

基于机器学习的奶牛深度图像身体区域精细分割方法

赵凯旋 李国强 何东健

(西北农林科技大学机械与电子工程学院, 杨凌 712100)

摘要: 奶牛目标各区域的精细分割和识别能够提供精确的奶牛形体细节信息,是奶牛体形评价、姿态检测、行为分析和理解的前提和基础。为实现深度图像中奶牛头、颈、躯干和四肢等身体区域的精确分割,提出一种基于深度图像特征和机器学习的奶牛目标各区域精细分割方法。该方法以每个像素点在不同采样半径下的带阈值LBP序列为深度特征值,设置分类约束条件,用决策树森林机器学习方法实现奶牛各区域的精细分类。对10头奶牛的288幅侧视深度图像进行试验,结果表明,当采样半径分段数为30,决策树训练至20层时,奶牛整体各像素点的平均识别率为95.15%,较传统深度图像特征值有更强的细节信息提取能力,可以用较少参数实现对复杂结构的精确识别。

关键词: 奶牛; 目标检测; 肢干分割; 深度图像; 机器学习

中图分类号: S24; S823.9⁺1 文献标识码: A 文章编号: 1000-1298(2017)04-0173-07

Fine Segment Method of Cows' Body Parts in Depth Images Based on Machine Learning

ZHAO Kaixuan LI Guoqiang HE Dongjian

(College of Mechanical and Electronic Engineering, Northwest A&F University, Yangling, Shaanxi 712100, China)

Abstract: The recognition of cows' body parts is essential for providing accurate details of the cows' shape, which is the fundamental prerequisite for locomotion scoring, posture detection and behavioral quantifications. The objective was to develop a robust depth feature in order to reduce the difficulty in building the classifier and detect cows' body parts with higher accuracy. Therefore, a method for segmenting cows' body parts was proposed, including the head, neck, body, forelimbs, hind limbs and tail, with high accuracy on the basis of depth image processing and machine learning. The local binary patterns of each pixel under several sampling radii were used as the features with which the filtering rules were designed, and a decision forest was trained and tested to classify the pixels into six groups. Furthermore, totally 288 depth images were captured from 30 cows; 150 images were randomly selected to build three decision trees, and the rest images were used for testing. The results showed that when the number of sampling radii and training layers were 30 and 20, respectively, the recognition rate reached 95.15%. Among the cows' body parts, the recognition rate of tail was 54.97%, and the minimum recognition rate of other parts was 89.22%. In some cases that tail was too close to trunk to segment tail from trunk by human marker, the decision trees recognized the tail successfully. The average recognition time for pixel were 0.38 ms and 0.25 ms, and the recognition time for cow target were 20.30 s and 15.25 s for the conventional method and new method, respectively. This LBP-based depth image feature was translation-invariant and rotation-invariant and had fewer parameters. The results showed that the new method proposed was more effective in recognizing small and complex structures of the cow target with higher accuracy. Compared with the typical depth image features, the new feature employed was capable of extracting the details of cows' body and recognizing complex parts more accurately with fewer parameters and simple model.

Key words: cows; target detection; body segment; depth image; machine learning

收稿日期: 2017-01-02 修回日期: 2017-02-11

基金项目: 国家自然科学基金项目(61473235)

作者简介: 赵凯旋(1989—),男,博士生,主要从事基于图像处理的动物行为分析方法研究,E-mail: zhx@nwsuaf.edu.cn

通信作者: 何东健(1957—),男,教授,博士生导师,主要从事生物图像分析及识别、智能化检测与控制研究,E-mail: hdj168@nwsuaf.edu.cn

引言

视频分析技术能够对动物的行为进行自动监测和理解,是提高养殖管理智能化水平的重要手段^[1],已越来越多地应用于奶牛智能化养殖^[2-4]。近年来,奶牛疫病检测^[5-7]、身体评分^[8-9]、行为信息智能获取^[10-13]等受到了研究人员的广泛关注。奶牛目标捕捉是行为理解、智能监控等领域的技术基础和前提^[14]。然而受环境背景、光照变化等因素的影响^[15-16],目标检测仍是一个具有挑战性的课题^[17]。近年来发展的深度图像为实现奶牛目标跟踪与分割提供了新的思路和方法。微软公司发布的基于结构光技术的多传感器3D图像采集设备Kinect,能够同时获取目标像素点RGBD数据。由于其具有成本低、精度高、布设方便等优点,已在奶牛姿态检测和识别等相关研究中应用^[18]。

