Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Факультет інформаційних радіотехнологій та технічного захисту інформації
(повна назва)

Кафедра медіаінженерії та інформаційних радіоелектронних систем
(повна назва)

КВАЛІФІКАЦІЙНА РОБОТА
Пояснювальна записка

рівень вищої освіти другий (магістерський)
(позначення документа)

Методи використання трекінгу у відеоконференціях
(тема)

Виконала:
студентка 2 курсу, групи СТМм-21-1
Анна ГРЕЧКО.
(прізвище, ініціали)

Спеціальність 171 Електроніка
(код і повна назва спеціальності)

Тип програми освітньо-професійна
(освітньо-професійна або освітньо-наукова)

Освітня програма Системи, технології і комп'ютерні засоби мультимедіа
(повна назва освітньої програми)

Керівник доц. Віталій ПОСОШЕНКО
(посада, прізвище, ініціали)

Допускається до захисту

Зав. кафедри _____ Володимир КАРТАШОВ
(підпис) (прізвище, ініціали)

2022 р.

Харківський національний університет радіоелектроніки

Факультет Інформаційних радіотехнологій та технічного захисту інформації
Кафедра Медіаінженерії та інформаційних радіоелектронних систем
Рівень вищої освіти другий (магістерський)
Спеціальність 171 Електроніка
(код і повна назва)
Тип програми освітньо-професійна
(освітньо-професійна або освітньо-наукова)
Освітня програма "Системи, технології і комп'ютерні засоби мультимедіа"

ЗАТВЕРДЖУЮ:

Зав. кафедри _____

(підпис)

«____»_____ 20 ___ р.

ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ

Студентці Гречко Анні Володимирівні
(прізвище, ім'я, по батькові)

1. Тема роботи Методи використання трекінгу у відеоконференціях
затверджена наказом по університету від " 21 " 11 2022 р. № 1503СТ
2. Термін подання студентом роботи 08.12.2022 р.
3. Вихідні дані до проекту (роботи)
1. Розглянути існуючі технології для трекінгу обличчя
2. Дослідити можливості використання сучасних алгоритмів для трекінгу обличчя в режимі реального часу
3. Створити відео та провести експеримент для визначення якості роботи трекеру при маркерному та безмаркерному методах
4. Перелік питань, що потрібно опрацювати в роботі
ВСТУП
1. Аналітичний огляд літературних джерел
2. Огляд існуючих систем для трекінгу обличчя з можливістю подальшої анімації
3. Розробка алгоритмів анімації тривимірних моделей
4. Створення відео та проведення досліджень про ефективність роботи різних методів трекінгу обличчя
5. Аналіз експериментальних даних
ВИСНОВКИ
ПЕРЕЛІК ДЖЕРЕЛ І ПОСИЛАНЬ
ДОДАТКИ

5. Перелік графічного матеріалу із зазначенням обов'язкових креслеників, схем, плакатів, комп'ютерних ілюстрацій:
1. Основні відомості про роботу; 2. Технологія ротоскопінгу; 3. Технологія захоплення руху; 4. Методи трекінгу; 5. Базові математичні моделі розпізнання; 6. Програмні пакети для трекінгу; 7. Трек-обробки обличчя в After Effects; 8. Заміна обличчя в After Effects; 9. Заміна обличчя у FaceRig; 10. Заміна обличчя у Snap Camera; 11. Схема розташування контрольних точок у Snap Camera; 12. Заміна обличчя у додатку Reface; 13-15. Створення та використання тривимірного обличчя у відео; 16, 18, 22. Трекінг обличчя в Faceware; 17, 19, 21 ,23. Порівняння траєкторій безмаркерного методу; 24. Висновки

КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів роботи	Термин виконання етапів роботи	Примітка
1.	Аналітичний огляд літературних джерел	21.11.22–24.11.22	
2.	Огляд існуючих систем для трекінгу обличчя з можливістю подальшої анімації	25.11.22–26.11.22	
3.	Розробка алгоритмів анімації тривимірних моделей	27.11.22–29.11.22	
4.	Створення відео та проведення досліджень про ефективність роботи різних методів трекінгу обличчя	30.11.22–01.12.22	
5.	Аналіз експериментальних даних	07.12.22–09.12.22	
6.	Графічна частина роботи	10.12.22–12.12.22	
7.	Перевірка керівником	13.12.22-14.12.22	
8.	Перевірка на академічний плагіат	14.12.22	
9.	Перевірка завідувачем кафедри, рецензування	15.12.22–17.12.22	

Дата видачі завдання 21.11.2022 р.

Студент _____ Анна ГРЕЧКО
(підпис)

Керівник роботи _____ доц. Віталій ПОСОШЕНКО
(підпис) (посада, прізвище, ініціали)

РЕФЕРАТ

Пояснювальна записка кваліфікаційної роботи має: 91 с., 82 рис., 2 додатка, 51 джерело.

MOTION CAPTURE, ТРЕКІНГ ОБЛИЧЧЯ, КОНТРОЛЬНА ТОЧКА, ПОХИБКА ТРАЄКТОРІЇ ТРЕКІНГУ, РОТОСКОПІНГ, АНІМАЦІЯ ТРИВИМІРНОЇ МОДЕЛІ, АНІМАЦІЯ ОБЛИЧЧЯ, FACEWARE, AFTER EFFECTS, SNAP CAMERA, DIPFAKE, REFACE

Об'єкт дослідження - системи для трекінгу обличчя.

Предмет дослідження – системи якісного трекінгу обличчя і можливості для їх використання в анімації тривимірних моделей.

Мета кваліфікаційної роботи – дослідити можливі засоби для трекінгу обличчя з подальшим використанням у відеоконференціях.

Методи дослідження – теоретичний аналіз, практичне дослідження, програмне моделювання, статистична обробка даних, аналіз отриманих результатів.

У даній роботі наведено етапи розвитку сучасної системи motion capture, яка використовується для анімації обличчя персонажів в кіно та комп’ютерних іграх, методи трекінгу обличчя (маркерний та безмаркерний, технічні засоби для їх реалізації, переваги та недоліки кожного з методів, математичні моделі трекінгу, програмні методи для трекінгу обличчя та порівняння їх ефективності, алгоритм трекінгу та заміни обличчя в програмному середовищі After Effects, додатки та програми для заміни обличчя та накладення ефектів у режимі реального часу та на заготовлених відео- та фотоматеріалах, алгоритм створення тривимірної моделі людини у середовищі Metahumans, інтеграція моделі в програмне забезпечення для створення ігор Unreal Engine, анімація обличчя моделі за допомогою спеціального мобільного додатку, результати експериментів з використанням маркерного та безмаркерного методів трекінгу

обличчя за допомогою плагіну Faceware, порівняння траєкторій руху та оцінки похибок між ними, висновки до кваліфікаційної роботи.

Актуальність: необхідність дослідження методів використання трекінгу в відеоконференціях – це лише перший крок. Досягнувши якісного результату трекінгу дозволить не тільки керувати мімікою обличчя тривимірної моделі під час онлайн-трансляцій у відео-чатах, а й створювати анатомічно достовірних персонажів для кіно та ігрового контенту. Це допоможе значно знизити собівартість розробки ігор, тому що не доведеться знімати кожного героя у костюмі motion capture. У кіноіндустрії дана технологія дозволить «продовжувати кар'єру» актора навіть після його смерті. Достатньо буде просто створити модель актора, відзняти його обличчя і реалістично анімувати. Саме тому, ця тема кваліфікаційної роботи є актуальною.

ABSTRACT

The explanatory note of the qualification work has: 91 pages, 82 figures, 2 appendix, 51 sources.

MOTION CAPTURE, FACE TRACKING, CONTROL POINT, TRAJECTORY ERROR, ROTOSCOPPING, 3D MODEL ANIMATION, FACE ANIMATION, FACEWARE, AFTER EFFECTS, SNAP CAMERA, DIPFAKE, REFACE

The object of research is face tracking systems.

The subject of research is systems for effective face tracking and possibilities for their use in the animation of three-dimensional models.

The purpose of the qualification work is to investigate possible means for face tracking with further use in video conferencing.

Research methods – theoretical analysis, practical research, software modeling, statistical data processing, analysis of the obtained results.

This work presents the stages of development of the modern motion capture system, which is used for the animation of the faces of characters in movies and computer games, methods of face tracking (marker and markerless, technical means for their implementation, advantages and disadvantages of each method, mathematical models of tracking, software methods for face tracking and comparing their effectiveness, the algorithm for tracking and replacing faces in the After Effects software environment, applications and programs for replacing faces and applying effects in real time and on prepared video and photo materials, the algorithm for creating a three-dimensional human model in the Metahumans environment , integration of the model into the software for creating games Unreal Engine, animation of the model's face using a special mobile application, results of experiments using marker and markerless face tracking methods using the Faceware plugin, comparison of movement trajectories and estimation of errors between them, conclusions to qualification work.

Relevance: The need to explore methods of using tracking in video conferencing is only the first step. Achieving a high-quality tracking result will allow not only to control facial expressions of a 3D model during online broadcasts in video chats, but also to create anatomically reliable characters for movies and game content. This will help to significantly reduce the cost of game development, because you will not have to shoot every character in a motion capture suit. In the film industry, this technology will allow the actor to "continue his career" even after his death. It will be enough to simply create a model of the actor, photograph his face and realistically animate it.

ПЕРЕЛІК УМОВНИХ СКОРОЧЕНЬ

FHD – Full High-Definition, 1920×1080

FPS – Frames per second

GIF – Graphics Interchange Format

HD – High-Definition, 1280×720

ID – Identifier

IP – Internet Protocol

LBPH – Local Binary Patterns Histograms

Mocap – Motion capture

OS (ОС) – Operation System (Операційна система)

VGA – Video Graphics Array, 640×480

Wi-Fi – Wireless Fidelity

ЗМІСТ

Перелік умовних скорочень 8

Вступ 11

1 ОГЛЯД ІСНУЮЧИХ ТЕХНОЛОГІЙ ДЛЯ ТРЕКІНГУ РУХІВ 13

1.1 Техніка ротоскопінгу та її вплив на подальший розвиток систем захоплення руху 13

1.2 Перший досвід захоплення руху 16

1.2 Базові поняття трекінгу 19

1.2.1 Маркерний метод трекінгу, трекпоінти та їх розташування на обличчі (Motion Bro) 19

1.2.2 Безмаркерний метод 25

1.3 Базові математичні моделі розпізнавання (трекінгу) 28

1.3.1 Ознаки Хаара 28

1.3.2 Методи Віоли-Джонса 30

1.4 Огляд програмних пакетів для трекінгу 32

1.4.1 Open CV Pyton 32

1.4.2 Faceware 33

1.4.4 After Effect (Motion Bro) 40

2 СТВОРЕННЯ ВІДЕО ТА ТЕСТУВАННЯ ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ 42

2.1 Етапи трек-обробки обличчя в After Effects 42

2.2 Техніка заміни обличчя під час онлайн-трансляції 47

2.2.1 Використання Snap Camera 49

2.2.2 Використання Google Meet 59

2.2.3 Використання Deepfake 61

3 СТВОРЕННЯ ТА ВИКОРИСТАННЯ ТРИВИМІРНОГО ОБЛИЧЧЯ У ВІДЕО 64

3.1 Створення надточної та реалістичної моделі людини, інтеграція її в програмне середовище для створення ігор та анімація обличчя 64

3.1.1 Створення моделі Metahumans 64

3.1.2 Інтеграція створеної моделі Metahumans в програмне середовище Unreal engine 67

3.1.3. Анімація обличчя за допомогою «Live Link Face» 75

3.2 Трекінг обличчя в Faceware 76

ВИСНОВКИ 86

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАНЬ 88

ДОДАТОК А 95

ДОДАТОК Б 107

ВСТУП

З початком карантину, коли всі сфери нашого життя невпинно переходили у цифровий режим, а пізніше і з початком повномасштабного вторгнення, відео-конференції, переговори та презентації стали невід'ємною частиною щоденного спілкування. Кінокомпанії змагаються між собою у створенні якіснішою комп'ютерної графіки. Ігрова індустрія з кожним роком покращує графіку, бажаючи максимально наблизити її до реального світу.

Така відома технологія як motion capture, свого часу була кращим рішенням для створення персонажів (фільмів чи ігор). Сьогодні прогрес не стоїть на місці і з'являються нові методи для анімації обличчя. У 2015 році вийшла сьома частина «Форсажу», яка була завершена вже після смерті головного актора – Пола Уокера. Тоді режисери прийняли рішення створити цифрову модель обличчя персонажа та анімувати його для використання у сценах, знятих крупним планом. Далі пішла компанія Marvel, яка працювала над створенням стрічки «Чорна пантера: Ваканда назавжди». Актор Чедвік Боусман помер від раку ще до початку зйомок. Аби не розчаровувати глядачів заміною, було прийнято рішення про створення цифрового двійника.

У кожному з цих підходів використовувались різні методи анімації обличчя. Але для того, щоб «приміряти маску» улюбленого героя, не обов'язково зніматися у багатобюджетному фільмі.

У даній кваліфікаційній роботі розглянуто декілька методів та програмних забезпечень для трекінгу та подальшої заміни обличчя або анімації моделі. Самостійно можна перетворитися на персонажу мультфільму або іншу людину і розважити співрозмовників під час відео-конференції. Тепер історію України може розповідати Нестор Літописець у режимі реального часу. Якщо ви втомилися, не маєте настрою чи просто хочете «закритися» від людей – достатньо просто накласти маску на обличчя і проводити онлайн-зустрічі.

Ціллю роботи є виявлення методів та програмного забезпечення для створення якісного контенту, який згодом може перерости відео-дзвінки і значно спростити роботу ігровій та кіноіндустрії.

Тому, тема кваліфікаційної роботи є актуальною.

1 ОГЛЯД ІСНУЮЧИХ ТЕХНОЛОГІЙ ДЛЯ ТРЕКІНГУ РУХІВ

Найвідомішою технологією для захвату руху є motion capture. Відстеження руху або захоплення руху розпочалося як інструмент фотограмметричного аналізу в дослідженнях біомеханіки в 1970-х і 1980-х роках і поширилося на освіту, навчання, спорт і нещодавно на комп'ютерну анімацію для телебачення, кіно та відеоігор по мірі розвитку технології.

1.1 Техніка ротоскопінгу та її вплив на подальший розвиток систем захоплення руху

Сучасні системи захоплення руху є продуктом століття тренувань, інновацій та обчислювальних процесів. Мосар народився задовго до того, як ми побачили Горлума на великому екрані у «Володарі перстнів», задовго до холодної війни, війни у В'єтнамі чи Другої світової війни. Це був 1915 рік, посеред Першої світової війни, коли аніматор Макс Флейшер розробив техніку під назвою Rotosciping і заклав основу для сучасної передової технології mocap.

Ротоскопінг був примітивним і трудомістким процесом, але це було необхідною відправною точкою для розвитку галузі. У методі Rotosciping аніматори стояли за скляним столом і кадр за кадром малювали проектований фільм живої дії, копіюючи дії акторів або тварин безпосередньо на намальований від руки світ. Ця техніка створювала плавні, реалістичні рухи, яких аніматори не могли досягти самостійно.

Першим повнометражним американським фільмом, який використовував ротоскопію, була «Білосніжка та сім гномів», яка дебютувала в 1939 році, і Disney використовував цю техніку в наступних фільмах, включаючи «Алісу в країні чудес» (рис. 1.1 – 1.2), «Сплячу красуню» та «Пітера Пен». До цього дана технологія застосовувалась у короткометражних фільмах. У 1914 році Максом Флейшером було винайдено і використано ротоскопінг у серії фільмів «Із чорнильниці»: його брат, Дейв Флейшер, наряджений в костюм клоуна, зіграв

персонажа фільму — клоуна Коко. Пізніше, на початку 30-х, Флейшер використовував ту ж техніку в мультфільмах про Бетті Буп [1].

Рисунок 1.1 – Використання технології Rotosciping у повнометражному анімаційному фільмі «Alice in Wonderland» компанії Disney, 1951 р

Рисунок 1.2 – Створення персонажа Божевільного Капелюшника за використання технології ротоскопінгу

Хоча до реальних систем mocap ще залишалися десятиліття, ротоскопіювання було доказом концепції перспективності цієї галузі.

Очевидно, що це був найдієвіший спосіб імітувати дії реальних людей в анімованих просторах.

Незважаючи на те, що ротоскопінг був винайдений ще на початку XX століття, зараз він також активно використовується. Замінивши художнє малювання на комп’ютерне, технологія активно використовується в виготовленні відеопродукції, ігровій індустрії і комп'ютерній анімації для створення фонового зображення або текстурних карт.

Стилізацію під ротоскопіювання доволі часто використовують кінематографісти — від телевізійних реклам до фільмів Голівуду. Ротоскопіювання використовується також в коміксах і в фотографії. Цей прийоми широко використовуються в комерційній ілюстрації.

Анімація серіалу «Скасування» (рис.1.3) була створена шляхом поєднання технологій захоплення руху і ротоскопіювання [2].

Рисунок 1.3 – Кадр з серіалу «Скасування» (2019-2021 рр), реж. Хіско Хулсінг

1.2 Перший досвід захоплення руху

У часи холодної війни піонер аналогової електронної анімації – Лі Гаррісон III починає експерименти з аналоговими схемами та електронно-променевими трубками. У 1959 році Гаррісон обшив спеціальний натільний костюм потенціометрами (регульованими резисторами) і зміг записувати та анімувати рухи актора в реальному часі на ЕПТ.

Це була елементарна установка (рис.1.4): анімований актор був, по суті, сяючою фігуркою, але це був перший випадок захоплення руху в реальному часі [3].

Рисунок 1.4 – Експеримент Лі Гаррісона III (1959 р)

1.3 Перехід до сучасної технології Motion capture

До 1980-х років аніматори використовували костюми, обшиті активними маркерами, і кілька великих камер для відстеження рухів акторів, в результаті чого цифрові зображення були набагато більш детальними та точними, ніж радіоактивні малюнки Гаррісона.

Але навіть у 1990-х роках кожна готова до mocap камера була досить габаритною (розміром приблизно з маленький холодильник), і аніматорам

доводилося вручну призначати кожен маркер у кожному кадрі для кожної сцени [4].

Це було майже такою ж кропіткою роботою, як і Rotoscoping.

До сьогодні, технологія захоплення рухів пройшла довгий шлях автоматизації і тепер процес виглядає так: захоплення рухів відбувається в приміщеннях або спеціально відбудованих декораціях, де немає сонячних променів, що важливо, оскільки дані отримують за допомогою інфрачервоних датчиків.

Але у фільму «Планета мавп: Революція» ця технологія зазнала змін. При зйомці вдень (рис.1.5) на місці, побудована на інфрачервоних променях, система дає збій — сонце «забиває» ці промені. Систему довелось модифікувати. Маркери стали активними, потім вони увімкнули радіосигнали, що підсилило надійність системи. Камери стали бездротовими та були заховані на дереві, а кінематографісти вирізалися на знімальній камері одночасно із записом захопленого руху [5].

Рисунок 1.5 – Використання технологій motion capture на відкритій місцевості при зйомках фільму «Планета мавп»

«Повідомлення та розуміння того, що таке захоплення рухів, було важко пояснити широкій публіці. Я завжди казав, що зйомка перформансу — це не вид акторської гри — це просто різні камери, які знімають актора. Ви буквально керуєте цифровою маскою» (рис.1.6). – каже Серкіс [6].

Рисунок 1.6 – Використання технологій motion capture для анімації обличчя персонажа Енді Серкіса у фільмі «Планета мавп»

1.2 Базові поняття трекінгу

1.2.1 Маркерний метод трекінгу, трекпоінти та їх розташування на обличчі (Motion Bro)

Одним із перших великих проектів, що використовують технологію захоплення рухів, був комп'ютерний мультфільм «Sinbad: Beyond the Veil of Mists» (1997 р). Усі людські персонажі були анімовані саме за допомогою motion capture з подальшою обробкою.

Система маркерів motion capture, де використовується спеціальне обладнання: на актора одягається костюм з датчиками, коли він відтворює потрібні рухи, то дані з датчиків фіксуються спеціальними камерами і надходять до комп'ютера, де зводяться в 3D модель, що відтворює дії актора, на основі якої пізніше (або в режимі реального часу) створюється анімація персонажа. Також цим методом створюється міміка актора.

На сьогоднішній день існує велика кількість маркерних систем захоплення рухів. Різниця між ними полягає у принципі передачі рухів.

Оптичні пасивні. На костюмі, що входить до комплекту такої системи, прикріплені датчики-маркери, які названі пасивними, тому що відображають тільки світло, що надійшло на них, але самі не світяться. У таких системах світло (інфрачервоне) на маркери посилається зі встановлених на камерах високочастотних стробоскопів (рис. 1.7) і, відбившись від маркерів, потрапляє назад в об'єктив камери, повідомляючи цим позицію маркера.

Рисунок 1.7 – Високочастотні стробоскопи компанії Vicon, що встановлені на камеру для оптичної пасивної маркерної системи

Мінус оптичних пасивних систем полягає в тривалості розміщення маркерів на акторі. Так само іноді при швидкому русі або близькому розташуванні маркерів одна до одної система може їх плутати (оскільки ця технологія не передбачає ідентифікації кожного маркера окремо) [7].

Активними названі оптичні системи, які замість світловідбивних маркерів, що кріпляться до костюма актора, використовують світлодіоди з інтегрованими процесорами та радіо-синхронізацією. Кожному світлодіоду призначається ID (ідентифікатор), що дозволяє системі не плутати маркери один з одним, а також впізнавати їх після того, як вони були перекриті і знову з'явилися в камері.

Зрештою, принцип роботи таких систем схожий з пасивними системами. Мінуси активних систем:

— відсутність можливості захоплення рухів та міміки обличчя;

— додатковий контролер, що кріпиться до актора і підключений до маркерів-світлодіодів, сковує його рухи;

— крихкість та відносно висока вартість маркерів-світлодіодів.

Механічні системи прямо стежать за згинами суглобів, для цього на актора надягається спеціальний механічний mocap-скелет (рис. 1.8), який повторює слідом за ним всі рухи. У комп'ютер у своїй передаються дані про кутах згинів всіх суглобів.

Мінуси механічних систем [8]:

— mocap-скелет, з додатковим контролером, прикріпленим до актора і підключеним до сенсорів згинів, а в деяких випадках і дроти, що тягнуться від кістяка, сильно обмежують рухи актора;
— відсутність можливості захоплення;
— рухів та міміки обличчя;
— рухів тісної взаємодії двох і більше акторів (боротьба, танці з підтримками і т.д.);
— рухів на підлозі - перекиди, падіння і т.д.;
— ризик поломки механіки за необережного використання.

Рисунок 1.8 – Механічний mocap-скелет

Гіроскопічні або інерційні системи для збору інформації про рух використовують мініатюрні гіроскопи та інерційні сенсори (рис.1.9), розташовані на тілі актора - так само як і маркери або магніти в інших mocap-системах. Дані з гіроскопів та сенсорів передаються до комп'ютера, де

відбувається їх обробка та запис. Система визначає як положення сенсора, а й кут його нахилу.

Мінуси гіроскопічних/інерційних систем:

— відсутність можливості захоплення рухів та міміки обличчя;

— додатковий контролер, прикріплений до актора і підключений до магнітних маркерів, або навіть зв'язування проводів, що тягнеться від актора до комп'ютера;висока вартість гіроскопів та інерційних сенсорів;

— для визначення положення актора у просторі потрібна додаткова міні-система (оптична чи магнітна) [9].

