

基于计算机视觉的四足动物行走步态参数自动提取方法

陈瑶¹, 张云伟^{1,2,3*}, 雷金辉^{1,3}, 黎丽¹

¹昆明理工大学信息工程与自动化学院, 云南 昆明 650500;

²昆明理工大学云南省人工智能重点实验室, 云南 昆明 650500;

³昆明理工大学云南省计算机技术应用重点实验室, 云南 昆明 650500

摘要 四足动物行走的运动特征自动识别在动物的仿生学、行为识别、疾病预测,以及个体身份识别等方面都有着非常广阔的应用前景。基于计算机视觉技术,结合深度学习方法,针对四足动物的行走运动,建立了一种四足动物行走步态参数自动提取方法。利用视频帧分解技术对采集的四足动物行走视频进行处理,得到各帧四足动物行走图像;基于改进 DeeplabV3+ 语义分割模型提取运动目标;在此基础上,结合四足动物行走步态及其时空特性分析,以目标体轮廓中心点-边界距离曲线为基础实现运动角点的检测及匹配。为有效地解决四足动物运动特征参数提取的问题,建立了一种基于四肢运动角点至参考点距离变化曲线的运动特征分析方法。实验结果表明,所提方法能较好地实现四足动物运动角点的检测,对犀牛、水牛、羊驼运动角点检测的最大误差分别为 32, 27, 19 pixel, 运动角点匹配关系准确,步态周期、步态频率的计算误差小于 2%,步态顺序输出正确,步幅的计算误差最大为 2.85%。

关键词 机器视觉; 深度学习; 语义分割; 运动角点检测; 步态参数提取

中图分类号 TP391.4

文献标志码 A

doi: 10.3788/LOP202259.0815006

Automatic Extraction Method for Gait Parameters of Quadruped Walking Based on Computer Vision

Chen Yao¹, Zhang Yunwei^{1,2,3*}, Lei Jinhui^{1,3}, Li li¹

¹Faculty of Information Engineering and Automation, Kunming University of Science and Technology, Kunming, Yunnan 650500, China;

²Yunnan Key Laboratory of Artificial Intelligence, Kunming University of Science and Technology, Kunming, Yunnan 650500, China;

³Yunnan Key Laboratory of Computer Technology Application, Kunming University of Science and Technology, Kunming, Yunnan 650500, China

Abstract The automatic recognition of motion features of quadruped walking could be widely used in animal bionics, behavior recognition, disease prediction, and individual identification recognition. In this paper, based on the computer vision technology and deep learning method, an automatic extraction method for the gait parameters of quadruped walking is established. At first, the walking image frames of quadrupeds can be obtained from the collected quadruped walking videos by using the video frame decomposition technology. Next, the moving object

收稿日期: 2021-03-08; 修回日期: 2021-04-23; 录用日期: 2021-04-28

基金项目: 国家自然科学基金(51365019)

通信作者: *zhangyunwei72@gmail.com

can be extracted via the improved semantic segmentation model DeeplabV3+. Then, according to the characteristic analysis of quadruped walking gait, the detection and matching of the motion corners can be realized based on the distance from the center point to the contour of the object. Finally, a method based on the distance from four limb motion corners to a fixed reference point is established to effectively solve the problem of quadruped motion feature parameter extraction. The experimental results show that the proposed method can give better results for the motion corner detection of quadrupeds. The maximum error is 32, 27, and 19 pixel in the motion corner detection for rhino, buffalo, and alpaca, respectively. The corners of four limbs match accurately. The results also show that the calculation error of the gait cycle and gait frequency is less than 2%, the gait sequence output is correct, and the maximum calculation error of the stride length is 2.85%.

Key words machine vision; deep learning; semantic segmentation; motion corner detection; gait parameter extraction

1 引言

在自然界中,运动是动物区别于其他生物的最根本特征,在动物的生存、繁衍和进化过程中发挥了重要作用。对动物的运动行为进行特征识别,能帮助人们进一步地发现和揭示动物运动的内在规律与本质特征,原始的做法通常是先对某一特定实验场景进行视频记录,然后由人通过肉眼观察,记录相关实验的结果。该方法存在人工成本高、工作强度大、缺乏客观性和实时性差等问题。目前,针对动物运动信息的数据获取主要采用两类方式:接触式佩戴传感器的方式^[1-2]和基于计算机视觉感知的方式^[3-4]。其中,接触式佩戴传感器的方式需要将各类传感器佩戴在动物身上进行数据采集,获取数据方便快捷,对数据的处理也相对容易。然而,该方法易使动物产生应激反应,干扰动物的正常生理状态,特别是对于野外动物较难实施,并且佩戴在动物身上的传感器受晃动、碰撞等因素的影响后会产生噪声数据,进而影响动物行为检测的准确性。相比之下,基于计算机视觉感知的方式能够较好地避免上述问题。信息感知是非接触式的,其实施也更加容易,进而受到了较为广泛的关注。

在将动物的运动特征应用于仿生机器人的研究中,张秀丽等^[5]受到婴儿爬行时以膝着地、以柔顺脊柱配合腿部的运动方式启发,设计了具有柔性脊柱和弹性膝关节的欠自由度四足爬行机器人 BabyBot。Daou 等^[6]通过解剖、视频影像以及压电传感器等方式获取了乌龟的几何特性和运动特性参数,使用 MD ADAMS 平台设计和模拟仿生乌龟的运动。在对动物步态分析方面, Kim 等^[7]针对目前开发的步行机器人没有足够能力适应不平坦地形的问題,采用红外摄像机捕获蜥蜴的运动数据,

并通过对蜥蜴的运动数据的分析,提出了一种基于空间的关节步态,并将其应用于蜥蜴的运动学模型中。刘波等^[8]提出了一种基于深度图像和生猪骨架端点分析的生猪步频特征提取方法,通过计算图像序列的完整步态与序列采集时间长度比值提取了生猪步频特征。实验结果表明,该方法对生猪步频分析的准确率达 82.1%。在对动物行为识别方面, Cangar 等^[9]提出了一种基于图像处理技术的母牛产犊前行为分析方法。通过对 8 头母牛产犊前 24 h 的视频监控图像进行分析,对母牛的站立、躺卧、进食和饮水行为进行识别分类。薛芳芳等^[10]利用目标检测算法定位牛的特征部位在图像中的位置,构建牛特征部位的空间关系向量,然后利用全连接神经网络对空间关系向量进行分类,实现了对牛的站立、卧躺和采食行为的识别。

综上所述,虽然目前已有较多学者利用视频影像信息开展对动物的运动特征识别的研究,并且该方法在仿生机器人、动物智能监控等方面得到了应用。然而,如何通过计算机视觉感知的方式对动物的运动特征进行自动分析识别进而获得准确的运动特征参数,这一问题仍有待解决。本文基于计算机视觉技术,结合深度学习方法,针对四足动物的行走运动,建立了一种四足动物行走步态参数自动提取方法,能准确识别和定位动物四肢运动角点,自动提取步态周期、步态频率、步态顺序、步幅、角点摆动角和运动速度等步态参数,并以水牛、犀牛和羊驼为实验对象,对该方法的有效性进行了验证。

2 实验数据采集

实验数据采集于云南省昆明市动物园,采集对象为四足行走动物。为保证拍摄视频的稳定性,采用

三脚架固定拍摄设备进行拍摄。拍摄设备设置在四足动物旁,视频采集系统如图 1 所示。其中,坐标系为摄像机坐标系 $O-XYZ$,摄像机距地面约 1.5 m。定义四足动物的运动方向与 $X-Y$ 平面的夹角为运动方向角 β ,运动方向远离 $X-Y$ 平面时为正,运动方向接近 $X-Y$ 平面时为负。摄像机镜头光圈为 $f/1.8$,等效焦距为 26 mm,镜头视角约为 72° ,对焦模式为自动对焦。在拍摄时,尽可能地完整录制四足动物的行走过程,以保证视频中四足动物四肢交替变换过程清晰可见。在不同天气情况下的早上、中午、傍晚进行数据采集,动物运动方向角 β 的范围为 $-33^\circ \sim 33^\circ$,进而获得光照角度、强度和不同视角的数据。数据采集对象包括水牛、犀牛和羊驼。在删除目标动物停顿或目标体不完整的视频后,共得到 20 段视频,每段视频时长在 4~15 s 之间,视频格式为 mp4,分辨率为 1280 pixel \times 720 pixel,帧速率为 30 frame/s。接着,选取其中 15 段视频用于制作数

据集,利用视频帧分解技术,每间隔 2 帧取 1 帧,进而得到 1573 幅四足动物行走侧视图图像,随机选择其中 700 幅图像作为原始图像。考虑到自然环境中,光照强度随时间变化的特点,分别以原始图像的亮度增强 10% 和降低 10% 的图像扩充数据集,最终得到 2100 幅图像。随机选取其中 1800 幅作为训练集,300 幅作为验证集。采用开源工具 Labelme 对训练数据集进行人工标注,用于模型的训练。图 2 为四足动物行走图像。



