

УДК 004.8

ОПИСАНИЕ ДВИЖЕНИЙ С ПОМОЩЬЮ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В ОБЛАСТИ 3D-АНИМАЦИИ

Мозговой Николай Васильевич,

Магистрант, Донской Государственный Технический Университет,
кафедра «Медиатехнологии»

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

E-mail: cav.unknown@gmail.com

09.04.02 Информационные системы и технологии

Ладоша Евгений Николаевич,

К.т.н, доцент, Донской Государственный Технический Университет,
кафедра «Медиатехнологии»

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

E-mail: eladosha@donstu.ru

Аннотация

С прогрессивным развитием технологий 3D-графики открывается больше перспектив и возможностей для реализации визуального контента. Тем не менее, рост потребности в трехмерном продукте опережает возможности специалистов и студий по его производству. Из-за чего заказчики сокращают сроки на выполнение работ, тем самым вынуждая разработчиков выпускать «сырой» материал. Поэтому, избавление от рутинной части работы сумело бы повысить эффективность труда специалистов.

Цель работы заключается в анализе методов по упрощению создания трехмерных анимаций. Проведено исследование различных подходов по реализации искусственного интеллекта в области трехмерной анимации, рассмотрен вопрос релевантности и перспективности применения метода процедурной анимации 3D-персонажей.

Ключевые слова: искусственный интеллект, 3D-анимация, процедурная анимация, машинное обучение.

VISUALIZATION OF MOTIONS USING ARTIFICIAL INTELLIGENCE IN THE FIELD OF 3D ANIMATION

Nikolay V. Mozgovoy,

Master's Degree student, Don State Technical University,
Department of "Media Technologies"

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

E-mail: cav.unknown@gmail.com

09.04.02 Information systems and technologies

Evgeny N. Ladasha,

Candidate of Technical Sciences, Associate Professor, Don State Technical University,
Media Technologies Department
344000, Russian Federation, Rostov-on-Don, Gagarina Pl. 1.
E-mail: eladosha@donstu.ru

ABSTRACT

With the progressive development of 3D graphics technology, there are more prospects and opportunities for the realization of visual content. Nevertheless, the growing demand for three-dimensional products outstrips the ability of specialists and studios to produce them. Because of this, customers are reducing the deadlines for work, thereby forcing developers to produce unfinished material. Therefore, getting rid of the routine part of the work would increase the efficiency of labor of specialists.

The purpose of the work is to analyze methods to simplify the creation of three-dimensional animations. The research of various approaches to the realization of artificial intelligence in the field of three-dimensional animation is carried out, the question of relevance and prospects of application of the method of procedural animation of 3D characters is considered.

Keywords: artificial intelligence, 3D animation, procedural animation, machine learning.

Наиболее длительным и ресурсозатратным процессом при разработке 3D-продукта является создание анимаций [1]. Традиционные методы производства движений объектов остаются незаменимыми инструментами для достижения желаемого результата, тем не менее существуют и альтернативные подходы для ускорения решения задач 3D-анимаций.

Большинство решений, упрощающих процесс создания трехмерной анимации с помощью машинного обучения, базируются на безмаркерном методе захвата движений на основе распознавания и оценки поз людей с видео, разработанных на основе YOLO или OpenCV. В частности, для устройств семейства IOS существует приложение, позволяющее делать мокап «МОСАРХ», аналогичный функционал имеет «Rokoko» и «Cascadeur». Подобные решения в значительной степени снижают стоимость производства по сравнению с стандартными подходами. Принцип работы заключается в покадровом определении расположения конечностей на двумерном изображении отслеживаемого объекта с последующим переносом изменений координат на трехмерного персонажа [2]. Однако, их алгоритмы не всегда обеспечивают точность отслеживания, требуя наличия определенных предусловий для создания анимации. Из наиболее явных недостатков можно отметить обязательное наличие актера, хорошо освещенной локации, большого разрешения видео для улучшения точности захвата, а также необходимость проводить «чистку» мокапа от артефактов ретаргетинга, убирая неестественные «подергивания» конечностей, вызванных несовершенством технологии. Поэтому, для создания окончательной, убедительной и реалистичной анимации по-прежнему требуется большое мастерство и художественный вклад для усовершенствования движений, но они могут прочно базироваться на данных мокапа.

