

基于视觉词典法的母牛产前行为识别*

温长吉^{1,2} 王生生² 赵昕¹ 王恣³ 马丽¹ 刘依婷²

(1. 吉林农业大学信息技术学院, 长春 130118;

2. 吉林大学符号计算与知识工程教育部重点实验室, 长春 130012;

3. 陶森大学计算机与信息科学系, 马里兰 21252)

摘要: 母牛活动方式及规律的信息获取是特定阶段进行必要人工干预的重要判据之一, 针对已有通过外置传感器获取信息手段的局限性, 提出一种改进时空局部二值模式用于特征描述, 构建视觉词典实现对视频中母牛基本行为的识别, 并最终通过统计产前特定基本行为发生频次来揭示临产期规律。试验表明所提方法对 90 组规定视角下母牛产前行走、侧卧和回望等典型行为的平均正确识别率约为 94.6%; 对 30 组水平随机视角下拍摄的上述 3 种行为平均正确识别率约为 88.3%。通过采集母牛临产期分娩前后约 13 h 的视频, 提出的算法识别母牛平均执行回望和侧卧的频次依次为 30 次和 21.2 次, 行为发生频次体现为前低后高, 在分娩前回望频次达到峰值 8.8 次, 侧卧频次明显减低至 2.2 次。试验结果证明, 视觉词典法能实现对母牛产期行为的识别与规律的量化描述。

关键词: 母牛 产前行为 视觉词典法 改进局部二值模式 时空兴趣点

中图分类号: TP391.41; S823.4 **文献标识码:** A **文章编号:** 1000-1298(2014)01-0266-09

引言

母牛临产前期行为是特定条件下生理和心理行为的反应, 一般表现为起卧频繁, 经常变换体位(漫步行走、侧卧和头部频繁回顾腹部等)等典型行为。因此该时期母牛行为活动方式及其规律的信息获取是进行必要人工干预的重要判据之一, 为有效降低难产率, 提高犊牛成活率提供帮助。

传统人工监测的方法具有主观性强、人工成本高、工作强度大、易疲劳和误诊率高等问题。为了克服人工监测方法的弊端, 国内外学者以自动化监测相关技术开展广泛研究^[1-5]。通过将三轴加速计、蓝牙设备等传感器部件安置在禽畜不同身体部位上, 结合数据挖掘方法对记录数据进行分析实现禽畜某些特定阶段生命体征和行为的识别, 取得较为满意的结果。但是在试验过程中, 研究人员也同时发现上述方法存在一定的局限性, 即侵入式采集方式可能影响动物的行为, 外置部件易受环境浸染, 体位遮挡干扰信号传输等。

视觉是信息获取的主要来源, 机器视觉技术具有实时、非侵入、免打扰等特点, 因此在作物种植等

领域得到广泛应用^[6-8]。但是, 该技术应用于禽畜养殖研究报道较少, 已有的研究也多集中于禽类等小型养殖领域^[9-11], 对于猪牛羊等大型畜牧类养殖的研究更是少有报道。

综上, 本文将视觉词典法用于母牛基本行为识别, 通过对产前基本行为发生频次规律的统计, 最终实现对产前行为特征识别与规律发现。为自动化肉奶牛精准养殖系统子模块——智能视觉监控的研发提供可实施技术支持。

1 材料与方法

1.1 试验材料

1.1.1 试验对象

试验视频采集自农业部现代农业产业技术体系—吉林农业大学校外肉牛繁育技术示范与推广基地, 采集时间跨度从 2011 年 3 月至 2012 年 9 月。视频采集对象为随机选取的 37 头改良西门塔尔牛, 分别为距人工预产期 7 d 左右转入单独分娩区进行饲养, 卧盘铺垫干牛粪末或干草垫, 并保证牛舍温度维持在 15~17℃, 该时期遵照体重喂饲的原则每天喂食 3 次。

收稿日期: 2013-07-28 修回日期: 2013-09-12

* 国家自然科学基金资助项目(61133011, 61170092)、国家高技术研究发展计划(863 计划)资助项目(2006AA10Z245)、吉林省科技发展规划资助项目(20100181, 201105068, 201101114, 20120665)、吉林大学科学前沿与交叉学科创新资助项目(200903178)和吉林农业大学科研启动基金青年基金资助项目(2012037)

作者简介: 温长吉, 讲师, 博士生, 主要从事模式识别、农业信息化研究, E-mail: chagou2006@163.com

通讯作者: 王生生, 教授, 博士生导师, 主要从事空间推理、模式识别、精准农业研究, E-mail: wss@jlu.edu.cn

1.1.2 试验环境

视频采集设备主要包括: Panasonic WV - CP240/G 型彩色 CCD 摄像机,分辨率为 752 像素(水平) × 582 像素(垂直),大华 DAHUADH - DVR0804LE - AS8 型硬盘刻录机,后端处理器为 Intel 酷睿 i5 - 3210 M,主频为 2.5 GHz,8 GB 内存,500 GB 硬盘便携式计算机。系统软件模块采用 IMAQ Vision,算法开发平台为 Matlab 2011b。

视频采集方式为固定成像设备分别从正前方、侧面、后侧和水平方向随机视角采集母牛侧卧、行走和回望行为视频。围产期产前行为视频采集单帧示意图如图 1 所示。

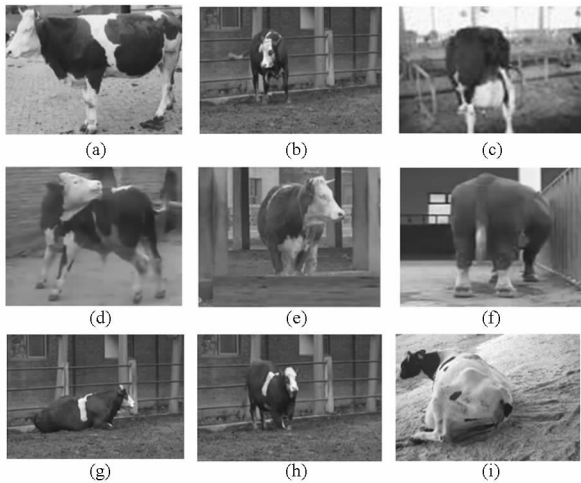


图 1 行为示意图

Fig.1 Behavior diagram

- (a) 行走侧面 (b) 行走正面 (c) 行走后面
- (d) 回望侧面 (e) 回望正面 (f) 回望后面
- (g) 侧卧侧面 (h) 侧卧正面 (i) 侧卧后面

