doi:10.6041/j.issn.1000-1298.2018.09.004

基于动态多特征变量的黄羽肉鸡跛行状态定量评价方法

沈明霞[1,2]　李嘉位[1,2]　陆明洲[1,2]　刘龙申[1,2]　孙玉文[1,2]　李　泊[1,2]
(1. 南京农业大学工学院, 南京 210031; 2. 江苏省智能化农业装备重点实验室, 南京 210031)

摘要: 肉鸡步态是肉鸡健康状态的重要表征,为实现黄羽肉鸡跛行步态的无损自动化快速分类识别,提出了一种基于多特征变量的肉鸡跛行定量评价方法。该方法从步态视频中提取肉鸡的速度、步幅、步幅差、步频、投影面积参数,拟合得出速度、步幅与投影面积具有相关性(决定系数分别为 0.805 1、0.793 5),据此定义肉鸡动态理想参数与异常指数,基于 C4.5 决策树模型,以速度异常指数、步幅异常指数、步幅差异常指数为分类特征,根据鸟类步态评分标准将肉鸡分为 GS0 ~ GS4 五类,实现对肉鸡跛行状态进行预警和判别。实验结果表明:该模型针对 GS0 ~ GS4 分类准确率依次为:66%、71%、74%、98%、95%,整体准确率为 78%。该模型可作为早期跛行的检测工具,为养殖自动化的实现和动物福利产业的升级提供支持。

关键词: 肉鸡跛行; 步态评分; 图像处理; 步幅差; 异常指数; 决策树
中图分类号: TP391; S831.4　　**文献标识码**: A　　**文章编号**: 1000-1298(2018)09-0035-10

Evaluation Method of Limping Status of Broilers Based on Dynamic Multi-feature Variables

SHEN Mingxia[1,2]　LI Jiawei[1,2]　LU Mingzhou[1,2]　LIU Longshen[1,2]　SUN Yuwen[1,2]　LI Bo[1,2]
(1. *College of Engineering*, *Nanjing Agricultural University*, *Nanjing* 210031, *China*
2. *Jiangsu Provincial Key Laboratory of Intelligent Agricultural Equipment*, *Nanjing* 210031, *China*)

Abstract: The bird gait score (GS) is an important tool for evaluating the gait status of broiler. GS0 ~ GS5 corresponds to the broilers whose limping level varies from low to high. The level of limping is used as an important indicator to measure the health of broilers. At present, traditional methods for gait assessment of broiler are mostly completed by visual inspection. The process is time-consuming with low standardization. The dynamic feature variables extracted from video were used to evaluate the gait status of broilers based on decision-tree, and a fast, stable and non-contact broiler gait evaluation method was explored. The experiment was conducted at Quanjiao Broiler Breeding Center of Wenzhou Group, from December 2017 to January 2018. A total of 260 broilers (GS0 ~ GS4) were selected. Each broiler was subjected to twice walking-experiments. The experiment was conducted in a special broiler walkway. Two cameras were placed on the opposite side of the walkway and at the top of walkway, and videos were collected horizontally and vertically. Each frame of the video underwent image reorganization, filtering for pretreatment in HSV space. The broiler projection area was calculated by the least squares ellipse fitting based on vertical image, and the dynamic parameters such as the walking speed, stride-length, stride-difference value, and walking steps of the broilers were calculated based on horizontal image. Based on the study of the dynamic parameters of GS0 broilers, the linear fitting relationship between walking speed, stride-length and projection area of broilers was obtained by the least square method, the coefficient of certainty was 0.805 1 and 0.793 5, respectively. According to the fitting results, based on the different top projection areas of the broiler, the ideal stride and ideal speed of the broiler were proposed. Then, according to the difference between the actual value and the ideal value of the parameters such as stride and speed, the abnormal index of dynamic parameters in broiler walking was defined. Taking the anomaly index, including speed, stride and step difference as training attributes, the

收稿日期: 2018-03-28　修回日期: 2018-05-20
基金项目: 国家重点研发计划项目(2017YFD0701602-2)、国家自然科学基金青年基金项目(61503187)和常州市科技支撑计划项目(CE20172005)
作者简介: 沈明霞(1964—),女,教授,博士生导师,主要从事机器视觉与信息农业研究,E-mail: mingxia@njau.edu.cn

C4.5 decision tree model was optimized for learning and post-pruning. Totally 520 data was verified by a 10-fold crossover method to obtain the classification result. The accuracy of GS0 ~ GS4 classification was 66%, 71%, 74%, 98% and 95%, and the overall accuracy was 78%. The above results showed that based on the dynamic multi-feature variables extracted from video and decision tree model, the quantitative evaluation of limping state of broilers can be achieved. The research result provided a method for assessing the degree of non-contact broilers with high accuracy. The method can be used as an early detection tool for identification and early warning for broilers limping, which provided support for the realization of farming automation and animal welfare industry upgrading, which had certain practical value.

Key words: broiler limping; gait score; image processing; step difference; anomaly index; decision tree

0 引言

跛行程度是对肉鸡腿部健康状态和损伤判断的标准化描述[1]。随着肉鸡集约化与规模化养殖的推广,培育期大幅缩短,但肉鸡腿部疾病发生率较为常见,肉鸡有明显跛行的占总量的 20.66%[2-3]。跛行肉鸡的存在比例与现代化养殖要求相悖,既降低了动物福利水平,又影响了肉鸡生产的品质与产量。

对跛行进行早期识别,是对肉鸡疾病进行预防的有效途径。肉鸡跛行经典评价方法是 KESTIN 等[4]提出的鸟类步态六级评分标准(Gait score, GS),其中 GS0 ~ GS5 6 个级别分别表示从腿部健康到严重病态的情况,如表 1 所示。

表 1 鸟类步态评分标准说明

Tab.1 Gait score introduction of six-class definition

步态评级	行走能力	特征描述
GS0	行走无异常	步态灵活,每只脚落地时都在重心下方无摇摆,能够单腿平衡,可完全掌握行进方向
GS1	有轻微缺陷	该缺陷难以明确定义,如因步幅过大产生的轻微不均衡步态,仅作为育种缺陷
GS2	有易识别缺陷	未妨碍移动或竞争资源,如单腿摇摆或跛行。不严重损害身体控制能力、加速度或速度
GS3	有明显缺陷	行走时单腿呈现严重劈叉或张开,身体控制能力、加速度和速度受到影响
GS4	有严重缺陷	受到严重威胁或驱动时才会行走。身体控制能力、加速度和速度受到严重影响
GS5	不能持续走路	很难站立,即使站立只能在翅膀的协助下或通过小腿在地上爬行实现运动

