

Подход, основанный на глубоком обучении с подкреплением, для создания поведения человека при ходьбе с помощью нервно-мышечной модели.

Ахил С Ананд¹, Гопин Чжао², Юбер Рот³ и Андре Зейфарт⁴

Аннотация. Модель походки, способная генерировать человеческое поведение при ходьбе как на кинематическом, так и на мышечном уровне, может быть очень полезной основой для разработки схем управления гуманоидами и носимыми роботами, такими как экзоскелеты и протезы. В этой работе мы продемонстрировали возможность использования подхода, основанного на глубоком обучении с подкреплением, для моделирования нервно-мышечной походки. Модель походки нижних конечностей, состоящая из семи сегментов, четырнадцати степеней свободы и двадцати двух мышц типа Хилла, была построена для отражения динамики ног человека и характеристик свойств мышц. Мы реализовали алгоритм оптимизации проксимальной политики, чтобы изучить сенсомоторные сопоставления (политику управления) и сгенерировать для модели поведение при ходьбе, подобное человеческому. Данные о захвате движений человека, моделях активации мышц и оценке метаболических затрат были включены в функцию вознаграждения за тренировку. Результаты показывают, что модель может точно воспроизводить кинематику человека и силы реакции опоры во время ходьбы. Он способен генерировать поведение человека при ходьбе в диапазоне скоростей от 0,6 м/с до 1,2 м/с. Он также способен выдерживать неожиданные изменения крутящего момента бедра во время ходьбы. Мы далее исследовали преимущества использования нейромышечной модели по сравнению с идеальной моделью, основанной на крутящем моменте сустава. Мы заметили, что нервно-мышечная модель более эффективна по сравнению с моделью крутящего момента.

ВВЕДЕНИЕ

Модель походки, способная воспроизводить движения человека, может помочь нам лучше понять схему управления движением человека, которую можно использовать для разработки двуногих роботов и носимых роботов (например, экзоскелетов, протезов и т. д.). Например, с помощью простой модели перевернутого маятника было показано, что походка на двух ногах, подобная человеческой, может быть достигнута пассивно (без активного контроля) из-за естественной динамики человеческого тела [1]. Также было обнаружено, что как ходьба, так и беговая походка человека могут быть описаны с помощью простой модели подпружиненного перевернутого маятника [2], [3]. Было разработано несколько роботов с ногами, которые успешно продемонстрировали преимущества управления этими моделями [1], [4], [5]. Кроме того, биоинспирированная концептуальная модель также демонстрирует преимущества для управления экзоскелетом [6]. Однако, хотя эти упрощенные шаблонные модели могут генерировать походку, подобную человеческой (например, с точки зрения центра

массовое движение, длину/частоту шага и т. д.), их способность воспроизводить богатое двигательное поведение, подобное человеческому (например, подъем по лестнице/склону, ускорение/замедление и т. д.), очень ограничено.

Недавно Сонг и Гейер [7] продемонстрировали, что разнообразные модели передвижения человека могут быть созданы с помощью сложной нервно-мышечной модели походки с использованием нейронной схемы, которая подчеркивает мышечные рефлексy. Они также продемонстрировали, что модель может вызывать немедленные изменения в мышечной активации некоторых групп мышц, подобные человеческим [8]. Однако архитектура нейронных цепей, использованных в модели, была создана вручную. Производительность модели можно было бы еще улучшить, если бы мы изучили схемы (соединения рефлекторных путей) с помощью систематического подхода. В [9] авторы разработали скелетно-мышечную модель нижних конечностей, основанную на контактно-инвариантной оптимизации, в первую очередь для анимации деятельности человека, управляемой

нижней частью тела. [10]–[12] представили различные подходы к мышечным контроллерам локомоции, в то время как [13], [14] представили подход, основанный на глубоком обучении с подкреплением путем изучения паттерна мышечной активации и паттерна крутящего момента суставов соответственно. Пэн и др. [15], [16] и Мерел и др. [10] показали, что глубокое обучение с подкреплением (deep-RL) является очень полезным подходом при разработке надежных контроллеров для сложных локомоторных систем. Они также продемонстрировали возможности глубокого RL в изучении широкого спектра сложных навыков передвижения с использованием кинематических данных. Еще одна близко соответствующая работа Peng et al. [17] продемонстрировала обучение 2D-движению на двух ногах, активируемому мышцами, с использованием глубокого RL. Они определили, что локальная обратная связь, обеспечиваемая параметризацией действий высокого уровня, может существенно повлиять на обучение, надежность и качество движения получаемых политик. Наша работа направлена на изучение индивидуального поведения при ходьбе при различных скоростях ходьбы путем непосредственного изучения

RL добился больших успехов в задачах дискретной области, таких как компьютерные игры [18], но передвижение человека необходимо решать как задачу непрерывного управления. RL имеет преимущество при решении задач с непрерывной областью с момента появления последних разработок в методах, основанных на градиенте политики [19]–[22]. Политические градиенты стали прорывом в непрерывной области, но все еще ограничивались многими факторами, такими как скорость обучения, эффективность выборки и т. д. Многие подходы пытались устранить эти недостатки, что привело к разработке таких алгоритмов, как TRPO [20], ACER [23], PPO [19] и т. д. Все эти методы имеют свои недостатки. ACER намного сложнее, чем PPO, требуя добавления кода для корректировок, не отвечающих политике, и буфера воспроизведения с очень

¹Ахил С. Ананд из отд. кафедры инженерной кибернетики в Норвежском техническом-натуровитенскаплиге университете - NTNU, 7034, Тронхейм, Норвегия. akhil.s.anand@ntnu.no

²Гопин Чжао работает в Лаборатории локомоции Lauflabor в Техническом университете Дармштадта, 64289, Дармштадт, Германия. zhao@sport.tu-darmstadt.de ³Хуберт Рот работает в Институте техники управления Университета Зигена, 57076, Зиген, Германия. Hubert.roth@uni-siegen.de ⁴Андр Зейфарт работает в Лаборатории локомоции Lauflabor в Техническом университете Дармштадта, 64289, Дармштадт, Германия. seyfarth@sport.tu-darmstadt.de

незначительное преимущество в результатах, протестированных на тесте Atari компанией Open-AI [24]. Принимая во внимание вышеупомянутые компромиссы, мы решили использовать алгоритм оптимизации проксимальной политики (PPO) [19] для решения нашей проблемы, связанной с передвижением человека.

