

doi: 10.17586/2226-1494-2023-23-4-786-794

УДК 004.9

Оптимизация систем отслеживания человека в виртуальной реальности на основе нейросетевого подхода

Артём Дмитриевич Обухов¹, Даниил Вячеславович Теселкин²

^{1,2} Тамбовский государственный технический университет, Тамбов, 392000, Российская Федерация

¹ obuhov.art@gmail.com, <https://orcid.org/0000-0002-3450-5213>

² teselkin.dv@tstu.ru, <https://orcid.org/0000-0002-1304-9490>

Аннотация

Введение. Рассмотрена проблема определения оптимального количества и расположения точек отслеживания на теле человека для обеспечения необходимой точности реконструкции кинематических параметров движений человека в виртуальном пространстве. Выполнена оптимизация системы отслеживания человека в виртуальной реальности для снижения объема передаваемой информации, вычислительной нагрузки и стоимости систем захвата движений за счет снижения количества физических датчиков. **Метод.** Поставлена задача оптимизации количества и расположения точек отслеживания на теле человека, необходимых для реконструкции виртуальной модели тела из ограниченного набора входных точек, с использованием численной аппроксимации функции регрессии. Разработан алгоритм сбора большого количества данных с модели тела человека в виртуальной сцене и с костюма захвата движений в реальном мире. **Основные результаты.** Получено наименьшее количество точек отслеживания тела человека и их расположение с использованием предложенного алгоритма. Обучены и протестированы различные топологии нейронных сетей, позволяющие аппроксимировать регрессионную зависимость между ограниченным по размеру вектором точек отслеживания (от 3 до 13) и вектором 18 виртуальных точек, используемых для полной реконструкции модели тела человека. Необходимая точность реконструкции кинематических параметров движений человека обеспечено при 5 и 7 входных точках.

Обсуждение. Предложенный подход позволил использовать 5 или 7 физических датчиков для построения модели тела человека и восстановления кинематических параметров его движений в виртуальной реальности. Подход может быть применен при решении задач инверсной кинематики с целью снижения количества физических датчиков, размещенных на поверхности исследуемого объекта, для упрощения процессов обработки и передачи информации. За счет объединения данных как с костюма захвата движений, так и с виртуального аватара значительно ускорен процесс сбора информации, расширен объем обучающей выборки и смоделированы различные паттерны движений тел пользователей.

Ключевые слова

виртуальная реальность, отслеживание движений человека, инверсная кинематика, оптимизация систем отслеживания и захвата движений, цифровое представление человека

Благодарности

Исследование выполнено за счет гранта Российского научного фонда (№ 22-71-10057, <https://rscf.ru/project/22-71-10057/>).

Ссылка для цитирования: Обухов А.Д., Теселкин Д.В. Оптимизация систем отслеживания человека в виртуальной реальности на основе нейросетевого подхода // Научно-технический вестник информационных технологий, механики и оптики. 2023. Т. 23, № 4. С. 786–794. doi: 10.17586/2226-1494-2023-23-4-786-794

Optimization of human tracking systems in virtual reality based on a neural network approach

Artem D. Obukhov¹, Daniil V. Teselkin²

^{1,2} Tambov State Technical University, Tambov, 392000, Russian Federation

¹ obuhov.art@gmail.com, <https://orcid.org/0000-0002-3450-5213>

² teselkin.dv@tstu.ru, <https://orcid.org/0000-0002-1304-9490>

© Обухов А.Д., Теселкин Д.В., 2023

Abstract

The problem of determining the optimal number and location of tracking points on the human body to ensure the necessary accuracy of reconstruction of kinematic parameters of human movements in virtual space is considered. Optimization of the human tracking system in virtual reality has been performed to reduce the amount of transmitted information, computational load and cost of motion capture systems by reducing the number of physical sensors. The task of optimizing the number and location of tracking points on the human body necessary for the reconstruction of a virtual body model from a limited set of input points using numerical approximation of the regression function is set. An algorithm has been developed for collecting a large amount of data from a human body model in a virtual scene and from a motion capture suit in the real world. The smallest number of human body tracking points and their location were obtained using the proposed algorithm. Various neural network topologies have been trained and tested to approximate the regression relationship between a vector of tracking points limited in size (from 3 to 13) and a vector of 18 virtual points used for the complete reconstruction of the human body model. The necessary accuracy of reconstruction of kinematic parameters of human movements is provided at 5 and 7 input points. The proposed approach made it possible to use 5 or 7 physical sensors to build a model of the human body and restore the kinetic parameters of its movements in virtual reality. The approach can be applied to solving inverse kinematics problems in order to reduce the number of physical sensors placed on the surface of the object under study, to simplify the processing and transmission of information. By combining data from both the motion capture suit and the virtual avatar, the process of collecting information has been significantly accelerated, the volume of the training sample has been expanded and various patterns of user body movements have been modeled.

Keywords

virtual reality, human movement tracking, inverse kinematics, optimization of motion tracking and capture systems, digital representation of a person

Acknowledgements

The research was carried out at the expense of the grant of the Russian Science Foundation (no. 22-71-10057, <https://rscf.ru/project/22-71-10057/>).