用深度图像进行姿态检测和识别时,首先需要分割头部、颈、躯干、四肢等区域,然后观测头部和四肢的动作,并对前后肢进行更细致的分析以识别姿态和行为。此外,奶牛身体区域的精细分割能够提高奶牛体形评价的自动化程度和精度。因此,需要一个鲁棒性强的奶牛肢干分割算法以区分奶牛身体各部分区域。目前针对奶牛的目标检测方法只能将视频图像中的奶牛作为一个整体进行分割^[17],奶牛身体区域的精细识别未见报道。人体区域精细分割的相关研究为解决这一问题提供了思路。SHOTTON等^[19]提出一种基于深度图像的人体肢干检测方法,该方法用预设的像素点深度值约束算子训练随机森林,然后预测未知像素点所属的身体区域,最后用mean shift方法估计人体关节并绘制骨架。ANTONIO等^[20]对上述方法进行了补充,使用图割理论(Graph-cuts theory)对随机森林算法得到的肢干分类结果进行了优化,且将时间和空间上邻近数据点的深度值用于确定目标边界。

上述2种算法中均使用同一种深度图像特征进行像素点识别,图像特征值原理简单易于实现,但每个特征值均包含4个未知量,每个像素点均需要多组偏置量进行约束性检测,因此导致分类器中可变参数较多、训练收敛速度慢。且该特征不具备旋转不变性,为得到较好的识别效果,需要大量不同姿态人体图像对模型进行训练。

为降低奶牛肢干分割与识别过程中模型训练的难度并提高识别精度,本文提出一种基于局部二值模式(Local binary patterns, LBP)原理的深度图像特征,该特征使用深度归一化后的带阈值LBP序列,具有深度图像目标细节信息提取能力。首先计算每

个像素点在不同采样半径下的统一LBP序列值,并以此作为分类条件,用决策树森林方法实现奶牛肢干的精细分类。

1 材料和方法

1.1 深度图像采集

2015年8月在陕西杨凌科元克隆股份有限公司奶牛养殖场,以荷斯坦奶牛为对象获取试验图像。奶牛挤奶结束后经过窄道行走至饮水池饮水,窄道两侧有高度为1.4 m的栏杆,奶牛、摄像机与围栏的相对位置如图1所示。在窄道一侧栏杆处布设Kinect V2体感器(Microsoft, 美国),Kinect V2安装于三角架上,传感器中心距离地面1.45 m,且深度图像传感器与奶牛行走方向平行。沿窄道边放置高度为0.1 m的长条形导向阻挡物,以保证奶牛行走至视野中心时与摄像机的距离约为2.5 m。由于奶牛从挤奶室行走至饮水池过程中已形成直线行走的习性,因此奶牛通过视野时始终与摄像机平行。阻挡物对后续图像分析与处理的影响可忽略不计。

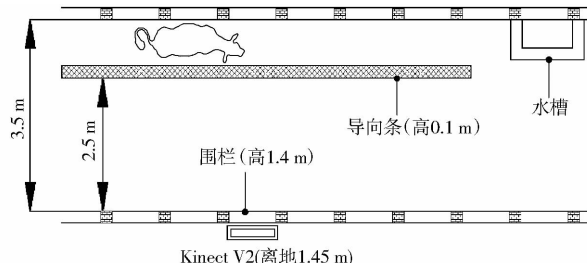


图1 试验设置与摄像机位置示意图

Fig. 1 Sketch of experiment setup and location of video camera

在室外晴天20:00—20:30时段对奶牛进行拍摄。通过便捷式计算机用Microsoft公司提供的Kinect Studio v2.0软件控制Kinect摄像机录制深度数据。奶牛个体全部出现在视频左侧时开始采集,持续采集到奶牛行走至视野右侧边缘。采集得到的码流数据保存在计算机硬盘内,该数据流文件中包含了每帧中颜色、深度、红外图像等信息。视频帧率为20帧/s,深度图像分辨率为512像素(水平)×424像素(垂直)。

数据处理计算机处理器为Intel Core i5-2400,主频为3.2 GHz,8 GB内存,500 GB硬盘。首先在Visual Studio 2012平台上将Kinect采集的数据流解析成深度数据文件并保存在硬盘中,然后用Matlab 2012a读取数据文件并进行后续算法的开发。

1.2 供试数据

对不同体形的奶牛,拍摄其在直线行走过程中的深度图像,奶牛经过通道耗时2 s左右。由于

帧率为 20 帧/s,每头奶牛通过视野过程中将产生约 40 幅图像。拍摄持续 7 d,挑选出无异常行为并连续行走的 10 头奶牛,剔除奶牛进入和离开视野时不完整的图像,最终共得到奶牛侧视深度图像 288 幅,平均每头奶牛约 30 幅。10 头奶牛中高大体形有 2 头、中等体形 5 头和矮小体形 3 头;严重跛行 1 头,中等跛行 3 头,正常行走 6 头。奶牛横穿摄像机视野的过程中,奶牛在图像中所处的位置以及四肢相对位置均不断发生变化。因此采集的数据对于奶牛体形、姿态和在图像中的位置均具有一定的多样性。