Целью данной статьи является исследование возможности использования глубокого RL для создания человеческой ходьбы с помощью нейромышечной модели походки. Здесь мы представляем основанный на глубоком RL подход к разработке модели человеческой походки, способной создавать индивидуальную трехмерную походку на кинематическом, кинетическом и мышечном уровнях. Хотя основное внимание в нашей работе уделяется разработке походки человека на основе глубокого RL, мы также изучаем преимущества обучения мышечному контролю над контролем на основе крутящего момента с точки зрения эффективности выборки.

Подход, использованный в этой статье, заключается в следующем: (i) проведение экспериментов на людях для сбора индивидуальных кинематических и кинетических данных и предоставление набора данных для глубокого RL, (ii) создание модели скелетно-мышечной походки для выполнения глубокого RL, (iii) проведение deep-RL для создания кинематики, подобной человеческой, и оптимизации энергии, (iv) тестирования модели на надежность и (v) сравнения выборочной эффективности управления на основе мышц и крутящего момента.

II. МЕТОДЫ

A. Эксперименты на людях

Эксперименты по ходьбе на беговой дорожке на людях проводились с одним субъектом (мужчина, 27 лет, рост 1,75 м, вес 66 кг) для получения данных о кинематике суставов нижних конечностей и силах реакции опоры (GRF) в диапазоне скоростей ходьбы от 0,6 м/с до 1,2 м/с. Субъект дал информированное согласие на эксперимент. Дизайн и протокол исследования были одобрены этическим комитетом Технического университета Дармштадта.

Экспериментальные данные были обработаны, чтобы подготовить набор данных для обучения с подкреплением. Всего набор данных содержит 1200 шагов ходьбы, каждый из которых производится с частотой 200 Гц. Один шаг в наборе данных содержит всю кинематику отдельных суставов нижней части тела от приземления одной стопы до приземления противоположной стопы.

B. Моделирование

Скелетно-мышечная модель, используемая в этом исследовании, представляет собой модель нижней конечности человека с семью сегментами и двадцатью двумя мышцами. Модель была принята из [7] и реализована в MuJoCo [25] для достижения высокой скорости моделирования. Затем модель была интегрирована с Open-AI Gym, чтобы облегчить внедрение глубокого RL. Модель имеет рост 1,8 м, вес 66 кг и четырнадцать степеней свободы (шесть глобальных степеней свободы туловища и восемь внутренних шарнирных степеней свободы). Модель на 5 см выше объекта, но мы считаем, что влияние столь небольшой разницы в высоте на кинематику сустава незначительно. Физические свойства, мышечно-сухожильные единицы (MTU) и свойства мышц аналогичны модели из [7], за исключением стопы. Стопа моделируется в виде кубоида (ширина 10 см, длина 25 см и высота 6 см) с четырьмя точками контакта с землей. Одиннадцать групп мышц каждой ноги показаны на рис. 1(а). Модель с контролем крутящего момента

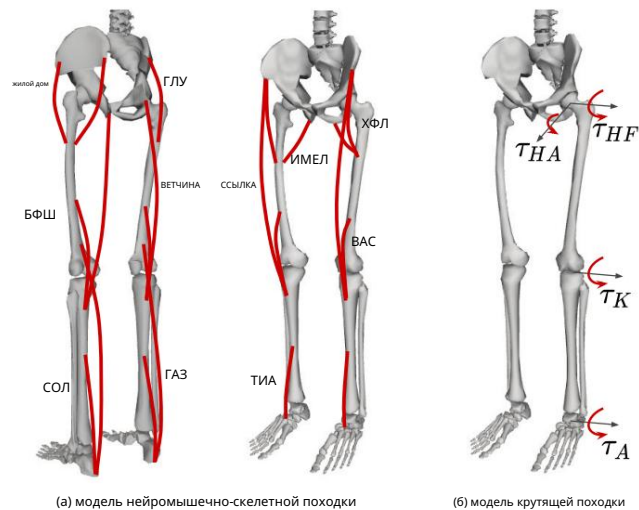


Рис. 1. Схема скелетно-мышечной модели и модели крутящего момента. (а) Модель скелетно-мышечной системы нижних конечностей с одиннадцатью группами мышц на ногу. Группы мышц включают отводящие бедра (HAB), приводящие бедра (HAD), сгибатели бедра (HFL), ягодичные мышцы (GLU), подколенные сухожилия (HAM), прямую мышцу бедра (REF), широкую мышцу бедра (VAS), короткую головку двуглавой мышцы бедра (BFSH), икроножная мышца (GAS), камбаловидная мышца (SOL) и большеберцовая мышца (TIA). HAM, REF и GAS — двусуставные мышцы. HAB, HAD, HFL, GLU, VAS, BFSH и TIA являются моносуставными мышцами. (б) Двухзвеновая модель с контролем крутящего момента нижних конечностей и 8 моментными приводами. Для каждой ноги имеется 4 привода крутящего момента суставов: (i) сгибание/разгибание бедра τ_{HF} , (ii) приведение/отведение бедра τ_{HA} , (iii) сгибание/разгибание колена τ_K и (iv) тыльное сгибание/разгибание голеностопного сустава τ_A . Значения крутящего момента могут быть как положительными, так и отрицательными.

также имеет такой же вес, размеры сегментов и степени свободы. Модель приводится в действие восемью идеальными моментными приводами (по четыре на каждую ногу), как показано на рис. 1 (б).