AYDIN 等[5]的研究证明肉鸡步态得分可以作为跛足衡量指标。对肉鸡进行 GS 人工分类时,主观性强,准确度差,易受到环境影响,而基于视频的跛行评价算法为肉鸡步态的自动判别提供了支持。NAAS 等[6]开发了一套协同算法,通过图像分析单只肉鸡的行走速度与加速度,实现对肉鸡进行 GS0 ~ GS5 的自动分类,在 GS3、GS4 等极端分数上取得了接近 100% 的准确率,但针对 GS1 和 GS2 的判断准确率仅为 50%。DAWKINS 等[7]通过使用光流算法来调查鸡群的行为,揭示了光流与肉鸡步态评分水平之间的关系。AYDIN[8]通过 3D 相机研究肉鸡的躺卧时间,提出了一种自动评估肉鸡跛足的方法,以评估肉鸡跛足的严重程度,得出肉鸡躺卧时间与传统肉鸡的步态评分之间存在着明显负相关的结论。在相近领域,劳凤丹等[9]通过贝叶斯模型基于 10 个特征量实现了对蛋鸡日常行为的分类。赵凯旋[10]通过奶牛行走视频中对四肢进行定位和跟踪,得出了奶牛运动曲线,并通过交叉验证的决策树建立了奶牛 3 分制步态评分模型,识别奶牛跛行总体精度达到 90.18%。肉鸡腿部健康与步态评价相关的研究,国内尚未见文献报道。

本文基于 C4.5 决策树模型,从肉鸡步态视频中提取速度、步频、步幅、步幅差等动态指标,并将其作为训练特征,以提高对 GS1、GS2 等轻微病态步态识别的准确率。最终以 GS 为依据,提出一种新的肉鸡步态评分鉴别与分类方法,为肉鸡健康预警提供技术支持。

1 材料和装置

1.1 实验对象

实验于 2017 年 12 月—2018 年 1 月在滁州市温氏集团牧业的全椒肉鸡养殖中心进行,共选取黄羽肉鸡 267 只,日龄在 40 ~ 45 d 之间,质量在 0.6 ~ 1.2 kg 之间,鸡舍温度(21 ± 5)℃。实验前将木质步行道放入鸡舍 7 d,以使肉鸡适应实验环境,消除实验时的环境应激性。以 KESTIN 步态评分标准[4]为依据,经过动物科学相关专家人工鉴定,选取肉鸡的步态分布结果如表 2 所示。其中 GS5 步态因为其走动能力极差,难以采集步态视频,且强迫其走动有违动物福利要求,故未被作为实验对象,实际参与

实验研究肉鸡 260 只，每只鸡记录 2 次行走视频，共记录步态数据 520 条。

表 2　实验鸡基本情况

Tab. 2　Basic condition of experimental broilers

步态评级	GS0	GS1	GS2	GS3	GS4
数量	61	66	58	42	33
平均质量/kg	0.9 ±0.1	0.9 ±0.1	0.9 ±0.3	0.8 ±0.2	0.8 ±0.2

1.2　实验平台

实验装置是由木质步行道、白色背景墙和摄像头组成的特制平台，步行道长 130 cm，宽 50 cm，高 20 cm，白色背景墙与步行道等长，高 65 cm，与木质步行道垂直放置。背景墙左右两侧各有一个标记为入口、出口的计时标记点。正对背景墙 100 cm 处放置一台摄像机进行全程视频采集。正对步行道地面高 120 cm 处放置摄像头，在肉鸡到达入口计时点时触发，进行连续 3 帧的图像数据采集。步态采集装置示意图如图 1 所示。

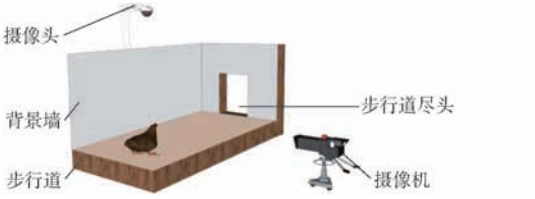

图 1　步态采集装置示意图

Fig. 1　Schematic of broiler gait collection device

1.3　视频采集

将单只肉鸡人为放置在步行道最左端起点，使其自由行走。因仅有步行道终点一个方向可以通向鸡群和日常生活环境，由于肉鸡趋同行为，大多数时候肉鸡可以沿这个方向通过直线路径走向步行道尽头。特别地，对于个别行走意愿不强烈的肉鸡个体，进行轻微驱赶后采集，但在步行过程中不予干预。采集时间为 09:00—14:00。

每只鸡在步行道上完成 5 次视频采集，从中选取 2 次步态相对自然的步态数据记录，以消除偶发应激性导致的异常数据影响。摄像机分辨率为 1 024 像素 ×768 像素，采集频帧为 60 f/s。本实验采用自然光源以避免外加光源导致的肉鸡应激反应。垂直摄像头采集的图像用来计算肉鸡投影面积；水平摄像机采集的视频用来获取肉鸡的行走速度、步幅、步幅差等动态特征参数。

1.4　距离标定实验

在本实验装置设定完毕后，分别在摄像头视野内不同的 10 个位置（水平位置 5 个，竖直位置 5 个）放置边长为 100 mm 的正四边形标定块。通过 K 值法进行距离标定实验。K 是物理长度与像素长度的比值。在已采集图像中，求出正四边形标定块边长对应的像素长度 A，标定块实际边长 B 为 100 mm，计算得出 $K = B/A$，标定结果如表 3 所示。

表 3　标定实验结果

Tab. 3　Calibration test results

类别	次序	A/像素	B/mm	K/(mm·像素$^{-1}$)
水平图像	1	58	100	1.72
	2	58	100	1.72
	3	61	100	1.64
	4	57	100	1.75
	5	63	100	1.59
	平均值	59.4	100	1.68
垂直图像	1	53	100	1.89
	2	57	100	1.75
	3	55	100	1.82
	4	54	100	1.82
	5	57	100	1.75
	平均值	55.2	100	1.81

图像处理过程中，像素为最小处理单位，根据实验结果，水平图像中设置距离换算标准为 1 像素对应 1.68 mm；竖直图像中设置换算标准为 1 像素对应 1.81 mm。

