

中华绒螯蟹背壳特征识别及其算法研究*

邵伟鹏^{1,2}, 李浩¹, 张炳良³, 王春^{4**}

(1. 安徽工业大学计算机科学与技术学院, 安徽 马鞍山 243000; 2. 安徽工业大学信息技术研究院, 安徽 马鞍山 243000;
3. 天邦食品股份有限公司, 浙江 余姚 315400; 4. 上海海洋大学农业农村部淡水水产种质资源重点实验室,
水产科学国家级实验教学示范中心, 上海水产养殖工程技术研究中心, 上海 201306)

摘要: 针对现阶段中华绒螯蟹(*Eriocheir sinensis*)防伪技术准确性差和可靠性低的问题, 本文提出一种基于图像匹配的蟹背壳特征识别算法。首先, 分割出具有唯一性的中华绒螯蟹背壳图像, 用 SURF 算法提取特征点, 然后使用双向 FLANN 算法与数据库中保存的特征点匹配, 以匹配点与最近欧式距离和次近欧式距离匹配点的位置关系作为误匹配检测的依据, 采用正确匹配对的占比验证其相似度, 以此实现河蟹单体的识别。研究表明, 在近似的拍摄条件下, 两幅属于同一只蟹的背壳图像相似度可达到 0.7 以上, 而属于不同蟹的背壳图像的相似度则偏低。本研究可有效解决目前使用绑在蟹足上的二维码或条形码等标识物作为中华绒螯蟹商品蟹溯源及其防伪技术准确性和可靠性差的问题, 为商品蟹的单体识别与精准溯源提供一项技术手段。

关键词: 中华绒螯蟹; 背壳; 特征识别; 图像处理; 相似度; 防伪查询

中图分类号: S985.2⁺2; TP391.4

文献标志码: A

文章编号: 1672-5174(2021)01-138-09

DOI: 10.16441/j.cnki.hdxsb.20180416

引用格式: 邵伟鹏, 李浩, 张炳良, 等. 中华绒螯蟹背壳特征识别及其算法研究[J]. 中国海洋大学学报(自然科学版), 2021, 51(1): 138-146.

TAI Wei-Peng, LI Hao, ZHANG Bing-Liang, et al. Research on the feature recognition and algorithm of the carapace of *Eriocheir sinensis*[J]. Periodical of Ocean University of China, 2021, 51(1): 138-146.

中华绒螯蟹(*Eriocheir sinensis*)又称河蟹、大闸蟹, 是中国淡水养殖业中的珍品, 至 2017 年底, 产业规模已超过 75 万 t^[1]。由于各地气候、水质、土壤等特征和使用的河蟹养殖方法的不同, 造成各地养殖的河蟹在品质上存在较大的差别, 也导致商品蟹的销售和利润受到了直接的影响。随着消费者对食品安全日益重视和产业的深度发展, 各产区陆续建设自己的商业品牌或以地理位置作为销售标志, 如远近闻名的“固城湖大闸蟹”和“阳澄湖大闸蟹”等, 以提高其河蟹的销售优势。出于防止产品被假冒以及对消费者权益和自身品牌保护的目的, 找出一种可以区分河蟹产地的方法是地方企业和政府一直想做的事情^[2]。目前, 用于商品蟹溯源和防伪的方法是在河蟹的蟹足上绑缚防伪标识, 如条形码、二维码等, 通过验证标识来实现防伪^[3,4]。周晓庆等曾开发了一套基于 QR 码防伪技术

和加密技术的防伪系统, 以“一物一码”的形式部分解决了商品防伪和产品溯源等问题^[5]。邓洋等就当前防伪技术的原理, 在数字防伪原理和典型数字防伪技术的基础上, 提出一种二维码和 RFID 防伪技术的组合应用^[6]。李清扬等对防伪原理以及防伪技术的发展趋势进行了总结, 其中认为信息网络防伪技术适用于河蟹的防伪, 即在河蟹身体上绑缚防伪标识, 消费者购得商品后使用终端, 将防伪标识信息传至验证端进行验证^[7]。然而以上方法验证的是防伪标识, 并非河蟹本体, 制假者仍可以设法将假品冒充真品查询后卖出, 造成真假产品混淆。张炳良等提出一种二维码和河蟹甲壳图像相结合的双重防伪体系, 但是仅提出了思路, 并未给出结果和验证^[8]。研究表明, 蟹壳表面分布着丰富的凸起、沟、脊、纹理等性状, 形成了河蟹的种质图案^[9]。虽然不同品种的河蟹背壳图案具有异质性且相

* 基金项目: 国家自然科学基金青年基金项目(61602007); 安徽省自然科学基金青年基金项目(1608085QF132); 安徽省自然科学基金项目(KJ2018ZD007); 上海市现代农业产业技术体系项目(沪农科产字(2018)第4号); 上海高校水产一流学科建设项目资助

Supported by the National Natural Science Foundation of China(61602007); the Anhui Province Natural Science Foundation for Youths(1608085QF132); the Natural Science Foundation of Anhui Province of China(KJ2018ZD007); the Agriculture Research System of Shanghai, China(Chinese Characters of Shanghai Agricultural Sciences(2018) No.4); the Shanghai University Fisheries First-Class Discipline Construction Projects

收稿日期: 2019-04-06; 修订日期: 2019-06-07

作者简介: 邵伟鹏(1979-), 男, 博士, 副教授, 主要研究方向: 物联网技术, 数字图像处理。E-mail: taiweipeng@ahut.edu.cn

** 通讯作者: E-mail: cwang@shou.edu.cn

对稳定^[10],但是按照适应生物学的观点:栖息环境的异质性会造成同一基因型的个体在表型特征方面存在明显的变化,物种性状的表达或改变常常受到其栖息环境的影响^[11-14],因此,理论上,生活或养殖在不同环境下的河蟹,其背壳上或多或少会留下其生活的痕迹,这为识别不同地区的河蟹提供了新的线索。本文通过数字图像处理技术,利用河蟹背壳所呈现的特征,提出一种基于 SURF(Speeded up robust features, SURF)和 FLANN(Fast library for approximate nearest neighbors, FLANN) 算法的方案,力求证明河蟹单体背壳图案具有唯一性,为河蟹个体识别技术的研制提供新的思路和方法,为真正实现河蟹品质溯源、防伪和品牌价值提升贡献参考资料。