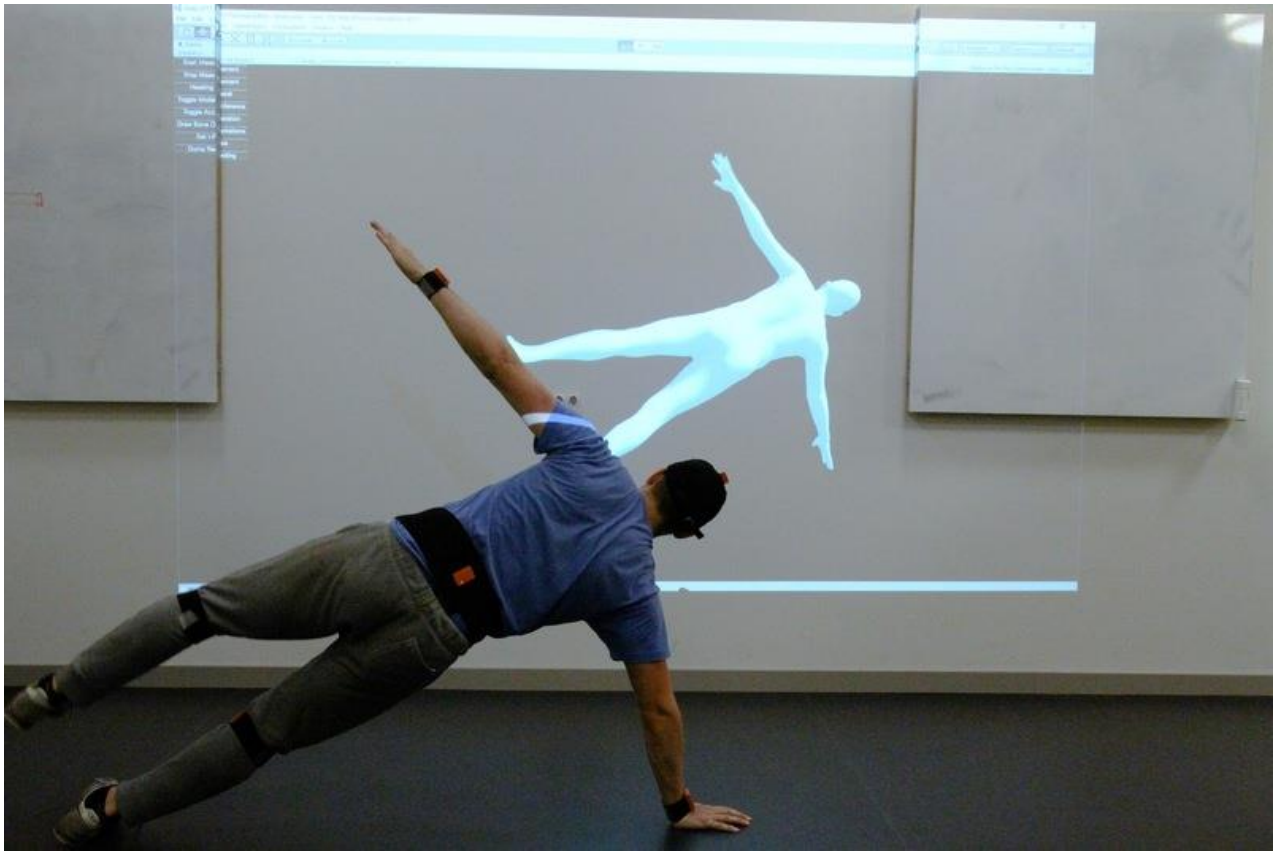

Рисунок 1.9 – Гіроскопічна система для технології motion capture

Магнітні системи, в яких маркерами є магніти, а камерами – ресивери, система вираховує їх позиції щодо спотворень магнітного потоку. Мінуси магнітних систем:

— магнітні системи схильні до магнітних та електричних перешкод від металевих предметів та оточення (електропроводки приміщення, оргтехніки, арматури в плитах будівлі);
— мінлива чутливість сенсорів залежно від їхнього положення в робочій зоні;
— найменша в порівнянні з оптичними системами робоча зона;
— відсутність можливості захоплення рухів та міміки обличчя;
— додатковий контролер, прикріплений до актора і підключений до магнітних маркерів, або навіть зв'язування проводів, що тягнеться від актора до комп'ютера;
— висока вартість магнітних маркерів.

Якщо ми кажемо про маркери для відтворення міміки персонажів, кіностудії використовують спеціальні світловідбиваючі сферичні маркери з клейкою основою (рис. 1.10) [10] або стікери з контрастними контрольними точками (рис. 1.11) [11].

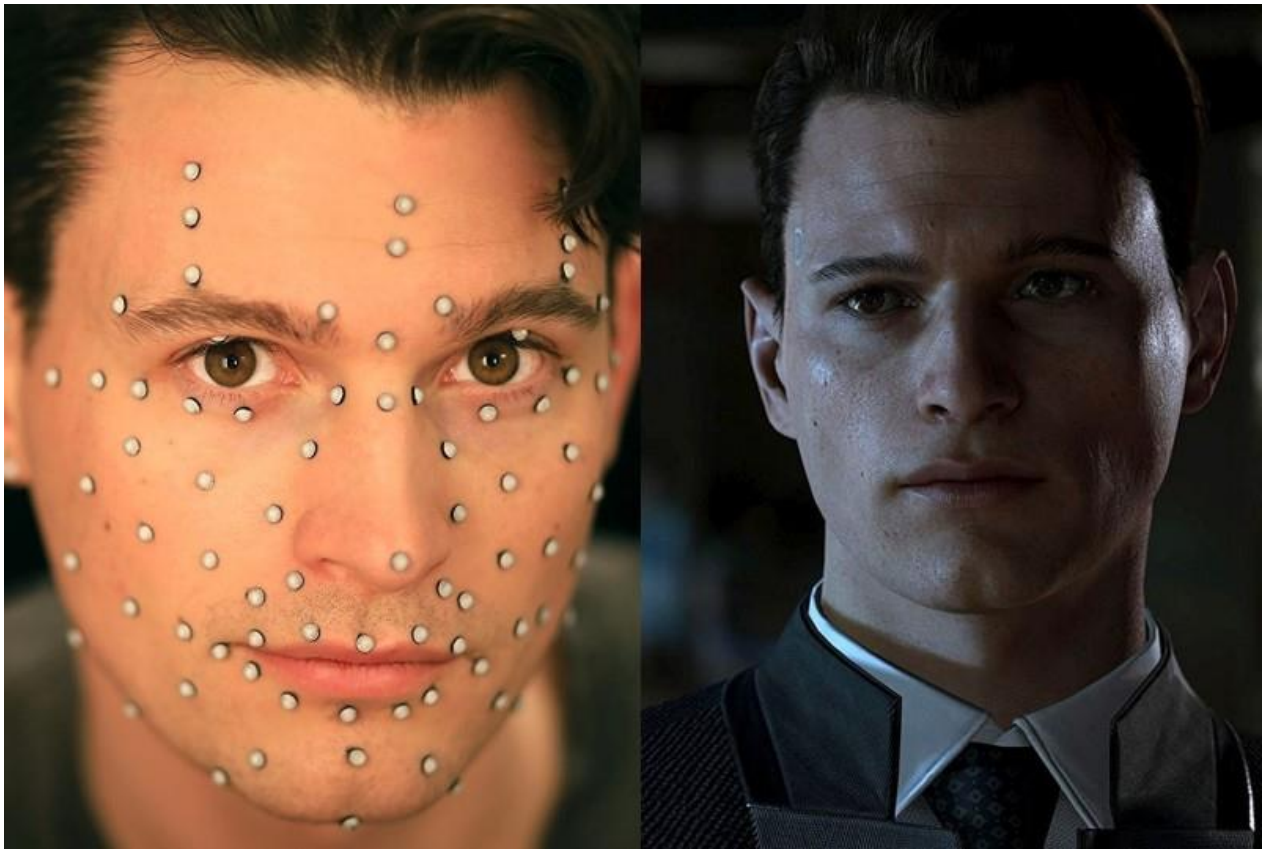

Рисунок 1.10 – Брайан Декарт під час зйомок для гри «Detroit: Become Human»

Рисунок 1.11 – Аарон Екхарт у ролі Харві Дента, також відомого як Two-Face, у фільмі "Темний лицар" 2008 року

У якості бюджетного варіанту можна використати фломастери або фарбу темних кольорів, намалювавши контрольні точки безпосередньо на обличчі (рис.1.12).

Рисунок 1.12 – Створення маркерів для обличчя у домашніх умовах [12]

1.2.2 Безмаркерний метод

Безмаркерний метод, який не потребує спеціальних датчиків чи спеціального костюма – ґрунтується на технологіях комп’ютерного зору та розпізнавання образів. Актор може зніматися у звичайному одязі, що сильно прискорює підготовку до зйомок і дозволяє знімати складні рухи (боротьба, падіння, стрибки, тощо) без ризику пошкодження датчиків або маркерів.

Безмаркерна технологія працює за особливими алгоритмами розпізнавання, де на обличчя актора, знятого камерою, накладається віртуальна «сітка». На цій сітці програмні алгоритми знаходять деякі опорні точки, якими визначають точне місце, якого буде «прив'язана» віртуальна модель. Перевага такої технології в тому, що об'єкти реального світу є маркерами самі по собі і для них не потрібно створювати спеціальних візуальних ідентифікаторів.

Ще одним представником безмаркерної технологіє є MotionScan, яка використовується, переважно, при створенні комп’ютерних ігор.

MotionScan — програмно-апаратна технологія, призначена для оптичного сканування обличчя людини та переведення відсканованих даних у цифровий формат. MotionScan дозволяє з високим ступенем достовірності відсканувати міміку обличчя людини та імпортувати її до програмних програм. MotionScan розроблена австралійською компанією Depth Analysis та була офіційно анонсована 4 березня 2010 року [13].

Технологія MotionScan використовує 32 камери високої роздільної здатності для захоплення реалістичного тривимірного руху з частотою 30 кадрів на секунду (рис.1.13). Апаратних засобів та обчислювальних потужностей системи MotionScan вистачає на захоплення та обробку 50 хвилин даних зйомки, при цьому на виході виходить високоякісна мультиплікація практично телевізійної якості. Технологія не вимагає використання спеціальних маркерів або флуоресціюючих покриттів, як цього вимагають деякі подібні технології, так само при обробці не потрібно ніякої ручної роботи з очищення зображень та підгонки тривимірних моделей. Для режисерів та операторів додатковою

перевагою технології MotionScan є можливість перегляду гри актора з будь-якого кута без необхідності використання додаткових камер та джерел освітлення [15].

Рисунок 1.13 – Студія з обладнанням для MotionScan

Першою грою, в якій було заявлено використання MotionScan, стала «L.A. Noire» розроблення студії Team Bondi (рис.1.14). Розробники розповідали про те, що жодна технологія захоплення міміки їм не підходила, оскільки ці технології не могли відсканувати міміку так, як вона виглядає в реальному житті. Творці гри говорили також і про те, що міміка персонажів сильно впливатиме на геймплей. Гравцю в ролі детектива на допитах доведеться визначати за мімікою підозрюваного, бреше він чи ні [14].

Рисунок 1.14 – Зліва направо: (1) Аарон Стейтон під час сеансу зйомки MotionScan (2) Початковий результат сканування (3) Скріншот у грі та кінцевий продукт Аарона Стейтона в ролі «Коула Фелпса» у грі «L.A. Noire» (4) Реальний вигляд Стейтона в його образі «Кена Косгроува» в «Божевільних».

Проаналізувавши ці два методи трекінгу обличчя (маркерний і безмаркерний), можна зробити наступні висновки:

1) Маркерний метод є більш доступним. Оскільки контрольні точки можна намалювати безпосередньо на обличчі підручними матеріалами. На відміну від безмаркерного метода, який потребує декілька десятків професіональних камер, з маркерним методом впорається звичайна камера, навіть телефонна.

2) Можна нанести маркери на ті частини обличчя, яким необхідно приділити найбільшу увагу. У програмних засобах безмаркерного методу є певний шаблон для контрольних точок, який не завжди підходить для поставлених цілей.

Тож, для подальшого дослідження було обрано маркерний метод трекінгу обличчя.

1.3 Базові математичні моделі розпізнавання (трекінгу)

Нині існує величезна різноманітність підходів для виявлення обличь. Найбільш досконалим методом розпізнавання на сьогоднішній день є згорточні нейронні мережі. Але вони досить складні у розробці, тому конкуренто спроможні рішення розробляються великими корпораціями (як, наприклад, розробка Facebook – DeepFace). У той же час існують алгоритми і рішення, які є сучасними і ефективними та можуть бути імплементовані без значних зусиль. У даної роботі досліджені ті методи, які можливо реалізувати у мобільних додатках при невеликих обсягах пам’яті. До них відносяться алгоритми Віоли-Джонса для пошуку об’єктів на зображені та алгоритми Eigenfaces і LBPH для розпізнавання обличь на фотографіях [16].

1.3.1 Ознаки Хаара

Інтегральне представлення зображення використовується для обчислення ознак Хаара. Ознаки Хаара – ознаки цифрового зображення, використовувані в розпізнаванні образів. Своєю назву вони зобов’язані інтуїтивним схожістю з вейвлетами Хаара. Ознаки Хаара використовувалися в першому детекторі облич, що працював в реальному часі [17].

Ознака Хаара складається з суміжних прямокутних областей. Вони позиціонуються на зображенні, далі сумуються інтенсивності пікселів в областях, після чого обчислюється різниця між сумами. Ця різниця і буде значенням певної ознаки, визначеного розміру, певним чином позиційованої на зображенні. Загальним для всіх зображень є те, що область в районі очей темніше, ніж область в районі щік. Отже загальною ознакою Хаара для обличь є 2 суміжних прямокутних регіону, що лежать на очах і щоках (рис. 1.15).

Рисунок 1.15 – Ознаки Хаара і приклад їх розташування на обличчі людини

Ключовою особливістю ознак Хаара є найбільша, в порівнянні з іншими ознаками, швидкість. Обчислюється значення такої ознаки за формулою:

$$F = X - Y, \quad (1.1)$$

де X – сума значень яскравості точок закриваються світлою частиною ознаки, Y – сума значень яскравості точок що закриваються темною частиною ознаки. Ознаки Хаара дають точкове значення перепаду яскравості по осі X і Y відповідно.

Алгоритм сканування вікна з ознаками виглядає так:

— обирається вікно сканування та використовувані ознаки для основних рис обличчя;

— вікно сканування послідовно просувається по зображенню з кроком в 1 клітинку вікна (припустимо, розмір самого вікна 24x24 клітинки);

— при скануванні зображення в кожному вікні обчислюється приблизно 200 000 варіантів розташування ознак, за рахунок зміни масштабу ознак і їх положення в вікні сканування;

— сканування проводиться послідовно для різних масштабів скануючого вікна;

— всі знайдені ознаки потрапляють до класифікатору, який фіксує наявність рота, носа, та іншої ознаки певної розмірності у певному регіоні і знаходить найближчий зразок у базі.

Класифікатор повинен реагувати тільки на певну, підмножину всіх ознак, якою він навчається з навчальної вибірки. У якості класифікатора може бути перцепторн, ієрархічні алгоритми та інші алгоритми класифікації.

Вхідні дані: розмір та розташування очей, розмір та розташування рота, розмір та розташування інших головних рис. Для поліпшення точності використовується технологія бустингу. Бустинг – комплекс методів, що сприяють підвищенню точності аналітичних моделей [18]. В його основі лежить побудова каскада класифікаторів, кожен з яких (крім першого) навчається на помилках попереднього. Наприклад, один з перших алгоритмів бустинга використовував каскад з 3-х моделей, перша з яких навчалася на всьому наборі даних, друга – на вибірці прикладів, в половині з яких перша дала правильні відповіді, а третя – на прикладах, де «відповіді» перших двох розійшлися.

Таким чином, має місце послідовна обробка прикладів каскадом класифікаторів, причому так, що завдання для кожного наступного стає важче. Результат визначається шляхом простого голосування: приклад відноситься до того класу, який виданий більшістю моделей каскаду.

1.3.2 Методи Віоли-Джонса

Метод, запропонований Віолою і Джонсом в 2001 році, став справжнім проривом в цій області. Цей метод набув великої популярності завдяки високій точності і серйозної теоретичної основі [18, 19, 20].

Основні принципи, на яких базується метод:

— використовуються зображення в інтегральному уявленні, що дозволяє швидко обчислювати необхідні об’єкти;
— використовуються ознаки Хаара, за допомогою яких відбувається пошук потрібного об’єкта (в цьому контексті, обличчя і його рис);
— використовується бустинг для вибору найбільш підходящих ознак для шуканого об’єкта на цій частині зображення;

— всі ознаки надходять на вхід класифікатора, який дає результат «так» або «ні»;

— використовуються каскади ознак для швидкого відкидання вікон, де не знайдено обличчя.

Найбільш популярний цей метод для пошуку об’єктів у відео потоці. Також цей детектор має вкрай низьку ймовірністю помилкового виявлення обличчя. Алгоритм навіть добре працює і розпізнає риси обличчя під невеликим кутом, приблизно до 30 градусів. При куті більше 30 градусів точність різко падає. Для того, щоб проводити будь-які дії з даними в методі Віоли-Джонса, використовується інтегральне представлення зображень. Однією з корисних особливостей інтегрального представлення є можливість дуже швидко вирахувати суму пікселів довільного прямокутника (або будь-який інший фігури, яку можна апроксимувати декількома прямокутниками).

Інтегральне представлення зображення – це матриця, що збігається за розмірами з вихідним зображенням. У кожному елементі її зберігається сума інтенсивності яскравості всіх пікселів, що знаходяться лівіше і вище даного елемента. Елементи матриці розраховуються за формулою:

$$L(x, y) = \sum_{i=0, j=0}^{i \le x, j \le y} I(i, j), \quad (1.2)$$

де I(i, j) – яскравість пікселя вихідного зображення.

Кожен елемент матриці L(x, y) є сумою пікселів в прямокутнику від (0,0) до (x,y), тобто значення кожного пікселя (x, y) дорівнює сумі значень всіх пікселів лівіше і вище даного пікселя (x, y). Розрахунок матриці займає лінійний час, пропорційне числу пікселів в зображенні, тому інтегральне зображення прораховується за один прохід. Розрахунок матриці можливий за формулою:

$$L(x, y) = I(x, y) - L(x-1, y-1) + L(x, y-1) + L(x-1, y) \quad (1.3)$$

Розглянуті математичні моделі можна використовувати при навчанні нейронних мереж розпізнаванню обличь із подальшим використанням в роботі БПЛА [21-39].

1.4 Огляд програмних пакетів для трекінгу

1.4.1 Open CV Pyton

OpenCV — це бібліотека, яка використовується для обробки зображень за допомогою мов програмування, таких як Python.

OpenCV є найпопулярнішою бібліотекою комп'ютерного зору. Спочатку написаний мовою C/C++, тепер він забезпечує прив’язки для Python.

OpenCV використовує алгоритми машинного навчання для пошуку обличь на зображенні. Оскільки обличчя дуже складні, немає жодного простого тесту, який скаже, знайдено воно чи ні. Натомість є тисячі дрібних шаблонів і особливостей, які повинні бути узгоджені. Алгоритми розбивають завдання ідентифікації обличчя на тисячі дрібніших завдань, кожне з яких легко вирішити. Ці завдання називають класифікаторами.

Для таких складних елементів, як обличчя, може бути 6000 або більше класифікаторів, усі вони мають збігатися, щоб обличчя було виявлено (звичайно, у межах похибок). Але в цьому полягає проблема: для виявлення обличчя алгоритм починається у верхньому лівому куті зображення та рухається вниз через невеликі блоки даних, переглядаючи кожен блок, постійно запитуючи: «Це обличчя? … Це обличчя? … Це обличчя?» Оскільки на блок припадає 6000 або більше тестів, можуть виникати мільйони обчислень, які призведуть до зупинки комп’ютера.

Щоб запобігти цьому, OpenCV використовує каскади. Подібно серії водоспадів, каскад OpenCV розбиває проблему виявлення обличь на кілька етапів. Для кожного блоку виконується дуже приблизний і швидкий тест. Якщо це проходить, виконується детальніший тест і так далі. Алгоритм може мати від 30 до 50 таких етапів або каскадів, і він виявить обличчя, лише якщо всі етапи пройдено [40].

Перевагою є те, що більша частина зображення повертатиме негатив на перших кількох етапах, а це означає, що алгоритм не витрачатиме час на

перевірку всіх 6000 функцій на ньому. Замість того, щоб зайняти години, розпізнавання обличчя тепер можна виконувати в режимі реального часу.

Хоча теорія може здатися складною, на практиці це досить легко. Самі каскади — це лише купа файлів XML, які містять дані OpenCV, які використовуються для виявлення об’єктів. Необхідно лише ініціалізувати написаний код потрібним каскадом, а потім він виконує роботу без участі людини.

Оскільки розпізнавання обличь є досить поширеним випадком, OpenCV постачається з низкою вбудованих каскадів для виявлення всього: від обличь до очей, рук і ніг.

Кілька нюансів бібліотеки OpenCV, на які варто звернути увагу:

— виявлення працює лише на зображеннях у градаціях сірого, тому важливо перевести кольорове зображення у градації сірого;

— функція detectMultiScale використовується для виявлення обличь. Вона має 3 аргументи — вхідне зображення, scaleFactor і minNeighbours. scaleFactor робить, менш зменшує розмір зображення з кожним масштабом. minNeighbours вважає, скільки сусідів має мати кожен кандидат-прямокутник, щоб зберегти його;

— faces містить список координат для прямокутних областей, де були знайдені обличчя. Ми використовуємо ці координати, щоб намалювати прямокутники на нашому зображенні [41].

1.4.2 Faceware

Faceware Technologies — компанія, що займається захопленням рухів обличчя, яка розробляє апаратне та програмне забезпечення для MoCap обличчя в режимі реального часу та офлайн. Faceware протягом багатьох років виробляє конвеєри захоплення руху обличчя. Спочатку команда припускала, що їхня робота буде використана для попереднього перегляду, але ігрова спільнота була першою спільнотою, яка дійсно відреагувала на інструменти компанії. Сьогодні

компанія пропонує ряд рішень, які поєднують програмне та апаратне забезпечення. Їхня технологія нейронної мережі Faceware Studio може відстежувати та анімувати будь-яке зображення обличчя, чи то з камери телефону в реальному часі, чи з попередньо записаного відео. Faceware Studio підключається до Unreal Engine через безкоштовний плагін Live Link, що робить його ідеальним способом застосування даних реалістичної анімації обличчя до персонажа MetaHuman у реальному часі.

На відміну від деяких рішень завдяки додаванню машинного навчання (ML), Faceware здатне забезпечувати високу точність реакцій обличчя за допомогою лише однієї камери, а не стереопари. Він також обходиться без інфрачервоного випромінювання, яке, теоретично, значно ускладнює розпізнавання облич через природні тіні, які відкидають плечі камери або підставка головного пристрою. Засновник компанії Кріс Джонс коментує свою розробку наступними словами: «Я хотів створити цю систему так, щоб вона була якомога куленепробивнішою на знімальному майданчику. Це просто HD-SDI, який просто «ввімкнено» — весь час. Коли ви ставите батарею, налаштувань камери немає, ви не можете її випадково вимкнути, змінити її налаштування при перезавантаженні».

Монокамера даної технології використовує машинне навчання для визначення положення щелепи, те, що інші системи намагаються вирішити за допомогою пари камер або стереопари в передній частині шолома. Наявність однієї камери не повинно дати таких хороших результатів, але ML вражаюче підвищує результати. Ще однією перевагою наявності лише однієї камери є менша вага. Важкі установки викликають підвищене напруження шиї, а додаткова вага вимагає вищого шолома, щоб боротися з хитанням навантаженої руки камери. Шолом Mark IV (рис.1.16), представлений компанією Faceware Technologies легкий і зручний [43].

Рисунок 1.16 – Шолом Mark IV з однією камерою для захвату обличчя від компанії Faceware Technologies

Відео з камери надходить як кольорове відео з низьким рівнем спотворень, навіть якщо камера розташована близько до обличчя художника (рис.1.17). Це дуже важливо для будь-якої подальшої перевірки. Записане відео стає дуже корисним довідковим матеріалом для аніматорів, які прагнуть налаштувати або перевірити вирази в офлайн-конвеєрах. Майже всі інші установки надають широко спотворене монохромне, примарне ІЧ-зображення як стандарт, за яким дуже важко оцінити продуктивність, коли пізніше воно використовується як еталон.

Рисунок 1.17 – Передача кольорового відео з камер на шоломах акторів

Головні камери Mark IV не оснащені вбудованим мікрофоном, але аудіо можна вставити в те саме джерело відео обличчя, підключивши аудіоканал до аудіовходів на задній панелі рекодера AJA Ki Pro Rack, який є частиною робочої системи.

У режимі реального часу можна захопити до 5 акторів бездротовим способом. Для дротових налаштувань немає обмежень, крім кількості доступних цифрових відеомагнітофонів.

Камера MarkIV записує до 1080 кадрів в секунду, але виробник часто рекомендує 720, цього буде достатньо для гарного захоплення. Програмне забезпечення Faceware не пов’язане або обмежене роздільною здатністю камер.