C. Реализация Deep-RL

В окончательной реализации с алгоритмом PPO отдельные сетевые архитектуры и гиперпараметры выбираются как для модели, основанной на мышцах, так и для модели, основанной на крутящем моменте, поскольку они различаются размером и характеристиками пространства состояния-действия. Пространство входных состояний для модели крутящего момента содержит положения суставов θ , угловые скорости суставов $\dot{\theta}$, GRFs g_{rf} и целевую скорость ходьбы v . Модель на основе мышц имеет дополнительные входные состояния, которыми являются мышечная сила f , длина мышцы l , скорость мышц BM и мышечная активация a . Входными данными для мышечной модели является мышечная стимуляция u . В случае модели крутящего момента входными данными являются крутящие моменты τ на 8 суставах.

Архитектура как политики, так и функции ценности определяется с использованием нейронных сетей, показанных на рис. 2. Все гиперпараметры процесса обучения показаны в таблице I. Алгоритм реализован с помощью 40 потоков (воркеров), которые собирают данные, воздействуя на окружающую среду, и обучают нейронную сеть на одном графическом процессоре. Псевдокод реализации показан в алгоритме 1. Вектор входного состояния и скалярные значения вознаграждения нормализуются с использованием его текущего среднего значения. И стандартное отклонение состояний ограничивается диапазоном $[-10, 10]$.

В алгоритме 1 θ и ϕ являются параметрами политики и функции стоимости (базовой линии). N — общее количество временных шагов,

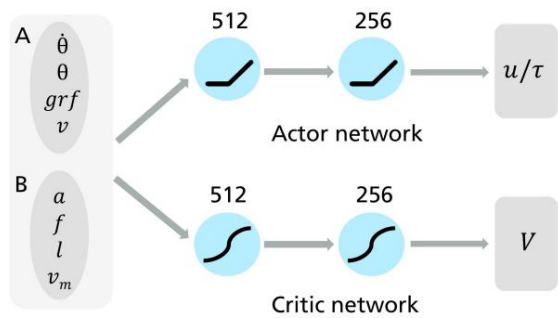


Рис. 2. Сеть состоит из двух полносвязных скрытых слоев размером 512 и 256 соответственно. Актор (политик) и критик имеют одну и ту же сетевую структуру. Функция активации для сетей актеров и критиков — это ReLU и tanh соответственно как для моделей мышц, так и для моделей крутящего момента. Входной набор A является входными данными для модели крутящего момента, а A, B вместе являются входными данными для модели мышц. u и τ обозначают мышечные активации и крутящие моменты суставов для моделей мышц и крутящих моментов соответственно. V обозначает государственную ценность от критика

ТАБЛИЦА I

ГИПЕРПАРАМЕТРЫ, ИСПОЛЬЗУЕМЫЕ ДЛЯ МЫШЕЧНЫХ МОДЕЛЕЙ И МОДЕЛЕЙ КРУТЯЩЕГО МОМЕНТА

Гиперпараметр	Модель мышц	Модель крутящего момента
Кол-во актеров N	40	40
Выборка на актера/эпизод n	128	128
мини-пакетов Число эпох	32	32
Коэффициент клипирования	0,2	0,2
β Параметр GAE λ	0,95	0,95
Коэффициент дисконтирования γ	0,99	0,99
Коэффициент функции ценности c1	0,5	0,5
Коэффициент энтропии c2	0 0,005 3e-5	0 0,005 5e-4
Скорость обучения lr	1e-7	1e-6

I и J — это количество подитераций с обновлениями политики и базовой линии (здесь функция значения) для пакета точек данных. T обозначает количество точек данных, собранных на одного работника. λ и α — коэффициент регуляризации KL и масштабирующий член соответственно.

D. Мышечная динамика и протоколы возмущений

Чтобы генерировать паттерны активации мышц, подобные человеческим, модель обучается с использованием экспериментальных данных на людях из [26]. Управление осуществляется путем ограничения стимуляции на соответствующих фазах цикла походки. Например, во второй половине фазы опоры мышцы HAD и HAM совершенно неактивны. Это свидетельство реализуется путем ограничения активаций HAD и HAM до 0 на этом этапе. Частота мышечного контроля составляет 5 кГц, а все сигналы сенсорной обратной связи (кинематика, мышечная динамика и силы реакции опоры), а также входные сигналы стимуляции для модели мышц задерживаются на 15 мс, чтобы имитировать задержку сенсорной обратной связи человека [7].

Модель тренируется с возмущениями крутящего момента суставов при движениях сгибания/разгибания бедра. На этапе тренировки к модели непрерывно в течение 50 мс (начиная со случайно выбранного временного шага в цикле походки) применяются случайные возмущения крутящего момента суставов с величиной в диапазоне от -5 Нм до 5 Нм. Максимум только одно возмущение

Алгоритм 1. Псевдокод для реализации PPO θ случайные веса

```
φ случайные веса для n
{1, . . . , N} do pw πθ
Запустить W рабочих
параллельно
pold πθ для i {1, . . . , I} do JP
P O(θ) =
Обновление θ методом
стохастического T πθ(v|st)
градиента относительно pold(at|st)A^ t λKL[pold|πθ]
JP P O(θ)
end for
для j {1, . . . , J} do LV (φ) =
t > t γ T T - T
rt Vφ (ст)
Обновите φ градиентным методом относительно конца
LV (φ) для
если KL [pold|πθ] > βhighKLtarget , то
λ αλ
иначе, если KL [pold|πθ] < βlowKLtarget , то λ λ/α
end if end
for
```

Рабочий

```
для t {1, . . . , T} do
Запускайте политику πθ , собирая {st, at, rt} r
Оценочная доходность Rt τ = τ γ T - T (st , at )
= Оценочная выгода A^ τ = R^ t Vφ (ст)
Сохраните информацию о траектории
в течение
```

применяется за цикл ходьбы в любом из тазобедренных суставов. Вероятность применения возмущения в любом цикле походки составляет 50%. Для проверки устойчивости к бедру непрерывно в течение 50 мс прикладываются гораздо большие возмущения крутящего момента сустава в диапазоне [-200, 200] Нм. Возмущения крутящего момента суставов выбраны так, чтобы имитировать ситуацию использования экзоскостюмов для облегчения/препятствования ходьбы.

