

基于 Adaboosting_SVM 算法的多特征蛋壳裂纹识别

熊利荣 谢 灿 祝志慧

华中农业大学工学院, 武汉 430070

摘要 以无裂纹蛋和裂纹蛋为测试对象,采用机器视觉技术和支持向量机等技术手段,分析无裂纹蛋和裂纹蛋在图像上的差异,提取特征参数,实现蛋壳裂纹的自动识别;针对蛋壳表面的亮斑,对预处理后的图像运行消除亮斑算法并进行区域标记。在此基础上,从5个不同视角提取13个能够表征无裂纹蛋和裂纹蛋的特征参数,分别是图像标记区域参数(区域标记数和标记点数)、几何特征参数(长轴和短轴)、基于 Freeman 链码的形状参数(形状数)、纹理特征参数(均值、标准偏差、平滑度、三阶矩、一致性、熵)和频谱特性参数(最大幅值和最大相位)。采用 Adaboosting 算子对上述特征参数进行优化,突出影响因子较大的参数组合,作为 SVM 的输入向量,建立蛋壳裂纹的识别模型。结果表明:该方法对蛋壳表面的亮斑、微小裂纹及普通裂纹均具有识别能力,模型正确率达 97.5%,符合蛋品企业对蛋壳裂纹检测的精度要求。

关键词 蛋壳; 裂纹; 机器视觉; 支持向量机

中图分类号 S 126; TP 391.41 **文献标识码** A **文章编号** 1000-2421(2015)02-0136-05

随着中国禽蛋加工业的发展和外贸出口的需要,研究高速、高效、高精度的禽蛋裂纹自动化检测方法和设备,实现蛋壳裂纹的自动检测和剔除,对促进农业现代化十分重要。目前,蛋品加工企业大多采用人工法检测蛋壳裂纹,不仅劳动强度大、生产效率低,而且检测精度受人为影响而得不到根本保证^[1]。近年来,机器视觉技术逐步成为品质检测的一种有效方法,在农产品检测中也有很多应用,如对葡萄、玉米、马铃薯、稻谷、棉花、黄瓜等农产品的品质检测,都取得了良好效果^[2-5]。但对蛋壳的裂纹检测,大多仍采用声学敲击方法,很少采用机器视觉技术^[6-7]。其原因是当蛋壳表面的裂纹极其微小、肉眼不可见时,机器视觉无法识别;当无裂纹蛋表面存有明显大面积亮斑时,机器视觉容易产生误判,将无裂纹蛋误判为裂纹蛋。针对这些问题,本试验选择蛋壳裂纹目标区域,以鸡蛋为测试对象,以机器视觉和图像处理技术为技术手段,寻找能更全面地描述无裂纹蛋和裂纹蛋差异的特征参数组合。同时,运用特征分类算法(feature classification algorithm) Adaboosting^[8]进行特征筛选,选择支持向量机(support vector machine, SVM)^[9]作为模型分类

器,提出了基于 Adaboosting_SVM 的分类建模方法,提高分类的精度和运行效率,旨在为实现蛋品裂纹的自动分级提供理论依据。

1 材料与方法

1.1 供试材料

鸡蛋:供试新鲜鸡蛋来自湖北省武汉市洪山区九峰山新亚养鸡场;主要仪器设备:图像采集硬件系统由光照系统(包括光室和光源)、照相机(德国产 UI-2230ME-C 工业相机)和计算机(Acer4820TG)组成。

1.2 检测方法

通过放大镜观察选取当天生产的、均无裂纹的新鲜鸡蛋 60 枚,统一编号备用;取其中 30 枚无裂纹蛋,在万能试验机(型号为 RGT2000-10)上进行压力试验,通过在蛋壳不同位置(如中间、大头和小头)施加不同的力产生不同裂纹类型(如线状裂纹、块状裂纹和网状裂纹等)、不同裂纹程度(如轻微裂纹和较大裂纹)和不同裂纹位置(如裂纹分别在大头、小头或中间)的裂纹蛋样品;另外 30 枚为无裂纹蛋样品。依次采集样品鸡蛋图像(图像格式为 BMP,位