深度图像中奶牛与摄像机的距离固定为 2.5 m 左右,且摄像机与目标之间无前景遮挡。因此使用深度阈值分割出奶牛整体目标。然后参照奶牛肢体分类方法手工将 288 幅图像中的奶牛标识为头、颈、前肢(包括左前肢和右前肢)、后肢(包括左后肢和右后肢)、尾巴、躯干等区域,从中随机选取 150 幅作为训练集样本,其余图像作为测试集样本。

1.3 传统的深度图像特征

传统方法中,用被检测像素点 2 个偏移位置的深度差值作为该像素点的特征值^[19-20],其计算公式为

$$f_{\phi}(I, \mathbf{x}) = d_I\left(\mathbf{x} + \frac{\mathbf{u}}{d_I(\mathbf{x})}\right) - d_I\left(\mathbf{x} + \frac{\mathbf{v}}{d_I(\mathbf{x})}\right) \quad (1)$$

式中 $d_I(\mathbf{x})$ ——图像 I 中像素点 \mathbf{x} 处的深度

\mathbf{u}, \mathbf{v} ——2 个相对于 \mathbf{x} 的偏移向量

ϕ —— \mathbf{u} 和 \mathbf{v} 的一组值

图 2 为 2 个特征在不同像素点上的检测原理。图中特征值 f_{ϕ_1} 向上检测该像素点与邻近点的关系。假设背景点的深度为 $+\infty$,由于偏移点分别在目标和背景上,则 f_{ϕ_1} 的值为 $+\infty$ 。 f_{ϕ_2} 能够检测出类似于四肢的细长形结构,对于图中检测薄片结构的 f_{ϕ_2} 特征值,2 个偏移点均落到背景上,因此 f_{ϕ_2} 的值为零。但对于同一特征值,当奶牛肢体发生旋转时,其值将发生变化,因此传统的深度图像特征值不具备旋转不变性。此外,传统方法中 \mathbf{u}, \mathbf{v} 值的选取具有很强的盲目性,通常使用随机数生成其取值空间里的可能值,然后对其进行检视和筛选,以确定有效的 \mathbf{u}, \mathbf{v} ,选取过程耗时耗力。

由式(1)可知,该深度图像特征值包含 2 个偏移向量共 4 个变量,用以计算目标像素点邻近的 2 个点的深度差值,以评估目标像素点所在的身体区域。结合式(1)和图 2 可知,传统的图像特征值只能提供目标区域的微弱信息,对分类器的性能要求较高,且需要大量训练数据以覆盖各结构的不同旋转角度,模型训练成本较高,难以保证精度。

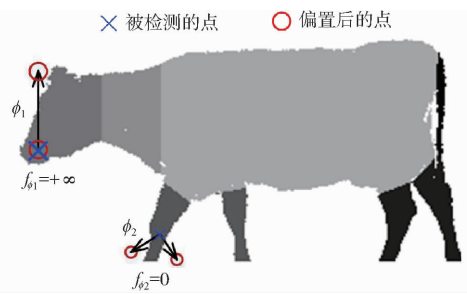


图 2 传统深度图像特征值的检测原理

Fig. 2 Detection principle of original depth image features

1.4 基于 LBP 的深度图像特征

考虑到式(1)中传统深度图像特征值不具备旋转不变性,且采样精度难以实现细小结构的识别,为更有效地提取深度图像中的像素点信息,本文提出使用基于 LBP 原理的深度图像局部二值模式(Deep image LBP, DILBP)。LBP 是一种典型的纹理特征值,其衡量了某一点邻近区域内灰度值跳变的模式^[16],原始的 LBP 算子定义为:在采样半径 $p \times p$ 的窗口内,以窗口中心像素为阈值,将相邻的 2^p 个像素的灰度与其进行比较,若周围像素值大于中心像素值,则该像素点的位置被标记为 1,否则为 0。

本文对原始 LBP 中的判定原则加以改进以提高 LBP 对几何结构的敏感性,同时对采样半径进行深度归一化处理,使其具有深度不变性。改进后的 DILBP 为

$$D_{p,R} = \sum_{p=0}^{p-1} s\left(\frac{|d_p - d_c|}{d_c}\right) 2^p \quad (2)$$