1 材料与方方法

1.1 实验材料

实验材料于 2017 年 12 月 3 日购自安徽省马鞍山市当涂县塘南水产批发市场,为养殖的中华绒螯蟹商品蟹。雌蟹 65 只,壳宽 80.0~110.0 mm,壳长 75.0~100.0 mm,体质量 140.0~160.0 g;雄蟹 65 只,壳宽 98.0~130.0 mm,壳长 95.0~125.0 mm,体质量 190.0~205.0 g。每只河蟹以流水号 001~130 进行编号,然后统一将编号牌绑定在河蟹右侧的步足上。

1.2 中华绒螯蟹背壳图像采集及图像库建立

从上述 130 只实验河蟹中随机选择 100 只,使用专用数据采集器^[8]采集图像。每只河蟹采集 20 幅图像,采用河蟹编号和采集次序的组合作为图像文件的命名方式,总共采集 2 000 幅图像,构成样本图像库(简称“样本库”)。采集时使用白色塑胶板作为背景板,并要求背壳没有污物和水渍,图像分辨率为 1 920×2 560。

用多种像素和多种型号的手机对 130 只实验蟹进行拍摄,拍摄时要求避免强光直射,蟹背部干燥且干净,同时拍摄的图像不可模糊。分别对每只河蟹拍摄 5 张图像,采用与样本库图像文件同样的命名方式对图像文件命名,共计 650 张图像,构成测试图像库(简称“测试库”)。

1.3 中华绒螯蟹背壳图像分割

河蟹背壳图像分割的流程图如图 1 所示。首先将采集到的图像转为灰度图,使用高斯滤波进行降噪处理,利用 sobel 算子获取图像的边缘信息^[15]。通过阈值分割,过滤一些细节信息,使蟹背部边缘轮廓更清晰。然后做灰度反转获得掩膜 mask1,同时做膨胀、空洞填充和腐蚀运算获得掩膜 mask2。两个掩膜做交运算,分割开前景与背景。腐蚀分割出来的前景图像,使蟹背壳与步足分离,然后空洞填充。根据连通域的大小筛选出背部区域,获得掩膜 mask3,对掩膜 mask3 做膨胀运算,让其恢复到原来的尺寸。最后与原灰度图

相交,即可获得河蟹背部的灰度图(见图 2)。

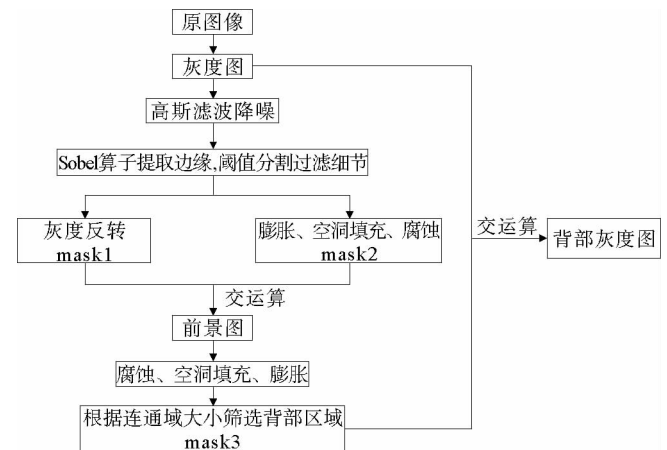


图 1 中华绒螯蟹甲壳图像分割流程图

Fig.1 Flow diagram of image segmentation on the carapace of *Eriocheir sinensis*

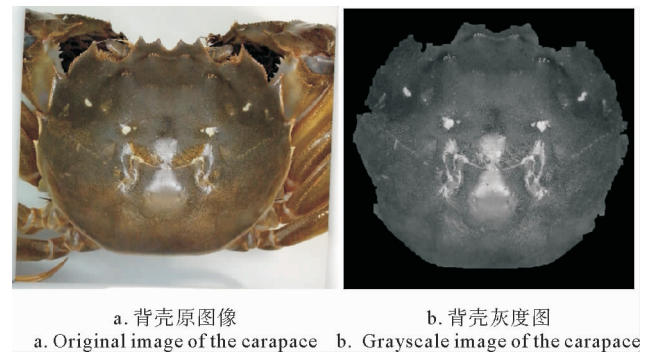


图 2 中华绒螯蟹甲壳图像分割

Fig.2 Image segmentation on the carapace of *Eriocheir sinensis*

1.4 基于 SURF 算法的中华绒螯蟹背壳特征提取

分布在河蟹背壳上的凸起、沟、脊和纹理等图案的灰度值与周围区域有着明显的差异,因此形成大量的角点和斑点等局部特征点(见图 3)。考虑到采集图像时光照、角度等因素的变化,本研究利用对图像旋转、平移、缩放和噪声等具有较好鲁棒性的 SURF 算法^[16],

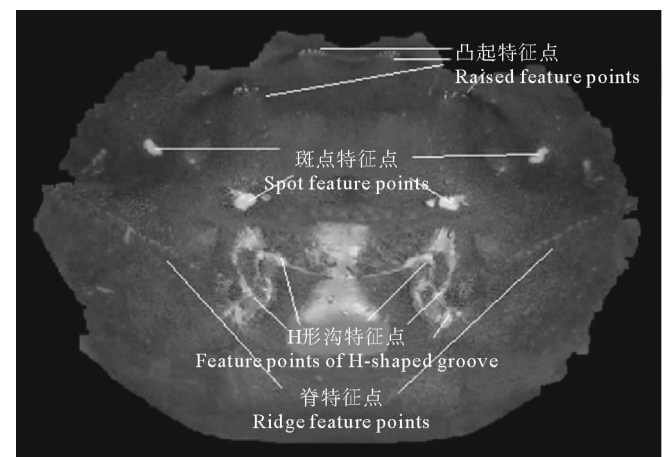


图 3 中华绒螯蟹甲壳图像特征点

Fig.3 Feature points on the carapace of *Eriocheir sinensis*

来提取河蟹背壳图像的特征。SURF 算法采用 Hessian 矩阵检测特征点^[17]。在不同尺度上检测特征点来实现尺度不变性,同时给每个特征点分配主方向以确保特征具有旋转不变性。

