

Онлайн-контроль обучения с подкреплением для Персонализация роботизированного протеза коленного сустава

Юэ Вэнь^{ID}, Дженни Си, научный сотрудник IEEE, Андреа Брандт^{ID},
Сян Гао и Хе (Хелен) Хуан^{ID}, Старший член IEEE

Аннотация. Роботизированные протезы выполняют большую функцию, чем пассивные протезы, но перед нами стоит задача настройки большого количество параметров управления для персонализации устройства для отдельных пользователей с ампутированными конечностями. Эту проблему нелегко решить с помощью традиционных конструкций управления или новейших робототехнических технологий. Обучение с подкреплением (RL) естественно привлекательно. Недавний, беспрецедентный успех AlphaZero продемонстрировал RL как реальное и крупномасштабное средство решения проблем. Однако настройка протеза проблема связана с несколькими нерешенными проблемами, такими как что у него нет известной и стабильной модели, непрерывные состояния и контроль над проблемой могут привести к проклятию размерности, а система человек-протез постоянно подвержен воздействию шума при измерении, изменениям окружающей среды и изменениям, вызванным человеческим телом. В этой статье мы продемонстрировали осуществимость прямого эвристического динамического программирования, подхода приближенного динамического программирования (ADP), позволяющего автоматически настройте 12 параметров роботизированного коленного протеза в соответствии с индивидуальными потребностями потребности пользователей-людей. Мы протестировали ADP-тюнер на двух субъектах (один здоровый субъект и один субъект с ампутированными конечностями), идущие по фиксированная скорость на беговой дорожке. ADP-тюнер научился достигать цели кинематика походки в среднем за 300 циклов походки или 10 мин ходьбы. Мы наблюдали улучшение производительности настройки ADP, когда перевел ранее изученный контроллер ADP на новый сеанс обучения по тому же предмету. В меру наших знаний, Наш подход к персонализации роботизированных протезов — это первая реализация онлайн-контроля обучения ADP для решения клинической проблемы. с участием человеческих субъектов.

Индексные термины — приближительное динамическое программирование (ADP), прямое эвристическое динамическое программирование (dHDP), подкрепление обучение (RL), роботизированный коленный протез.

I. ВВЕДЕНИЕ

Роботизированные протезы по сравнению с традиционными достигшими устройствами показали большие перспективы. для дальнейшего улучшения мобильности людей с более низкими ампутация конечностей [1]–[5]. Обычно управление роботизированным протезом

Рукопись получена 1 мая 2018 г.; рассмотрено 26 сентября 2018 г.; принял 19 декабря 2018 г. Работа выполнена при частичной поддержке Национального научного фонда. Фонд в рамках Гранта 1563454, Гранта 1563921, Гранта 1808752 и Гранта 1808898. Эту статью рекомендовал заместитель редактора Х. Чжан. (Соответствующие авторы: Хэ (Хелен) Хуан; Дженни Си.)

Ю. Вэнь, А. Брандт и Х. Хуан из Объединенного комитета UNC/NCSU. Департамент биомедицинской инженерии Университета штата Северная Каролина, Роли, Северная Каролина 27695 США, и Университет Северной Каролины в Чапеле. Хилл, Чапел-Хилл, Северная Каролина 27599, США (электронная почта: hhuang11@ncsu.edu).

Дж. Си и С. Гао работают на кафедре электротехники, компьютеров, и энергетическая инженерия, Университет штата Аризона, Темпе, Аризона 85281, США (электронная почта: si@asu.edu).

Доступны цветные версии одного или нескольких рисунков в этой статье. онлайн по адресу <http://ieeexplore.ieee.org>.

Цифровой идентификатор объекта 10.1109/TCYB.2019.2890974

состоит из конечного автомата и контроллера низкого уровня

для регулирования импеданса протезного сустава. Существующая робототехника Контроллеры протезов полагаются на большое количество настраиваемых параметров (т.е. 12–15 для коленных протезов [3], [5], [6] и 9–15 для протезов голеностопного сустава [3], [4]) для одного режима передвижения, например ходьбы по ровной земле. Количество параметры растут с увеличением количества включенных локомоций режимы увеличиваются. Эти параметры управления должны быть персонализированы с учетом индивидуальных особенностей пользователя, таких как рост, вес, и физические способности. В настоящее время в клиниках контроль протезирования параметры персонализируются вручную [7], [8], что можно время, трудозатраты и затраты.

Исследователи попытались повысить эффективность настройка протеза с помощью трех основных подходов. Первый подход заключается в оценке параметров импеданса управления с помощью либо скелетно-мышечная модель [9], либо измерения биологического импеданса суставов [10], [11]. Однако эти методы не были валидированы для реального контроля протеза. Второе решение не касается непосредственно настройки параметров, но направлена на уменьшение количества управляющих параметров [7], [12]. Третий метод обеспечивает автоматическую настройку параметров путем кодирования решений протезистов [13], что может занять много времени. выполнять и потенциально предвзято со стороны отдельных протезистов опыт. Поэтому нам нужны новые подходы для решения этой проблемы. Проблема настройки параметров протеза.

Персонализация носимых роботов, то есть роботизированных протезов и экзоскелетов, требует оптимальных адаптивных решений управления. Коллер и др. [14] использовали метод градиентного спуска для оптимизации время появления экзоскелета голеностопного сустава для повышения работоспособности (AB) эффективность походки человека. Чжан и др. [15] использовали стратегию эволюции для оптимизации четырех контрольных параметров голеностопного сустава. экзоскелет. Дин и др. [16] применила байесовскую оптимизацию. определить два контрольных параметра помощи в разгибании бедра. Эти методы перспективны, но они не были использованы для потенциальная персонализация роботизированных протезов, поскольку их трудно масштабировать до многомерного (5) пространства параметров, адаптироваться к изменяющимся условиям (например, изменению веса) или контролировать выбранный показатель эффективности в повседневной жизни (например, метаболический расходы).

Обучение с подкреплением (RL) может служить альтернативой. подход к персонализации протезов нижних конечностей. Хотя он победил двухтысячелетние человеческие знания ГО, научившись играть в игру за несколько часов. RL еще нет применялись в клинических ситуациях с большей сложностью и человеческие взаимодействия. Например, контроль над носимой робототехникой создает дополнительную проблему проклятия.

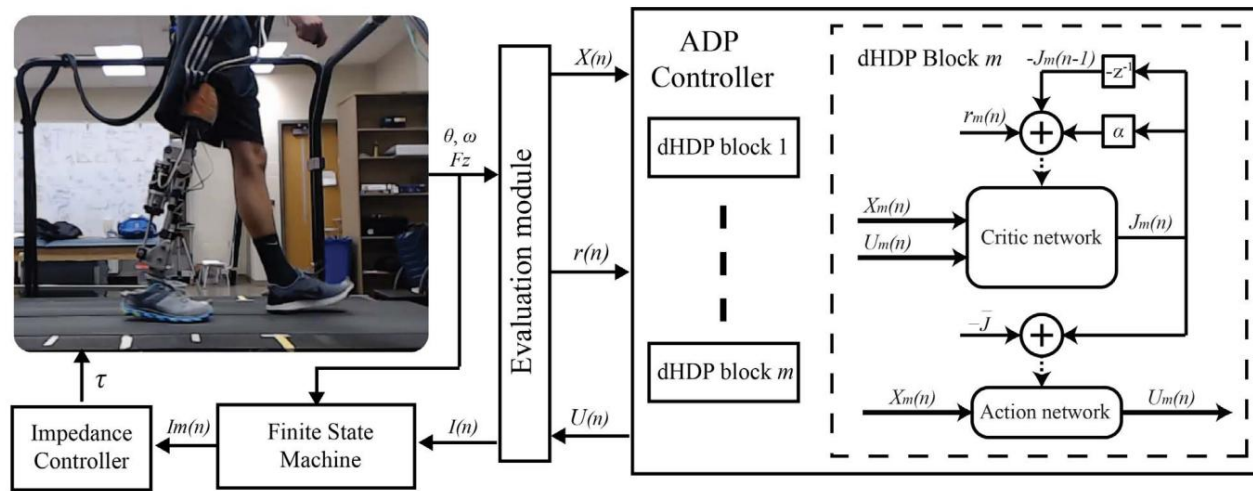


Рис. 1. Блок-схема ADP-тюнера — схемы настройки параметров автоматического управления коленным суставом робота с помощью dHDP с участием человека с ампутированной конечностью. Система управления обучением работает в трех различных временных масштабах: 1) контроллер импеданса в реальном времени обеспечивает выходные данные с частотой 100 Гц для регулирования крутящего момента сустава; 2) конечный автомат работает с частотой походки (обозначенной индексом времени g) с четырьмя фазами на цикл ходьбы; и 3) генерируемое dHDP управление обновляется $I_{m,p}$ каждые несколько шагов (обозначается индексом времени p) для обновления параметров импеданса. Соответствующие переменные на рисунке определены и обсуждаются в разделах II и III. ADP-тюнер состоит из четырех блоков dHDP ($m = 1, 2, 3, 4$), соответствующих четырем фазам походки в автоматном регуляторе импеданса.

высокой размерности в пространствах непрерывного состояния и управления/действия, а также требования достижения оптимальных показателей производительности в условиях неопределенности системы, шума измерений и неожиданных помех. Приближенное динамическое программирование (ADP) [17]–[19] является синонимом RL, особенно в сообществах исследователей систем управления и операций, и оно показало большие перспективы для решения вышеупомянутых проблем.