1.2 基本方法

依据专家经验和观察分析,母牛临产前期的基本行为一般表现为起卧频繁和频繁回顾腹部,因此

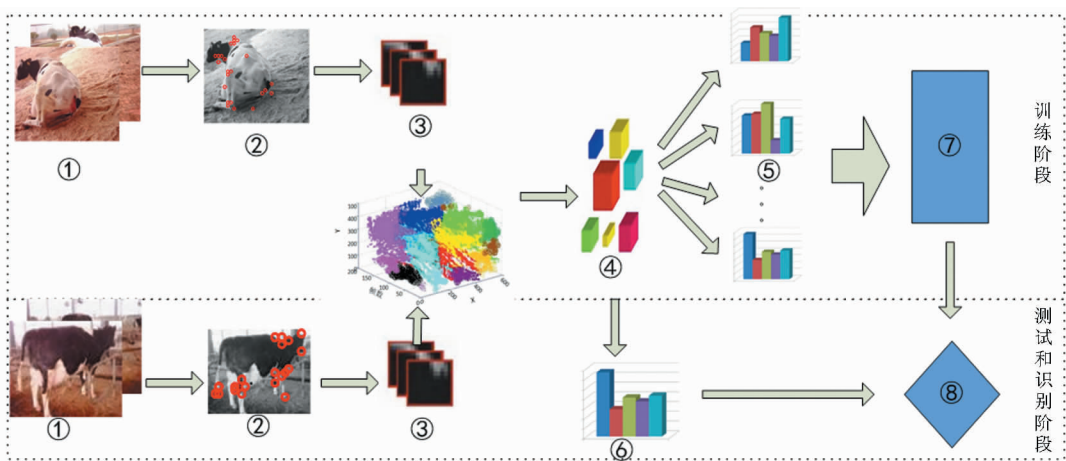
有效识别视频中基本行为在临产前的频次变化规律是母牛产前行为识别与分析的关键。首先对基本行为(行走、侧卧和回望)利用视觉词典法进行识别;然后在基本行为准确识别的基础上,结合视频分割方法对采集的母牛临产前测试视频中的基本行为进行识别;最后利用识别结果统计单位时间内母牛产前上述基本行为变化的频次,从而实现母牛临产前行为规律的识别与分析。具体方法步骤为:①提出一种改进时空局部二值模式特征描述子对时空兴趣点提取的区域进行特征描述。②对训练视频提取的局部特征量进行聚合,聚合中心作为视觉单词,所有视觉单词构成视觉词典。③用视觉单词描述视频行为,构建视觉单词直方图,用以训练分类器,并用于侧卧、行走和回望基本行为的分类识别。④对母牛产前测试视频单帧图像进行特征量化描述和聚类,以行为类别为标签实现视频的语义分割。⑤在上述基本行为识别的基础上,利用已训练分类器实现对视频段行为进行分类识别,并对单位时间内基本行为发生频次进行统计。实现母牛产前行为规律的识别与分析。具体算法流程图如图 2 所示。

1.2.1 视频行为特征提取与描述

为了有效对视频中基本行为进行特征提取与描述,提出利用改进时空局部二值模式算子对时空兴趣点进行量化描述的方法。

1.2.1.1 基本理论

时空兴趣点作为一种局部特征检测算子,由于在视频表示上的稀疏性,运动尺度、视角等的鲁棒性,因此在人体行为识别领域得到广泛的应用。单个时空兴趣点体现的是视频中局部区域灰度值的突变,因此无法描述行为的全局特性;而由视频中检测到的时空兴趣点集合,由于不同行为在执行过程中



①视频输入 ②兴趣点检测 ③利用 MLBP - TOP 特征表示 ④视觉单词 ⑤基本行为 ⑥待识别行为 ⑦训练分类器 ⑧分类器

图 2 识别方法流程图

Fig.2 Flowchart of recognition

引起影像中对应位置像素灰度在时空域 XYT 上发生变化,因此对于不同行为,其检测的时空兴趣点集在时空域上的分布也具有较大的差异,可以最大限度地区别视频影像中的不同行为,因此用于行为识别的初始特征提取。Laptev、Dollar 和 Willems 等^[12-14]对于时空兴趣点(Spatio-temporal interest point, STIP)的研究被视为该领域的经典之作,后续研究多数是对现有方法的改进。本文对母牛产前行为的时空兴趣点检测采用 Dollar 提出的方法,该算子通过二维高斯平滑核和一维时域 Gabor 滤波器的参数尺度变换,可以较好地检测出视频中由于行为执行而产生的局部显著区域。该检测子响应函数计算式为

$$\begin{cases} R_{ef} = f(x, y, t)gh_{ev} + f(x, y, t)gh_{od} \\ h_{ev}(t; \tau, \omega) = -\cos(2\pi t\omega) e^{-t^2/\tau^2} \\ h_{od}(t; \tau, \omega) = -\sin(2\pi t\omega) e^{-t^2/\tau^2} \end{cases} \quad (1)$$

