

基于改进 Hu 矩算法的 AGV 字符识别研究

文生平^{1,2}, 冯泽锋^{1,2}, 洪培烽^{1,2}, 张施华^{1,2}

(1. 华南理工大学 广东省高分子先进制造技术及装备重点实验室, 广州 510640;

2. 华南理工大学 聚合物成型加工工程教育部重点实验室, 广州 510640)

摘要: 视觉导引 AGV 以识别标识符为基础实现工位转换; Hu 矩算法可以提取字符的特征参数, 结合机器学习分类算法, 对特征值进行提取并分类, 其运算速度快, 但是识别准确率低; 分析其原理公式, 发现由于阶数高导致数值范围大, 数据相近度差; 基于此, 对 Hu 矩公式进行改进, 构造了新的算法; 通过对 UCI 字符数据库进行检测, 并在视觉导引 AGV 平台上进行在线测试, 验证了改进后的 Hu 矩算法可以更快地提取特征值, 并大大提高机器学习算法字符分类准确率, 说明改进后的 Hu 矩算法在视觉导引 AGV 的字符识别上具有较好的实用性。

关键词: Hu 矩; 特征参数; 字符分类; AGV

Research on AGV Character Recognition Based on Improved Hu Moment Algorithm

Wen Shengping^{1,2}, Feng Zefeng^{1,2}, Hong Peifeng^{1,2}, Zhang Shihua^{1,2}

(1. Guangdong Provincial Key Laboratory of Technique and Equipment for Macromolecular Advanced Manufacturing,

South China University of Technology, Guangzhou 510640, China; 2. Ministerial Key Laboratory of Polymer

Processing Engineering, South China University of Technology, Guangzhou 510640, China)

Abstract: Visual guided AGV realizes position conversion based on identifiers recognition. The Hu moment algorithm is able to extract the feature parameters of characters and, combined with the machine learning classification algorithms, to classify them. The operation speed of Hu moment algorithm is fast, but its recognition accuracy is low. By analyzing its principle formulas, it was found that the large numerical range is due to the high order in formulas, which lead to a poor data closeness. Based on this, the Hu moment formulas were improved and new algorithm was constructed. Through detecting the UCI character database and taking an online testing on AGV, it is proved that the improved Hu moment algorithm can extract the feature faster and improve the accuracy of character classification of machine learning algorithm, indicating that the improved Hu moment algorithm has good practicability in character recognition of visual AGV.

Keywords: Hu moment; feature parameters; character classification; AGV

0 引言

视觉导引 AGV 在运动过程中, 需要在特定的工位完成不同运动状态的转换, 如加速、减速、转弯、停车等。路径标识符以其制作简便、布置灵活、易于识别等优势, 逐渐成为视觉导引机器人普遍应用的导引方式^[1]。

字符识别是一门检测并分类字符的技术^[2], 是目前主要存在的两种标识符识别技术之一, 另一类是控制标识符识别。控制标识符通常是各类几何形状, 如矩形、圆形、梯形等, 这类标识符的识别是检测出不同标识符水平截线段宽度的变化规律; 字符标识符通常包括 0~9 和 A~Z 等字符, 在不同工位上对其赋予不同含义以实现机器人的自动导引, 在移动机器人视觉导引字符识别和智能交通的车牌识别中都得到了更广泛的应用^[3]。

标识符的识别过程步骤主要包括: 图像采集及预处理、感兴趣区域的选取、特征提取、标识符分类。

用特征提取算法对字符的变化规律进行检测, 检测结果用特征向量表示。特征提取算法主要包括传统的模板匹配法、傅立叶描述子法、自回归模型法和不变矩法等, 其中不变矩法应用最广。1962 年, M. K. Hu 教授提出了几何不变 Hu 矩, 理论推导证明了 Hu 矩具有比例不变性、旋转不变性和平移不变性^[4]。

完成字符的特征提取后, 进而对特征进行分类。基于机器学习的 SVM^[5-6]和 KNN^[7-8]分类算法是应用最广泛的分类器, 其实用性已经得到了广泛的证明。