2　动态特征提取

将肉鸡步态视频按时间顺序转换成图像序列，将每帧储存为 JPG 文件，每帧图像尺寸为 1 920 像素 ×1 080 像素。

2.1　图像预处理

针对黄羽肉鸡，其鸡体各部分毛色差异较大，在 RGB 图像中将完整鸡体进行提取有一定困难。本实验中，HSV 空间下鸡体各部分均更易于与背景分离。

本文采用 HSV 空间对图像进行处理[11]，实验中发现 S 分量能较好地将肉鸡身体与背景分离，但肉鸡部分暗色羽毛（如尾部、翅膀等）特征不明显。而在 V 分量中能较好地提取暗色羽毛部分，则以 S 通道为基础，叠加 V 通道部分特征，可以取得理想的效果，如图 2 所示。

肉鸡场环境复杂，无法通过背景减差法[12-13]进行目标提取。本文通过 OTSU 算法[14]进行自适应二值化处理后得到二值图，如图 2d、2e，并可消除大部分无关背景。将二值化后的 S、V 分量图重新组合。最后通过连通域面积阈值过滤与开闭运算[15]，填补图像空隙，消除非必要连通域，得到完整的肉鸡图像，如图 2g 所示。

2.2　图像分割

通过团序列检测算法[16]快速将肉鸡腿部与身

图 2 肉鸡步态序列图像处理过程

Fig. 2 Process of broiler sequence gait image

体进行分割。逐行遍历最大的连通域图像,把每行中不间断像素组成的序列标记为一个“团”,标记第 i 行团的起点为 T_{si}、终点为 T_{ei},每帧图像团序列总长度计算公式为

$$\Delta T = \sum_{i=1}^{n} (T_{ei} - T_{si}) \tag{1}$$

团序列分布图是以图像列数为自变量的离散函数,通过样条函数分段拟合[17]得到样条曲线,在拟合样条曲线中通过极值点识别到团序列变化过程中的转折点,在该列处以 10 像素为宽度进行图像分割,从而分离肉鸡的身体部分与腿部部分,团序列分布如图 3 所示,分割结果如图 4 所示。之后对腿部图像进行转置团序列处理,从而实现将肉鸡的两条腿进行分割,方法相同不再赘述。

图 3 为针对图 2g 的团序列识别结果,并识别出图像团序列函数的极小值点,即图 3 中第 418 行的团序列数,将此行作为分割点,将 413 行到 423 行的数据清除,并将 0 ~ 412 行与 424 ~ 469 行的像素数据分别进行提取,从而实现肉鸡身体与腿部的分离,分割结果如图 4 所示。

2.3 质心提取

从分离的肉鸡身体、前腿、后腿图像块中获得肉鸡身体和腿部的质心点,作为肉鸡步态关键特征提取基础。

本文选择一阶矩[18]和零阶矩的方法计算图像的质心,针对二值图像,其零阶矩 $\boldsymbol{M}_{00}$ 以及一阶矩 $\boldsymbol{M}_{01}$、$\boldsymbol{M}_{10}$ 分别为

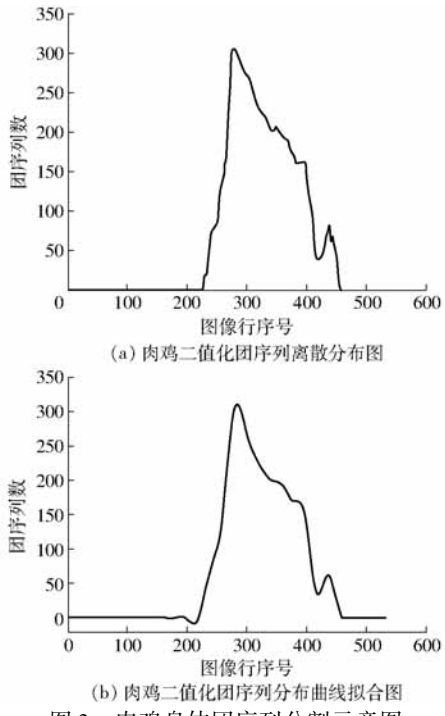

图 3 肉鸡身体团序列分割示意图

Fig. 3 Schematic diagrams of broiler body segmentation

图 4 肉鸡身体与腿部分离示意图

Fig. 4 Diagram of broiler body and leg separation

$$\boldsymbol{M}_{00} = \sum_{i=1}^{I}\sum_{j=1}^{J} V(i,j) \tag{2}$$

$$\begin{cases} \boldsymbol{M}_{01} = \sum_{i=1}^{I} iV(i,j) \\ \boldsymbol{M}_{10} = \sum_{j=1}^{J} jV(i,j) \end{cases} \tag{3}$$

式中 $V(i,j)$——坐标点(i,j)的像素值

图像质心坐标为 $X_0 = \dfrac{\sum p_i x_i}{\sum p_i}$,$Y_0 = \dfrac{\sum p_i y_i}{\sum p_i}$,其中 x_i 和 y_i 表示像素点坐标,p_i 表示该点的像素值。

根据式(2)、(3),分别从分离出的鸡身、左侧腿、右侧腿图像中提取其鸡身质心 W、左脚质心 P 和右脚质心 Q,并将分割图像重新组合,在原图中得到 W、P、Q 3 点的坐标,如图 5 所示。

2.4 肉鸡步态关键特征定义

本文把视频转换成连续单帧图像进行处理[19],以 GS[4]为标准,从中提取肉鸡行进速度、步幅均值、步幅差均值、步频等典型动态特征,特征定义及描述

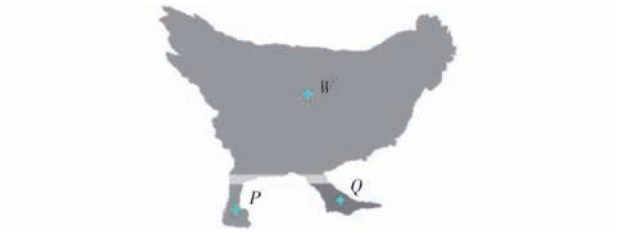

图 5　肉鸡各部分质心位置示意图

Fig. 5　Schematic diagram of broilers centroid

如表 4 所示。

表 4　关键特征定义及描述

Tab. 4　Key feature definition and description

变量	描述	符号	单位	计算方法
行进速度	步行道长度与肉鸡通过时间的比值	V	cm/s	计时标记点间距 $s=110$ cm，肉鸡通过所用时间为 t，则 $V=\frac{110}{t}$
步数	行走步数	n	次	步数 n 为单次实验中，肉鸡双脚同时落地的次数
行走步幅	肉鸡每一步向前方行进距离	S	cm	两只脚同时落地时，鸡脚质心点 x 坐标的差值，取全程平均值
步幅差	肉鸡行走时左、右腿之间的步幅差	D	cm	相邻 2 个行走步幅之间的差值，取全程平均值
步频	行进速度和步长的比值	F	s^{-1}	肉鸡平均每秒行走的步数，取行进速度与步长的比值，即 $F=\frac{V}{S}$