图 2 四足动物行走图像
Fig. 2 Images of quadruped walking

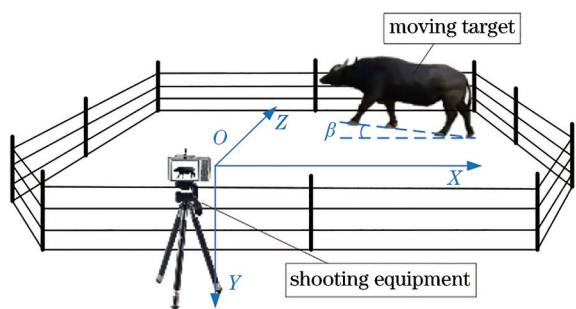


图 1 视频采集系统
Fig. 1 Video acquisition system

3 算法模型

3.1 行走运动特征的自动识别流程

四足动物行走步态参数的自动提取流程如图 3 所示,主要包括 4 个关键步骤:1) 利用改进的深度学习 DeeplabV3+ 语义分割模型提取运动目标;2) 以目标轮廓中心点与边缘点间的距离曲线定位四足动物运动角点坐标;3) 四足动物运动角点坐标的重构与匹配;4) 通过四肢运动角点与参考点的距离变化曲线提取行走步态参数

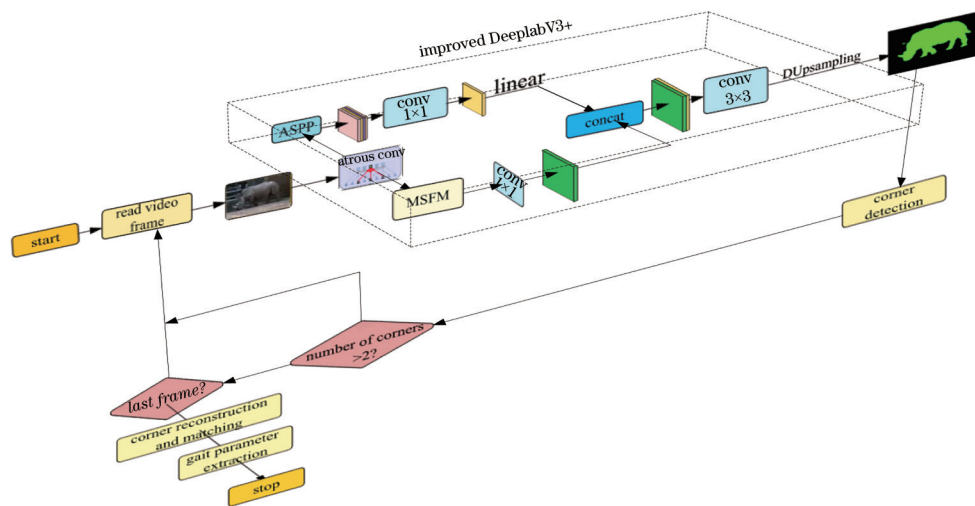


图 3 运动特征自动识别流程图
Fig. 3 Flow chart of automatic identification of motion feature

3.2 运动目标提取

图像语义分割是计算机对图像像素级的分类,即对输入图像的每个像素指定一个与目标相对应的标签并输出一张与原图相对应的分类结果图。为解决实际场景中复杂背景下的四足动物运动目标提取问题,选用基于全卷积的图像语义分割模型进行目标的检测。利用基于全卷积的语义分割模型对输入图像的大小没有要求的特点,实现了输入和输出都是图像的端对端语义分割^[11]。

自 2015 年全卷积神经网络被提出以来,众多研究者在其基础上提出了如 U-net^[12]、Segnet^[13]、PSPNet^[14]和 Deeplab 系列^[15-18]算法等多种高性能的语义分割模型。为了选取效果更好的算法模型,采用迁移学习的思想,在预训练权值的基础上,使用自建的四足动物数据集优化网络参数,对 U-net、Segnet、PSPNet 和 DeeplabV3+ 语义分割模型进行了分析对比。为保证算法的实时性和对比一致性,对 4 种模型均使用轻量化深度神经网络 MobileNet^[19]进行图像特征提取。实验所用计算机的配置为:Windows 操作系统,i5 9400F 处理器(2.9 GHz),深度学习框架为 TensorFlow1.11, GPU 为 GTX1660s,显存为 6 GB。PPM(Pyramid Pooling Module)中 pool_size 和 strides 分别设置为

18, 9, 6, 3; ASPP(Atrous Spatial Pyramid Pooling)中 rate 分别设置为 6, 12, 18; 初始学习率为 0.001, batch_size 设置为 2, epochs 设置为 50, 优化器为 Adam。从结果来看,DeeplabV3+ 模型获得了更好的分割效果,但其对目标边缘信息的感知能力较弱,不利于后续四足动物运动角点的提取。

DeeplabV3+ 模型结构如图 4 所示,该模型采用编码-解码结构。在编码阶段,利用深度卷积神经网络(DCNN)和 ASPP 模块提取图像特征,其中 ASPP 模块采用不同空洞率的空洞卷积和全局平均池化并行提取输入特征图的多尺度上下文信息。在解码器阶段,将低层特征和经过 ASPP 后的特征进行融合;最后经过线性插值将低分辨率的图像恢复到原图大小,从而实现对输入图像像素级的分类。DeeplabV3+ 在高层的语义特征提取的过程中多次进行下采样操作,损失了大量的空间以及细节信息,而简单线性插值法对较低分辨率的特征图的恢复能力有限,是导致分割结果不理想的主要原因。为解决上述问题、进一步提高分割性能,本文使用 Xception^[20]作为特征提取网络,针对四足动物运动目标提取对 DeeplabV3+ 结构进行两个方面的改进:1) 设计了一个多尺度特征信息的融合模块;2) 采用 DUpsampling^[21]改进 DeeplabV3+ 的线性插值上采样。

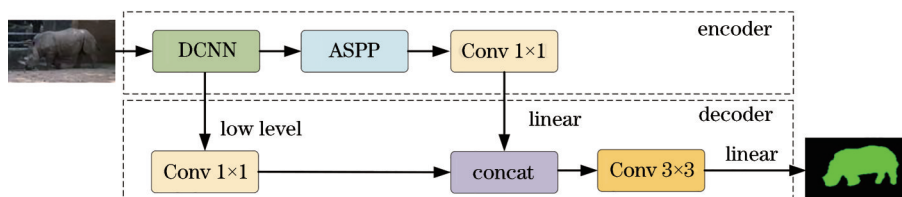


图4 DeeplabV3+结构

Fig. 4 Structure of DeeplabV3+

为融合多尺度的特征信息,本文借鉴 HRNet^[22]思想,在 DeeplabV3+ 上构建多尺度信息融合模块。在 HRNet 的残差单元 Bottleneck 中,输入特征图经学习单元和直连单元的相加融合后,被输入到下一模块中,其中学习单元分别经 1×1、3×3、1×1 三次卷积后提取图像特征,直连单元根据学习单元输入、输出特征的通道数的情况进行 1×1 卷积的添加。Bottleneck 有效减少了训练过程中梯度爆炸的情况,有利于网络层次的加深,用于提取高层次的图像特征。为减少卷积过程中的参数量,本文将 Bottleneck 中学习单元中的 3×3 卷积替换为深度可分离卷积,构建 Sep Bottleneck 模块。DeeplabV3+ 在解码过程中对输入图像进行了 4 次下采样,但解

码过程只用到了第 2 次下采样的特征信息。为更好地利用不同尺度的特征信息进行图像分割,本文将特征提取网络中前三次下采样后的信息输入到所构建的三输入单输出多尺度特征融合模块(MSFM)中,如图 5 所示。该模块共由 4 个基本单元构成,每个单元为三输入、三输出的结构,不同分辨率的特征图分别经 Sep Bottleneck 进行多尺度特征交叉融合增强,最后将输出结果变换到同样尺寸进行通道拼接。

在 DeeplabV3+ 中采用双线性插值的方法将低分辨率特征图映射为较高分辨率的特征图。但是这种方法对各像素之间的相关性考虑不足,导致恢复后的结果欠佳,且往往对于越高级的特征,引入

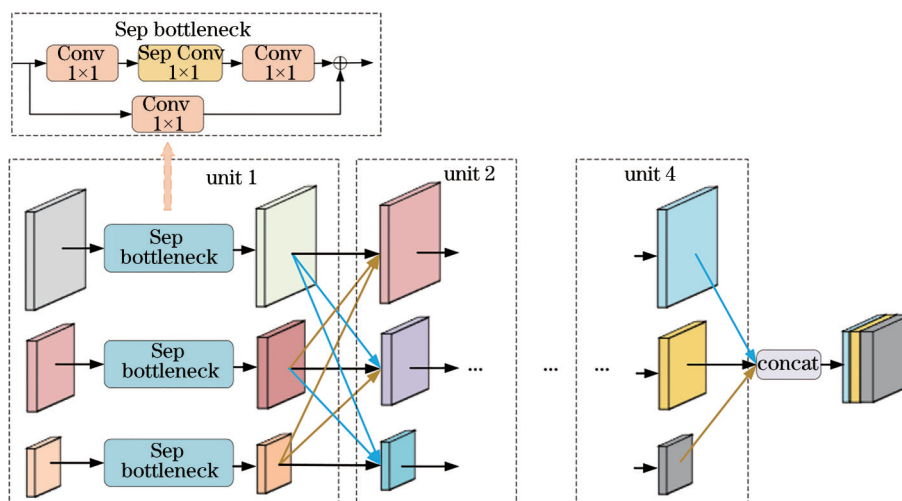


图 5 多尺度特征融合模块

Fig. 5 Multiscale feature fusion module

的噪声信息就会越多。而DUpsampling可以较好地解决上述问题,在运算量减少的同时,降低了上采样操作对特征图分辨率的要求,显著地提高图像分割的精度。故本文采用DUpsampling替换DeeplabV3+上采样方式。DUpsampling原理如

