

Měření přesnosti Metahuman Animatoru oproti Rokoko facial motion capturing systému

Vojtěch Benešovský

Bakalářská práce
2025



Univerzita Tomáše Bati ve Zlíně
Fakulta multimediálních komunikací

Univerzita Tomáše Bati ve Zlíně
Fakulta multimediálních komunikací
Ateliér Audiovize

Akademický rok: 2024/2025

ZADÁNÍ BAKALÁŘSKÉ PRÁCE

(projektu, uměleckého díla, uměleckého výkonu)

Jméno a příjmení:	Vojtěch Benešovský
Osobní číslo:	K21232
Studijní program:	B0211P310005 Teorie a praxe audiovizuální tvorby
Specializace:	Vizuální efekty
Forma studia:	Prezenční
Téma práce:	1. Teoretická část: Měření přesnosti Unreal Metahuman Animatoru oproti Rokoko facial motion capturing systému 2. Praktická část: Významná triková spolupráce na bakalářském filmu (nezbytná délka závislá na náročnosti technologie po konzultaci a schválení Výrobní komisí AAV). VFX na souboru audiovizuálních děl, oficiálně schváleného před odevzdáním Výrobní komisí AAV, ve výstupní kvalitě uvedené ve Výrobní knize AAV, v minimální délce 4 minuty. Samostatné audiovizuální dílo založené na využití VFX v délce odpovídající námětu a náročnosti technologie, v minimální délce 3 minuty. Varianta musí být schválena před odevzdáním Výrobní komisí ateliéru Audiovizuální tvorba. viz Zásady pro vypracování

Zásady pro vypracování

1. Teoretická část:

Rozsah práce: minimálně 15 normostran textu bez započítání obsahu, rejstříku a obrazových příloh.

Formální podoba: Jednotná formální úprava teoretické části práce, její uložení a zpřístupnění se řídí aktuální verzí příslušné směrnice rektora. Student odevzdává 1 ks fyzické (tištěné) práce v pevné vazbě. Tištěná verze práce obsahuje originální ?Zadání DP/BP? včetně příslušných podpisů a studentem podepsané Prohlášení o původnosti práce. Práce v elektronické podobě obsahuje nascanované ?Zadání DP/BP? se všemi formálními náležitostmi a také nepodepsané Prohlášení studenta o původnosti práce. Plný text elektronické verze ve formátu PDF/A a případné přílohy (zkomprimované do jednoho zip souboru) student odevzdá nahráním do IS/STAG a do příslušné složky na NAS-AAV (viz níže).

Pokyny k vypracování: prostudujte a analyzujte dostupné materiály z profesního hlediska a formulujte závěry a získané vědomosti do podoby akademického/odborného textu.

2. Praktická část:

Přípustné varianty praktické části:

1) Významná triková spolupráce na bakalářském filmu (nezbytná délka závislá na náročnosti technologie po konzultaci a schválení Výrobní komisí AAV).

2) VFX na souboru audiovizuálních děl, oficiálně schváleného před odevzdáním Výrobní komisí AAV, ve výstupní kvalitě uvedené ve Výrobní knize AAV, v minimální délce 4 minuty.

3) Samostatné audiovizuální dílo založené na využití VFX v délce odpovídající námětu a náročnosti technologie, v minimální délce 3 minuty. Varianta musí být schválena před odevzdáním Výrobní komisí ateliéru Audiovizuální tvorba.

Další požadované materiály praktické části:

a) VFX breakdown (?making-of?, ?behind-the-scenes?) k předloženému audiovizuálnímu dílu. (var. 1, 2, 3).

b) Písemná explikace z pohledu dané specializace. Minimální rozsah 2 normostrany (var. 1, 2, 3).

c) Anotace (var. 1, 2, 3).

d) Storyboard, případně animatik (var. 1, 2, 3).

e) Štábová listina (var. 1, 2).

V případě, že je dílo autorským počinem nebo není součástí praktické části SZS studenta Produkce, je nutné dodržet doložení požadovaných materiálů a?h dle zadání specializace Produkce. Tato data odevzdává za projekt vždy jeden člověk. Nezbytná je konzultace s vedením AAV.

Všechny odevzdávané materiály musí splňovat vnitřní technické normy dle Výrobní knihy AAV pro odevzdávání prací a musí být řádně popsány (jméno, název, logo fakulty, formát, rozlišení). Součástí závěrečné práce je vytištěný a podepsaný formulář "Údaje o bakalářské práci studenta".