For citation: Obukhov A.D., Teselkin D.V. Optimization of human tracking systems in virtual reality based on a neural network approach. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2023, vol. 23, no. 4, pp. 786–794 (in Russian). doi: 10.17586/2226-1494-2023-23-4-786-794

Введение

На текущем этапе развития систем виртуальной реальности (VR) одной из актуальных проблем является точное воспроизведение тела пользователя в виртуальном пространстве [1]. Под виртуальным пространством подразумевается смоделированная среда, имитирующая реальное или вымышленное окружение посредством трехмерных моделей, анимации, звуковых и визуальных эффектов, а также обеспечивающее пользователю возможность взаимодействия с виртуальными объектами с помощью специального оборудования, например систем VR [2]. Базовый комплект поставки большинства систем VR включает шлем и два контроллера для перемещения и организации взаимодействия с объектами. Отслеживание действий может производиться по камерам, установленным в шлеме, с использованием внешних датчиков и базовых станций (технология Lighthouse) или с применением костюмов захвата движений. В первом случае возможно использование технологий компьютерного зрения, однако они обеспечивают качественное воспроизведение только рук, в частности ладоней и пальцев, распознавание остальных частей тела из-за специфики расположения камер затруднено. Использование технологии Lighthouse позволяет точно позиционировать шлем, контроллеры и дополнительные датчики (трекеры). Недостатком данного подхода является неудобство одновременного закрепления на поверхности тела человека большого количества датчиков. Третий подход основан на использовании дорогостоящих и зачастую чувствительных к внешним электромагнитным помехам костюмов с множеством инерциальных датчиков, что значительно ограничивает их использование в массовых решениях.

Разработчики программного обеспечения для VR осуществляют решение проблемы визуализации тела человека в виртуальном пространстве двумя основными методами [3]: с использованием прямой и инверсной кинематики. В первом случае требуется достаточный объем информации обо всех ключевых точках тела пользователя, так как кинематика моделируется сверху вниз, начиная, например от плеча, далее к локтю и заканчивая кистью. В инверсной кинематике, наоборот, перемещение дочерних сегментов приводит к изменению положения родительских, т. е. алгоритм рассчитывает положение и ориентацию плеча и локтя, исходя из положения и ориентации кисти. Первый подход используется в костюмах захвата движений, второй — в условиях ограниченного объема информации о точках тела человека, например, когда разработчику доступно только положение кистей пользователя, по которому восстанавливаются координаты точек рук, туловища и ног (учитывая текущую высоту шлема).

В настоящей работе рассмотрена проблема отслеживания положения тела человека для последующей реконструкции в VR кинематических параметров его движений с учетом минимизации количества отслеживаемых точек.

Обзор существующих исследований по отслеживанию человека в VR

На первом этапе рассмотрены существующие направления исследований в данной области.

В работе [4] использовано пять дополнительных закрепленных на спине, пятках и локтях человека трекеров, что в совокупности с двумя контроллерами и шлемом составляет 8 ключевых точек, на основе которых

разработчики решили задачу инверсной кинематики. Полученная система обладает достаточно низкой задержкой и высокой точностью по субъективной оценке участников контрольной группы.

Работа [5] основана на анализе походки человека и сопоставлении положения головы в шлеме VR с фазами шага, что позволило использовать только одну точку для восстановления кинематики ног с использованием различных подходов (Threshold, метода на основе корреляции Пирсона, Support Vector Machine, Bidirectional Long Short-Term Memory (BLSTM)). Архитектура BLSTM обеспечила наименьшую погрешность, однако недостаточно производительна для применения в режиме реального времени. По этой причине наиболее предпочтителен метод, основанный на расчете корреляции Пирсона между ускорением головы и ног. Данный подход не является универсальным и ограничен сценарием равномерной ходьбы.

Ряд научных работ сконцентрирован на более точной реконструкции только отдельных частей тела, например рук и пальцев. Для этого могут использоваться комбинации из перчаток VR и дополнительных акселерометров для каждого пальца [6] или сочетание камер с датчиком глубины и нейронных сетей для более точного восстановления трехмерных координат [7]. Нейронные сети также применены для успешной реконструкции трехмерной модели тела человека по двумерному изображению, на котором сначала распознаются основные точки скелета, а потом осуществляется их преобразование в трехмерное представление модели тела с коррекцией углов поворота сегментов [8].

Применение методов машинного обучения для аппроксимации функций регрессии при решении задач инверсной кинематики осуществляется достаточно давно, но ранние работы [9] в этой области отмечали низкую точность и производительность данных подходов. Развитие программного и аппаратного обеспечения в сфере машинного обучения позволило обеспечить прогресс по данному направлению: сочетания различных нейронных сетей используются для реализации инверсной кинематики роботов [10], прогнозирования положения тела с учетом окружающей среды на основе виртуального скелета [11] или набора датчиков [12]. В [13] рассмотрено применение нейронных сетей для восстановления ключевых точек силуэта человека при удалении от 10 до 30 % данных.

Таким образом, проведенный анализ показал, что для решения задачи инверсной кинематики оправдано использование небольшого количества датчиков и методов численной аппроксимации функций регрессии (например, нейронных сетей различной топологии) в качестве инструмента восстановления недостающих точек виртуальной модели тела человека. Определение минимального количества точек отслеживания позволяет снизить объем передаваемой информации, вычислительную нагрузку, а в случае использования трекеров VR — стоимость реализации системы отслеживания движений.

Постановка задачи оптимизации количества и расположения точек отслеживания тела человека в VR

Для достижения цели работы на первом этапе выполнен анализ предметной области. Далее формально определены основные объекты и поставлена задача определения оптимального количества и расположения точек отслеживания на теле человека.