图6所示,其中 F 是由神经网络提取到的特征图, h 、 w 、 c 表示特征图的长、宽及通道数,将特征图中每个 $1 \times c$ 维度的像素乘上一个 $c \times n$ 的权值矩阵 W ,得到 $1 \times n$ 的特征表示,最后通过重组完成上采样操作,得到 $2h \times 2w \times (n/4)$ 的高分辨率特征图 R 。

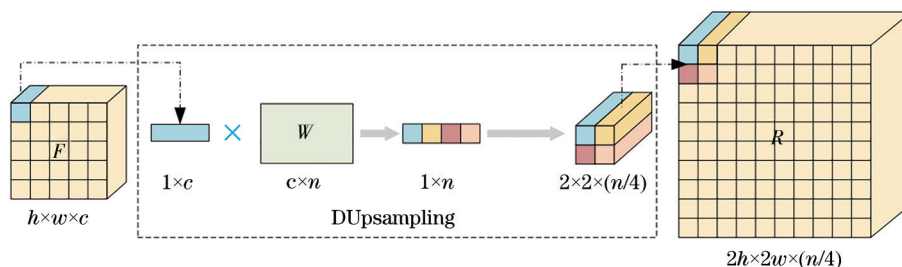


图 6 DUpsampling

Fig. 6 DUpsampling

改进后的DeeplabV3+语义分割模型结构如图7所示。图8(a)为原图。U-net、Segnet、PSPNet和Deeplab这4种模型的语义分割结果如图8(b)~(e)所示。在同样的实验环境下,对同一数据进行测试,分割结果如图8(f)所示。可以看出,改进后的DeeplabV3+模型对目标的分割能力有了进一步的提升,模型对目标的细节信息有了更强的感知能力,更好地恢复了目标的边界信息。进一步地,将分割图像转化为二值图像,提取图像中的连通区域,去除区域面积小于预设值的连通区域,获得完整的二值化运动目标图像,并提取目标的边缘轮廓曲线。

3.3 运动角点检测及匹配

在四足动物行走时,考虑到其四肢末端具有定位性能准确、重复检测率高、噪声鲁棒性好、计算方

便等特点,本文选择四足动物的四肢末端作为运动角点,将运动角点定义为目标轮廓中心点与四肢边缘点距离曲线上的局部极大值,共定义4个运动角点。靠近四足动物头部的两肢称为前肢,靠近四足动物尾部的两肢称为后肢。运动角点检测及匹配流程如图9所示。

1) 运动角点提取

根据前面运动角点的定义,可沿着轮廓线搜索得到各个运动角点。实际上,四足动物行走时,运动角点通常位于轮廓中心点下方的轮廓线上。为了提高搜索效率、减小搜索误差,本文结合四足动物角点摆动角的概念对运动角点搜索范围进行优化。

首先给出四足动物角点摆动角的定义,如图10所示,轮廓中心点用 O 表示,第 i 帧图像的目标轮廓

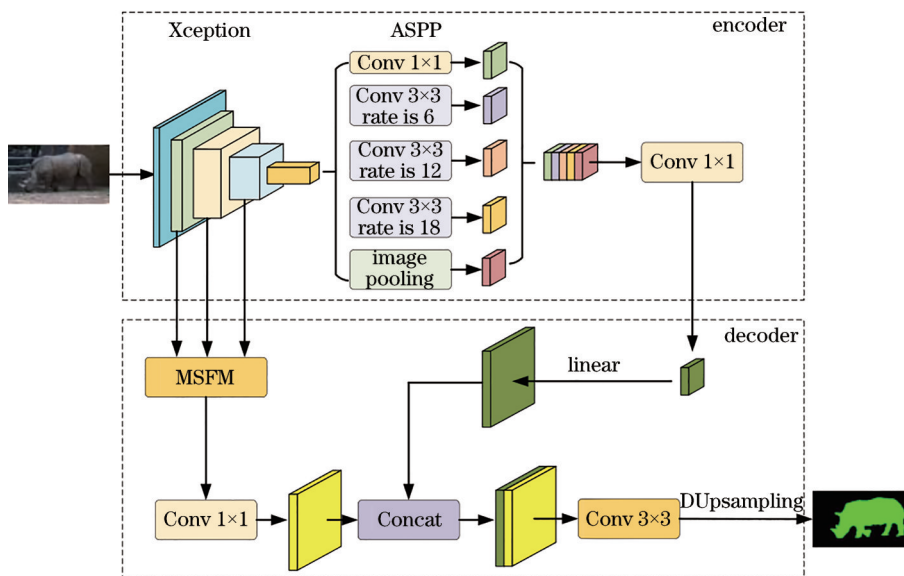


图7 改进 DeeplabV3+

Fig. 7 Improved DeeplabV3+

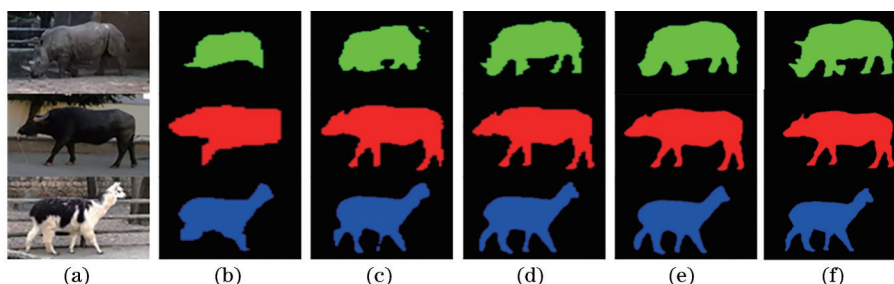


图 8 语义分割效果对比。(a)原图;(b)PSPNet;(c)Segnet;(d)Unet;(e)DeeplabV3+;(f)改进的 DeeplabV3+Xception
Fig. 8 Comparison of effects of semantic segmentation. (a) Original image; (b) PSPNet; (c) Segnet; (d) Unet; (e) DeeplabV3+;

(f) improved DeeplabV3+Xception

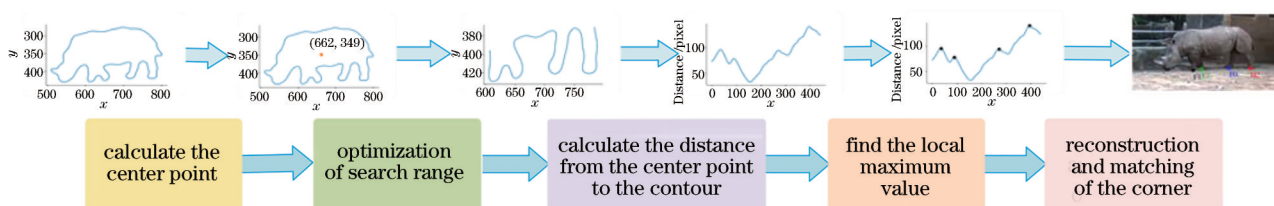


图 9 运动角点检测及匹配流程图

Fig. 9 Flow chart of motion corner detection and matching

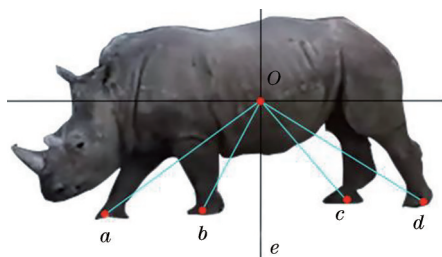


图 10 角点摆动角示意图

Fig. 10 Schematic diagram of corner swing angle

中心点坐标设为 (x_c^i, y_c^i) , 过轮廓中心点 O 作垂直中线, 连接运动角点 a, b, c, d 与轮廓中心点 O , 定义运动角点至轮廓中心点 O 的连线与垂直中线之间的夹角为角点摆动角, 逆时针方向为正, 顺时针方向为负。设第 i 帧图像 4 个运动角点的坐标为 $(x_{pj}^i, y_{pj}^i) (j=0, 1, 2, 3)$, 对应角点的摆动角为 $\alpha_j^i (j=0, 1, 2, 3)$, 则有

$$\alpha_j^i = \arctan \frac{x_{pj}^i - x_c^i}{y_{pj}^i - y_c^i} \quad (j=0, 1, 2, 3), \quad -\frac{\pi}{2} < \alpha_j^i < \frac{\pi}{2} \quad (1)$$