Адаптивные критические дизайны представляют собой серию дизайнов ADP, первоначально предложенных Вербосом [20]–[22]. За последнее десятилетие адаптивный критический дизайн был разработан и широко применяется для робастного управления [23], оптимального управления [24]–[26] и событийно-управляемых приложений [27]–[29]. Зависимое от действий эвристическое динамическое программирование (ADHDP) похоже на Q-обучение, но обладает многообещающей масштабируемостью [30]. Новые разработки в области ADHDP (например, нейронная Q (NFQ), NFQ с непрерывными действиями (NFQCA) [31], прямое эвристическое динамическое программирование (прямое HDP или dHDP) [32], прямая модель управления обучением [33]), и нечеткое адаптивное динамическое программирование [25]) появились и продемонстрировали свою пригодность для решения сложных и реалистичных задач управления обучением. Более того, алгоритмы dHDP и NFQCA (отмеченные как пакетный вариант dHDP [34]) связаны, пожалуй, с большинством продемонстрированных применений RL-управления для задач непрерывного состояния и управления [34]–[43]. Поэтому целью данной статьи является реализация dHDP [32] в режиме реального времени для онлайн-контроля обучения с целью адаптивной настройки параметров импеданса протезного колена.

Прежде чем приступить к реальным экспериментам с участием людей, мы провели симуляционное исследование [44]. Мы разработали ADP-тюнер для протеза коленного сустава и проверили этот контроль на установленной вычислительной модели OpenSim [45] для динамического моделирования походки людей с ампутированными конечностями. Мы сравнили dHDP с NFQCA. Результаты нашего моделирования показали, что контроллер dHDP позволил смоделированной модели человека с ампутированной конечностью научиться ходить за меньшее количество циклов походки и с более высоким уровнем успеха, чем NFQCA [44]. Хотя это интересно и многообещающе, неизвестно

как dHDP работает с реальным человеком в процессе. Это связано с тем, что модель OpenSim игнорирует реакцию человека на действия протеза, естественную вариативность походки и, самое главное, безопасность.

Этот документ, представленный здесь, включает следующие основные вклады.

- 1) Насколько нам известно, это первое исследование, в котором реализован контроллер обучения ADP для реальной ситуации, такой как персонализация роботизированных протезов для людей. Это приложение является новым в области реабилитации.
- 2) Немодальный dHDP был адаптирован для эффективного использования данных и времени для этого приложения и был реализован для автоматической настройки 12 параметров импеданса посредством взаимодействия с системой человек-протез в режиме онлайн.
- 3) Исследование впервые продемонстрировало, что предлагаемый контроль на основе RL осуществим и при дальнейшем развитии потенциально может стать безопасным и практичным для клинического использования.

Оставшаяся часть статьи организована следующим образом.

В разделе II описывается система «человек-протез» и формулируется проблема настройки/конфигурации человеческого протеза.

В разделе III представлена конструкция ADP-тюнера для онлайн-управления протезом коленного сустава. В разделе IV подробно рассматриваются соображения дизайна для реальных людей. В разделе V поясняется экспериментальная оценка ADP-тюнера. Результаты представлены в разделе VI. Замечания и обсуждения представлены в разделе VII, за которыми следуют выводы в разделе VIII.

II. ПОСТАНОВКА ЗАДАЧИ УПРАВЛЕНИЯ ПРОТЕЗИРОВАНИЕМ КОЛЕНА

На рис. 1 показан предлагаемый нами подход к автоматической настройке параметров управления протезом коленного сустава с участием человека.

В этом разделе мы представляем систему человека-протеза, а именно человека с ампутированной конечностью, носящего роботизированный коленный протез.

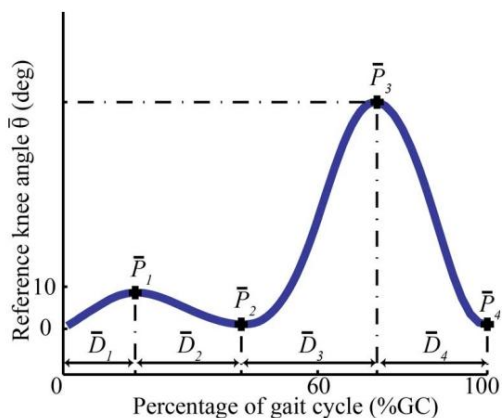


Рис. 2. Характеристика кинематики колена, близкой к нормальной, в течение одного в качестве цели контроля обучения использовался цикл походки, где \bar{P}^m обозначает угол функция, а \bar{P}^m указывает на функцию продолжительности. Указывается индекс фазы по $m = 1, 2, 3, 4$. Старт на 0% и финиш на 100% - это удар пяткой. событий, а 60 % - приблизительно время начала игры.

A. Конфигурация человеческого протеза

И механический интерфейс, и параметры управления

Роботизированный коленный протез должен быть персонализирован для каждого пользователя.

Люди различаются по своему физическому состоянию, например, по росту, вес и мышечная сила. Во-первых, длина пилона, выравнивание протеза и посадка гильзы, которая

Интерфейсы пользователя и протеза должны быть настроены с помощью дипломированный протезист. Затем параметры управления коленным роботом должен быть настроен для оказания индивидуальной помощи с колена протез. Предлагаемая нами автоматическая настройка, реализованная как дополнительное управление на основе RL, показана на рис. 1.

B. Импеданс конечного автомата протезного колена Контроллер

Конечный автомат управления импедансом (FSM-IC, рис. 1) является признанной основой для управления роботизированными коленными протезами [3], [5], [6], [46]. На основании контакта ноги с землей и движения коленного сустава, один цикл походки делится на четыре фазы (соответствующие $m = 1, \dots, 4$ на рис. 1): 1) фаза опорного сгибания (STF, $m = 1$); 2) фаза расширения стойки (СТЭ, $m = 2$); 3) фаза качания-сгибания (SWF, $m = 3$); и 4) фаза расширения качания (SWE, $m = 4$). Фазовые переходы могут быть вызваны измерениями тензодатчика и датчик угла в протезном устройстве. Затем предоставляются соответствующие параметры импеданса I_m , как описано в (1).

к контроллеру импеданса

$$I_m = [k_m, \theta_{em}, b_m]^T. \quad (1)$$

На каждом этапе m роботизированное колено регулируется другой контроллер импеданса (2) для создания фазозависимых динамические свойства. Контроллер импеданса контролирует положение коленного сустава θ и скорость $\dot{\theta}$ и управляет коленным суставом. крутящий момент τ в реальном времени на основе трех параметров импеданса: 1) жесткость k ; 2) демпфирование b ; и 3) положение равновесия θ_e

$$\tau_m = k_m(\theta - \theta_{em}) + b_m\dot{\theta}. \quad (2)$$

Таким образом, при четырех фазах ходьбы имеется 12 полных импедансов. параметры, которые необходимо настроить для каждого режима передвижения.

C. Представление кинематики колена

Кинематика коленного сустава робота измеряется датчиком угла, установленным на вращательном шарнире. Датчик угла считывает ноль, когда коленный сустав разгибается до положения голени линия с бедром и положительное значение в градусах, когда коленный сустав согнут. Обычно траектория угла коленного сустава в одном цикле ходьбы имеет локальный максимум во время сгибания стойки и сгибание при махе, а также локальный минимум во время разгибания и разгибания при махе (рис. 2). Пиковое значение каждой фазы в первую очередь определяется параметрами импеданса в этой фазе. Поэтому мы представили кинематику коленного сустава за один цикл ходьбы с четырьмя парами значений максимального угла. P и соответствующие им значения длительности D : $[P_m, D_m]$, где $m = 1, 2, 3, 4$. Аналогично мы извлекли те же функции из нормативная кинематика колена [47] в качестве целевых признаков, обозначаемых как $[P^m, D^m]$ (рис. 2).

D. Процесс настройки системы «человек-протез»

Процесс настройки построен на широко используемой схеме FSM-IC, и его цель состоит в том, чтобы найти набор импедансов. параметры, которые позволяют системе человек-протез генерировать нормативную целевую кинематику коленного сустава. Как упоминалось ранее, на каждой фазе походки действовали три параметра импеданса, и соответственно, были извлечены кинематические характеристики колена во время каждой фазы походки. Для удобства обсуждения мы будем с этого момента опустить индекс m для фазы походки.

Для системы человек-протез управляющими входами являются параметры импеданса $I(n)$, а выходы представляют собой характеристики $x(n)$ кинематики протезного колена

$$\begin{aligned} I(n) &= [k(n), \theta_e(n), b(n)]^T \\ x(n) &= [P(n), D(n)]^T \end{aligned} \quad (3)$$

где n обозначает индекс обновления каждого обновления параметра, то есть каждые семь циклов ходьбы.

В процедуре настройки параметры импеданса равны обновлено как

$$I(n) = I(n-1) + \beta U(n-1) \quad (4)$$

где U обозначает действия ADP-тюнера, $\beta \in \mathbb{R}^{3 \times 1}$ являются коэффициентами масштабирования для присвоения физических величин действий и является произведением Адамара двух векторов.

Состояния системы человек-протез, используемой в контроллер обучения определяются как

$$X(n) = \gamma \quad x^T(n) \quad \bar{x}^T(n), \quad x^T(n) - x^T(n-1) \quad (5)$$

где $\gamma \in \mathbb{R}^{4 \times 1}$ — вектор масштабных коэффициентов для нормализации состояния $[1, 1]^T$ и \bar{x}^T являются признаками $[P^-, D^-]^T$ кинематика целевого колена. Ошибки признаков $x(n) - \bar{x}$ отражают расстояние до целевой кинематики колена и изменение функции скорости $x(n) - x(n-1)$ получают динамическое изменение в течение Процедура настройки.

В процессе настройки действия от ADP-тюнера настраиваются на параметры импеданса, которые являются непрерывными, а состояния ADP-тюнера получаются из особенности кинематики колена, которые также являются непрерывными. Таким образом, процесс настройки человеческого протеза имеет непрерывные состояния и элементы управления. Уравнения (3)–(5) реализуются в

«модуль оценки» (рис. 1) как интерфейс между системой человек-протез и ADP-тюнером. Кроме того, модуль оценки включает в себя сигналы усиления, подаваемые в ADP-тюнер на основе выходных сигналов системы человек-протез (описанной в разделе IV-A).

