

УДК 519.254

# ОЦЕНИВАНИЕ КИНЕМАТИЧЕСКИХ ХАРАКТЕРИСТИК ПОДВИЖНЫХ ОБЪЕКТОВ СЛОЖНОЙ ФОРМЫ ЧЕРЕЗ ИХ ЛОКАЛИЗАЦИЮ НА ВИДЕОПОСЛЕДОВАТЕЛЬНОСТИ АЛГОРИТМАМИ ГЛУБОКОГО ОБУЧЕНИЯ

**В.А. Галкин**

*Институт проблем управления им. В.А. Трапезникова РАН*  
Россия, 117997, Москва, Профсоюзная ул., 65  
E-mail: galkin.vsevolod@gmail.com

**А.В. Макаренко**

*Институт проблем управления им. В.А. Трапезникова РАН*  
Россия, 117997, Москва, Профсоюзная ул., 65  
E-mail: avm.science@mail.ru

**Ключевые слова:** подвижные объекты, кинематические характеристики, оценивание, компьютерное зрение, глубокое обучение, локализация объектов, Фурье анализ, походка.

**Аннотация:** Продемонстрирован подход к оцениванию и анализу кинематических характеристик движущихся объектов по ограничивающему прямоугольнику, формируемому при локализации наблюдаемых объектов сложной формы. Отличительной чертой работы является отсутствие необходимости разработки шарнирной модели изучаемого объекта. Предложенный подход продемонстрирован на задаче анализа походки сельскохозяйственных животных. Проведено предварительное обучение нейросетевого детектора YoloV8 на видеороликах, сформированы треки объектов с использованием алгоритма SORT. На выборке данных вычислены оценки скоростей и ускорений объектов с использованием алгоритма «total variation regularization» и получены их Фурье-спектры. Установлено статистически значимое различие спектров в диапазоне [1, 2] Гц для групп животных различающихся состоянием опорно-двигательного аппарата и предварительно разделенных экспертами.

## 1. Введение

Развитие науки и техники, и, в частности, методов компьютерного зрения и глубокого обучения, позволяют решать в автоматическом режиме все более сложные задачи, в том числе такие, как оценивание и анализ кинематических характеристик подвижных объектов сложной формы. Мейнстрим-задачей в направлении является анализ походки человека [1] и ряд практических приложений на ее основе:

обнаружение аномалий и заболеваний, биометрия, мониторинг активности и другие. Стоит отметить, что анализом походки человека актуальность направления не ограничивается. Прикладной анализ походки животных является не менее релевантной задачей.

В данном направлении методы разделяют на две группы по типу источников данных:

1. датчики прямого контакта – инерциальные системы (гироскопы, акселерометры) [2], электро гониометры [1], оптико-электронные системы [1], магнитные системы [3], специализированные платформы [4] и др.
2. бесконтактные датчики – камеры машинного зрения [5].

Использование датчиков прямого контакта для съема данных требует установки сенсоров на изучаемый объект, что в свою очередь приводит к ряду проблем, в том числе масштабирования решения. Методы, основанные на камерах машинного зрения позволяют производить анализ дистанционно и бесконтактно. По этой причине, они более универсальны и экономически оправданы.

Превалирующее число работ по компьютерному зрению описывают сложную техническую и алгоритмическую реализацию системы, накладывают требования на составление шарнирной модели объекта, описывающей кинематику объекта по компонентам. Проблематика работы с шарнирной моделью заключается в двух основных факторов:

1. экспертное составление шарнирной модели.
2. обучение нейросетевой модели на задачу оценку позы [6].

Составление шарнирной модели, особенно в сложных случаях, приводит к необходимости задействовать группы экспертов предметной области. Обучение нейросетевой модели для автоматической генерации шарнирной модели изучаемого объекта по видеопоследовательности требует наличия размеченной выборки данных; кроме того, процесс обучения модели, как правило, осложняется рядом негативных моментов: низкое освещение порождает смаз, сложная форма объекта, низкий контраст изображения, и т.п. Привлечение экспертов предметной области, зачастую приводит к экономическим и временным издержкам, и проблематике несогласованности их оценок [7].