Uložení na NAS:

Ve složce na NAS-AAV, označené ?Bakalářská / Magisterská práce? uložte:

1. Teoretickou práci ve formátu PDF/A a případné přílohy (zkomprimované do jednoho zip souboru) dle specifikací výše.

2. Vytvořte podsložku Praktická práce, která bude obsahovat materiály částí a- h. Řádně nazvaný film/absolventské dílo odevzdávejte ve formátech splňujících vnitřní technické normy AAV pro odevzdávání prací.

3. Vytvořte podsložku s názvem Katalog, která bude obsahovat "Podklady pro katalog FMK UTB ve Zlíně": 10 kusů obrazové dokumentace praktické části závěrečné práce pro využití v publikacích FMK. Formát pro bitmapové podklady: JPEG, barevný prostor RGB, rozlišení 300 dpi, 250 mm delší strana. Formáty pro vektory: AI, EPS, PDF. Loga a texty v křivkách. V samostatném textovém souboru uveďte jméno a příjmení, login do portálu UTB, obor (ateliér), typ práce, přesný název práce v češtině i v angličtině, rok obhajoby, osobní e-mail, osobní web, telefon. Přiložte svou osobní fotografii v tiskovém rozlišení.

Forma zpracování bakalářské práce: **tištěná/elektronická**

Seznam doporučené literatury:

KITAGAWA, Midori a WINDSOR, Brian. MoCap for artists: workflow and techniques for motion capture. Amsterdam: Elsevier/Focal Press, c2008. ISBN 978-0-240-81000-3.
DOWER, John a Pascal LANGDALE. Performing for Motion Capture: A Guide for Practitioners. Methuen Drama, 2022. ISBN 978-1350211254.
MENACHE, Alberto. Understanding Motion Capture for Computer Animation. 2nd. Morgan Kaufmann, 2010, 276 s. ISBN 978-0-12-381496-8.
ZWERMANN, Susan a Jeffrey A. OKUN, ed. The VES Handbook of Visual Effects: Industry Standard VFX Practices and Procedures. Focal Press, 2010, 948 s. ISBN 978-0240812427.
TOBON, Ricardo, RESTREPO, Alex, ed. The Mocap Book: A Practical Guide to the Art of Motion Capture. Foris Force, 2010, 288 s. ISBN 978-0615293066.
PARENT, Rick, David S. EBERT a Mark V. PAULY. Computer Animation Complete: All-in-One: Learn Motion Capture, Characteristic, Point-Based, and Maya Winning Techniques. Morgan Kaufmann, 2009, 544 s. ISBN 978-0123750785.
ZHANG, Manyu. Application of Performance Motion Capture Technology in Film and Television Performance Animation [online]. Tianjin Academy of Fine Arts Tianjin, CHINA, 2013 [cit. 2023-10-25]. Dostupné z: <https://www.atlantispress.com/article/5878.pdf>. Druh práce neznámý. Tianjin Academy of Fine Arts Tianjin, Dept. Animation of Art.
LIU, Bobby. Capture on the Virtual Stage: a Guide to Successful Motion Capture Production. Medium.com [online]. 2023 [cit. 2023-10-25]. Dostupné z: <https://bunby8motion913.medium.com/capture-on-the-virtual-stage-a-guide-to-successful-capture-production-99ab1dbd867f>.

Vedoucí bakalářské práce: **ak. mal. Boris Masník**
Ateliér Audiovize

Konzultant bakalářské práce: **MgA. Marcel Legindi, Ph.D.**
Ateliér Audiovize

Datum zadání bakalářské práce: **2. prosince 2024**

Termín odevzdání bakalářské práce: **16. května 2025**

L.S.