1.5 基于双向 FLANN 算法的特征匹配

由于使用的分割背壳图像的算法会丢失一部分边

缘信息,而边缘中有丰富的特征点。为了减少不必要的误匹配对,需要对 SURF 算法提取到的特征点做一些预处理,删除掉离背部边缘较近的一些特征点,如图 4 所示。

1.5.1 双向 FLANN 算法 利用双向 FLANN 算法获取准确可靠的匹配点对^[16]。假设两幅图像分别用

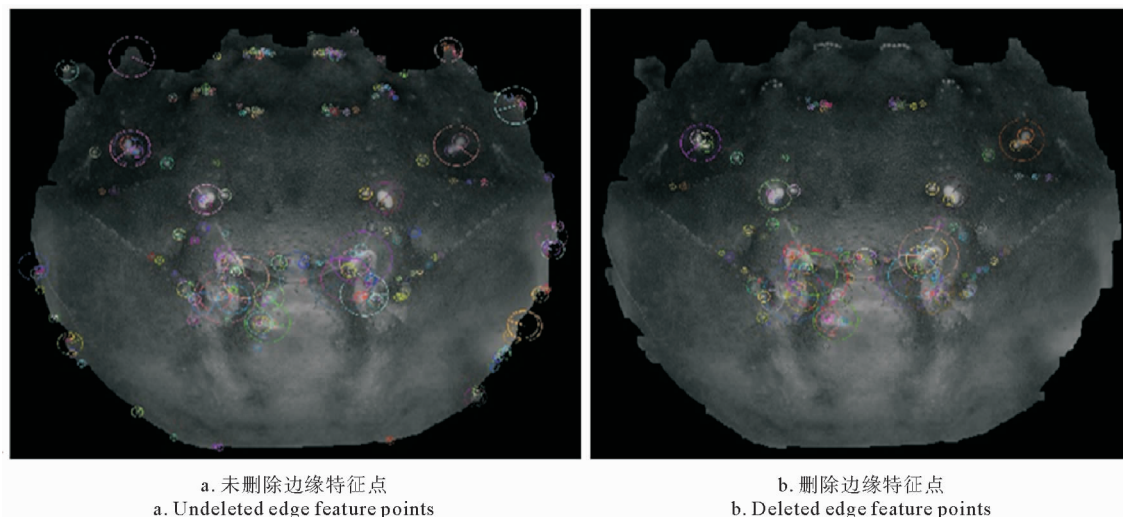


图 4 中华绒螯蟹甲壳删除边缘特征点对比图

Fig.4 Contrast images of deleting edge feature points on the carapace of *Eriocheir sinensis*

I_1 和 I_2 表示,双向 FLANN 算法具体流程如下:

(1)通过 FLANN 算法找到图像 I_1 中的特征点 m_1 在图像 I_2 中的匹配点 m_2 ,记作匹配对 (m_1, m_2) 。然后用同样的方法找到图像 I_2 中的特征点 m_2 在图像 I_1 中的匹配点 m_3 ,记作匹配对 (m_2, m_3) 。

(2)判断特征点 m_2 在不同方向的匹配过程中找到的对应特征点 m_1, m_3 是否相同。若特征点 m_1, m_3 是图像 I_1 中的同一特征点,则判定匹配结果正确。反之则认为误匹配。

1.5.2 匹配点对正误检测及相似度计算 由于图像之间是刚性变换,对于所有正确的匹配点对,任意两个匹配点对的距离比应该是相等的或者近似相等,而且距离比值也应该与两幅图像的真实尺度比近似相等^[18]。因此,本文采用一种以最近欧式距离点和次近欧式距离点构建矢量坐标系,判断匹配对在该坐标系上的位置关系是否相似的检测方法检测正确匹配对。对匹配成功的匹配对按照欧式距离从小到大排序,取出最近欧式距离和次近欧式距离匹配对。计算所有特征点到最近欧式距离点的距离与最近欧式距离到次近欧式距离点的距离的比例。计算每个特征点在以最近欧式距离点和次近欧式距离点构建的矢量坐标系上的角度。判断距离比例和角度差值,如果分别小于一定的差值,则判定匹配点对为正确匹配对。以正确匹配对在全部分匹配对中的占比衡量两幅图像的相似度。

1.6 不同相似度条件下正确匹配与漏匹配之间的关系

从前述样本库中选取 20 只蟹的图像,采用不同相似度阈值进行匹配试验,统计正确匹配对和漏匹配对个数,拟合其相关关系。其中,“正确匹配”是指匹配结果中正确匹配数目占总匹配数目的平均比例,“漏匹配”是指本属于同一只蟹的图像却没有归于匹配结果的数目占属于该蟹的图像数目的平均比例。

1.7 相似度验证实验

从上述样本库中以随机抽样的方式选出一些图像,其中有属于同一只蟹的和不属于同一只蟹的图像。计算选取图像间的相似度并统计。“相似度”或“相似度阈值”是“正确匹配对数”与“筛选后匹配对数”的比率(值),“筛选后匹配对数”是采用双向 FLANN 算法得到匹配对数,按照欧式距离从小到大排序,筛选排名前 80%的匹配对数后的数目。本研究通过寻找合适的相似度阈值来确定两幅图像的相似度。若匹配对数较少,则认定两幅图像相似度为 0;若相似度大于某一阈值,则判定两幅图像属于同一只蟹。

