

<https://doi.org/10.26160/2474-5901-2026-53-102-110>

## ДЕТЕКТИРОВАНИЕ УДАРОВ В НАСТОЛЬНОМ ТЕННИСЕ НА ОСНОВЕ МЕТОДОВ КОМПЬЮТЕРНОГО ЗРЕНИЯ

*Примаченко П.Н., Хабаров С.П.*

*Санкт-Петербургский государственный лесотехнический университет  
имени С.М. Кирова, Санкт-Петербург, Россия*

**Ключевые слова:** настольный теннис, компьютерное зрение, анализ видео, оценка позы человека, детектирование ударов, MediaPipe Pose, спортивная аналитика.

**Аннотация.** Предложен метод автоматического детектирования ударов в настольном теннисе на основе анализа видео и оценки позы. Ключевые точки тела извлекаются через MediaPipe Pose, нормированная скорость кисти вычисляется по координатам рабочей руки, удары определяются по локальным максимумам скорости. Разработан программный прототип, который формирует выходное видео с наложенной скелетной моделью и таблицу обнаруженных ударов. Эксперимент на 136-секундной видеозаписи показал покрытие позы 98,37% и обнаружение 46 из 47 ударов, подтверждая практическую пригодность подхода для автоматизированной разметки тренировочных видео.

## SHOT DETECTION IN TABLE TENNIS BASED ON COMPUTER VISION METHODS

*Primachenko P.N., Khabarov S.P.*

*Saint-Petersburg State Forest Technical University, Saint-Petersburg, Russia*

**Keywords:** table tennis, computer vision, video analysis, human pose estimation, shot detection, MediaPipe Pose, sports analytics.

**Abstract.** A method for automatic shot detection in table tennis is proposed based on video analysis and pose estimation. Body keypoints are extracted using MediaPipe Pose, normalized wrist velocity is computed from the playing hand coordinates, and shots are detected based on local velocity maxima. A software prototype was developed that generates output video with an overlaid skeletal model and a table of detected shots. An experiment on a 136-second video recording showed 98.37% pose coverage and detection of 46 out of 47 shots, confirming the practical viability of the approach for automated annotation of training videos.

### Введение

Автоматический анализ спортивных видеозаписей является актуальным направлением исследований в области компьютерного зрения. Для настольного тенниса особый интерес представляет выделение ударных действий игрока, поскольку такая информация может использоваться для анализа техники, разметки тренировочных эпизодов и создания интеллектуальных систем поддержки тренера.

Традиционно анализ ударов выполняется тренером визуально, что требует значительных временных затрат и зависит от субъективной оценки. В связи с этим актуальной является разработка программных средств, способных автоматически обрабатывать видеозаписи и обнаруживать игровые действия спортсмена.

Задачи анализа движения человека по видеоданным широко исследуются в области компьютерного зрения. Классические подходы включают методы, основанные на анализе оптического потока и трекинге объектов [1]. В спортивной биомеханике широко применяются системы захвата движения, обеспечивающие высокую точность измерений, однако требующие специализированного оборудования [2]. В последние годы активно развиваются методы оценки позы человека на основе глубокого обучения, позволяющие извлекать координаты ключевых точек тела непосредственно из видеопоследовательностей [3, 4]. Именно этот подход выбран в работе для детектирования ударов по видеозаписям тренировок.

В отличие от существующих подходов, в работе предлагается адаптивный алгоритм детектирования на основе нормированной скорости и робастной статистики, не требующий калибровки камеры или специализированного оборудования.

Целью данной работы является разработка метода автоматического детектирования ударов в настольном теннисе на основе видеозаписей с использованием оценки позы игрока. Общая схема предлагаемого алгоритма представлена на рисунке 1.



Рис. 1. Общая схема алгоритма детектирования ударов

## 1. Постановка задачи

Задача заключается в автоматическом выделении дискретных моментов ударов из видеопоследовательности на основе анализа кинематики движений игрока. Математически она сводится к отображению видеопотока во множество временных меток, соответствующих локальным максимумам нормированной скорости кисти. Пусть задана видеопоследовательность:

$$V = \{F_i\}_{i=1}^N,$$

где  $F_i$  –  $i$ -й видеокادر, а  $N$  – количество кадров.