选取一个阈值 A_a , 使其满足以下条件:

$$A_a \leq \frac{\pi}{2} - \max \left\{ \max_{k=1}^i |\alpha_0^k|, \max_{k=1}^i |\alpha_1^k|, \max_{k=1}^i |\alpha_2^k|, \max_{k=1}^i |\alpha_3^k| \right\}$$

则第 $i+1$ 帧时, 优化的角点搜索范围为边缘轮廓线上轮廓点至中心点 O 的连线与垂直中线的夹角在 $\left[-\frac{\pi}{2} + A_a, \frac{\pi}{2} - A_a\right]$ 之间的轮廓点集, 初始时 A_a 可根据检测对象由经验预设。

设当前图像为第 i 帧图像, 则目标轮廓中心点至边缘轮廓点 j 的距离 d_j^i 的计算公式为

$$d_j^i = \sqrt{(x_j^i - x_c^i)^2 + (y_j^i - y_c^i)^2}, \quad j=1, 2, \dots, n_i, \quad (2)$$

式中: (x_j^i, y_j^i) 为第 i 帧指定范围内边缘轮廓点 j 的坐标; n_i 为第 i 帧指定范围内边缘轮廓点数量。

按照(2)式沿逆时针方向依次计算出优化搜索范围内各个轮廓点到轮廓中心点的欧氏距离, 即可得到目标轮廓中心点至边缘轮廓点距离的曲线, 如图 11 所示, 该距离用数字序列 $\{d_j^i | j=1, 2, \dots, n_i\}$ 表示。

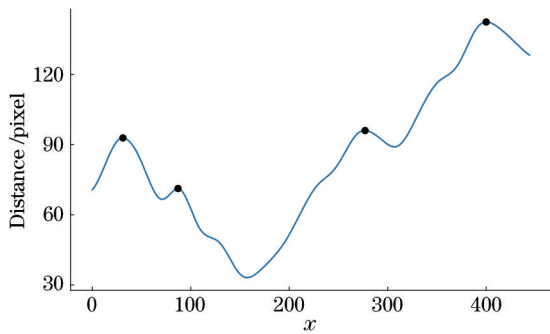


图 11 局部极大值

Fig. 11 Local maximum values

利用差分法来检测数字序列 $\{d_j^i | j=1, 2, \dots, n_i\}$ 中的局部极大值点, 从而获得运动角点的坐标值。差分法局部极大值点检测算法步骤如下:

(1) 对数字序列 $\{d_j^i | j=1, 2, \dots, n_i\}$ 中相邻数据作差, 得到数字序列 $\{\Delta d_j^i | j=1, 2, \dots, n_i-1\}$; 对差分数字序列重新赋值, 即: 如果 Δd_j^i 小于 0, 则将 Δd_j^i 赋值为 -1; 如果 Δd_j^i 大于等于 0, 则将 Δd_j^i 赋值为 1。

(2) 对数字序列 $\{\Delta d_j^i | j=1, 2, \dots, n_i-1\}$ 中相邻数据作差, 得到数字序列 $\{\Delta^2 d_j^i | j=1, 2, \dots, n_i-2\}$ 。

(3) 如果 $\Delta^2 d_j^i$ 小于 0, 则 d_{j+1}^i 为局部极大值, 进一步求得运动角点的坐标值为 (x_{j+1}^i, y_{j+1}^i) 。

同理, 按照上述方法, 即可搜索出所有运动角点的坐标值。对比分析图 10、图 11 可知, 图 11 中的 4 个局部极大值点对应着四足动物前肢和后肢的 4 个运动角点。

2) 运动角点匹配

通过前面步骤已经可以获得各图像帧中的运动角点序列, 但各图像帧中四肢分别对应哪个运动角点还不明确, 需要通过运动角点匹配确定各运动角点与四肢间的隶属关系。由于四足动物在行走过程中, 前肢或者后肢多次交替前行, 无法直接确定四肢上角点的隶属关系。图 12(a) 为犀牛角点无重合的情况。此外, 四足动物在行走过程中存在前肢左右运动角点重合或者后肢左右运动角点重合的情况, 如图 12(b)~(d) 所示, 此时的运动角点坐标只有 3 组, 存在坐标数据缺失问题。因此, 运动角点匹配面临 3 个关键问题: (1) 如何解决运动角点局部重合时的坐标数据缺失问题? (2) 如何将各图像帧中提取的运动角点正确分配到四肢? (3) 如何区分四肢运动角点的同侧位置关系? 为了便于描述, 下面以视频图像中四足动物自右向左的行走方向为例进行说明。



图 12 四足动物角点检测。(a)犀牛角点无重合;(b)犀牛前肢角点重合;(c)水牛后肢角点重合;(d)羊驼前肢角点重合
Fig. 12 Corner detection of quadrupeds. (a) Corners of rhino without overlap; (b) forelimb corners of rhino with overlap; (c) hindlimb corners of buffalo with overlap; (d) forelimb corners of alpaca with overlap

为解决运动角点坐标的数据缺失问题,本文提出了一种运动角点坐标的数据分类重构方法。将当前第 i 帧图像下 4 组运动角点坐标按照图像坐标系下 x 值的升序排列的集合用 $\{C_{0(i)}, C_{1(i)}, C_{2(i)}, C_{3(i)}\}$ 表示。没有运动角点重合时,如图 12(a) 所示,可以提取到 4 组有效的运动角点坐标,将其按照图像坐标系下 x 值升序排列后就得到 $\{C_{0(i)}, C_{1(i)}, C_{2(i)}, C_{3(i)}\}$ 。有运动角点重合时,只能提取 3 组有效的运动角点坐标,将这 3 组运动角点坐标按照其图像坐标系下 x 值进行升序排序,设排序

后的坐标集合为 $\{C'_{0(i)}, C'_{1(i)}, C'_{2(i)}\}$ 。存在两种情况:1) 如果是前肢左右运动角点重合,如图 12(b) 所示,则缺失的角点坐标应包含在 $C'_{0(i)}$ 中,有 $\{C_{0(i)}, C_{1(i)}, C_{2(i)}, C_{3(i)}\} = \{C'_{0(i)}, C'_{0(i)}, C'_{1(i)}, C'_{2(i)}\}$; 2) 如果是后肢左右运动角点重合,如图 12(c) 所示,则缺失的角点坐标应包含在 $C'_{2(i)}$ 中,有: $\{C_{0(i)}, C_{1(i)}, C_{2(i)}, C_{3(i)}\} = \{C'_{0(i)}, C'_{1(i)}, C'_{2(i)}, C'_{2(i)}\}$ 。可以根据当前帧与第 $i-1$ 帧图像中对应运动角点的距离之和最小原则来确定是情况 1) 还是情况 2), 具体计算方法为

$$\begin{cases} \rho_1 = d(C_{0(i-1)}, C'_{0(i)}) + d(C_{1(i-1)}, C'_{0(i)}) + d(C_{2(i-1)}, C'_{1(i)}) + d(C_{3(i-1)}, C'_{2(i)}) \\ \rho_2 = d(C_{0(i-1)}, C'_{0(i)}) + d(C_{1(i-1)}, C'_{1(i)}) + d(C_{2(i-1)}, C'_{2(i)}) + d(C_{3(i-1)}, C'_{2(i)}) \\ C_{0(i)} = C'_{0(i)}, C_{1(i)} = C'_{0(i)}, C_{2(i)} = C'_{1(i)}, C_{3(i)} = C'_{2(i)} \text{ if } \rho_1 < \rho_2 \\ C_{0(i)} = C'_{0(i)}, C_{1(i)} = C'_{1(i)}, C_{2(i)} = C'_{2(i)}, C_{3(i)} = C'_{2(i)} \text{ if } \rho_2 < \rho_1 \end{cases}, \quad (3)$$

式中: d 表示括号中两个坐标点之间的距离; ρ_1 表示假设情况 1) 时第 i 帧与第 $i-1$ 帧图像中对应运动角点距离之和, ρ_2 表示假设情况 2) 时第 i 帧与第 $i-1$ 帧图像中对应运动角点的距离之和。

对于如何将各图像帧中提取的运动角点正确分配到四肢的问题,本文建立了一种基于特定交换规则的分配方法。首先,按照初始状态时四肢与头部的距离从近到远,依次用前肢 1、前肢 2、后肢 1、后肢 2 代表四肢,并将其在各帧视频图像下的运动角点坐标集合用 $G_0 = \{g_{1(i)} | i = 1, 2, 3 \dots\}$ 、 $G_1 = \{g_{2(i)} | i = 1, 2, 3 \dots\}$ 、 $G_2 = \{g_{3(i)} | i = 1, 2, 3 \dots\}$ 、 $G_3 = \{g_{4(i)} | i = 1, 2, 3 \dots\}$ 表示,其元素 $g_{1(i)}$ 、 $g_{2(i)}$ 、 $g_{3(i)}$ 、 $g_{4(i)}$ 分别表示前肢 1、前肢 2、后肢 1、后肢 2 在第 i 帧图像下的运动角点坐标。设第 1 帧图像为有效初始帧,即第 1 帧图像中没有运动角点重合,由于目标行走方向为自右向左,可直接将 $\{C_{0(1)}, C_{1(1)}, C_{2(1)}, C_{3(1)}\}$ 顺序分配给 $g_{1(1)}$ 、 $g_{2(1)}$ 、 $g_{3(1)}$ 、 $g_{4(1)}$ 。由于四足动物在行走过程中四肢前后摆腿交替进行,两前肢的两个运动角点在图像坐标系下 x 方向的相对距离 $d_x(g_{1(i)}, g_{2(i)})$ 存在逐渐减小至零然后逐渐增大的变化趋势,两后肢的两个运动角点在图像坐标系下 x 方向的相对距离 $d_x(g_{3(i)}, g_{4(i)})$ 也存在同样的变化趋势。当两前肢或两后肢的运动角点在 x 方向上的相对距离为零时,两运动角点坐标 x 值的升序排列位置将发生交换,需要将当前四肢与运动角点坐标的