2 结果与分析

2.1 滤除错误匹配点正确率曲线

图 5 为采用不同的差值分别使用距离比例和角度两种约束滤除错误匹配点正确率的对比分析。由图 5 可以发现,当差值取值在 0~0.2 内,使用距离比例和

角度两种约束,滤除错误匹配点的正确率是 1,表明两种约束的差值取值在此范围内可有效进行错误匹配对的滤除。当差值在 0.2~1 范围内时,使用距离比例和角度两种约束滤除错误匹配对的效果较差。由图 5 可知,当距离比例和角度的差值为 0.3 时,滤除错误匹配点的正确率才能到 0.9 及以上。但是如果拍摄角度差异较大,两幅属于同一只蟹的图像使用该范围内的差值滤除错误匹配对,会存在正确匹配对被误判为错误匹配对的情况。

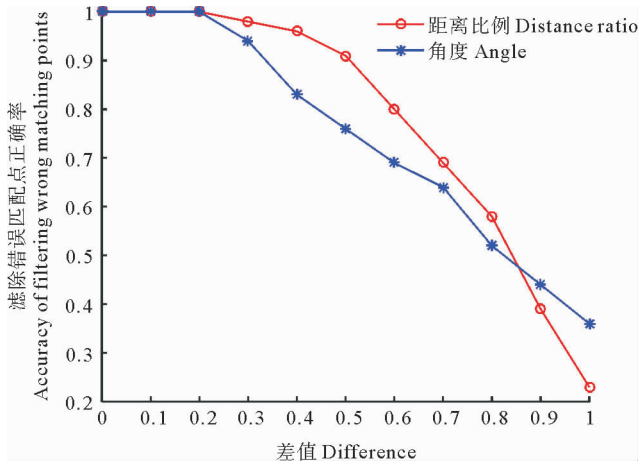


图 5 滤除错误匹配点正确率曲线

Fig.5 Correct rate curves of filtering wrong matching points

2.2 滤除正确匹配点曲线及两种约束条件差值界限的判定

图 6 是在不同约束条件、不同差值下滤除正确匹配点的情况。当距离比例的差值取值范围为 0~0.5 和角度的差值取值在 0~0.2 之间时,有部分正确匹配对被误判为错误匹配对;然而距离比例差值在 0.5~1 范围内和角度差值在 0.2~1 范围内时,滤除正确匹配点的数目为 0,即说明没有误判的情况。显然,

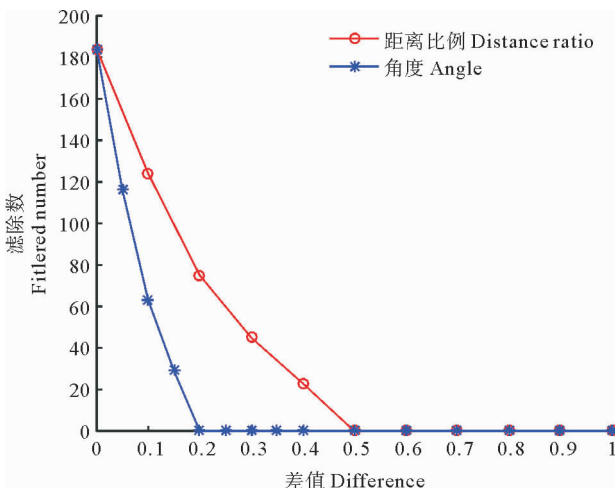


图 6 滤除正确匹配点曲线

Fig.6 Curves of filtering correct matching points

本文算法中距离比例的差值的判定界限选为 0.5,角度差值选为 0.2。

2.3 不同相似度条件下正确匹配与漏匹配之间的关系

图 7 所呈现的是从样本库中选取 20 只蟹,采用不同的相似度阈值进行匹配的试验结果。由图 7 可知,当相似度阈值在 0~0.5 范围内,正确匹配占比较小,表明误判的情况较为严重。阈值在 0.7~1 范围内时,正确匹配占比高于 90%。但与漏匹配比例曲线(虚线)结合分析,当阈值在 0.8~1 范围内,漏匹配的情况较为严重,即本算法并没有将本属于同一只蟹的图像归于到匹配结果中。漏匹配比例在相似度阈值取值在 0~0.7 内时低于 10%,在 0.8~1 范围内时却能达到 60%以上。因此,经过综合分析,本研究相似度阈值选为 0.7。

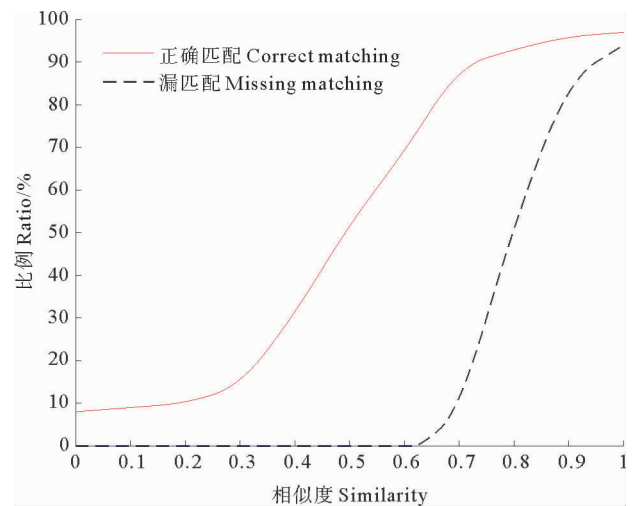


图 7 不同相似度条件下正确匹配与漏匹配试验结果曲线

Fig.7 Test results curves for correct and missing match with different similarities