Hu 矩算法具有提取特征快速但准确率低的特点, 但其三个几何不变性具有很高的实际过程应用价值。为此, 本文分析 Hu 矩算法原理, 对其进行改进, 结合 SVM 和 KNN 算法的快速分类, 研究一种基于 Hu 矩的机器学习算法分类器。

1 AGV 系统及图像预处理

1.1 AGV 系统

移动机器人需要通过识别并跟踪导引线, 以及识别标

收稿日期: 2019-10-16; 修回日期: 2019-11-05。

基金项目: 广州市科技计划项目(201802010008)。

作者简介: 文生平(1966-), 男, 湖北荆州人, 工学博士, 教授, 主要从事工业装备的智能控制与机器视觉方向的研究。

识符以获得控制指令来实现自主导航，本文的实验对象是广州市井源机电设备有限公司设计和制备的 AGV，其实物图如图 1 所示。



图 1 移动机器人

移动机器人系统包括了机器人本体、驱动模块、车轮系统、控制模块、光源模块、视觉信息采集模块、图像处理单元、无线客户端、电源装置和安全防护装置等。移动机器人在导航行进的过程中在光源的辅助下由相机拍摄采集图像，图像经过处理后将提取到的图像信息转换为辅助决策信息，再输入到控制模块中以获得控制决策，将控制指令输出给驱动模块，改变驱动电机的转速，进而改变机器人驱动轮的状态，以控制机器人的移动速度、方向以及运行状态。

1.2 图像预处理

移动机器人通过识别并跟踪导航线可以实现指定轨迹的运动，但是光有导航线不足以实现指令下达与复杂任务分派。所以还需要在导航线上配置标识符，以标识符作为指令的载体，在机器人识别到特定的标识符时，能够按照标识符所代表的指令进行动作，实现更加复杂的功能。图像处理算法流程如图 2 所示。

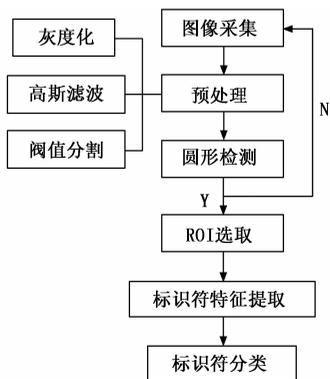


图 2 图像处理算法流程图

移动机器人实时采集到的图像为 RGB 三通道彩色图像，每个像素点的取值范围为 $256 \times 256 \times 256$ ，数据量大，直接对其进行特征提取的处理速度慢，不利于实时检测的需求。由于应用场景的图像颜色需求低，因此对图像进行灰度化处理，在保留图像信息的同时，将三通道图像降为单通道。本文使用加权法灰度化，其根据人眼对红绿蓝的感光程度对图像降维，经过加权法灰度化后得到的图像亮

度比较均匀，计算公式为：

$$Y = 0.3 * R + 0.6 * G + 0.1 * B \quad (1)$$

为消除采集图像上的噪声，需对其进行滤波处理。常用的滤波处理包括中值滤波、均值滤波、方框滤波和高斯滤波。中值滤波以像素点邻域 3×3 区域的像素值的中值作为该点的像素值；均值滤波以像素点邻域 3×3 区域的所有像素值的均值作为该点的像素值；方框滤波在均值滤波的基础上对每个像素点进行加权平均；高斯滤波相当于将图像与正态分布函数做卷积，其高斯核函数如下：

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

考虑到本研究场景中 AGV 的导航需要对导航带的边缘线和细节特征进行特征提取，因此对导航带图像的边缘信息和细节的完整性和清晰度有较高的要求。因此采用高斯滤波作为本研究的图像滤波方法。

为消除采集图像中存在的大量无用背景信息，需要对其进行阈值分割处理。阈值分割的思想是将图像上像素值大于给定阈值的像素点归类。以 $f(x,y)$ 表示灰度图，最佳分割阈值为 T ，则分割结果如下：

$$g(x,y) = \begin{cases} 255, & f(x,y) > T \\ 0, & f(x,y) \leq T \end{cases} \quad (3)$$