Но, существуют и альтернативные методы, использующие ИИ для создания анимаций. Например, Learned Motion Matching, в котором используются три нейронных сети для определения наиболее подходящей позы, за счет поиска ближайших соседей, или

отрывка движения, в целях создания непрерывного движения персонажа. Данная система была разработана компанией Ubisoft, применялась в таких игровых проектах как The Last of Us II (гибридное применение) и Assassin's Creed. Принцип синтеза анимации основывается на том, что каждые N-кадров осуществляется поиск наиболее подходящего отрывка движения, иными словами, позы и, в зависимости от траектории движения, прогнозируется наклон и проекция на поверхность передвижения [3]. В качестве исходных данных для обучения используются необработанные, зачастую продолжительные по времени, записи перемещения координат захвата движения, снятые с реальных актеров. Технология Motion Capture фиксирует движения окружающей человека действительности и переводит данные о перемещении объектов в цифровую среду, в которой, впоследствии можно манипулировать запечатленной информацией. Появление подхода вызвано потребностью воспроизведения реалистичных физических движений. Применение длительных сессий motion capture в обучающей выборке обусловлено возможностью обратиться к этим данным для определения наиболее уместного и соответствующего пользовательскому вводу отрывка движения. Сама анимация представляет собой последовательность поз, которые хранятся в датасете, а для поиска подходящего компонента используется индекс кадра. Первая из нейронных сетей, применяемых в данном подходе, названа Decompressor, призванная координировать и преобразовывать текущую позу в результирующую. Вторая – Stepper, отвечает за прогнозирование последующей позы, то есть возможных изменений в координатах передвижения объекта. Заключительная сеть Projector корректирует переход текущей позы в результирующую с учетом прогноза траектории, полученного в Stepper. Вышеупомянутые сети обучались с помощью PyTorch, используя алгоритм оптимизации RAdam. Итоговая модель не содержит координатные или иные метаданные в памяти, в отличие от классического Motion Matching алгоритма. Подобное решение наследует все недостатки метода захвата движений: необходимость использовать исключительно реальные данные, влекущие за собой подготовку к съемкам актеров и постобработку полученных результатов.

Метод Masked Motion Model, состоящий из двух модулей, базируется на архитектуре Transformers, токенизирующей данные об движении отдельно для верхней и нижней частей тела 3D-персонажа. Генерация движения обеспечивается за счет ввода текстового промта желаемой анимации с сохранением семантической информации о действии модели. В качестве датасета выступают как синтетические (искусственно сформированные), так и реальные данные. Именно благодаря этому набор предобученных действий может быть масштабирован. Данный подход обеспечивает широкую вариативность финальной анимации, учитывая каждый токен в контексте как предыдущего, так и последующего токена [4, 5]. Также метод на архитектуре Transformers способен заполнять промежуточные кадры, добавляющие плавность, бесшовность и правдоподобность движению вдоль заданного промптом направления. Из-за удобства архитектуры модель можно конвертировать в кросс-платформенный формат ONNX, который поддерживается многими крупными игровыми движками и 3D-пакетами. Данный метод способен испытывать трудности при генерации движений, радикально меняющих основное положение тела, например, вращение. Тем не менее, иерархическая структура восприятия токенов позволяет получать оптимальный результат анимации за короткий промежуток времени. В основе Masked Motion Model лежит MotionCLIP – 3D автокодер движений, обученный воссоздавать образы с помощью естественного языка [6].

Альтернативой всем методам способна выступить процедурная анимация – подход, достигающийся за счет строгих компьютерных алгоритмов и программных математических вычислений. Процедурное придание трехмерному объекту движения, в отличие традиционных подходов, использует инверсивную кинематику для автоматического

позиционирования соединений костей рига, благодаря чему, обеспечивает более разнообразную последовательность действий и взаимодействий в реальном времени с особенностями виртуальной окружающей среды [7]. Это позволяет допускать серию различных сценариев поведения 3D-объекта. Для создания процедурной анимации аниматор определяет набор операций, необходимых для достижения движения объекта, задаются исходные условия, которые приспособляются путем регуляции физических параметров в целях установления гибкости контроля и изменения позиции, формы и движения объекта. Приоритетом процедурализма становится способ решения рабочих задач, которые затруднительны, объемны или требуют множества итераций с большой точностью для выполнения вручную. В отличие от решений безмаркерного захвата движений на базе машинного обучения процедурный подход не требует предварительных условий для создания анимаций, может быть адаптивен под любого трехмерного персонажа, а также позволяет изменять последовательность действий и их направленность в режиме реального времени. Данный способ будет уместен как для получения инновационного результата, так и для решения повторяющихся задач.