其中, R_{ef} 是响应函数,该函数值对应为输入视频图像在空间域 XY 和时间域 t 处局部灰度值突变的位置点,该位置点即为时空兴趣点的空间位置坐标。 $f(x, y, t)$ 是输入视频图像, g 为 2 维空间域方向的高斯平滑核函数,用于对视频帧图像在空间域上进行平滑噪声,方差是 σ^2 ,该参数用于控制空间域,判定局部点是否作为备选兴趣点进行标注。 h_{ev} 和 h_{od} 为 1 维时域 Gabor 滤波器, τ 用以描述时间维 T 上的尺度, $\omega = \tau/4$ 。通过尺度参数 σ 和 τ 将检测得到的兴趣点拓展为兴趣区域,从而更好地获取不同行为的局部信息。

对于已检测到的兴趣点及其区域需要利用特征量化手段对该区域信息进行必要的量化描述。本文提出一种改进的时空局部二值模式对该区域进行量化描述,以便后续利用分类器对该行为的量化特征进行分类。局部二值模式^[15](Local binary pattern, LBP)是一种简单高效的灰度图像局部纹理特征描述算子。该算子利用邻域窗口内中心点像素值作为阈值,利用邻域周边像素值与中心像素值进行比较建立的二进制串转换为十进制数作为该邻域中心像素值的量化。具体计算表达式为

$$\begin{cases} V_{LBP}(I_c) = \sum_{i=0}^{p-1} 2^i S(I_i - I_c) \\ S(A) = \begin{cases} 1 & (A \geq 0) \\ 0 & (A < 0) \end{cases} \end{cases} \quad (2)$$

式中 I_c, I_i ——邻域内中心像素点和相邻像素点的灰度

V_{LBP} ——中心像素点二值模式值

p ——邻域像素点数

为了有效对序列图像进行特征描述,Zhao 等^[16]在经典 LBP 基础上提出一种时空局部二值模式特征描述子(Local binary patterns on three orthogonal planes, LBP-TOP),该特征描述子将视频图像视为 XY 帧在时间维度 T 上的连续变化序列,通过计算 XY, XT, YT 3 个正交面上的 LBP 值,然后串联用以对区域特征进行描述。

1.2.1.2 改进 LBP-TOP

视频在拍摄过程中受光照,以及信号捕捉、记录和传输过程中噪声干扰等影响,从而对 LBP-TOP 纹理特征量化描述和应用产生一定影响。针对该问题,本文在借鉴 Liao 等^[17]改进方法基础上提出一种改进时空局部二值模式特征描述算子(Modified local binary patterns on three orthogonal planes, MLBP-TOP)用于描述兴趣区域。算法基本思想为:首先依据邻域内像素值差异最小化原则,在 LBP 邻域内为每个像素计算确定一个权值,用赋权后的值作为邻域对应位置像素值;然后通过邻域内像素赋权值均值与方差的线性加权统计量作为阈值,从而实现中心阈值的自适应调整。算法实现如下:

(1) 邻域内像素权值计算。计算表达式为

$$\begin{cases} \min D_{Dif} = (g_c(x_c, y_c, t_c) - \sum_{q=0}^{p-1} w_{j,q} g_{j,q}(x, y, t))^2 \\ \text{s. t } \sum_{q=0}^{p-1} w_{j,q} = 1 \end{cases} \quad (j=0, 1, 2) \quad (3)$$

式中 D_{Dif} ——差异值

$g_{j,q}(x, y, t)$ ——窗口邻域像素灰度值

$g_c(x_c, y_c, t_c)$ ——窗口中心点像素灰度值, $j=0, 1, 2$, 顺次表示 XY, XT, YT 各坐标面

p ——邻域像素点数

$w_{j,p}$ ——对应像素点处的权值

对上述约束表达式利用有约束拉格朗日求极值方法确定邻域像素对应位置的权值。

(2) MLBP-TOP 调整中心阈值。具体计算表达式为

$$V_{MLBP-TOP} = \begin{cases} 1 & (g_{j,q}(x, y, t) \geq t_{th}) \\ 0 & (g_{j,q}(x, y, t) < t_{th}) \end{cases} \quad (4)$$

其中 $t_{th} = M_{mean} + ae_{std}$ ($a \in (0, 1)$)

式中 $V_{MLBP-TOP}$ ——中心二值模式值

M_{mean}, e_{std} ——像素值加权均值和方差

$$M_{mean} = \frac{\sum_{q=0}^p w_{j,q} g_{j,q}(x, y, t)}{p} \quad (5)$$

$$e_{\text{std}} = \sqrt{\sum_{q=0}^{p-1} (g_{j,q}(x, y, t) - \mu)^2} \quad (6)$$

其中

$$\mu = \frac{\sum_{q=0}^p g_{j,q}(x, y, t)}{p}$$

1.2.1.3 特征提取与描述

视频行为特征提取与描述具体执行步骤如下:

(1) 时空兴趣点检测。在给定视频利用 Dollar 提出的检测算子提取兴趣区域, 其中尺度参数依据原文献设定 $\sigma = 2, \tau = 3.6$ 。