分配关系进行交换,具体方法如下:

1) if $d_x(g_{1(i)}, g_{2(i)}) = 0$,

$$\begin{cases} \text{case 1: } \begin{cases} g_{1(i)} = C_{0(i)} \xrightarrow{\text{exchange}} g_{1(i+1)} = C_{1(i+1)} \\ g_{2(i)} = C_{1(i)} \xrightarrow{\quad \quad \quad} g_{2(i+1)} = C_{0(i+1)} \end{cases}; \\ \text{case 2: } \begin{cases} g_{1(i)} = C_{1(i)} \xrightarrow{\text{exchange}} g_{1(i+1)} = C_{0(i+1)} \\ g_{2(i)} = C_{0(i)} \xrightarrow{\quad \quad \quad} g_{2(i+1)} = C_{1(i+1)} \end{cases} \end{cases}$$

2) if $d_x(g_{3(i)}, g_{4(i)}) = 0$,

$$\begin{cases} \text{case 1: } \begin{cases} g_{3(i)} = C_{2(i)} \xrightarrow{\text{exchange}} g_{3(i+1)} = C_{3(i+1)} \\ g_{4(i)} = C_{3(i)} \xrightarrow{\quad \quad \quad} g_{4(i+1)} = C_{2(i+1)} \end{cases}; \\ \text{case 2: } \begin{cases} g_{3(i)} = C_{3(i)} \xrightarrow{\text{exchange}} g_{3(i+1)} = C_{2(i+1)} \\ g_{4(i)} = C_{2(i)} \xrightarrow{\quad \quad \quad} g_{4(i+1)} = C_{3(i+1)} \end{cases} \end{cases}$$

利用这一方法对后续图像帧进行处理,以保持各运动角点与四肢分组间的正确匹配关系。

对于如何区分四肢运动角点的同侧位置关系的问题,本文使用对比前后肢运动角点相对距离的方法进行判断。通过观察发现,在四足动物行走过程中,位于同侧的前后肢运动角点与处于对角上非同侧的前后肢运动角点相比,其相对距离的变化趋势以及最小值存在较为明显的区别。按照前文的描述,在四足动物行走方向为自右向左的情况下,运动角点经过重构和匹配处理后, G_0 、 G_1 属于前肢, G_2 、 G_3 属于后肢。此时存在两种可能:1) G_0 与 G_2 属于同一侧, G_1 与 G_3 属于另一侧;2) G_0 与 G_3 属于同一侧, G_1 与 G_2 属于另一侧。分别计算 $d_{\min}(G_0, G_2)$ 和 $d_{\min}(G_0, G_3)$, 如果 $d_{\min}(G_0, G_2) > d_{\min}(G_0, G_3)$, 则判

定 G_0 与 G_3 属于同一侧, G_0 与 G_2 属于另一侧; 反之, 则 G_0 与 G_2 属于同一侧, G_1 与 G_3 属于另一侧。

同理, 如果视频图像中四足动物行走方向为自左向右, 如图 12(d) 所示, 可以根据上述思路和步骤给出对应的运动角点匹配方法。

3.4 运动特征识别

目前针对四足动物步态的研究大多基于人工实现定量分析, 为有效地解决四足动物行走步态参数的自动提取的问题, 本文基于对四足动物行走运动的分析, 建立了一种基于四肢运动角点至参考点的距离变化曲线的分析方法, 该方法计算简单、高效, 可有效解决四足动物步态参数的自动提取问题。

本文方法以图像原点为参考点, 分别计算视频中 4 个运动角点与参考点的欧氏距离, 首先根据距离变化曲线判断四足动物行走的左右方向, 即: 运动角点与图像参考点的距离逐渐减小表明四足动物自左向右运动; 反之, 表明四足动物自右向左运动。由左右方向可确定四足动物的前后肢。四足动物在运动过程中, 各肢均存在支撑相和摆动相两个状态, 这两个状态交替出现、周期变化。根据本文方法: 如果蹄部处于支撑相, 则运动角点到参考点的距离不会发生变化; 如果蹄部处于摆动相, 则运动角点到参考点的距离将发生变化。为便于说明, 对处于同一侧的运动角点用相同的数字编号表示。假设四足动物自右向左运动且 G_0 和 G_2 属于同一侧, 分别用前肢 1 和后肢 1 表示, G_1 和 G_3 属于另一侧, 分别用前肢 2 和后肢 2 表示。理论上, 视频序列中四肢运动角点的距离变化曲线可用图 13 表示。

1) 步态周期和步态频率

四肢中某一肢的运动角点从离开地面开始运

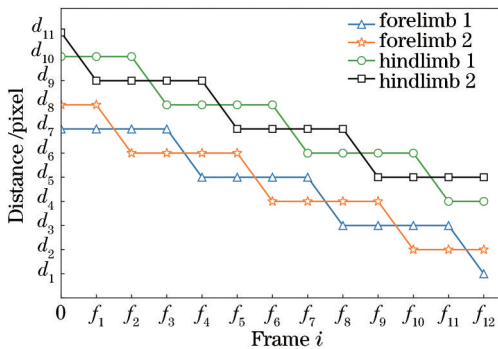


图 13 角点距离变化曲线

Fig. 13 Curves of corner distance variation

动至此角点并再次开始运动的时间间隔称为这一肢的步态周期, 步态周期与步态频率互为倒数关系。由图 13 可知, 通过分析运动角点距离变化曲线的纵坐标是否有变化, 即可判断各肢运动角点是处于支撑相还是摆动相, 进而获得各肢的步态周期。按照这一思路, 对运动角点的距离变化曲线进行均值滤波, 之后采用一阶差分法进行二值化, 0 代表支撑相, 1 代表摆动相。设前肢 1 的原始数据序列为 $u_{F1}(i)$, 则

$$\Delta u_{F1}(i) = u_{F1}(i+1) - u_{F1}(i), 1 \leq i \leq n-1, (4)$$

式中: $\Delta u_{F1}(i)$ 为前肢 1 距离变化差分。

根据一阶差分数据设定经验阈值 R , 若 $|\Delta u_{F1}(i)| \geq R$, 则 $\Delta u_{F1}(i)$ 中第 i 点等于 1, 反之等于 0。按照相同方法对前肢 2、后肢 1、后肢 2 运动角点的距离变化曲线进行二值化后, 得到各肢的步态周期图, 如图 14 所示, 进而计算得到各肢的步态周期和步态频率。

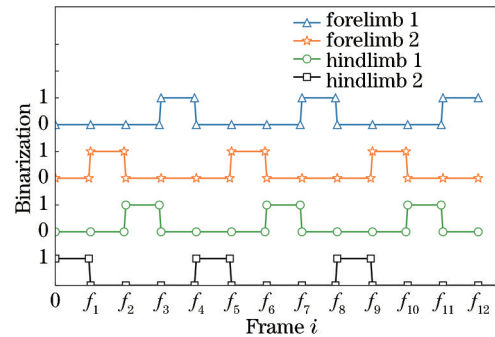


图 14 各肢步态周期图

Fig. 14 Gait cycle of each limb

以前肢 1 为例, 在图 14 中: 前肢 1 从 f_3 帧处开始运动至 f_4 帧, $f_3 \sim f_4$ 帧期间处于摆动相, 在 f_4 帧处前肢 1 落地, $f_4 \sim f_7$ 帧处于支撑相, 在 f_7 帧处前肢 1 再次开始运动, 重复 $f_3 \sim f_7$ 帧的过程, 即 $f_3 \sim f_7$ 为前肢 1 第一个完整的周期。前肢 2、后肢 1、后肢 2 呈现与前肢 1 类似的规律。

2) 步态顺序

步态顺序是四足动物步态研究的基础, 而行走步态是四足动物在不同自然条件下最常使用的一种低速运动方式。所谓步态顺序, 即四足动物四肢按照一定的顺序依次摆动。但四足动物在运动过程中四肢间的摆动过程极短, 同时存在时间和空间的变换, 运动情况复杂, 这给步态顺序的提取带来了一定的困难。基于上述问题, 本文方法在有限二维空间信息的基础上, 将四足动物的四肢运动状态