Инициализация случайного состояния (RSI) и раннее завершение (ET) являются очень полезными методами обучения с подкреплением [15]. Мы разделили каждую траекторию в наборе обучающих данных на 10 равных временных интервалов. RSI реализуется путем случайного выбора траекторий во время обучения из эталонного кинематического набора данных и определения начального состояния путем случайного выбора из 10 равных интервалов траектории. Идея ЭТ реализуется путем прекращения эпизода, если кинематическая ошибка превышает заданный предел. Более конкретно, ЭТ прекращается, если вертикальное положение таза ниже 0,8 м или выше 1,4 м, что соответствует нежелательным движениям падения и прыжка соответственно.

Е. Формирование вознаграждения

Функция вознаграждения для модели мышц и крутящего момента содержит члены, которые поощряют имитацию кинематической траектории, непрерывную стабильную ходьбу и достижение целевой скорости.

В мышечную модель включен дополнительный термин вознаграждения за метаболические затраты, но модель крутящего момента изучается без использования какого-либо термина минимизации крутящего момента.

$$r = wlrl + wkrk + wmrn + wvrn \quad (1)$$

где r — награда, rl — бонус жизни, rk — бонус кинематического поведения, rn — метаболический бонус и rv — бонус целевой скорости. $wl=1$, $wk=4$, $wn=4$, $wv=1$ — веса rl , rk , rn и rv соответственно. Все эти индивидуальные бонусы имеют значения от 0 до 1. Общая награда r находится в диапазоне от 0 до 10. Бонус жизни rl обозначает награду за ходьбу без падения. Состояние падения возникает, когда вертикальное положение таза выходит за пределы $[0,8, 1,4]$ м. Член rk определяет вознаграждение за имитацию желаемой траектории. Индивидуальные ошибки положения и скорости между моделью и экспериментальными данными рассчитываются для каждого шага выборки. Этими ошибками являются: • Ошибка вектора положения стопы e_{rl} , которая обозначает квадрат

разницы между вектором положения стопы модели и эталонными данными траектории движения человека. $e_{rl} = [c(s_{rl}(t) - s_{rl}^{*}(t))]^2$ Здесь $s_{rl}(t)$ и $s_{rl}^{*}(t)$ — вектор

положения стопы модели и $s_{rl}^{*}(t)$ — эталонные данные соответственно в момент времени t и коэффициент масштабирования, $c=30$.

- Ошибка положения COM таза e_{rn} , которая обозначает квадрат разницы между вектором положения COM таза модели и справочными данными.

$$e_{rn} = [c(s_{rn}(t) - s_{rn}^{*}(t))]^2 \quad (3)$$

Здесь $s_{rn}(t)$ и $s_{rn}^{*}(t)$ — вектор положения COM таза модели и справочные данные соответственно в момент времени t и $c=20$.

- Ошибка скорости Pelvis COM e_{rv} , которая обозначает квадрат разницы между вектором скорости Pelvis COM модели и справочными данными.

$$e_{rv} = [c(s_{rv}(t) - s_{rv}^{*}(t))]^2 \quad (4)$$

Здесь $s_{rv}(t)$ и $s_{rv}^{*}(t)$ — вектор скорости таза COM модели и справочных данных соответственно в момент времени t и $c=2$.

- Ошибка углового положения шарнира e_{ar} , которая обозначает квадрат разницы между всеми углами шарнира модели и справочными данными.

$$e_{ar} = [c(\theta_{ar}(t) - \theta_{ar}^{*}(t))]^2 \quad (5)$$

Здесь $\theta_{ar}(t)$ и $\theta_{ar}^{*}(t)$ — это массив всех совместных углов модели и справочных данных соответственно в момент времени t и $c=12$.

- Ошибка совместной угловой скорости e_{av} , которая обозначает квадрат разницы между всеми общими угловыми скоростями модели и справочными данными.

$$e_{av} = [c(\dot{\theta}_{av}(t) - \dot{\theta}_{av}^{*}(t))]^2 \quad (6)$$

$\dot{\theta}_{av}(t)$ и $\dot{\theta}_{av}^{*}(t)$ — массив всех совместных угловых скоростей модели и справочных данных соответственно в момент времени t и $c=0,1$.

Все эти отдельные ошибки объединяются в один вектор ошибок E следующим образом:

$$E = [e_{rl}, e_{rk}, e_{rn}, e_{ar}, e_{av}] \quad (7)$$

E преобразуется в отрицательную экспоненту, а полученные члены суммируются, чтобы получить скалярное значение T :

$$T = e^{-E} \quad (8)$$

Термин rk обозначает, насколько велико значение T по сравнению с предельным значением 28. Оно вычисляется следующим образом:

$$rk = \frac{T - T_{limit}}{T_{max} - T_{limit}} \quad \text{где } T_{limit} = 28, T_{max} = 35 \quad (9)$$

Значение rk находится в диапазоне от 0 до 1, где 1 обозначает точную имитацию совместной траектории, а 0 соответствует максимально допустимому отклонению, определяемому T_{limit} . T_{limit} также используется в качестве критерия досрочного прекращения действия (ET). Другими словами, ET сработает, если $T < T_{limit}$.

Скорость метаболизма r для скелетно-мышечной модели оценивается на основе состояний мышц согласно работе Александера [27]. Метаболическая энергия на этапе отбора проб преобразуется в значение от 0 до 1 путем принятия отрицательной экспоненты с коэффициентом $1/30$. Значение $1/30$ выбрано путем наблюдения за диапазоном r во время обучения. Расчет rm следующий:

$$rm = e^{-r/30} \quad (10)$$

Член rv является функцией разницы между средним значением экспериментальной скорости ходьбы v_r и средним значением модельной скорости ходьбы v_p . ($e_{rv} = (v_r - v_p)^2$)

$$rv = \frac{e_{rv}}{e_{rv} + 1} \quad \text{где } e_{rv} = (v_r - v_p)^2 \quad (11)$$

Коэффициент $c = 50$.