Варто зазначити, що камера знімає портрет, а не ландшафт, тому обличчя актора знімається «збоку» (рис.1.18), щоб максимально використати роздільну здатність [43].

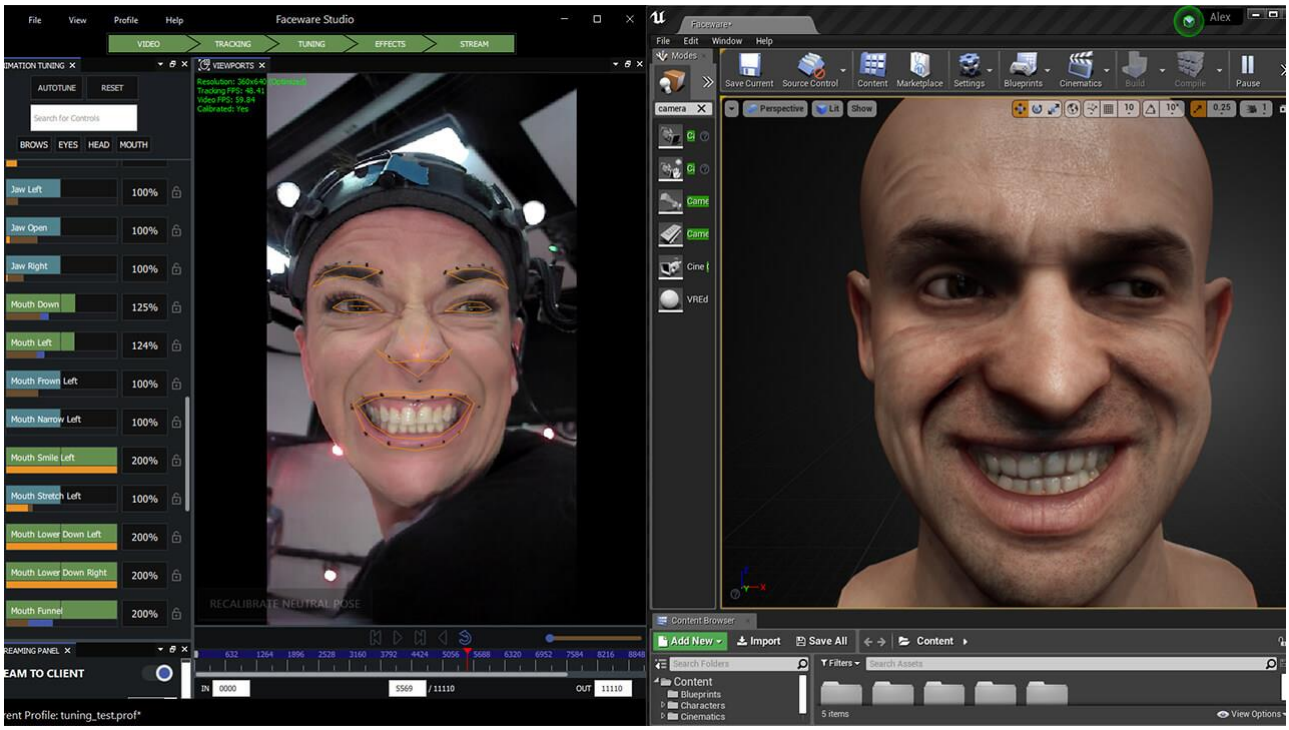

Рисунок 1.18 – Запис відео з шолома акторки та використання її міміки для анімування обличчя персонажа в системі Faceware

1.4.3 Технологія Deepfakes

Deepfakes (поєднання «глибокого навчання» та «підробки») — це синтетичні засоби масової інформації, у яких людину на наявному зображенні чи відеозаписі замінюють на іншу подобу. Хоча акт створення підробленого контенту не є новим, глибокі фейки використовують потужні методи машинного навчання та штучного інтелекту для маніпулювання або створення візуального та аудіоконтенту, яким легше ввести в оману.

Глибинне навчання – це метод навчання комп’ютерів і програмного забезпечення, натхненний тим, як навчається органічний мозок. Він передбачає, що системи знову і знову виконують певні завдання, іноді абсолютно без нагляду з боку людей, щоб дізнатися, як найкраще перетворити певні вхідні дані на бажаний результат. У deepfakes це означає зміну обличчя однієї людини на обличчя іншої, у спосіб, який редактор-людина може не придумати або не зможе виявити.

Спочатку технологію Deepfakes використовували в злочинних цілях: фейкових новинах, містифікаціях, залякуваннях та фінансовому шахрайстві. Це викликало реакцію промисловості та уряду щодо виявлення та обмеження їх використання.

Як приклад, на початку війни проти російських загарбників, усі новинні канали сигналізували про можливу появу відео за участі Володимира Зеленського, де він має капітулювати [44]. Це відео було зроблено за допомогою технології Deepfakes (рис. 1.19).

Придивившись, можна побачити, що якість підставного обличчя відрізняється від тіла. Не витримані пропорції, не виконана корекція кольорів. Тобто, недостатньо просто «начіпити» маску на обличчя та примусити її повторяти заданий текст. Необхідно також працювати з комплексною картинкою.

Рисунок 1.19 – Порівняння оригінального фото Президента України Володимира Зеленського та діпфейку, створеного РФ

Причина, за якою зміна обличчя є основною категорією дипфейків, полягає в тому, що всі обличчя мають спільні риси, навіть у людей дуже різної зовнішності, і доки ці елементи є узгодженими, наш мозок схильний вірити тому, що вони бачать. Наприклад, якщо накласти обличчя акторів одне на одне, програмне забезпечення дізналося б, які частини оригінальних обличь були специфічними для кожного актора, їх слід змінити (колір очей, форма рота), а які були частиною сцени і повинні зберігатися (щоб підняти брови і відкрити рот). Після обміну у вас повинно залишитися обличчя з характерними рисами одного актора, але з оригінальними манерами іншого (рис.1.20) [45].

Тут вже видно, що змінилась форма брів, зіниць, форма обличчя стала менш округлою. Нижня губа майже не змінилась, носо-губна западина стала більш видовженою, що притаманно чоловічому обличчю. В цілому, можна зробити висновок, що експеримент вдався.

Рисунок 1.20 – Емі Адамс у фільмі «Людина зі сталі» та дипфейк, створений YouTube-каналом derpfakes, який накладає обличчя актора Ніколаса Кейджа на обличчя Адамс

Ніщо не перешкоджає створенню механізму deepfakes, який створюватиме враження, ніби люди тримають у руках зброю, хоча вони його не мають, або змушуватиме їх виглядати вищими, ніж вони є, але це не впливає на сильні сторони глибокого навчання та, ймовірно, буде легше досягти за допомогою інших засобів.

Слід також зазначити, що, хоча технологія дуже вражаюча, ідея про те, що дипфейк можна буквально не відрізнити від реального відеоматеріалу, все ще є лише теоретичною. Приклади, подібні до наведеного вище, явно фальшиві та невідповідні, тоді як спроба фактично обдурити за допомогою глибоких фейків потребує неймовірно вдосконаленого програмного забезпечення, ретельних даних і способу імітувати голос суб'єкта, а також його обличчя.

Далі в кваліфікаційній роботі буде розглянуто симбіоз технологій motion capture та deepfakes для створення анімованих масок, що будуть застосовуватись в відео-конференціях.

У даному випадку, motion capture буде використано для надання фейковому обличчю емоцій, природної рухливості, пластичності та виразності. Для трекінгу ключових маркерів буде використовуватись програмне забезпечення компанії Apple для ідентифікації обличчя – face ID.

1.4.4 After Effect (Motion Bro)

Також для трекінгу обличчя можна використовувати спеціальний плагін Motion Bro для Adobe After Effects. В цьому плагіні є розширення Face Tools, яке дозволяє автоматично призначати маркери на обличчі і керувати ними при редакції відео-файлів.

При створенні трекпоінтів у Face Tools, використовується наступний принцип: декілька точок для визначення форми обличчя, чотири точки для зони вуст, три точки для позначення початку, вершини та кінчика брів, в зоні очей також використовують 3 контрольні точки – внутрішній кут, зовнішній та положення зіниці (рис. 1.21 – 1.22).

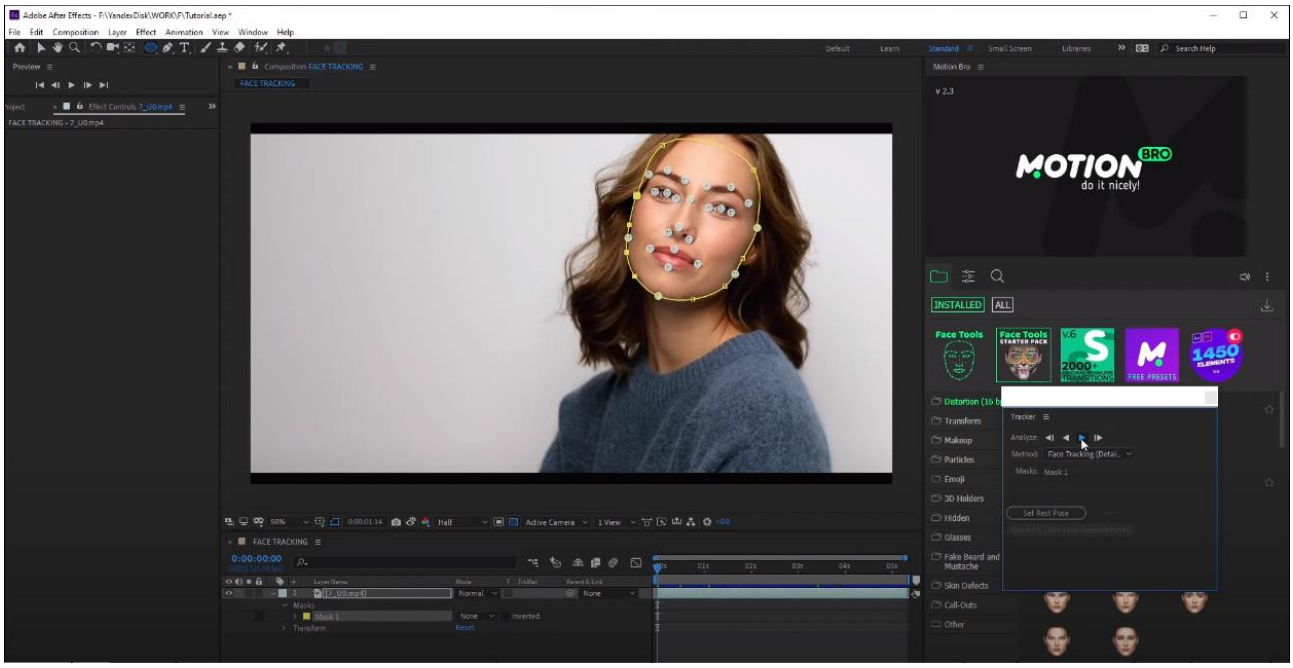

Рисунок 1.21 – Автоматичне налаштування маркерів обличчя у розширенні Face Tools плагіну Motion Bro

Рисунок 1.22 – Розподіл трекпоінтів обличчя у розширенні Face Tools

Якщо брати технологію з безпосереднім нанесенням трекпоінтів на обличчя, увага приділяється зоні чола, вилиць, щік, підборіддя та міжбрів'я. В такому випадку, чим більше, тип краще. Якщо деякі маркери виявились зайвими, їх просто не чіпають під час пост-обробки.

2 СТВОРЕННЯ ВІДЕО ТА ТЕСТУВАННЯ ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ

2.1 Етапи трек-обробки обличчя в After Effects

Якщо в нас є готове відео (рис.2.1), в якому необхідно замінити обличчя (накласти інше), необхідно виконати наступні етапи:

1. Прибираємо дефекти та нерівності шкіри. Обрати оптимальний колірний тон обличчя, прибрати усі недоліки (висипи, нерівність шкіри, широкі пори тощо). Цей етап реалізується за допомогою інструменту «Cosmo II».

За допомогою двох регуляторів «Skin Offset» і «Skin Tolerance» намагаємося виділити якнайбільшу площу обличчя. Обираємо на обличчі точку «середнього тону» - тобто не засвічену і не затемнену ділянку. Тон шкіри по всьому обличчю вирівнюється.

2. Згладжування шкіри. Щоб прибрати всі візуальне нерівності на обличчі, за допомогою інструмента «Skin Smooth» Робимо вирівнювання.

3. Повертаємо реалістичність. В процесі згладжування шкіра може втратити свою реалістичність, виглядати «замиленою» та штучною. Тому, використавши інструмент «Restore detail», відновлюємо втрачені деталі: шорсткість шкіри, контури, веснянки тощо.

4. Додання штучного шуму (рис.2.2). Якщо результат картинки після відновлення деталей не задовольняє, можна використати інструмент штучного шуму. Він створює на шкірі ефект матової фотографії. Тобто з’являється дрібнодисперсна зернистість картинки, що імітує пори на шкірі і виглядає досить природньо.

5. Кольорова корекція (рис 2.3). Якщо з’являється якийсь побічний відтінок (зелений, блакитний, жовтий та ін.), робимо корекцію кольору.

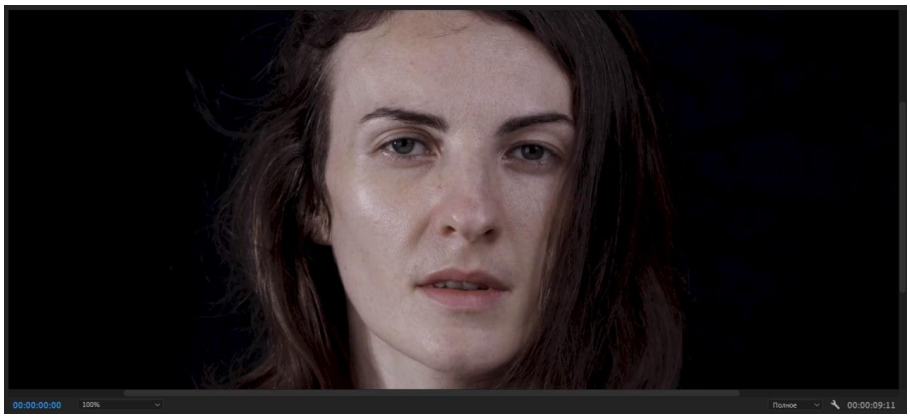

Рисунок 2.1 – Вихідне зображення, яке потребує корекції перед накладанням іншого обличчя

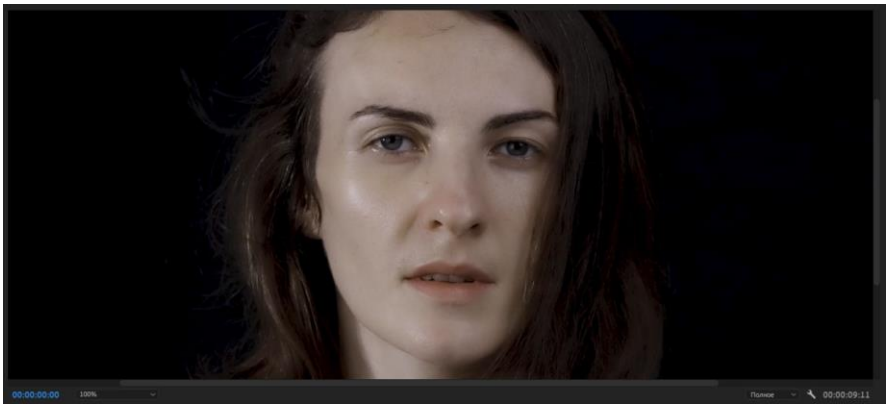

Рисунок 2.2 – Вихідне зображення після відновлення деталей і накладання штучного шуму

Рисунок 2.3 – Зображення після кольорової корекції, готове до накладання іншого обличчя

Після завершення підготовки обличчя, можна переходити до безпосередньої заміни обличчя або накладання ефектів, передбачених плагіном Motion Bro.

Для початку, необхідно створити маску (рис. 2.4) нашого обличчя і зробити її трекінг. Програма сама призначає контрольні точки для трекінгу в межах зазначеної маски (рис. 2.5). Задаємо точку спокою і можемо використати один із вбудованих ефектів деформації обличчя (рис. 2.6).

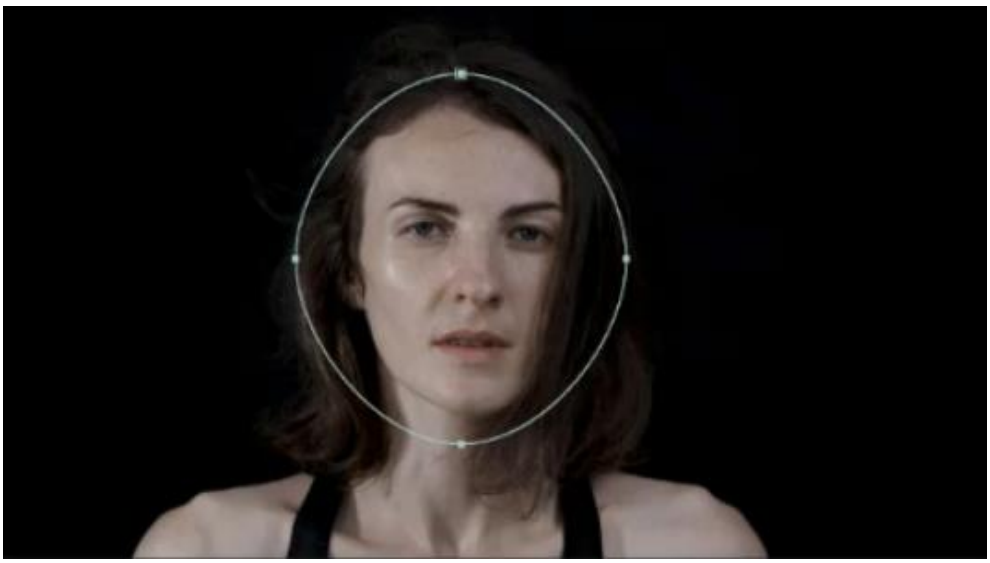

Рисунок 2.4 – Створення маски для призначення контрольних точок

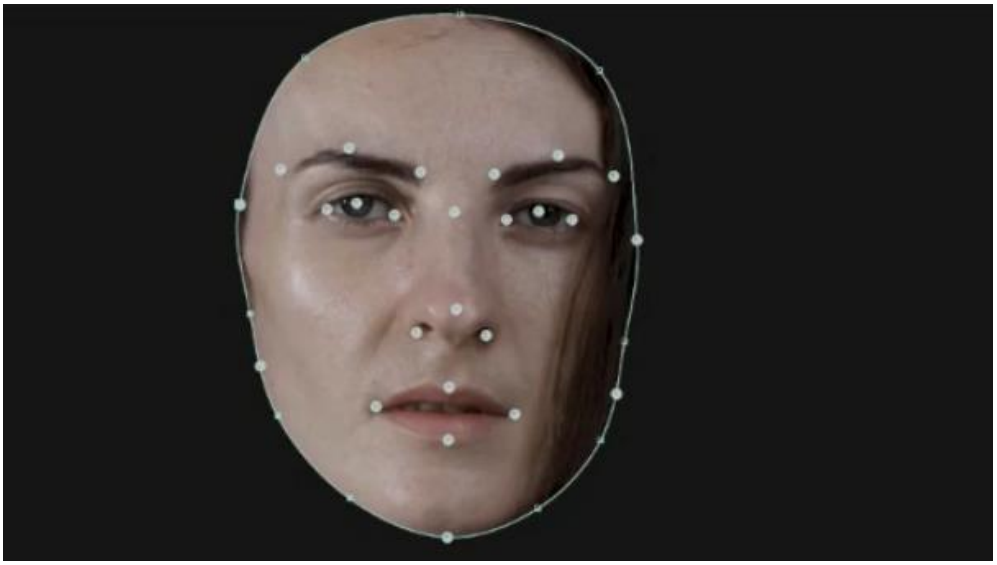

Рисунок 2.5 – Контрольні точки для трекінгу, автоматично призначені програмою

Рисунок 2.6 – Деформація обличчя за допомогою стандартного ефекту плагіну Motion Bro

За бажанням, можна зайти в налаштування ефектів і погратися з параметрами. Такий спосіб накладання, дозволяє змінювати картинку в часі і повторювати кожен рух вихідного зображення на відео. Завдяки контрольним точкам для трекінгу, накладена маска-ефект будуть повторювати міміку використаного оригіналу.

Якщо ми хочемо замінити вихідне зображення на обличчя іншої людини, процес ускладнюється.

У вкладку «Replacement face» поміщуємо обличчя, на яке хочемо замінити і піддягнути його до синіх контурів, що розміщені на екрані (рис.2.7).

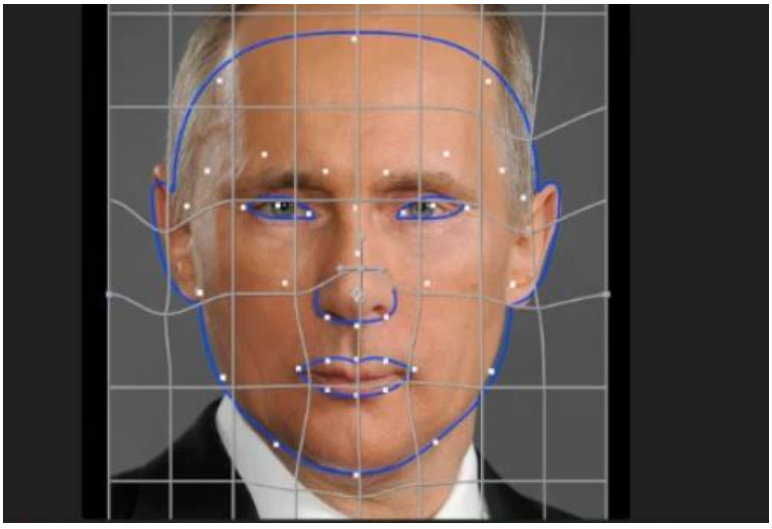

Рисунок 2.7 – Підготовка зображення для заміни обличчя у вкладці «Replacement face»

Після цього, переходимо до етапу ріґґінгу – призначення точок подальшої деформації. Загалом, програма пропонує нам схему із 32-ма точками прив’язки (рис. 2.8), які ми маємо створити на своїй моделі.

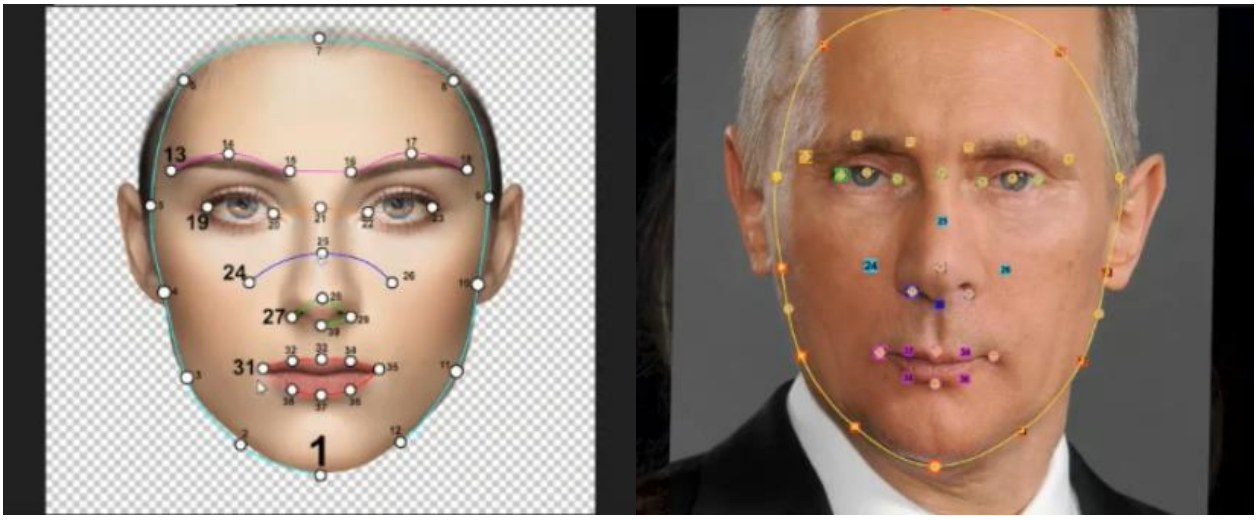

Рисунок 2.8 – Схема розташування точок подальшої деформації та перенесення їх на зображення для накладання

Після нанесення точок деформації, необхідно їх «активувати». У вкладці інструменту «Булавка» обираємо пункт «Маріонетка» і послідовно з’єднуємо між собою усі 32 точки (рис. 2.9). Важливо робити це саме послідовно: від 1 до 32, щоб правильно відбувся процес трекінгу.