Отличительная черта настоящей работы заключается в оценивании кинематических характеристик объектов сложной формы по ограничивающему прямоугольнику. С одной стороны, «грубое» представление объекта не позволяет выполнить глубокой анализ кинематики, с другой стороны, исключается необходимость шарнирной модели, что позволяет оперативно получить результаты, которые в некоторых приложениях являются достаточными. Предлагаемый подход демонстрируется на задаче анализа походки сельскохозяйственных животных.

## 2. Набор данных

Набор данных – множество видеороликов с проходами животных. Каждый видеоролик – это трек – содержит проход одного животного. Сбор данных

осуществлялся с помощью камеры высокого разрешения с частотой кадров 20 fps. Сырые видеоролики были предварительно обработаны, в частности, были устранены искажения дисторсии [8]. Общее количество треков в наборе данных равно 1 267.

По каждому треку имеется автоматическая разметка, ограничивающими прямоугольниками – решение задачи локализации. Следует отметить, что разметка для решения задачи локализации существенно проще, нежели для задачи оценивания позы. Для генерации ограничивающих рамок была использована предобученная на задаче локализации нейронная сеть YoloV8 [9] и алгоритм слежения (треккер) SORT [10]. Точность предобученной модели на тестовой выборке составляет 0.98 по метрике  $F_1$  и 0.927 по метрике IoU.

### 3. Эксперимент

#### 3.1. Оценка скоростей и ускорений

Как известно, операция численного дифференцирования является некорректной [11]. Для вычисления скорости и ускорения воспользуемся регуляризацией посредством алгоритма «total variation regularization» [12] со значением гиперпараметра  $\gamma = 0.025$ . В работе [12] экспериментально показано, что использование данного алгоритма на зашумленных данных в сравнении с прямой разностной схемой, погрешность вычисления производной по метрике RMS уменьшается более чем в 8 раз. Дополнительно данные фильтруются за счет усреднения – через вычисление центра прямоугольника. В силу ограниченности места для изложения, приведем только выборочный результат (см. рис. 1).

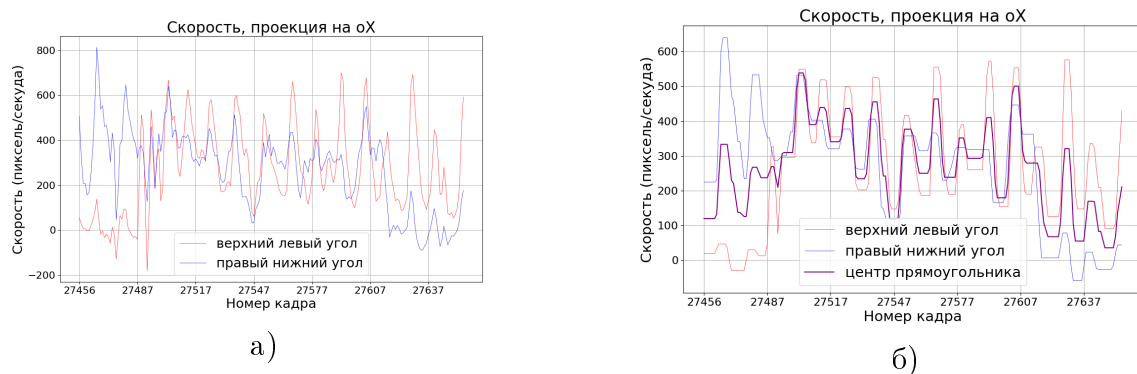


Рис. 1. Проекция скорости на  $oX$  объекта (один трек): а) использование разностной схемы первого порядка, б) использование алгоритма «total variation regularization»

Из приведенных результатов видно, что треки содержат некие, достаточно периодичные колебательные компоненты. Выполним спектральный анализ набора данных.