式中 $D_{p,R}$ ——深度图像局部二值模式(DILBP)

P ——采样点数 R ——采样半径

d_p ——第 p 个采样点的深度

d_c ——中心点的深度

s ——判定函数

由于摄像机成像平面与奶牛侧身基本平行,因此在深度图像中属于奶牛同一身体区域像素点的深度差异较小,而目标与背景之间的差值较大。故本文用带阈值的判定函数,以判定采样点与中心点是否同属于奶牛身体区域。

$$s(x) = \begin{cases} 1 & (x \leq T) \\ 0 & (x > T) \end{cases} \quad (3)$$

式中 T ——同一身体区域内的深度差异度阈值

本文使用统一的 LBP 模式,因此特征值计算结果为 58 个固定的模式之一^[21]。DILBP 模式中的采样半径为

$$R = \frac{R_i}{d_I(\mathbf{x})} \quad (4)$$

式中 R_i ——真实空间内的采样半径

真实空间内的半径 R_i 乘以 $1/d_I(\mathbf{x})$ 以确保图像

特征具有深度不变性,能够消除目标与摄像机之间距离对检测结果的影响。本文 DILBP 序列中的 1 和 0 分别表示该采样点与中心点是否属于同一肢干。如图 3 所示,对于某一肢干,其内部的像素点在特定的采样半径 R 下,计算得到的 DILBP 序列基本相同。然而不同肢干之间,其特定的采样半径和 DILBP 序列均不同。因此,多个采样半径下的 DILBP 序列值,能够将不同身体区域内的像素点进行分类。

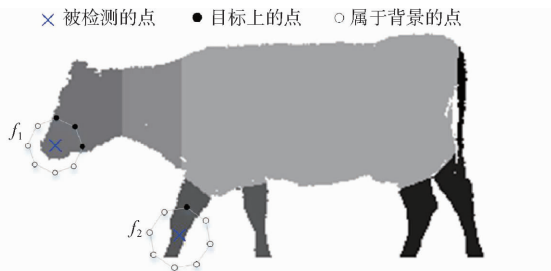


图 3 基于 DILBP 模式的深度图像特征检测原理
Fig. 3 Detection principle of depth image features based on DILBP

由图 3 可知,当被检测的四肢发生旋转时,特定的采样点总是落在肢干区域内,由于统一的 LBP 模式计算 LBP 序列内 0 和 1 的跳变次数,因此本文提出的基于 LBP 的深度图像特征值继承了 LBP 的旋转不变性。

传统的深度图像特征使用 4 个参数检测目标点与邻近 2 个偏置点的位置关系。而本文提出的 DILBP 图像特征仅通过一个参数即可对目标点周围的 8 个点进行检测,且检测结果为固定的 58 个模式值,降低了参数值的选取难度。

1.5 随机决策森林

随机决策森林是一种经典的机器学习方法。考虑到奶牛目标肢干分割属于先验的模式识别问题,因此用 DILBP 特征构建随机决策森林。对于给定的带分类标签的奶牛目标深度图像,决策树的构建过程如下^[20]:

(1) 决策树中第 i 个分支节点包含一组判定参数 R_i 和 L_i ,其值为各自取值空间内的随机数。决策树训练完成后,每个分支节点参数的 R 和 L 将取得最优值,能够将流入该节点的样本点划分成 2 组,且分类前后的熵增益最大。

(2) 设 Q 为训练图像数据集中所有参与训练样的像素点的集合, Q 对于分类标签 c 的概率分布计算公式为

$$P_Q(c) = \frac{h_Q(c)}{|Q|} \quad (c \in C) \quad (5)$$

式中 $h_Q(c)$ ——集合 Q 相对于标签 c 的直方图

$|Q|$ ——集合 Q 中元素的个数

P_Q ——集合 Q 的概率分布函数

C —— Q 中的标签集合

(3) 对于第 i 个未被处理的节点,若该节点内的样本点属于同一类,说明该节点是叶子节点,直接跳过处理;否则该节点为分支节点。假设流入该节点的样本集为 Q_i ,则随机选取一组 R 和 L 参数值将 Q_i 划分为 Q_L 和 Q_R 2 部分,即

$$\begin{cases} Q_L(\phi) = \{ (I, \mathbf{x}) \mid \text{DILBP}_{p,R}(I, \mathbf{x}) = L \} & (\phi = (R, L)) \\ Q_R(\phi) = C_{Q_i}(Q_L) \end{cases} \quad (6)$$

式中 $\text{DILBP}_{p,R}(I, \mathbf{x})$ ——按式(2)计算图像 I 中像素点 \mathbf{x} 处的深度图像局部二值模式特征值