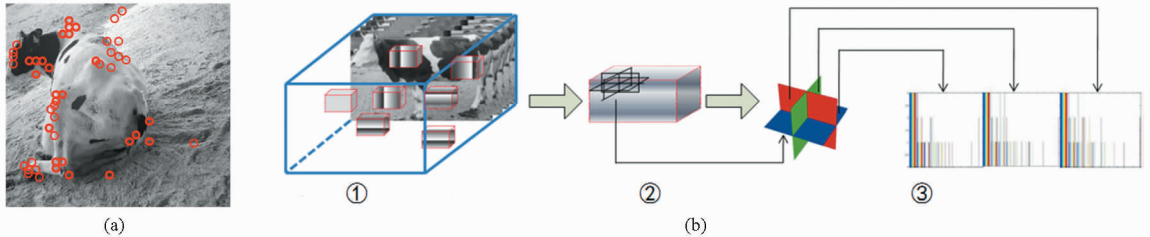


图 3 特征提取与描述示意图

Fig. 3 Diagram of feature abstraction and representation

(a) 兴趣点检测 (b) 特征提取与描述

1.2.2 构建视觉词典

视觉单词和视觉词典概念引申自文本分析方法^[18]。在文本分析中, 单词是文本中的基本单元, 单词的全体构成一个词典, 文本中每个主题可以通过多个单词及其对应出现频次进行量化描述。在本文行为识别中, 视觉单词是全部训练视频上初始提取时空兴趣点量化聚类后形成的聚类中心, 是相似时空兴趣点的抽象量化表示, 用于降低相似度较高兴趣点的共现性, 作为行为判定的基本单元; 视觉词典是所有视觉单词 (聚类中心) 构成的集合; 每个训练视频和测试视频, 将检测到的兴趣点利用最近邻视觉单词进行标识, 通过统计视觉单词出现以及相应频次, 构建视觉直方图作为对应行为的量化特征描述, 并用于后续训练分类器实现行为分类识别。

传统构建视觉词典的方法一般选 k -means 方法。该方法操作简单, 但是初始聚类数 k 的选取和初始聚类中心的选定都会对视觉词典特征表示与应用识别产生负面的影响。因此本文利用 Liu 等^[19]提出的通过最大化互信息构建最小化类内差异最大化类间差异的紧致视觉单词学习算法。该算法的基本实现分为两步: 首先利用 k -means 算法对初始特征量进行聚类, 本文初始 $k = 100$ 形成元视觉单词; 然后利用贪婪算法对近邻元视觉单词进行合并, 合并依据原则满足表达式为

$$\max(I(\hat{X}; Y) - \lambda^{-1}I(\hat{X}; X)) \quad (7)$$

其中, $X = \{x_1, x_2, \dots, x_n\}$ 为元视觉单词集, $\hat{X} = \{\hat{x}_1, \hat{x}_2, \dots, \hat{x}_m\}$ 为聚合后的紧致视觉单词集, 且 $m \ll n$,

(2) 特征描述。在已检测时空兴趣区域上利用 MLBP-TOP 算子对每个兴趣区域进行量化描述, 计算后得到的串联灰度直方图作为该兴趣区域特征的量化结果。其中, 本文算法中的参数设定 $p = 8, XY, YT, XT$ 3 个坐标面上的邻域半径 $R_x = R_y = R_t = 1$, 其中 $R_t = 1$ 相当于在时间维度上, 窗口跨度前后为 1 帧。按照上述参数设定, 检测得到的兴趣区域量化结果为一个 768 维的特征向量。该步骤算法示意图如图 3 所示。图中, ①为时空兴趣点提取, ②为利用 MLBP-TOP 特征表示, ③为构建特征直方图。

$Y = \{y_1, y_2, \dots, y_N\}$ 为视频行为为类别标签。I 函数为互信息表达式, $\lambda = 0.5$ 为 Lagrange 乘子, $I(\hat{X}; Y)$ 对应互信息函数值越大, 表示聚合后的视觉单词集 \hat{X} 类内描述性越强, 类间差异性越大; $I(\hat{X}; X)$ 函数值越小, 表示聚合后视觉单词集 \hat{X} 相比元视觉单词集越紧致。具体方法及实现参见文献^[19]。该步最终得到的 \hat{X} 即为建立在训练视频集上的视觉单词集合, 用于构建视觉词典。

1.2.3 基本行为识别

通过构建视觉单词和视觉词典, 每类行为都可以通过视觉词典中的单词表示为一个视觉单词直方图, 利用训练视频对最近邻分类器进行训练, 训练结果用于对测试视频中母牛基本行为进行分类识别。相似度量标准选用 χ^2 距离, 计算表达式为

$$d_{\chi^2}(H_1, H_2) = \sum_{i=1}^n \frac{(\text{bin}_1(i) - \text{bin}_2(i))^2}{\text{bin}_1(i) + \text{bin}_2(i)} \quad (8)$$

其中 $H_j = \text{bin}_j(i)$ 表示视觉直方图, i 为直方图维数。

1.2.4 视频语义分割

对于长视频的语义分割是对视频中基本行为进行识别和频次统计的关键。视频语义分割涉及两方面的基本内容: 首先是单帧图像的特征量化表示; 其次是对量化表示后的视频帧进行聚类, 继而实现视频的语义分割。具体执行步骤如下:

(1) 帧图像特征表示。LBP 算子可以较好地描述图像的局部细节特征, 因此本文将 LBP 算子用于单帧图像的特征描述。在量化描述过程中, 首先对

单帧图像进行 4×4 区域划分,并在每个图像区域块上利用 LBP 算子进行特征描述,计算该区域块的特征直方图;然后将 4×4 共 16 个区域块的特征直方图进行串联作为该单帧图像的特征描述。其中 LBP 算子参数设定为:邻域像素数为 $p = 8$,邻域半径 $R = 1$ 。单帧图像特征量化示意图如图 4 所示。

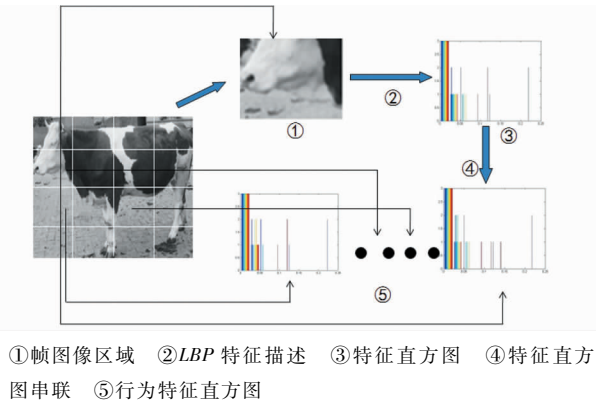


图 4 帧图像特征提取与描述

Fig. 4 Feature abstraction and presentation of frame images

(2) 视频分割。对于上述量化后的视频帧图像,利用 k -means 方法进行聚类,聚类后形成的连续视频段即为分割结果,每段视频中包含独立的基本行为。尽管聚类过程中由于初始聚类中心的选取会影响每类行为对应视频段的起止帧及数量,但是考虑视频段对应行为在视觉过渡过程中本身就存在界限模糊的现象,因此包含基本行为的视频段起止帧不会对后续行为识别产生较为明显的影响。