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 [8]. 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 промпты как основу для генерации анимации, сумеют затмить решения, в основе которых лежит компьютерное зрение.

По результатам исследования были проанализированы методы, упрощающие создание анимаций, релевантность применения процедурной 3D-анимации. В перспективе, к системе процедурной анимации возможно добавить ИИ, который бы опирался на генерацию того или иного движения на основе пользовательских промптов.

Список литературы:

1. Мусони П. и др. Перенос функционального скелета модели // Сборник трудов АСМ по компьютерной графике и интерактивным технологиям. – 2021. – Т. 4. – №. 3. – С. 1-15.
2. Егоров М. С. и др. РАЗРАБОТКА ПРОГРАММНОГО МОДУЛЯ ОБРАБОТКИ ВИДЕОДАНЫХ НЕЙРОННЫМИ СЕТЯМИ ДЛЯ ПОЛУЧЕНИЯ КЛЮЧЕВЫХ ТОЧЕК ПОЛОЖЕНИЯ ТЕЛА ЧЕЛОВЕКА В ПРОСТРАНСТВЕ // Успехи в химии и химической технологии. – 2023. – Т. 37. – №. 4 (266). – С. 36-38.
3. Холден Д. и др. Обучаемое соответствие движений // АСМ Материалы по вопросам графики (ТОГ). – 2020. – Т. 39. – №. 4. – С. 53: 1-53: 12.
4. Пинуоаунтапонг Е. и др. МММ: Генеративная модель движения с маскированием // arXiv preprint arXiv:2312.03596. – 2023.
5. Гуо С. и др. MoMask: Генеративное моделирование трехмерных движений человека с помощью маскирования // arXiv preprint arXiv:2312.00063. – 2023.

6. Тевет Г. и др. Motionclip: Демонстрация генерации человеческих движений в клиповом пространстве //Европейская конференция по компьютерному зрению. – Cham : Springer Nature Switzerland, 2022. – С. 358-374.
7. Янг П. С. и др. Генерация гуманоидных движений робота на основе подхода обратной кинематики процедурной анимации //2022 IEEE/SICE Международный симпозиум по системной интеграции (SII). – IEEE, 2022. – С. 491-498.
8. Джордон Д. и др. Синтетические данные – что это, зачем и как? //arXiv preprint arXiv:2205.03257. – 2022.

References:

1. Musoni P. et al. A functional skeleton transfer //Proceedings of the ACM on Computer Graphics and Interactive Techniques. – 2021. – Т. 4. – №. 3. – P. 1-15.
2. Egorov M. S. et al. DEVELOPMENT OF A SOFTWARE MODULE FOR PROCESSING VIDEODATA BY NEURON NETWORKS FOR RECOGNIZING KEY POINTS OF HUMAN BODY POSITION IN SPACE // Advances in Chemistry and Chemical Technology. - 2023. - Т. 37. - №. 4 (266). - P. 36-38.
3. Holden D. et al. Learned motion matching //ACM Transactions on Graphics (TOG). – 2020. – Т. 39. – №. 4. – P. 53: 1-53: 12.
4. Pinyoanuntapong E. et al. MMM: Generative Masked Motion Model //arXiv preprint arXiv:2312.03596. – 2023.
5. Guo C. et al. MoMask: Generative Masked Modeling of 3D Human Motions //arXiv preprint arXiv:2312.00063. – 2023.
6. Tevet G. et al. Motionclip: Exposing human motion generation to clip space //European Conference on Computer Vision. – Cham : Springer Nature Switzerland, 2022. – P. 358-374.
7. Yang P. C. et al. Generating Humanoid Robot Motions based on a Procedural Animation IK Rig Method //2022 IEEE/SICE International Symposium on System Integration (SII). – IEEE, 2022. – P. 491-498.
8. Jordon J. et al. Synthetic Data--what, why and how? //arXiv preprint arXiv:2205.03257. – 2022.