收稿日期:2014-04-28

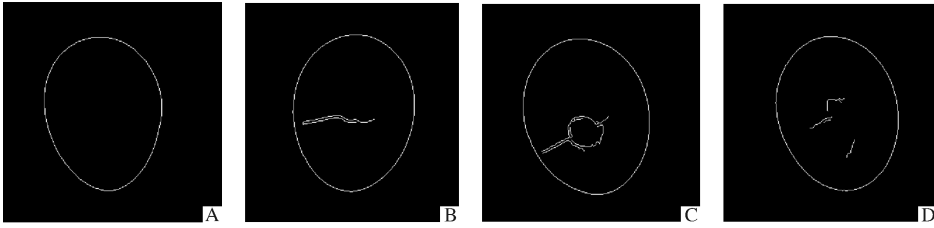
基金项目:国家自然科学基金项目(51105160);华中农业大学博士启动基金项目(52902-0900206027)

熊利荣,博士,副教授。研究方向:农业电气化与自动化。E-mail:lgj1234_73@sina.com

通信作者:祝志慧,博士,副教授。研究方向:农业电气化与自动化。E-mail:zzh@hzau.edu.cn

深度为 24 bit),对图像进行灰度化、图像增强等常规预处理后,采用 Canny 算子提取边界信息。图 1 为无裂纹蛋和裂纹蛋样本图像,不仅包含了裂纹区域,也包含了外轮廓。将图 1 与样品蛋外轮廓二值

化图像进行减运算,得到本试验的目标区域。对于无裂纹蛋,目标区域不含白色像素点或含极少白色像素点(由于蛋壳表面亮斑引起);对于裂纹蛋,目标区域为裂纹信息。



A:无裂纹蛋 No crack egg; B:线状裂纹蛋 Line crack egg; C:网状裂纹蛋 A net crack egg; D:散点状裂纹蛋 Spot crack egg.

图 1 样品蛋图像预处理后的结果

Fig.1 The result after pre-procession of the samples

2 结果与分析

2.1 亮斑消除和目标区域连通区域的标记

考虑到部分无裂纹蛋表面存在较多亮斑,预处理后在目标区域会留下类似裂纹的像素而导致误判,因此,本试验提出了一种消除蛋壳斑痕的区域标记算法,通过寻找区域面积阈值 T_1 ,将小于阈值的区域像素由 1 变为 0,从而得到边缘检测以及去除斑痕阈值的最优组合。

算法步骤如下:

(1)选择初始阈值 $T_1 = 10$;

(2)对边缘检测后的二值图像(背景用“0”像素表示,而目标区域用“1”像素表示)进行二次扫描来实现区域连通标记;

(3)计算每个样品鸡蛋的区域标记数目 Num_k ($k=1,2,\dots,30$),表示第 k 个鸡蛋的标记点数;

(4)统计各样品鸡蛋的各标记区域中的标记点数 $L(j,k)$, ($j=1,2,\dots,Num_k, k=1,2,\dots,30$) 表示第 k 个鸡蛋的第 j 个标记区域的标记点数;

(5)比较标记点数 $L(j,k)$ 与阈值 T_1 ,如果 $L(j,k) < T_1$,则将第 j 个标记区域的像素值全部赋

值为 0;

(6)重新对处理后的鸡蛋进行标记,计算新的标记点数 $L(j,k)$ 和标记区域数目 Num_k ,计算无损鸡蛋 $Num_{k1} = 0$ ($k1=1,2,\dots,15$) 的个数 $N1$,以及裂纹鸡蛋 $Num_{k2} \neq 0$ ($k2=1,2,\dots,15$) 的个数 $N2$,计算精度 $p1 = \frac{N1}{15}, p2 = \frac{N2}{15}$;