Реализованный алгоритм можно найти здесь. ¹

III. ПОЛУЧЕННЫЕ РЕЗУЛЬТАТЫ

A. Воспроизведение походки человека

Обученная модель скелетно-мышечной походки способна генерировать кинематику суставов, подобную человеческой, активацию мышц и GRF в диапазоне скоростей от 0,6 м/с до 1,2 м/с. Он может поддерживать заранее заданную целевую скорость ходьбы с погрешностью 0,1 м/с, когда он инициализируется желаемой скоростью. Кинематическое поведение, генерируемое мышечной моделью при скорости ходьбы 1,2 м/с, по сравнению с экспериментальными данными показано на рис. 3. Сравнение показывает тесную корреляцию (значения корреляции R от 0,82 до 0,98) между моделью и экспериментальными данными. данные для всех углов суставов нижних конечностей. По сравнению с углами суставов, характеры угловых скоростей суставов модели менее похожи на экспериментальные данные (значения R находятся между 0,46 и 0,92). Модель способна воспроизводить асимметричные характеристики данных человека (например, угол лобно-бедренного сустава). 1 Кинематические результаты можно просмотреть на видео

¹http://lauflabor.ifs-tud.de/doku.php?id=проекты:projects_learnwalk

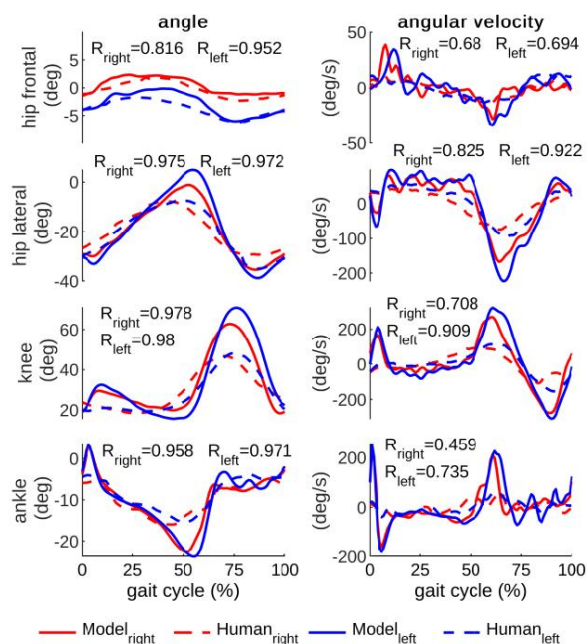


Рис. 3. Угол суставов и угловая скорость модели скелетно-мышечной походки (сплошные линии) и экспериментальных данных человека (пунктирные линии) при скорости ходьбы 1,2 м/с. Данные представляют собой среднее значение 100 шагов устойчивой ходьбы в течение цикла походки (от касания до касания). Красным и синим цветом обозначены данные суставов правой и левой ноги соответственно. Rright и Rleft обозначают значения взаимной корреляции (R) для суставов правой и левой ноги соответственно.

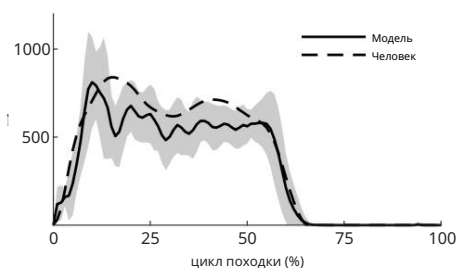


Рис. 4. Вертикальная GRF модели скелетно-мышечной походки (сплошная линия) и экспериментальных данных человека (пунктирная линия) во время ходьбы. Среднее (более 100 шагов ходьбы) GRF для скорости ходьбы 1,2 м/с показано для модели скелетно-мышечной походки. Полоса ошибок обозначает ± 1 стандартное отклонение.

GRF, генерируемые моделью походки, не такие плавные, как в экспериментальных данных на людях. Среднее значение (более 100 шагов ходьбы) вертикальной GRF из модели походки (модель на основе мышц) имеет высокую корреляцию с экспериментальными данными со значением R 0,98 при скорости ходьбы 1,2 м/с (показано на рис. 4). Модель походки обучена генерировать мышечные активации, подобные человеческим, путем ограничения стимуляции и оптимизации с минимальными метаболическими затратами. Паттерны мышечной активации, генерируемые моделью походки, показаны на рис. 5. Общие паттерны мышечной активации аналогичны экспериментальным данным, найденным в литературе [26].

Модель скелетно-мышечной походки обеспечивает уверенную ходьбу. Он может восстановиться после возмущения до 200 Н·м на бедре. Реакция на возмущение, демонстрируемая моделью походки

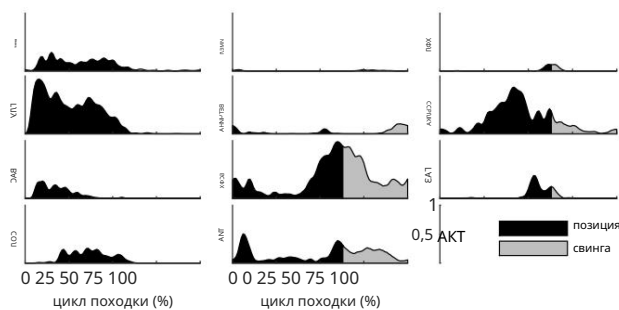


Рис. 5. Мышечная активация скелетно-мышечной модели походки при скорости ходьбы 1,2 м/с. Фаза опоры и фаза поворота обозначены черно-серой зоной.

сильно зависит от времени возникновения возмущения. Для условий приземления и взлета реакция на возмущение является случайной, поскольку она не демонстрирует ожидаемого поведения модели, которая становится все более нестабильной с увеличением величины возмущения. Например, производительность падает, когда к поворотной ноге прикладывают высокие крутящие моменты.

В. Мышечный контроль против контроля крутящего момента

После обучения модель походки на основе крутящего момента также способна имитировать кинематику суставов человека (результат примера показан на рис. 6). По сравнению с моделью, основанной на крутящем моменте, модель, основанная на мышцах, демонстрирует немного лучшую производительность при воспроизведении кинематических данных человека. Например, средние значения R (слева и справа) для модели на основе мышц выше, чем для модели на основе крутящего момента для приведения/отведения бедра, сгибания бедра, сгибания колена и угла голеностопного сустава при скорости ходьбы 1,2 м/с.