С применением модели оценки позы для каждого кадра извлекается конфигурация тела в виде множества ключевых точек:

$$P_i = p_i^{(k)}_{k=1}^K,$$

$$p_i^{(k)} = (x_i^{(k)}, y_i^{(k)}, v_i^{(k)}).$$

где  $x_i^{(k)}$ ,  $y_i^{(k)}$  – пространственные координаты  $i$ -й точки в системе изображения- координаты точки в изображении,  $v_i^{(k)} \in [0,1]$  – оценка достоверности детектирования,  $K$  – количество ключевых точек модели.

На основе траектории рабочей руки  $\omega_i = (x_i^{(\omega)}, y_i^{(\omega)})$  формируется временной ряд скоростей, нормированный относительно масштаба фигуры игрока. Требуется определить множество моментов ударов:

$$E = \{e_j\}_{j=1}^M, \quad e_i \in \{1, \dots, N\},$$

где  $M$  – количество детектированных событий, а каждый элемент  $e_j$  – это момент времени, соответствующий ударному действию.

Таким образом, задача формулируется как поиск оптимального подмножества индексов кадров  $E \subset \{1, \dots, N\}$ , максимизирующего точность выделения ударных событий при заданных ограничениях на ложные срабатывания и временное разрешение алгоритма.

## 2. Алгоритм детектирования ударов

На основе результатов оценки позы извлекаются координаты ключевой точки «кисть» рабочей руки игрока. Выбор конкретной руки осуществляется на основе предварительного анализа видеозаписи или по признаку максимальной кинематической активности в начальном эпизоде.

Для обеспечения устойчивости траектории используются только те кадры, в которых оценка достоверности детектирования превышает заданный порог  $v_{min}$ ; при кратковременных потерях ключевой точки применяется линейная интерполяция соседних валидных позиций.

Пусть в  $i$ -м кадре положение кисти задаётся вектором  $w_i = (x_i, y_i)$ . Тогда траектория движения кисти представляется в виде временной последовательности:

$$W = \{w_j\}_{j=1}^N,$$

где  $N$  – общее число кадров видеопоследовательности.

Пример исходного кадра видеозаписи приведён на рисунке 2, а результат оценки позы с наложением скелетной модели с использованием MediaPipe Pose – на рисунке 3.

*Оценка скорости движения.* Ударные действия в настольном теннисе характеризуются резким ускорением дистального сегмента кисти в момент контакта с мячом, что проявляется в виде выраженных локальных максимумов скорости на кинематических кривых [5, 6]. Для количественной оценки динамики движения вычисляется мгновенная скорость перемещения кисти между последовательными кадрами. При известной частоте съёмки  $f_{ps}$  временной шаг составляет

$$\Delta t = \frac{1}{f_{ps}}.$$

Скорость движения кисти между соседними кадрами определяется следующим выражением:

$$S_i = \frac{\sqrt{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2}}{\Delta t}.$$



Рис. 2. Исходный кадр видеозаписи без разметки

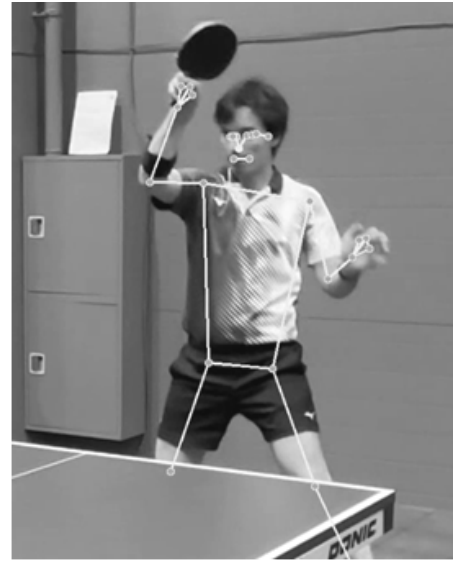


Рис. 3. Пример оценки позы с использованием MediaPipe Pose

Поскольку оценки позы, получаемые нейросетевыми моделями, подвержены высокочастотному шуму, непосредственное использование сырой скорости приводит к ложным срабатываниям и размытию пиков. Для повышения устойчивости сигнала применяется сглаживание скользящим средним [1] с окном ширины  $2k + 1$ :

$$\tilde{S}_i = \frac{1}{2k + 1} \cdot \sum_{j=i-k}^{i+k} S_j,$$

где параметр  $k$  выбирается экспериментально в зависимости от частоты кадров и характерной длительности ударного движения. Сглаженная последовательность  $\{\tilde{S}_i\}_{i=1}^N$  служит основой для дальнейшего масштабирования и детектирования ударных событий.