简化为摆动相和支撑相,同时完整地保留了时间信息,精确地还原四足动物行走运动过程中的摆腿顺序。如图 14 所示,在 $0 \sim f_1$ 帧时,后肢 2 率先开始运动,后肢 1、前肢 1、前肢 2 处于支撑相;在 $f_1 \sim f_2$ 帧,前肢 2 运动,前肢 1、后肢 1、后肢 2 处于支撑相;在 $f_2 \sim f_3$ 帧,后肢 1 运动,前肢 1、前肢 2、后肢 2 处于支撑相;在 $f_3 \sim f_4$ 帧,前肢 1 运动,前肢 2、后肢 1、后肢 2 处于支撑相。其步态顺序为:后肢 2,前肢 2,后肢 1,前肢 1...,腿部主要以三腿支撑为主。

3) 步幅

定义四足动物四肢中某一肢的运动角点从离开地面到此运动角点接触地面跨越的距离为步幅。从二维图像中得到的距离是以 pixel 为单位,本文采用简单标定法来实现空间距离单位的转换,设水平和垂直的距离转换比例分别为 L 、 H 。

在图 14 中,前肢 1 在第一周期内的步幅用 d_{s1} 表示,计算方式如下:设前肢 1 在 f_3 帧图像中的坐标为 (x_{F1}^3, y_{F1}^3) ,在 f_4 帧图像中的坐标为 (x_{F1}^4, y_{F1}^4) ,则有

$$d_{s1} = \sqrt{[(x_{F1}^4 - x_{F1}^3) \cdot L]^2 + [(y_{F1}^4 - y_{F1}^3) \cdot H]^2}, \quad (5)$$

前肢 2、后肢 1、后肢 2 的步幅可按相同方法计算。

表 1 角点检测精度对比

Table 1 Comparison of corner detection accuracy

Type	DeeplabV3+		Improved DeeplabV3+	
	Accuracy / %	Deviation value / pixel	Accuracy / %	Deviation value / pixel
Rhino	80.2	38	84.6	32
Buffalo	82.4	36	85.9	27
Alpaca	93.7	28	100.0	19

利用本文提出的角点检测及匹配方法对犀牛、水牛、羊驼的行走视频进行处理,将匹配后的角点数据标记于原图像,如图 15 所示,图中 F1、F2、H1、H2 分别表示四足动物的前肢 1、前肢 2、后肢 1、后肢 2。由实验结果可知,本文方法即使在运动目标存在角度偏移的情况下也能够准确地提取到四足动物的运动角点,在其四肢交替前行的过程中亦能实现对运动角点关系的准确匹配。

2) 步态周期和步态频率

犀牛、水牛、羊驼四肢的运动角点距离变化曲线如图 16、18、20 所示,进行二值化后,犀牛、水牛、羊驼各肢的步态周期分别如图 17、19、21 所示。

4 实验与分析

4.1 步态参数提取

在实验中,挑选三段分别包含完整步态周期的犀牛、水牛、羊驼行走运动的视频作为实验测试集进行分析处理。采用犀牛、水牛、羊驼的最小外接矩形面积与图像总面积的比值(分别约为 6.80%、6.53%、15%)度量实验对象的尺度。

1) 角点检测及匹配

为验证本文方法对运动角点检测的有效性,分别以改进前后的 DeeplabV3+ 进行角点检测和对比实验。角点检测的准确度评价方法为:以 pixel 为单位,计算视频帧中不同方法检测的运动角点坐标与人工标定角点坐标之间的欧氏距离小于 20 pixel 的运动角点个数占运动角点总数的百分比及检测角点与人工标记角点的最大偏离像素。由表 1 可知,改进后的 DeeplabV3+ 对犀牛、水牛、羊驼运动角点坐标的定位精度都有了不同程度的提升,表明本文方法对于四足动物语义分割的改进效果明显。此外,作为对比,文献[4]采用基于 ViBe 算法提取奶牛四肢并通过连通区域最小坐标进行牛蹄定位的方法,在阈值为 20 pixel 时精确度为 73.8%,最大偏离像素为 49 pixel。相比之下,本文方法具有更高的精度。

由图 16、18 可知,犀牛、水牛的运动角点与图像原点的距离逐渐减小,说明犀牛、水牛自左向右运动,进一步可以确定犀牛、水牛的前肢和后肢。由图 17、19、21 可以直接计算出犀牛、水牛、羊驼四肢的步态周期和步态频率的平均值,为验证方法的准确性,通过人工统计数据并进行对比,结果如表 2 所示。结果表明,步态周期和频率的相对误差不超过 2.00%,说明本文方法可以有效地提取四足动物的步态周期和频率。进一步分析可知:表 2 中犀牛各肢步态周期、频率的波动较大,频率误差是最大的。结合犀牛的实际情况进行分析,发现此时犀牛前方出现障碍物,导致犀牛运动速度降低。对比

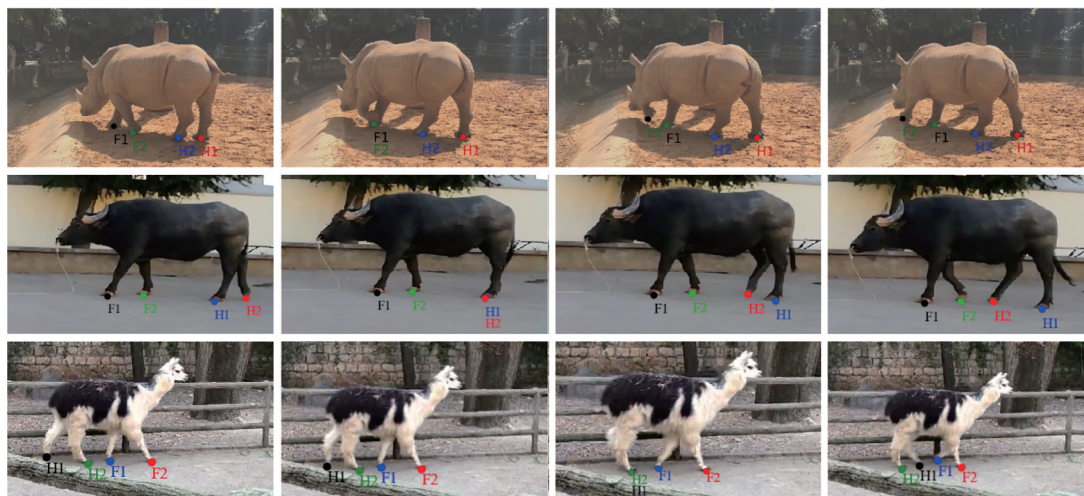


图 15 运动角点检测及匹配

Fig. 15 Motion corner detection and matching

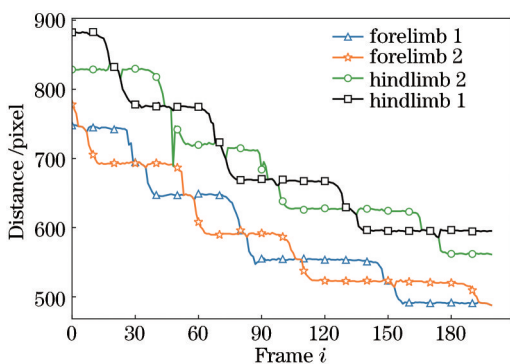


图 16 犀牛角点距离变化曲线

Fig. 16 Rhino's corner distance variation curves

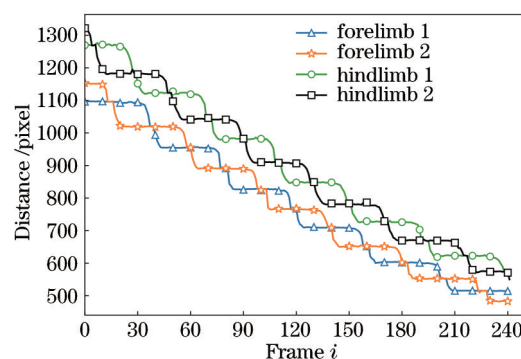


图 18 水牛角点距离变化曲线

Fig. 18 Buffalo's corner distance variation curves

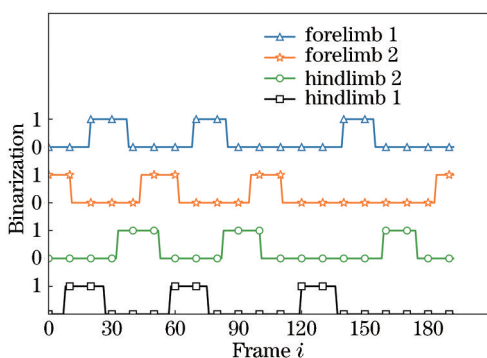


图 17 犀牛各肢步态周期图

Fig. 17 Rhino's gait cycle of each limb

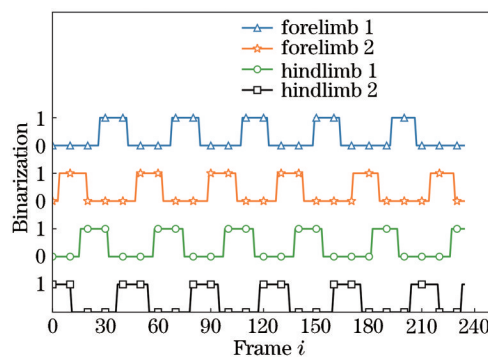


图 19 水牛各肢步态周期图

Fig. 19 Buffalo's gait cycle of each limb

图 17 可知,在犀牛速度减小的过程中,摆动相所占帧数基本不变,但支撑相所占帧数明显增多,即负荷因数(单腿着地时间与步态周期的比值)变大。

3) 步态顺序

水牛、犀牛、羊驼行走时四肢的步态顺序如图 22(a)~(c)所示,与实际步态顺序是完全一致

的,说明步态顺序的识别是准确的。从图中还可看出,三种动物的步态顺序遵循共同的规律,即某一后肢运动后,其同侧前肢开始运动,然后切换另一侧后肢运动,接着另一侧前肢开始运动,按照这一规律周而复始、循环进行。初始运动足可以是前肢也可以是后肢,但最终的步态顺序会趋于一致。此外,通过对犀牛、水牛、羊驼行走步态的分析可知,