Пусть задан набор исходных точек (физических трекеров VR, датчиков, виртуальных маркеров и др.), соответствующих ключевым зонам человеческого тела и позволяющих осуществить его последующую реконструкцию в виртуальном пространстве. Обозначим количество датчиков (трекеров) через TK . Каждый трекер $tr_i \in TR$ определим набором реальных точек с координатами по трем осям:

$$TR = \{tr_i\}, \\ tr_i = \{x_i, y_i, z_i\}.$$

Обозначим через $\mathbf{R}$ вектор множества значений координат всех трекеров:

$$\mathbf{R} = \{x_1, y_1, z_1, \dots, x_{TK}, y_{TK}, z_{TK}\}.$$

Размер $\mathbf{R}$ обозначим как N . Так как отслеживается положение трекеров по трем осям, то $N = 3TK$. Пусть $\mathbf{R}$ с размером N однозначно определяет цифровое представление тела человека (виртуальный аватар) с необходимой точностью (т. е. позволяет реализовать все необходимые части модели виртуального тела, их соединение и воспроизвести кинематические параметры процесса движения тела человека).

Оптимизация системы отслеживания тела человека состоит в определении минимального размера N вектора $\mathbf{R}$, т. е. нахождения количества трекеров, позволяющих обеспечить необходимую точность реконструкции цифрового представления человека. Так как уменьшение $\mathbf{R}$ при реализации инверсной кинематики модели тела человека приведет к увеличению погрешности при определении координат частей тела виртуального аватара, необходимо восстановить по наблюдениям регрессионную зависимость $\mathbf{F}$ вида:

$$\mathbf{F}(\mathbf{R}) = \mathbf{VR},$$

которая отображает пространство $\mathbb{R}^N$ реализации $\mathbf{R}$ в некоторое пространство $\mathbb{R}^K$, в котором реализованы величины $\mathbf{VR}$ (вектора виртуальных точек), определенным образом вычисляемые на основании величин $\mathbf{R}$. В качестве аппроксимации регрессии $\mathbf{F}$ использована некоторая NN определенной топологии. Топология указанной нейронной сети зависит от длины N вектора $\mathbf{R}$, которая определяет размерность входного слоя нейронной сети, скрытые слои для различных NN могут быть одинаковыми.

В формализованном виде задача оптимизации заключается в следующем: необходимо найти минимальный размер N входного вектора $\mathbf{R}$, а также определить параметры и топологию NN_N для каждого N , при которых среднеквадратичное отклонение восстановленных

элементов **VR** от истинных значений будет минимальным.

Для решения поставленной задачи необходимо сформировать обучающую выборку достаточных размеров и обучить набор NN_N для каждого выбранного N .

Алгоритм формирования данных для обучения нейронных сетей

Для формирования достаточного объема обучающей выборки необходим сбор информации, например, с костюмов захвата движений [14], которые позволяют сформировать полное представление о положении и кинематических параметрах модели тела человека в виртуальном пространстве. Отметим, что такой подход отличается значительными временными затратами и необходимостью привлечения большого количества участников с различными физическими параметрами. Главным недостатком является невозможность смоделировать в реальном мире ряд труднодоступных паттернов перемещения и небезопасных движений.

Решение данной проблемы — объединение данных от двух источников: систем захвата движений и виртуального аватара из игровой сцены [15]. Так как процедура построения цифрового представления модели тела человека в виртуальном пространстве осуществляется на основе набора точек в метрической системе координат, то для конечной реализации виртуальной модели тела человека источник данных не будет иметь принципиальной разницы. Предложенный подход дает возможность значительно ускорить процесс сбора данных за счет записи с повышенной частотой кадров и параллельно на нескольких виртуальных сценах.

Проведенный анализ виртуальных моделей тела человека показал, что для его реконструкции достаточно 18 ключевых точек. Это соответствует многим коммерческим решениям, например, Perception Neuron, основанном на 18 сенсорах [16]. Таким образом, при записи данных необходимо фиксировать 54 значения **R**. Расположение этих точек должно соответствовать узлам соединения основных сегментов тела человека: кисть, локоть, плечо, крайние точки стопы, колено, тазобедренный сустав, крайние и центральная точка спины, голова.

На следующем этапе осуществлен выбор сочетаний трекеров на основе эмпирического подхода в порядке их приоритетности и удобства закрепления на поверхности тела человека: начиная с крайних точек на руках и ногах, заканчивая промежуточными положениями (на коленях, локтях, спине и т. д.). В итоге сформулировано 6 предпочтительных наборов, начиная от простейшего с 3 точками и заканчивая максимальным с 13 точками:

- 3 точки: голова, правая и левая кисть;
- 5 точек: 3 точки с добавлением правой и левой стопы (зона пяты);
- 7 точек: 5 точек с добавлением правого и левого локтя;
- 9 точек: 7 точек с добавлением правого и левого колена;
- 11 точек: 9 точек с добавлением правого и левого тазобедренного сустава;

— 13 точек: 11 точек с добавлением верхней и нижней точек спины.

После обучения нейронных сетей на собранной базе данных осуществляется их тестирование на контрольном наборе и визуализация модели тела человека для оценки величины отклонения положения восстановленных точек от реальных. Дополнительно рассчитывается среднеквадратичное отклонение координат восстановленных точек позы от ожидаемых (среднее по всем точкам и максимальное по всему телу человека).

Результаты экспериментальных исследований

На первом этапе экспериментальных исследований осуществлен сбор данных с различных источников: типовые движения тела нескольких людей с разными физическими параметрами и воспроизведенные анимации в виртуальной сцене (источник анимаций — CMU Graphics Lab Motion Capture Database¹). В совокупности это позволило собрать 367 млн записей 18 трекеров (54 точки). В качестве обучающей выборки использовано 3583 тыс. записей, в качестве контрольной выборки для тестирования — 80 тыс. записей.