2.4 相似度验证实验结果

以随机抽样的方式从样本库中选取部分图像(其中有属于同一只蟹的和不属于同一只蟹的图像),并计算其相似度,统计实验结果见表 1。根据表 1 中单向 FLANN 匹配数,会发现单向 FLANN 算法是在匹配图像中的所有特征点中找到与待匹配图像中特征点相似的特征点。因此会存在一对多的匹配情况。如序号 2 的实验,蟹 1 为待匹配图像而蟹 2 为匹配图像时,匹配对数等于蟹 1 图像的特征点数,相反方向匹配时匹配对数则等于蟹 2 图像的特征点数。而匹配对数在采用双向 FLANN 算法后却只有 49。所以可以通过双向 FLANN 算法剔除错误匹配对。属于同一只蟹的背壳图像计算的相似度(正确匹配数占筛选后匹配数的比例)较大。如序号 3 为 0.913,序号 6 为 0.778。而不同蟹的背壳图像计算的相似度较小。如序号 1 为 0.164,序号 5 为 0.143。由实验结果可以发现,本文算法中的相似度可以作为判断两幅图像是否属于同只蟹的依据。

表 1 部分图像相似度结果
Table 1 The similarity of partial images

序号 ^①	蟹 1 的 SURF 特 征点数 ^②	蟹 2 的 SURF 特 征点数 ^③	单向 FLANN 匹配数 ^④		双向 FLANN 匹配数 ^⑤	筛选后 匹配数 ^⑥	正确 匹配数 ^⑦	相似度 ^⑧	匹配结果 (同一只蟹) ^⑨	真实结果 (同一只蟹) ^⑩
			1→2	2→1						
1	294	325	294	325	137	110	18	0.164	否 No	否 No
2	297	322	297	322	49	39	3	0.077	否 No	否 No
3	213	194	213	194	57	46	42	0.913	是 Yes	是 Yes
4	189	253	189	253	80	64	58	0.906	是 Yes	是 Yes
5	342	276	342	276	35	28	4	0.143	否 No	否 No
6	169	189	169	189	45	36	28	0.778	是 Yes	是 Yes
7	261	220	261	220	11	9	2	0.222	否 No	否 No
8	391	148	391	148	111	89	74	0.832	是 Yes	是 Yes
9	148	284	148	284	38	31	8	0.258	否 No	否 No
10	275	360	275	360	120	96	4	0.042	否 No	否 No

Notes:①Sequence number; ②Number of first crab’s SURF feature points; ③Number of second crab’s SURF feature points; ④Number of unidirectional FLANN matches; ⑤Number of bidirectional FLANN matches; ⑥Number of matches after filtering; ⑦Number of correct matches; ⑧Similarity; ⑨Matching result(the same crab); ⑩Real result(the same crab).

2.5 样本库中同一只蟹图像的一致性检验

从样本库中随机挑选 12 只蟹,采用本文算法从样本库中筛选出与每只蟹对应的 20 幅图像的相似度在 0.7 以上的图像(不包括该图像)进行一致性检验,汇总结果见表 2。由表 2 发现,大部分河蟹图像可以从挑选

的图像库中正确找到属于该蟹的其他 19 幅图像。但是对于一小部分蟹,如蟹 22 之第 4、5 幅,有漏匹配的情况。由表 2 中计算的一致性均能达到 96.0%以上,即表明样本库里同一只蟹的图像基本都能找到属于该蟹的其他图像。

表 2 样本库图像一致性检验
Table 2 Consistency check of sample library images

河蟹编号 Crab SN	图像编号 Image index																				一致性/% Consistency
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
7	18	19	19	19	19	19	17	19	19	19	19	18	18	19	19	19	19	19	19	19	98.68
19	19	19	18	19	19	18	18	18	17	18	18	17	19	19	18	17	19	19	18	19	96.32
22	19	19	19	18	17	19	18	19	19	19	18	19	19	19	19	18	19	19	19	19	98.42
29	19	19	17	18	19	18	17	19	18	19	19	19	18	18	19	19	19	18	18	19	97.11
35	18	19	18	17	19	19	18	19	18	19	19	19	19	18	19	19	19	18	19	18	97.63
42	18	18	19	18	19	18	19	19	19	19	19	19	18	19	19	18	19	18	19	19	98.16
58	19	18	18	18	18	19	19	19	18	18	19	19	19	18	19	19	18	19	19	19	97.89
66	19	19	17	18	18	19	18	18	19	18	19	18	19	19	19	18	19	19	19	18	97.37
75	19	18	19	19	19	19	19	19	19	19	19	18	19	19	19	19	19	19	18	19	99.21
87	18	18	18	19	19	18	18	19	18	19	18	19	18	18	18	19	19	18	19	19	97.11
93	19	19	18	19	18	18	18	17	18	19	18	18	19	18	19	19	19	18	18	18	96.58
98	18	17	18	18	19	19	19	17	19	19	18	19	19	18	18	18	19	19	18	19	96.84

注:表中,“河蟹编号”是蟹螯足上的编号,“图像编号”是采集该蟹的次序,“一致性”是指该编号蟹的 20 幅图像的平均匹配数除以 19 的百分比。
Notes:Crab SN is referred to serial number of the number card tied to the chela of crabs. Image index means sequence number on shooting a crab. Consistency is referred to the percentage of average match number divided by 19 in all of the crab’s 20 images.