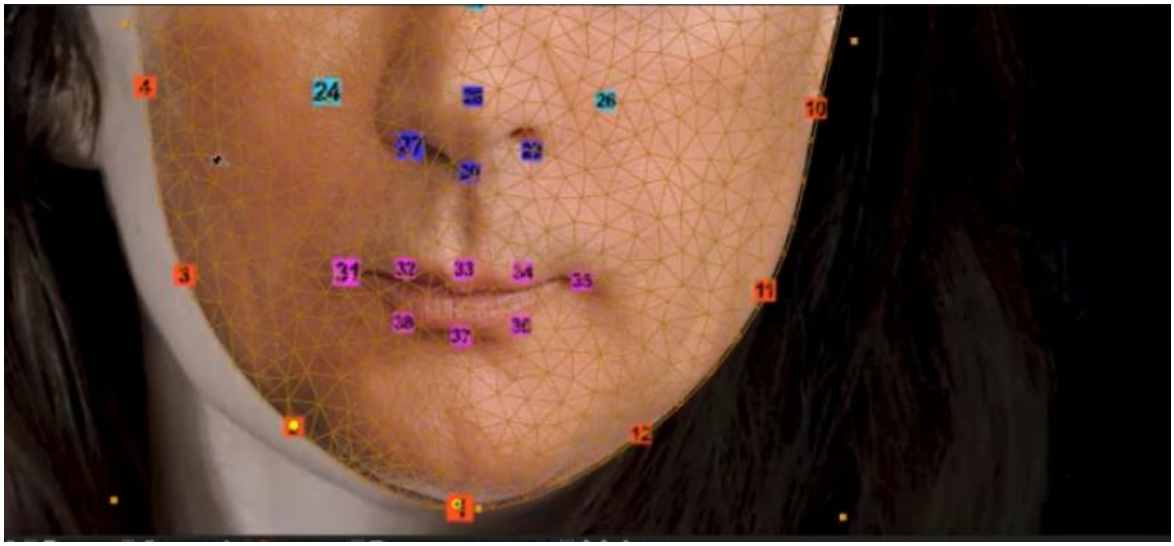

Рисунок 2.9 – Послідовна «активація» точок подальшої деформації обличчя

Після цього, виконуємо процес ріґтінгу і насолоджуємося отриманим результатом.

За бажанням, можна вирізати фрагменти обличчя (губи і очі), проінвертувати створену маску і тоді на відео в накладеному обличчі будуть відображатись фрагменти вихідного зображення.

2.2 Техніка заміни обличчя під час онлайн-трансляції

Коли мова стосується онлайн-трансляцій, процес заміни ускладнюється. На даний момент існує додаток «FaceRig», який дозволяє перетворитися на різних тварин, монстрів або людей. Для того, аби використовувати цей додаток, має бути в наявності камера та мікрофон. Через онлайн-сервіс цифрового розповсюдження комп'ютерних ігор та програм «Steam» завантажується додаток «FaceRig». Далі за допомогою спеціальних опцій налаштовується під розмір обличчя, додаток робить скан і потім з’являється можливість керувати різними монстрами. Додаток буде зчитувати рухи обличчя і передавати його на екран, але вже в іншому вигляді. Також важливо підключити мікрофон, щоб синхронізувати аудіо-доріжку і міміку обличчя персонажу.

«FaceRig» має наступні системні вимоги для встановлення і коректної роботи:

- ОС: Windows 10
- процесор: Intel Core i5-6600 3.3 GHz or equivalent
- оперативна пам'ять: 1 GB ОЗУ
- відеокарта: NVIDIA GeForce GTX980 or equivalent
- мережа: Широкосмугове підключення до інтернету
- місце на диску: 1 GB

Додатково уточнюється, що використання додатку можливо лише на стаціонарних ПК та з широкоформатними веб-камерами. Мобільни пристрї, такі я к телефони, планшет, нетбуки на ноутбуки не підходять [46].

Також, «FaceRig» надає можливість змінювати фон навколишнього середовища. На початку розробки програми, необхідно було проводити трансляцію на тлі хромокею. Наразі, ця технологія оптимізована і дозволяє виділяти обличчя людини на будь-якому тлі.

Таким чином, використовуючи заміну обличчя на заздалегідь заготованому відео, ми маємо можливість зробити кольорову корекцію зображення, замінити окремі фрагменти обличчя, додати ефекти, накласти макіяж тощо. Під час онлайн-трансляції, можна замінити ведучого на будь-якого персонажа з бібліотеки (рис. 2.10 - 2.11), або створити його самостійно. Також, можна підібрати відповідний фон.

Головним нюансом роботи додатку «FaceRig» є те, що велика увага приділяється відстежуванню руху очей. Це головні контрольні точки, які керують персонажем. Особливо, якщо ведучий повертає голову вбік. Частина контрольних точок у такому положенні втрачається, але положення хоча б одного ока завжди можна відстежити.

Якщо було дотримано усіх технічних вимог, то навіть при різких рухах немає затримки на обробку зображення.

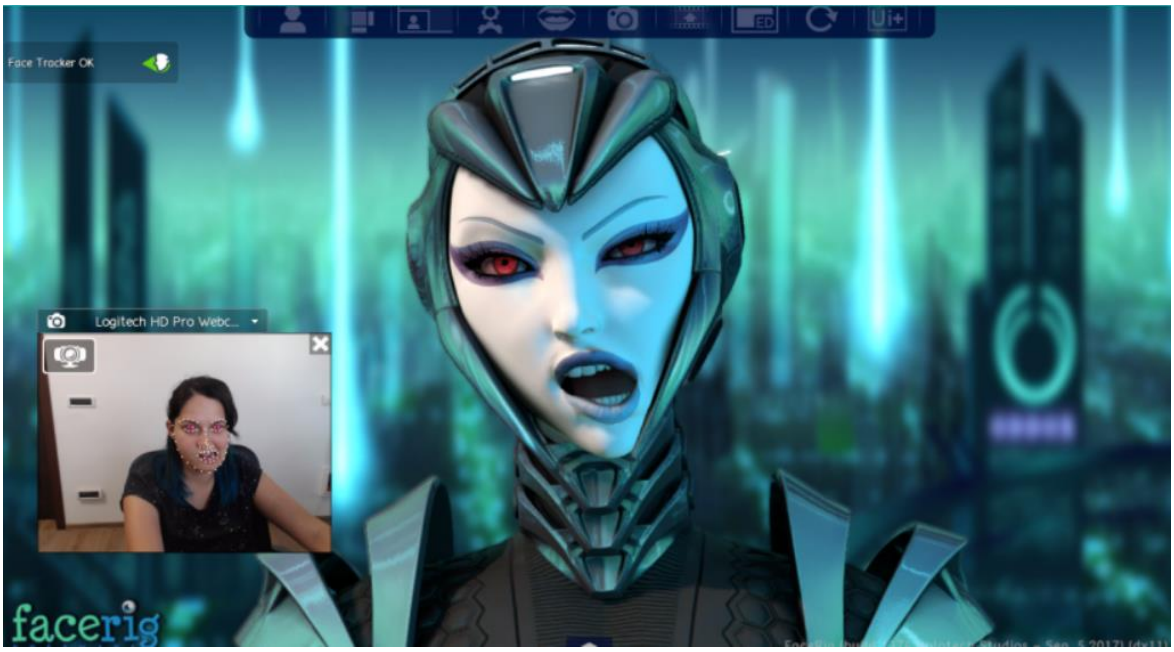

Рисунок 2.10 – Використання додатку «FaceRig» для заміни обличчя під час онлайн-трансляції

Рисунок 2.11 – Використання додатку «FaceRig» для заміни обличчя під час онлайн-трансляції для групи людей одночасно

2.2.1 Використання Snap Camera

Snap Camera – безкоштовна програма для веб-камери Windows. Це якісний цифровий продукт від творців знаменитого Snapchat. Він дозволяє замінити обличчя користувача на безліч позитивних образів (рис. 2.12) – «об’єктивів» – («виростити» що-небудь на голові, перетворитися на моркву або картоплю, яка говорить, змінити стать тощо) практично у будь-якому сучасному засобі відеозв'язку (Zoom, Skype, GoogleMeet та ін.).

Таким чином можна значно підвищити собі та співрозмовнику настрій під час спілкування в мережі.

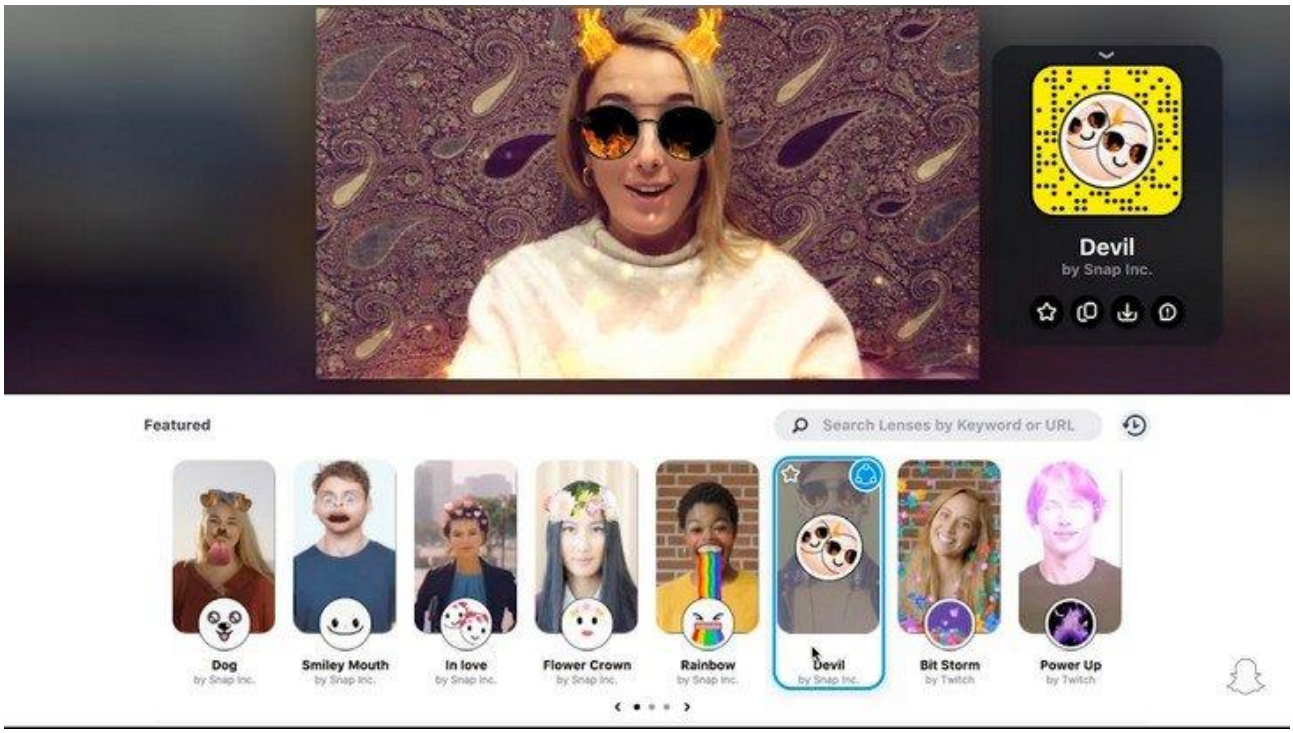

Рисунок 2.12 – Використання об’єктиву «Devil» у програмі Snap Camera

При встановленні Snap Camera створює віртуальну веб-камеру на комп’ютері. Коли Snap Camera відкрита, вона приймає дані фізичної веб-камери, а потім доповнює його вибраним об’єктивом. Потім це доповнене відео виводиться на віртуальну веб-камеру Snap Camera. У програмі відеозв’язку, яка підтримує введення з веб-камери, достатньо просто вибрати Snap Camera зі списку веб-камер, щоб використтати доповнене відео.

Для успішного використання Snap Camera, комп’ютер повинен мати наступні системні вимоги [47]:

— операційна система виключно Windows 10 з розрядністю 64-біт (macOS 10.13);
— процесор не нижче Intel Core i3 2.5 ГГц або AMD FX 4300 2.6 ГГц;
— оперативної пам'яті щонайменше 4 Гб;
— відеосистема (карта) не нижче Intel HD Graphics 4000/Nvidia GeForce 710/AMD Radeon HD 6450 з обов'язковою підтримкою специфікації OpenGL 4.1 і вище;
— роздільна здатність екрана не менше 1280×768.

Якість відео при використанні застосунку Snap Camera, якість відео залежить виключно від параметрів вбудованої або підключеної веб-камери.

В технічній документації додатку вимоги до освітлення не вказані. Провівши експеримент при чотирьох рівнях освітлення (3, 10, 50 та 100 lx), можна зробити наступні висновки: загалом, рівень освітлення не принциповий, поки видно контрольні точки, за якими відбувається трекінг обличчя (рис.2.13 – 2.25). Але, чим вище освітлення, тим краще проявляється ефект об’єктива. Як тільки рівень освітлення падає нижче 3 lx, об’єктиви перестають працювати.

Рисунок 2.13 – Накладання об’єктиву з маскою LiL Yass Grinch (від Paige Piskin) при освітленні 3 lx

Рисунок 2.14 – Накладання об’єктиву з ефектом Christmas Freckles Lens (від Snap Inc.) при освітленні 3 lx

Рисунок 2.15 – Накладання об’єктиву з маскою Snowman Head Lens (від Snap Inc.) при освітленні 3 lx

Рисунок 2.16 – Накладання об’єктиву з маскою LiL Yass Grinch (від Paige Piskin) при освітленні 10 lx

Рисунок 2.17 – Накладання об’єктиву з ефектом Christmas Freckles Lens (від Snap Inc.) при освітленні 10 lx

Рисунок 2.18 – Накладання об’єктиву з маскою Snowman Head Lens (від Snap Inc.) при освітленні 10 lx

Рисунок 2.19 – Накладання об’єктиву з маскою LiL Yass Grinch (від Paige Piskin) при освітленні 50 lx

Рисунок 2.20 – Накладання об’єктиву з ефектом Christmas Freckles Lens (від Snap Inc.) при освітленні 50 lx

Рисунок 2.21 – Накладання об’єктиву з маскою Snowman Head Lens (від Snap Inc.) при освітленні 50 lx

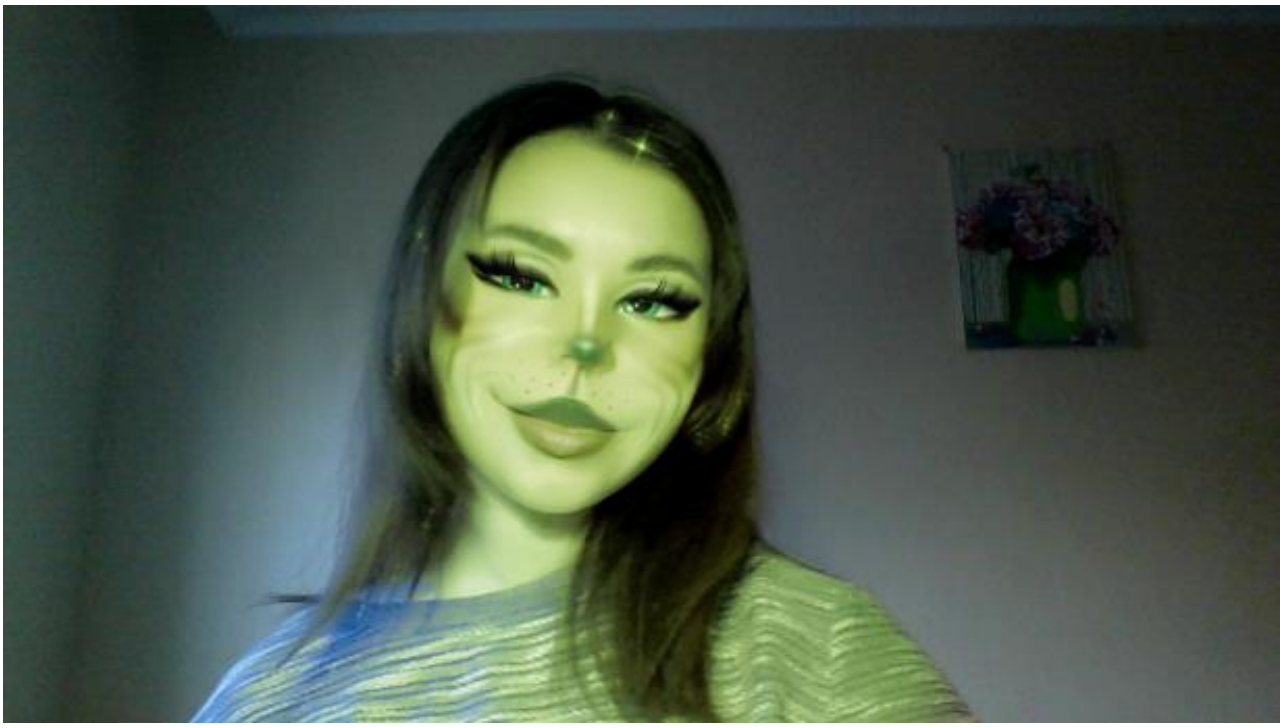

Рисунок 2.22 – Накладання об’єктиву з маскою LiL Yass Grinch (від Paige Piskin) при освітленні 100 lx

Рисунок 2.23 – Накладання об’єктиву з ефектом Christmas Freckles Lens (від Snap Inc.) при освітленні 50 lx

Рисунок 2.24 – Накладання об’єктиву з маскою Snowman Head Lens (від Snap Inc.) при освітленні 100 lx

Контрольні точки для відстежування обличчя в додатку Snap Camera розташовані наступним чином: дві на очах (на зіницях), одна на носі та по одній на губах.

Якщо взяти, наприклад, об’єктив з ефектом Heart Cheeks Lens (рис. 2.25), він використовує три контрольні точки: дві на зіницях і третю на кінчику носа. Щоб впевнитись в цьому, будемо почергово закривати кожну імовірну контрольну точку та дивитись, скільки мінімально має бути точок, для відображення об’єктиву (рис. 2.26).

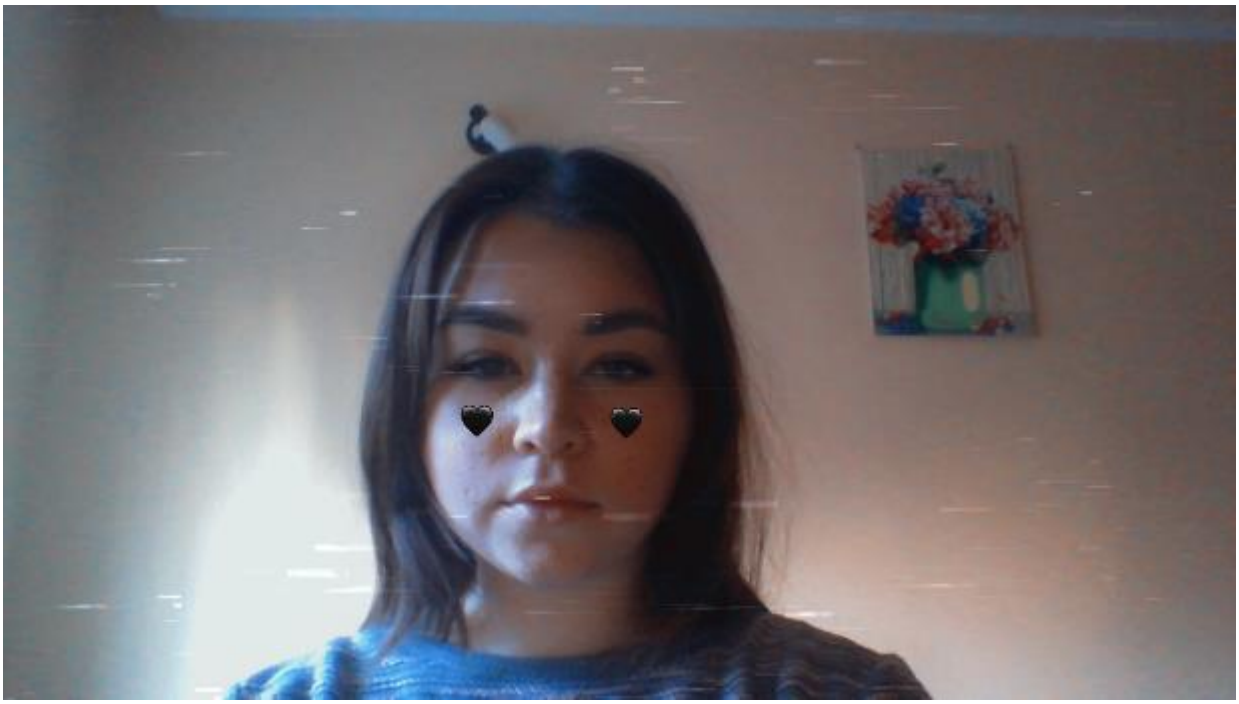

Рисунок 2.25 – Накладання об’єктиву з ефектом Heart Cheeks Lens (від Snap Inc.)

Рисунок 2.26 – Накладання об’єктиву з ефектом Heart Cheeks Lens (від Snap Inc.)

Таким чином, можна стверджувати, що поки програма бачить хоча б одну контрольну точку, об’єктив буде працювати.

Якщо жодної контрольної точки не видно, об’єктив не працює (рис.2.27). Тобто, додаток Snap Camera не використовує контури обличчя для відстежування рухів. Для кожного об’єктиву є свої контрольні точки (мінімум одна).

Рисунок 2.27 – Накладання об’єктиву з ефектом Heart Cheeks Lens (від Snap Inc.)

Затримка на обробку. Якщо системні параметри відповідають вимогам, то затримки на обробку зображення з накладеним об’єктивом немає. Але оскільки параметри даної відеокарти не відповідають, то присутня затримка приблизно 500 мс.

2.2.2 Використання Google Meet

Зі Snap Camera можна працювати і в Google Meet. Проте, цей засіб для відеоконференцій має власні фільтри та ефекти.

Якщо ви хочете приховати те, що знаходиться за спиною, або просто порадувати співрозмовників, ви можете розмити фон або помістити на задній

план картинку, зображення з ефектом присутності або відео (рис.2.28). Також можна додавати стилі: змінювати освітлення, застосовувати колірні фільтри та стилізацію (рис.2.29) [48].

При застосуванні ефекту автоматично вмикається камера. При використанні ефектів і фонів може уповільнити роботу пристрою.

Щоб не заважати передачі відео в реальному часі на пристрої, можна обрати такі ефекти, як розмиття фону і оптимізація яскравості, в хмарі. Це не уповільнює роботу процесора і дозволяє заощаджувати заряд батареї та оптимізувати якість відео-зустрічі.