其中行走步幅 S 与步幅差 D 的计算公式为

$$S=\frac{\sum_{a=2}^{n}|x_a-x_{a-1}|}{n-1}\quad(n\geqslant 2)\tag{4}$$

$$D=\frac{\sum_{a=3}^{n}(|x_a-x_{a-1}|-|x_{a-1}-x_{a-2}|)}{n-2}\quad(n\geqslant 3)\tag{5}$$

式中　a——肉鸡的脚落地次数

x_a——第 a 次落地的脚质心横坐标

本文中双脚同时落地状态，指两脚质心 y 坐标相同时的帧。若相邻两次双脚同时落地，每只脚质心的 x 坐标变化之和不超过 10 像素，认为是同一步，不进行重复记录。

3　讨论与分析

行进速度、步幅均值、步幅差均值是描述肉鸡步态的主要特征量，实验中发现 GS 评分同一级别的肉鸡，其特征量仍有较大差异，导致 GS 各级分类时容易发生混淆。AYDIN 等[20]的研究发现特征量差异是由于肉鸡之间体积、体重、体型等客观差异导致的。为消除肉鸡个体之间形态差异造成的数据异常波动，需预先对动态数据进行投影处理。

3.1　投影依据选择

在二维图像特征中，随着肉鸡生长，肉鸡身体长度、宽度会同步增加，二者均可以作为肉鸡体重的关联特征。由于肉鸡行走时，头、尾部会产生伸缩状移动，MORTENSEN 等[21]基于对白羽肉鸡的研究，认为在预测肉鸡体重时，肉鸡的宽度特征比长度更加稳定。同时观察发现，肉鸡羽毛分布情况对水平投影图像影响较大；俯视投影图像中，受羽毛影响较小。

为验证本实验中反映黄羽肉鸡体重情况的最佳变量，选取相同质量（(1.16 ±0.02) kg）的健康肉鸡 59 只，测量其长度、宽度、俯视投影面积、水平投影面积，实验结果如图 6 所示。

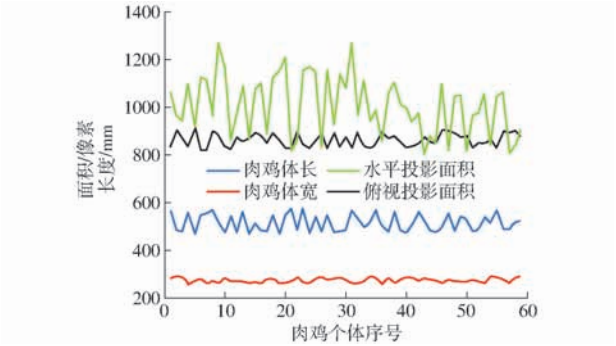

图 6　肉鸡形态参数分布

Fig. 6　Broiler shape parameter distribution

实验结果表明，对于相同质量的肉鸡，其肉鸡体长、肉鸡体宽、水平投影面积、俯视投影面积分布的标准差分别为：35.09 mm、9.63 mm、124.28 像素、27.21 像素。由此可得出，在衡量肉鸡体重的参数中，肉鸡体宽比肉鸡体长更稳定，俯视投影比水平投影更稳定，与 MORTENSEN 等的研究一致。

基于上述结论，本文以肉鸡俯视投影面积为依据对主要特征量进行投影处理，使 GS 评分中同级肉鸡的动态特征量更加集中。

3.2　俯视投影面积提取

本文将肉鸡俯视投影面积作为动态特征数据进行投影处理的核心依据，俯视投影面积提取过程如下。

3.2.1　垂直俯视图像预处理

垂直图像由竖直摄像头采集。图像预处理方法与水平摄像头相同，采用 HSV 分量组合的方法，不再赘述。

3.2.2　投影图像椭圆拟合

肉鸡俯视投影图近似纺锤状，肉鸡行走过程中头部与尾部有前伸或后缩的动作，会产生图像的不

规则形状误差。为进一步提升精确度,需要对俯视投影图像进行拟合处理。MORTENSEN 等[21]在通过肉鸡俯视图像预测肉鸡体重时,使用了“最大内切圆”的拟合方法。根据实验观察,年轻肉鸡一般体型较长,年老肉鸡体型较圆,此时“最大内切圆”方法对年轻肉鸡预测误差较大。

本文将“最大内切圆”优化为基于最小二乘法的椭圆拟合方法对肉鸡的俯视投影图像进行处理,椭圆拟合方法考虑了肉鸡整体的形态特征,更好地降低了体型差异导致的误差,并可以消除肉鸡行进中头部尾部晃动对面积计算的影响。设平面内任意圆锥曲线为

$$\alpha x^2+\beta xy+\zeta y^2+\eta x+\varphi y+\lambda=0 \tag{6}$$

式中 α、β、ζ、η、φ、λ——椭圆方程参数

设 $P_i(x_i,y_i)(i=1,2,\cdots,N)$ 为肉鸡俯视图像轮廓上的 $N(N\geqslant 6)$ 个测量点,依据最小二乘法原理,目标函数为

$$f(\alpha,\beta,\zeta,\eta,\varphi,\lambda)=\sum_{i=1}^{N}(\alpha x^2+\beta xy+\zeta y^2+\eta x+\varphi y+\lambda)^2 \tag{7}$$

为使 f 最小,令

$$\frac{\partial f}{\partial\alpha}=\frac{\partial f}{\partial\beta}=\frac{\partial f}{\partial\zeta}=\frac{\partial f}{\partial\eta}=\frac{\partial f}{\partial\varphi}=\frac{\partial f}{\partial\lambda}=0$$

解得椭圆系数 α、β、ζ、η、φ、λ,即可得到拟合曲线。特别地,为避免零解,同时把多倍的解看作同一个椭圆,设定制约条件为 $\alpha+\zeta=1$。本实验中使用椭圆拟合改进算法[22-23],结合本实验情况,采用如下方案:

(1)在所有样本点中随机选取 6 个,用最小二乘法得到椭圆参数 α、β、ζ、η、φ、λ。

(2)遍历所有样本点,判断每个样本点和已知椭圆的距离,如果小于 15 像素,标记该点为匹配点。

(3)循环执行步骤(1)、(2)400 次,找到匹配点最多的椭圆,其椭圆参数即为所求参数,拟合结果如图 7 所示。

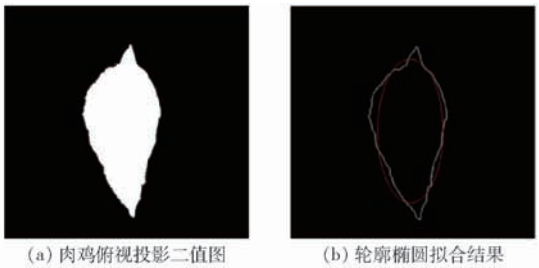

图 7 椭圆拟合示意图

Fig. 7 Ellipse fitting diagrams

3.2.3 椭圆面积计算

将圆锥曲线一般方程转换为几何参数表示,即由椭圆中心(x_0、y_0)、长轴长度 δ、短轴长度 ε、长轴转角 θ 5 个变量表示。其中 δ 可近似视为肉鸡的身长,ε 可近似视为肉鸡胸部宽度,计算方法为

$$\begin{cases}\delta=2\sqrt{\dfrac{-2\lambda}{\alpha+\zeta-\sqrt{\beta^2+\left(\dfrac{\alpha-\zeta}{\lambda}\right)^2}}}\\ \varepsilon=2\sqrt{\dfrac{-2\lambda}{\alpha+\zeta+\sqrt{\beta^2+\left(\dfrac{\alpha-\zeta}{\lambda}\right)^2}}}\end{cases} \tag{8}$$

由积分公式得出,投影椭圆面积为

$$H=\pi\varepsilon\delta \tag{9}$$

3.3 数据处理

本实验中,选取健康状态(GS0)肉鸡 61 只,采集其步态信息与拟合椭圆面积共 122 次(每只 2 次),以得到肉鸡理想动态参数。数据分析结果如图 8 所示。

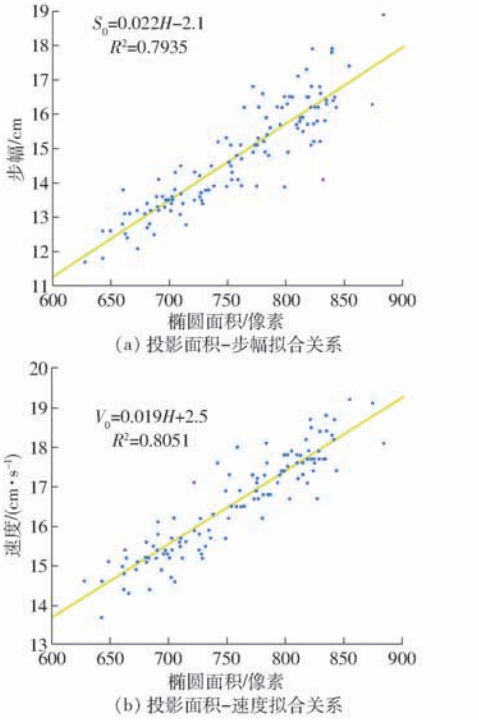

图 8 肉鸡投影面积-动态特征散点图

Fig. 8 Scatter plot of broiler projection area and dynamic characteristic

从图中可以看出,肉鸡步幅、速度与拟合椭圆的面积近似呈正相关关系。基于本实验数据,拟合得出理想步幅 S_0、理想速度 V_0 与投影面积 H 的近似关系,决定系数为 0.793 5 与 0.805 1,具有较好的统计学意义。为方便统计,更好地反映动态数据偏离正常值的程度,引入步幅异常指数 S_u、速度异常指数 V_u、步幅差异常指数 D_u,定义为

$$S_u=100\left(1+\frac{S-S_0}{S_0}\right) \tag{10}$$

$$V_u = 100\left(1 + \frac{V - V_0}{V_0}\right) \tag{11}$$

$$D_u = 100\frac{D}{S} \tag{12}$$

步幅异常指数和速度异常指数等于 100 表示实际值与理想值相等的最佳状态，速度异常指数为零表示没有步幅差的理想状态。异常指数能反映肉鸡动态参数实际情况，同时最大程度消除肉鸡个体形态学差异对数据的影响，一定程度上提高判断算法适应性。

3.4　数据差异性验证

将不同 GS 级别的肉鸡平均行进速度、步幅均值、步幅差均值与每只肉鸡的速度异常指数、步幅异常指数、步幅差异常指数的均值进行统计，得到数据如表 5 所示。

表 5　原始特征与异常指数对比

Tab. 5　Comparison between original features and abnormal index

评级	V/(cm·s^{-1})	V_u	S/cm	S_u	D/cm	D_u	F/s^{-1}
GS0	16.64	100.53	14.58	99.92	1.81	12.54	1.15
GS1	16.05	98.18	14.51	101.18	2.84	19.82	1.11
GS2	15.45	91.25	13.79	92.23	4.78	34.93	1.12
GS3	10.91	66.96	11.59	81.13	5.01	43.88	0.94
GS4	3.47	24.69	4.89	41.34	1.76	42.19	0.71

从表 5 对比可以看出，通过异常指数评价肉鸡步态时，可消除肉鸡个体形态对数据的影响，并保留原始数据中各级别动态参数的差异，同时将差异进行更为合理的缩放，使各级之间区分更加明显，特征独立性更强。全数据异常指数分布箱线图如图 9 所示。

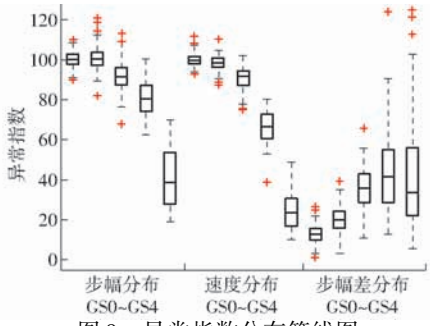

图 9　异常指数分布箱线图

Fig. 9　Box plot of abnormal index distribution

从以上数据可看出，随着腿部疾病级别加重，速度、步幅、步频等数值总体呈下降趋势。特征数据中，除 GS3 与 GS4 之间速度异常指数差异较大外，其他各级别各特征之间数据有很大部分的重叠，难以对个体数据通过简单统计学方法进行分类。

4　步态评分模型

各级步态之间数据存在一定的统计学差异，而这些差异不足以直接作为评判各个 GS 级别的标准。DAWKINS 等[7]的研究表明，速度和加速度等动态参数与步态评分之间没有简单的联系，本文使用 C4.5 决策树模型对肉鸡动态特征异常指数进行训练，并进行步态级别分类。