На предварительном этапе рассмотрено несколько топологий нейронных сетей, часть из которых из-за недостаточного или избыточного количества нейронов и слоев не достигла необходимых значений точности. Для выбора оптимальной топологии нейронной сети проведен анализ следующих гиперпараметров:

- наличие слоев разреживания (dropout): {да, нет};
- увеличение количества нейронов на каждом последующем слое: {да, нет};
- тип слоев: {Dense, Conv1D, RNN, LSTM};
- количество скрытых слоев: {1, 2, 3, 4, 5, 6, 7, 8, 9, 10};
- количество нейронов в скрытом слое: {20, 50, 100, 200, 400, 800, 1600};
- количество эпох {5, 10, 15, 25, 50}.

Результаты выбора наилучших гиперпараметров представлены в табл. 1. Для выбора параметров в целях ускорения процесса обучения использована ограниченная выборка из 5000 элементов тренировочного набора и фиксированный вход из трех точек модели человека. В качестве функции потерь используется среднеквадратичная ошибка — Mean Square Error (MSE). Для выбора каждого гиперпараметра выполнена фиксация остальных элементов и полный перебор значений текущего гиперпараметра. Далее осуществлен переход к следующему параметру с фиксацией наилучшего значения предыдущих параметров.

Наилучшие показатели получены с применением многослойной сети из пяти полносвязных (Dense) слоев с количеством нейронов от 400 до 1600, а также двумя слоями разреживания. Для финальной архитектуры создано 6 моделей, размер выходного слоя каждой — 54 нейрона, входного — варьируется от 9 (для трех то-

¹ B. Hahne. The daz-friendly bvh release of CMU motion capture database, 2010 [Электронный ресурс]. URL: <https://sites.google.com/a/cgspeed.com/cgspeed/motion-capture/daz-friendly-release> (дата обращения: 01.09.2023).

Таблица 1. Сравнение различных архитектур нейронных сетей

Table 1. Comparison of different neural network architectures

Гиперпараметр	Выбранное значение	Значение ошибки MSE	Анализ альтернативных вариантов
Наличие слоев разреживания (dropout)	да	0,008	При выборе «нет» MSE выше (0,010)
Увеличение количества нейронов на каждом последующем слое	да	0,009	При выборе «нет» MSE выше (0,014)
Тип слоев	Dense	0,009	С учетом структуры данных все альтернативные типы слоев показали сравнимые результаты (от 0,010 до 0,012), не превосходящие Dense, поэтому выбран наиболее производительный и простой тип слоев
Количество скрытых слоев	5	0,008	MSE выше при меньшем количестве слоев. При увеличении количества слоев растет сложность модели без значительного улучшения MSE. После 7–8 слоев начинается рост MSE
Количество нейронов в скрытом слое	400, 400, 800, 800, 1600	0,006	Получены сравнимые результаты для всех вариантов до 400 нейронов включительно, дальше наблюдается ухудшение. Меньшее количество нейронов не выбрано для универсальности архитектуры на большом наборе данных и при различном количестве входных точек
Количество эпох	10	0,005	После 10-й эпохи процесс обучения замедлился, скорость уменьшения MSE значительно упала

чек) до 39 (для 13 точек). Процесс обучения выполнен в течение 10 эпох и представлен на рис. 1. В результате можно сделать вывод, что NN_3 обладает наименьшей точностью, но, так как в процессе обучения точность обученных нейронных сетей соизмерима, необходимы дополнительные исследования для оценки полученных результатов. Для проверки качества нейронных сетей и величины отклонения восстановленных точек от реальных проведена реконструкция виртуальной модели тела человека. Ее результаты для всех нейронных сетей показаны на рис. 2. Красные линии соответствуют

эталонным значениям точкам модели тела человека, а синие — восстановленным с применением соответствующей i -ой нейронной сети.

Для контрольных данных рассчитано среднее отклонение по всем точкам полученных виртуальных моделей тела пользователя (в метрах), максимальное отклонение среди всех точек тела, а также средние отклонения этих величин в выборке. Результаты представлены в табл. 2.

Анализ производительности обученных нейронных сетей показал, что среднее время реконструкции