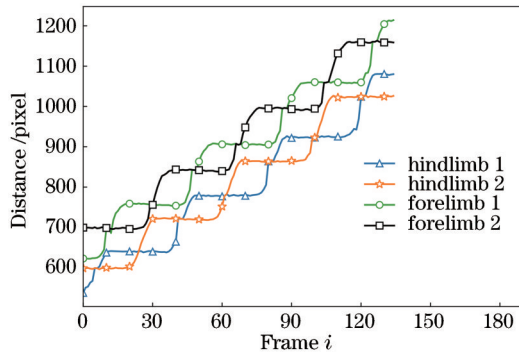


图 20 羊驼角点距离变化曲线

Fig. 20 Alpaca's corner distance variation curves

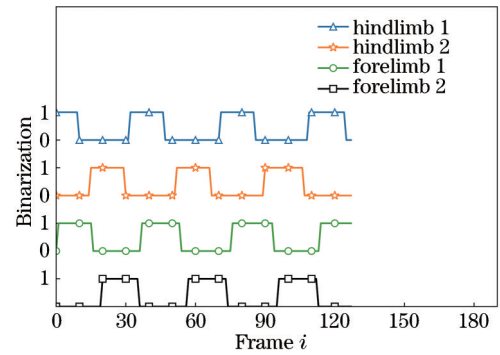


图 21 羊驼各肢步态周期图

Fig. 21 Alpaca's gait cycle of each limb

在行走过程中,由于四肢的摆动相很少重叠,可以获得较好的运动稳定性。三者腿部支撑主要以三腿支撑为主,但同侧两肢运动状态切换以及对角两

肢运动状态切换时,可能会出现短暂的同侧两肢和对角两肢离地的情况,未出现腾空和单腿支撑的情况。

表 2 周期和频率值

Table 2 Values of cycle and frequency

Type	Limb	Rhino		Buffalo		Alpaca	
		T/s	f/Hz	T/s	f/Hz	T/s	f/Hz
Method of this article	Forelimb 1	2.00	0.50	1.37	0.73	1.28	0.78
	Forelimb 2	2.33	0.43	1.39	0.72	1.28	0.78
	Hindlimb 1	2.10	0.48	1.38	0.72	1.28	0.78
	Hindlimb 2	1.87	0.53	1.39	0.72	1.28	0.78
	Average value	2.08	0.49	1.38	0.72	1.28	0.78
Manual calculation	Forelimb 1	1.97	0.51	1.38	0.72	1.26	0.79
	Forelimb 2	2.13	0.47	1.39	0.72	1.27	0.79
	Hindlimb 1	2.07	0.48	1.36	0.74	1.30	0.77
	Hindlimb 2	1.93	0.52	1.38	0.72	1.21	0.82
	Average value	2.07	0.50	1.38	0.73	1.26	0.79
Error / %		0.50	2.00	0.00	1.37	1.56	1.27

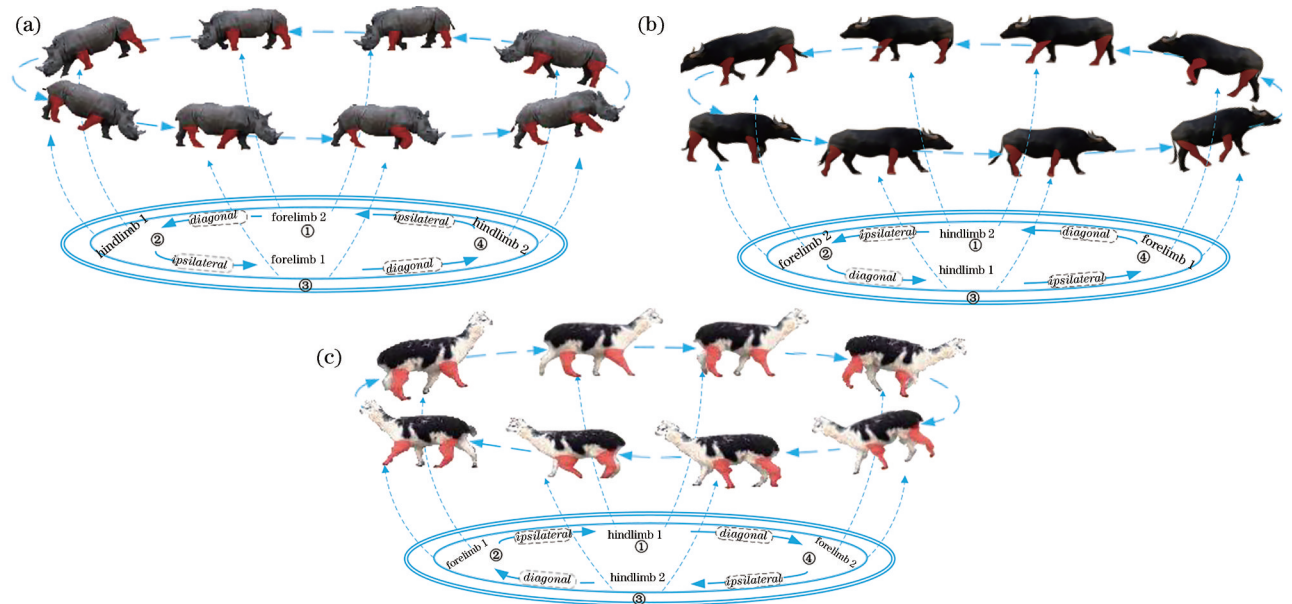


图 22 步态顺序。(a)水牛步态顺序;(b)犀牛步态顺序;(c)羊驼步态顺序

Fig. 22 Gait sequences. (a) Gait sequence of buffalo; (b) gait sequence of rhino; (c) gait sequence of alpaca

4) 角点摆动角

依次计算出视频序列每一帧图片中的角点摆动角,如图 23 所示。从图中可见,犀牛、水牛、羊驼右后肢的角点摆动角均呈周期性变化,其变化周期与步态周期具有相关性。三个实验对象的角点摆动角均出现了跨越正负的情况,说明四足动物在行走过程中,

运动角点可能会超过轮廓中心点。由于个体的差异性,不同四足动物间的角点摆动角的变化范围也不相同。其中体型较小的羊驼的四肢角点摆动角的变化范围都在 40° 左右。而体型较大的犀牛、水牛前后肢的角点摆动角的变化范围却存在差异,相对前肢角点摆动角而言,后肢角点摆动角的变化范围更小。

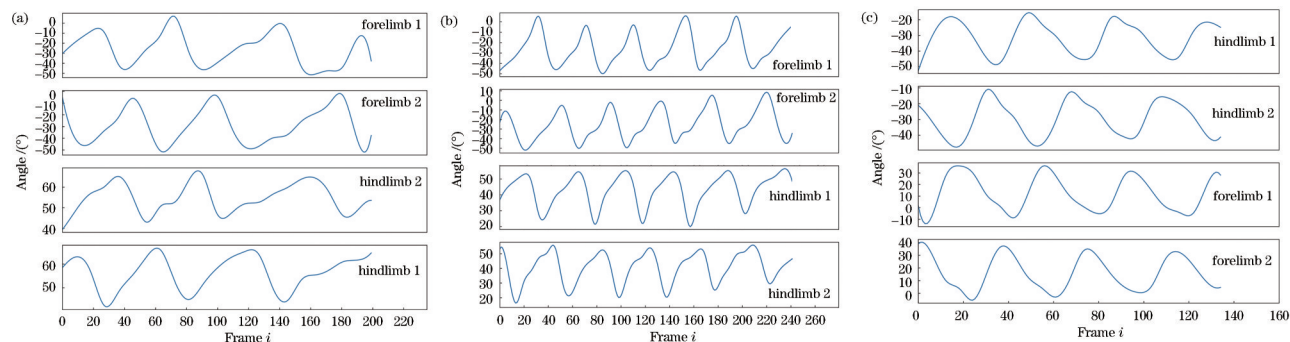


图 23 角点摆动角。(a)犀牛角点摆动角;(b)水牛角点摆动角;(c)羊驼角点摆动角

Fig. 23 Corner swing angles. (a) Corner swing angle of rhino; (b) corner swing angle of buffalo; (c) corner swing angle of alpaca

5) 步幅

采用前述方法计算的步幅如表 3 所示。本文方法与人工统计相比,步幅计算误差的最大值为 2.85%,说明本文方法对步幅的计算较为准确。犀牛、

水牛和羊驼各自四肢间步幅比较接近,相近的幅度使得其在行走的过程中获得较好的平稳性。但由于各自体型的差异,实验对象间的步幅存在较大差异。一般而言,体型越高、越长,对应的步幅就会越大。

