

УДК 004.8

ОПТИМИЗАЦИЯ ПРОЦЕССА ГЕНЕРАЦИИ 3D-АНИМАЦИЙ С ПОМОЩЬЮ НЕЙРОННЫХ СЕТЕЙ: РОЛЬ СБОРА И ОБРАБОТКИ ДАННЫХ ДЛЯ ЭФФЕКТИВНОГО СОЗДАНИЯ ТРЕХМЕРНОГО КОНТЕНТА

Мозговой Николай Васильевич,

Магистрант, Донской государственной технической университет,

кафедра «Институт сквозных технологий»

344000, РФ, г. Ростов - на - Дону, пл. Гагарина 1

E-mail: cav.unknown@gmail.com

09.04.02 Интеллектуальные медиатехнологии

Аннотация

В статье рассматриваются способы по оптимизации нейросетевых методов генерации трехмерных анимаций при выборе представления и обработки данных для обучения. Данный аспект крайне важен, так как в эпоху больших данных образовалась необходимость в точных и масштабируемых методах сбора данных, направленных на улучшение существующих результатов функционирования нейросетей.

Исследуется потенциал использования исходно-синтетических данных, генерации новых признаков для улучшения качества анимаций, работы не с единственным табличным представлением данных, а со множеством. Анализируются перспективы и релевантность использования искусственного интеллекта в области трехмерной анимации гуманоидных персонажей.

Ключевые слова: искусственный интеллект, синтетические данные, нейронные сети, трансформеры, автоматизация, 3D-анимация.

OPTIMIZING 3D-ANIMATION GENERATION PIPELINES VIA NEURAL NETWORKS: THE ROLE OF DATA MINING AND PROCESSING FOR EFFICIENT THREE-DIMENSIONAL CONTENT CREATION

Nikolai V. Mozgovoi,

Master's student, Don State Technical University,

Department of «Institute of End-to-End Technologies»

344000, Russian Federation, Rostov - on - Don, pl. Gagarina 1

E-mail: cav.unknown@gamil.com

09.04.02 Intellectual Media Technologies

ABSTRACT

The paper discusses ways to optimize neural network methods for generating three-dimensional animations when selecting the representation and processing of training data. This aspect is extremely important because in the era of big data there is a need for accurate and scalable data acquisition methods aimed at improving existing neural network results.

The potential of using source-synthetic data, generating new features to improve the quality of animations, working with multiple rather than a single tabular representation of data is explored. The prospects and relevance of using artificial intelligence in the field of three-dimensional animation of humanoid characters are analyzed.

Keywords: artificial intelligence, synthetic data, neural networks, transformers, automation, 3D animation.

Данные как совокупность зафиксированных сведений для последующего анализа, интерпретации и обработки являются неотъемлемым элементом при построении любого проекта по машинному обучению или нейросети. Для того, чтобы искусственный интеллект сумел обработать и предоставить релевантный, предсказуемый результат, ей необходимо передать n -примеров, на которых она сумела бы обучиться. Эталонные образцы требуют предобработки, углубленного анализа, генерации новых признаков для улучшения сходимости модели в общем, так и повышения точности выдаваемых результатов. Допустимо искать бесплатные, общедоступные и заранее структурированные наборы данных на специализированных, академических ресурсах и в публичных репозиториях [1, 2, 3]. В зависимости от задачи, возможно столкнуться с ситуацией, когда подготовленных данных недостаточно или их вовсе нет, то альтернативой станет сбор данных вручную или автоматически, например, с помощью парсинга веб-ресурсов или генерации на основе существующего датасета.

Цель работы заключается в анализе методов по оптимизации сбора и обработки данных для последующего обучения нейросетевых моделей вида seq2seq в области генерации 3D-анимаций. Проведено исследование различных подходов, таких как использование исходно-синтетических данных, комбинация данных из внешних источников, а также необходимость выделения новых признаков для ускорения и упрощения процесса создания трехмерного движения персонажей.