2.6 测试库与样本库的匹配试验结果

表 3 所示为样本库与测试库中图像的匹配试验结果。从表 3 中可以发现,对于样本库有存储图像的河蟹的测试库图像,其匹配正确率能达到 80%以上,即说明采用本文算法可以较准确的从样本库里找到属于该蟹的图像。然而对于没有在样本库中存储图像的河蟹,如编号 26、85 以及 111 的河蟹的匹配结果数为 0,即在样本库中没有匹配到图像。根据匹配结果数与正确匹配数,可以发现使用本文算法得到的匹配结果正

确率为 100%。由于部分情况下存在漏匹配,因此会导致匹配结果数占样本库数目的比例偏低。针对漏匹配问题,通过观察发生漏匹配现象的图像,发现其主要原因是拍摄的角度存在着较大差异,导致同一只蟹的两幅图像形变差异较大,因此可以通过矫正拍摄角度等手段减少漏匹配。针对一只河蟹的溯源问题,可以对其采集多幅图像建立样本库,在匹配过程中如果样本库中的一幅或者多幅图像匹配成功,即可认定该蟹属于正品蟹。

表 3 部分测试库图像匹配结果
Table 3 Matching results of partial test library images

螃蟹编号 Crab index	匹配结果数					样本库数目	正确匹配数					正确率/%				
	Number of matching results					Number of the crab's images in sample library	Number of correct matching results					Correct rate				
	1	2	3	4	5		1	2	3	4	5	1	2	3	4	5
8	19	20	20	19	20	20	19	20	20	19	20	95	100	100	95	100
13	20	19	20	18	20	20	20	19	20	18	20	100	95	100	90	100
26	0	0	0	0	0	0	—	—	—	—	—	—	—	—	—	—
37	20	19	18	19	20	20	20	19	18	19	20	100	95	90	95	100
64	18	20	19	20	20	20	18	20	19	20	20	90	100	95	100	100
85	0	0	0	0	0	0	—	—	—	—	—	—	—	—	—	—
98	20	20	19	20	20	20	20	20	19	20	20	100	100	95	100	100
103	17	20	20	20	20	20	17	20	20	20	20	85	100	100	100	100
111	0	0	0	0	0	0	—	—	—	—	—	—	—	—	—	—
129	20	20	20	20	20	20	20	20	20	20	20	100	100	100	100	100
130	19	18	20	19	20	20	19	18	20	19	20	95	90	100	95	100

注:“样本库数目”是指样本库中属于该试验蟹图像的数目;“匹配结果数”是指属于该蟹的 5 幅测试库图像在样本库中的匹配数目。
Notes: Number of the crab's images in sample library is referred to the number of this very testing crab belonging to this sample library.
Number of matching results means the matching number in the sample library of 5 images of this testing crab belonging to the test library.

3 讨论

3.1 关于中华绒螯蟹背壳特征识别的唯一性

生物识别是依据个体之间具有较高稳定性且具有唯一性的生物特征对其进行识别与身份的认证,是生物识别技术应用于个体身份鉴定的前提^[19]。因此,要识别出某区域或某产地的河蟹,必须找到其稳定性兼具唯一性的个体特征。

一方面,像河蟹背甲上的额齿和侧齿、隆起的嵴等可数或可量性状属于物种(种质)性状,并不会因为生活场景的影响而明显减少数量,表现出较高的稳定性;另一方面,河蟹无论是在自然条件下还是在养殖环境中,喜欢攀爬、掘洞、潜伏于沙或泥中^[9],但因为个体生活轨迹的不同而出现不同程度的磨损是真实存在的。Idaszkin Y L 等采用几何形态测量法,研究了在不同栖

息环境下角突弓蟹(*Cyrtograpsus angulatus*)背壳形态的变化。研究结果表明,虽然采样点只相隔 3 km,但由于角突弓蟹栖息地的差异,其背壳特征出现了变化^[20]。因此,理论上讲,蟹个体背甲特征应该都是不一样的,即个体的特征具有唯一性。另外,螃蟹背部图案因其复杂性,无法仿造。在市场流通期间,也不会发生改变。所以,作者也有意识地选择了这些容易反映河蟹生活状况的性状来加以研究,提取相应的识别特征。

3.2 关于中华绒螯蟹背壳特征提取的算法与图像匹配

如何提取河蟹个体背壳的差异性特征,是图像匹配的关键步骤,也是提高不同河蟹个体识别度的前提。研究表明,图像特征划分为点、线和面三种。点特征提取是其中经常使用的一种图像特征提取算法^[17]。作者注意到,分布在河蟹背壳上的凸起、沟、脊和纹理等图案的灰度值与周围区域有着明显的差异,因此形成了

大量的角点和斑点等局部特征点,从而构成河蟹背壳的众多几何特征(见图3)。

常用的特征提取主要有 Harris^[21]、ORB^[22]、SIFT^[23]、SURF^[24]等算法,但都存在特征信息少和误匹配率高的问题^[18]。在特征点的匹配过程中如果采用基于最近邻的匹配算法,会由于 SURF 算法提取出来的特征向量是高维向量而导致计算量过大,从而影响算法的整体速度^[24],但是考虑到该算法对图像旋转、平移、缩放和噪声等具有较好鲁棒性^[24],本研究还是选择此算法来提取河蟹背壳图像的特征。

图像特征提取与匹配一直是计算机视觉中的关键问题,在靶标检测、物体或生物个体识别、图像配准等实际应用中发挥着重要作用^[18]。Muja 和 Lowe 提出一种快速近似最近邻查找(Fast library for approximate nearest neighbors, FLANN)搜索的图像匹配算法,此算法使用多重随机 KD-TREE 和分层 K-means 树,能够根据用户采用的数据集和欲达到的精确度确定参数值和最佳算法,使搜索速率获得显著的提高^[25]。冯亦东等研究表明,基于 SURF 特征提取和 FLANN 搜索的图像匹配算法,能解决目前其它图像特征匹配算法中存在的图像提取特征量少、特征误匹配率高和匹配速度慢的问题^[16,26]。本研究依据 SURF 特征提取和 FLANN 搜索的图像匹配算法,通过对河蟹背壳某些特征点的删减,实现了两幅图像之间的良好匹配,即使图像间出现噪声、平移、尺度缩放、亮度变化、遮挡、旋转等情况。