AGV 系统拍摄图像过程在光源的辅助下，使得目标部分和背景部分在稳定的光照下能够比较好地区分开，可以使用基于全局的阈值分割方法。本文采用 OTSU 阈值分割法^[9]，其基本思想是最大化类间方差，且能够自动确定阈值，可以很好的保留标识符部分的信息。

移动机器人运行过程采集的图片及其经过加权灰度化、高斯滤波、OTSU 阈值分割的效果如图 3 所示。

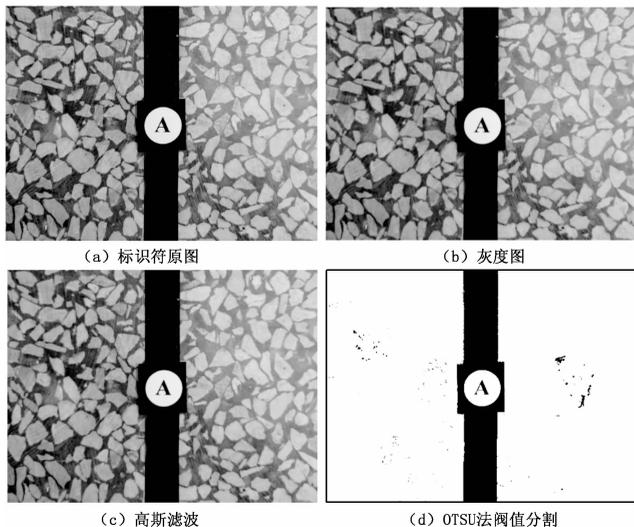


图 3 图像预处理

由于在整张图像中标识符占的比例较小，机器学习算法对小目标的处理精度存在不足。因此对标识符进行识别前，需要将标识符的位置大致定位且对其分割，再通过机器学习算法来实现分类识别。根据上述标识符图像的特点，

利用 Hough 圆变换算法^[10], 以快速定位圆心位置, 然后通过设置合适的偏移量获取包含标识符在内的矩形 ROI (感兴趣区域), 最后将其从整个图像中分割出来, 其算法原理如图 4 所示。

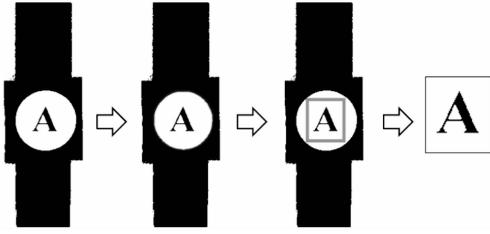


图 4 ROI 区域选取

2 改进 Hu 矩算法简介

2.1 Hu 矩^[10]简介

灰度图像的分布可以用连续函数 $f(x, y)$ 来表示, 图像的 $(p + q)$ 阶二维原点矩则表示为:

$$m_{pq} = \iint_{-\infty}^{+\infty} x^p y^q f(x, y) dx dy \quad p, q = 0, 1, 2, \dots \quad (4)$$

当图像的比例发生变化, 矩不变; 当图像进行旋转或者平移时, 随之改变。图像矩心不受像素变化的影响, 利用矩心构造出具有比例和平移不变性的中心矩。

$(p + q)$ 阶中心矩表示为:

$$\mu_{pq} = \iint_{-\infty}^{+\infty} (x - x_0)^p (y - y_0)^q f(x, y) dx dy \quad (5)$$

其中: (x_0, y_0) 为矩心, 其公式表示为 $x_0 = \frac{m_{10}}{m_{00}}$, $y_0 = \frac{m_{01}}{m_{00}}$ 。

对中心矩归一化处理, 得到具有旋转不变性的矩:

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_r}, r = \frac{p + q + 2}{2} \quad (6)$$

M. K. Hu 利用二阶和三阶的归一化中心矩构造了七个不变矩, 证明在连续图像条件下得到的特征矩可以保持缩放、旋转和平移的不变性, 七个不变矩公式如式 (7) ~ (13) 所示:

$$M_1 = \eta_{20} + \eta_{02} \quad (7)$$

$$M_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \quad (8)$$

$$M_3 = (\eta_{20} - 3\eta_{12})^2 + 3(\eta_{21} + \eta_{03})^2 \quad (9)$$

$$M_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \quad (10)$$

$$M_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \quad (11)$$

$$M_6 = (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \quad (12)$$

$$M_7 = (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{12} - \eta_{30})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \quad (13)$$