При сборе данных необходимо оценивать природу данных, источник и лицензию на использование того или иного материала [4]. Существуют различные виды лицензий, позволяющих использовать данные как в академических, так и в коммерческих целях. Несоблюдение права на использование способно повлечь за собой критические последствия: невозможность публикации проекта, использовании и распространении генерируемого материала, судебные иски. Для минимизации подобных рисков следует внимательно знакомиться с лицензией на использование тех или иных данных.

Можно пронаблюдать тенденцию, связанную с тем, что для упрощения и ускорения процесса создания анимации, ИИ используется для возможности большей масштабируемости данных, а также в нем начинают все больше задействовать исходно-синтетические данные. Именно эта направленность исключит вопросы, связанные с политикой авторского права исходных данных, на которых обучалась модель, сохранит их конфиденциальность и даст им более широкое использование [5]. Применимо к области трехмерной анимации, станет возможно исследовать сложные движения, не требующие наличия специалиста для выполнения того или иного физического трюка, камер с высоким разрешением. В перспективе, модели, использующие текстовые промты как основу для

генерации анимации, сумеют затмить решения, в основе которых лежит компьютерное зрение.

Данные можно подразделить на две группы, в зависимости от природы происхождения: естественные и синтетические. Под первым подразумеваются данные, собранные непосредственно в реальном мире, включая текст, изображения, видео и аудио. Под вторым – данные, созданные алгоритмическим путем, либо исключительно в цифровом пространстве, на основе окружающей действительности, которые призваны лишь отражать её, при этом не имея материального начала (т.е. искажающие реалистичное представление источника). Естественные описывают без прикрас, в том виде, в котором они были получены, с присущими им дефектами и неточностями в выборке, тем не менее, они в наиболее явном виде отображают свою область и конкретный случай. Однако, из-за присущей им ограниченности и неполноты в практических приложениях возникают такие проблемы, как дисбаланс данных и дискриминация данных [6]. Поскольку удовлетворить спрос, опираясь только на реальные данные, сложно, исследователи начинают применять различные методы генерации синтетических данных на основе существующих реальных данных. Эти методы варьируются от традиционных статистических моделей до современных передовых методик, основанных на моделях глубокого обучения [7, 8].

В отличие от реальных данных, имеющих естественную природу происхождения, а соответственно и наследующих все ограничения окружающего мира: недостаток информации, конечность материала, наличие определенных условий для выполнения, синтетические данные (искусственно сформированные) могут быть легко масштабированы, и способны составить данные в редких или сложных сценариях, которых крайне затруднительно или вовсе невозможно достичь, тем самым устраняя проблемы разнообразия в некоторых наборах данных.

Тем не менее, процесс создания синтетических наборов также не обходится без нетривиальных препятствий, в особенности того, что затрагивает отношение справедливости и репрезентативности распределения данных, и потенциально может ограничить эффективность обобщения моделей [9]. Данные изъяны способны повлиять на производительность моделей и привести к предвзятости и дискриминации в рамках выполнения реальных задачах. В рамках задачи по созданию 3D-анимаций – необходимые данные, т.к. традиционный метод создания движений создается в цифровой среде, лишь пытаясь адаптировать действия из окружающей среды. Синтетические данные следует рассматривать как инструмент для ускорения выстраивания «конвейера исследований», но, в итоге, любые результаты обязательно должны быть оценены и, при необходимости, доработаны в рамках реального мира.

Наиболее типичным представлением данных для генерации трехмерных анимаций являются изображения, по которым определяется положение тела человека, распознается и оценивается та или иная поза, которая в последствии переносит и применяет полученные координаты точек к ригу (скелету) 3D объекта [10]. Подобные решения реализуются на базе сверточных нейросетей (YOLO) или Transformer (ViT). Тем не менее, наличие предикатов к требуемым данным, делает данный подход достаточно дорогим в производстве.

При подготовке датасета для последующей генерации трехмерных анимаций гуманоидных персонажей надлежит руководствоваться вышеупомянутыми принципами, а именно:

- Найти источник данных и ознакомиться с лицензией на использование;
- Спарсить или скачать, собрать данные, формализовав их;
- Предобработать данные на основе поставленной задачи, сгенерировать новые признаки.