1.2.5 规律识别与分析

在上述视频分割的基础上,首先利用前述方法分别对每个视频段上包括行走、侧卧和回望基本行为进行识别,然后统计单位时间段内侧卧和回望发生的频次,通过频次变化实现对母牛产前行为规律的量化分析。

2 试验

进行 3 组试验,其中第 1、2 组试验“水平固定视角行为识别”和“水平随机视角行为识别”用于本文方法基本行为识别准确性的验证;第 3 组试验“规律量化与分析”用于产前行为规律的量化分析。

2.1 试验设定

于 2011 年 3 月至 2012 年 9 月期间采集 37 头产前母牛行走、侧卧和回视的 120 组视频,其中指定视角(正前方、侧面和后面)视频 90 组,每组行为在规定视角下对应 15 个视频;水平方向随机视角采集 30 组视频,每种行为对应 10 个视频,并人工对视频行为类别进行标注,该组视频用于基本行为识别。另外一组测试视频通过截取母牛临产前 10 h 和产

后 3 h 时间段获取,该组视频用于产前规律性识别与分析。

2.2 试验结果与分析

2.2.1 水平固定视角行为识别

为了验证算法性能,本组试验在指定视角下对 3 类 6 种行为进行识别。首先对于指定视角的 90 组视频,训练视频数依次设定为 5、8 和 10,剩余视频依次用于测试,试验结果采用 5 次试验取均值的方式。首先,使用时空兴趣点进行兴趣区域检测,利用 MLBP-TOP 算子对兴趣区域进行量化描述,从而得到行为的局部量化特征,每个量化特征根据 MLBP-TOP 参数设定,计算结果为一个 $256 \times 3 = 768$ 维的特征向量;然后利用文献[19]提出的算法计算得到 18 个聚合中心,每个聚合中心对应一个视频单词,每个视觉单词用索引号 1, 2, ..., 18 进行标识,并最终得到由 18 个视觉单词构成的词典;最后用视觉单词直方图描述指定视角下的每类行为,利用最近邻分类器进行分类识别。

图 5 给出初始兴趣点聚合成最终视觉单词的结果示意图,坐标为量化特征对应兴趣点中心位置坐标。

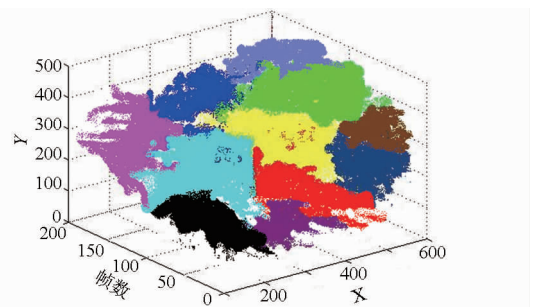


图 5 时空兴趣点聚合示意图

Fig. 5 Diagram of clustered STIPs in spatio-temporal XYT space

图 6 给出利用视觉单词直方图进行行为描述的示意图。图中横坐标表示视觉单词索引号,纵坐标为归一化后的视觉单词出现的频率。对视觉直方图进行归一化的目的在于,考虑到个体在行为执行过程中可能出现同一行为的执行时间长短不一,在视频图像中体现为该行为所出现的图像帧数不同,继而引起描述对应行为的视频段中视觉单词出现的次数发生差异,通过归一化可以有效保证视觉单词直方图对于行为特征描述的一致性。

为测试识别结果与训练视频数的关系,试验中顺次选取训练视频数为 5、8 和 10。识别结果表明,本文提出的方法在小训练样本下(训练样本数为 5),行走和侧卧部分视角下识别准确率已经达到 80% 左右,随着训练样本数的提升(训练样本数为

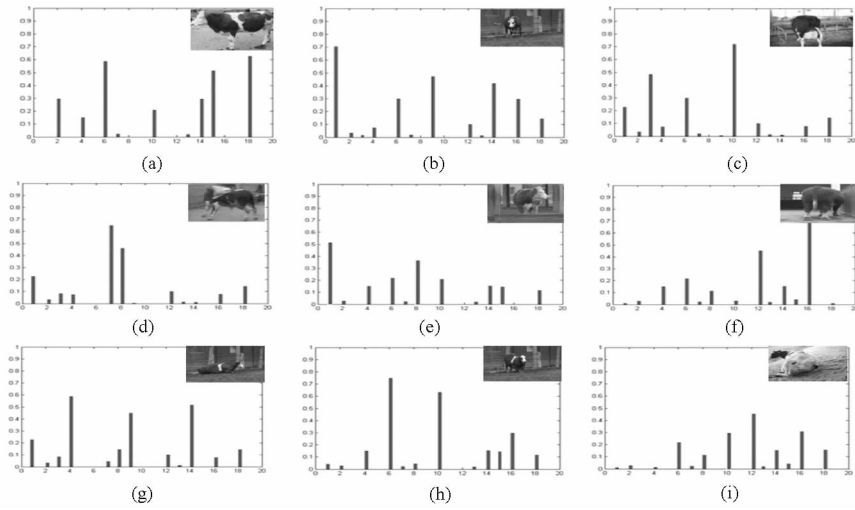


图 6 视觉单词行为描述直方图

Fig. 6 Example diagram of video-words for behavior representation

(a) 行走侧面 (b) 行走正面 (c) 行走后面 (d) 回望侧面 (e) 回望正面 (f) 回望后面 (g) 侧卧侧面 (h) 侧卧正面 (i) 侧卧后面

10), 3 种视角下行为识别准确率平均达到 94.6%, 本文方法适合小样本识别。图 7 为训练样本数为 10 时, 识别结果准确率混淆矩阵。从结果中可以进一步发现, 侧卧行为在 3 种视角下实现 100% 的识别, 分析原因在于侧卧从体态动作上明显与行走和回望存在巨大差异; 识别准确率相对较低的是行走后面为 85%, 误识别发生在回望后面为 15%, 分析其原因在于二者单独从后侧视角进行识别由于体态动作特征相似, 因此引起误识别。其余行为识别准确率都达到 90% 以上。