Ход обучения изображен на кривых обучения на рис. 7. Мышечная модель достигает средней доходности 1688 после 10 миллионов временных шагов. Но модель крутящего момента смогла достичь только среднего возврата 1497 после 27,65 миллионов временных шагов. Хотя одинаковые средние значения доходности для обеих моделей не соответствуют одному и тому же поведению, доходность сопоставима, поскольку вознаграждения за один шаг для обеих моделей масштабируются от 0 до 10.

IV. ДИСКУССИИ

Мы получили сенсорно-моторную карту поведения человека при ходьбе на уровне спинного мозга с помощью искусственной нейронной сети. Обученная модель мышц и крутящего момента способна точно отслеживать углы суставов на основе экспериментальных данных. При сравнении результатов нашей модели при скорости 1,2 м/с с моделью Сонга и Гейера [7] наша модель имеет значения R 0,832 и 0,946 для левой и правой ноги соответственно по сравнению с 0,54 в их модели при воспроизведении движения приведения/отведения бедра. Также для движений лодыжки (фронтальная плоскость) наша модель имеет значение R 0,96 по сравнению с 0,46 в их модели при скорости 1,2 м/с. Наша модель также могла очень близко изучить асимметрию левого и правого человека, что проявляется в движении бедер во фронтальной плоскости.

Это значительные улучшения по сравнению с существующими моделями походки.

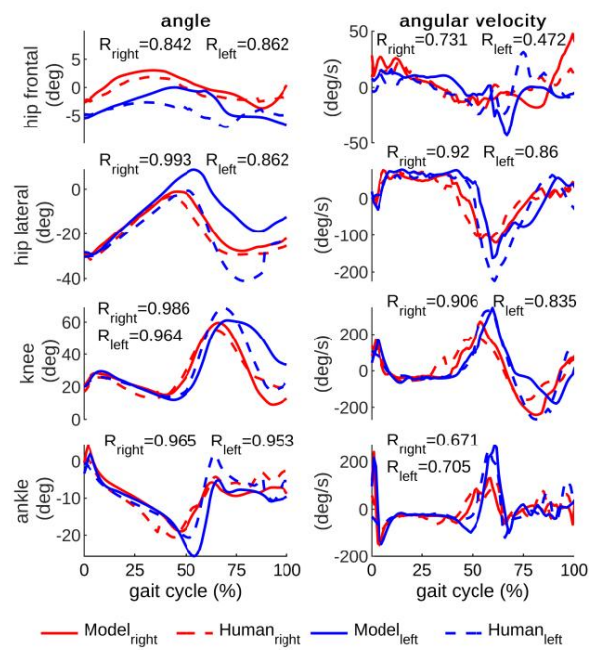


Рис. 6. Угол сустава и угловая скорость модели походки, основанной на крутящем моменте (сплошные линии) и экспериментальные данные на человеке (пунктирные линии) при скорости ходьбы 1,2 м/с. Красным и синим цветом обозначены данные суставов правой и левой ноги соответственно. R обозначает значение взаимной корреляции.

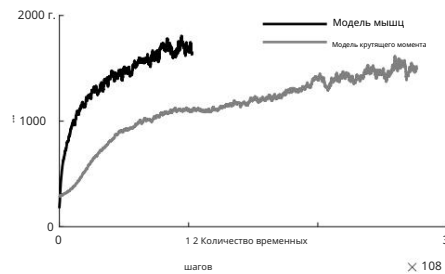


Рис. 7. Кривая обучения модели мышц (сзади) и модели крутящего момента (серого цвета) при тренировке.

Люди склонны ходить с предпочтительной частотой шагов и с заданной скоростью, чтобы минимизировать метаболические затраты [28]. Таким образом, оптимизируя модель для минимальных метаболических затрат, она позволяет модели имитировать энергетику человека. Мышечные паттерны были оптимизированы с использованием минимизации метаболических затрат, а также путем ограничения входных данных стимуляции по стандартным справочным данным по активации мышц человека. Мышцы HAV, HAD, HFL, GLU, VAS, SOL и TIA хорошо коррелируют с паттерном активации мышц человека. Напротив, паттерны активации групп мышц HAM, REF, BFSH и GAS явно отклоняются от человеческих паттернов. Низкая активация GAS компенсируется высокой активацией BFSH, поскольку эти мышцы вместе способствуют сгибанию колена. Аналогичным образом, более высокие уровни активации группы мышц REF в нашей модели могут соответствовать очень низким уровням активации группы мышц HAM, поскольку обе они способствуют разгибанию бедра. Динамика мышц отличается от данных на людях, поскольку метаболические затраты не достигли глобального минимума.

Это потребует дальнейшей оптимизации.

Профиль GRF является важной характеристикой локомоции человека [29]–[31]. Это видно по процедуре обучения, поскольку обратная связь GRF оказывает большое влияние на окончательную политику. Профиль GRF, созданный моделью, не такой гладкий, как экспериментальные данные на людях. Это может быть связано с жесткой моделью контакта с землей в моделировании. Эту проблему можно решить, используя более реалистичную модель контакта с землей.

Трехсегментная модель стопы, используемая в моделях OpenSim, может быть адаптирована, обеспечивая стопе гибкость плавного приземления, переворота и отталкивания по сравнению с одиночным жестким элементом стопы с четырьмя точками контакта, как в этой работе.

По сравнению с моделью, основанной на крутящем моменте, модель, основанная на мышцах, освоила политику, которая обеспечивает более высокую эффективность выборки во время тренировки. Этот результат аналогичен результатам других исследований, показывающих преимущества мышечного контроля для задач локомоции по сравнению с контролем, основанным на крутящем моменте [15], [3]. В этом исследовании не проводилось детального сравнения эффективности управления на основе крутящего момента и контроля мышц при передвижении на двух ногах. Но дальнейшие аспекты контроля мышц и крутящего момента будут рассмотрены в будущих исследованиях.