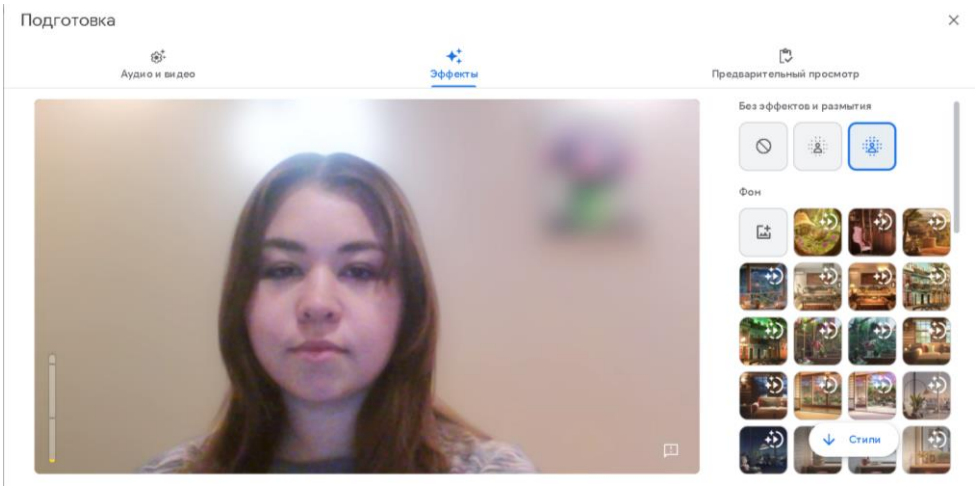

Рисунок 2.28 – Накладання ефекту розмиття на зображення для відео-трансляції в Google Meet

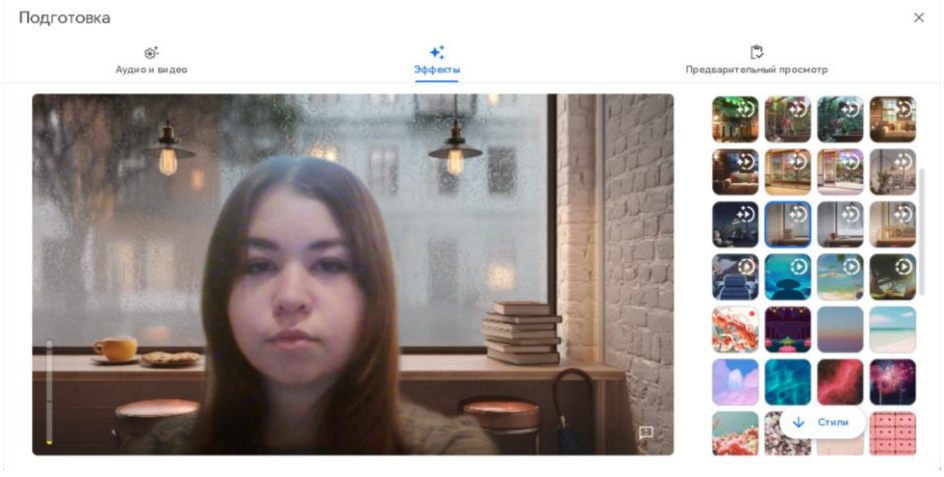

Рисунок 2.29 – Накладання фону на зображення для відео-трансляції в Google Meet

2.2.3 Використання Deepfake

На відміну від попередніх програмних продуктів, технологія Deepfake є значно складнішою. Вона працює завдяки використанню так званої генеративно-змагальної нейромережі (GAN). Одна частина алгоритму вчиться на реальних фотографіях певного об'єкта і створює зображення, "змагаючись" з другою частиною алгоритму, поки та не почне плутати копію з оригіналом.

Для того, аби автоматизувати процес створення матеріалів за допомогою даної технології, було розроблено декілька додатків, наприклад: український стартап Reface, Zao (Китай), DeepFaceLab, Deepfakes web (США).

Розглянемо процес створення зображення Deepfake на прикладі мобільного додатку Reface. На першому етапі нам необхідно зробити фото свого обличчя (яке програма буде накладати на матеріали).

Основні вимоги:

— не посміхатись;

— не заплющувати очі;

— мати гарне освітлення, щоб не було затемнених ділянок на обличчі (затемнені ділянки негативно позначяться під час процесу накладання).

На відміну від інших програм для заміни обличчя , Reface не використовує 3D-модель вашого обличчя. «Цей підхід ми припинили більше року тому. Я говорю тут про 3D-маски, реконструкцію обличчя та змішування текстур. Це рішення може працювати лише з попередньо зібраним і попередньо обробленим вмістом. І може бути нестабільним у виразі обличчя та поворотах», – сказав генеральний директор Reface.

Додаток Reface працює в реальному часі зі статичними форматами (рис.2.30 – 2.32) або GIF. В іншому широка аудиторія залишається непричетною до створення такого контенту та споживає deepfake-відео, яке випускають відомі продакшн-студії або вільні художники на YouTube [49].

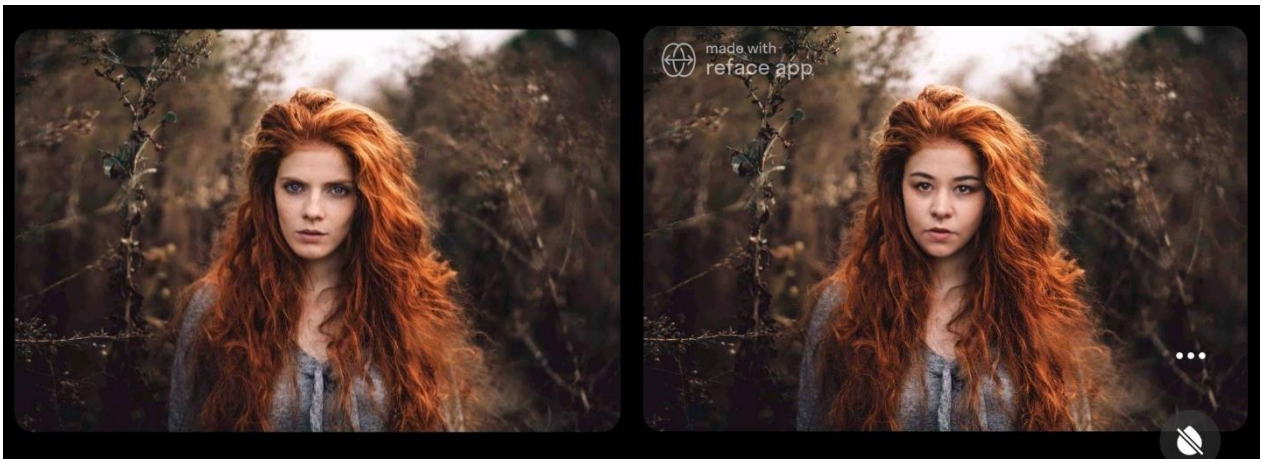

Рисунок 2.30 – Результат накладання мого обличчя на фото в додатку Reface

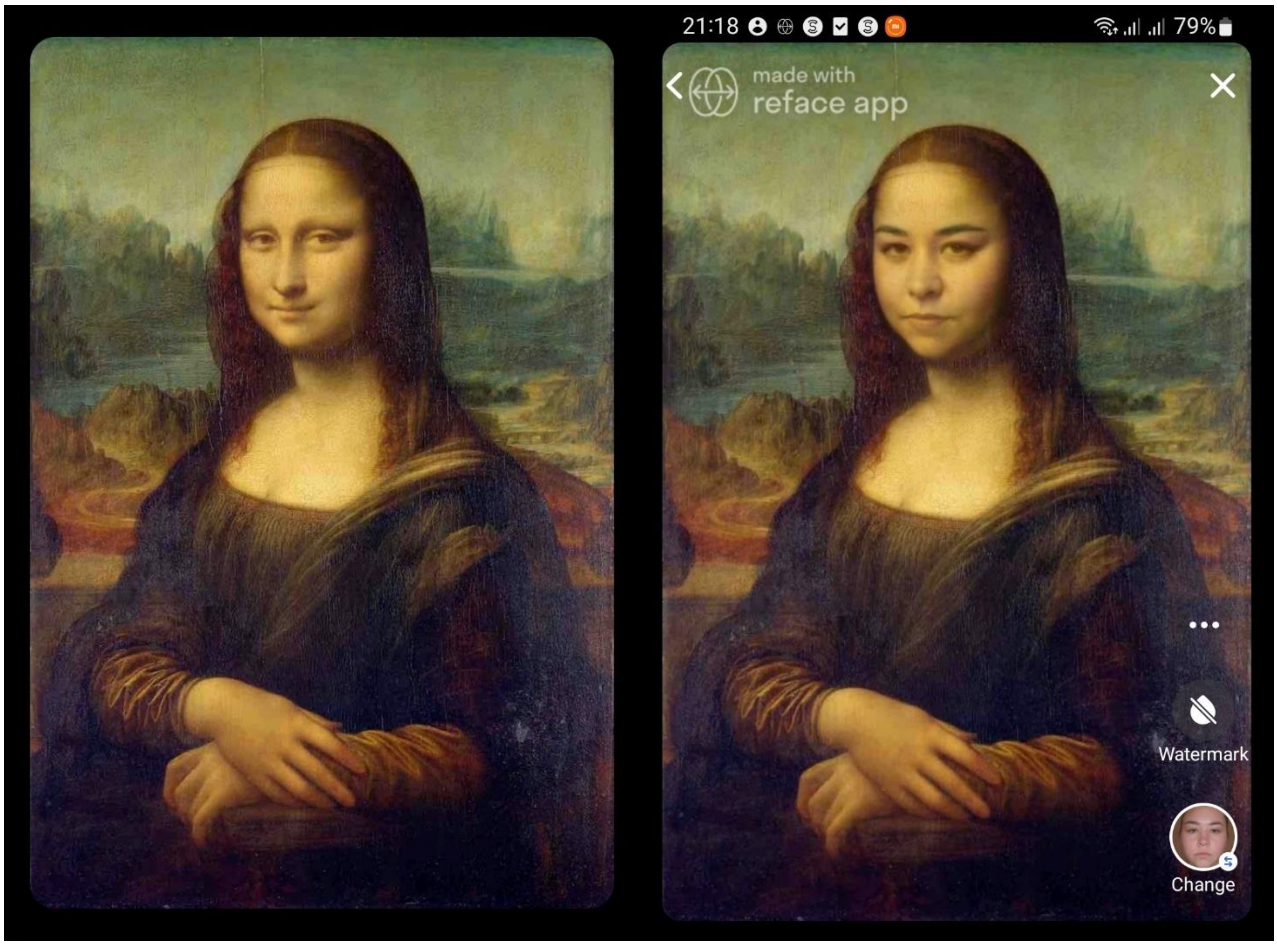

Рисунок 2.31 – Результат накладання мого обличчя на портрет Леонардо да Вінчі – «Джоконда» в додатку Reface

Рисунок 2.32 – Результат накладання мого обличчя на фото в додатку Reface

Затримка на обробку фото становить приблизно 30-40 секунд. Все залежить від ракурсу та деталей на обличчі. При обробці GIF-файлу, затримка може бути 60-90 секунд. Час також залежить від кількості кадрів з накладеним обличчям, ракурсів та тривалості файлу.

Наразі немає можливості використовувати технологію deepfake в онлайн-трансляції. Алгоритм потребує багато часу (до кількох тижнів) на навчання нейронної мережі. Також, для цього необхідно мати великі технічні ресурси (потужні відеокарти, додаткова пам’ять тощо). Тому, сьогодні deepfake доступний для звичайних користувачів як забавка.

3 СТВОРЕННЯ ТА ВИКОРИСТАННЯ ТРИВИМІРНОГО ОБЛИЧЧЯ У ВІДЕО

3.1 Створення надточної та реалістичної моделі людини, інтеграція її в програмне середовище для створення ігор та анімація обличчя

Американська компанія, що займається розробкою комп'ютерних ігор та програмного забезпечення Epic Games, випустила хмарний додаток для створення реалістичних моделей MetaHuman Creator.

У MetaHuman Creator розробники можуть створювати своїх персонажів для подальшого використання в движку Unreal Engine. Є можливість імпортувати моделі з інших програм, наприклад Autodesk Maya або Blender, з наступним редагуванням. Широкий набір інструментів дозволяє повністю налаштувати зовнішній вигляд віртуального персонажа: від зубів та волосся до макіяжу та татуювань. На сьогодні, користувачам доступно понад 50 готових моделей. Також, їх можна створювати самостійно.

3.1.1 Створення моделі Metahumans

Для того, аби використовувати даний програмний продукт, необхідно мати обліковий запис в Epic Games. Тоді на сайті Metahumans Creator можна скористатись усіма можливостями ресурсу. На сайті представлені кілька десятків створених моделей (рис.3.1), які, за бажанням, можна змінити. Наприклад, змінити відтінок шкіри (рис.3.2), зачіску (рис.3.3), статуру людини, можна накласти макіяж, змінити колір очей та зубів. Також, можна змінити середовище (рис.3.4) та позиціонування камери.

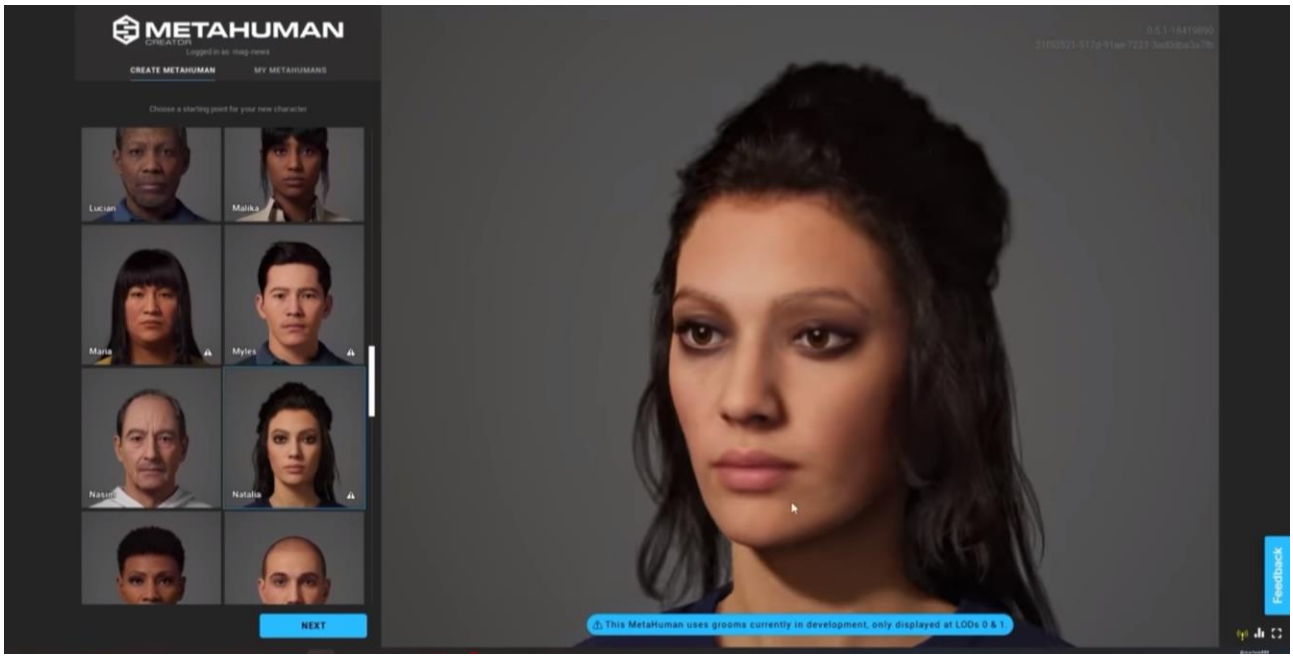

Рисунок 3.1 – Колекція моделей, представлених на сайті Metahumans

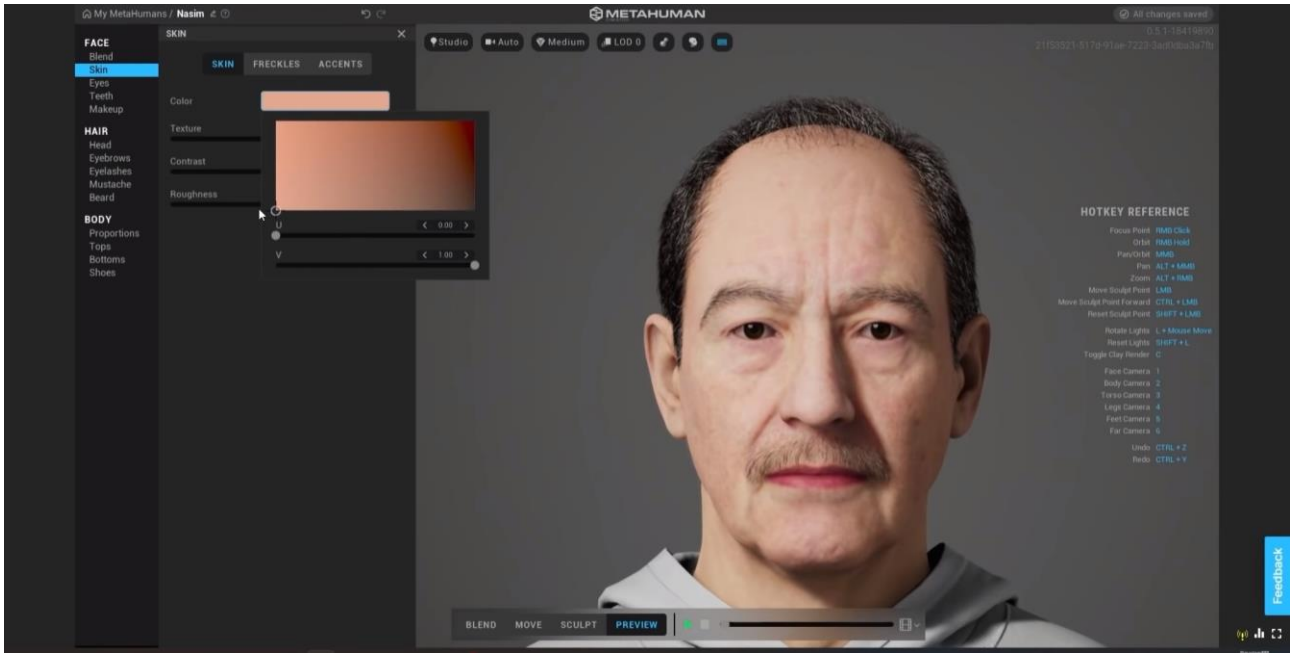

Рисунок 3.2 – Зміна відтінку кольору обраної моделі

Рисунок 3.3 – Зміна зачіски обраної моделі

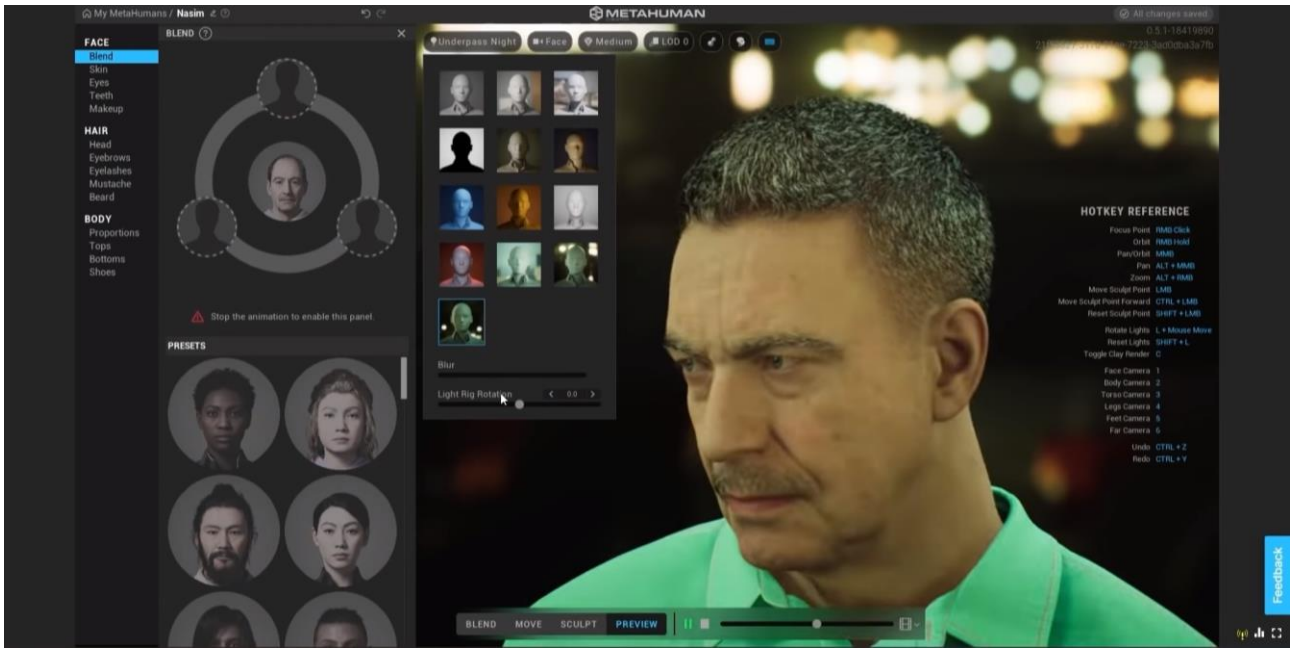

Рисунок 3.4 – Зміна оточення навколо обраної моделі

Усі створені користувачем персонажі зберігаються у базі Metahumans у вкладці «Мої моделі».

3.1.2 Інтеграція створеної моделі Metahumans в програмне середовище Unreal engine

Для того, щоб перенести створеного персонажа в програмне середовище для конструювання ігор Unreal engine, необхідно встановити декілька програмних продуктів.

Встановлюємо Epic Games Launcher. Якщо до цього можна було працювати в браузері, то для подальшої роботи необхідно встановити додаток на комп’ютер.

Завантажуємо з магазину Unreal engine 5 (рис.3.5) оскільки тільки ця версія дозволяє комфортно і якісно працювати з моделями Metahuman та встановлюємо її на комп’ютер. Для якісної роботи в середовищі Unreal engine 5 необхідні значні потужності комп’ютерної техніки.

Мінімальні системні вимоги розраховані на кадрову частоту в районі 15-20 fps і включають в себе [50]:

— відеокарти Nvidia GeForce GTX 1080 або AMD Radeon RX Vega 64;

— 8-ядерний процесор із частотою від 3 ГГц;

— 32 ГБ оперативної пам'яті;

— 100 ГБ вільного місця.

Але для забезпечення 30 кадрів за секунду потрібні:

— 12-ядерний процесор із частотою від 3,4 ГГц;

— GeForce RTX 2080 або Radeon 5700 XT;

— 64 ГБ ОЗУ.

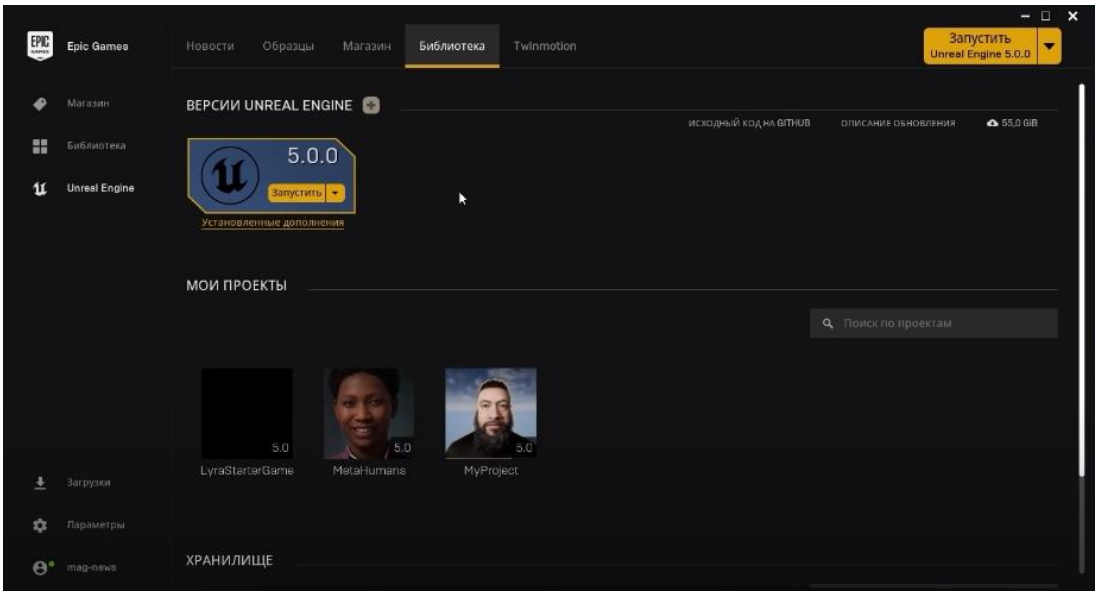

Рисунок 3.5 – Завантаження програмного середовища для конструювання ігор Unreal engine з магазину Epic Games Launcher

Завантажуємо та встановлюємо плагін Quixel Bridge (рис.3.6). Він надає повнофункціональний доступ до бібліотеки Megascans у редакторі рівнів [51]. Це дозволяє переглядати колекції, шукати певні ресурси та додавати їх до своїх проектів Unreal Engine. Завантаження моделі з Metahuman можливо якраз завдяки цьому плагіну.