体位上出现一定的旋转, 因此给准确识别带来比较大的困难。进行 5 次试验取平均值的方式得到正确识别率混淆矩阵见图 8。试验结果中行走和回望的识别率为 85%, 其中行走在 5 次试验中误识别视频全部识别为回望; 回望在 5 次试验中, 误识别视频有 4 次记录误分为行走, 1 次误分为侧卧。分析误识别的原因在于, 行走和回望本身行为特征较为相近, 附加拍摄时体位变换更具随机性。综合分析本文方法整体平均识别准确率仍然达到 88.3%。进一步验证本文方法的有效性。

行走侧面	1.00	.00	.00	.00	.00	.00	.00	.00	.00
侧卧侧面	.00	1.00	.00	.00	.00	.00	.00	.00	.00
回望侧面	.04	.00	.96	.00	.00	.00	.00	.00	.00
行走正面	.00	.00	.00	.90	.00	.10	.00	.00	.00
侧卧正面	.00	.00	.00	.00	1.00	.00	.00	.00	.00
回望正面	.00	.00	.00	.10	.00	.90	.00	.00	.00
行走后面	.00	.00	.00	.00	.00	.00	.85	.00	.15
侧卧后面	.00	.00	.00	.00	.00	.00	.00	1.00	.00
回望后面	.00	.00	.00	.00	.00	.00	.10	.00	.90
行走侧面									
侧卧侧面									
回望侧面									
行走正面									
侧卧正面									
回望正面									
行走后面									
侧卧后面									
回望后面									

图 7 固定视角行为识别结果

Fig. 7 Result of fixed perspective behavior recognition

2.2.2 水平随机视角行为识别

该试验是对第 1 组试验的拓展, 利用第 1 组试验聚合形成的 18 个视觉单词构建的视觉词典对水平随机视角获取每类行为各 10 组视频共计 30 组视频进行识别。由于视角为水平方向随机采集, 视频视角不再是单一的侧面、正面和后面, 而在实际母牛

行走	.85	.00	.15
侧卧	.05	.95	.00
回望	.10	.05	.85
	行走	侧卧	回望

图 8 随机视角行为识别结果

Fig. 8 Result of random perspective behavior recognition

2.2.3 规律量化与分析

母牛临产前期行为一般以频繁侧卧和回望腹部最为典型, 在量化标准上体现为单位时间段内行为发生频次的不断提升。因此本组试验在第 1 组试验和第 2 组试验实现行走、侧卧和回望基本行为为准确识别的基础上, 通过对母牛临产期前后采集视频中包括行走、侧卧和回望等行为在单位时间段内发生频次统计, 实现产前行为规律的量化描述。试验对象为随机选取临产期前 10 h 和产后 3 h 的 5 组视频采集数据作为试验视频进行分析。测试视频采集依据专家经验和实际观察, 上述产前 10 h 和产后 3 h 这个时间段内个体接近临产到产后行为变化较为明显。

试验过程中为实现视频中各种基本行为的识别

和频次统计,首先利用 LBP 算子对视频帧图像进行量化描述,然后通过 *k*-means 方法对给定视频通过聚类进行语义分割,初始聚类数设定 $k = 5$ 。设定原则依据基本行为除了本文已识别明显揭示产前特征的包括行走、侧卧和回望等基本行为外,还有矗立(静止)、饮食等常见基本行为。选取其中一组测试视频的聚类分割效果图如图 9 所示。为了方便说明临产前后基本行为频次变化规律,该组图截取的是在前 10 h 和产后 3 h 时间段内随机选取等长的临产前期、临产末期和临产后期的聚类分割结果图。同时,为了验证本文算法的有效性,对于聚类分割的视频段识别结果采用人工验证方式进行验证,并在图中进行对应段标注,算法自动识别与人工识别结果相一致,用打对勾的方式标注。图中横坐标指示视

频帧索引数(量级为 10^4 帧),纵坐标为聚类索引数。截图试验结果表明,包括行走、侧卧和回望基本行为得到有效识别,验证了本文算法在连续视频中对本基本行为识别的有效性。同时识别结果中也存在误识别和无法识别的问题,分析发现该类行为主要出现在“矗立”上,其原因在于,首先在前述基本行为识别过程中没有对“矗立”和“饮食”进行单独识别学习;另外由于矗立行为多发生在运动行为之后的静止状态,例如侧卧行为之后牛处于静卧状态,或者行走、回望行为之后的静立状态,因此会对相应行为产生误分或无法识别的结果。但是,通过对比本文算法识别结果和人工观测结果,上述误识别或无法识别结果基本没有对后续频次规律统计产生比较大的影响。因此本文算法在测试的长视频上验证是有效的。