$C_{Q_i}(Q_L)$ ——计算 Q_L 在 Q_i 中补集的函数

然后用式(5)分别计算 Q_L 和 Q_R 的概率分布函数得到 P_{Q_L} 和 P_{Q_R} 。

(4) 估计当前节点的最佳分类参数 ϕ^* ,该参数保证分类前后的信息熵增益取得最大值。分类前后的信息熵增益计算公式为

$$\begin{cases} H(Q) = - \sum_{c \in C} P_Q(c) \ln(P_Q(c)) \\ G_Q(\phi) = H(Q) - \frac{|Q_L(\phi)|}{|Q|} H(Q_L(\phi)) - \frac{|Q_R(\phi)|}{|Q|} H(Q_R(\phi)) \end{cases} \quad (7)$$

式中 $H(Q)$ ——集合 Q 的信息熵

$G_Q(\phi)$ ——使用参数 ϕ 对集合 Q 进行划分前后的信息熵增益

将 ϕ 的所有可能值代入式(7),最终通过比较得到最佳的分类参数 ϕ^* ,其值使得 $G_Q(\phi)$ 在其定义域内取得最大值。

(5) 在每个新生成的节点中执行步骤(3)、(4),直到决策树层数达到预设值时中止训练,未处理的节点视为叶子节点,并保存该节点样本的概率分布函数。

一旦训练完成,对于某一个待识别的像素点,从树的根节点开始,根据式(6)和该节点所保存的分类参数判定该像素点的路径,直到该像素点到达某一叶子节点,并得出分类结果。

当参与决策树训练的样本数过多时,训练过程将消耗大量的计算资源。故将训练数据随机划分为奇数组,分别训练多个决策树构成决策森林。将待检测样本分别输入每个决策树,得到每棵树最终叶子节点处的概率分布函数,然后对其进行合并

$$P(c|I, \mathbf{x}) = \frac{1}{T} \sum_{i=1}^T P_i(c|I, \mathbf{x}) \quad (c \in C) \quad (8)$$

式中 $P_i(c|I, \mathbf{x})$ ——图像 I 中像素点 \mathbf{x} 在第 i 棵树中最终达到的叶子节点所保存概率分布函数

最终, 样本的识别结果为合并后的概率分布函数中所占比例最大的类。

2 试验结果与分析

2.1 奶牛各区域识别结果

将训练集中 150 幅图像平均分成 3 组分别训练 3 棵决策树, 每幅图像中奶牛目标上的像素点均参与训练。测试集图像中每个像素点均进入 3 棵树进行决策得到的结果按式 (8) 进行合并后, 得到最终的分类结果。构建决策树过程中, 采样半径的最大值为奶牛的体高, 采样半径的分段数 B 决定了深度特征值的精度。分段数最大, 候选的分类含参数越多。决策树的最大层数 Y 决定了分类的精度, 层数越大则决策树对于训练样本的识别精度越高, 但决策树的泛化能力将会降低。为确定最优训练参数, 对不同参数组合进行了比较试验, 表 1 为本文算法在不同参数下的平均识别率。

表 1 本文算法在不同参数组合下的平均识别率

训练层数 Y	分段数 B			
	10	20	30	40
10	91.41	92.13	93.10	93.16
15	92.57	94.37	94.94*	94.99*
20	92.22	94.71	95.15*	95.17*

注: * 识别率最高的前 4 种组合。

由表 1 可知, 提高采样半径 R 的分段数 B 能提高检测精度, 但在训练层数为 10 层时, 分段数 40 相对分段数 30 的识别效果提升并不明显。这是由于分段数较多导致决策树中待训练的参数增加, 而较浅的决策树难以训练出最优分类参数。为进一步分析分段数和训练层数对识别结果的影响, 表 2 给出了平均识别率最高的前 4 种组合下的奶牛身体各区域识别结果。

由表 2 可知, 当决策树由 15 层训练到 20 层后, 奶牛头部识别精度有微小提升, 颈、躯干等简单粗大结构的识别率均有所下降, 前肢、后肢和尾巴等细小复杂结构的识别率得到明显提升。尤其是尾巴的识别率提高超过 10 个百分点。表明多层决策树对细小结构的识别能力较强, 但对粗大结构的泛化能力受到影响。

表 2 4 种组合下奶牛各区域的识别结果

Tab.2 Recognition results of cows' body parts under

参数组合	four combinations					奶牛整体	
	头	颈	躯干	前肢	后肢		尾巴
$B=30, Y=15$	93.20	90.13	97.93	91.14	86.96	44.07	94.94
$B=30, Y=20$	93.43	89.22	97.80	91.61	90.69	54.97	95.15
$B=40, Y=15$	93.21	90.56	98.01	91.34	85.99	45.45	94.99
$B=40, Y=20$	93.47	89.81	97.81	92.23	89.21	56.62	95.17