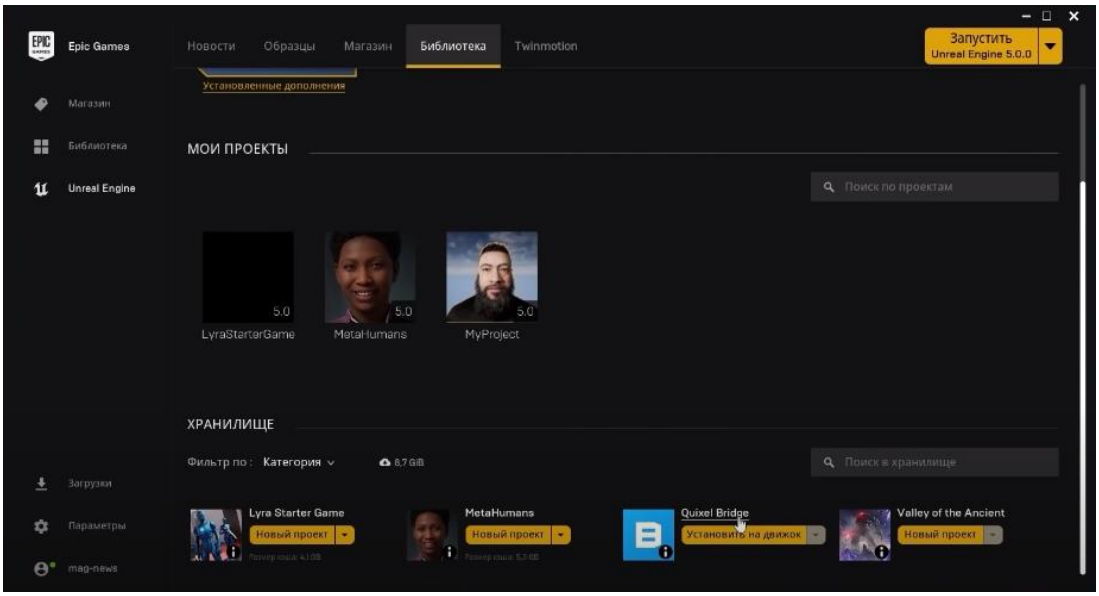

Рисунок 3.6 – Встановлення плагіну Quixel Bridge для інтеграції моделей Metahumans з магазину Epic Games Launcher

Останнім підготовчим етапом завантажуємо усі моделі Metahumans із сайту на наш комп'ютер. Зробити це можна, натиснувши на кнопку «Новий проект» і зберегти його (рис.3.7). Там будуть міститись усі моделі, які ми адаптували під свій проект.

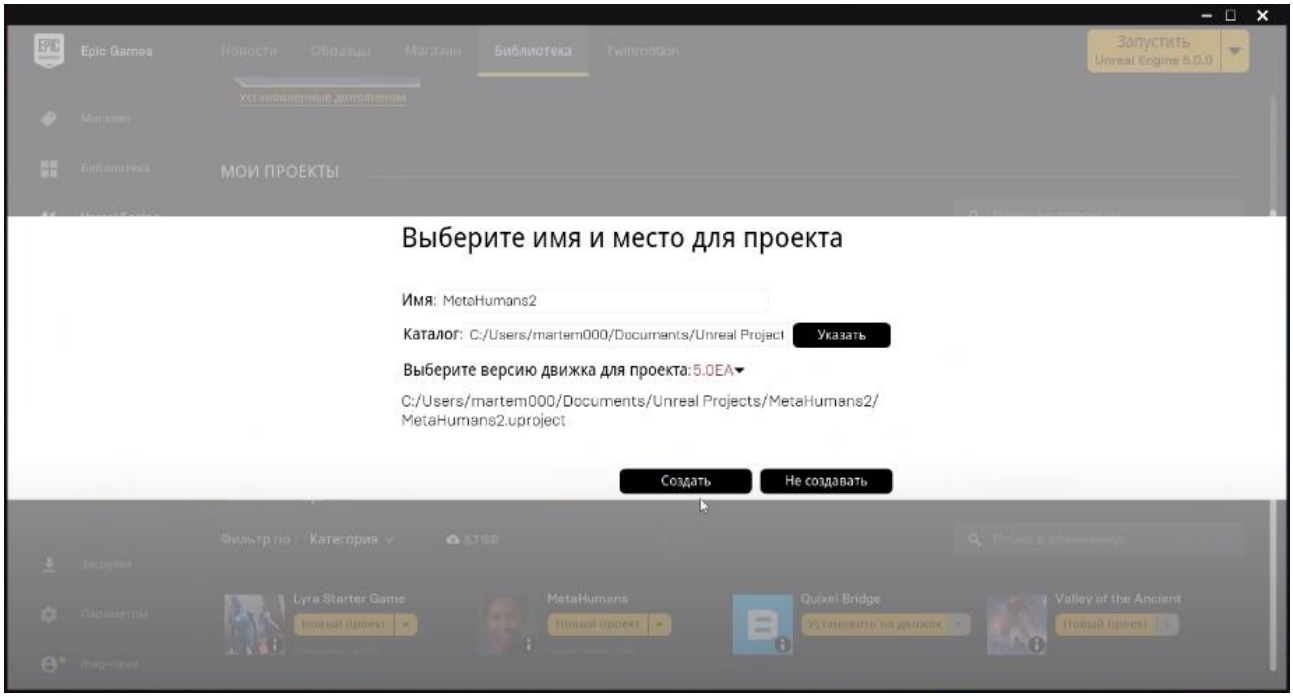

Рисунок 3.7 – Завантаження проекту Metahumans з усіма вбудованими моделями та створеними власноруч користувачем

Відкривши Unreal Engine, бачимо, що Metahumans вже всередині (рис.3.8) і переходимо безпосередньо до інтеграції моделі в нову сцену.

Через вкладку Games створюємо «порожню» сцену. Порожня вона умовно, тому що там за замовчуванням розміщенні об'єкти. Стандартна «порожня» сцена являє собою платформу із меблями (два крісла та стіл між ними) (рис.3.9).

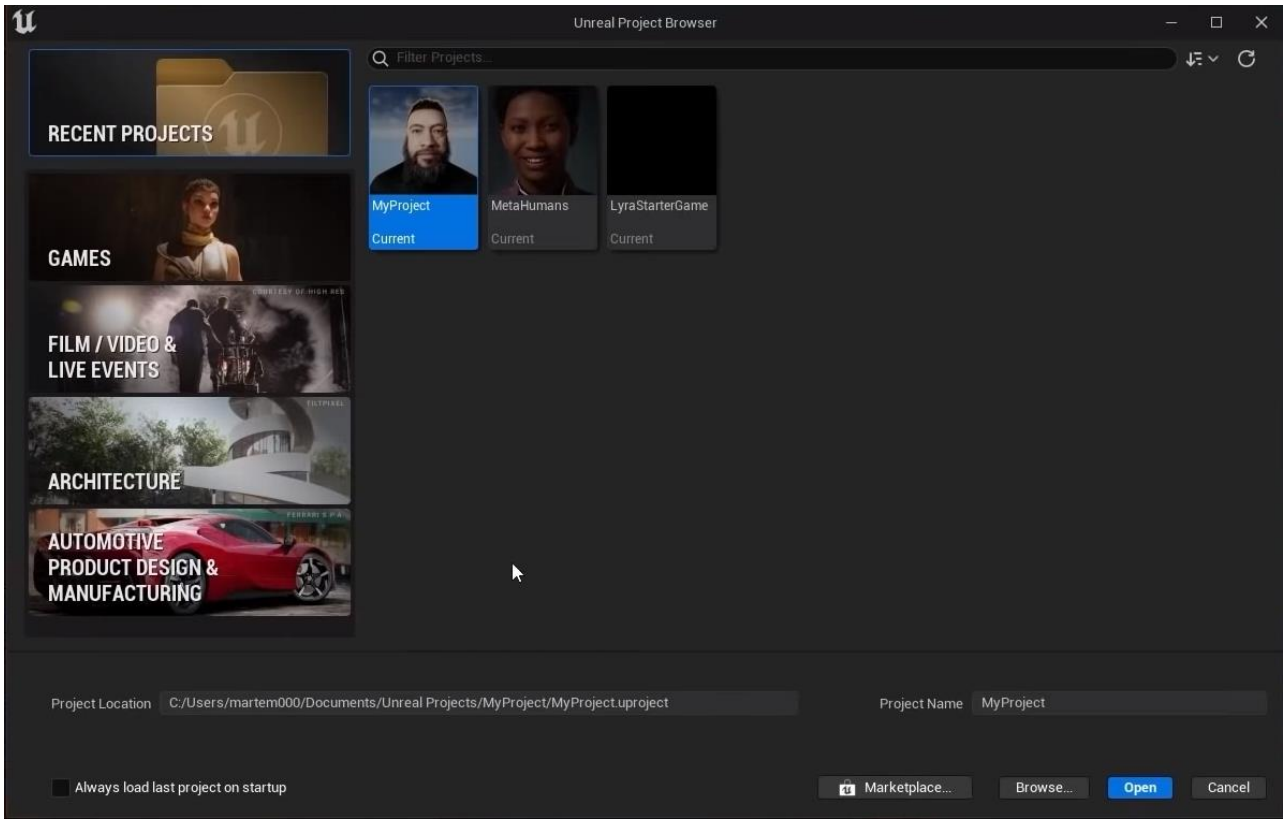

Рисунок 3.8 – Бібліотека Metahumans у середовищі Unreal engine 5

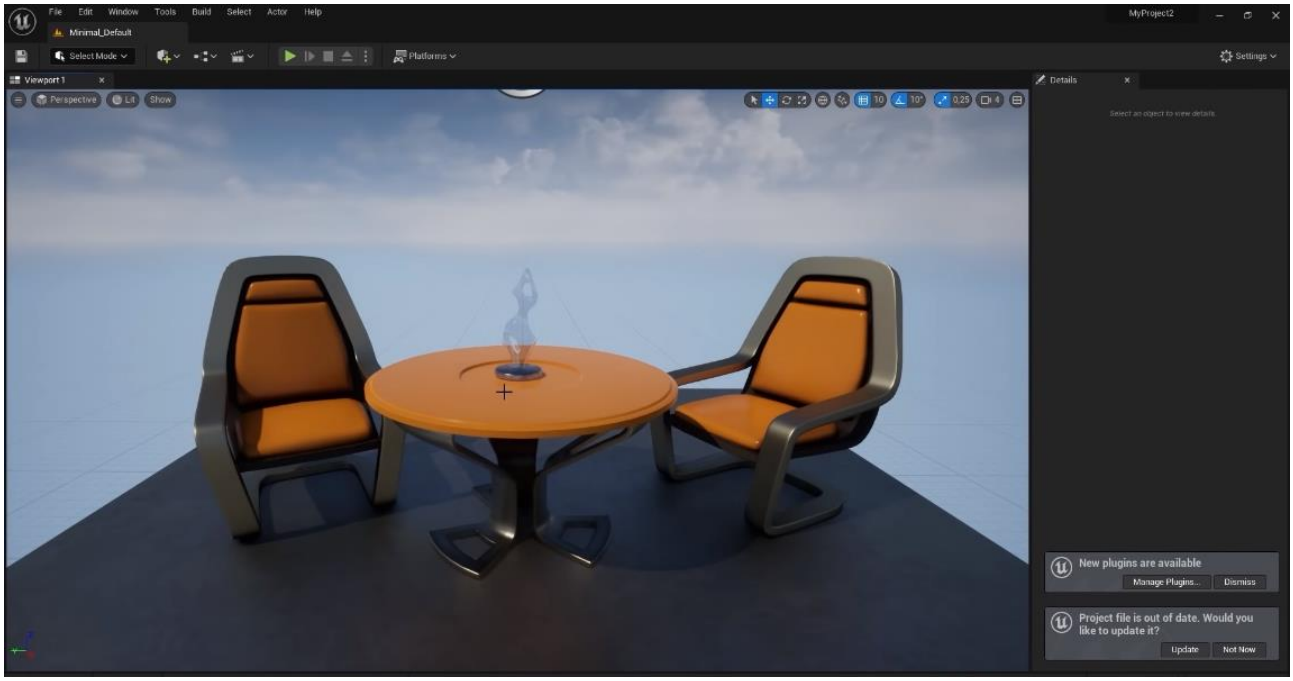

Рисунок 3.9 – Стандартна нова сцена в середовищі Unreal Engine

Можемо змінювати кут повороту камери відносно сцени, додавати в неї об’єкти зі встановлених бібліотек (стандартних або додатково завантажених та

підключених), переміщувати об’єкти в межах сцени, додавати нові камери та джерела світла (і розміщувати їх за особистим бажанням).

Для того, щоб помістити у створену сцену модель, необхідно на верхній робочій панель програми натиснути кнопку «Додати» і обрати плагін Quixel Bridge. Коли відкриється вікно плагіну, необхідно в ньому додатково увійти в свій обліковий запис для імпорту моделі. Щойно виконано вхід в акаунт, з’являється вкладка My Metahumans, де розміщенні створені нами раніше моделі (рис.3.10). Після того, як модель обрана – є можливість змінити якість моделі (низька, середня та висока). Максимальне значення сягає 8К. ЇЇ доцільно використовувати, якщо потужності комп’ютера відповідають системним вимогам. Щоб спростити та прискорити процес інтеграції моделі в програмне середовище Unreal engine, краще обирати низьку якість. Потім завантажуємо обрану модель. Цей процес може зайняти від 20 хвилин, до 2-3 годин. Час завантаження залежить від того, чи була згенерована модель на сервері Metahumans, після того, як ми її створили, чи ще ні. У другому випадку, модель буде згенеровано, тобто додано усі зміни, які вносив автор, проведено рендерінг та ін. операції безпосередньо під час завантаження.

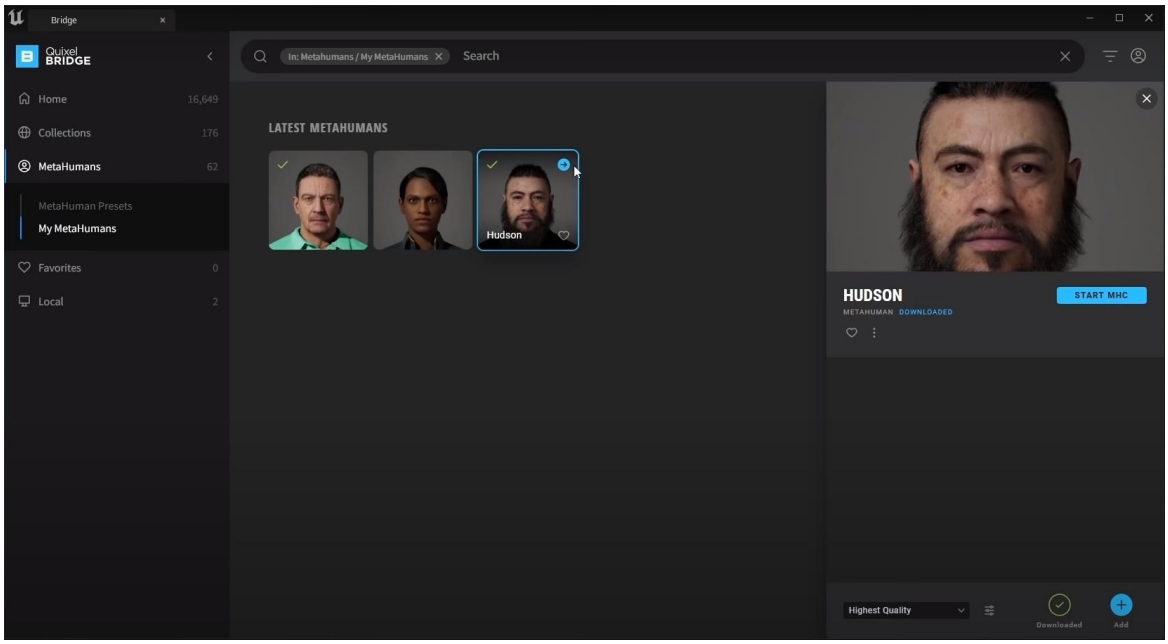

Рисунок 3.10 – Вибір моделі для завантаження та подальшого переміщення у сцену

Натиснувши кнопку Add, імпортуємо обрану модель у сцену. Процес імпорту наразі відображається некоректно (вікно зі статусом завантаження постійно зникає з екрану), проте переміщення моделі відбувається і може сягнути 5-10 хв. Спостерігати за процесом можна через диспетчер завдань, але це не обов’язково.

Після завершення імпорту, модель з’являється в каталозі папок, де містяться усі елементи бібліотеки з моделями (рис.3.11). Тепер просто переміщуємо файл з цією моделлю у сцену (рис.3.12).

Рисунок 3.11 – Каталог файлів із об’єктами бібліотеки Metahumans та щойно завантажена модель

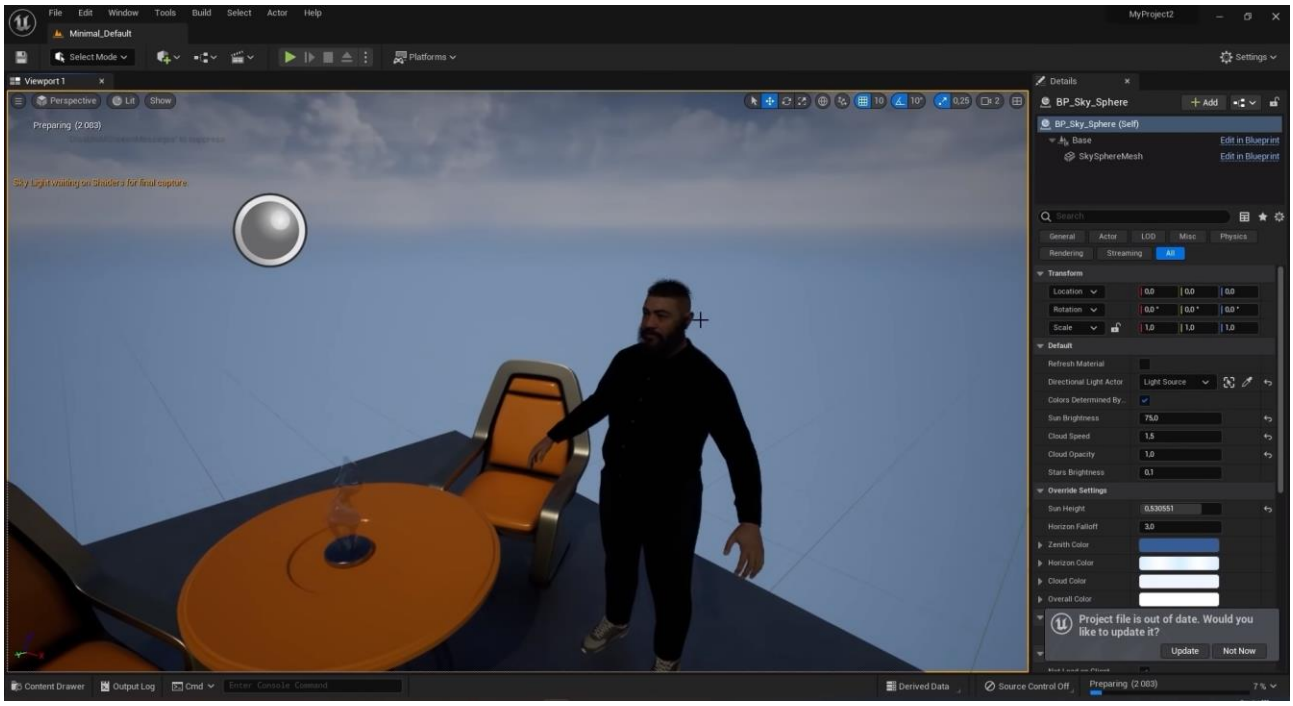

Рисунок 3.12 – Створена особисто модель людини, інтегрована в стандартну сцену середовища Unreal engine

Важливо інколи під час роботи зберігати проект, щоб створювати «точки відліку». Якщо щось піде не так, завжди буде можливість повернутись до попереднього етапу і спробувати знову.

За необхідності, можна додати джерело світла, щоб освітити нашу модель і сцену загалом. В даному прикладі було використано стельову лампу зі світлом білого кольору та яскравістю 20 000 кд/м^2. Ці параметри можна змінити у меню керування, вкладки «Color» та «Brightness» відповідно (рис.3.13).

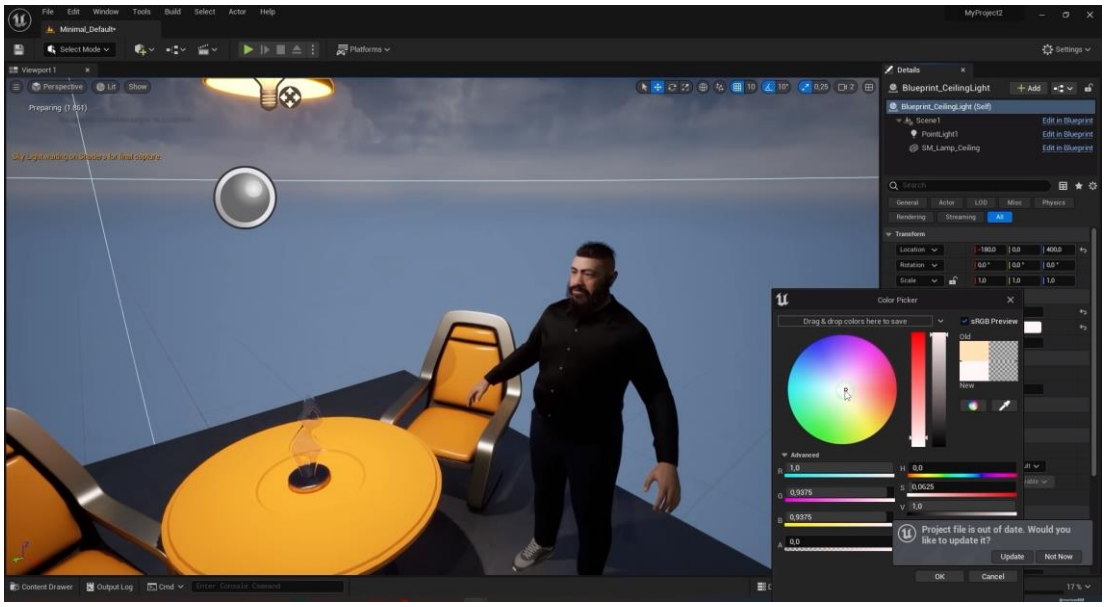

Рисунок 3.13 – Стандартна сцена з інтегрованою моделлю та доданим джерелом білого світла яскравістю 20 000 кд/м^2

Для перегляду обличчя і подальшої анімації, можна змінити кут огляду камери, її поворот, та відстань до об’єкту (рис.3.14). Після завершення інтеграції моделі, можна переходити до анімації обличчя моделі за допомогою спеціального мобільного додатку «Live Link Face» для операційної системи Apple.

Рисунок 3.14 – Обличчя інтегрованої моделі з доданим освітлення, підготовлене для подальшої анімації

3.1.3. Анімація обличчя за допомогою «Live Link Face»

Анімацію обличчя моделі Metahumans безпосередньо в середовищі Unreal Engine 5 можна виконати за допомогою мобільного додатку для iPhone – «Live Link Face».

Додаток дозволяє:

— передавати високоякісні вирази обличчя персонажам і візуалізувати їх за допомогою рендерингу в реальному часі в Unreal Engine;

— записувати дані відстеження обличчя, які можна додатково налаштувати в інструментах анімації для досягнення кінцевої продуктивності та зібрати в Sequencer Unreal Engine;

— знімати професійний рівень продуктивності за допомогою інтегрованого робочого процесу сцени.

Для початку анімації, необхідно підключити камеру з додатку до проекту в Unreal Engine 5. Для цього

1. Встановлюємо додаток на телефон
2. Якщо комп’ютер був підключений до мережі Інтернет, необхідно перепідключити його до Wi-Fi.
3. У параметрах мережі Wi-Fi знаходимо IP-4 адрес, щоб через нього підключити телефон.
4. У налаштуваннях додатку «» в пункті «Targets» задаємо IP-адресу нашої Wi-Fi мережі. Адреса порту додається за замовчуванням – 11111.
5. Тепер можемо повертатись до додатку і анімувати обличчя моделі.
6. У вікні Unreal Engine, в лівому меню у вкладці «Default», пункт «LLink Face Subj» обираємо iPhone.
7. Видаляємо стандартну камеру, яка використовувалась в сцені, натискаємо play і тепер персонаж відтворює усю міміку обличчя користувача, зчитуючи її через вбудовану камеру iPhone (в тому числі, повторює траєкторію руху зіниць).

Результатом такої роботи є – побудована тривимірна сцена, яку можна використовувати в подальшій обробці з реалістичною анімацією обличчя та трансляції на будь-яких зустрічах або стрімінгових платформах.

3.2 Трекінг обличчя в Faceware

Для створення реалістичної анімації тривимірної моделі необхідно задатися вхідними параметрами: розширення відео та метод відстеження (маркерний або безмаркерний).