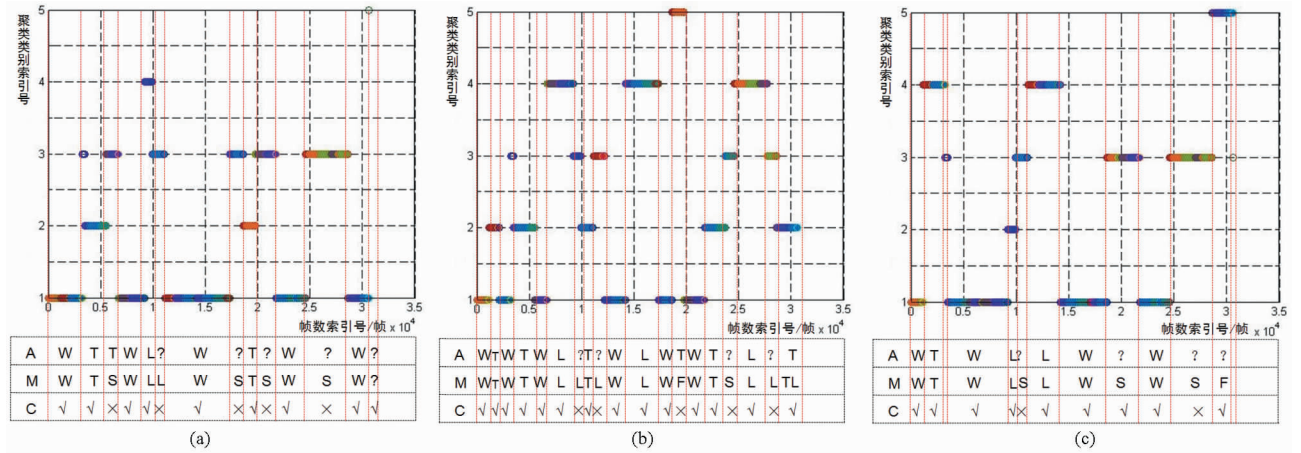


图 9 母牛临产期前后视频分割识别结果

Fig. 9 Result of video segmentation and recognition during sow-behavior

(a) 临产前期 (b) 临产末期 (c) 产后初期

A:自动识别 M:人工识别 C:识别结果一致性 W:行走 T:回望 L:侧卧 S:矗立 F:饮食 √:结果一致 ×:结果不一致 ?:结果未知

表 1 列出在上述视频聚类分段识别的基础上,对连续视频中出现的回望和侧卧发生的频次进行统计的结果。统计时间以 2 h 为基本时间单位,统计结果为 5 次试验取均值。试验结果表明,本文方法对产前行为规律的识别结果与人工识别结果在数据趋势变化上相一致,即临产前 10 h 时间段内,母牛回望频次明显增加,数据呈现规律为前低后高,产后降低,临产前高峰频次算法识别均值为 8.8 次,人工识别为 10.4 次,出现差异的原因是由于不同个体在行为执行过程中的时长差异;侧卧行为发生频次规律为前低后高,但是在接近分娩前算法识别频次降低,平均频次为 2.2 次,结合人工矫正其原因在于分娩前母牛大多时间采用侧卧体位,该行为执行时间跨度较大,所以执行频次较低。

综上,试验表明母牛在临产前的行为规律为频繁发生侧卧和回望行为,临近分娩期回望频次明显

增加,而侧卧频次明显减少。上述识别规律基本与养殖专家和行为观察相一致。规律示意图如图 10 所示。其中横坐标为时间间隔,纵坐标为行为发生频次数,每个时间段内的识别结果为算法识别与人工识别结果。

3 结束语

提出一种基于改进时空局部二值模式对时空

表 1 识别结果统计表

Tab. 1 Statistical table of recognition result

	回望(自动)	回望(人工)	侧卧(自动)	侧卧(人工)
0~2 h	3.4	4.4	2.2	2.2
2~4 h	3.6	4.6	4.2	4.6
4~6 h	4.2	5.6	4.4	4.2
6~8 h	5.4	7.6	5.2	5.6
8~10 h	8.8	10.4	2.2	2.6
10~13 h	4.6	6.2	3.2	3.4

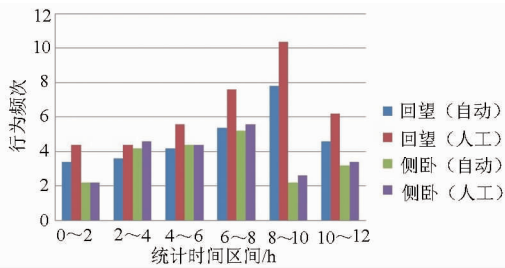


图 10 产前行为规律量化结果

Fig. 10 Quantifiable results of cow-behavior rule

趣点进行特征提取和量化描述,用于构建视觉词典的方法。利用视觉词典构建行为描述直方图,并应用于围产期产前母牛行为识别和规律发现。本文方法对 90 组规定视角下母牛产前行走、侧卧和回望行为的平均正确识别率依次为:行走正面、侧面和后面依次为 100%、90%、85%;侧卧正面、侧面和后面识

别准确率为 100%;回望正面、侧面和后面依次为 96%、90%、90%。上述规定视角下总体行为平均识别准确率为 94.6%。对 30 组水平随机视角下拍摄的 3 种行为平均正确识别率依次为 85%、95% 和 85%,总体平均识别准确率为 88.3%。在母牛临产期长视频行为规律识别中,本文方法识别侧卧和回望发生频次趋势规律与人工识别结果相一致,在分娩前后的 13 h 内,算法识别母牛平均执行回望和侧卧的频次为 30 次和 21.2 次,行为规律体现为前低后高,在分娩前回望频次达到峰值,但是侧卧频次却出现明显减低,因此该变化预示母牛分娩期的临近。试验结果表明,本文方法可以较为有效地为母牛临产期及时进行人工干预提供重要的判定依据。机器视觉方法是对已有佩戴传感器进行行为识别技术的有效补充。