2.2 改进 Hu 矩算法

Hu 矩提取标识符特征速度非常快, 但是准确率不高。对 Hu 矩原理公式 (7) ~ (13) 进行分析, 可以发现, M_1

~ M_4 阶数低, M_5 ~ M_7 阶数高。由数学原理可知, 阶数越高, 数值的变化范围越大。为了观察 M_1 ~ M_7 的变化情况, 用如图 5 所示的 9 张背景、字体大小、形状和方向都不一样的数字 1 来测试。

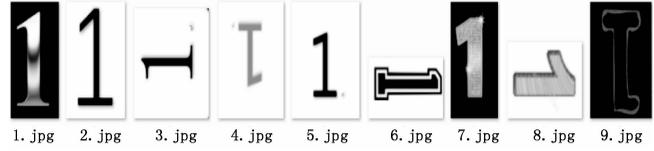


图 5 数字 1 的测试集

按以下步骤进行 Hu 矩测试实验:

- 1) 读取图像;
- 2) 对图像进行灰度化处理、高斯滤波去噪和 OTSU 分割处理;
- 3) 使用 Hough 快速圆变换检测标识符的定位圆;
- 4) 分割出举行 ROI 区域;
- 5) 用 Hu 矩算法检测标识符的特征属性, 得到特征向量

利用处理得到的特征向量数据生成对数图形, 如图 6 所示是特征值曲线:

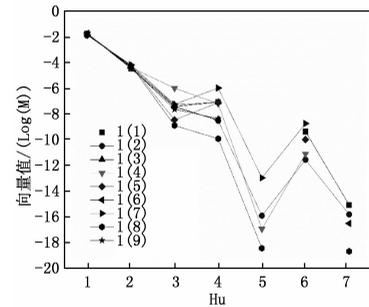


图 6 特征参数图形

分析图 6, 发现 9 个样本的 M_1 、 M_2 基本相等, M_3 、 M_4 虽然有点分散, 也集中在一定范围。但是从第五个参数开始, 跳动性大, 有的数据已经不落在图形范围里, 说明从第五个参数开始, 参数不稳定。10 组测试集的情况相似, 可以得出的结论是, Hu 矩前四个参数的可靠性好, 后三个参数不稳定。研究 7 个参数, 发现 M_1 、 M_2 的次幂较低, M_3 、 M_4 的次幂次之, M_5 、 M_6 、 M_7 的次幂较高, 根据数学原理, 次幂越高, 数据变化范围越大, 所以猜测特征数值的跳动性与参数的次幂高低有直接的关系。基于此, 对 Hu 矩进行改进, 参考 M_1 、 M_2 公式的特点, 改造 M_3 、 M_4 , 旨在得到可靠性更好的特征参数。

反复试验, 最后确定的特征参数如公式 (14) ~ (17) 所示。

$$M_1^* = \eta_{20} + \eta_{02} \quad (14)$$

$$M_2^* = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \quad (15)$$

$$M_3^* = 3(\eta_{21} + \eta_{03})^2 + 4\eta_{11}^2 \quad (16)$$

$$M_4^* = (\eta_{30} + \eta_{12})^2 + \eta_{20} + \eta_{02} \quad (17)$$

对上述测试集进行测试，得到的特征曲线如图 7 所示，四个特征值都更加稳定，特征数据的方差变小。

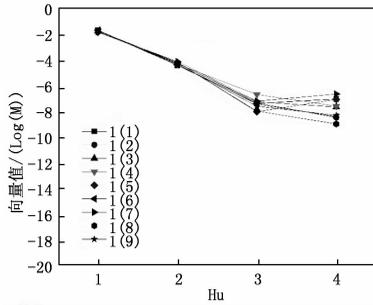


图 7 改进的特征参数图形

3 改进 Hu 矩算法性能验证

3.1 离线运算速度测试

为了验证改进 Hu 矩算法的提取速度，对 UCI 机器学习数据库的 2 520 张图像进行测试，2 520 张图像由 0~9 和 A~Z 的 36 个字符各 70 张组成。先后用原 Hu 矩算法和改进的 Hu 矩算法分别测试，计算时间如表 1 所示。