Оскільки в програмній документації плагіну Faceware достеменно не вказано, який метод відстеження буде працювати якісніше, надалі розглянемо обидва варіанти.

Розширення вхідного відео впливає лише на швидкість трекінгу, оскільки 3Д модель можливо побудувати у будь-якому розширенні (рис.3.15) для подальшої обробки та трансляції.

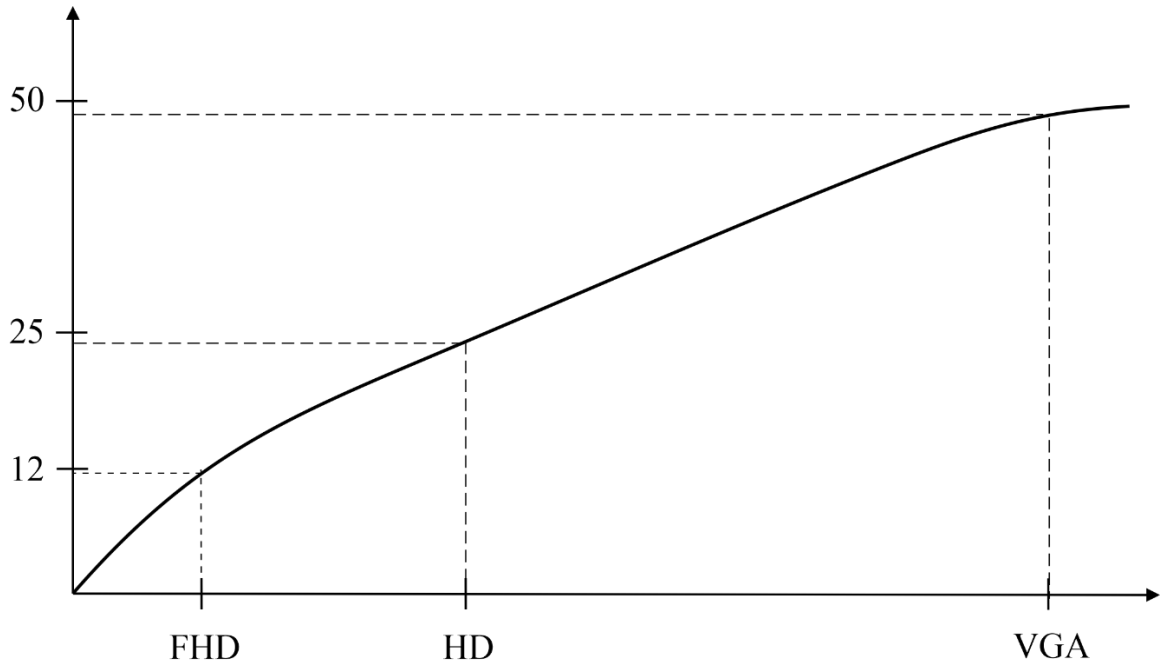

Рисунок 3.15 – Графік залежності формату розширення відео від залежності частоти кадрів відео відтворення

Цей графік залежності справедливий для комп’ютера з системними параметрами:

— процесор i7-6700;

— відеокарта GT-640;

— ОЗУ 16 GB DDR 4.

Для аналізу якості роботи плагіну з безмаркерним методом відстежування, було використано 3 відео з розширенням VGA і тривалістю 1,5 хв. Скріншот обробки відео у програмі Faceware наведено на рис. 3.16.

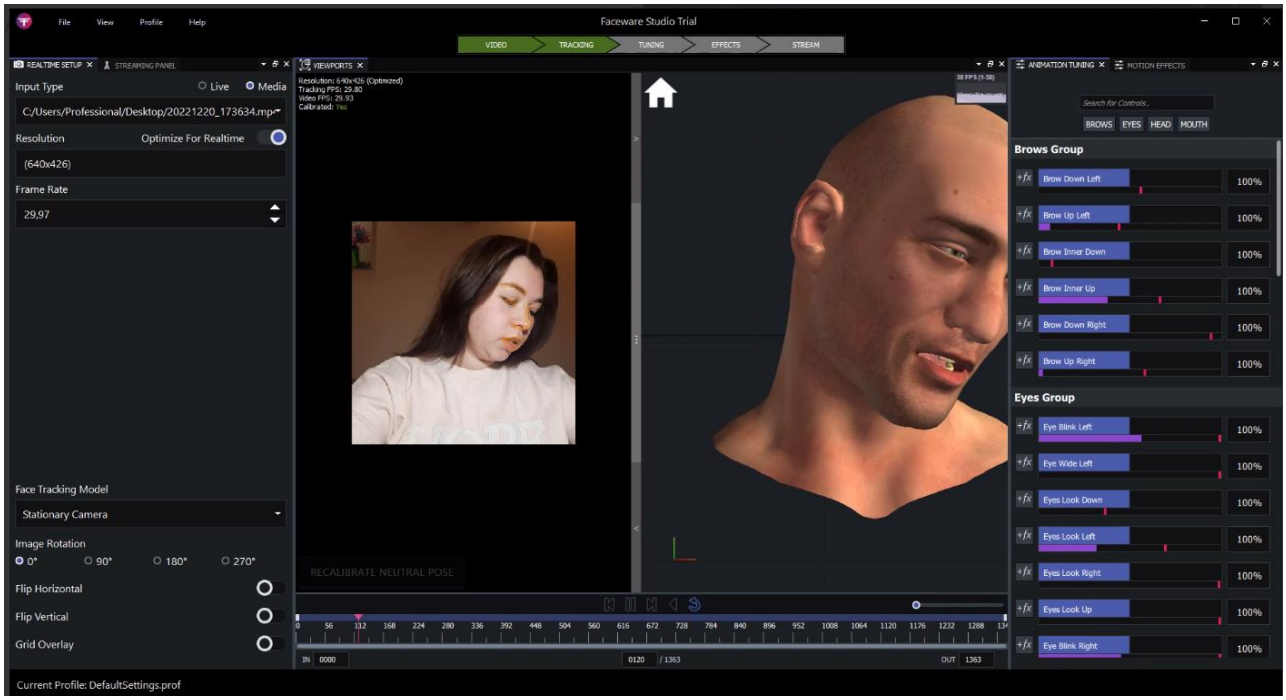

Рисунок 3.16 – Загальне вікно програми трекінгу

Для подальшого аналізу отриманого результату у програмі Faceware будуть використовуватися короткі (кілька секунд) фрагменти.

Обробка отриманих результатів відбувалась в програмному середовищі After Effects, за допомогою вбудованого трекера. Розширення відео складає 1920x1080 pixel. Інтервал аналізу – 4 с та похибка аналізу з використанням вбудованого трекеру, склала 1,25 pixel (рис.3.17).

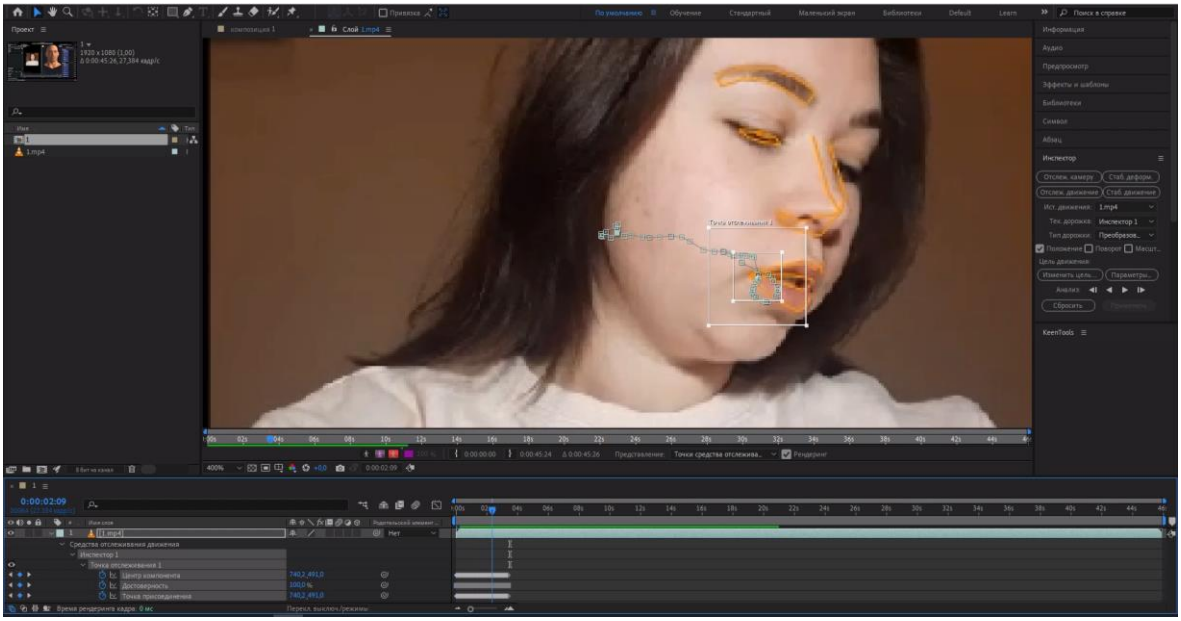

Рисунок 3.17 – Еталонна траєкторія руху контрольної точки

Для відстеження траєкторії руху контрольної точки, використовувались 2 змінні:

1 – внутрішній квадрати, який використовується безпосередньо для відстеження заданої точки.

2 – зовнішній квадрат, вказує траєкторію зміщення (на скільки точка може зміщуватися у наступному кадрі).

Чим більше буде площина другого квадрату, тим більше часу буде витрачено на відстеження точки.

Аналогічні дії повторюємо для аналізу траєкторії руху тривимірної моделі (рис.3.18).

Після отримання траєкторій руху контрольної точки у програмі After Effects можна візуально оцінити відхилення траєкторій руху тривимірної моделі по відношенню до траєкторії руху людини (рис.3.19). Максимальне відхилення між траєкторіями руху складає 40 pixel.

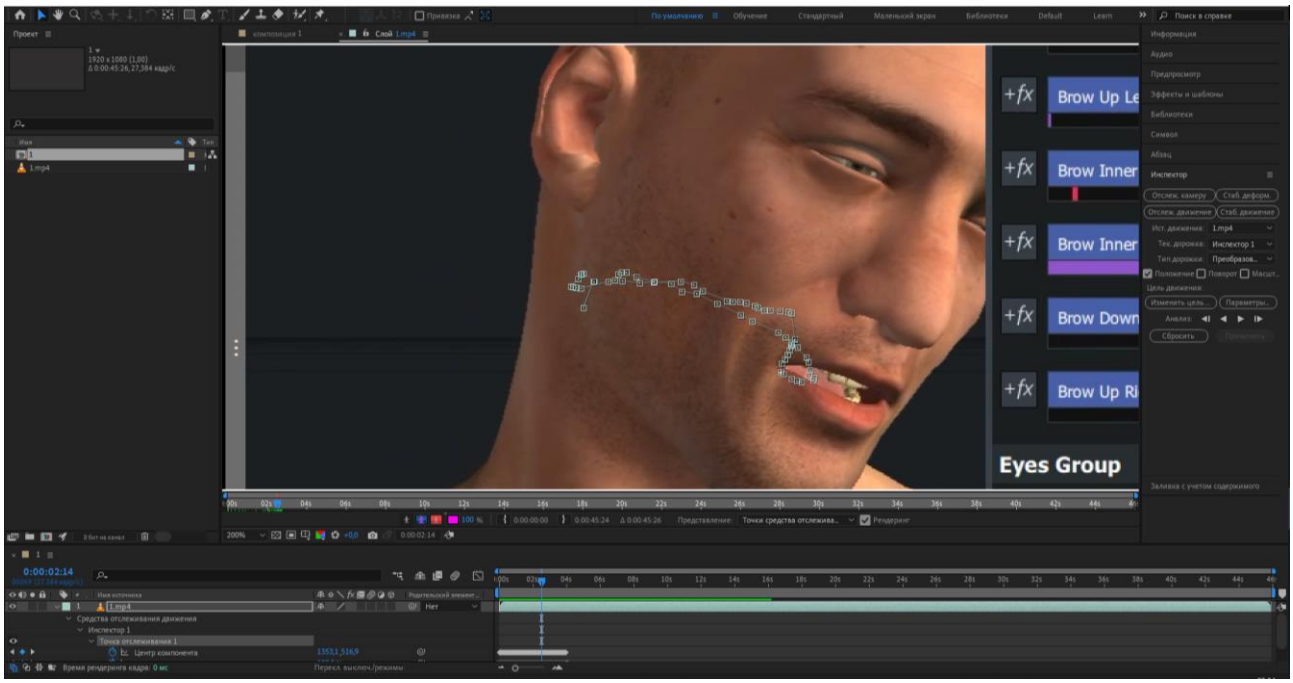

Рисунок 3.18 – Траєкторія руху контрольної точки тривимірної моделі

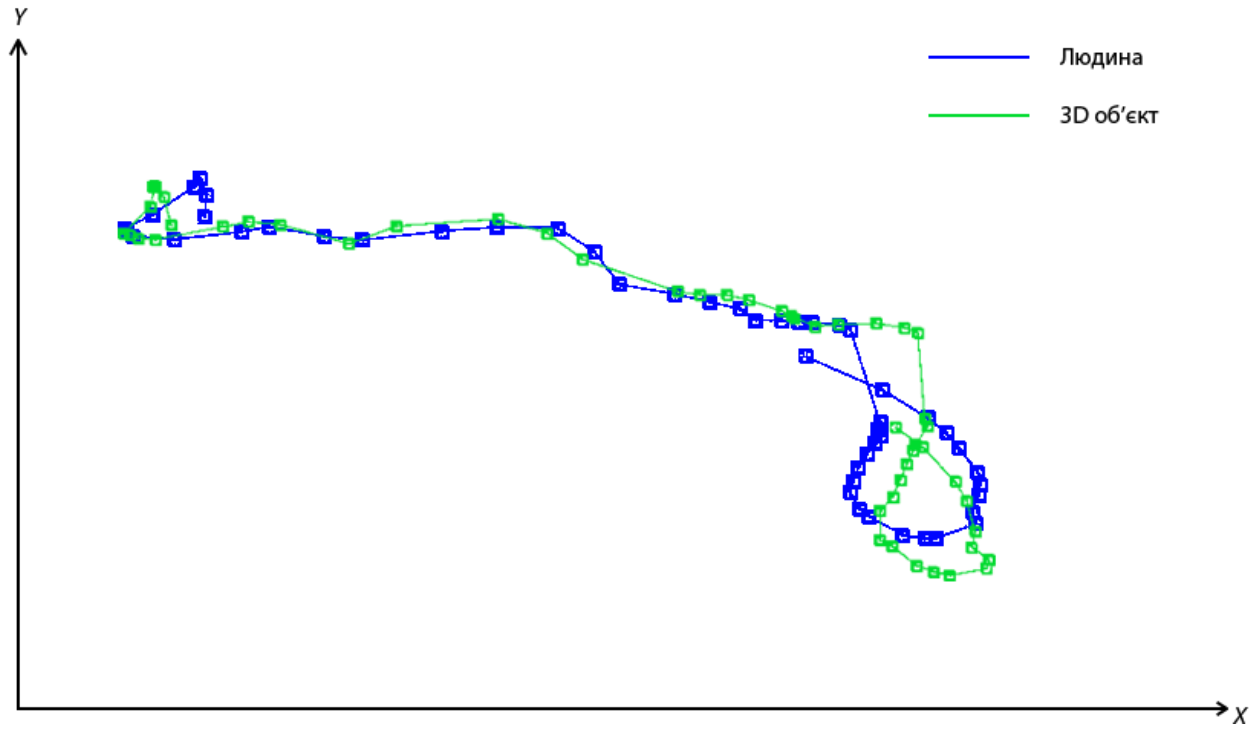

Рисунок 3.19 – Суміщені траєкторії руху контрольної точки (людини та тривимірної моделі)

Отримані траєкторії руху дозволяють оцінити похибку, що складає 40 pixel. Для формату зображення 16:9, розширенням 1920x1080 відхилення руху 40 pixel не призводить до видимих спотворень.

Наступним етапом було розглянуто відео, у якому присутній момент нахилу голови (рис.3.20-3.21), тому що в такому випадку змінюється об’єм та площина об’єкту, що відстежується. А це, в свою чергу, може призвести до помилки під час роботи.

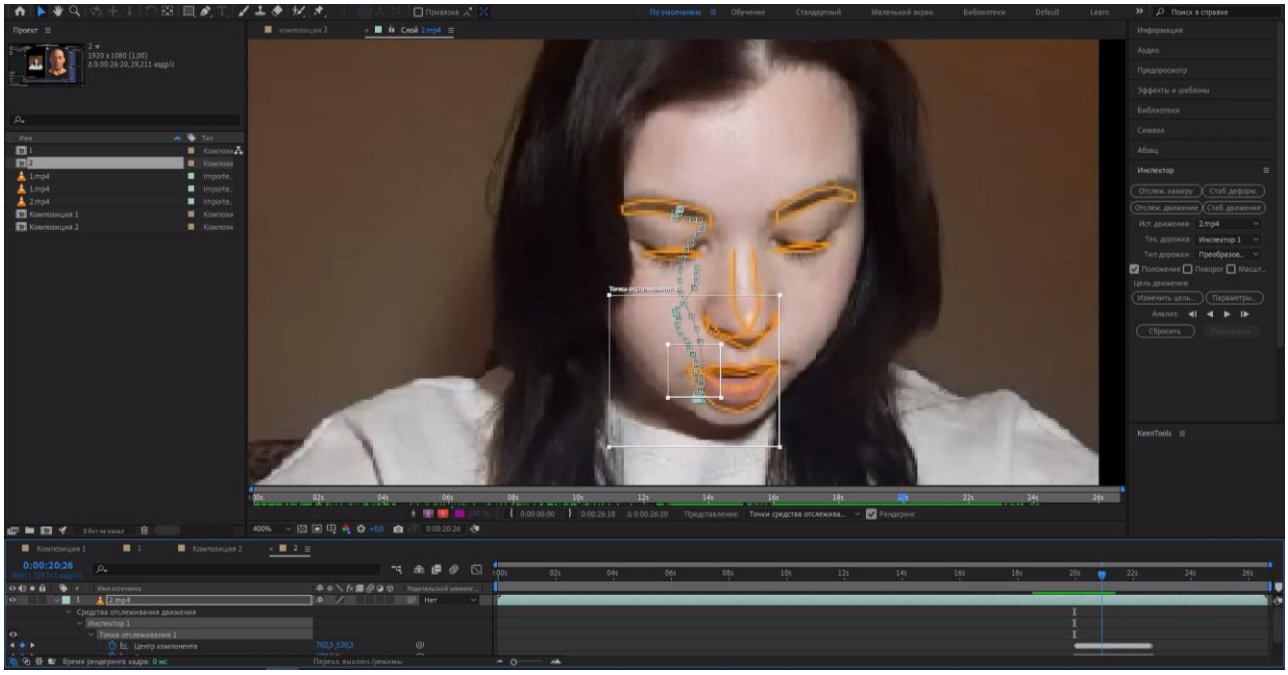

Рисунок 3.20 – Еталонна траєкторія руху в момент нахилу голови людини

Рисунок 3.21 – Траєкторія руху тривимірної моделі в момент нахилу голови

Як видно з графіку на рис. 3.22, похибка аналізу траєкторій з використанням вбудованого трекеру, склала 1,25 pixel.

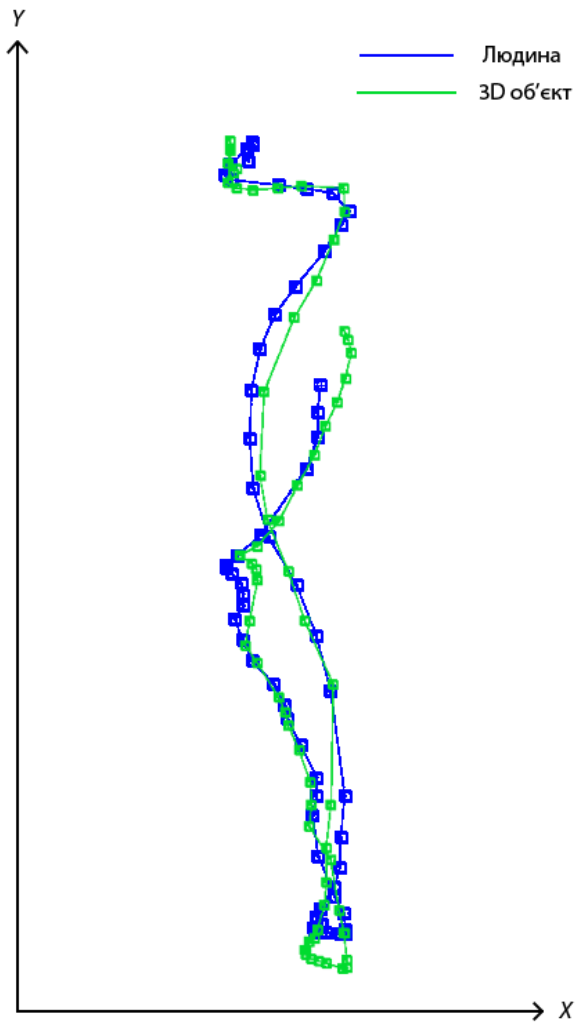

Рисунок 3.22 – Суміщені траєкторії руху контрольної точки (людини та тривимірної моделі) в момент нахилу голови

Для третього експерименту в межах дослідження роботи безмаркерного методу було обрано відео-фрагмент тривалістю 2,3 с, а саме – момент повороту голови вбік. Траєкторії руху контрольної точки приведено на рис.3.23-3.24, на графіку нижче (рис.3.25) наведено суміщені траєкторії руху для візуальної оцінки. Максимальне відхилення між траєкторіями руху склало 140 pixel. Проте, не дивлячись на таке велике значення відхилення траєкторій, анімація тривимірної моделі була плавної, гармонійною, без різких стрибків та

ривків. Тож, можна стверджувати, що в даному випадку безмаркерний метод впорався із завданням.