参 考 文 献

- Cornou C, Lundbye Christensen S. Classifying sows' activity types from acceleration patterns: an application of the multi-process Kalman filter[J]. *Applied Animal Behavior Science*, 2008, 111(3):262 ~ 273.
- Cornou C, Lundbye-Christensen S. Classification of sows' activity types from acceleration patterns using univariate and multivariate models [J]. *Computers and Electronics in Agriculture*, 2010, 72(2):53 ~ 60.
- Hugo Jair Escalante, Sara V R, Jorge Cordero, et al. Sow-activity classification from acceleration patterns: a machine learning approach [J]. *Computers and Electronics in Agriculture*, 2013, 93(4), 17 ~ 23.
- 刘龙申,沈明霞,姚文,等. 基于加速传感器的母猪产前行为特征采集与分析[J]. *农业机械学报*,2013,44(3):192 ~ 196,191. Liu Longshen, Shen Mingxia, Yao Wen, et al. Acquisition and analysis of sows' behavior before farrowing based on acceleration sensor[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2013, 44(3):192 ~ 196,191. (in Chinese)
- 尹令,刘财兴,洪添胜,等. 基于无线传感网络的母牛行为特征监测系统研究[J]. *农业工程学报*,2010,26(3):203 ~ 209. Yin Ling, Liu Caixing, Hong Tiansheng, et al. Design of system for monitoring dairy cattle's behavioral features based on wireless sensor networks[J]. *Transactions of the CSAE*, 2010, 26(3): 203 ~ 209. (in Chinese)
- Boaz Zion. The use of computer vision technologies in aquaculture—a review[J]. *Computers and Electronics in Agriculture*, 2012, 88(10):125 ~ 132.
- 温长吉,王生生,于合龙,等. 基于改进蜂群算法优化神经网络的玉米病害图像分割[J]. *农业工程学报*,2013,29(13):142 ~ 149. Wen Changji, Wang Shengsheng, Yu Helong, et al. Image segmentation method for maize diseases based on pulse coupled neural network with modified artificial bee algorithm [J]. *Transactions of the CSAE*,2013,29(13):142 ~ 149. (in Chinese)
- 彭辉,吴鹏飞,翟瑞芳,等. 基于视觉图像的重叠果实图像分割方法[J]. *农业机械学报*,2012,43(6):165 ~ 173. Peng Hui, Wu Pengfei, Zhai Ruifang, et al. Image segmentation algorithm for overlapping fruits based on disparity map[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2012,43(6):165 ~ 173. (in Chinese)
- Xue Xinwei, Thomas C H. Feature fusion for basic behavior unit segmentation from video sequences[J]. *Robotics and Autonomous Systems*,2009,57(3):239 ~ 248.
- Dawkins M S, Lee Hyoung-joo, Waitt C D, et al. Optical flow patterns in broiler chicken flocks as automated measures of behaviour and gait[J]. *Applied Animal Behavior Science*, 2009, 119(3 ~ 4): 203 ~ 209.
- 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)
- Ivan Laptev. On space-time interest points [J]. *International Journal of Computer Vision*,2005,64(2 ~ 3):103 ~ 127.
- Dollar P, Rabaud V, Cottrell G, et al. Behavior recognition via sparse spatio-temporal features[C]//2nd Joint IEEE International Workshop on, VS-PETS., 2005:65 ~ 72.
- Geert Willems, Tinne Tuytelaas, Luc VanGool. An efficient dense and scale-invariant spatio-temporal interest point detector[C]. *Computer Vision-ECCV 2008, Lecture Notes in Computer Science (LNCS)*, 2008,5303:650 ~ 663.
- Ojala T, Pietikainen M, Harwood D. A comparative study of texture measures with classification based on feature distributions [J]. *Pattern Recognition*, 1996, 29(1):51 ~ 59.
- Zhao Guoying, Pietikainen M. Dynamic texture recognition using local binary patterns with an application to facial expressions

[J]. IEEE transactions on Pattern Analysis and Machine Learning, 2007, 29(6):915 ~ 928.

- 17 Liao W H. Region description using extended local ternary patterns [C] // 2010 20th International Conference on Pattern Recognition, 2010:1 003 ~ 1 006.
- 18 Glaser T, Zelnik-Manor L. Incorporating temporal context in bag-of-words models[C] // Computer Vision Workshops, 2011 IEEE International Conference on. , 2011:1 562 ~ 1 569.
- 19 Liu Jingen, Shah M. Learning human actions via information maximization[C] //2008 IEEE Conference on Computer Vision and Pattern Recognition, 2008:1 ~ 8.

Visual Dictionary for Cows Sow Behavior Recognition

Wen Changji^{1,2} Wang Shengsheng² Zhao Xin¹ Wang Min³ Ma Li¹ Liu Yiting²

(1. College of Information and Technology, Jilin Agricultural University, Changchun 130118, China

2. Key Laboratory of Symbolic Computing and Knowledge Engineering of Ministry of Education,
Jilin University, Changchun 130012, China

3. Department of Computer and Information Sciences, Towson University, Maryland 21252, U. S. A.)

Abstract: Information acquisition of the cow behavior modes and regulars are very important. It is one of judgments for the manual intervention in a particular period. There is certain limitation in the way of obtaining information by using external sensors. So, firstly a modified spatial-temporal local binary pattern for feature presentation was proposed. Secondly through building a visual dictionary, the cow sow-behavior were recognized on the test video. Finally, the basic rule of cow sow-behavior was studied through the statistics occurrence frequency of specific behaviors. For verifying the efficiency of the proposed method, different experimental settings were tested for recognizing the typical sow-activities such as walking, side-lying and look-backing. One was tested on 90 groups of videos under given visual angles and the other on 30 groups of videos under the random perspective ; the experiment results show that the average accuracy is 94.6% and 88.3% , respectively. The other side, a thirteen-hours video before and after delivery was using to sow-activity recognition and frequency count. The experiment results show that the average occurrence frequencies of look-backing and side-lying are 30 and 21.2, respectively. The occurrence frequencies of above two kinds of behaviors are former-low to after-high. Before delivery, the frequency of look-backing reaches to maximum 8.8 and side-lying reduces to 2.2. The results show that this method proposed reveals the basic principle of cow sow-behavior.

Key words: Cows Sow behavior Visual dictionary Modified local binary pattern Spatio-temporal interest point