(7)判断 $p1, p2$ 是否都大于精度 $p=0.98$,若大于 p ,则 T_1 数选择较合理;若小于 p ,则 $T_1 = T_1 + 5$;重复上述步骤,终获得的阈值 $T_1 = 40$,同时得到每个样品蛋的 Num_k 和 $L(j,k)$ 。为了方便起见,令 $Num_k = x_1, L(j,k) = x_2$ 。

由于篇幅有限,本试验仅列出 30 个样品蛋的 x_1 和 x_2 (表 1)。由表 1 可知,无裂纹蛋的 x_1 和 x_2 基本为 0,去除亮斑效果较好。需要指出的是,8 号样品蛋的蛋壳表面存在大量黑斑,属于蛋壳异常蛋,因此 x_1 和 x_2 出现异常。

上述样品蛋壳消除亮斑后的特征区域将作为后续提取参数的基础,这种方法可从根本上消除亮斑对后续特征参数的影响,直接提高特征参数的准确性。

表 1 去除亮斑后的 30 枚样品蛋目标区域的 x_1 和 x_2

Table 1 x_1 and x_2 of the target area for 30 samples removed bright spot

No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
x_1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
x_2	0	0	0	0	0	0	0	41	0	0	0	0	0	0	0
No.	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45
x_1	4	1	1	4	1	3	1	2	1	2	6	2	2	1	2
x_2	312	109	123	255	143	256	353	242	40	216	673	120	118	124	191

2.2 目标区域几何特征参数的提取

图像的几何特征参数是区分一个图像最基本也是最常用的特征信息之一。

本试验提取蛋壳裂纹目标区域的长轴 x_3 和短轴 x_4 (表 2)。由表 2 可知,无裂纹蛋 x_3 和 x_4 均比裂纹蛋小很多。

表 2 30 枚样品蛋特征区域 x_3 和 x_4

Table 2 x_3 和 x_4 of the target area for 30 samples

No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
x_3	0	0	0	0	0	0	0	21	0	0	0	0	0	0	0
x_4	0	0	0	0	0	0	0	13	0	0	0	0	0	0	0
No.	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45
x_3	36	55	46	59	75	65	105	102	25	54	21	43	41	83	82
x_4	12	6	28	7	42	26	56	21	14	40	10	9	3	15	12

2.3 基于 Freeman 链码的形状数的提取

长度,称为形状数的阶,记为 x_5 ,它可作为闭合边界的周长。

Freeman 链码通过一个指定长度与方向的直线段的连接序列来表示一条边界,可用来表示目标区域边界,称为形状数^[10]。本试验提取形状数序列的

表 3 为 30 枚样品蛋目标区域 x_5 。由表 3 可知,无裂纹蛋 x_5 比裂纹蛋小很多。

表 3 30 枚样品蛋特征区域 x_5

Table 3 x_5 of the target area for 30 samples

No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
x_5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
No.	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45
x_5	128	117	175	140	233	277	332	266	93	224	387	140	140	197	209

2.4 目标区域纹理特征参数的提取

表 4 30 枚样品蛋特征区域 x_6 、 x_7 、 x_8 、 x_9 、 x_{10} 和 x_{11}

Table 4 $x_6, x_7, x_8, x_9, x_{10}$ and x_{11} of the target area for 30 samples

运用灰度直方图的统计属性可得到图像的某些纹理特征参数,本试验采用以下 6 个统计特征值:(1)均值 x_6 ,反应灰度图像的平均亮度;(2)标准偏差 x_7 ,反应平均对比度;(3)平滑度 x_8 ,反应区域中亮度的相对平滑度;(4)三阶矩 x_9 ,反应灰度直方图的偏斜量;(5)一致性 x_{10} ,当所有灰度值相等时,该值最大并从此处开始减小;(6)熵 x_{11} ,反应度量的随机性。