Mgr. Josef Kocourek, Ph.D.
děkan

MgA. Irena Pustina Kocí, Ph.D., MBA
vedoucí ateliéru

Ve Zlíně dne 2. prosince 2024

PROHLÁŠENÍ AUTORA BAKALÁŘSKÉ PRÁCE

Beru na vědomí, že

- odevzdáním bakalářské práce souhlasím se zveřejněním své práce podle zákona č. 111/1998 Sb., v platném znění bez ohledu na výsledek obhajoby;
- bakalářská práce bude uložena v elektronické podobě v univerzitním informačním systému a bude dostupná k nahlédnutí;
- jedno vyhotovení bakalářské práce v listinné podobě bude ponecháno Univerzitě Tomáše Bati ve Zlíně k uložení;
- na moji bakalářskou práci se plně vztahuje zákon č. 121/2000 Sb. o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon) ve znění pozdějších právních předpisů, zejm. § 35 odst. 3;
- podle § 60 odst. 1 autorského zákona má Univerzita Tomáše Bati ve Zlíně právo na uzavření licenční smlouvy o užití školního díla v rozsahu § 12 odst. 4 autorského zákona;
- podle § 60 odst. 2 a 3 mohu užít své dílo – bakalářskou práci – nebo poskytnout licenci k jejímu využití jen s předchozím písemným souhlasem Univerzity Tomáše Bati ve Zlíně, která je oprávněna v takovém případě ode mne požadovat přiměřený příspěvek na úhradu nákladů, které byly Univerzitou Tomáše Bati ve Zlíně na vytvoření díla vynaloženy (až do jejich skutečné výše);
- pokud bylo k vypracování bakalářské práce využito softwaru poskytnutého Univerzitou Tomáše Bati ve Zlíně nebo jinými subjekty pouze ke studijním a výzkumným účelům (tj. k nekomerčnímu využití), nelze výsledky bakalářské práce využít ke komerčním účelům;
- pokud je výstupem bakalářské práce jakýkoliv softwarový produkt, považují se za součást práce rovněž i zdrojové kódy, popř. soubory, ze kterých se projekt skládá; neodevzdání této součásti může být důvodem k neobhájení práce.

Prohlašuji, že

- jsem na bakalářské práci pracoval(a) samostatně a použitou literaturu jsem řádně citoval(a); v případě publikace výsledků budu uveden(a) jako spoluautor;
- odevzdaná verze bakalářské práce a verze elektronická nahraná do IS/STAG jsou obsahově totožné.

Ve Zlíně, dne

.....

podpis autora

ABSTRAKT

Práce se okrajově věnuje základnímu definování typů motion capture. Její hlavní cíl je ale čtenáře seznámit s možností záznamu facial capture za pomoci softwaru jako je Rokoko Studio a MetaHuman Animator, kde následně u obou těchto typů zkoumá přesnost jejich snímání za pomoci měřeného experimentu.

Klíčová slova: facial motion capture, Unreal Engine, Rokoko, Metahuman Animator, Live Link

ABSTRACT

The thesis briefly outlines the fundamental types of motion capture; however, its primary objective is to introduce the reader to the possibilities of facial capture using software such as Rokoko Studio and MetaHuman Animator. It then proceeds to examine the accuracy of both systems through a controlled experimental measurement.

Keywords: facial motion capture, Unreal Engine, Rokoko, Metahuman Animator, Live Link

Prohlašuji, že odevzdaná verze bakalářské/diplomové práce a verze elektronická nahraná do IS/STAG jsou totožné.

OBSAH

ÚVOD.....	11
I TEORETICKÁ ČÁST	12
1 MOTION CAPTURE.....	13
1.1 BODY MOTION CAPTURE.....	13
1.2 FACIAL MOTION CAPTURE.....	14
1.2.1 Marker-based performance capture.....	15
1.2.2 Markerless performance capture	17
2 SOFTWARE K ZÁZNAMU FACIAL CAPTURE.....	20
2.1 ROKOKO.....	21
2.2 UNREAL ENGINE 5.....	22
II PRAKTICKÁ ČÁST	24
3 ZADÁNÍ EXPERIMENTU	25
3.1 VÝZKUMNÉ OTÁZKY	26
III PROJEKTOVÁ ČÁST.....	28
4 EXPERIMENT	29
4.1 REALITY CAPTURE.....	29
4.2 TRACKOVACÍ BODY NA RTY	32
4.3 KEENTOOLS PLUGIN.....	33
4.4 METAHUMAN IDENTITY	34
4.5 NASTAVENÍ TELEFONŮ	38
5 DEFINICE POUŽITÝCH SOFTWAREŮ	39
5.1 ROKOKO STUDIO LEGACY.....	39
5.2 METAHUMAN ANIMATOR.....	40
5.2.1 Blendshapes.....	41
6 PŘENOS DAT	43
6.1.1 LiveLink Rokoko	43
6.1.2 Nastavení sekvence	44
7 NAHRÁNÍ DAT	46
7.1 VYTVOŘENÍ ANIMACE V METAHUMAN ANIMATORU	47
7.2 TRACKING	49
8 MĚŘENÍ NASBÍRANÝCH DAT	51
8.1 SKRIPT.....	51
8.2 PŘESUNUTÍ HODNOT DO TABULEK	52
9 ANALÝZA MĚŘENÍ.....	53