4.1　训练参数

选取速度异常指数 V_u、步幅异常指数 S_u、步幅差异常指数 D_u 作为关键属性，针对每只肉鸡，把每只肉鸡的 GS 评分作为评价结果，形成最优决策树。

训练集组成方式如下：训练集由 4 列组成，每列依次为 X_1、X_2、X_3、Y_{lable} 4 个变量，前 3 列为特征属性，对应速度异常指数 V_u、步幅异常指数 S_u、步幅差异常指数 D_u，第 4 列为标签变量，为肉鸡的 GS 人工评分结果，由 GS0 ~ GS4 表示。数据集的每行对应一条训练数据。

4.2　算法模型

C4.5 决策树作为传统 ID3 决策树的优化算法，是使用信息增益率作为分裂属性指标进行决断的[24-25]。

信息熵是度量样本纯度的常用指标与计算信息增益的基础[26]。若肉鸡步态样本集合 T 中，第 k 类步态样本所占比例为 $p_k(k=1,2,3,4,5)$，则 T 的信息熵定义为

$$\text{Ent}(T) = -\sum_{k=1}^{5} p_k \text{lb} p_k \tag{13}$$

针对肉鸡速度属性 V 而言，有 M 个不同的取值 $\{V_1, V_2, \cdots, V_M\}$，若使用 V 属性对样本 T 进行划分，则会产生 M 个分支点，其中第 m 个分支点包括了 T 中所有在属性 V 上取值为 V_m 的样本，记为 T_m。则信息增益定义为

$$\text{Gain}(T,V) = \text{Ent}(T) - \sum_{m=1}^{M} \frac{|T_m|}{|T|}\text{Ent}(T_m) \tag{14}$$

本文中肉鸡步态相关的速度、步幅、步幅差等动态参数为连续值。以速度参数 V 为例，将 $\{V_1, V_2, \cdots, V_M\}$ 进行升序排列，基于划分点 t 将 T 分为子集 T_t^- 和 T_t^+，对于相邻属性取值 V_i 和 V_{i+1} 而言，t 在区间 $[V_i, V_{i+1})$ 中取任意值所产生划分结果相同。针对连续属性 V，设定

$$T_V = \frac{V_i + V_{i+1}}{2} \quad (1 \leqslant i \leqslant M-1) \tag{15}$$

此时，针对数据集纯度提升最高划分点，信息增益定义为

$$\text{Gain}(T,V) = \max_{t \in T_V} \text{Gain}(T,V,t) =$$

$$\max_{t \in T_V}\left(\text{Ent}(T) - \sum_{\delta \in \{-,+\}} \frac{|T_t^\delta|}{|T|}\text{Ent}(T_t^\delta)\right) \tag{16}$$

信息增益越大，表示按照属性 V 进行划分所获

得的“纯度增益”越大。当分支点过多时,取值数目更多的属性信息增益将变大,影响分类结果。使用增益率可以解决这一问题[27],信息增益率定义为

$$\mathrm{Gain}_{\mathrm{ratio}}(T,V) = \frac{\mathrm{Gain}(T,V)}{P(V)} \tag{17}$$

其中

$$P(V) = -\sum_{m=1}^{M}\frac{|T_m|}{|T|}\mathrm{lb}\frac{T_m}{T} \tag{18}$$

属性 V 的取值数目 M 越多,$P(V)$ 的值也会越大[28],从而消除取值数目对结果的影响。

4.3 训练过程

针对肉鸡步态评价方法,以速度、步幅、步幅差等连续动态属性为基础,进行决策树生成。特别地,步频是步幅和速度的派生变量,非独立变量,不单独作为决策树生长的属性。决策树生长流程如下:

(1)创建根节点,基于速度、步幅差、步幅等属性增益率,选择节点分裂属性。

(2)建立新节点,进行数据集的划分。

(3)判断节点是否达到生长停止条件,如果达到,决策树建立终止;否则重复步骤(1)、(2)。为防止过拟合,在本实验中设定样本数低于 5 时,终止递归。

(4)用 REP 后剪枝处理[29],自下而上处理每个节点,直到进一步修剪有害为止。修剪过程如下:①删除以此节点为根的子树,让节点变为叶子节点。②赋予这个节点最常见分类,若修剪后验证集的性能未出现降低,正式删除此节点,即以叶节点代替子树。

决策树生长流程图如图 10 所示。

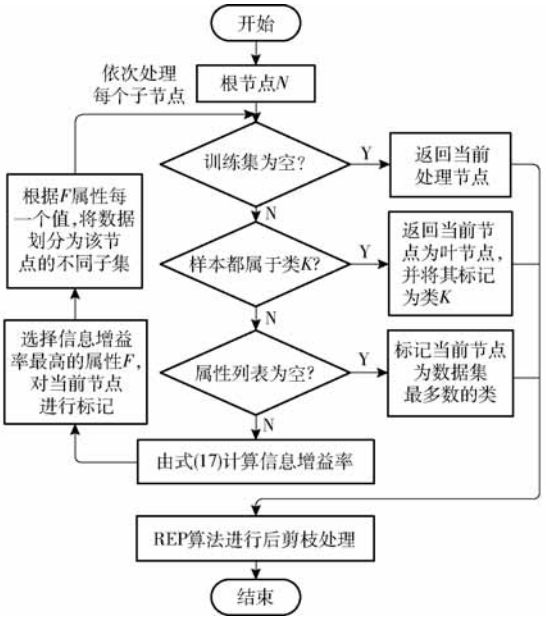

图 10 决策树生长流程图

Fig. 10 Flow chart of decision tree growth

4.4 训练方法

HENG[30]通过大量数据集、使用不同学习技术进行的研究表明,10 折交叉验证是可以获得较好误差估计的恰当选择,相比其他验证方式有简便高效的优势。本文采取“十折交叉验证”法将肉鸡步态数据集 T 随机分为 10 个包,轮流选取其中 9 个作为训练数据集 U,另外 1 个包作为测试数据集 I。训练结果如表 6 所示。

表 6 十折交叉法验证结果

Tab. 6 10-fold cross validation results

评级	训练次序	使用样本数量(含交叉验证)	准确率/%
GS0	训练集	1 098	66
	测试集	122	
	正确数	81	
GS1	训练集	1 188	71
	测试集	132	
	正确数	94	
GS2	训练集	1 044	74
	测试集	116	
	正确数	86	
GS3	训练集	756	98
	测试集	84	
	正确数	82	
GS4	训练集	594	95
	测试集	66	
	正确数	63	
总体	训练集	4 680	78
	测试集	520	
	正确数	406	

肉鸡的步幅差作为新的统计量首次加入到步态评分模型中,对易混淆级别的分类准确度做出了较高贡献。使用决策树 C4.5 模型识别的结果中,GS0 ~ GS2 等级别中分类准确率分别达到了 66%、71% 与 74%,弥补了以往 GS0 ~ GS2 级别区分准确率低的问题。