在相同的训练层数下, 分段数分别为 30 和 40 时, 各区域的识别精度差异很小, 然而后肢的识别率反而下降, 表明分段数越多, 决策树泛化能力随层数的下降速度越快。

由于尾巴所包含的奶牛身体及行为信息较少, 故本文以分段数为 30、层数为 20 作为最优参数, 进一步研究决策树的分割性能。图 4 为对不同体形和姿态下的奶牛各肢干人工分割及决策树分割结果。从上到下分别为高体深壮硕正常行走奶牛 A、低体深瘦小中等跛行奶牛 B、样本中较少的严重跛行奶牛 C。

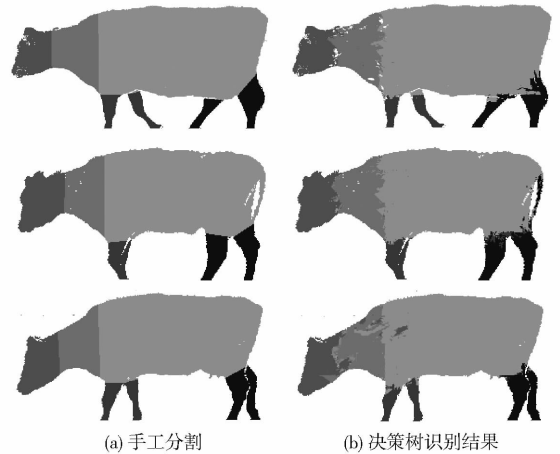


图 4 不同奶牛的分割效果示例图

Fig.4 Recognition results of different cows

图 4 中 A、C 奶牛尾巴已经被人工去除, 而奶牛 B 由于尾巴过于靠近躯干, 故人工分割时均未将其分割出来。但决策树能正确地将奶牛 B 大部分尾巴像素点识别出来。奶牛 A、B、C 体深、头部位置均不同, 用决策树均得到较高的识别精度, 表明本文算法对于奶牛的体形、姿态具有很好的鲁棒性。四肢错误识别的像素点主要集中于四肢与躯干接合处的过渡区域。

2.2 识别率与各区域样本比例的关系

表 3 为奶牛身体各区域的识别率与该区域所占奶牛整体目标比例的比较结果。

由表 3 可知, 奶牛身体各区域所占比例和识别率呈正相关关系, 占奶牛目标比例最大的躯干区域

表3 奶牛身体各区域所占比例和识别率

Tab.3 Recognition rate and proportion of

cows' body parts							%
区域名称	头	颈	躯干	前肢	后肢	尾巴	
区域比例	9.38	11.75	66.48	4.36	7.25	0.76	
识别率	93.43	89.22	97.80	91.61	90.69	54.97	

识别率最高;占奶牛目标比例最小的尾巴区域识别率最低。这是由于在训练分类器时,各个类别内样本数分布不均造成的。当训练决策树时,训练图像中所有像素点均参与训练,因此流入第1个节点的数据中躯干区域样本点比例最大,由于决策树总是以最大熵增加值作为评估依据,因此占绝对数量的躯干类将在前几层被优先分割出来。虽然颈部所占比例略高于头部,但识别率比头部低4.21个百分点。这是由于颈部作为头部和躯干的连接区域,形状复杂且变形度较大,同时颈部靠近躯干区域,受大容量样本的影响大,因此识别率低于头部。后肢受尾巴的影响,识别率略低于前肢。此外,由于四肢、尾巴等区域自身的像素点数较少,少量错误识别的像素点即会导致较大的识别错误率。

尾巴区域的识别率低于55%,这是由于参与训练的尾巴区域的像素点太少所致。

2.3 与 Kinect 原始方法的比较

为比较本文方法相对于 Kinect 原始方法的目标分割效果,根据文献[14]中的深度图像特征值,用相同的训练数据和测试数据进行对比试验。为保证与本文算法相近的参数数量,试验中传统方法的偏置点为40组,待比较阈值为20个。2种方法的分割结果如表4所示。

