

УДК 519.688

Обзор систем захвата движения для лицевой анимации моделей Unreal Engine Metahuman

Витюков Федор Андреевич

fedor.vityukov@bmstu.ru

МГТУ им. Н.Э. Баумана, Москва, Россия

Работа посвящена исследованию программных и аппаратных систем захвата движения (motion capture) для создания лицевой анимации персонажей в Unreal Engine. Важное требование к рассматриваемым продуктам — доступность для малых компаний и минимизация затрат при высоком качестве создаваемой анимации. Рассмотрены методы захвата движения на основе двух видеокамер, мобильных устройств на Android и iOS со специализированным программным обеспечением. Результаты исследования показали, что программно-аппаратные комплексы на основе iPhone 11-й модели и выше и бесплатного приложения Live Link Face в сочетании с визуализацией на Unreal Engine 5 и генерацией персонажей в MetaHuman дают лучший результат.

Ключевые слова: захват движения, motion capture, лицевая анимация, character animation, Unreal Engine, Android, iOS

Введение. Создание лицевой анимации — одна из задач, возникающих при разработке современных интерактивных приложений с 3D-графикой. Для ее решения используются программно-аппаратные комплексы разной ценовой категории и уровня сложности. Ключевым критерием работы систем захвата движения является точность, определяемая соответствием исходного движения и результата, воспроизводимого на экране после импорта и обработки записанного массива данных.

Материалы и методы. Существует множество программно-аппаратных комплексов, позволяющих решать задачу захвата движения для создания лицевой анимации. Разные ценовые категории систем создают широкий охват клиентов, способных их использовать: от крупного бизнеса со своими студиями до малых компаний и энтузиастов, способных проводить данные работы в неспециализированных помещениях после подготовки аппаратуры. Серьезной проблемой является отсутствие открытой детальной информации о методах работы профессиональных студий по данной тематике.

В базовом ценовом сегменте для малого бизнеса и энтузиастов представлены следующие варианты.

1. Две видеокамеры, расположенные на фиксированном расстоянии друг от друга. Специализированное программное обеспечение проводит калибровку видеокамер и захват движения для анимации.

2. Камера на мобильном устройстве с ОС Android и специализированное программное обеспечение.

3. Камера на мобильном устройстве с ОС iOS и специализированное программное обеспечение.

4. Специальный шлем, на котором установлены камеры на определенном расстоянии от лица, связанный с вычислительным комплексом, производящим запись и обработку данных.

В рамках данной работы была поставлена цель исследовать результаты захвата движения лица, полученные в вариантах 1–3. В ходе исследования были задействованы 14 программных комплексов, подключенных в различных сочетаниях к аппаратному обеспечению из данных вариантов.

Из бесплатных программных комплексов были исследованы:

- Live Link Face [1];
- iFacialMocap [2];
- VSeeFace [3];
- StrongTrack;
- FaceShift;
- DSSbodytrackor2.2;
- MediaPipe4U [4].

Из платных программных комплексов были исследованы:

- iFacialMocapPoweredbyNVIDIA;
- Faceware;
- MocapForAll;
- LuppetX;
- Dollars MoCap;
- Avatary;
- WebcamMotionCapture.

Для визуализации результата используется движок Unreal Engine. Благодаря Unreal Engine MetaHuman Creator есть возможность создавать качественные, анатомически корректные, высокодетализированные модели людей и их лиц.

Известно, что существует решение для iOS и iPhone в виде программы Live Link Face, позволяющей в реальном времени транслировать движения лица в анимацию мимики трехмерной модели, созданной в MetaHuman Creator. Проблема в том, что экосистема iOS является закрытой и требует больших вложений, так как стоимость iPhone и сопутствующих устройств достаточно велика, по сравнению с решениями, построенными на Android, либо на системе из нескольких камер. Поэтому было принято решение провести поиск программ, работающих на ОС Android и Windows и дающих качественный результат в связке с моделями, созданными в Unreal Engine MetaHuman Creator.

Результаты. Было проведено исследование всех вышеперечисленных программных комплексов. Кроме работы с мимикой также оценивалась возможность захвата движения рук и пальцев.

Ниже приведены выводы по отдельно взятым программным комплексам из исследуемого списка. CPU, на котором проводились исследования, — Intel Core i3-12100F.

iFacialMocap: поддерживает только iPhone. Учитывая существование Live Link Face, нет смысла работать с данной программой.

VSeeFace: достаточно развитое программное решение, поддерживает ОС Windows, но требует установки дополнительного программного обеспечения для трансляции захвата движения в модель MetaHuman на Unreal Engine. Результат захвата движения ниже среднего, но подойдет для любительских целей. В связке с Unreal Engine MetaHuman использование данной программы не имеет смысла для решения задачи качественного захвата движения.