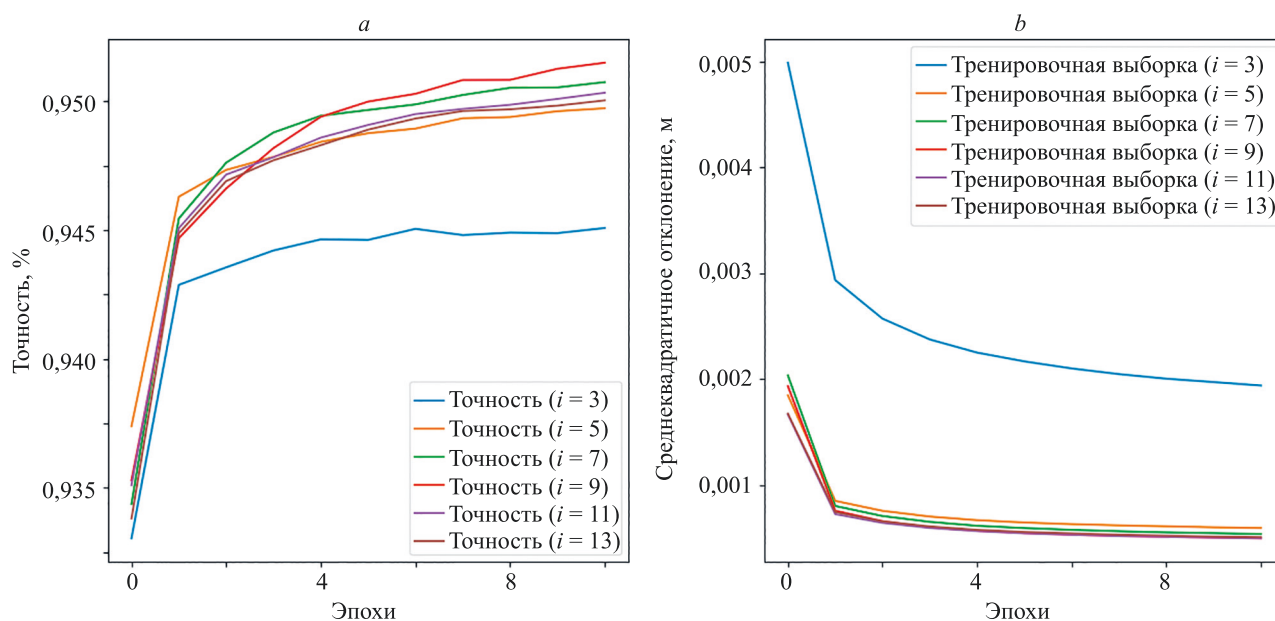


Рис. 1. Процесс обучения нейронной сети: точность нейронной сети (a); значения функций потерь (MSE) (b)

Fig. 1. Neural network training process: neural network accuracy (a); loss function values (MSE) (b)

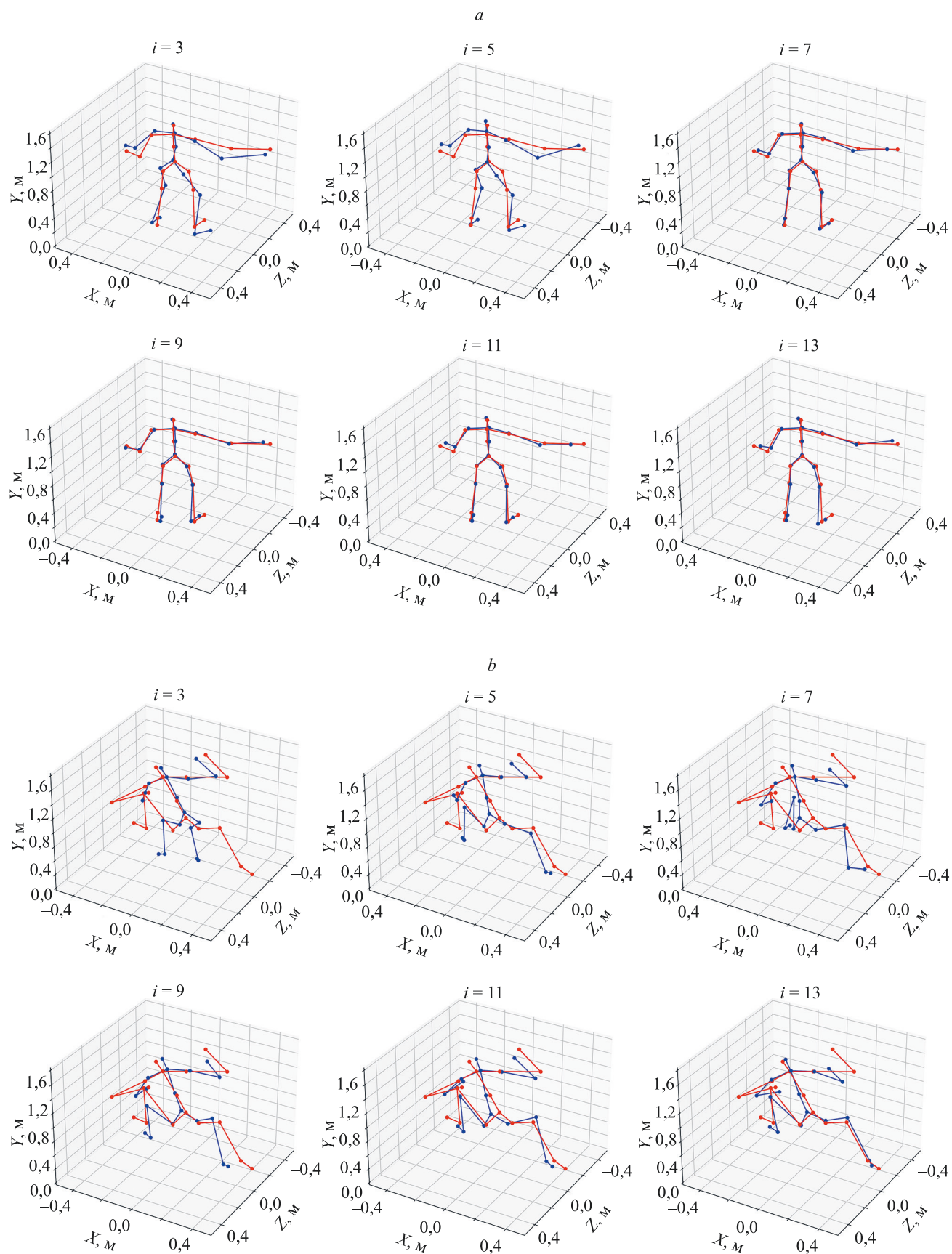


Рис. 2. Реконструкции виртуальных моделей тела для различных нейронных сетей: первая (а) и вторая (б) позы