III. ADP-ТЮНЕР

Для данной задачи настройки параметров импеданса протеза человека мы реализовали ADP-тюнер с четырьмя параллельными блоками dHDP, соответствующими четырем фазам походки: 1) STF; 2) СТЭ; 3) SWF; 4) SWE (рис. 1). Каждый блок dHDP принимал четыре переменные состояния в (5) и настраивал три параметра импеданса для соответствующей фазы. Все блоки dHDP были идентичны, включая одну нейронную сеть действия (ANN) и одну нейронную сеть критика (CNN). Таким образом, без потери общности, мы представляем подробную реализацию dHDP без номеров фаз.

A. Функция полезности/сигнал подкрепления.

Сигнал подкрепления $r(n)$ определяется как мгновенная стоимость, которая определяется из системы человек-протез.

$$r(n) = \begin{cases} 1, & \text{если } x(n) \in [B_l, B_u] \\ -0,8, & \text{если } S > 4 \\ 0, & \text{в противном случае} \end{cases} \quad (6)$$

где $[B_l, B_u]$ обозначает границы безопасности, как определено в Разделе IV-A, S — штрафной балл, а сигнал подкрепления $-0,8$ налагается на блок ADP, когда значение S больше 4, указывая блок dHDP продолжает подстраивать параметр импеданса в неблагоприятном направлении (т.е. увеличивая угловую ошибку и/или ошибку длительности) [44]. Когда сигнал подкрепления равен -1 , параметры импеданса системы человек-протез сбрасываются до значений по умолчанию.

Общая стоимость эксплуатации на этапе настройки ADP n определяется выражением

$$J(n) = r(n+1) + \alpha r(n+2) + \dots + \alpha N r(n+N+1) + \dots \quad (7)$$

где α — ставка дисконтирования ($0 < \alpha < 1$), а N бесконечно. Его можно переписать как

$$J(n) = r(n+1) + \alpha J(n+1). \quad (8)$$

Б. Критическая нейронная сеть

CNN состояла из трех слоев нейронов (7-7-1) с двумя слоями весов, принимала состояние X $R \times 4$ и действие U $R \times 4$ в качестве входных данных и прогнозировала общую стоимость реализации. Дж

$$\hat{J}(n) = Wc2(n) \phi(Wc1(n) X^T(n), U^T(n)) \quad (9)$$

где $Wc1$ $R \times 7$ — весовая матрица между входным и скрытым слоями, а $Wc2$ $R \times 1$ — весовая матрица между скрытым и выходным слоями. И

$$\phi(v) = \frac{1}{1 + \exp(-v)} \quad (10)$$

$$= Wc1(n) X^T(n), U^T(n) hc1(n) = \quad (11)$$

$$\phi(vc1(n)) \quad (12)$$

где $\phi(\cdot)$ — тангенсо-сигмовидная функция активации, а $hc1$ — выход скрытого слоя.

Ошибка прогноза ec CNN может быть записана как

$$ec(n) = \alpha \hat{J}(n) - J(n). \quad (13)$$

Чтобы исправить ошибку прогноза, целью обновления веса было минимизировать квадрат ошибки прогнозирования Ec , обозначенный как

$$E_c(n) = \frac{1}{2} (ec(n))^2. \quad (14)$$

Правило обновления веса для CNN основывалось на градиенте адаптации, предоставленная

$$W(n+1) = W(n) + W(n). \quad (15)$$

Обновления веса матрицы скрытого слоя $Wc2$ были

$$Wc2(n) = Ic(n) \frac{Ec(n)}{Wc2(n)} = lc(n) \frac{ec(n) \hat{J}(n)}{Wc2(n) ec(n)} \quad (16)$$

Обновления веса матрицы входного слоя $Wc1$ были

$$Wc1(n) = Ic(n) \frac{Ec(n)}{Wc1(n)} hc1(n) \frac{Ec(n) ec(n) hc1(n) vc1(n)}{ec(n) vc1(n)} Wc1(n) = lc(n) \frac{J(n)}{Wc1(n)} \quad (17)$$

где $Ic > 0$ — скорость обучения CNN.

C. Нейронная сеть действия

ИНС состояла из трех слоев нейронов (4-7-3) с двумя слоями весов, принимала состояние X $R \times 4$ от системы человек-протез и выдавала действия U $R \times 4$ для корректировки импедансные параметры системы человек-протез

$$U(n) = \phi(Wa2(n) \phi(Wa1(n) X(n))) \quad (18)$$

где $Wa1$ $R \times 7$ и $Wa2$ $R \times 3$ — весовые матрицы, а $\phi(\cdot)$ — тангенсо-сигмоидальная функция активации скрытого слоя и выходного слоя.

Согласно нашей формулировке задачи, целью адаптации ИНС было обратное распространение ошибки между желаемой конечной целью, обозначаемой J^* , и аппроксимированной общей стоимостью реализации \hat{J} . А J^* был установлен на 0, что означает «успех». Таким образом, целью обновления политики было минимизировать абсолютное значение общей стоимости использования до 0. Правило обновления веса для ИНС должно было минимизировать следующую ошибку производительности:

$$Ea(n) = \frac{1}{2} (\hat{J}(n) - J^*)^2. \quad (19)$$

Аналогично, весовая матрица была обновлена на основе градиентного спуска.

$$W(n+1) = W(n) + W(n). \quad (20)$$

Обновления веса матрицы скрытого слоя $Wa2$ были

$$Wa2(n) = Ia(n) \frac{Ea(n)}{Wa2(n)}. \quad (21)$$

Алгоритм 1. Онлайн-ADP-настройка параметров импеданса.
для роботизированного коленного протеза

Инициализация системы человек-протез: $I(0)$, $x(0)$ и
Случайная инициализация весов ANN и CNN.

Шаг 1: Обновление значения

Получите состояние $X(n)$ из (5) и сигнал подкрепления $r(n)$ из (6)
Обновите веса CNN, используя (13)–(17).

Шаг 2: Улучшение политики.

Обновите веса ИНС, используя (19)–(22).
Рассчитайте $U(n)$ из (18) и обновите $I(n)$ с помощью (4).
Сбросить $I(n)$, если $r(n) == -1$ из (6).

Перейдите к шагу 1 до тех пор, пока не появятся критерии прекращения (раздел IV-E).

Обновления веса матрицы входного слоя W_{a1} были

$$W_{a1}(n) = I_{a1}(n) + \frac{E_{a1}(n)}{W_{a1}(n)} \quad (22)$$

где $I_{a1} > 0$ — скорость обучения ИНС.

Вышеупомянутые обновления весов ANN и CNN, а также реализация ADP-тюнера обобщены в алгоритме 1. Веса как ANN, так и CNN были инициализированы равномерно распределенными случайными числами от -1 до 1. В мягких и интуитивно понятных условиях dHDP с дисконтированная стоимость обладает свойством равномерной предельной ограниченности [48].

IV. АСПЕКТЫ ПРОЕКТИРОВАНИЯ ОНЛАЙН- ОБУЧЕНИЯ ДЛЯ ЛЮДЕЙ

Исследования на людях отличаются от симуляционных исследований, и поэтому мы модифицировали и внедрили алгоритм настройки ADP с учетом реальных условий для людей, носящих протезную ногу.

A. Границы

Безопасности Для протезов, несущих вес, безопасность является первоочередной задачей, поэтому мы включили ограничения, гарантирующие, что выходные данные системы человек-протез остаются в безопасном диапазоне [обозначены $[B_l, B_u]$ в (6)]. Во-первых, чтобы избежать потенциального вреда пользователю с ампутированной конечностью, мы установили границы ошибок характеристик в 1,5 раза больше стандартного отклонения средних кинематических характеристик колена людей, идущих без протеза (т. е. STF 10,5 градусов, STE 7,5 градусов, SWF 9 градусов, а SWE 6 градусов [49]). Во-вторых, чтобы избежать столкновений механических частей протеза, которые могут повредить роботизированный протез, мы установили ограничения на диапазон движения до -5 градусов и 60 градусов. Эти ограничения определяли диапазон исследования для контроллера ADP, чтобы избежать повреждения или вреда системе человек-протез. Когда характеристики превышали эти диапазоны, мы сбрасывали параметры управления до значений по умолчанию, определенных в начале каждого экспериментального сеанса, которые, как известно, приводили к безопасной эксплуатации. В то же время на ADP-тюнер был отправлен сигнал подкрепления -1.

B. Надежное извлечение признаков

Шум сигнала датчика неизбежен в реальных протезах, поэтому мы внедрили надежный метод извлечения признаков для извлечения характеристик угла коленного сустава. В действительности траектория угла коленного сустава не идеальна главным образом по двум причинам: 1) неизбежный шум в показаниях датчика угла и 2) почти пологая траектория угла в некоторых местах цикла походки, где датчик

показания оставались стабильными. В этих условиях временной признак D сильно менялся при получении пиковых значений и продолжительности цикла ходьбы. Чтобы решить эту проблему, мы сначала определили минимальные или максимальные характеристики $[P^-, D^-]$ на траектории угла коленного сустава, где i обозначает индекс выборки датчика (100 Гц). Для каждой выборки θ_j траектории угла коленного сустава имеются две особенности $[P_j, D_j]$. Мы выбрали и использовали функции с индексом j для замены $[P^-, D^-]$, где

$$j = \arg \min [P_j, D^-] \quad (23)$$

и индекс j находится в пределах $[i - 10, i + 10]$, а соответствующий угловой признак P_j находится в пределах $[P^-, D^-]$. Это необходимо для поиска надежных и репрезентативных характеристик длительности на основе данных датчика реального времени. меры.

C. Человеческая изменчивость

Чтобы смягчить неизбежные изменения походки человека от шага к шагу, ADP-тюнер обрабатывал характеристики системы человека-протеза для каждого цикла походки, но обновления контроля производились для каждых семи циклов походки. То есть люди ходили с обновленным набором параметров импеданса в течение семи циклов походки. Если угловые характеристики конкретного цикла походки превышали 1,5 стандартных отклонения от среднего значения семи циклов походки, это считалось выбросом и удалялось. Это исключило влияние случайных спотыканий или несбалансированных циклов походки на обновления управления.