*Нормализация по масштабу.* Для обеспечения инвариантности алгоритма к изменению расстояния до камеры, фокусного расстояния объектива и разрешения видеозаписи скорость движения нормируется на характерный размер фигуры игрока. В качестве реперного масштаба используется расстояние между плечевыми суставами [2, 5]. Пусть на  $i$ -м кадре координаты левого и правого плеча заданы как

$$l_i = (x_i^{(ls)}, y_i^{(ls)}), r_i = (x_i^{(rs)}, y_i^{(rs)}).$$

Евклидово расстояние между ними вычисляется по формуле:

$$d_i = \|r_i - l_i\|_2 = \sqrt{(x_i^{(ls)} - x_i^{(rs)})^2 + (y_i^{(ls)} - y_i^{(rs)})^2}.$$

Для предотвращения численной неустойчивости нормировка применяется только при условии  $d_i > \varepsilon$ , где  $\varepsilon$  – малая константа (например,  $\varepsilon = 10$  пикселей), отсекающая кадры с ошибочной или потерянной оценкой позы. При условии  $d_i > \varepsilon$  нормированная скорость определяется как:

$$\hat{s}_i = \frac{\tilde{s}_i}{d_i}.$$

*Статистическая оценка порога.* Использование фиксированных порогов скорости не обеспечивает устойчивой работы алгоритма на разнородных видеоданных. Для автоматической адаптации к конкретному видеофрагменту применяется робастный статистический подход на основе медианы и медианного абсолютного отклонения (MAD) [1]. Пусть  $\{\hat{s}_i\}_{i=1}^N$  – последовательность нормированных скоростей. Оцениваются центральная тенденция и разброс:

$$m = \text{median}(\hat{s}_i),$$

$$MAD = \text{median}(|\hat{s}_i - m|).$$

Для приведения  $MAD$  к масштабу стандартного отклонения нормального распределения используется корректирующий множитель 1.4826:

$$\sigma \approx 1,4826 \cdot MAD.$$

Порог детектирования ударов задаётся как:

$$T = m + k_\sigma \cdot \sigma,$$

где  $k_\sigma$  – эмпирический коэффициент, регулирующий чувствительность алгоритма (обычно  $k_\sigma \in [2.5, 4.0]$ ). Данный подход минимизирует влияние выбросов, вызванных кратковременными ошибками трекинга или неигровыми движениями, обеспечивая стабильное выделение значимых пиков скорости без ручной калибровки под каждое видео.

*Детектирование ударных действий.* На основе нормированной последовательности скорости  $\{\hat{s}_i\}_{i=1}^N$  формируется множество моментов ударов  $E = \{e_j\}_{j=1}^M$ . Кадр  $i$  идентифицируется как кандидат на ударное действие при одновременном выполнении трёх условий:

- Локальный экстремум:  $\hat{s}_i > \hat{s}_{i-1} \wedge \hat{s}_i > \hat{s}_{i+1}$ , что соответствует фазе максимального разгона кисти перед контактом с мячом.

- Превышение адаптивного порога:  $\hat{s}_i \geq T$ , это условие отсекает фоновые и подготовительные движения.

- Достоверность оценки позы:  $v_i^{(wrist)} \geq v_{\min}$ , гарантирующее, что детекция основана на надёжно локализованной ключевой точке.

Для исключения кратных срабатываний в пределах одного физического удара применяется временной фильтр непериодического подавления. Между соседними детекциями задаётся минимальный интервал  $\Delta_{\min}$ , определяемый физиологически возможной длительностью цикла ударного движения:

$$|e_j - e_{j-1}| \geq \Delta_{\min},$$

где  $\Delta_{\min}$  выбирается эмпирически в зависимости от частоты съёмки (например,  $\Delta_{\min} \approx 0,25 \div 0,35 f_{fps}$ ).