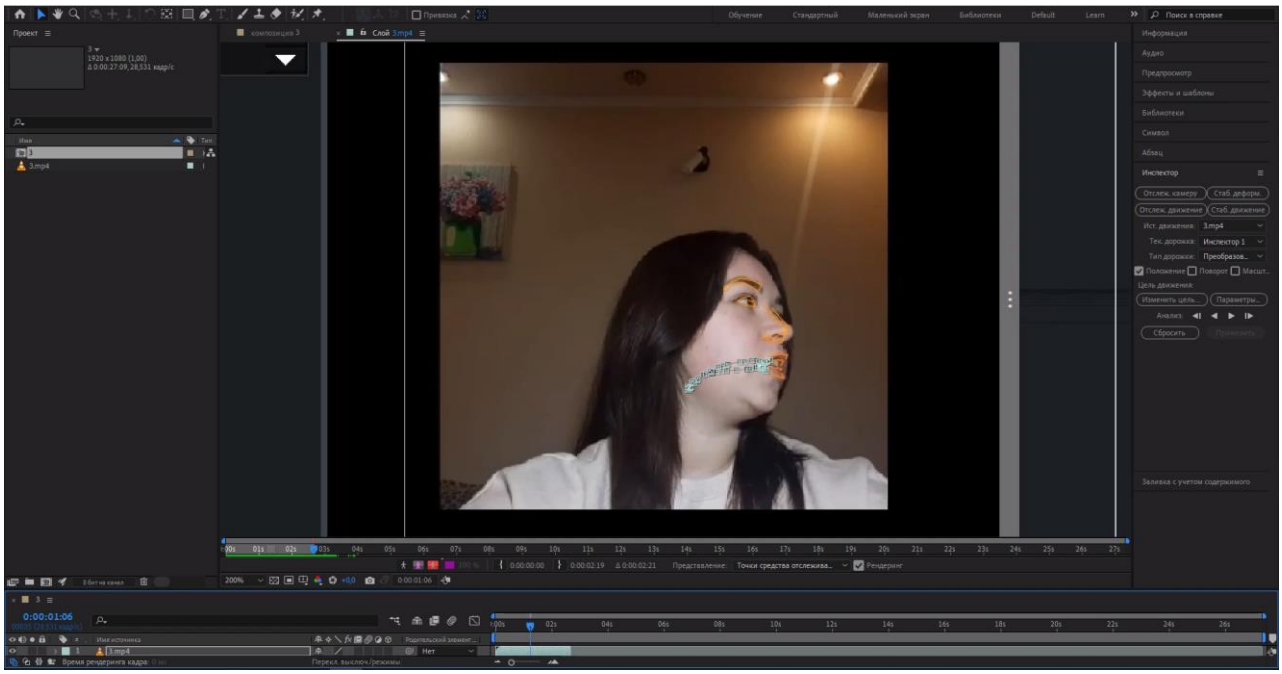

Рисунок 3.23 – Еталонна траєкторія руху контрольної точки при повороті голови

Рисунок 3.24 – Траєкторія руху контрольної точки тривимірної моделі при повороті голови

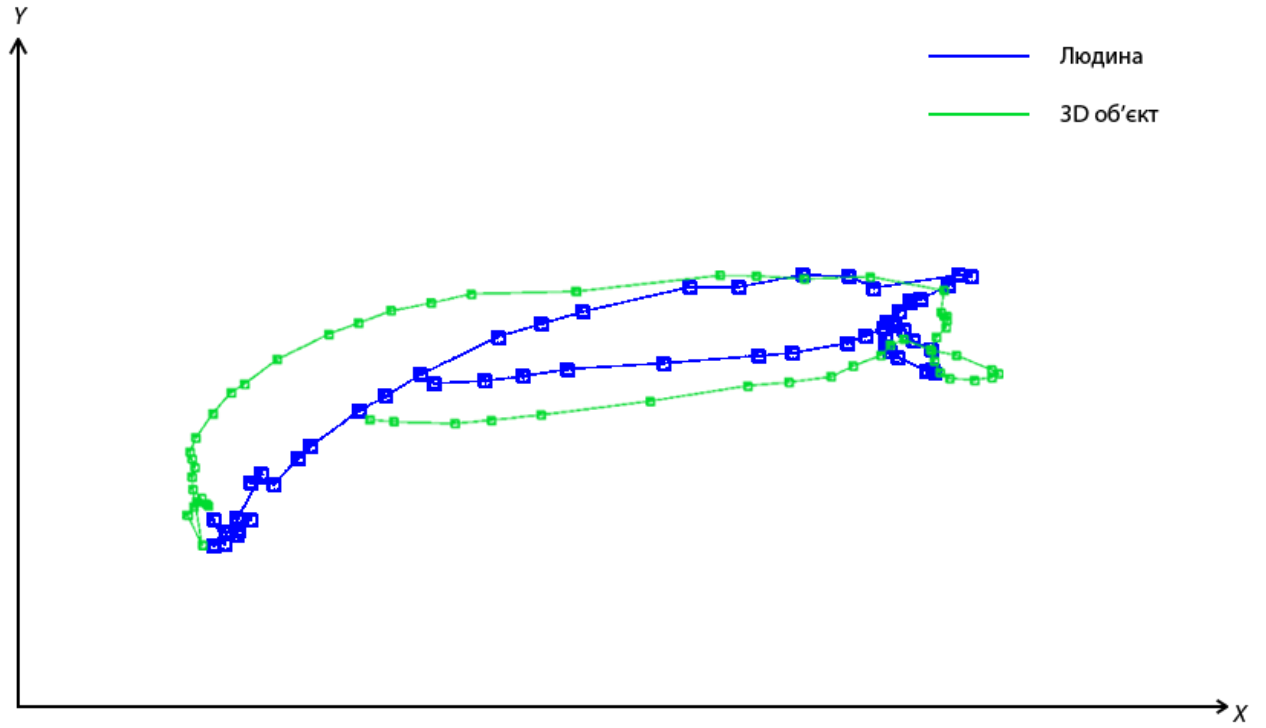

Рисунок 3.25 – Суміщені траєкторії руху контрольної точки (людини та тривимірної моделі) в момент повороту голови

Наступний експеримент був присвячений дослідженню маркерного методу при анімації тривимірної моделі. Спочатку було нанесено контрастні сині маркери (акрилова фарба) на обличчя, відповідно до схеми трекінгових точок, що використовується плагіном Faceware (рис.1.22). Форма маркерів може бути задовільною, головне, щоб можна було до неї прив'язати контрольну точку.

На рис. 3.26-3.27 наведено траєкторії руху контрольної точки. На рис.3.28 наведено суміщені траєкторії руху для візуальної оцінки. Якщо в попередніх випадках максимальне відхилення між траєкторіями руху складала 40 pixel, то при використанні маркерного методу, цей параметр сягає 200 pixel.

Під час дослідження процесу трекінгу маркерним методом було виявлено, що при будь-якому повороті голови процес трекінгу зупинявся. Це повністю підтверджує ефективність використання планарного методу (Віоли-Джонса) трекінгу у програмі Faceware. Саме він проявив себе як стабільний, якісний та надійний спосіб трекінгу контрольних точок та подальшої анімації обличчя

тривимірної моделі. У зв’язку з цим, було прийняти рішення не витрачати зайвий час на безрезультатні дослідження маркерного методу.

Рисунок 3.26 – Еталонна траєкторія руху контрольної точки при маркерному методі

Рисунок 3.27 – Траєкторія руху контрольної точки тривимірної моделі при маркерному методі

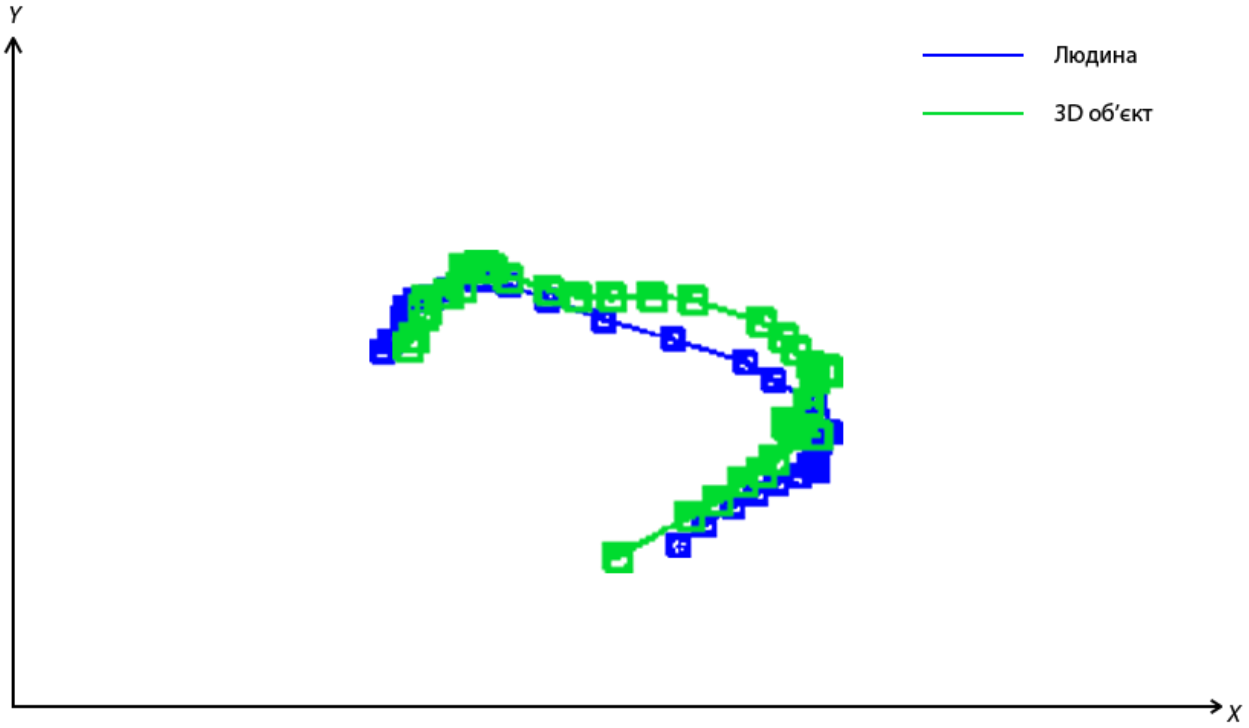

Рисунок 3.28 – Суміщені траєкторії руху контрольної точки (людини та тривимірної моделі) при використанні маркерного методу

ВИСНОВКИ

У даній кваліфікаційній роботі другого (магістерського) рівню освіти на тему «Методи використання трекінгу у відео-конференціях» було розглянуто усі існуючі технології трекінгу обличчя та тіла людини для створення кіно та комп'ютерних ігор. На протязі усього часу найбільшої популярності набула технологія motion capture. Окрім значних переваг, цей спосіб захвату рухів ще й дуже складний: спеціальні костюми, маркери, камери, багато часу на постобробку. До того ж, на нього йде значна частина бюджету проєкту.

Розвиток комп'ютерних технологій та створення нових програмних продуктів дозволили значно спростити процес трекінгу обличчя людини для подальшої анімації персонажів. Метою кваліфікаційної роботи було дослідження існуючих методів та програм для трекінгу обличчя з подальшим використанням у відеоконференціях.

У роботі були розглянуті такі технології, як: deepfake, FaceRig (для заміни обличчя тривимірним персонажем), Snap Camera, плагін Faceware для After Effects та Metahumans на Unreal engine.

Майже кожен з цих програмних засобів можна використовувати для заміни обличчя під час відеоконференцій (в Zoom, Google Meet, Skype та ін.). Їх ефективність буде залежати від вихідного зображення, рівня освітлення і системних параметрів комп'ютера користувача.

Третій розділ кваліфікаційної роботи був присвячений експериментам по виявленню якісного способу трекінгу обличчя для подальшої анімації тривимірної моделі. Під час дослідження було знято декілька відео-фрагментів для аналізу маркерного та безмаркерного методу відстеження контрольних точок на обличчі. Були зняті траєкторії руху контрольної точки (правий кут рота) на обличчі у відео та тривимірної моделі. Самі дослідження було проведено в програмному середовищі After Effects за допомогою плагіну Faceware.

Загалом, безмаркерний метод проявив себе як стабільний та надійний спосіб трекінгу контрольних точок. Навіть при повороті голови в профіль, похибка між траєкторіями голови та моделі не перевищувати 1,5 pixel.

Експеримент з маркерним методом було зірвано майже одразу після початку. Нажаль, при невеличкому повороті голови програмні контрольні точки, що прив’язуються до маркерів на обличчі, одразу «злітали». Тривимірна модель рухалася ривками, не відстежувала рухів губами та постійно зупинялась.

Такі результати експериментів свідчать, що при роботі з програмним пакетом After Effects найкраще працює планарний метод розпізнання обличчя, що базується на математичному методі розпізнання Віоли-Джонса.

Перспективи для даної теми кваліфікаційної роботи досить великі. Досягнувши якісних результатів в трекінгу обличчя, можна переходити до анімації тривимірних моделей персонажів для кіно- та ігрової індустрії. Такий спосіб анімації дозволить акторам «грати» у фільмах навіть після їх смерті.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАНЬ

1. Rotoscoping: From Early Animation to Blockbuster VFX: стаття. URL: https://www.rocketstock.com/blog/rotoscoping-from-early-animation-to-blockbuster-vfx (дата звернення: 21.11.22).
2. How the World of ‘Undone’ Came to Life Through Rotoscope: стаття. URL: https://www.indiewire.com/2019/09/undone-came-to-life-through-rotoscope-1202172916/ (дата звернення: 21.11.22)
3. This Is What 1970s Motion Capture Tech Looked Like: стаття. URL: https://www.vice.com/en/article/wnkbzz/this-is-what-1970s-motion-capture-tech-looked-like (дата звернення: 21.11.22)
4. 100 years of motion-capture technology: стаття. URL: https://www.engadget.com/2018-05-25-motion-capture-history-video-vicon-siren.html (дата звернення: 21.11.22)
5. Filmmakers went totally ape with 'Dawn' motion-capture: стаття. URL: https://www.usatoday.com/story/life/movies/2014/07/08/dawn-planet-apes-serkis-motion-capture-suits/11146633 (дата звернення: 21.11.22)
6. Andy Serkis on His Most Groundbreaking Motion-Capture Performances: стаття. URL: https://www.vulture.com/2018/12/andy-serkis-on-his-breakthrough-motion-capture-performances.html (дата звернення: 21.11.22)
7. Захоплення руху: стаття. URL: https:// thelastofus.fandom.com/ru/wiki/Захват_движения (дата звернення: 22.11.22).
8. Захоплення руху: стаття. URL:https://uk.wikipedia.org/wiki/Захоплення_руху (дата звернення: 22.11.22).
9. Inertial Motion Capture: стаття. URL: https://ps.is.mpg.de/pages/inertial-motion-capture (дата звернення: 22.11.22).
10. Detroit: Become Human Voice Cast: стаття. URL: https://www.vg247.com/detroit-become-human-guide-tips-tricks-and-controls?page=2 (дата звернення: 22.11.22).

11. Як створюють найнеймовірніші спецефекти у кіно: стаття. URL: https://good-news.net.ua/ru/razvlechenyia/kyno/kak-sozdaiut-samye-neveroiatnye-spetseffekty-v-kyno/ (дата звернення: 22.11.22).
12. Blender 2.8 Face Motion Capture Guide: відео. URL: https://www.youtube.com/watch?v=uNK8S19OSmA&ab_channel=I%27mBreakingNews (дата звернення: 22.11.22).
13. MotionScan: стаття. URL: https://ru.wikipedia.org/wiki/MotionScan (дата звернення: 22.11.22).
14. Нова технологія захоплення рухів MotionScan змусить відеогру бути схожою на фільм: стаття. URL: https://dailytechinfo.org/infotech/1939-novaya-texnologiya-zaxvata-dvizhenij-motionscan-zastavit-videoigru-byt-poxozhej-na-film.html (дата звернення: 22.11.22).
15. Технології захоплення руху: Ігри: стаття. URL: https://www.youtube.com/watch?v=Ikv5OmyEXEA&ab_channel=%D0%98%D0%B3%D1%80%D0%BE%D0%BC%D0%B0%D0%BD%D0%B8%D1%8F%D0%9A%D0%B8%D0%BD%D0%BE (дата звернення: 22.11.22).
16. Левенець Т. В., Кравець І. О.Дослідження методів розпізнавання обличь при використанні мобільних технологій. Наукові праці. Комп'ютерні технології. 2016. URL: https://journals.indexcopernicus.com/api/file/viewByFileId/156415.pdf (дата звернення: 22.11.22).
17. Turk, M.A. Face Recognition Using Eigenfaces / M.A. Turk, A.P. Pentland // In: Proc. IEEE, 1991. – P. 586–591. 114. Vapnik, V. Support-vector networks / V. Vapnik, C. Cortes // Machine Learning, 1995. – Vol. 20. – Issue 3. – P. 273–297.
18. Ole, H.J. Implementing the Viola-Jones face detection algorithm: Master's thesis / H.J. Ole. – DTU, 2008. 94. OpenCV [Електроний ресурс]. – Режим доступу: http://old.kv.by/index2001513402.htm. Дата обращения: 19.04.2012.

19. Viola, P. Fast and Robust Classification Using Symmetric AdaBoost and a Detector Cascade / P. Viola, M. Jones. // In: Proc. Neural Information Processing Systems, Vancouver, 2001. – P. 1311–1318.
20. Viola, P. Rapid object detection using a boosted cascade of simple features / P. Viola, M.J. Jones // IEEE Conf. On Computer Vision and Pattern Recognition, Kauai, Hawaii, USA, 2001. – V. 1. – P. 511–518.
21. V. M. Kartashov, V. N. Oleynikov, S. A. Sheyko, I. V. Koryttsev, S. I. Babkin, O. V. Zubkov, "Peculiarities of small unmanned aerial vehicles detection and recognition," Telecommunications and Radio Engineering, 2019, V. 78, Iss. 9, pp. 771–781. DOI: 10.1615/TelecomRadEng.v78.i9.30.
22. V. N. Oleynikov, O. V. Zubkov, V. M. Kartashov, I. V. Korytsev, S. I. Babkin, S.A. Sheiko, "Investigation of detection and recognition efficiency of small unmanned aerial vehicles on their acoustic radiation," Telecommunications and Radio Engineering, 2019, V. 78, Iss. 9, pp. 759–770. DOI: 10.1615/TelecomRadEng.v78.i9.20.
23. Kartashov, V.M., Sidorov, G.I., Sheiko, S.A., Kolendovska, M.M., Sergienko O.Yu. Principles of Construction and Assessment of technical Characteristics of multi-Frequency atmospheric Sodar in the Humidity Measurement Mode / Telecommunications and Radio Engineering.- New York. - 2020.- Vol. 79, №4.- P.323-333. (стаття). DOI: 10.1615/TelecomRadEng.v79.i4.50.
24. Kartashov, V.M., Oleynikov V.N, Zubkov, O.V., Korytsev I.V., Babkin, S. I., Sheiko, S.A.,. Kolendovskaya, M.M. Spatial-temporal Processing of acoustic Signals of Unmanned Aerial Vehicles; Telecommunications and Radio Engineering, 2020. Vol. 79, Iss, 9, pp.769-780.
25. V.M. Semenets, V.M. Kartashov, V.I. Leonidov. Features of Acoustic Noise of Small Unmanned Aerial Vehicles // Telecommunications and Radio Engineering.- New York. - 2020.- Vol. 79, №11.- P. 985-995. DOI: 10.1615/TelecomRadEng.v79.i11.80 (стаття).
26. Oleynikov V.N., Kartashov, V.M., Babkin, S. I., Zubkov, O.V., Korytsev I.V., Sheiko, S.A.,. Seleznov I.S. Structure and Parameter Unmanned Aerial Vehicles

Sound Fields/ Telecommunications and Radio Engineering.- New York. - 2020.- Vol. 79, №17.- P.1539-1550. DOI: 10.1615/TelecomRadEng.v79.i17.50 (стаття).

27. В.А. Тихонов, В.М. Карташов, В.М. Олейніков, В.И. Леонидов, Л.П. Тимошенко, И.С. Селезнеов, Н.В. Рыбников. Обнаружение-распознавание беспилотных летательных аппаратов с использованием составной модели авторегрессии их акустического излучения// Вісник НТУУ «КПІ». Радіотехніка. Радіоапаратобудування. - 2020.- Вип. №81. – С.38-46. (Web of Science)
28. Карташов В.М., Корытцев И.В., Олейников В.Н., Зубков О.В., Шейко С.А., Бабкин С.И., Левский Н.А., Селезнев И.С. Алгоритмы пеленгации беспилотных летательных аппаратов по их акустическому излучению// Радиотехника. (Харьков). — 2019. — Вып. 196. — С. 22-31.
29. Карташов В.М., Харченко О.И., Чумаков В.И. Использование эффекта стохастического резонанса для анализа спектров акустического излучения малых беспилотных летательных аппаратов // Радиотехника. (Харьков). — 2019. — Вып. 197. — С. 100-106.
30. Карташов В.М., Сидоров Г.І., Толстих Є.Г., Шаповалов С.В. Акустичний вимірювач швидкості вітру в атмосферному прикордонному шарі// Радиотехника. (Харьков). — 2019. — Вып. 199. – С. 54-58.
31. Карташов В.М., Посошенко В.А., Цехмистро Р.И., Тимошенко Л.П., Колендовская М.М. Методы ориентации, навигации и контроля мобильных робототехнических платформ// Радиотехника. (Харьков). — 2019. — Вып. 199. – С. 38-44.
32. Олейников В.Н., Зубков О.В, Карташов В.М., Корытцев И.В., Бабкин С.И., Шейко С.А, Селезнев И.С. Экспериментальная оценка эффективности аглоритмов пеленгования беспилотных летательных аппаратов по акустическому излучению// Радиотехника. (Харьков). — 2019. — Вып. 199. – С. 29-37.
33. Карташов В.М., Олейников В.Н., Колендовская М.М., Тимошенко Л.П., Капуста А.И., Рыбников Н.В. Комплексирование изображений при

обнаружении беспилотных летательных аппаратов // Радиотехника. (Харьков). — 2020. — Вып. 201. — С. -120-129.

34. Карташов В.М., Олейников В.Н., Воронин В.В., Рябуха В.П., Капуста А.И., Рыбников Н.В., Селезнев И.С. Методы комплексной обработки и интерпретации радиолокационных, акустических, оптических и инфракрасных сигналов беспилотных летательных аппаратов // Радиотехника. (Харьков). — 2020. — Вып. 202. — С. 173-182-. DOI:10.30837/rt.2020.3.202.19
35. Карташов В.М., Корытцев И.В., Олейников В.Н., Зубков О.В., Шейко С.А., Бабкин С.И., Обработка сигналов при пеленгации и определении дальности до малоразмерных БПЛА в оптическом и инфракрасном диапазонАХ // Радиотехника. (Харьков). — 2020. — Вып. 202. — С. 125-135. DOI:10.30837/rt.2020.3.202.13
36. Карташов В.М., Корытцев И.В., Олейников В.Н., Зубков О.В., Шейко С.А., Бабкин С.И. ОПТИКО-ЭЛЕКТРОННЫЕ МЕТОДЫ ОБНАРУЖЕНИЯ ВОЗДУШНЫХ ОБЪЕКТОВ И ИЗМЕРЕНИЯ ИХ КООРДИНАТ// Радиотехника. (Харьков). — 2020. — Вып. 202. — С. 153-59. DOI:10.30837/rt.2020.3.202.16
37. Карташов В.М., Корытцев И.В., Олейников В.Н., Зубков О.В., Шейко С.А., Бабкин С.И.. Эффективность детектирования и распознавания изображений дронов по видеопотоку стационарной видеокамеры // Радиотехника. (Харьков). — 2020. — Вып. 202. — С. 136-146. DOI:10.30837/rt.2020.3.202.14
38. Рябуха В.П., Карташов В.М. Методы обнаружения-распознавания радиолокационных, акустических, оптических и инфракрасных сигналов беспилотных летательных аппаратов / В.П. Рябуха, В.М. Карташов// Известия высших учебных заведений. Радиоэлектроника. — 2020. — Т. 63, № 11. — С. 1–35.
39. В.М. Карташов, В.Н. Олейников, В.И. Леонидов, канд. Техн. Наук, В.В. Воронин, А.И. Капуста, И.С. Селезнев, Е.В. Першин/ Комплексная обработка сигналов интегрированной системы наблюдения беспилотных летательных

аппаратов с использованием целеуказания//Радиотехника. (Харьков). - 2020. - Вып. 203. - С. 1-13.

40. Face Recognition with Python, in Under 25 Lines of Code: стаття. URL: https://realpython.com/face-recognition-with-python/ (дата звернення: 23.11.22).
41. Face Detection in 2 Minutes using OpenCV & Python: стаття. URL: https://towardsdatascience.com/face-detection-in-2-minutes-using-opencv-python-90f89d7c0f81 (дата звернення: 23.11.22).
42. Learn how to create real-time facial animation for your MetaHumans: стаття. URL: https://www.unrealengine.com/en-US/blog/learn-how-to-create-real-time-facial-animation-for-your-metahumans (дата звернення: 23.11.22).
43. Faceware Studio in 2021: стаття. URL: https://www.fxguide.com/quicktakes/faceware-studio-in-2021/ (дата звернення: 23.11.22).
44. A Zelensky deepfake was quickly defeated. The next one might not be: стаття. URL: https://www.wired.com/story/zelensky-deepfake-facebook-twitter-playbook/ (дата звернення: 24.11.22).
45. What is a deepfake? : стаття. URL: https://www.smh.com.au/technology/what-is-the-difference-between-a-fake-and-a-deepfake-20200729-p55ghi.html (дата звернення: 24.11.22).
46. Animaze by FaceRig: стаття.URL: https://store.steampowered.com/app/1364390/Animaze_by_FaceRig/?l=ukrainian (дата звернення: 25.11.22).
47. Snap Camera. Get starting: стаття. URL: https://snapcamera.snapchat.com/guides/getting-started (дата звернення: 25.11.22).
48. Як змінювати фон та включати візуальні ефекти в Google Meet: стаття.URL:https://support.google.com/meet/answer/10058482?hl=ru&co=GENIE.Platform%3DDesktop (дата звернення: 24.11.22).

49. Reface: стаття.URL:https://uk.wikipedia.org/wiki/Reface (дата звернення: 25.11.22).
50. Epic Games випустила Unreal Engine 5 у ранній доступ і відкрила технодемо Valley Of The Ancient: стаття.URL: https://www.overclockers.ua/news/games/2021-05-27/129274/ (дата звернення: 26.11.22).
51. Quixel Bridge Plugin for Unreal Engine: стаття.URL: https://docs.unrealengine.com/5.0/en-US/quixel-bridge-plugin-for-unreal-engine/ (дата звернення: 27.11.22).