D. Предотвращение неправильного сигнала

подкрепления. Как упоминалось ранее, особенности одной фазы походки влияют на последующие фазы. Чтобы избежать распространения ошибочного сигнала подкрепления, мы предоставили сигнал подкрепления -1 только тому блоку dHDP, который демонстрировал выходящие за пределы характеристики угла/длительности. Если одновременно генерировались несколько сигналов подкрепления отказа, мы расставляли приоритеты (от высокого к низкому) сигналу подкрепления обратной связи в следующем порядке: STF, SWF, SWE, STE. Другими словами, если несколько фаз генерировали сигналы подкрепления -1 в одной и той же итерации настройки, мы применяли сигнал подкрепления -1 к блоку dHDP, который имел более высокий приоритет.

E. Критерии прекращения действия

Для практических применений с участием человека в контуре управления в реальном времени необходимы критерии завершения, чтобы избежать утомления человека в процедуре настройки. Процедура настройки была ограничена 70 итерациями настройки (т.е. $7 \times 70 = 490$ циклов походки) и завершалась раньше, если настройка была успешной. Поскольку системы «человек-протез» сильно нелинейны, уязвимы к шуму и помехам и подвержены неопределенности, мы ввели диапазон допуска мкм ($m = 1, \dots, 4$ обозначает четыре фазы походки) для приемлемых диапазонов характеристик. ошибки, что в 1,5 раза превышало стандартное отклонение характеристик при более чем 15 циклах походки без дополнительных входных сигналов управления импедансом. Настройка параметров на данном этапе считается успешной, если особенности этого этапа соответствуют критерию допуска как минимум для трех из пяти предыдущих итераций настройки. Когда все четыре этапа пройдены успешно, процедура настройки считается успешной и, следовательно, прекращается.

V. ПОСТАНОВКА ЭКСПЕРИМЕНТА

A. Участники

Институциональный наблюдательный совет Северного университета Каролина в Чепел-Хилл одобрила этот документ. Один мужчина АВ субъект (возраст: 41 год, рост: 178 см, вес: 70 кг) и один мужчина, субъект с односторонней трансфеморальной ампутацией (TF) (возраст: 21 год лет, рост: 183 см, вес: 66 кг, время с момента ампутации: шесть лет) были приняты на работу. Оба испытуемых предоставили письменные и информированное согласие перед экспериментами.

B. Установка протеза и предметное обучение

Сертифицированный протезист выровнял роботизированный коленный протез для каждого предмета. Субъект TF использовал свой собственный разъем, а Субъект АВ использовал L-образный адаптер (с согнутой на 90 градусов одной ногой). градусов) ходить с роботизированным коленным протезом [49]. Каждый испытуемый посещал лабораторию как минимум для пяти двухчасовых тренировок, включая время подготовки и отдыха, чтобы попрактиковаться в ходьбе. с нашим роботизированным коленным протезом на оборудованной беговой дорожке (Bertec Corporation, Колумбус, Огайо, США) со скоростью 0,6 м/с. На первой тренировке параметры импеданса были настраиваются вручную на основе наблюдения за походкой субъекта и вербальная обратная связь субъекта, аналогичная процессу настройки в клинике. На второй тренировке физ. Терапевт научил субъекта TF уменьшать необоснованную компенсацию при ходьбе с роботизированным коленом. Предметы разрешили держаться за поручни беговой дорожки для равновесия, когда нужный. Субъект был готов к тестированию, как только он смог комфортно ходить в течение трех минут, не держась за поручни.

C. Протокол эксперимента

Мы провели три сессии тестирования (более трёх дней) для каждому предмету, чтобы оценить успеваемость наивного ADP и дополнительная четвертая сессия тестирования с субъектом TF для оценки производительности опытного ADP в настройка протезов.

1) Инициализация ADP-тюнера и параметров импеданса: ADP-тюнер не будет работать, если ANN и CNN инициализируются случайным образом. ADP-тюнер считается опытным, если ИНС и CNN были переведены из ранее успешного сессия. Мы случайным образом выбрали начальные параметры импеданса. из диапазона, полученного в наших предыдущих экспериментах, проведенных на 15 людях [13], но полученное колено движение не было оптимизировано для цели. Мы исключили исходные наборы параметров, которые: 1) не позволяли испытуемому ходить. не держась за поручни; 2) создан протез колена кинематика, которая была слишком близка к целевой кинематике колена (т. е. среднеквадратическая ошибка (RMSE) между этими двумя траектория коленей за один цикл ходьбы составляла менее 4 градусов); или 3) сгенерированные кинематические характеристики колена были вне безопасности диапазон.

2) Сеансы тестирования с помощью Naï ve ADP-Tuner: в каждом из В течение трех сеансов тестирования мы сначала предоставили испытуемому трехминутное время на акклиматизацию для ходьбы с вновь инициализированным наивный ADP-тюнер и параметры управления. Затем испытуемый ходил по беговой дорожке со скоростью 0,6 м/с не более семи раз. сегменты, каждый из которых длился не более 3-х минут ходьбы

периоды. После каждого сегмента следовал 3-минутный отдых. Эти периоды отдыха типичны в клинических условиях и предотвращают потенциальные мешающие эффекты усталости. Для всех периодов ходьбы мы записывали временные ряды данных кинематики коленного сустава от датчик угла и сила нагрузки от тензодатчика.

Первые 30 с первого периода ходьбы служили нашим условием «преднастройки», при которой ADP-тюнер был не включен. тем не менее, параметры импеданса оставались постоянными (т.е. исходными случайно выбранными параметрами импеданса). Последние 30 с их последнего периода ходьбы послужило нашим условием «пост-настройки» для оценки производительности, при котором ADP-тюнер был отключен, и параметры импеданса снова сохранились. константа (т.е. параметры импеданса были на конечном уровне) параметры, предоставляемые ADP-тюнером).

Во время всех остальных периодов ходьбы мы просили испытуемых последовательно ходить по беговой дорожке, пока ADP контроллер был включен и итеративно обновил протез параметры импеданса. Каждое обновление (определяемое как обучение ADP) итерация) выполнялась для каждых семи циклов походки. Как сказал ранее это было сделано для уменьшения ступенчатой ​​вариабельности коленного сустава. кинематические особенности угла пика и длительности фазы.

Мы приостанавливали работу ADP-тюнера во время каждого периода отдыха, чтобы предотвратить потеря полученной информации.

Мы завершили сеанс тестирования, когда один из двух были выполнены критерии остановки: 1) сессия тестирования достигла 70 итерации обучения, чтобы избежать утомления субъекта или 2) ошибок все четыре угловых характеристики находились в пределах соответствующего диапазона допуска μ для трех из пяти предыдущих методов обучения ADP. итерации.

3) Сеанс тестирования с опытным ADP-тюнером: чтобы оценить знание ранее изученных ADP-тюнеров.

Чтобы сделать обучение более эффективным, мы провели дополнительное тестирование с испытуемым TF в другой день с тем же протоколом. Вместо этого мы начали с опытного ADP, который использовал сетевые коэффициенты ANN и CNN. получено из предыдущего сеанса, который сгенерировал самый низкий RMSE.

D. Анализ данных

Данные временных рядов кинематики роботизированного колена были сегментированы на циклы походки на основе событий удара пяткой. (рис. 2), а затем были нормализованы до 100 образцов на походку. цикл.

Точность наивного и опытного ADP-тюнера оценивалось по среднеквадратическому отклонению между измеренной и целевой кинематикой колена и ошибками характеристик, полученными в каждом итерация настройки. Для сравнения предварительной настройки и последующей настройки производительность, усредненное среднеквадратичное отклонение кинематики колена и особенности ошибок 20 циклов походки при предварительной настройке и после настройки Условия были рассчитаны и сравнены.

Эффективность данных количественно оценивалась по количеству обучений. итераций в каждом сеансе тестирования. Эффективность времени количественно оценивалась по продолжительности ходьбы субъекта в каждом сеансе тестирования.

Наконец-то была продемонстрирована стабильность работы ADP-тюнера настроенными параметрами импеданса коленного сустава и кинематикой коленного сустава (т. е. RMSE и ошибками характеристик, усредненными по семи походкам). циклов внутри каждой итерации) на протяжении итераций обучения.

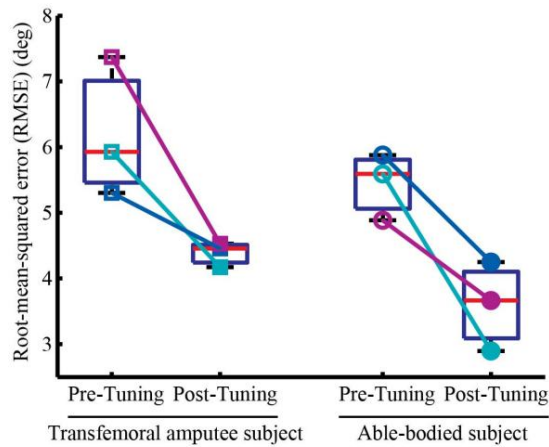


Рис. 3. Сравнение кинематики колена по RMSE между предварительной и последующей настройкой в ходе нескольких сеансов тестирования. Квадратные маркеры обозначают сеансы тестирования субъекта TF, а круговые маркеры представляют сеансы тестирования субъекта АВ. Открытый маркер представляет состояние предварительной настройки, а закрытый маркер представляет состояние после настройки.

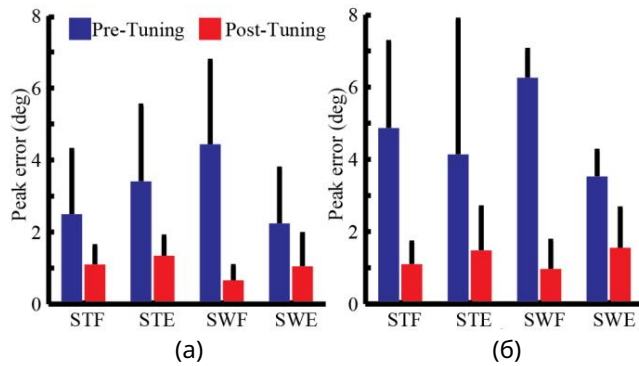


Рис. 4. Сравнение пиковых ошибок между условиями предварительной настройки и после настройки субъекта TF (а) и субъекта АВ (б) на каждой фазе. Каждая полоса представляет собой среднюю ошибку трех сеансов тестирования, а полоса ошибок обозначает одно стандартное отклонение от среднего значения.