No.	x_6	x_7	x_8	x_9	x_{10}	x_{11}
1	0	0	0	0	0	0
2	2	0	0	0	0	1
3	0	0	0	0	0	0
4	0	0	0	0	0	0
5	143	35	93	47	22	1
6	2	0	0	0	0	1
7	0	0	0	0	0	0
8	0	0	0	0	0	0
9	315	42	137	69	30	1
10	5	0	0	0	5	1
11	0	0	0	0	0	0
12	0	0	0	0	0	0
13	0	0	0	0	0	1
14	0	0	0	0	0	0
15	1	0	0	0	0	0
31	183	37	105	53	24	1
32	10	18	24	0	7	0
33	12	19	27	0	7	0
34	130	34	89	45	21	1
35	16	20	32	16	9	1
36	142	35	93	47	21	1
37	129	34	88	45	21	0
38	52	27	56	29	14	1
39	6	16	0	0	6	1
40	77	30	68	35	17	1
41	531	48	177	90	37	0
42	16	20	32	16	9	1
43	28	23	41	21	11	1
44	15	20	30	0	8	1
45	33	24	45	23	11	1

考虑到各参数数值上的差异,将其均放大 2.5 倍,得到各参数如表 4 所示。表 4 中,对 x_6 、 x_7 、 x_8 、 x_9 、 x_{10} 这 5 个参数而言,无裂纹蛋较裂纹蛋数值小很多, x_{11} 的规律则不明显。

2.5 目标区域谱纹理特征参数的提取

作目标区域的频域特性曲线,分别取最大幅值和最大相位,放大取整后,记作 x_{12} 、 x_{13} ,如表 5。由表 5 可知,无裂纹蛋和裂纹蛋在幅值和相位上存在差异。

2.6 基于多重特征筛选的裂纹禽蛋识别

1)数据归一化。由于所提取的特征是从不同侧面描述无裂纹蛋和裂纹蛋特征,不同特征向量间的物理意义不一样,且不同特征分量的幅度区间也

表 5 30 枚样品蛋目标区域 x_{12} 和 x_{13}

Table 5 x_{12} and x_{13} of the target area for 30 samples

No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
x_{12}	0	0	0	0	3	0	0	0	2	0	0	0	0	0	0
x_{13}	0	1	0	0	2	1	0	0	3	1	0	0	0	0	1
No.	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45
x_{12}	2	1	1	2	1	2	2	1	1	2	3	1	1	1	1
x_{13}	3	3	1	2	2	3	3	3	1	3	4	1	2	2	3

不一样,因此,建立分类模型之前,对特征向量(x_1 、 x_2 、 x_3 、 x_4 、 x_5 、 x_6 、 x_7 、 x_8 、 x_9 、 x_{10} 、 x_{11} 、 x_{12} 、 x_{13})进行归一化处理。

2)基于 Adaboosting_SVM 的分类器模型。本试验运用 Adaboosting 算法进行特征参数筛选。根据每

个特征参数在分类训练过程中的错误率和收敛情况,确定当前特征参数对分类结果的影响因子,将影响因子较大的特征参数突出,将影响因子较小的特征参数忽略或者削弱,去除冗余特征,提高目标识别正确率。Adaboosting 算法的实现步骤如图 2 所示。