3.3 关于不同相似度条件下正确匹配与漏匹配之间的关系

“相似度”或“相似度阈值”是“正确匹配对数”与“筛选后匹配对数”的比率(值)。图5、6分别是滤除错误匹配点正确率曲线以及滤除正确匹配点曲线,与图7中使用不同相似度时正确匹配和漏匹配的实验曲线结合分析,发现在相似度采用不同阈值时出现了所谓的“漏匹配”现象。作者分析,可能由于采集河蟹背壳图像时拍摄的角度过于倾斜、有强光照射等环境因素的影响,导致属于同一只蟹的图像呈现较大的差异,从而出现“漏匹配”。因此,相似度取值越大,判断两幅图像属于同一只蟹的标准就越高,极端情况是相似度为1时,两幅完全一样的图像才能匹配成功。由图7可以发现,相似度取值在0~0.5范围内,正确匹配所占的比例偏小,即存在误判不同蟹图像属于同一只蟹的情况。经过大量的对比研究,本实验结果得出相似度阈值为0.7,即两幅蟹壳图像相似度大于等于此阈值,即判断其为同一只蟹。反之则为不同蟹。作者相信,随着实验的深入及样本数量的不断提高,相似度阈值将更为精确。另外,对于使用不同型号和不同像素的手

机拍摄的河蟹图像,由表3的匹配结果可以发现,无法在已建立的样本库中匹配到结果。因此,对于市场中没有录入数据库的假冒河蟹,通过使用本文算法在数据库中匹配,可以有效的识别出该蟹是否为正宗的品牌蟹。

4 结语

本文采用 SURF 算法进行中华绒螯蟹甲壳图像特征点的提取,然后使用双向 FLANN 算法完成特征点的匹配,最后以每个匹配成功的特征点基于最近欧式距离和次近欧式距离匹配点的位置关系作为检测匹配对正误的依据,以正确匹配对数占总匹配对数的比例来判断两幅图像是否属于同一只蟹。以100只中华绒螯蟹的图像建立样本库,并用不同手机对这100只蟹以及另外的30只蟹进行拍摄,建立测试库。对样本库和测试库使用本文算法进行实验,实验结果表明属于同一只蟹的图像使用本文算法计算得到的相似度均在70%以上,而不同蟹图像的相似度普遍偏低,因此选定相似度阈值为0.7。本文算法对市场上商品蟹拍摄的图像和该蟹生产地存储的源图像进行匹配,对于在生产地建立过图像库的河蟹,可以准确的从图像库中找到属于同一只蟹的图像,而没有建立图像库的商品蟹,则无法匹配到结果,以此可以有效的避免目前使用的河蟹防伪方法的缺陷,实现河蟹的识别和溯源,达到防伪的目的。

参考文献:

- [1] 农业农村部渔业渔政管理局,全国水产技术推广总站,中国水产学会. 2018 中国渔业统计年鉴[M]. 第1版. 北京:中国农业出版社,2018.
Administration of Fisheries of the Ministry of Agriculture and Rural Affairs of the People's Republic of China, National Fisheries Technology Extension Center, China Society of Fisheries. 2018 China Fishery Statistical Yearbook[M]. First Version. Beijing: Agriculture Press of China, 2018.
- [2] 王超,杨德利. 市场环境下的大闸蟹品牌创立与维护探讨[J]. 山西农业科学, 2011, 39(12):1323-1327.
Wang C, Yang D L. Approach to brand building and maintenance for Yangcheng Lake hairy crabs under market environment[J]. Journal of Shanxi Agricultural Science, 2011, 39(12): 1323-1327.
- [3] 付延松,胡更生,王群,等. 一种基于多重信息加密的二维码防伪方法[J]. 印刷质量与标准化, 2014(4): 28-31.
Fu Y S, Hu G S, Wang Qun, et al. A two-dimensional code anti-counterfeiting method based on multiple information encryption[J]. Printing Quality & Standardization, 2014(4): 28-31.
- [4] 黄颖为,龚小超. 二维条码技术及其在防伪中的应用[J]. 中国品牌与防伪, 2007(7): 60-63.
Huang Y W, Gong X C. Two-dimensional bar code technology and its application in anti-counterfeiting[J]. China Brand and Anti-counterfeiting, 2007(7): 60-63.