VI. ПОЛУЧЕННЫЕ РЕЗУЛЬТАТЫ

В качестве меры точности ADP-тюнера среднеквадратическое отклонение угла роботизированного колена (по сравнению с целевым) в среднем по сеансам тестирования и испытуемым уменьшилось с $5,83 \pm 0,85$ градуса до $3,99 \pm 0,62$ градуса (рис. 3, результаты отдельных испытуемых). Все ошибки угловых характеристик уменьшились после настройки ADP-тюнером (рис. 4). Ошибки характеристик продолжительности не показали последовательной тенденции (рис. 5) для этих двух субъектов.

Эта изменчивость ошибок характеристик длительности не была неожиданностью, поскольку: 1) продолжительность каждой фазы частично контролируется пользователем человеческого протеза и 2) наш алгоритм ADP обеспечивал большую гибкость (или относительно больший допустимый диапазон) ошибок характеристик длительности, чем ошибки угловых характеристик для достижения цели и полной настройки.

В качестве меры эффективности данных и времени ADP-тюнеру потребовалось в среднем 43 ± 10 итераций обучения, чтобы найти «оптимальные» параметры импеданса, что в среднем составляет 300 циклов походки и 10 ± 2 минуты ходьбы. Данные и эффективность времени были одинаковыми для испытуемых (АВ: 45 ± 9 итераций и субъект с ампутированной конечностью: 41 ± 12 итераций).

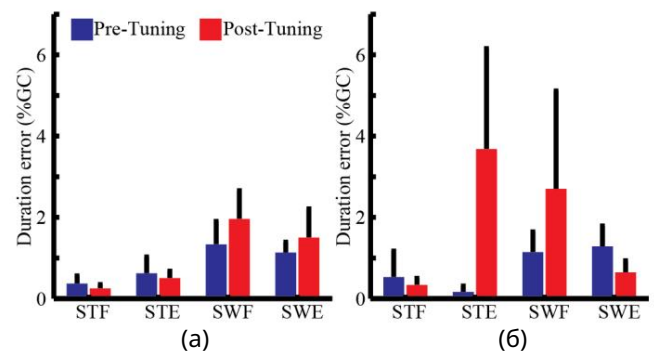


Рис. 5. Сравнение ошибок длительности между преднастройочными и постнастройочными состояниями субъекта TF (а) и субъекта АВ (б) для каждой фазы. Каждая полоса представляет собой среднюю ошибку трех сеансов тестирования, а полоса ошибок обозначает одно стандартное отклонение от среднего значения.

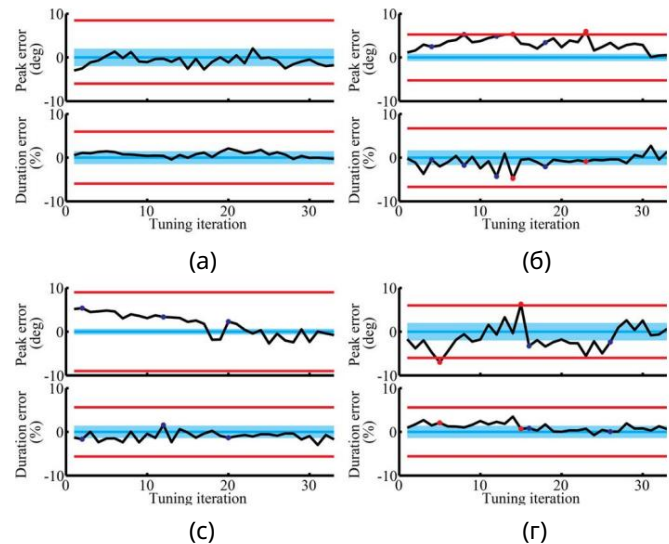


Рис. 6. Пиковая ошибка и ошибка длительности во время четырех фаз для типичной процедуры настройки. а – STF, б – STE, в – SWF и г – SWE. Красные точки обозначали времена, когда применялись сигналы подкрепления -1, а синие точки — времена, когда применялись сигналы подкрепления -0,8. Горизонтальные синие области с центром в нуле обозначают диапазоны допусков для каждого элемента. Парные горизонтальные красные линии обозначают разрешенные максимальные и минимальные пределы исследования для каждого объекта.

Как ошибки характеристик, так и параметры импеданса в целом стабилизировались в период после настройки (рис. 6 и 7, показаны репрезентативные испытания). В частности, стабилизировались как ошибки характеристик, так и параметры импеданса фаз походки сгибания и разгибания. Однако для сгибания и разгибания стойки ошибки характеристик стабилизировались, в то время как параметры импеданса все еще медленно менялись. Окончательные параметры импеданса, выбранные ADP-тюнером, чтобы позволить пользователю ходить с почти нормальным движением колена или целевым профилем колена, не были одинаковыми во всех сеансах тестирования (таблица I).

В целом параметры жесткости и параметры демпфирования на фазах опоры ($2,33 \pm 0,56$ Нм/град, $0,13 \pm 0,05$ Нм/град) были выше, чем на фазах поворота ($0,95 \pm 0,83$ Нм/град, $0,03 \pm 0,02$ Нм/град). В опытом сеансе тестирования ADP для всех четырех фаз походки ошибки как угла, так и продолжительности имели тенденцию к уменьшению к нулю (рис. 8(а) и (б)).

Значение J^* сети CNN уменьшалось в ходе настройки

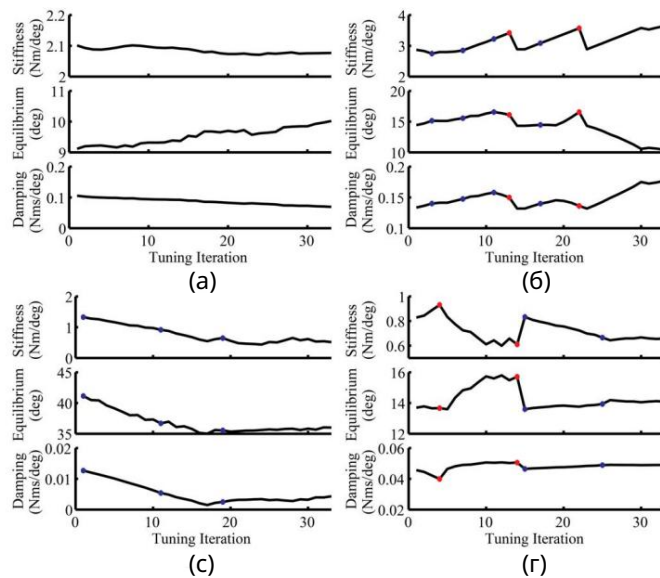


Рис. 7. Параметры импеданса четырех фаз во время типичной процедуры настройки. а – STF, б – STE, в – SWF и г – SWE. Значения красных и синих точек такие же, как на рис. 6.

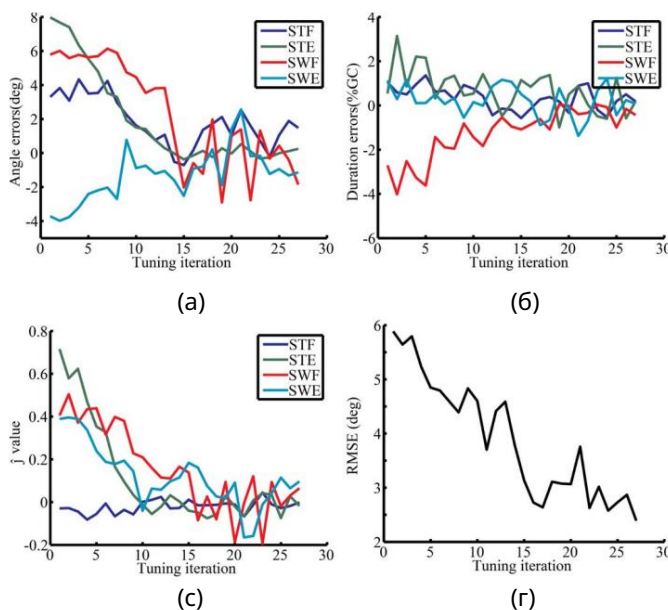


Рис. 8. Результаты онлайн-оценки изученного автотюнера ADP. (а) Тенденции угловой ошибки на протяжении итераций настройки. (б) Тенденции изменения продолжительности на итерациях настройки. (в) Изменение значений J^* по мере обучения. (г) RMSE на итерациях настройки.

итерация [рис. 8(с)], а среднеквадратическое отклонение кинематики роботизированного колена уменьшилось с 5,9 градусов до 2,5 градусов от до и после настройки [рис. 8(г)]. В этом случае опытному ADP-тюнеру потребовалось 28 итераций (около 7 минут), чтобы найти 12 оптимальных параметров импеданса. Во время этого сеанса тестирования с опытным/обученным ADP не возникло никакого дополнительного сигнала подкрепления.

VII. ОБСУЖДЕНИЕ

Целью данной статьи является исследование возможности создания нового RL-подход для персонализации управления

роботизированный коленный протез. Всего 12 параметров импеданса были настроены одновременно с помощью нашего ADP-тюнера для двух человек. Здесь мы рассмотрим последствия и остающиеся проблемы нашего предлагаемого подхода на основе RL для достижения нашей проектной цели по автоматической настройке роботизированных протезов для людей с ампутированными конечностями.

A. Реализуемость и надежность

Точность ADP-тюнера по достижению целевого профиля угла колена как для каждой фазы походки (рис. 4 и 5), так и для всего цикла походки (рис. 3) указывает на то, что с помощью ADP-тюнера можно оптимизировать большое количество протезов. параметры управления. В этом исследовании ADP-тюнер отрегулировал параметры импеданса, чтобы позволить обоим испытуемым последовательно ходить, приближая к нормальной кинематике колена. Кроме того, ADP-тюнер надежно воспроизводил аналогичные результаты для всех сеансов тестирования, каждый из которых начинался с разных, случайно инициализированных весовых матриц ANN и CNN (т. е. без предварительных знаний, встроенных в контроллер обучения) и параметров импеданса.