Итоговое множество моментов ударов формируется последовательным сканированием временного ряда скоростей с применением указанных логических и временных фильтров.

### 3. Программная реализация и результаты экспериментов

Прототип программы разработан на языке Python с использованием библиотеки OpenCV для обработки видеопоследовательностей [7]. Для оценки позы игрока применялась модель MediaPipe Pose, обеспечивающая детектирование ключевых точек тела в реальном времени [3, 4]. Численные расчёты и агрегация данных осуществляются средствами библиотек NumPy и Pandas. Результатом работы программы является:

- видео с наложенной скелетной моделью игрока;
- таблица обнаруженных ударных событий.

Алгоритм реализован в виде консольного приложения, которое принимает на вход видеофайл тренировки. На каждом кадре извлекаются координаты ключевых точек, вычисляется нормированная скорость кисти рабочей руки, после чего ударные события детектируются по локальным максимумам скорости с применением робастных статистик, что согласуется с биомеханическими подходами к анализу движений [2, 5]. Параметры алгоритма (коэффициент сглаживания, пороговые значения, минимальный интервал между ударами) вынесены в конфигурационный файл для обеспечения воспроизводимости экспериментов.

Экспериментальная проверка проводилась на видеозаписи тренировки по настольному теннису длительностью 136,49 с и разрешением 1920×1080 пикселей. Всего видео содержало 4044 кадра. В процессе обработки модель оценки позы успешно определила ключевые точки тела игрока на 3978 кадрах, что соответствует 98,37% от общего числа кадров. График нормированной скорости кисти и отмеченных ударных событий приведен на рисунке 4.

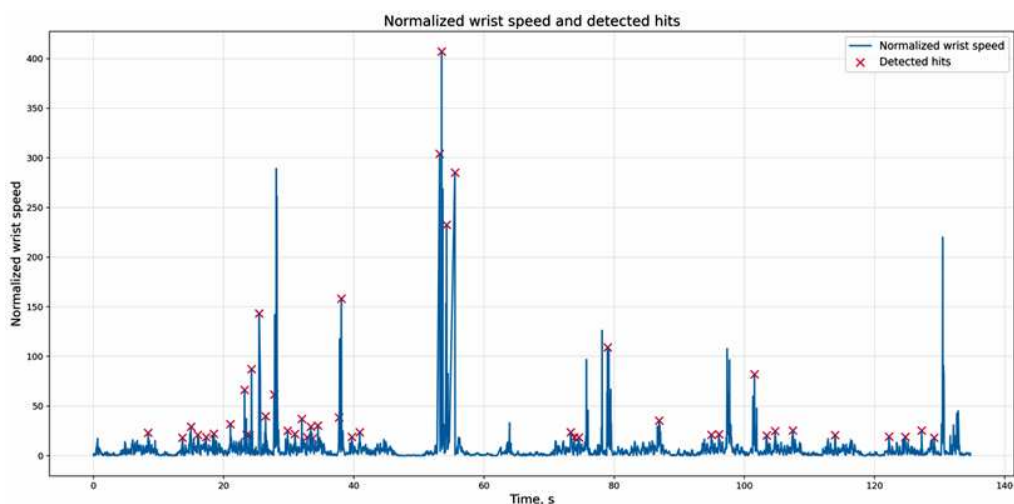


Рис. 4. График нормированной скорости кисти и отмеченных ударов

В результате работы алгоритма было автоматически выделено 46 кандидатов на ударные действия. При ручном просмотре видеозаписи установлено, что фактическое количество ударов справа без учета подач составило 47. Таким образом, предложенный метод продемонстрировал высокую чувствительность к выявлению моментов ударных действий на рассматриваемом видео фрагменте и может применяться для автоматизированной предварительной разметки тренировочных записей.