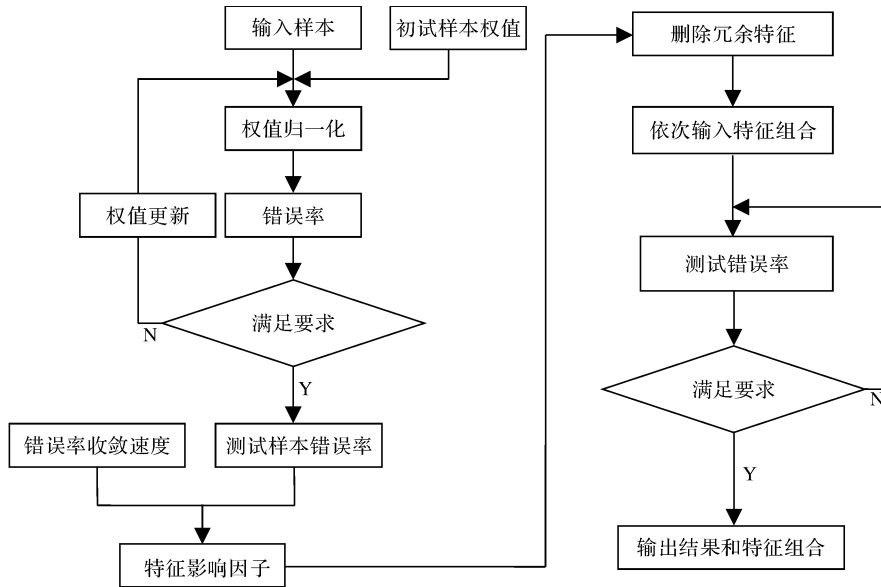


图 2 Adaboosting 算法的实现步骤

Fig. 2 The implementation steps of the Adaboosting algorithm

运用 Adaboosting 算子计算前述 13 个参数的特征因子,其中, x_1 、 x_2 、 x_3 和 x_4 的特征因子均为 0.042,其他 9 个特征参数的特征因子均为 0.083。对上述特征因子进行特征筛选后,选出最优特征组合为 (x_{10} 、 x_{12})。

3)两种识别模型检测结果对比。取全变量 (x_1 、 x_2 、 x_3 、 x_4 、 x_5 、 x_6 、 x_7 、 x_8 、 x_9 、 x_{10} 、 x_{11} 、 x_{12} 、 x_{13})建立 SVM 识别模型,进行蛋壳裂纹的识别。同时,选择特征向量(x_{10} 、 x_{12})作为 SVM 的输入向量,建立 Adaboosting_SVM 识别模型,对同一批样品蛋进行识别(表 7)。由表 7 可知,经过 Adaboosting 算法特征筛选后,系统的识别性能比未筛选时的性能优越,不仅提高了准确率,运行效率也得到很大提高。

表 7 两种模型识别结果(样本数 60)

Table 7 Identification result of two models (samples 60)

模型 Models	正确样本个数 Right samples	正确率/% Accuracy rate	运行时间/s Running time
SVM	58	93.33	0.005 24
Adaboosting_SVM	60	100.00	0.001 99

4)模型验证。为了验证模型的适用性和正确性,本试验重新采集 80 个新样本,无裂纹蛋和裂纹

蛋各 40 个。分别运用 SVM 和 Adaboosting_SVM 方法进行分类,结果如表 8 所示。由表 8 可知,表 8 的检测结果与表 7 一致。

表 8 两种模型验证结果(样本数 80)

Table 8 Verification result of two models (samples 80)

方法 Methods	正确样本个数 Right samples	正确率/% Accuracy rate	运行时间/s Running time
SVM	76	95.00	0.009 40
Adaboosting_SVM	78	97.50	0.001 10

3 讨论

本试验提出了一种结合图像处理、多个特征参数提取、Adaboosting 算法和支持向量机等技术手段进行蛋壳裂纹识别的方法。以图像处理技术为基础,从不同侧面提取表征无裂纹蛋和裂纹蛋的 13 个特征参数,并尝试了 Adaboosting 算法对特征参数进行筛选,从而建立 SVM_Adaboosting 识别模型,准确率达到 97.5%,识别时间也大大缩短。

本试验存在的不足之处:第一,当裂纹出现在图像的背面或两头时,无法使用图像处理方法进行裂纹检测,以后可结合声学技术,或采用多相机或图像旋转方式采集一枚蛋的多幅图像。采集一枚样品蛋

