文方法与传统 SIFT 方法对比实验图二

从图 6 和图 7 可以看出, 这两幅图像对都存在较大的变形及光照变化和噪声. 本文使用不同的阈值参数, 分别用本文方法和传统的 SIFT 算法进行匹配实验, 实验结果如图 6(阈值取 0.5)和图 7(阈值取 0.6)所示, 其中图 6(a)和图 7(a)为 SIFT 使用欧式距离的方法, 图 6(b)和图 7(b)为 SIFT 使用扩散距离的方法.

从图 6 和图 7 中还可以看出, 本文方法比传统 SIFT 方法具有更高的匹配精度, 原因是由于传统的 SIFT 方法在计算两点之间的距离时采用欧氏距离, 而欧氏距离仅仅是像素灰度差异的累加, 若一个较小的图像形变就可能使欧氏距离变化较大, 导致匹配的误差较大. 而本文的方法采用的扩散距离, 可以避免该缺点. 因此, 本文方法比传统 SIFT 方法具有更高的匹配精度.

同时, 为了更加有效地对图像特征匹配方法进行评价, 本文引入了文献[7]中介绍的受试者工作特征曲线 ROC(Receiver Operator Characteristic), 其中横轴为假阳性率 FPR(False Positive Rate), 纵轴为真阳性率 TPR(True Positive Rate), 它们的范围都为(0-1).

在两幅图像的匹配中, 真阳性率=正确的匹配数/

第一幅图像中的特征点数; 假阳性率=不正确的匹配数/两幅图像的特征点数.

利用 ROC 曲线下方面积大小对匹配方法进行评价, 面积越大, 算法的性能越好, 匹配率越高.

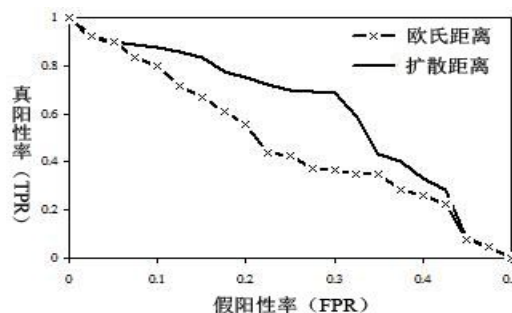


图 8 ROC 曲线比较

从图 8 中可以看出, 本文方法进行匹配所产生的 ROC 曲线下方面积比使用欧氏距离进行匹配所产生的 ROC 曲线下方面积大, 可知匹配率比原来的 SIFT 算法更高, 原因是由于本文采用扩散距离, 而传统方法采用的是欧氏距离.

5 结语

虽然传统的 SIFT 特征点具有尺度不变性, 有很高的独特性, 其特征描述符对旋转, 光照和视角的变化也具有较好的鲁棒性. 但是, SIFT 算法对图像形变比较敏感, 本文利用扩散距离代替传统的欧氏距离, 在某种程度上克服了传统 SIFT 方法的缺点. 实验结果表明, 使用扩散距离进行特征匹配具有很高的鲁棒性并提高了匹配的精度.

参考文献

- 1 Lowe DG. Distinctive image features from scale-invariant keypoints. International Journal of Compute Vision, 2004,60 (2):91-110.
- 2 Ke Y, Sukthankar R. PCA-SIFT: A more distinctive representation for local image descriptors. IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Washington DC:IEEE Computer Society, 2004. 506-513.
- 3 Mikolajczyk K, Schmid C. A performance evaluation of local descriptors. Pattern Analysis and Machine Intelligence, 2005,

(下转第 52 页)

由图 6 中实验可得,在路径不断扩大的情况下,小车仍旧能够顺利跟踪路径,LCD 显示屏上实时提取的路径可以得到,模糊控制器能够拟合不完整的路径,使得小车沿着最佳路径行驶。

小车在行驶中丢失路径的情况下,实验如图 7 所示.可以看出,小车丢失路径情况下能够较好的重新寻找路径,达到较好的控制效果.表 2 为控制器针对 ARM 优化前后小车性能比较结果。

表 2 优化前后运行速度比较

算法比较	优化前	优化后
直道行驶	0.5m/s	1.6m/s
弯道行驶	0.3m/s	0.9m/s
自动寻路	22s	12s

通过表 2 可以明显看出,经过优化后,小车直道行驶和弯道行驶的速度大大增加,在丢失路径的情况下初次寻得路径的时间大大缩短,实时性和鲁棒性明显提高。

表 3 为采用最大偏离度、最小偏离度、平均偏离度对多点预瞄模糊 PID 控制效果与单纯的多点预瞄 PID 控制对比结果.对比可得,应用本文所设计的控制器控制效果远远好于采用单纯多点预瞄控制器。

表 3 控制效果评价指标比较

评价指标	\bar{d}_p	$d_{p\max}$	$d_{p\min}$
本文控制器	4.025	6.496	3.018
单纯多点预瞄控制器	5.686	8.245	4.358

5 结语

本设计在硬件上使用 S3C6440 作为嵌入式开发板,采用视觉导航方式,实现了一款能够自动寻迹的智能导航轮式机器人.设计在传统预瞄控制基础上,融合模糊控制器与 PID 控制器,使机器人在路径缺失与路径丢失的情况下仍然能够自动寻找与拟合完整路径,具有更强的适应性与鲁棒性.软件上,针对嵌入式处理器的性能,对图像处理算法进行了优化,使得系统完全能够满足实时处理要求。

参考文献

- 1 朱少平,钱富才.超混沌系统同步非线性反馈控制.计算机工程与应用,2011,47(1):50-52,55.
- 2 董文杰,霍伟.链式系统轨迹跟踪控制.自动化学报,2000,26(3):310-316
- 3 Zhang K,Wu YX, Lu XQ. In Dynamic modeling and simulation for nonholonomic welding mobile robot. Journal of Cent South Univ Technol, 2007:05-0679-06.
- 4 Pourboghrat F, Karlsson MP. Adaptive control of dynamic mobile robots with nonholonomic constraints. Computers and Electrical Engineering, 2002,28:241-253.
- 5 Yu DY, Xue H. Application of fuzzy control method to AGV. Proc. of the 2003 IEEE International Conference on Robotics, Intelligent Systems and Signal Processing. 2003. 768-772.
- 6 毛佳,苑森淼.嵌入式直流 PWM 伺服控制系统的设计研究.计算机测量与控制,2003,6(11):420-422.

(上接第 96 页)

- 27(10):1615-1630.
- 4 Bay H, Ess A, Tuytelaars T, et al. SURF: Speeded Up Robust Features. Computer Vision and Image Understanding, 2008, 110(3):346-359.
 - 5 Ling H, Okada K. Diffusion Distance for Histogram Comparison. IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE Computer Society, 2006. 246-253.

- 6 Lu JY, Dorsey JL, Rushmeier H. Dominant Texture and Diffusion Distance Manifolds. EUROGRAPHICS, 2009, 28(2):667-676.
- 7 Fawcett T. An introduction to ROC analysis. Pattern Recognition Letters, 2006,27:861-874.
- 8 Ling HB. Image DataSet. <http://www.dabi.temple.edu/~hbling/data/RD-cvpr06.zip>,<http://www.dabi.temple.edu/~hbling/data/SD-cvpr06.zip>.