4.5 结果与分析

4.5.1 分类误差

本研究中在 1、7、8、10 训练轮次中出现了 GS3 与 GS4 中 1 或 2 例明显的错分结果,经调取原始数据研究发现,在对肉鸡鸡体进行分割时,若肉鸡步行中因腿部疾病导致身体匍匐在地面行走,某些情况下会导致团序列检测错误,从而发生错误分割致使提取到了错误的动态信息。

从整体测试结果来看,GS3 和 GS4 等极端分值的区分准确率仍显著高于 GS0 ~ GS2 等级别。通过训练结果的混淆矩阵看出,GS0 级别被误判为 GS1 级别的样本占到误判为其他级别的 90.4%,GS1 级别被误判为 GS0 级别的样本占到误判为其他级别的 72.6%,说明 GS0 和 GS1 之间的相互误判比较

高，这是因为其步态特征相近，即使综合新引入的步幅差变量来看，GS0 和 GS1 中的样本也存在很大的重合度。经调查得知，在实际人工操作中区分 GS0 与 GS1 的依据是“轻微缺陷引起的肉鸡不均衡步态”，而这种不均衡不仅体现在本文关注的肉鸡前进步态的水平方向（步幅、步幅差等），还会存在于垂直方向[31]（左右摇摆等），针对垂直方向特征本文暂未做研究。

4.5.2　基于分类结果的肉鸡步态健康预警

根据肉鸡步态评分标准，被评判为 GS0、GS1 的状态均不会影响肉鸡的生活与其对资源的争夺，而 GS2 ~ GS5 分别表示已经对肉鸡健康产生了从轻微到严重的影响[4]。大多数腿部出现疾病的病鸡都会出现从 GS1 ~ GS4 逐级恶化的现象[32]。一般被发现时肉鸡已经达到 GS3 及以上的严重病态，此时给予治疗已经效果不佳。

结合本文方法，对识别为 GS1 与 GS2 级别的肉鸡分别发出低级和高级疾病预警；GS3、GS4 级别肉鸡发出一级、二级疾病报告，从而实现提前发现肉鸡的腿部疾病，并争取最优时间对有发病可能的肉鸡实施预防和干预。

5　结束语

以决策树 C4.5 模型为基础提出了一种肉鸡步态自动定量评价方法，基于肉鸡步行的理想参数与异常指数，实现了对步态得分（GS）的预测。本文方法对 GS0 ~ GS4 评分的准确率依次为：66%、71%、74%、98%、95%，整体准确率为 78%。将 GS0、GS1、GS2 级别的鉴别准确率从 50% 提升到 70% 左右，为肉鸡早期腿部疾病的识别奠定基础。

参　考　文　献

1　THORP B H, DUFF S R. Effect of unilateral weight-bearing on pelvic limb development in broiler fowls: vascular studies[J]. Research in Veterinary Science, 1988, 44(2): 164 - 174.

2　BESSEI W. Welfare of broilers: a review[J]. Worlds Poultry Science Journal, 2006, 62(3): 455 - 466.

3　李鹏飞，姜敏，史超颖，等. 规模化养殖场肉鸡腿病情况调查[J]. 畜牧与兽医, 2016, 48(6): 116 - 118.
LI Pengfei, JIANG Min, SHI Chaoying, et al. An investigation of broiler leg disorders in large-scale farms[J]. Animal Husbandry and Veterinary Medicine, 2016, 48(6): 116 - 118. (in Chinese)

4　KESTIN S, KNOWLES T, TINCH A, et al. Prevalence of leg weakness in broiler chickens and its relationship with genotype[J]. Veterinary Record, 1992, 131(9): 190 - 192.

5　AYDIN A, BAHR C, BERCKMANS D. Automatic classification of measures of lying to assess the lameness of broilers[J]. Animal Welfare, 2015, 24(3): 2666 - 2670.

6　NAAS I D A, LOZANO L C M, MEHDIZADEH S A, et al. Paraconsistent logic used for estimating the gait score of broiler chickens[J/OL]. Biosystems Engineering, 2017. ISSN 1537 - 5110. Available online 8 December 2017. https://doi.org/10.1016/j.biosystemseng.2017.11.012.

7　DAWKINS M S, RUSSEL C, STEPHEN J R. Flock behaviour and chicken welfare[J]. Animal Behaviour, 2012, 84(1): 219 - 223.

8　AYDIN A. Using 3D vision camera system to automatically assess the level of inactivity in broiler chickens[J]. Computers and Electronics in Agriculture, 2017, 135: 4 - 10.

9　劳凤丹，滕光辉，李军，等. 机器视觉识别单只蛋鸡行为的方法[J]. 农业工程学报, 2012, 28(24): 157 - 163.
LAO Fengdan, TENG Guanghui, LI Jun, et al. Behavior recognition method for individual laying hen based on computer vision [J]. Transactions of the CSAE, 2012, 28(24): 157 - 163. (in Chinese)

10　赵凯旋. 基于机器视觉的奶牛个体信息感知及行为分析[D]. 杨凌：西北农林科技大学, 2017: 61 - 72
ZHAO Kaixuan. Dairy cows' information perception and behavior analysis based on machine vision[D]. Yangling: Northwest Agriculture and Forestry University, 2017: 61 - 72. (in Chinese)

11　毕敏娜，张铁民，庄晓霖，等. 基于色差信息多色彩模型的黄羽鸡快速分割方法[J/OL]. 农业机械学报, 2016, 47(12): 293 - 298, 308. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx? flag=1&file_no=20161236&journal_id=jcsam. DOI: 10.6041/j.issn.1000-1298.2016.12.036.
BI Minna, ZHANG Tiemin, ZHUANG Xiaolin, et al. Fast segmentation method of yellow feather chicken based on difference of color information in different color models[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2016, 47(12): 293 - 298, 308. (in Chinese)

12　刘龙申，沈明霞，柏广宇，等. 基于机器视觉的母猪分娩检测方法研究[J/OL]. 农业机械学报, 2014, 45(3): 237 - 242. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx? flag=1&file_no=20140339&journal_id=jcsam. DOI: 10.6041/j.issn.1000-1298.2014.03.039.
LIU Longshen, SHEN Mingxia, BO Guangyu, et al. Sows parturition detection method based on machine vision [J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2014, 45(3): 237 - 242. (in Chinese)

13　SERGEANT D, BOYLE R, FORBES M. Computer visual tracking of poultry[J]. Computers and Electronics in Agricultrue, 1998, 21: 1 - 18.