#### 3.2. Оценка спектра

Сформированные из исходного набора данных сигналы стандартизировались, дополнялись нулевыми отсчетами до одинаковой длины. К преобразованным

сигналам применялось окно Ханна с корректирующим коэффициентом  $\sqrt{8}/3$ .

Для анализа выборки высчитывался усредненный спектр по всем сигналам согласно формуле:

$$G(f_k) = \frac{1}{n_d N \Delta t} \sum_{i=1}^{n_d} |c X_i(f_k)|^2,$$

где:  $X_i(f_k)$  – преобразование Фурье отдельного сигнала;  $c$  – корректирующий коэффициент для окна Ханна;  $n_d$  – количество сигналов;  $N$  – число отсчетов в последовательности;  $\Delta t = 0.05$  с – шаг между отсчетами, согласно FPS.

Выборка данных также была предварительно разделена экспертами предметной области на две группы, различающиеся состоянием опорно-двигательного аппарата. В работе для них введены обозначения: «Группа 1» и «Группа 2».

Спектры проекций скоростей и ускорений на оси  $\{oX, oY\}$  приведены на рисунке 2. Для визуализации и анализа, осуществлялось преобразование шкалы спектров в дБ:  $\hat{G} = 10 \lg G$  относительно уровня 1.0.

Согласно эмпирическим оценкам, частота движений («шагов») изучаемых объектов равна  $\sim 1.25$  Гц. Следовательно, в диапазоне  $[1, 2]$  Гц на реальной спектрограмме ожидается пик. Экспериментальные результаты подтверждают эмпирические оценки, см. рисунок 2 (целевая область отмечена зеленым).

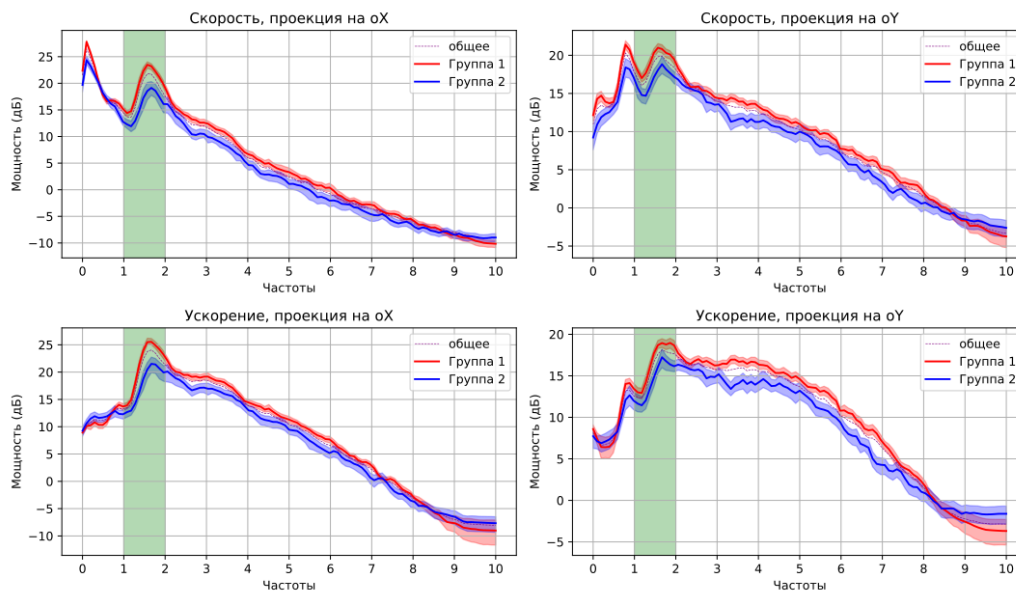


Рис. 2. Спектры проекций скоростей и ускорений на оси  $\{oX, oY\}$  с разделением по выделенным группам

Как видно из рисунка 2, выделенные экспертами группы изучаемых объектов являются разделимыми. Для подтверждения статистической значимости результата, были построены доверительные интервалы двумя методами:

1. с доверительной вероятностью 0.95.
2. с использованием бутстреп выборок и границ по уровням  $\{5, 95\}$  перцентилей.