Дополнительно был выполнен анализ распределения автоматически присвоенных классов ударов. Наибольшее число детекций пришлось на класс *forehand topspin*, тогда как остальные классы были представлены заметно реже. Основные результаты эксперимента представлены на рисунке 5, а распределение автоматически обнаруженных классов ударов на рисунке 6.

| event         | frame | time_sec           | wrist_speed_norm   | stroke_side | stroke_type | stroke_label |
|---------------|-------|--------------------|--------------------|-------------|-------------|--------------|
| hit_candidate | 626   | 20.88381265406738  | 67.87259879448754  | bh          | underspin   | bh_underspin |
| hit_candidate | 656   | 21.884634346754314 | 74.0026986743826   | bh          | underspin   | bh_underspin |
| hit_candidate | 779   | 25.98800328677075  | 258.3704120777315  | fh          | underspin   | fh_underspin |
| hit_candidate | 896   | 29.891207888249795 | 58.26042857264242  | bh          | topspin     | bh_topspin   |
| hit_candidate | 963   | 32.12637633525062  | 122.50298409035494 | bh          | underspin   | bh_underspin |
| hit_candidate | 998   | 33.29400164338537  | 140.5479337723269  | bh          | underspin   | bh_underspin |
| hit_candidate | 1032  | 34.428266228430566 | 82.14868762587272  | bh          | underspin   | bh_underspin |
| hit_candidate | 1059  | 35.32900575184881  | 87.33714487101517  | fh          | underspin   | fh_underspin |
| hit_candidate | 1091  | 36.39654889071488  | 58.84976944868788  | bh          | underspin   | bh_underspin |

Рис. 5. Основные результаты эксперимента

Из скриншота, представленного на рисунке 5, следует, что все записи имеют статус *hit\_candidate* – это означает, что алгоритм поместил эти моменты как потенциальные удары на основе превышения порога скорости. Поля *frame* и *time\_sec* показывают хронологию, и, судя по соотношениям (например, кадр  $626 \approx 20,88$  с), частота видео составляет примерно 30 FPS. Анализ классификации ударов (*stroke\_side*) – Backhand (bh): 6 ударов (66%) и Forehand (fh): 3 удара (33%), позволяет сделать вывод, что в данном фрагменте доминирует игра слева. Это может быть специфика упражнения (например, игра только по диагонали слева) или то, что камера лучше видит левую руку игрока.

Из анализа вращения (*stroke\_type*) – Underspin (нижнее вращение): 8 ударов и Topspin (верхнее вращение): 1 удар, следует доминирование нижнего вращения. В реальном теннисе это обычно означает «накат» или «срезку». Если это была активная атака, то классификатор мог ошибочно отнести быстрые плоские удары к *underspin* из-за траектории движения кисти (сверху вниз или прямо). Это указывает на то, что модуль классификации надо рассматривать как предварительную эвристическую надстройку над основным детектором ударов.

Анализ распределения автоматически обнаруженных классов ударов показал доминирование категории *forehand topspin* (30 детекций, 65% от общего числа), что свидетельствует о преобладании атакующих действий правой рукой в рассматриваемом тренировочном эпизоде. Суммарно удары

справа (forehand) составили 80% от всех зарегистрированных событий, тогда как удары слева (backhand) – лишь 20%. По типу вращения топспины превалируют над нижним вращением в соотношении 34 к 12, что характеризует данный фрагмент тренировки как атакующий с высокой интенсивностью ударных действий.