表 1 特征提取算法处理结果

算法	总处理时间/s	平均时间/ms
Hu 矩	12124	4.81
改进 Hu 矩	6274	2.49

表 1 的实验结果表明，改进后的 Hu 矩算法明显比原 Hu 矩算法运算时间短，一张图像的处理时间几乎节省了一半，在大规模的图像样本中，改进后的算法提取速度更快。

3.2 离线提取精度测试

为了验证改进 Hu 矩算法的提取精度，用 C-SVC 和 KNN 分类算法对其提取后的特征进行分类。

考虑 Hu 矩的特征参数远少于样本数量，按照经验法，C-SVC 分类算法的核函数选择高斯型。本文采用阎晓娜等人修正后的高斯核函数，如式 (17) 所示，此修正后的高斯核函数被证明可以很好地保留数据鲁棒性强的优点和弥补处理时间长的不足^[11]。

$$K(x_i, x_j) = G \cdot \left[\exp\left(\frac{\sigma^2}{\|x_i - x_j\| + \gamma^2}\right) + \lambda \right] \quad (18)$$

式中，G 是幅度调节参数，σ 是带宽变量，γ 是位移参数，λ 是微调变量。

对于 2 520 个样本集，其中 1 800 个样本作为训练集，用于生成 C-SVC 分类器；360 个样本作为验证集，对已经生成的 C-SVC 分类器反复调参，用于参数的调整，最终确定的高斯核函数参数是 G = 2.1, σ = 0.3, γ = 0.2, λ = 0.75；剩下 360 个样本作为测试集，用于测试分类器性能，其测试结果如表 2 所示。

表 2 C-SVC 分类结果对比

算法	检出数量	真实数量	准确率/%
Hu 矩	360	323	89.72
改进 Hu 矩	360	358	99.44

KNN 分类算法的关键是 K 值的选择，一般取小于 18，没有固定理论指导最优 K 值的选择。多次取值，得到的测试结果如表 3 和表 4 所示。

表 3 原 Hu 矩的 KNN 分类结果对比

K 值	检出数量	真实数量	准确率/%
1	360	270	75
3	360	275	76.39
5	360	278	77.22
7	360	279	77.5
9	360	288	80
11	360	291	80.83
13	360	289	80.27
15	360	281	78.06
17	360	278	77.22

表 4 改进 Hu 矩的 KNN 分类结果对比

K 值	检出数量	真实数量	准确率/%
1	360	319	88.6
3	360	323	89.72
5	360	327	90.83
7	360	331	91.94
9	360	337	93.61
11	360	337	93.61
13	360	335	93.05
15	360	331	91.94
17	360	328	91.11

表 2~4 的实验结果表明，无论是用 C-SVC 分类还是 KNN 分类，改进后的 Hu 矩算法的检测准确率都明显比原 Hu 矩算法的高，说明改进后的 Hu 矩算法提取的特征值更有利于分类，其提取精度更高。

3.3 移动机器人实时在线测试

布置移动机器人的黑色导引带及标识符如图 8 所示，每个标识符之间相隔越 2 m。移动机器人每次从数字 0 与字母 Z 这段导航带的中点开始运行，环绕引导带行进回到起点结束。

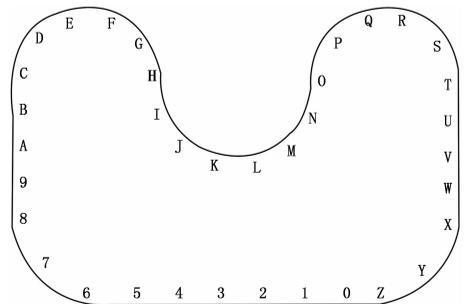


图 8 导航路线布置示意图

特征提取算法先后采用 Hu 矩算法及改进的 Hu 矩算法，分类器先后采用 C-SVC 分类器和 KNN 分类器。分别让机器人以 20 m/min、30 m/min、40 m/min、50 m/min 的

(下转第 251 页)