的多幅图像,可分别进行裂纹检测。只要任何一幅识别结果为裂纹蛋,则该样品判为裂纹蛋,以解决裂纹在背面或两头的问题;第二,将静态的研究成果运用到实验室现有的动态在线检测系统中时,对动态图像的恢复很重要。如何消除边界模糊,并使图像采集的时间与带速匹配,也是值得注意的;第三,如何消除样本自身复杂性造成的误差,如鸡蛋的形状、蛋壳厚度、蛋壳强度、裂纹大小和分布等不同因素,也是后续研究的重点。

参 考 文 献

- [1] 马美湖. 蛋与蛋制品加工学[M]. 北京:中国农业出版社,2005.
- [2] BARMES M, DUCKETT T, CIELNIAK G, et al. Visual detection of blemishes in potatoes using minimalist boosted classifiers [J]. *Journal of Food Engineering*, 2010, 98(3): 339-346.
- [3] LYV N, LI B, CHEN S, et al. Comparison of color techniques to measure the color of parboiled rice [J]. *Journal of Cereal Science*, 2009, 50(2): 262-265.
- [4] LI D L, YANG W Z, WANG S L. Classification of foreign fibers in cotton lint using machine vision and multi-class support vector machine [J]. *Computers and Electronics in Agriculture*, 2010, 74(2): 274-279.
- [5] CLEMENT J, NOVAS N, GAZQUEZ J A, et al. An active contour computer algorithm for the classification of cucumbers [J]. *Computers and Electronics in Agriculture*, 2013, 92: 75-81.
- [6] SUN L, BI X K, LIN H, et al. On-line detection of eggshell crack based on acoustic resonance analysis [J]. *Journal of Food Engineering*, 2013, 116(1): 240-245.
- [7] DENG X Y, WANG Q H, CHEN H, et al. Eggshell crack detection using a wavelet-based support vector machine [J]. *Computers and Electronics in Agriculture*, 2010, 70(1): 135-143.
- [8] OWUSU B, ZHAN Y Z, MAO Q R, et al. A neural-AdaBoost based facial expression recognition system [J]. *Expert Systems with Applications*, 2014, 41(7): 3383-3390.
- [9] 白鹏, 张喜斌, 张斌. 支持向量机理论及工程应用实例[M]. 西安:电子科技大学出版社, 2008:13-20.
- [10] GONZALEZ R C, WOODS R E. *Digital image processing* [M]. 2nd ed. London: Pearson Higher Education Press, 2002.

An identification method based on multi-feature and Adaboosting_SVM of eggshell crack

XIONG Li-rong XIE Can ZHU Zhi-hui

College of Engineering, Huazhong Agricultural University, Wuhan 430070, China

Abstract With good eggs and crack eggs as experimental subjects, the machine vision and support vector machines (SVM) were used to study the differences between good eggs and crack eggs, and multi-feature parameters were extracted to achieve automatic recognition of crack eggs. Firstly, an algorithm would be run to eliminate bright spots on the preprocessed image of the surface of eggs before marking them by region. Secondly, 13 characteristic parameters from five different domains to identify the good eggs and crack eggs were extracted, and these parameters were as follows: the marked region parameters of images (the number of markers and the marker area points), the geometric parameters (the major axis and the minor axis), the shape parameters based on Freeman chain code (the shape number), the texture parameters (the mean, the standard deviation, the smoothness, the third moment, the uniformity and the entropy) and the spectral parameters (the maximum amplitude and the maximum phase). Thirdly, to highlight the greater impact factors between 13 parameters and to shorten the detection time, adaboosting algorithm was used to optimize the above parameters, which was the input vector of SVM. Finally, the recognition model was built by SVM. The results indicated that the accuracy rate of the recognition model was 97.5%, which could meet the requirements of enterprises basically.

Key words eggshell; crack; machine vision; support vector machine (SVM)