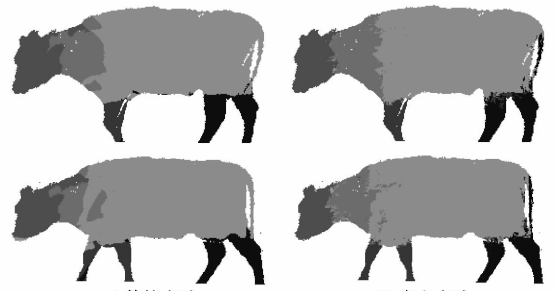
表4 传统方法与本文方法识别结果比较

Tab.4 Comparison of recognition results of original method and proposed method

方法	头	颈	躯干	前肢	后肢	尾巴	奶牛整体
传统方法	97.17	78.89	97.50	72.33	88.57	13.14	92.72
本文方法	93.43	89.22	97.80	91.61	90.69	54.97	95.15

如表4所示,传统方法中头部的识别精度高于本文方法,躯干识别精度与本文接近。本文方法中颈、前肢的识别精度提高了10个百分点以上。尤其是尾巴识别精度由13.14%提高到54.97%。2种方法的识别效果示例如图5所示。

由图5可知,传统方法中虽然头部的识别精度较高,但头与颈结合处出现了较多的交叉分割,不利于定位头颈结合关节。两头奶牛的识别结果显示,传统方法均将尾巴错误识别为躯干。综上所述,本文方法交叉识别少,较传统方法具有更强的细节信息提取能力,能够以较少的参数,实现对复杂结构的



(a) 传统方法

(b) 本文方法

图5 传统方法和本文方法识别效果示例

Fig.5 Two examples of recognition results of two methods

精确识别。

为比较传统方法与本文方法的效率,每个像素点进入决策树森林到得出识别结果所需要的时间定义为单个像素点的识别耗时,奶牛目标上的所有像素点的总体耗时定义为奶牛目标的识别耗时。传统方法与本文算法的平均训练耗时、决策树节点数量以及平均识别耗时如表5所示。

表5 传统方法与本文方法训练与识别效率比较

Tab.5 Comparison of training and recognition efficiency of original method and proposed method

方法	训练耗时/ h	节点数量	像素点 耗时/ms	奶牛目标 识别耗时/s
传统方法	11	20 111	0.38	20.30
本文方法	13	15 107	0.25	15.25

由于本文方法中图像特征较传统方法复杂,因此训练耗时较长。但由于参数利用率高,决策树中的节点数量较少,因此能够以较短路径达到叶子节点,故识别效率更高。由于单幅图像中所有的像素点均参与识别,故图像的识别耗时为15s左右。实际进行奶牛各区域实时检测时,可随机选取图像中10%的像素点进行检测,即可得到奶牛各区域的大体轮廓,则可将单幅图像识别耗时降低至1.5s,且用C++语言实现识别算法可进一步提高识别效率。

3 结论

(1)提出一种深度归一化后的带阈值LBP序列作为深度图像特征,与传统深度特征值^[14]相比,该特征值具有平移和旋转不变性,且参数利用率高,能以较少的训练图像实现较高识别精度。

(2)用本文深度图像特征值作为分类条件构建决策树森林,可实现奶牛肢干的精细分割和识别。奶牛各身体区域平均识别率为95.15%,除尾巴外的奶牛各个主要身体区域均保持较高的识别率。