В силу ограниченности места для изложения, приведем результаты построения первым методом. Из приведенных на рисунке 2 границ видно, что в

диапазоне [1, 2] Гц границы доверительных интервалов не пересекаются. Таким образом, в данном диапазоне частот, разделение групп объектов является статистически значимым.

## 4. Заключение

В работе представлен подход по оценке и анализу кинематических характеристик подвижных объектов сложной формы по ограничивающему прямоугольнику, формируемому в задаче локализации. Подход продемонстрирован на примере задачи анализа походки сельскохозяйственных животных. Было проведено предварительное обучение нейросетевого детектора YoloV8 на локализацию объектов, по видеоряду сформированы треки с использованием алгоритма SORT. На выборке данных были вычислены оценки проекций скоростей и ускорений с использованием алгоритма «total variation regularization» и получены их Фурье-спектры. Было установлено статистически значимое различие спектров в диапазоне [1, 2] Гц для двух групп животных, различающихся состоянием опорно-двигательного аппарата (априорное разделение животных на группы произведено экспертами предметной области).

## Список литературы

1. Sethi D., Bharti S., Prakash C. A comprehensive survey on gait analysis: History, parameters, approaches, pose estimation, and future work // *Artificial Intelligence in Medicine*. 2022. Vol. 129. P. 102314.
2. Petraglia F., et al. Inertial sensors versus standard systems in gait analysis: A systematic review and meta-analysis // *European journal of physical and rehabilitation medicine*. 2019. Vol. 55. No. 2. P. 268-280.
3. Cutti A.G., Ferrari A., Garofalo P., Raggi M., Cappello A., Ferrari A. 'Outwalk': a protocol for clinical gait analysis based on inertial and magnetic sensors // *Medical biological engineering computing*. 2010. Vol. 48. P. 17-25.
4. <https://www.tekscan.com/products-solutions/animal-gait-analysis> (дата обращения: 15.01.2024).
5. Fan C., et al. OpenGait: Revisiting Gait Recognition Towards Better Practicality // *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2023. P. 9707-9716.
6. Zheng C. et al. Deep learning-based human pose estimation: A survey // *ACM Computing Surveys*. 2023. Vol. 56. No. 1. P. 1-37.
7. Андрейчик Н.Л., Макаренко А.В. Методика предварительного анализа и повышения непротиворечивости датасета, маркируемого ограниченным числом экспертов-разметчиков // *Труды 16-й Всероссийской мультikonференции по проблемам управления (МКПУ-2023, Волгоград)*. Волгоград: ВолгГТУ, 2023. Т. 2. С. 38-41.
8. Носков Н.О., Галкин В.А., Макаренко А.В. Исследование возможности автоматизации процесса калибровки камеры по видеоданным на основе коэффициентов кривизны // *Труды 18-ой Всероссийской школы-конференции молодых ученых «Управление большими системами» (УБС'2022, Челябинск)*. Челябинск: Издательский центр ЮУрГУ, 2022. С. 128-134.
9. <https://github.com/ultralytics/ultralytics> (дата обращения: 11.01.2024).
10. Bewley A., et al. Simple online and realtime tracking // *2016 IEEE international conference on image processing (ICIP)*. IEEE, 2016.
11. Самарский А.А., Гулин А.В. Численные методы. 1989.
12. Van Breugel F., et al. PyNumDiff: A Python package for numerical differentiation of noisy time-series data // *Journal of Open Source Software*. 2022. Vol. 7. No. 71. P. 4078.