Исследование реакции на возмущения для проверки надежности модели дало неожиданные результаты. Поведение отклика не показывает никакого соответствия со стабильностью модели и моментами возмущений, а скорее является в высшей степени случайным. Например, модель становится нестабильной после 20 шагов при приложении возмущающего момента 40 Нм, но модель может пройти до 100 шагов после приложения возмущающего момента 150 Нм. Причина этой случайной реакции возмущения еще не ясна и требует рассмотрения в будущих исследованиях. Это может быть результатом изученной нелинейной политики и пассивной динамики мышц, что делает модель способной стабилизироваться даже при очень высоких крутящих моментах возмущений в бедре, таких как ± 200 Нм.

Мышечная модель могла выдержать такие сильные возмущения, хотя она тренируется с очень низким крутящим моментом возмущения, составляющим ± 5 Нм. Это связано с тем, что мышечный контроль использует пассивную динамику скелетно-мышечной структуры для обучения надежной политике. Но модель не смогла изучить поведение человека при ускорении/замедлении во время ходьбы. Это связано с отсутствием поведения ускорения/замедления в обучающих данных. Это можно улучшить, используя тренировочные данные, полученные в ходе хорошо спланированного эксперимента, в котором испытуемого просят ускоряться и замедляться между различными скоростями во время ходьбы.

Модели OpenSim не использовались в этом исследовании из-за вычислительной сложности. Такие модели могут быть полезны для будущих исследований с использованием более крупных вычислительных ресурсов. В этом исследовании мы используем нервно-мышечную модель общего назначения (например, параметры мышц, длина сегмента и распределение массы были получены из литературы), поскольку трудно получить конкретные параметры субъекта. Модель может лучше воспроизводить индивидуализированную походку, если параметры модели индивидуализированы. Модель крутящего момента не была оптимизирована с точки зрения энергоэффективности, поскольку включение термина энергоэффективности не привело бы к повышению эффективности выборки. Подробный

Анализ управления крутящим моментом и контролем мышц может быть рассмотрен в будущем исследовании.

ВЫВОДЫ

В этом исследовании мы исследовали идею изучения модели походки для выполнения человеческой ходьбы с использованием глубокого RL. Для изучения модели походки была разработана функция вознаграждения на основе кинематики, целевой скорости ходьбы, стабильности и метаболических затрат. Обученная модель походки способна точно воспроизводить кинематику и кинетику ходьбы человека при определенной целевой скорости. Результаты демонстрируют преимущества моделирования походки человека с помощью глубокого RL. Результаты также показывают, что управление на основе мышц превосходит управление на основе крутящего момента с точки зрения эффективности обучения выборке. Наша будущая цель — разработать индивидуальную модель ходьбы при ходьбе на основе глубокого RL, которая была бы более надежной и способна воспроизводить реакцию человека на неожиданные возмущения. С помощью изученной модели мы стремимся определить оптимальные схемы управления вспомогательными устройствами человека, такими как экзоскелеты и протезы нижних конечностей.

ПОДТВЕРЖДЕНИЕ

Мы благодарим доктора Мартина Гриммера (ТУ Дармштадта) за помощь в доработке рукописи и проведении экспериментов на людях. Эта работа была частично поддержана проектом ERA, финансируемым DFG, в рамках гранта № AN307/2-1 и гранта № SE1042/29-1.

ИСПОЛЬЗОВАННАЯ ЛИТЕРАТУРА

[1] Т. МакГир, «Пассивная динамическая ходьба» , Международный журнал исследований робототехники, том. 9, нет. 2, стр. 62–82, 1990.

[2] Р. Бликхан, «Модель пружинной массы для бега и прыжков» , Журнал биомеханики, вып. 22, нет. 1112, стр. 1217 – 1227, 1989.

[3] Х. Гейер, А. Зейфарт и Р. Бликхан, «Податливое поведение ног объясняет основную динамику ходьбы и бега» . Слушания. Биологические науки / Королевское общество, вып. 273, нет. 1603, стр. 2861–7, ноябрь 2006 г.

[4] С. Коллинз, А. Руина, Р. Тедрейк и М. Виссе, «Эффективные двуногие роботы на основе пассивно-динамических ходунков» , Science, vol. 307, нет. 5712, стр. 1082–1085, 2005.

[5] К. Хубицки, Дж. Граймс, М. Джонс, Д. Реневски, А. Спрвиц, А. Абате и Дж. Херст, «Атриас: проектирование и проверка беспривязного 3D-подпружиненного двуногого устройства» робот» , Международный журнал исследований робототехники, том. 35, нет. 12, стр. 1497–1521, 2016.

[6] Г. Чжао, М. Шарбафи, М. Влаттерс, Э. ван Ассельдонк и А. Зейфарт, «Управление на основе обратной связи по силе ног, основанное на шаблонной модели, может помочь человеку при ходьбе» , Международная конференция по реабилитационной робототехнике (ICORR) 2017 г. , июль 2017 г., стр. 473–478.

[7] С. Сонг и Х. Гейер, «Нейральная схема, которая подчеркивает спинальную обратную связь, генерирует разнообразные модели передвижения человека» , Журнал физиологии, том. 593, нет. 16, стр. 3493–3511, 2015. [8] —, «Оценка нейромеханической модели управления ходьбой с использованием экспериментов с возмущениями» , Frontiers in Computational Neuroscience, vol. 11, с. 15, 2017.

[9] И. Мордатч, Дж. М. Ван, Э. Тодоров и В. Колтун, «Анимация нижних конечностей человека с использованием контактно-инвариантной оптимизации» , ACM Transactions on Graphics (TOG), vol. 32, нет. 6, с. 203, 2013.

[10] Дж. Мерел, А. Ахуджа, В. Фам, С. Туньясуванакул, С. Лю, Д. Тиру-мала, Н. Хесс и Г. Уэйн, «Иерархический зрительно-моторный контроль гуманоидов» , препринт arXiv arXiv: 1811.09656, 2018 г.

[11] Дж. М. Ван, С. Р. Хамнер, С. Л. Дельп и В. Колтун, «Оптимизация контроллеров локомоции с использованием биологических приводов и целей» , Транзакции АСМ на графике, том. 31, нет. 4, 2012.

[12] Т. Гейтенбек, М. Ван Де Панн и А. Ф. Ван Дер Стаппен, «Гибкое передвижение двуногих существ с помощью мышц» , ACM Transactions on Graphics (TOG), vol. 32, нет. 6, с. 206, 2013.