(3)本文方法交叉识别少,较 Kinect 原始方法^[14]具有更强的细节信息提取能力,能够实现对复杂结构的精确识别。

参 考 文 献

- 1 何东健,刘冬,赵凯旋. 精准畜牧业中动物信息智能感知与行为检测研究进展[J/OL]. 农业机械学报, 2016, 47(5): 231 – 244. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20160532&journal_id=jcsam. DOI: 10.6041/j.issn.1000-1298.2016.05.032.
- HE Dongjian, LIU Dong, ZHAO Kaixuan. Review of perceiving animal information and behavior in precision livestock farming[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2016, 47(5): 231 – 244. (in Chinese)
- 2 HOFFMANN G, SCHMIDT M, AMMON C, et al. Monitoring the body temperature of cows and calves using video recordings from an infrared thermography camera[J]. Veterinary Research Communications, 2013, 37(2): 91 – 99.
- 3 CHAPINAL N, TUCKER C B. Validation of an automated method to count steps while cows stand on a weighing platform and its application as a measure to detect lameness[J]. Journal of Dairy Science, 2012, 95(11): 6523 – 6528.
- 4 BRUYERE P, HETREAU T, PONSART C, et al. Can video cameras replace visual estrus detection in dairy cows[J]. Theriogenology, 2012, 77(3): 525 – 530.
- 5 VIAZZI S, BAHR C, SCHLAGETER-TELLO A, et al. Analysis of individual classification of lameness using automatic measurement of back posture in dairy cattle[J]. Journal of Dairy Science, 2013, 96(1): 257 – 266.
- 6 PLUK A, BAHR C, POURSAHERI A, et al. Automatic measurement of touch and release angles of the fetlock joint for lameness detection in dairy cattle using vision techniques[J]. Journal of Dairy Science, 2012, 95(4): 1738 – 1748.
- 7 POURSAHERI A, BAHR C, PLUK A, et al. Real-time automatic lameness detection based on back posture extraction in dairy cattle: shape analysis of cow with image processing techniques[J]. Computer and Electronics in Agriculture, 2010, 74(1): 110 – 119.
- 8 AZZARO G, CACCAMO M, FERGUSON J D, et al. Objective estimation of body condition score by modeling cow body shape from digital images[J]. Journal of Dairy Science, 2011, 94(4): 2126 – 2137.
- 9 BEWLEY J M, PEACOCK A M, LEWIS O, et al. Potential for estimation of body condition scores in dairy cattle from digital images[J]. Journal of Dairy Science, 2008, 91(9): 3439 – 3453.
- 10 PORTO S M C, ARCIDIACONO C, ANGUZZA U, et al. A computer vision-based system for the automatic detection of lying behaviour of dairy cows in free-stall barns[J]. Biosystems Engineering, 2013, 115(2): 184 – 194.
- 11 MIEDEMA H, COCKRAM M, DWYER C, et al. Behavioural predictors of the start of normal and dystocic calving in dairy cows and heifers[J]. Applied Animal Behaviour Science, 2011, 132(1 – 2): 14 – 19.
- 12 CANGAR O, LEROY T, GUARINO M, et al. Automatic real-time monitoring of locomotion and posture behaviour of pregnant cows prior to calving using online image analysis[J]. Computers and Electronics in Agriculture, 2008, 64(1): 53 – 60.
- 13 赵凯旋,何东健,王恩泽. 基于视频分析的奶牛呼吸频率与异常检测[J/OL]. 农业机械学报, 2014, 45(10): 258 – 263. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20141040&journal_id=jcsam. DOI:10.6041/j.issn.1000-1298.2014.10.040.
- ZHAO Kaixuan, HE Dongjian, WANG Enze. Detection of breathing rate and abnormality of dairy cattle based on video analysis[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2014, 45(10): 258 – 263. (in Chinese)
- 14 赵凯旋,何东健. 基于卷积神经网络的奶牛个体身份识别方法[J]. 农业工程学报, 2015, 31(5): 181 – 187.
- ZHAO Kaixuan, HE Dongjian. Recognition of individual dairy cattle based on convolutional neural networks[J]. Transactions of the CSAE, 2015, 31(5): 181 – 187. (in Chinese)
- 15 刘冬,赵凯旋,何东健. 基于混合高斯模型的移动奶牛目标实时检测[J/OL]. 农业机械学报, 2016, 47(5): 288 – 294. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20160539&journal_id=jcsam. DOI:10.6041/j.issn.1000-1298.2016.05.039.
- LIU Dong, ZHAO Kaixuan, HE Dongjian. Real-time target detection for moving cows based on gaussian mixture model[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2016, 47(5): 288 – 294. (in Chinese)
- 16 何东健,孟凡昌,赵凯旋,等. 基于视频分析的犊牛基本行为识别[J/OL]. 农业机械学报, 2016, 47(9): 294 – 300. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20160940&journal_id=jcsam. DOI:10.6041/j.issn.1000-1298.2016.09.040.
- HE Dongjian, MENG Fanchang, ZHAO Kaixuan, et al. Recognition of calf basic behaviors based on video analysis[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2016, 47(9): 294 – 300. (in Chinese)
- 17 ZHAO K, HE D. Target detection method for moving cows based on background subtraction[J]. International Journal of Agricultural and Biological Engineering, 2015, 8(1): 42 – 49.
- 18 VIAZZI S, BAHR C, Van HERTEM T, et al. Comparison of a three-dimensional and two-dimensional camera system for automated measurement of back posture in dairy cows[J]. Computers and Electronics in Agriculture, 2014, 100(1): 139 – 147.
- 19 SHOTTON J, SHARP T, KIPMAN A, et al. Real-time human pose recognition in parts from single depth images[J]. Communications of the ACM, 2013, 56(1): 116 – 124.
- 20 ANTONIO H V, NADEZHDA Z, ALEXANDER M, et al. Graph cuts optimization for multi-limb human segmentation in depth maps[C]//2012 IEEE Conference on Computer Vision & Pattern Recognition, 2012:726 – 732.
- 21 NANNI L, BRAHNAM S, LUMINI A. A simple method for improving local binary patterns by considering non-uniform patterns[J]. Pattern Recognition, 2012, 45(10): 3844 – 3852.