Fig. 2. Reconstructions of virtual body models for various neural networks: the first (a) and second (b) poses

Таблица 2. Сравнение погрешности нейронных сетей на тестовой выборке

Table 2. Comparison of the error of neural networks on a test sample

NN_i	Среднее отклонение по всем точкам модели тела человека (среднее по выборке $\pm$ отклонение), м	Максимальное отклонение среди всех точек модели тела человека (среднее по выборке $\pm$ отклонение), м
3	$0,0771 \pm 0,0459$	$0,2358 \pm 0,1290$
5	$0,0420 \pm 0,0236$	$0,1021 \pm 0,0559$
7	$0,0306 \pm 0,0186$	$0,0761 \pm 0,0485$
9	$0,0298 \pm 0,0198$	$0,0723 \pm 0,0474$
11	$0,0295 \pm 0,0186$	$0,0737 \pm 0,0406$
13	$0,0299 \pm 0,0168$	$0,0745 \pm 0,0376$

модели тела человека составляет от 1,6 до 2,8 мс, что обеспечивает необходимое быстродействие (до 300 вызовов в секунду) и не повлияет на работоспособность систем VR.

Проведенные исследования показали, что использование трех точек не является достаточным, так как обученная таким образом NN_3 не может восстанавливать корректное положение ног. NN_5 точнее позиционирует нижние конечности. Использование 7 точек повышает точность реконструкции при сложных положениях рук, особенно в локтевой зоне. Далее применение 9 и более точек оказывает некоторое положительное влияние, но оно не является значительным и визуально значимым.

Полученные нейронные сети могут использоваться для решения следующих задач:

- NN_3 : вариант для упрощенных систем захвата движений, в которых не требуется отслеживание ног и точное позиционирование рук;
- NN_5 : оптимальное решение для большинства систем VR благодаря возможности достаточно точно позиционировать все конечности пользователя;
- NN_7 : решение для систем, требующих высокой точности позиционирования верхних конечностей, обеспечивающее сравнимую с более сложными моделями погрешность восстановления.

С учетом необходимости минимизации количества трекеров для широкого спектра систем VR оптимальная точность реконструкции тела человека достигается при 5 и 7 входных точках.

Интеграция обученных нейронных сетей в среды разработки систем VR требует изменения их архитектуры путем добавления дополнительных модулей.

Рассмотрим их специфику.

Модуль формирования цифрового представления человека используется для взаимодействия с нейронной сетью. В ходе запуска модуля указывается количество и расположение входных трекеров в зависимости от текущих настроек системы. Изначально подмодуль настроен для работы с игровым движком Unity и библиотекой SteamVR для взаимодействия с трекерами VR, но может быть модифицирован, так как предоставляет интерфейс для получения точек трекеров извне и передачу восстановленного тела человека в произвольную систему.

Модуль инверсной кинематики осуществляет достраивание виртуального аватара на основе совокупности точек, добавляя необходимые связи между

ними. Также модуль учитывает некоторые физические ограничения, что позволяет корректировать положение точек, если они выходят за границы допустимых пределов.

Модуль визуализации формирует виртуальную сцену и аватар пользователя, построенный на основе данных от модуля инверсной кинематики.

Функционирование системы тогда основано не на прямом использовании точек от трекеров VR, а на их отправке в модуль формирования цифрового представления, получение восстановленного набора точек модели тела человека, формирование виртуального аватара и изменение положения итоговой модели человека. Сформированный таким образом аватар будет максимально приближен к реальному положению тела с учетом минимального количества отслеживаемых точек.

Обученные нейронные сети и проект, реализующий интеграцию предложенной концепции по реконструкции модели тела человека, размещен в репозитории GitHub¹.

Заключение

Рассмотрена задача минимизации точек для обеспечения необходимой точности реконструкции кинематических параметров движений человека в виртуальном пространстве. Использование специализированных костюмов позволяет эффективно решать данный класс задач, однако подобный подход связан с большими затратами, а оборудование достаточно чувствительно к помехам и неудобно для повседневного использования. Использование же небольшого количества датчиков (трекеров виртуальной реальности, контроллеров) для отслеживания ключевых точек тела человека является перспективным направлением с точки зрения материальных и вычислительных затрат. Для точной реконструкции тела человека необходимо использование 18 точек отслеживания (датчиков), что не является оптимальным решением с экономической и эргономической точки зрения (датчики имеют достаточно большой размер).

Предложен подход, основанный на численной аппроксимации регрессионной зависимости между огра-

¹ Репозиторий GitHub DigitalShadowInMotion [Электронный ресурс]. URL: <https://github.com/Obukhov-Artem/DigitalShadowInMotion> (дата обращения: 22.05.2023).

ническим по размеру вектором точек отслеживания и вектором 18 виртуальных точек, используемых для полной реконструкции модели тела человека. В рамках подхода сформулирована регрессионная зависимость между реальными и виртуальными точками, поставлена задача оптимизации количества трекеров. Выполнены экспериментальные исследования и обучены топологии нейронных сетей для различных наборов входных точек (от 3 до 13). Оптимальная точность восстановления точек получена при 5 и 7 входных точках, что позво-

ляет упростить реализацию систем захвата движений, построение виртуальной модели тела человека и восстановление кинетических параметров его движений в виртуальной реальности. Новизна исследования также включает подход к сбору данных не только с реального человека, но и с виртуальной модели, что позволило ускорить процесс сбора информации для обучения нейронных сетей и увеличить объем данных за счет моделирования различных паттернов движения тела человека в виртуальном пространстве.