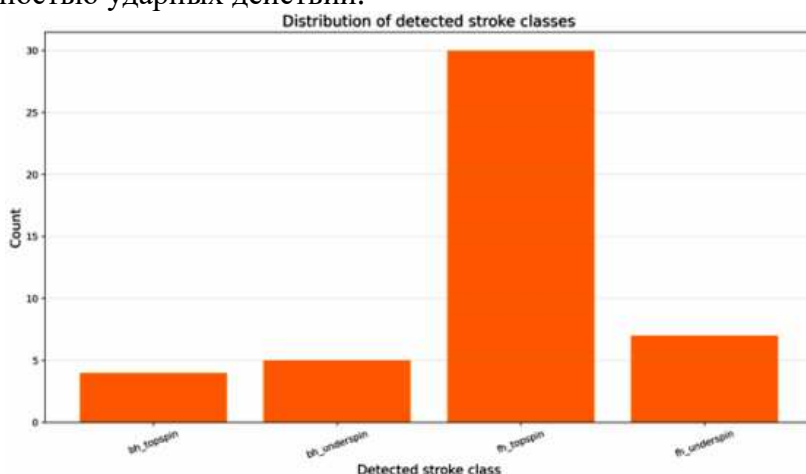


Рис. 6. Диаграмма распределения автоматически обнаруженных классов ударов

Выявленный дисбаланс в распределении классов может быть обусловлен как объективными факторами (специфика тренировочного упражнения, технический арсенал игрока, ракурс съёмки), так и ограничениями применяемого алгоритма классификации. В частности, недостаточное количество детекций backhand ударов может указывать на необходимость совершенствования метода распознавания стороны удара, возможно, за счёт учёта дополнительных кинематических признаков или применения мультимодального анализа.

Полученные результаты подтверждают работоспособность предложенного подхода к автоматическому детектированию и классификации ударов, однако демонстрируют потребность в расширении обучающей выборки и проведении валидации с привлечением экспертной разметки для повышения точности распознавания менее представленных классов ударов. Установлено, что задача детектирования факта удара решается заметно устойчивее, чем задача точной классификации конкретного типа удара. Это объясняется тем, что пик скорости руки является достаточно устойчивым признаком, в то время как разделение типов ударов чувствительно к ракурсу съёмки, качеству оценки позы и индивидуальным особенностям техники игрока. Различные типы ударов могут иметь сходные профили скорости, но различаться по пространственным характеристикам движения [5, 6].

### Заключение

В статье предложен метод автоматического детектирования ударов в настольном теннисе на основе анализа видеопоследовательностей и оценки позы игрока. Метод основан на выделении ключевых точек тела, вычислении

нормированной скорости рабочей руки и детектировании локальных максимумов скорости. Разработанный программный прототип позволяет формировать видео с наложенной позой игрока и таблицу обнаруженных ударов.

На экспериментальной видеозаписи длительностью 136,49 с алгоритм обеспечил извлечение позы на 98,37% кадров и обнаружил 46 ударных событий при 47 фактически наблюдаемых ударах справа без учета подач. Полученные результаты показывают, что предложенный подход может использоваться как основа для автоматизированной разметки тренировочных видеозаписей и дальнейшего построения интеллектуальных систем анализа техники игры в настольном теннисе.

### Список литературы / References

1. Szeliski R. Computer Vision: Algorithms and Applications. – 2nd ed. – Springer, 2022. – 925 с.
2. Winter D.A. Biomechanics and Motor Control of Human Movement. – 4th ed. – Wiley, 2009. – 384 p.
3. Bazarevsky V., Grishchenko I., Raveendran K., Zhu T., Zhang F., Grundmann M. BlazePose: On-device Real-time Body Pose Tracking // arXiv. – 2020. – URL: <https://arxiv.org/abs/2006.10204>.
4. Grishchenko I., Bazarevsky V., Zanzir A., Bazavan E.G., Zanzir M., Yee R., Raveendran K., Zhdanovich M., Grundmann M., Sminchisescu C. BlazePose GHUM Holistic: Real-time 3D Human Landmarks and Pose Estimation // arXiv. – 2022. – URL: <https://arxiv.org/abs/2206.11678>.
5. Bartlett R. Introduction to Sports Biomechanics: Analysing Human Movement Patterns. – 2nd ed. – Routledge, 2007. – 304 p.
6. Wong, D. W.-C., Lee, W. C.-C., Lam, W.-K. Biomechanics of Table Tennis: A Systematic Scoping Review of Playing Levels and Maneuvers // Applied Sciences. 2020, vol. 10, no. 15, p 5203. – doi.org/10.3390/app10155203.
7. Howse J., Minichino J. Learning OpenCV 4 Computer Vision with Python 3. – 3rd ed. – Packt Publishing, 2020. – 372 p.

|                                                                                            |                                                                                         |
|--------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------|
| <b>Примаченко Павел Николаевич</b> – магистрант                                            | <b>Primachenko Pavel Nikolaevich</b> – master's student                                 |
| <b>Хабаров Сергей Петрович</b> – кандидат технических наук, доцент<br>serg.Habarov@mail.ru | <b>Khabarov Sergey Petrovich</b> – candidate of technical sciences, associate professor |

*Received 16.04.2026*