Вариации окончательных параметров импеданса после настройки ADP показали, что несколько комбинаций параметров импеданса дают аналогичную кинематику протеза (таблица I). Это неудивительно, поскольку согласно (2) крутящий момент двигателя недоопределен комбинацией трех параметров импеданса. Было бы интересно в будущем изучить оптимальную комбинацию параметров управления с учетом дополнительных ограничений или целей оптимизации.

Несмотря на то, что кинематика протезного колена измерялась исключительно с помощью протеза, она по своей сути представляла собой объединенное усилие как человека, так и машины или

контроллер протеза. Основываясь на наших результатах, мы постулируем, что пики сгибания/разгибания колена робота в первую очередь зависят от параметров импеданса и, следовательно, от нашего ADP-тюнера (рис. 4), но на продолжительность каждой фазы походки может доминировать пользователь-человек (рис. 5). Субъекты могли контролировать время своих шагов, вероятно, потому, что они могли контролировать, когда ставить и поднимать протез стопы на беговую дорожку или снимать ее с ипсилатерального бедра и всего тела.

При управлении роботизированными протезами с обратной связью сигналы обратной связи должны реагировать на управляющее воздействие. Поэтому мы считаем, что использование кинематики колена в качестве состояния обратной связи и оптимизации было разумным в качестве первого шага, но вопросы относительно соответствующей цели управления остаются открытыми. Мы планируем систематически изучать эту тему в будущих исследованиях.

B. Эффективность.

Начав без каких-либо предварительных знаний или модели установки, наш ADP-настройщик смог собрать информацию и понять, как одновременно настроить 12 параметров управления за 10 минут одного тестового сеанса или 300 циклов ходьбы для обоих. предметы. Для справки: передовой метод настройки экспертной системы требовал как минимум трех дней систематической записи решений по настройке экспертов-людей и передачи этих знаний в компьютерный алгоритм, которому затем потребовалось 96 циклов походки для настройки параметров импеданса [13]. Однако обратите внимание, что эта киберэкспертная система является субъективной (т. е. предвзятой в зависимости от опыта протезистов) и негибкой, когда система вводит данные.

ТАБЛИЦА I
ПАРАМЕТРЫ ПОСЛЕ НАСТРОЙКИ ИМПЕДАНСА ТРЕХ СЕАНСОВ ИСПЫТАНИЯ ДЛЯ ДВУХ ОБЪЕКТОВ

| Phase | Stance Flexion | | | Stance Extension | | | Swing Flexion | | | Swing Extension | | |
|----------------------------------|----------------|------------|-------|------------------|------------|-------|---------------|------------|-------|-----------------|------------|-------|
| Impedance Parameter ^a | k | θ_e | b | k | θ_e | b | k | θ_e | b | k | θ_e | b |
| TF S1 | 2.077 | 10.020 | 0.069 | 3.644 | 10.399 | 0.177 | 0.515 | 36.000 | 0.004 | 0.658 | 14.096 | 0.049 |
| TF S2 | 2.423 | 9.842 | 0.089 | 2.534 | 13.366 | 0.165 | 0.740 | 35.098 | 0.020 | 0.609 | 12.546 | 0.046 |
| TF S3 | 2.068 | 11.099 | 0.103 | 2.493 | 13.749 | 0.177 | 0.601 | 32.779 | 0.006 | 0.657 | 12.814 | 0.066 |
| AB S1 | 1.948 | 9.914 | 0.066 | 2.542 | 12.875 | 0.133 | 1.853 | 36.754 | 0.013 | 0.967 | 24.068 | 0.011 |
| AB S2 | 2.783 | 8.241 | 0.067 | 2.288 | 14.010 | 0.127 | 0.305 | 52.407 | 0.019 | 0.522 | 16.768 | 0.014 |
| AB S3 | 1.578 | 10.774 | 0.107 | 1.571 | 19.581 | 0.237 | 3.294 | 36.444 | 0.042 | 0.625 | 16.334 | 0.057 |

^a k (Nm/deg) is the stiffness coefficient; θ_e (deg) is the equilibrium position; b (Nms/deg) is the damping coefficient.

и изменения вывода. Наш ADP-тюнер объективен и гибок по структуре. Более того, опытному ADP-настройщику (т.е. обладающему некоторыми предварительными знаниями) потребовалось всего 210 циклов ходьбы без дополнительных поддерживающих сигналов для обучения, демонстрируя, что полученные знания могут быть эффективно перенесены для настройки параметров импеданса. Поэтому мы считаем, что ADP-тюнер позволяет эффективно использовать время и данные для потенциального клинического использования.

В повседневной жизни ADP-тюнер потенциально может обрабатывать медленные изменения, например, изменение веса тела. Для изменений требований окружающей среды, таких как переход от ходьбы по ровной земле к подъему по пандусу или лестнице, ADP-тюнер потенциально может найти оптимальные параметры управления для каждого режима передвижения (например, ходьба по пандусу и подъем по лестнице), что может занять больше времени, но может хранить параметры импеданса и переключать параметры, когда пользователь сталкивается с изменением задачи в реальной жизни. Мы изучим это в нашей дальнейшей работе.

C. Результат обучения

ADP-тюнер обучался посредством поддерживающих сигналов (рис. 6 и 7, цветные точечные символы) и мог настраивать параметры импеданса, что, в свою очередь, уменьшало ошибки угловых характеристик, чтобы соответствовать соответствующей допуску на ошибку. В конце процедуры настройки ошибки функций также сохранялись в пределах допуска по крайней мере в течение трех из пяти предыдущих итераций обучения ADP, чтобы завершить сеанс настройки.

Ошибки признаков явно сходились к нулю в двух из четырех фаз (рис. 6(в) и (г)) и соответствующие параметры импеданса [рис. 7(с) и (d)] стабилизировались. Эти результаты показывают, что ADP-тюнер способен генерировать конвергентную политику для этих фаз походки. Однако на оставшихся двух этапах параметры импеданса все еще адаптировались, но ошибки характеристик находились в пределах допустимых значений. Эти результаты заставляют нас полагать, что ошибки характеристик не очень чувствительны к определенным параметрам импеданса или комбинациям параметров. Это явление также может быть вызвано нашими критериями остановки, составляющими максимум 70 итераций настройки, которые применяются для того, чтобы сделать эту статью практичной для клинического применения и предотвратить утомление людей с ампутированными конечностями. В будущем, чтобы быстро достичь конвергентной политики, мы могли бы решить эту проблему, добавляя небольшие искажения к параметрам импеданса, когда ошибки функции приближаются к нулю, чтобы проверить свойства конвергенции ADP-тюнера и позволяя ADP-тюнеру накапливать больше обучающий опыт.

Наконец, мы продемонстрировали, что опытный ADP-настройщик, после всего лишь взаимодействия с системой человеческого протеза в течение одного сеанса тестирования, эффективно усвоил знания по настройке для

достиж цельевой кинематики колена. Поскольку как человеческое влияние, так и межфазное влияние способствуют движению роботизированного колена [49], мы ожидали, что ошибки как угла, так и продолжительности будут колебаться около нуля [рис. 8(а) и (б)]. Кроме того, опытный ADP настраивал параметры управления протезом быстрее, чем наивный ADP. Этот потрясающий результат открывает возможность сделать наш контроллер протеза адаптивным к пользователям в их повседневной жизни.

D. Значение результатов

В этой статье ADP-тюнер не имел предварительных знаний о: 1) структуре контроллера импеданса и 2) механических свойствах роботизированного коленного протеза. Единственной информацией, которую наблюдал ADP, было состояние системы человек-протез посредством измерений угла протезного колена и сигналов подкрепления, когда характеристики/функции выходили за пределы допустимого диапазона исследования.

Таким образом, конструкция ADP-тюнера потенциально может быть применена к протезам коленного сустава с различными механическими конструкциями и методами управления и даже, возможно, расширена до задачи настройки параметров управления для протезов голеностопного сустава и экзоскелетов.

Кроме того, наш метод может быть применен к другим целям контроля для достижения поведенческих целей. Например, если целевая кинематика колена заключается в создании большего угла сгибания при повороте для обеспечения просвета стопы, опытный ADP-настройщик потенциально может быстро настроить параметры импеданса для достижения новой цели. Таким образом, наша политика обучения управлению может значительно улучшить процесс настройки/персонализации роботизированных протезов, а также адаптируемость протеза к изменениям внутри пользователя и его окружающей среды.

E. Ограничения и будущая работа

В этой статье мы сосредоточились на демонстрации возможности управления на основе RL для автоматической настройки роботизированных протезов. Люди различаются по своему физическому состоянию и поведению, и в своей повседневной жизни они взаимодействуют с разными ландшафтами. Чтобы раскрыть все возможности этого многообещающего подхода, нам необходимо дополнительно оценить ADP-тюнер с участием большего количества людей и большего количества двигательных задач (например, подъем или спуск по лестнице и ходьба по траве), чтобы закрепить надежность этого RL-основанный подход.

Еще одним ограничением является то, что в этом исследовании для проверки концепции мы выбрали в качестве цели настройки нормативную кинематику протезного коленного сустава, которая не может быть идеально согласована с каждым вариантом.

предпочтения человека с ампутированной конечностью и общая эффективность системы человек-протез. Однако с точки зрения настройки параметров эта статья доказала, что ADP-тюнер может автоматически найти набор параметров импеданса для достижения заданной цели настройки. Будущие исследования будут сосредоточены на определении желаемых целевых характеристик колена, которые улучшают характеристики походки человека с протезом, таких как индекс симметрии походки, запас устойчивости или даже субъективные предпочтения пользователя.