Литература

1. Obukhov A.D., Volkov A.A., Vekhteva N.A., Teselkin D.V., Arkhipov A.E. Human motion capture algorithm for creating digital shadows of the movement process // *Journal of Physics: Conference Series*. 2022. V. 2388. N 1. P. 012033. <https://doi.org/10.1088/1742-6596/2388/1/012033>
2. Azarby S., Rice A. Understanding the effects of virtual reality system usage on spatial perception: The potential impacts of immersive virtual reality on spatial design decisions // *Sustainability*. 2022. V. 14. N 16. P. 10326. <https://doi.org/10.3390/su141610326>
3. Parger M., Mueller J.H., Schmalstieg D., Steinberger M. Human upper-body inverse kinematics for increased embodiment in consumer-grade virtual reality // *Proc. of the 24th ACM Symposium on Virtual Reality Software and Technology*. 2018. P. 1–10. <https://doi.org/10.1145/3281505.3281529>
4. Caserman P., Garcia-Agundez A., Konrad R., Göbel S., Steinmetz R. Real-time body tracking in virtual reality using a Vive tracker // *Virtual Reality*. 2019. V. 23. N 2. P. 155–168. <https://doi.org/10.1007/s10055-018-0374-z>
5. Feigl T., Gruner L., Mutschler C., Roth D. Real-time gait reconstruction for virtual reality using a single sensor // *Proc. of the 2020 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. 2020. P. 84–89. <https://doi.org/10.1109/ismar-adjunct51615.2020.00037>
6. Liu H., Zhang Z., Xie X., Zhu Y., Liu Y., Wang Y., Zhu S.-C. High-fidelity grasping in virtual reality using a glove-based system // *Proc. of the 2019 International Conference on Robotics and Automation (ICRA)*. 2019. P. 5180–5186. <https://doi.org/10.1109/icra.2019.8794230>
7. Liu R., Liu C. Human motion prediction using adaptable recurrent neural networks and inverse kinematics // *IEEE Control Systems Letters*. 2021. V. 5. N 5. P. 1651–1656. <https://doi.org/10.1109/lcssys.2020.3042609>
8. Li J., Xu C., Chen Z., Bian S., Yang L., Lu C. Hybrik: A hybrid analytical-neural inverse kinematics solution for 3D human pose and shape estimation // *Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 2021. P. 3383–3392. <https://doi.org/10.1109/cvpr46437.2021.00339>
9. Oyama E., Agah A., MacDorman K.F., Maeda T., Tachi S. A modular neural network architecture for inverse kinematics model learning // *Neurocomputing*. 2001. V. 38–40. P. 797–805. [https://doi.org/10.1016/s0925-2312\(01\)00416-7](https://doi.org/10.1016/s0925-2312(01)00416-7)
10. Bai Y., Luo M., Pang F. An algorithm for solving robot inverse kinematics based on FOA optimized BP neural network // *Applied Sciences*. 2021. V. 11. N 15. P. 7129. <https://doi.org/10.3390/app11157129>
11. Kratzer P., Toussaint M., Mainprice J. Prediction of human full-body movements with motion optimization and recurrent neural networks // *Proc. of the 2020 IEEE International Conference on Robotics and Automation (ICRA)*. 2020. P. 1792–1798. <https://doi.org/10.1109/icra40945.2020.9197290>
12. Bataineh M., Timothy M., Karim A.-M., Jasbir A. Neural network for dynamic human motion prediction // *Expert Systems with Applications*. 2016. V. 48. P. 26–34. <https://doi.org/10.1016/j.eswa.2015.11.020>
13. Kucherenko T., Beskow J., Kjellström H. A neural network approach to missing marker reconstruction in human motion capture // *arXiv*. 2018. arXiv:1803.02665. <https://doi.org/10.48550/arXiv.1803.02665>