Наиболее адаптивным решением сумеет стать комбинация нескольких подходов: использовать готовые анимации на специализированных ресурсах, распространяющихся по лицензии MIT и CC-BY, а также воспользоваться возможностью генерации синтетических данных, при сборе достаточного количества исходного материала. В качестве внешнего источника стоит рассмотреть ресурс Maximo, предоставляющий базовые анимации бега, ходьбы, прыжков по лицензии CC-BY, а в качестве внутреннего ресурса – синтез новых данных, на основе имеющихся.

Выделение признаков из анимаций будет достигаться в программном обеспечении для работы с 3D-графикой Blender с помощью Python и модуля bpy. Данный модуль является Python API (Application Programming Interface) для Blender, позволяющий взаимодействовать с 3D-редактором с помощью кода, что в свою очередь, способно упростить и ускорить рутинные процессы, а в данном случае – вычленив необходимые данные, которые скрыты от пользователя.

Эти признаки будут иметь крайне значительное влияние и полноценно описывают любое движение: координаты положения и вращения демонстрируют расположение костей и всего объекта в трехмерной сцене, ограничения определяют диапазон значений, за пределы которого не сумеет переместиться или деформироваться кость, длина кости позволяет оценить размер рига, что впоследствии будет важным фактором при применении генерации на разных гуманоидных моделях, величина которых могут не соответствовать друг другу.

Дополнительно, возможно внедрить новый признак – зависимость координат для сглаживания движений, в целях достижения более плавных анимаций. Для вычисления данного признака задействуется кубическая сплайн-интерполяция, использующая кубические полиномы для интерполяции данных между узловыми точками. В рамках задачи, формируются кубические сплайны для интерполяции координат и вращений костей между кадрами анимации по следующей формуле:

$$S_i(x) = a_i + b_i * (x - x_i) + c_i * (x - x_i)^2 + d_i(x - x_i)^3$$

где $S_i(x)$ – значение сплайна на интервале $[x_i, x_i + 1]$;

где a_i, b_i, c_i, d_i – коэффициенты полинома;

Для каждого интервала $[x_i, x_i+1]$ коэффициенты a_i, b_i, c_i, d_i вычисляются так, чтобы обеспечить непрерывность функции и её первой и второй производных на всех узлах [11].

После выделения всех ключевых признаков, необходимо будет предобработать “сырые” данные. Так как задействованы будут несколько различных источники, то количество активных костей и их наименований способны отличаться, соответственно, требуется нормализовать их численность и названия для достижения однородности данных. Следует отметить, что лицевые анимации и сложные движения пальцами рук и ног – рассматриваться не будут, строки содержащие данные значения будут удалены. Останутся исключительно движения корпуса тела и конечностей гуманоидного персонажа. Так как результатом выделения признаков из одной анимации – будет один файл в формате csv, надлежит использовать библиотеку `pytorch tabular`, позволяющую взаимодействовать с таблицами, как со строками, т.е. передавать в модель несколько таблиц (единицу данных) для обучения с помощью `pytorch`.

По результатам исследования были проанализированы методы, упрощающие сбор и предобработку данных для генерации трехмерных анимаций с помощью нейросетей, релевантность применения исходно-синтетических данных. В перспективе, данные можно обогащать новыми признаками, на основе имеющихся, при достижении оптимального количества собранных анимаций возможно генерировать новые, которые, в последствии, станет возможным добавить в обучающую выборку.

Список литературы:

1. Тахердуст Х. Методы и инструменты сбора данных для исследований; пошаговое руководство по выбору техники сбора данных для академических и бизнес-проектов //Международный журнал академических исследований в менеджменте (IJARM). – 2021. – Т. 10. – №. 1. – С. 10-38.
2. Чапман А. и др. Поиск датасета: обзор //Журнал VLDB. – 2020. – Т. 29. – №. 1. – С. 251-272.
3. Рох У., Хео Джи., Ванг С. Е. Обзор сбора данных для машинного обучения: перспектива интеграции больших данных и ИИ //IEEE Транзакции по инженерии знаний и данных. – 2019. – Т. 33. – №. 4. – С. 1328-1347.
4. Демьянченко Д. А., Григоренко Д. С. ПОНЯТИЕ " СВОБОДНОЙ" ЛИЦЕНЗИИ И ПРОБЛЕМЫ ЕЁ ПРАВОВОЙ ИМПЛЕМЕНТАЦИИ В РОССИЙСКОЕ ГРАЖДАНСКОЕ ЗАКОНОДАТЕЛЬСТВО //Инновации в науке. – 2016. – №. 4-2 (53). – С. 132-138.
5. Джордон Дж. и др. Синтетические данные--что, зачем и почему? //arXiv препринт arXiv:2205.03257. – 2022.
6. Хао С. и др. Синтетические данные в ИИ: Проблемы, применение и этические последствия //arXiv препринт arXiv:2401.01629. – 2024.].
7. Копылов Д. А., Агешин Е. С., Хомутская О. В. ФОРМИРОВАНИЕ СИНТЕТИЧЕСКИХ ДАННЫХ ДЛЯ ОБУЧЕНИЯ СИСТЕМЫ КОМПЬЮТЕРНОГО ЗРЕНИЯ //Автоматизация и моделирование в проектировании и управлении. – 2022. – №. 4 (18). – С. 18-28.
8. Беляева О. В., Перминов А. И., Козлов И. С. Использование синтетических данных для тонкой настройки моделей сегментации документов //Труды Института системного программирования РАН. – 2020. – Т. 32. – №. 4. – С. 189-202.
9. Лу У. и др. Машинное обучение для генерации синтетических данных с помощью машинного //arXiv препринт arXiv:2302.04062. – 2023.
10. Мозговой, Н. В. Описание движений с помощью искусственного интеллекта в области 3D-анимации / Н. В. Мозговой, Е. Н. Ладоса // Оригинальные исследования. – 2024. – Т. 14, № 3. – С. 164-168.
11. Дуркан С. и др. Кубические-сплайны //arXiv препринт arXiv:1906.02145. – 2019.

References:

1. Taherdoost H. Data collection methods and tools for research; a step-by-step guide to choose data collection technique for academic and business research projects //International Journal of Academic Research in Management (IJARM). – 2021. – Т. 10. – №. 1. – С. 10-38.

2. Chapman A. et al. Dataset search: a survey //The VLDB Journal. - 2020. - Т. 29. - №. 1. - С. 251-272.
3. Roh Y., Heo G., Whang S. E. A survey on data collection for machine learning: a big data-ai integration perspective //IEEE Transactions on Knowledge and Data Engineering. - 2019. - Т. 33. - №. 4. - С. 1328-1347.
4. Demyanchenko D. A., Grigorenko D. S. The concept of “free” license and the problems of its legal implementation in the Russian civil rights organization. FREE” LICENSE AND PROBLEMS OF HIS LEGAL IMPLEMENTATION IN RUSSIAN CIVIL LAW //Innovations in Science. - 2016. - №. 4-2 (53). - P. 132-138.
5. Jordon J. et al. Synthetic Data--what, why and how? //arXiv preprint arXiv:2205.03257. - 2022.
6. Hao S. et al. Synthetic data in AI: Challenges, applications, and ethical implications //arXiv preprint arXiv:2401.01629. - 2024.].
7. Kopylov D. A., Ageshin E. S., Khomutskaya O. V. FORMATION OF SYNTHETIC DATA FOR TRAINING OF COMPUTER SIGHT SYSTEM // Automation and modeling in design and management. - 2022. - №. 4 (18). - P. 18-28.
8. Belyaeva O. V. V. V., Perminov A. I., Kozlov I. S. Using synthetic data for fine-tuning document segmentation models // Proceedings of the Institute of System Programming of the Russian Academy of Sciences. - 2020. - Т. 32. - №. 4. - P. 189-202.
9. Lu Y. et al. Machine learning for synthetic data generation: a review //arXiv preprint arXiv:2302.04062. - 2023.
10. Mozgovoi, N. V. Description of movements with the help of artificial intelligence in the field of 3D-animation / N. V. Mozgovoi, E. N. Ladosha // Original Research. - 2024. - Т. 14, № 3. - P. 164-168.
11. Durkan C. et al. Cubic-spline flows //arXiv preprint arXiv:1906.02145. - 2019.