DSSbodytrackor2.2: программа дает довольно точное отслеживание лица, глаз, рук, пальцев с обычной камеры. Но все движения 3D-модели идут с большой задержкой. При этом программа дает загрузку процессора до 50 %, что может не позволить совместить ее использование с работой других программ в реальном времени. Поддерживает открытый формат OSC. В связке с Unreal Engine MetaHuman использование данной программы не имеет смысла для решения задачи качественного захвата движения.

MediaPipe4U: самое рабочее бесплатное решение, транслирующее данные напрямую в Unreal Engine и подключающееся к моделям MetaHuman. Качество трекинга — ниже среднего, потребление CPU около 50 %. Вариант для системы самого начального уровня, использующий изображение с обычной камеры. По-настоящему качественный результат с помощью данного ПО получить практически невозможно.

Далее рассмотрим платные программные комплексы.

Dollars MoCap: транслирует данные напрямую в Unreal Engine в 3D-модель MetaHuman, без необходимости в дополнительном ПО. Захват лицевой анимации производится с помощью web-камеры, но результату не хватает детализации, теряются многие аспекты мимики. Поэтому можно назвать его неудовлетворительным, по сравнению с другими профессиональными аналогами. При этом захват движения тела, рук и пальцев — на сравнительно хорошем уровне, учитывая, что источник данных — только одна web-камера. Также большим минусом данного ПО является потребление до 60 % времени работы CPU.

Avatary: профессиональное программное обеспечение высокого уровня, работающее только со специальным шлемом от фирмы-производителя данного ПО, несмотря на информацию на официальном сайте о том, что оно способно работать с любой видеокамерой.

WebcamMotionCapture: программа обеспечивает сравнительно качественный трекинг пальцев относительно аналогов и загружает CPU на 25 %. Для стриминга данных в Unreal Engine может потребоваться установка дополнительного ПО.

Обсуждение полученных результатов. Полученные результаты говорят о том, что важными факторами для качественного захвата мимики являются:

1) Источник данных: кроме плоского изображения, получаемого с обычной камеры, для улучшения работы модели распознавания движения необходимо наличие дополнительной камеры с возможностью считывания глубины.

iPhone, начиная с модели iPhone X либо iPad Pro 3-го поколения [5] обладают камерой TrueDepth, поддерживающей данный функционал.

2) Полученные при считывании данные должны быть интерпретированы и адаптированы под 3D-модели MetaHuman. Без глубокого понимания устройства 3D-модели MetaHuman добиться качественного результата, используя сторонние решения, достаточно сложно.

Стоит отметить, что в Live Link Face есть два режима работы: стриминг данных в реальном времени и запись данных с последующей длительной постобработкой в MetaHuman Animator. Использование режимов зависит от требуемого результата и вычислительных возможностей аппаратной базы. Для качественных детализированных CG-роликов требуется постобработка в MetaHuman Animator. Для стриминга в реальном времени может быть достаточно стандартного режима. Он требует очень хорошего освещения помещения, в котором ведется запись, чтобы минимизировать «дрожание» считываемых с лица данных.

Проводить работу можно начиная с модели iPhone X, но лучший результат система дает с iPhone 11 и выше.

Заключение. В ходе исследования было выявлено, что программно-аппаратные комплексы на основе iPhone, начиная с 11 модели, и бесплатного приложения Live Link Face в сочетании с визуализацией на Unreal Engine 5 и генерацией персонажей в MetaHuman, дают лучший результат в рамках решения задач захвата движения лица и качественного рендеринга.

Список источников

- [1] Animating with Live Link. URL: <https://dev.epicgames.com/documentation/en-us/metahuman/animating-metahumans/livelinek> (accessed October 09, 2024).
- [2] iFacialMocap. URL: <https://www.ifacialmocap.com/tutorial/unrealengine> (accessed October 9, 2024).
- [3] VSeeFace. URL: <https://www.vseeface.icu/#virtual-camera> (accessed October 9, 2024).
- [4] MediaPipe4U. URL: <https://github.com/endink/Mediapipe4u-plugin> (accessed October 9, 2024).
- [5] Animating with Live Link. URL: <https://dev.epicgames.com/documentation/en-us/metahuman/animating-metahumans/livelinek> (accessed October 9, 2024).

Motion capture systems review for facial animation of unreal engine MetaHuman models

Vityukov Fedor A.

fedor.vityukov@bmstu.ru

BMSTU, Moscow, Russia

This paper is devoted to the study of software and hardware motion capture systems for creating facial animation of characters in Unreal Engine. An important requirement for the considered products is availability for small companies and minimization of costs at high quality of created animation. Motion capture methods based on two video cameras, Android and iOS mobile devices with specialized software are considered. The results of the study showed that complexes based on iPhone 11 model and above and free Live Link Face application in combination with visualization on Unreal Engine 5 and character generation in MetaHuman give the best result.

Keywords: motion capture, facial animation, character animation, Unreal Engine, Android, iOS