VIII. ЗАКЛЮЧЕНИЕ

В данной работе мы обеспечили существенный скачок вперед от традиционной трудоемкой и трудоемкой ручной настройки параметров управления протезом. Мы разработали новый подход к управлению на основе RL для автоматической настройки 12 параметров импеданса роботизированного коленного протеза. Новая концепция была проверена на одном субъекте АВ и одном TF в ходе нескольких сеансов тестирования. Многообещающие результаты показали, что ADP-тюнер является возможным и безопасным методом автоматической настройки большого количества параметров управления в рамках данной статьи. Алгоритм эффективно обучается посредством взаимодействия с системой «человек-протез» в режиме реального времени, без каких-либо предварительных знаний по настройке со стороны обученного врача или эксперта на местах. Обучение также не требует предварительной модели системы человека-протеза.

Результаты этой статьи могут привести к созданию новой системы управления протезами, позволяющей персонализировать роботизированные протезы, оптимизировать работу системы человек-протез при походке и сделать протез адаптивным к пользователям. В будущем мы продолжим изучать конструкции управления обучением, которые автоматически соответствуют различным целям производительности человеческого протеза, таким как симметрия походки, стабильность и даже восприятие пользователя.

ПОДТВЕРЖДЕНИЕ

Авторы выражают благодарность М. Лю, доктору философии. и С. Хуанг, доктор философии. за помощь в планировании эксперимента Д. Франкене, СРО, и М. Соярсу, РТ по обучению субъектов и установке эксперимента, а также обоим участникам.

ИСПОЛЬЗОВАННАЯ ЛИТЕРАТУРА

[1] Дж. К. Хитт, Т. Г. Шугар, М. Холгейт и Р. Беллман, «Активный протез стопы-лодыжки с биомеханической регенерацией энергии», J. Med. Устройство., вып. 4, нет. 1 марта 2010 г., ст. нет. 011003.

[2] Л. Амброзич и др., «CYBERLEGS: ориентированный на пользователя роботизированный трансформальный протез с контролем всего тела», IEEE Robot. Автомат. Маг., вып. 21, нет. 4, стр. 82–93, декабрь 2014 г.

[3] Ф. Суп, Х. А. Варол, Дж. Митчелл, Т. Дж. Уитроу и М. Гольдфарб, «Предварительные оценки автономного антропоморфного трансбедренного протеза», IEEE/ASME Trans. Мехатроника, вып. 14, нет. 6, стр. 667–676, декабрь 2009 г.

[4] С. К. Ау, Дж. Вебер и Х. Херр, «Электрический протез лодыжки и стопы улучшает метаболическую экономию при ходьбе», IEEE Trans. Робот., вып. 25, нет. 1, стр. 51–66, февраль 2009 г.

[5] Э. Дж. Роуз, Л. М. Муни и Х. М. Герр, «Сцепляемый последовательно-эластичный привод: значение для конструкции протеза коленного сустава», Int. Джей Робот. Рез., том. 33, нет. 13, стр. 1611–1625, октябрь 2014 г.

[6] М. Лю, Ф. Чжан, П. Дацерис и Х. Хуанг, «Улучшение конечного управления импедансом активного трансбедренного протеза с использованием правил перехода состояний на основе Демпстера-Шейфера», J. Intell. l. Робот. сист., вып. 76, нет. 3–4, стр. 461–474, декабрь 2014 г.

[7] А. М. Саймон и др., «Конфигурация механического протеза колена и голеностопного сустава для людей с трансформальной ампутацией в пяти конкретных режимах передвижения», PLoS ONE, vol. 9, нет. 6 июня 2014 г., ст. нет. e99387.

[8] А. Брандт, Ю. Вэнь, М. Лю, Дж. Столлинкс и Х. Х. Хуанг, «Взаимодействие между людьми с трансформальной ампутированной конечностью и механическим коленным протезом во время перевозки грузов», Sci. Респ., вып. 7, нет. 1 ноября 2017 г., ст. нет. 14480.

[9] С. Пфайфер, Х. Вальери, М. Хардеггер, Р. Ринер и Э. Дж. Перро, «Оценка жесткости колена на основе модели», IEEE Trans. Биомед. англ., вып. 59, нет. 9, стр. 2604–2612, сентябрь 2012 г.

[10] Э. Дж. Роуз, Л. Дж. Харгроув, Э. Дж. Перро и Т. А. Куикен, «Оценка импеданса голеностопного сустава человека во время фазы опоры при ходьбе», IEEE Trans. Нейронная система. Реабилитация. англ., вып. 22, нет. 4, стр. 870–878, июль 2014 г.

[11] М. Р. Такер, К. Широта, О. Ламберси, Дж. С. Зульцер и Р. Гассерт, «Проектирование и характеристика экзоскелета для возмущения колена во время ходьбы», IEEE Trans. Биомед. англ., вып. 64, нет. 10, стр. 2331–2343, октябрь 2017 г.

[12] Р.Д. Грегг, Т. Лензи, Л.Дж. Харгроув и Дж.В. Сенсингер, «Управление виртуальным ограничением силового протеза ноги: от моделирования до экспериментов с людьми с трансформальной ампутацией», IEEE Trans. Робот., вып. 30, нет. 6, стр. 1455–1471, декабрь 2014 г.

[13] Х. Хуанг, Д.Л. Крауч, М. Лю, Г.С. Савицкий и Д. Ван, «Киберэкспертная система для автоматической настройки параметров управления импедансом протеза с электроприводом», Ann. l. Биомед. англ., вып. 44, нет. 5, стр. 1613–1624, сентябрь 2016 г.

[14] Дж. Р. Коллер, Д. Х. Гейтс, Д. П. Феррис и К. Д. Реми, «Оптимизация вспомогательных роботизированных устройств «Тело в цикле»: проверочное исследование», в Proc. Робот. наук. Сист. XII, Анн-Арбор, Мичиган, США, июнь 2016 г., стр. 1–10.

[15] Дж. Чжан и др., «Оптимизация помощи экзоскелета при ходьбе человеком в процессе ходьбы», Science, vol. 356, нет. 6344, стр. 1280–1284, июнь 2017 г. [Онлайн]. Доступно: <http://www.sciencemag.org/lookup/doi/10.1126/science.aal5054> [16] Ю. Динг, М. Ким, С. Куиндерсма и К. Дж. Уолш, «Человеческая оптимизация в цикле поддержка бедра с помощью мягкого экзоскостюма во время ходьбы», Sci. Робот., вып. 3, нет. 15 февраля 2018 г., ст. нет. eaar5438.

[17] Д.П. Берцекас и Ю.Н. Цицелис, Нейро-динамическое программирование. Бельмонт, Массачусетс, США: Athena Sci., 1996.

[18] Дж. Си, А. Г. Барто, У. Б. Пауэлл и Д. Вунш, Справочник по обучению и приближенному динамическому программированию. Хобокен, Нью-Джерси, США: Wiley, 2004.

[19] У.Б. Пауэлл, Приближительное динамическое программирование: решение проклятий размерности, 2-е изд., Д. Болдинг, изд. Хобокен, Нью-Джерси, США: Wiley, 2011.

[20] Дж. Вербос, «Меню проектов для обучения с подкреплением с течением времени», в книге «Нейронные сети для управления». Кембридж, Массачусетс, США: MIT Press, 1990, гл. 3, стр. 67–95.

[21] П.Дж. Вербос, «За пределами регрессии: новые инструменты прогнозирования и анализа в поведенческих науках», Ph.D. диссертация, Гарвардский университет, Кембридж, Массачусетс, США, 1974.

[22] П.Дж. Вербос, «Создание и понимание адаптивных систем: статистический/числовой подход к автоматизации производства и исследованиям мозга», IEEE Trans. Сист. Человек Киберн., том. 17, нет. 1, стр. 7–20, январь 1987 г.

[23] Д. Ван, Х. Хе и Д. Лю, «Адаптивный критический нелинейный устойчивый контроль: обзор», IEEE Trans. Киберн., вып. 47, нет. 10, стр. 3429–3451, октябрь 2017 г.

[24] Х. Чжан, Л. Цуй, С. Чжан и Ю. Луо, «Управляемое данными устойчивое приближенное нелинейное отслеживание управления для неизвестных общих нелинейных систем с использованием метода адаптивного динамического программирования», IEEE Trans. Нейронная сеть., том. 22, нет. 12, стр. 2226–2236, декабрь 2011 г.

[25] Х. Чжан, Дж. Чжан, Г.-Х. Ян и Ю. Луо, «Оптимальное координационное управление на основе лидера для задачи консенсуса многоагентных дифференциальных игр посредством нечеткого адаптивного динамического программирования», IEEE Trans. Нечеткая система., вып. 23, нет. 1, стр. 152–163, февраль 2015 г.

[26] Дж. Чжан, Х. Чжан и Т. Фэн, «Распределенное оптимальное консенсусное управление для нелинейной мультиагентной системы с неизвестной динамикой», IEEE Trans. Нейронная сеть. Учитесь. сист., вып. 29, нет. 8, стр. 3339–3348, август 2018 г.

[27] Д. Ван и Д. Лю, «Обучение и гарантированный контроль затрат с помощью реализации адаптивной критики на основе событий», IEEE Trans. Нейронная сеть. Учитесь. сист., вып. 29, нет. 12, стр. 6004–6014, декабрь 2018 г.

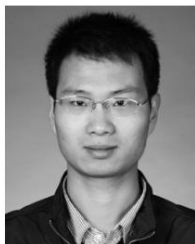
[28] Д. Ван, Х. Хе, К. Чжун и Д. Лю, «Управляемое событиями нелинейное оптимальное регулирование с дисконтированием с использованием приложения энергосистемы», IEEE Trans. Индийская электрон., вып. 64, нет. 10, стр. 8177–8186, октябрь 2017 г.

[29] Ю. Ван, В. К. Чжэн и Х. Чжан, «Динамическое событийно-ориентированное управление нелинейными стохастическими системами», IEEE Trans. Автомат. Контроль, том. 62, нет. 12, стр. 6544–6551, декабрь 2017 г.

30. Прохоров Д.В., Вунш Д.К. Адаптивные критические конструкции // IEEE Trans. Нейронная сеть., том. 8, нет. 5, стр. 997–1007, сентябрь 1997 г.