- [5] 周晓庆, 陆涛, 焦强. 基于 QR 码的药品防伪系统[J]. 计算机与现代化, 2015(1): 122-126.
Zhou X Q, Lu Tao, Jiao Qiang. Drug anti-counterfeiting system based on QR code[J]. Computer and Modernization, 2015(1): 122-126.
- [6] 邓洋, 李征. 数字化防伪技术的原理及应用研究[J]. 包装工程, 2015, 36(9): 139-143.
Deng Yang, Li Zheng. Principle and application of digital anti-counterfeit technology[J]. Packaging Engineering, 2015, 36(9): 139-143.
- [7] 李清杨. 防伪原理及防伪技术的发展趋势[J]. 安防科技, 2003(1): 19-20; 22.
Li Q Y. Anti-counterfeit principle and development trend of anti-counterfeiting technology[J]. Safety & Security Technology, 2003(1): 19-20; 22.
- [8] 张炳良, 练迪, 任海良. 基于蟹壳图像的中华绒螯蟹防伪体系研究[J]. 科学养鱼, 2014(2): 77-78.
Zhang B L, Lian D, Ren H L. Study on anti-counterfeiting system of chinese mitten crab based on crab shell images[J]. Scientific Fish Farming, 2014(2): 77-78.
- [9] 王武, 王成辉, 马旭洲. 河蟹生态养殖[M]. 第二版. 北京: 中国农业出版社, 2013.
Wang W, Wang C H, Ma X Z. Ecological Culture of Chinese Mitten Crab[M]. Second edition. Beijing: Agriculture Press of China, 2013.
- [10] 王成辉, 李思发. 中华绒螯蟹种质研究进展[J]. 中国水产科学, 2002, 9(1): 82-86.
Wang C H, Li S F. Advances in studies on germplasm in Chinese mitten crab, *Eriocheir sinensis*[J]. Journal of Fishery Sciences of China, 2002, 9(1): 82-86.
- [11] Boucot, Arthur J. Ecophenotypic or genotypic?[J]. Nature, 1982, 296(5858): 609-610.
- [12] Miner B G, Sultan S E, Morgan S G, et al. Ecological consequences of phenotypic plasticity[J]. Trends in Ecology & Evolution, 2005, 20(12): 1-692.
- [13] Zhao Y J, Qing H, Zhao C J, et al. Phenotypic plasticity of *Spartina alterniflora* and *Phragmites australis* response to nitrogen addition and intraspecific competition[J]. Hydrobiologia, 2010, 637(1): 143-155.
- [14] Sánchez R, Sepúlveda R D, Brante A, et al. Spatial pattern of genetic and morphological diversity in the direct developer *Acanthina monodon* (Gastropoda: Mollusca)[J]. Marine Ecology Progress, 2012, 434(4): 121-131.
- [15] 褚翔, 祝连庆, 娄小平, 等. 基于改进 Sobel 算子的动态自动调焦算法研究[J]. 应用光学, 2017, 38(2): 237-242.
Chu X, Zhu L Q, Lou X P, et al. Dynamic auto focus algorithm based on improved Sobel operator[J]. Journal of Applied Optics, 2017, 38(2): 237-242.
- [16] 冯亦东, 孙跃. 基于 SURF 特征提取和 FLANN 搜索的图像匹配算法[J]. 图学学报, 2015, 36(4): 650-654.
Feng Y D, Sun Y. Image matching algorithm based on SURF feature extraction and FLANN search[J]. Journal of Graphics, 2015, 36(4): 650-654.
- [17] 寇冬晨, 韩瑜, 顾浩. 基于 Hessian 矩阵的角点自动检测算法[J]. 指挥控制与仿真, 2018, 40(2): 118-122.
Kou D C, Han Y, Gu H. An angle point automatic detection algorithm based on Hessian Matrix[J]. Command Control & Simulation, 2018, 40(2): 118-122.
- [18] 单小军, 唐婷. 图像匹配中误匹配点检测技术综述[J]. 计算机应用研究, 2015, 32(9): 2561-2565; 2571.
Dan X J, Tang P. Review of false matching points detection methods for image matching[J]. Application Research of Computers, 2015, 32(9): 2561-2565; 2571.
- [19] 张铎. 生物识别技术基础[M]. 武汉: 武汉大学出版社, 2009.
Zhang D. The Basis of Biometric Identification Technology[M]. Wuhan: Wuhan University Press, 2009.
- [20] Idaszkin Y L, Márquez F, Nocera A C. Habitat-specific shape variation in the carapace of the crab *Cyrtograpsus angulatus*[J]. Journal of Zoology, 2013, 290(2): 117-126.
- [21] Harris C G, Stephens M J. A combined corner and edge detector[C].//Proceedings of the 4th Alvey Vision Conference. Manchester, England; [s.n], 1988.
- [22] Rublee E, Rabaud V, Konolige K, et al. ORB: An Efficient Alternative to SIFT or SURF[C]. [s.l.]: IEEE International Conference on Computer Vision, 2011: 2564-2571.
- [23] Lowe D. Distinctive image features from scale-invariant key points[J]. International Journal of Computer Vision, 2004, 60(2): 91-110.
- [24] Bay H, Tuytelaars T, Van Gool L. SURF: Speed up Robust Features[C]. [s.l.]: European Conference on Computer Vision, 2006: 404-417.
- [25] Muja M, Lowe D G. Fast approximate nearest neighbors with automatic algorithm configuration[C].// Proceedings of IEEE Conference on Computer Vision Theory and Applications. Lisbon. Portugal: IEEE Computer Society, 2009: 331-340.
- [26] 赵璐璐, 耿国华, 李康, 等. 基于 SURF 和快速近似最近邻搜索的图像匹配算法[J]. 计算机应用研究, 2013, 30(3): 921-923.
Zhao L L, Geng G H, Li Kang, et al. Images matching algorithm based on SURF and fast approximate nearest neighbor search[J]. Application Research of Computers, 2013, 30(3): 921-923.

Research on the Feature Recognition and Algorithm of the Carapace of *Eriocheir sinensis*

TAI Wei-Peng^{1,2}, LI Hao¹, ZHANG Bing-Liang³, WANG Chun⁴

(1. College of Computer Science and Technology, Anhui University of Technology, Ma'anshan 243000, China; 2. Research Institute of Information Technology, Anhui University of Technology, Ma'anshan 243000, China; 3. Tech-Bank Food Co., Ltd., Yuyao 315400, China; 4. Laboratory of Freshwater Fisheries Germplasm Resource, Ministry of Agriculture, National Demonstration Center for Experimental Fisheries Education, Shanghai Engineering Research Center of Aquaculture, Shanghai 201306, China)

Abstract: Aiming at the poor accuracy and lower reliability of current anti-counterfeiting methods of Chinese mitten crab *Eriocheir sinensis*, This paper suggests a carapace feature recognition algorithm for Chinese mitten crab based on image matching. Firstly, the unique carapace image of Chinese mitten crab is segmented on which feature points are extracted by means of SURF (Speeded up robust features), then these feature points are matched up with those stored in the database by virtue of the two-way FLANN (Fast library for approximate nearest neighbors) algorithm. According to the positional relationship between these matching points, the nearest Euclidean distance's and the second nearest Euclidean distance's matching points, it is used as the basis of mismatch detection. And the ratio of the correct matching point pair is used to verify its similarity so as to realize the individual identification of a crab. This findings show that the similarity of the two carapace images belonging to the same crab is higher than the threshold value of 0.7 and lower for other two different crabs under the similar conditions of image acquisition. This research can solve the problem of poor technical accuracy and low reliability of a crab's traceability and anti-counterfeiting using barcodes or two-dimensional barcodes attaching to a crab's chelae, and can contribute a technique for dealing with accurate individual identification of merchandise crabs to realize the purpose of anti-counterfeiting and Accurate traceability.

Key words: *Eriocheir sinensis*; carapace; feature recognition; image processing; similarity; anti-counterfeiting check

责任编辑 朱宝象