14　汤一平，王伟羊，朱威，等. 基于机器视觉的茶陇识别与采茶机导航方法[J/OL]. 农业机械学报, 2016, 47(1): 45 - 50. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx? flag=1&file_no=20160107&journal_id=jcsam. DOI: 10.6041/j.issn.1000-1298.2016.01.007.

TANG Yiping, WANG Weiyang, ZHU Wei, et al. Tea ridge identification and navigation method for tea-plucking machine based on machine vision[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2016, 47(1): 45 - 50. (in Chinese)

15 崔永杰,苏帅,王霞霞,等. 基于机器视觉的自然环境中猕猴桃识别与特征提取[J/OL]. 农业机械学报,2013,44(5):247 - 252. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx? flag=1&file_no=20130543&journal_id=jcsam. DOI:10.6041/j.issn.1000-1298.2013.05.043.
CUI Yongjie, SU Shuai, WANG Xiaxia, et al. Recognition and feature extraction of kiwifruit in natural environment based on machine vision[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2013, 44(5): 247 - 252. (in Chinese)

16 张弛,沈明霞,刘龙申,等. 基于机器视觉的新生仔猪目标识别方法研究与实现[J]. 南京农业大学学报,2017,40(1):169 - 175.
ZHANG Chi, SHEN Mingxia, LIU Longshen, et al. Newborn piglets recognition method based on machine vision [J]. Journal of Nanjing Agricultural University 2017, 40(1): 169 - 175. (in Chinese)

17 MITCHELL D R G, BERG J A. Development of an ellipse fitting method with which to analyse selected area electron diffraction patterns[J]. Ultramicroscopy , 2016, 160(1): 140 - 145.

18 杨倩,高晓阳,武季玲,等. 基于颜色和纹理特征的大麦主要病害识别研究[J]. 中国农业大学学报,2013,18(5):129 - 135.
YANG Qian, GAO Xiaoyang, WU Jiling, et al. Identification of barley diseases based on texture color feature[J]. Journal of China Agricultural University 2013, 18(5): 129 - 135. (in Chinese)

19 肖德琴,冯爱晶,杨秋妹,等. 基于视频追踪的猪只运动快速检测方法[J/OL]. 农业机械学报,2016,47(10):351 - 357,331. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx? flag=1&file_no=20161045&journal_id=jcsam. DOI:10.6041/j.issn.1000-1298.2016.10.045.
XIAO Deqin, FENG Aijing, YANG Qiumei , et al. Fast motion detection for pigs based on video tracking [J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2016, 47(10): 351 - 357, 331. (in Chinese)

20 AYDIN A O, CANGAR O, OZCAN S E, et al. Application of a fully automatic analysis tool to assess the activity of broiler chickens with different gait scores[J]. Computers and Electronics in Agriculture, 2010, 73(2): 194 - 199.

21 MORTENSEN A K, LISOUSKII P, AHREND P. Weight prediction of broiler chickens using 3D computer vision[J]. Computers and Electronics in Agriculture, 2016, 123: 319 - 326.

22 闫蓓,王斌,李媛. 基于最小二乘法的椭圆拟合改进算法[J]. 北京航空航天大学学报,2008,34(3):295 - 298.
YAN Bei, WANG Bin, LI Yuan. Opitmal ellipse fitting method based on least square principle[J]. Journal of Beijing University of Aeronautics and Astronautics, 2008, 34(3): 295 - 298. (in Chinese)

23 王辉,雷雨春,康峰,等. 基于机器视觉的樱桃外径检测[J/OL]. 农业机械学报,2012,43(增刊):246 - 249. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx? flag=1&file_no=2012s51&journal_id=jcsam. DOI:10.6041/j.issn.1000-1298.2012.S0.051.
WANG Hui, LEI Yuchun, KANG Feng, et al. Size detection for cherry fruit based on machine vision[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2012, 43(Supp.): 246 - 249. (in Chinese)

24 尚朝轩,王品,韩壮志,等. 基于类决策树分类的特征层融合识别算法[J]. 控制与决策,2016,31(6):1009 - 1014.
SHANG Chaoxuan, WANG Pin, HAN Zhuangzhi, et al. Feature-level fusion recognition algorithm based on analogy decision tree classification[J]. Control and Decision, 2016, 31(6): 1009 - 1014. (in Chinese)

25 周静平,李存军,史磊刚,等. 基于决策树和面向对象的作物分布信息遥感提取[J/OL]. 农业机械学报,2016,47(9):318 - 326,333. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx? flag=1&file_no=20160943&journal_id=jcsam. DOI:10.6041/j.issn.1000-1298.2016.09.043.
ZHOU Jingping, LI Cunjun, SHI Leigang, et al. Crops distribution information extracted by remote sensing based on decision tree and object-oriented method[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2016, 47(9): 318 - 326, 333. (in Chinese)

26 LIU Xiaoqian, LI Qianmu, LI Tao, et al. Differentially private classification with decision tree ensemble[J]. Applied Soft Computing, 2018, 62(1): 807 - 816.

27 WANG Yisen, XIA Shutao, WU Jia. A less-greedy two-term Tsallis Entropy Information Metric approach for decision tree classification[J]. Knowledge-Based Systems, 2017, 120(3): 34 - 42.

28 WANG Lihong, LI Qiang, YU Yanwei, et al. Region compatibility based stability assessment for decision trees[J]. Expert Systems with Applications , 2018, 150(9): 112 - 128.

29 周美琴,陈诗旭,袁鼎荣,等. 一种单位代价收益决策树剪枝算法[J]. 计算机工程与科学,2016,38(5):1023 - 1030.
ZHOU Meiqin, CHEN Shixu, YUAN Dingrong, et al. A pruning algorithm of decision tree based on unit cost gains[J]. Computer Engineering and Design, 2016, 38(5): 1023 - 1030. (in Chinese)

30 HENG L. Shrinkage tuning parameter selection in precision matrices estimation[J]. Journal of Statistical Planning and Inference, 2011, 141(8): 2839 - 2848.

31 AYDIN A. Development of an early detection system for lameness of broilers using computer vision [J]. Computers and Electronics in Agriculture, 2017, 136: 140 - 146.

32 郭盛洋,王连峰,李佳. 肉鸡腿病的发病原因和防治措施[J]. 现代畜牧科技,2017(5):146.
GUO Shengyang, WANG Lianfeng, LI Jia. The cause of broiler chicken leg disease and prevention measures[J]. Modern Livestock Science and Technology, 2017(5): 146. (in Chinese)