References

1. Obukhov A.D., Volkov A.A., Vekhteva N.A., Teselkin D.V., Arkhipov A.E. Human motion capture algorithm for creating digital shadows of the movement process. *Journal of Physics: Conference Series*. IOP Publishing, 2022, vol. 2388, no. 1, pp. 012033. <https://doi.org/10.1088/1742-6596/2388/1/012033>
2. Azarby S., Rice A. Understanding the effects of virtual reality system usage on spatial perception: The potential impacts of immersive virtual reality on spatial design decisions. *Sustainability*, 2022, vol. 14, no. 16, pp. 10326. <https://doi.org/10.3390/su141610326>
3. Parger M., Mueller J.H., Schmalstieg D., Steinberger M. Human upper-body inverse kinematics for increased embodiment in consumer-grade virtual reality. *Proc. of the 24th ACM Symposium on Virtual Reality Software and Technology*, 2018, pp. 1–10. <https://doi.org/10.1145/3281505.3281529>
4. Caserman P., Garcia-Agundez A., Konrad R., Göbel S., Steinmetz R. Real-time body tracking in virtual reality using a Vive tracker. *Virtual Reality*, 2019, vol. 23, no. 2, pp. 155–168. <https://doi.org/10.1007/s10055-018-0374-z>
5. Feigl T., Gruner L., Mutschler C., Roth D. Real-time gait reconstruction for virtual reality using a single sensor. *Proc. of the 2020 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, 2020, pp. 84–89. <https://doi.org/10.1109/ismar-adjunct51615.2020.00037>
6. Liu H., Zhang Z., Xie X., Zhu Y., Liu Y., Wang Y., Zhu S.-C. High-fidelity grasping in virtual reality using a glove-based system. *Proc. of the 2019 International Conference on Robotics and Automation (ICRA)*, 2019, pp. 5180–5186. <https://doi.org/10.1109/icra.2019.8794230>
7. Liu R., Liu C. Human motion prediction using adaptable recurrent neural networks and inverse kinematics. *IEEE Control Systems Letters*, 2021, vol. 5, no. 5, pp. 1651–1656. <https://doi.org/10.1109/lcssys.2020.3042609>
8. Li J., Xu C., Chen Z., Bian S., Yang L., Lu C. Hybrik: A hybrid analytical-neural inverse kinematics solution for 3D human pose and shape estimation. *Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2021, pp. 3383–3392. <https://doi.org/10.1109/cvpr46437.2021.00339>
9. Oyama E., Agah A., MacDorman K.F., Maeda T., Tachi S. A modular neural network architecture for inverse kinematics model learning. *Neurocomputing*, 2001, vol. 38–40, pp. 797–805. [https://doi.org/10.1016/s0925-2312\(01\)00416-7](https://doi.org/10.1016/s0925-2312(01)00416-7)
10. Bai Y., Luo M., Pang F. An algorithm for solving robot inverse kinematics based on FOA optimized BP neural network. *Applied Sciences*, 2021, vol. 11, no. 15, pp. 7129. <https://doi.org/10.3390/app11157129>
11. Kratzer P., Toussaint M., Mainprice J. Prediction of human full-body movements with motion optimization and recurrent neural networks. *Proc. of the 2020 IEEE International Conference on Robotics and Automation (ICRA)*, 2020, pp. 1792–1798. <https://doi.org/10.1109/icra40945.2020.9197290>
12. Bataineh M., Timothy M., Karim A.-M., Jasbir A. Neural network for dynamic human motion prediction. *Expert Systems with Applications*, 2016, vol. 48, pp. 26–34. <https://doi.org/10.1016/j.eswa.2015.11.020>
13. Kucherenko T., Beskow J., Kjellström H. A neural network approach to missing marker reconstruction in human motion capture. *arXiv*, 2018, arXiv:1803.02665. <https://doi.org/10.48550/arXiv.1803.02665>
14. Geigel J., Schweppe M. Motion capture for realtime control of virtual actors in live, distributed, theatrical performances. *Proc. of the 2011*

14. Geigel J., Schweppe M. Motion capture for realtime control of virtual actors in live, distributed, theatrical performances // Proc. of the 2011 IEEE International Conference on Automatic Face & Gesture Recognition (FG). 2011. P. 774–779. <https://doi.org/10.1109/FG.2011.5771347>
15. Degen R., Tauber A., Nüßgen A., Irmer M., Klein F., Schyr C., Leijon M., Ruschitzka M. Methodical approach to integrate human movement diversity in real-time into a virtual test field for highly automated vehicle systems // Journal of Transportation Technologies. 2022. V. 12. N 3. P. 296–309. <https://doi.org/10.4236/jtts.2022.123018>
16. Sers R., Forrester S., Moss E., Ward S., Ma J., Zecca M. Validity of the Perception Neuron inertial motion capture system for upper body motion analysis // Measurement. 2020. V. 149. P. 107024. <https://doi.org/10.1016/j.measurement.2019.107024>
- IEEE International Conference on Automatic Face & Gesture Recognition (FG), 2011, pp. 774–779. <https://doi.org/10.1109/FG.2011.5771347>
15. Degen R., Tauber A., Nüßgen A., Irmer M., Klein F., Schyr C., Leijon M., Ruschitzka M. Methodical approach to integrate human movement diversity in real-time into a virtual test field for highly automated vehicle systems. *Journal of Transportation Technologies*, 2022, vol. 12, no. 3, pp. 296–309. <https://doi.org/10.4236/jtts.2022.123018>
16. Sers R., Forrester S., Moss E., Ward S., Ma J., Zecca M. Validity of the Perception Neuron inertial motion capture system for upper body motion analysis. *Measurement*, 2020, vol. 149, pp. 107024. <https://doi.org/10.1016/j.measurement.2019.107024>

Авторы

Обухов Артём Дмитриевич — доктор технических наук, доцент, Тамбовский государственный технический университет, Тамбов, 392000, Российская Федерация, [sc 56104232400](https://orcid.org/0000-0002-3450-5213), <https://orcid.org/0000-0002-3450-5213>, obuhov.art@gmail.com

Теселкин Даниил Вячеславович — студент, Тамбовский государственный технический университет, Тамбов, 392000, Российская Федерация, [sc 57362498400](https://orcid.org/0000-0002-1304-9490), <https://orcid.org/0000-0002-1304-9490>, teselkin.dv@tstu.ru

Статья поступила в редакцию 07.02.2023
Одобрена после рецензирования 05.06.2023
Принята к печати 18.07.2023

Authors

Artem D. Obukhov — D.Sc., Associate Professor, Tambov State Technical University, Tambov, 392000, Russian Federation, [sc 56104232400](https://orcid.org/0000-0002-3450-5213), <https://orcid.org/0000-0002-3450-5213>, obuhov.art@gmail.com

Daniil V. Teselkin — Student, Tambov State Technical University, Tambov, 392000, Russian Federation, [sc 57362498400](https://orcid.org/0000-0002-1304-9490), <https://orcid.org/0000-0002-1304-9490>, teselkin.dv@tstu.ru

Received 07.02.2023
Approved after reviewing 05.06.2023
Accepted 18.07.2023



Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»