[31] М. Ридмиллер, «Итерация Q с нейронной подгонкой — первый опыт эффективного метода обучения с нейронным подкреплением», в Proc. 16 евро. Конф. Мах. ЖЖ, 2005, стр. 317–328.

- [32] Ж. Си и Ю.-Т. Ван, «Контроль онлайн-обучения посредством ассоциации и подкрепления», IEEE Trans. Нейронная сеть., том. 12, нет. 2, стр. 264–276, март 2001 г.
- [33] М.И. Джордан и Р.А. Джейкобс, «Обучение управлению нестабильной системой с помощью прямого моделирования», в сб. Proc. Adv. Нейронная инф. Процесс. Сист., 1990, стр. 324–331.
- [34] Р. Хафнер и М. Ридмиллер, «Обучение с подкреплением в контроле с обратной связью: проблемы и ориентиры в области управления техническими процессами», Mach. Учись., вып. 84, нет. 1–2, стр. 137–169, июль 2011 г.
- [35] М. Ридмиллер, М. Монтемерло и Х. Далкамп, «Научимся водить настоящую машину за 20 минут», в Proc. Передний. Конверг. Биология. Инф. Технол., 2007, стр. 645–650.
- [36] М. Ридмиллер, Т. Габель, Р. Хафнер и С. Ланге, «Обучение с подкреплением для роботизированного футбола», Auton. Роботы, вып. 27, нет. 1, стр. 55–73, июль 2009 г.
- [37] Т. Габель и М. Ридмиллер, «Адаптивное реактивное планирование цеха с помощью агентов обучения с подкреплением», Int. Ж. Инф. Технол. Интел. Вычисл., вып. 24, нет. 4, стр. 1–30, 2008 г.
- [38] Р. Хафнер и М. Ридмиллер, «Контроллеры обучения с нейронным подкреплением для реального приложения робота», в Proc. IEEE Международный. Конф. Робот. Autom., Рим, Италия, апрель 2007 г., стр. 2098–2103.
- [39] Р. Эннс и Дж. Си, «Управление балансировкой и отслеживанием вертолета с использованием прямого нейронного динамического программирования», IEEE Trans. Нейронная сеть., том. 14, нет. 4, стр. 929–939, июль 2003 г.
- [40] Л. Янг, Дж. Си, К. С. Цакалис и А. Родригес, «Прямое эвристическое динамическое программирование для нелинейного управления отслеживанием с фильтрацией ошибок отслеживания», IEEE Trans. Сист. Мужик. Киберн. Б, Киберн., том. 39, нет. 6, стр. 1617–1622, декабрь 2009 г.
- [41] К. Лу, Дж. Си и К. Се, «Прямое эвристическое динамическое программирование для гашения колебаний в большой энергосистеме», IEEE Trans. Сист. Мужик. Киберн. Б, Киберн., том. 38, нет. 4, стр. 1008–1013, август 2008 г.
- [42] Р. Эннс и Дж. Си, «Реконфигурация управления полетом вертолета при отказах привода несущего винта», J. Guid. Контроль. Дин., вып. 26, нет. 4, стр. 572–584, июль 2003 г.
- [43] W. Guo и др., «Проектирование и применение дополнительного контроллера ADP для онлайн-обучения для управления частотой энергосистемы с крупномасштабной интеграцией энергии ветра», IEEE Trans. Нейронная сеть. Учиться. сист., вып. 27, нет. 8, стр. 1748–1761, август 2016 г.
- [В сети]. Доступно: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=7124490>.
- [44] Ю. Вэнь, Дж. Си, С. Гао, С. Хуан и Х. Х. Хуан, «Новая система управления протезами нижних конечностей с электроприводом, основанная на адаптивном динамическом программировании», IEEE Trans. Нейронная сеть. Учиться. сист., вып. 28, нет. 9, стр. 2215–2220, сентябрь 2017 г.
- [45] С. Л. Дельп и др., «OpenSim: программное обеспечение с открытым исходным кодом для создания и анализа динамического моделирования движения», IEEE Trans. Биомед. англ., вып. 54, нет. 11, стр. 1940–1950, ноябрь 2007 г.
- [46] Б. Э. Лоусон, Х. А. Варол, А. Хафф, Э. Эрдемр и М. Гольдфарб, «Управление подъемом и спуском по лестнице с помощью трансформального протеза с электроприводом», IEEE Trans. Нейронная система. Реабилитация. англ., вып. 21, нет. 3, стр. 466–473, май 2013 г.
- [47] М. П. Кадаба, Х. К. Рамакришнан и М. Е. Вуттен, «Измерение кинематики нижних конечностей во время горизонтальной ходьбы», Дж. Ортоп. Рез., том. 8, нет. 3, стр. 383–392, май 1990 г.
- [48] Ф. Лю, Дж. Сунь, Дж. Си, В. Го и С. Мэй, «Результат об ограниченности для прямого эвристического динамического программирования», Neural Netw., vol. 32, стр. 229–235, август 2012 г., doi: [10.1016/j.neunet.2012.02.005](https://doi.org/10.1016/j.neunet.2012.02.005).
- [49] Ю. Вэнь, М. Лю, Дж. Си и Х. Х. Хуанг, «Адаптивное управление приводными трансформальными протезами на основе адаптивного динамического программирования», в Proc. 38-й год. Межд. Конф. IEEE англ. Мед. Биол. Soc., Орландо, Флорида, США, 2016 г., стр. 5071–5074.



Юэ Вэнь получил степень бакалавра в области автоматизации в Уханьском технологическом университете, Ухань, Китай, в 2011 году и степень магистра в области теории управления и техники в Уацжунском университете науки и технологий, Ухань, в 2014 году. В настоящее время он учится в доктор философии Степень в Объединенном факультете биомедицинской инженерии NCSU/UNC, Университете штата Северная Каролина, Роли, Северная Каролина, США, и Университете Северной Каролины в Чапел-Хилл, Чапел-Хилл, Северная Каролина, США.

Его текущие исследовательские интересы включают адаптивное управление бионическими конечностями и вспомогательные роботизированные устройства, машинное обучение и анализ движений человека.



Дженни Си (F08) получила степени бакалавра и магистра электротехники в Университете Цинхуа, Пекин, Китай, а также степень доктора философии. Степень Университета Нотр-Дам, Нотр-Дам, Индиана, США.

С 1991 года она работала преподавателем кафедры электротехники Университета штата Аризона, Темпе, Аризона, США. Она консультировала Intel, Санта-Клара, Калифорния, США, Государственную службу Аризоны, Феникс, Аризона, США, и Medtronic, Дублин, Ирландия. В настоящее время ее научные интересы включают управление на основе обучения с подкреплением, а именно адаптивное динамическое программирование

с использованием методов машинного обучения и аппроксимации нейронных сетей. Она также занимается фундаментальными нейробиологическими исследованиями лобной коры и ее роли в обучении поведению с использованием электрофизиологических методов.

Доктор Си была лауреатом премии NSF/президентского факультета Белого дома в 1995 году и премии Motorola за выдающиеся инженерные достижения в 1995 году. Она является почетным преподавателем Общества вычислительной разведки IEEE.

Она работала в исполнительных советах нескольких профессиональных организаций и комитетах международных конференций. С 2009 по 2012 год она была вице-президентом по образованию Общества вычислительной разведки IEEE. Она была советником Директории по социальному поведению и экономике NSF. Она работала в нескольких комиссиях по рассмотрению предложений. В прошлом она работала заместителем редактора журналов IEEE TRANSACTIONS ON SEMICONDUCTOR MANUFACTURING и IEEE TRANSACTIONS ON AUTOMATIC CONTROL, а также редактором журнала Neural Networks. В настоящее время она является заместителем редактора журнала IEEE ТРАНЗАКЦИИ НА НЕЙРОННЫХ СЕТЯХ.



Андреа Брандт получила степень бакалавра наук. Степень по математике получила в Университете Северной Каролины в Чапел-Хилл, Чапел-Хилл, Северная Каролина, США, в 2013 году.

В настоящее время она работает над докторской диссертацией. Степень получила на Объединенном факультете биомедицинской инженерии NCSU/UNC, Университете штата Северная Каролина, Роли, Северная Каролина, США, и Университете Северной Каролины в Чапел-Хилл.

Ее текущие исследовательские интересы включают в себя протезы, подхода для людей с ампутированными конечностями и реабилитация.



Сян Гао получил степень бакалавра в области автоматизации в Университете науки и технологий Уацжун, Ухань, Китай, в 2011 году и степень магистра в Университете штата Аризона, Темпе, Аризона, США, в 2014 году, где в настоящее время работает над докторской диссертацией. Степень в Школе электротехники, вычислительной техники и энергетики.

Его текущие исследовательские интересы включают машинное обучение, нейронные сети и реабилитационную робототехнику.



Он (Хелен) Хуан (S'03-M'06-SM'12) получил степень доктора философии. Степень в области биомедицинской инженерии, Университет штата Аризона, Темпе, Аризона, США.

Она работала научным сотрудником в области нейронной инженерии в Реабилитационном институте Чикаго/Северо-Западном университете, Эванстон, Иллинойс, США. В настоящее время она является профессором Объединенного факультета биомедицинской инженерии NCSU/UNC и директором отдела разработки замкнутого цикла для усовершенствованной реабилитации Университета штата Северная Каролина, Роли, Северная Каролина, США, и Университета Северной Каролины в Чапел-Хилл, Чапел-Хилл, Северная Каролина, США.

В настоящее время ее исследовательские интересы включают нейро-машинные интерфейсы для протезов конечностей и экзоскелетов, взаимодействие человека и робота, адаптивное и оптимальное управление носимыми роботами и управление движениями человека.

Доктор Хуанг была лауреатом премии Delsys за инновации в электромиографии, стипендии Мэри Э. Свитцер совместно с NIDILRR, премии NSF CAREER Award, а также была названа стипендиатом факультета штата Северная Каролина в 2015 году. Она также является членом Общества нейробиологии и нейробиологии. Общество биомедицинской инженерии.