[13] Ю. Ли, М. С. Парк, Т. Квон и Дж. Ли, «Управление локомоцией многомышечных гуманоидов» , ACM Transactions on Graphics (TOG), том. 33, нет. 6, с. 218, 2014.

[14] Ю. Цзян, Т. Ван Боуве, Ф. Де Гроот и К.К. Лю, «Синтез биологически реалистичных движений человека с использованием совместного крутящего момента» , препринт arXiv arXiv:1904.13041, 2019.

[15] XB Peng, P. Abbeel, S. Levine и M. van de Panne, «Deepmimic: глубокое обучение с подкреплением физических навыков персонажей на основе примеров» , препринт arXiv arXiv:1804.02717, 2018.

[16] XB Peng, G. Berseth, K. Yin и M. Van De Panne, «Deeploco: навыки динамического передвижения с использованием иерархического глубокого обучения с подкреплением» , ACM Transactions on Graphics (TOG), vol. 36, нет. 4, с. 41, 2017.

[17] XB Peng и M. van de Panne, «Изучение навыков передвижения с помощью deeprl: имеет ли значение выбор пространства действия?» в материалах симпозиума ACM SIGGRAPH/Eurographics по компьютерной анимации. AKM, 2017, с. 12.

[18] X.C. Чанг, М.К. Фу, Дж. Ху и С.И. Маркус, «Google Deep Minds AlphaGo» , OR/MS Today, vol. 43, нет. 5, стр. 24–29, 2016.

[19] Дж. Шульман, Ф. Вольски, П. Дхаривал, А. Рэдфорд и О. Климов, «Алгоритмы оптимизации проксимальной политики» , препринт arXiv arXiv:1707.06347, 2017.

[20] Дж. Шульман, С. Левин, П. Аббил, М. Джордан и П. Мориц, «Оптимизация политики доверительного региона» , Международная конференция по машинному обучению, 2015 г., стр. 1889–1897.

[21] Т.П. Лилликрап, Дж.Дж. Хант, А. Притцель, Н. Хесс, Т. Эрез, Ю. Тасса, Д. Сильвер и Д. Вьерстра, «Непрерывный контроль с глубоким обучением с подкреплением» , препринт arXiv arXiv:1509.02971. , 2015.

[22] В. Мних, А. П. Бадиа, М. Мирза, А. Грейвс, Т. Лилликрап, Т. Харли, Д. Сильвер и К. Кавучуоглу, «Асинхронные методы для глубокого обучения с подкреплением» , Международная конференция по машинному обучению, 2016, стр. 1928–1937.

[23] З. Ван, В. Бапст, Н. Хесс, В. Мних, Р. Мунос, К. Кавучуоглу и Н. де Фрейтас, «Образец эффективного актера-критика с воспроизведением опыта» , препринт arXiv arXiv:1611.01224, 2016.

[24] П. Дхаривал, К. Гессе, О. Климов, А. Никол, М. Плапперт, А. Рэдфорд, Дж. Шульман, С. Сидор, Ю. Ву и П. Жохов, «Базовые линии Openai» , <https://github.com/openai/baselines>, 2017 г.

[25] Э. Тодоров, Т. Эрес и Ю. Тасса, «MuJoCo: физический двигатель для управления на основе моделей» , Международная конференция IEEE/RSJ по интеллектуальным роботам и системам, октябрь 2012 г., стр. 5026–5033. .

[26] Дж. Перри, Дж. Бернфилд и Дж. Бернфилд, Анализ походки: нормальная и Патологическая функция. СЛАК, 2010.

[27] Р.М. Александр, «Оптимальная конструкция мышц для колебательных движений» , Журнал теоретической биологии, том. 184, нет. 3, стр. 253–259, 1997.

[28] М.И. Зарру, Ф.Н. Тодд и Х.Дж. Ралстон, «Оптимизация расхода энергии во время ходьбы» , Европейский журнал прикладной физиологии и профессиональной физиологии, том. 33, нет. 4, стр. 293–306, декабрь 1974 г. [Онлайн]. Доступно: <https://doi.org/10.1007/BF00430237> [29] К. Хираи, М. Хироэ, Ю. Хайкава и Т. Такенака, «Развитие гуманоидного робота Honda» , в журнале Robotics and Automation, 1998 г. .

Слушания. Международная конференция IEEE 1998 г., вып. 2. IEEE, 1998, стр. 1321–1326.

[30] Дж. П. Хантер, Р. Н. Маршалл и П. Дж. Макнейр, «Взаимосвязь между импульсом силы реакции опоры и кинематикой ускорения при спринтерском беге» , Журнал прикладной биомеханики, том. 21, нет. 1, стр. 31–43, 2005.

[31] Т. С. Келлер, А. Вайсбергер, Дж. Рэй, С. Хасан, Р. Шиави и Д. Шпенглер, «Взаимосвязь между вертикальной силой реакции опоры и скоростью во время ходьбы, медленного бега трусцой и бега» , Клиническая биомеханика. , том. 11, нет. 5, стр. 253–259, 1996.

[32] А. Круз Руис, К. Понтонье, Н. Проност и Г. Дюмон, «Мышечное управление для анимации персонажей» , в журнале Computer Graphics Forum, vol. 36, нет. 6. Интернет-библиотека Wiley, 2017, стр. 122–147.

[33] Дж. Шредер, К. Кавамура, Т. Гокель и Р. Диллманн, «Улучшенное управление гуманоидной рукой, приводимой в движение пневматическими приводами» , Proceedings of Humanoids 2003, 2003.

[34] Т. Комура, Ю. Синагава и Т.Л. Кунии, «Мышечный контроллер человеческого тела с прямой связью» , в журнале Computer Graphics Forum, vol. 16, нет. 3. Интернет-библиотека Wiley, 1997, стр. C165–C176.

[35] Х. Шен, «Нелинейное управление на основе моделей пневматических сервосистем с искусственными мышцами» , Control Engineering Practice, vol. 18, нет. 3, стр. 311–317, 2010.