表 3 水牛、犀牛和羊驼的步幅对比

Table 3 Comparison of stride lengths of rhino, buffalo, and alpaca

Limb	Rhino		Buffalo		Alpaca	
	Method of this article	Manual calculation	Method of this article	Manual calculation	Method of this article	Manual calculation
Forelimb 1	1.43	1.35	1.15	1.12	0.55	0.57
Forelimb 2	1.33	1.22	1.05	1.10	0.52	0.50
Hindlimb 1	1.40	1.45	1.10	1.15	0.50	0.51
Hindlimb 2	1.44	1.38	1.13	1.17	0.52	0.54
Average value	1.40	1.36	1.11	1.14	0.52	0.53
Average error /%	2.85		2.63		1.89	

4.2 目标尺度变化对角点定位的影响

在拍摄动物时,受拍摄距离变化及不同动物体型差异的影响,图像中动物目标尺度会有所变化。为了分析动物体型差异的影响,本文以体型较小的犬(萨摩耶,体高约为 0.5 m,体长约为 0.6 m)为实验对象,调整拍摄距离,获得不同尺度的目标图像,进行角点检测实验。实际测得三组视频中犬的最小外接矩形面积与图像总面积的比值 R 分别约为 1.06%、3.25%、6.24%。角点检测结果如图 24 所示,当动物尺度较小时,分割后的前景图像中四肢

易出现缺失、粘贴过度等情况,此时算法不能准确检测到运动角点。当 $R=3.25\%$ 时,算法可以更好地区分犬的四肢,其运动角点则可以被检测到,与人工标注点相比,最大偏离像素为 23 pixel。当 $R=6.24\%$ 时,运动角点与人工标注点的最大偏离像素减小到 19 pixel。实验结果表明,本文方法具有一定的尺度适应性,在动物目标尺度较大的情况下准确度更高,当动物体型较小时,图像采集易造成目标尺度过小、运动角点定位错误。



图 24 不同尺度下的角点检测。(a) $R=1.06\%$; (b) $R=3.25\%$; (c) $R=6.24\%$

Fig. 24 Corner detection for different scales. (a) $R=1.06\%$; (b) $R=3.25\%$; (c) $R=6.24\%$

5 结 论

将计算机视觉技术和深度学习方法相结合,针对四足动物的行走运动,建立了一种四足动物行走步态参数的自动提取方法,在不接触检测目标的情况下,可以准确地识别和定位动物四肢运动角点,自动提取步态周期、步态频率、步态顺序、步幅、角点摆动角、运动速度等步态参数。其中运动速度可根据运动角点或轮廓中点的移动距离与时间的比值获得,本文未进行详细说明。本文方法解决了在二维图像中对四足动物四肢蹄部的运动轨迹进行追踪的问题,具有较高的鲁棒性。最后,以犀牛、水牛、羊驼为实验对象,对本文方法的有效性进行了验证。实验结果表明,本文方法对四足动物运动角点的检测快速有效,对犀牛、水牛、羊驼运动角点检测的最大误差分别为 32, 27, 19 pixel, 运动角点匹配关系准确,步态周期、步态频率的计算误差小于 2%,步态顺序输出正确,步幅的计算误差最大为 2.85%。

通过实验分析发现,受视觉感知信息方式的局限性,当拍摄角度倾斜过大时,动物四肢的运动轨迹会发生重叠遮挡,造成本文算法失效;当动物目标尺度较小时,可能出现角点定位错误的问题。同时,本文目前所做工作主要是针对典型四足动物的行走运动,后续研究将继续改进完善,针对其他运动模式开展研究。

参 考 文 献

- [1] Thorup V M, Munksgaard L, Robert P E, et al. Lameness detection via leg-mounted accelerometers on dairy cows on four commercial farms[J]. *Animal*, 2015, 9(10): 1704-1712.
- [2] Maertens W, Vangeyte J, Baert J, et al. Development of a real time cow gait tracking and analysing tool to assess lameness using a pressure sensitive walkway: the GAITWISE system[J]. *Biosystems Engineering*, 2011, 110(1): 29-39.
- [3] Yang Q M, Xiao D Q, Zhang G X. Automatic pig drinking behavior recognition with machine vision[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2018, 49(6): 232-238.
杨秋妹, 肖德琴, 张根兴. 猪只饮水行为机器视觉自动识别[J]. *农业机械学报*, 2018, 49(6): 232-238.
- [4] Kang X, Zhang X D, Liu G, et al. Hoof location method of lame dairy cows based on machine vision [J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2019, 50(S1): 276-282.
康熙, 张旭东, 刘刚, 等. 基于机器视觉的跛行奶牛蹄蹄定位方法[J]. *农业机械学报*, 2019, 50(S1): 276-282.
- [5] Zhang X L, Liang Y. A baby-mimic insufficient-DOF quadruped crawling robot[J]. *Robot*, 2016, 38(4): 458-466.
张秀丽, 梁艳. 一种仿婴儿欠自由度四足爬行机器人[J]. *机器人*, 2016, 38(4): 458-466.
- [6] Daou H E, Libourel P A, Renous S, et al. Methods and experimental protocols to design a simulated bio-mimetic quadruped robot[J]. *International Journal of Advanced Robotic Systems*, 2013, 10(5): 256.
- [7] Kim C H, Shin H C, Lee H H. Trotting gait analysis of a lizard using motion capture[C]//13th International Conference on Control, Automation and Systems, October 20-23, 2013, Gwangju, Korea (South). New York: IEEE Press, 2013: 1247-1251.
- [8] Liu B, Zhu W X, Yang J J, et al. Extracting of pig gait frequency feature based on depth image and pig skeleton endpoints analysis[J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2014, 30(10): 131-137.
刘波, 朱伟兴, 杨建军, 等. 基于深度图像和生猪骨架端点分析的生猪步频特征提取[J]. *农业工程学报*, 2014, 30(10): 131-137.
- [9] Cangar Ö, 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.
- [10] Xue F F, Wang Y M, Li Q. Recognition of cattle

- daily behavior based on spatial relationship of feature parts[J]. *Laser & Optoelectronics Progress*, 2021, 58(22): 2215007.
- 薛芳芳, 王月明, 李琦. 基于特征部位空间关系的牛日常行为识别[J]. *激光与光电子学进展*, 2021, 58(22): 2215007.
- [11] Cheng X Y, Zhao L Z, Hu Q, et al. Real-time semantic segmentation based on dilated convolution smoothing and lightweight up-sampling[J]. *Laser & Optoelectronics Progress*, 2020, 57(2): 021017.
- 程晓悦, 赵龙章, 胡穹, 等. 基于膨胀卷积平滑及轻型上采样的实时语义分割[J]. *激光与光电子学进展*, 2020, 57(2): 021017.
- [12] Zhang W X, Zhu Z C, Zhang Y H, et al. Cell image segmentation method based on residual block and attention mechanism[J]. *Acta Optica Sinica*, 2020, 40(17): 1710001.
- 张文秀, 朱振才, 张永合, 等. 基于残差块和注意力机制的细胞图像分割方法[J]. *光学学报*, 2020, 40(17): 1710001.
- [13] Badrinarayanan V, Kendall A, Cipolla R. SegNet: a deep convolutional encoder-decoder architecture for image segmentation[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(12): 2481-2495.
- [14] Zhao H S, Shi J P, Qi X J, et al. Pyramid scene parsing network[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 6230-6239.
- [15] Chen L C, Papandreou G, Kokkinos I, et al. Semantic image segmentation with deep convolutional nets and fully connected CRFs[EB/OL]. (2014-12-22) [2021-02-01]. <https://arxiv.org/abs/1412.7062v2>.
- [16] Chen L C, Papandreou G, Kokkinos I, et al. DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018, 40(4): 834-848.
- [17] Chen L C, Papandreou G, Schroff F, et al. Rethinking atrous convolution for semantic image segmentation[EB/OL]. (2017-06-17) [2021-03-06]. <https://arxiv.org/abs/1706.05587>.
- [18] Chen L C, Zhu Y, Papandreou G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation[EB/OL]. (2018-08-22) [2020-03-06]. <http://export.arxiv.org/abs/1802.02611>.
- [19] Howard A G, Zhu M, Chen B, et al. MobileNets: efficient convolutional neural networks for mobile vision applications[EB/OL]. (2017-04-17) [2021-03-06]. <https://arxiv.org/abs/1704.04861>.
- [20] Chollet F. Xception: deep learning with depthwise separable convolutions[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 1800-1807.
- [21] Tian Z, He T, Shen C H, et al. Decoders matter for semantic segmentation: data-dependent decoding enables flexible feature aggregation[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 15-20, 2019, Long Beach, CA, USA. New York: IEEE Press, 2019: 3121-3130.
- [22] Sun K, Xiao B, Liu D, et al. Deep high-resolution representation learning for human pose estimation [C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 15-20, 2019, Long Beach, CA, USA. New York: IEEE Press, 2019: 5686-5696.