10 ZÁVĚR	56
SEZNAM POUŽITÉ LITERATURY	57
SEZNAM OBRÁZKŮ	60
SEZNAM TABULEK.....	61
SEZNAM POUŽITÝCH SYMBOLŮ A ZKRATEK	62

ÚVOD

Technologie nejen v oblasti vizuálních efektů ale i všude jinde se vyvíjí mílovými kroky vpřed. To, co bylo před pěti lety moderní je dnes v oblasti technologií už zastaralé, nahrazené mnohem vyspělejšími produkty, nabízejícím za levnější cenu klidně i více funkcí. Skvělým příkladem jsou mobilní telefony, kde kvůli vysoké kompetitivnosti a velkým investicím vznikají každý rok modely, vybaveny třeba kamerovými systémy, které na úkor své velikosti pořizují výborné snímky, nastavující vyšší a vyšší nároky na DSLR fotoaparáty.

Inovace se odehrávají také na poli motion capture, zejména v odvětví záznamu tváře člověka. Ta byla vždy jednou z nejtěžších partií lidského těla, protože jakákoli emoce, kterou svým obličejem vyjádříme je souzněním několika svalů najednou, nehledě na jemné mimické svaly okolo očí, které ani my sami nevnímáme. Naše okolí je ale vidí, a jsou to právě tyto svaly, které například určují, zda náš smích není falešný ale upřímný. S těmito výzvami se už nějakou řadu let potýkají i studia velkých filmových i herních společností, a za tu dobu přinesly širokou škálu poznatků, jak nejlépe zachytit lidský obličej.

Jejich výsledky můžeme dnes vidět v téměř každém filmu či AAA videohře, kde hlavní roli hrají také digitální postavy s mimikou opravdových herců. Slouží přes všechnu svou složitost, vytvořenou týmy specialistů, matematiků a inženýrů, stejnému účelu jako klasický natočený záběr, a to především k vyprávění příběhu. To je celé jádro všech těchto technologií, přenést něco co už známe a co považujeme za samozřejmost do jiného světa v ideálně nezměněné podobě.

Divák je totiž nemilosrdný soudce, a jakákoliv viditelná nedokonalost ho chtě nechť upoutá, protože neodpovídá stejnému vzorci, jako v realitě. Nároky na produkci přirozených výrazů v tváři jsou tudíž obrovské. Naštěstí se díky stále modernějším a výkonnějším přístrojům daří tyto nároky vesměs splnit, a to nejen na půdě profesionálů, kteří k tomu mají uzpůsobené vybavení, ale nově také několik těchto technologií mohou využít zcela zdarma i běžní lidé za pomoci obyčejného mobilního telefonu.

Tato bakalářská práce se právě těmito systémům, poskytnutým společnostmi Epic Games a Rokoko zabývá, definuje jejich vlastnosti a srovnává jejich přesnost.

I TEORETICKÁ ČÁST

1 MOTION CAPTURE

1.1 BODY MOTION CAPTURE

Motion capture je proces zakódování pohybu z reálného světa na digitální médium v trojrozměrném prostoru. Jednoduše řečeno, je to technika pro měření pohybů. Jakmile je záznam hotov, může poskytnout ovládání digitální postavy nebo vdechnout život neživým objektům.

Performance capture je termín označující podkategorii motion capture, kde je nahráván jak pohyb hercova těla, tak i jeho obličej.

Motion capture se v odvětví vizuálních efektů stalo široce využívanou metodou záznamu pohybu pro svou schopnost snímat vysoce věrohodnou animaci postavy v časovém rámci, ve kterém je stejná činnost za použití jiných metod prakticky nedosažitelná.

Existuje nespočet technologií motion capture (piezoelektrické, magnetické, inerciální, radiofrekvenční), které se různými způsoby snaží o stejný cíl, a to co nejvěrohodněji zaznamenat pohyb, nejrozšířenější jsou ale systémy optické.

Termín pasivní optický systém se odkazuje na bílé kuličky, které jsou zaznamenávány vysokorychlostními digitálními kamerami. Tyto kuličky, známé jako *markery*, jsou nazývané pasivními proto, jelikož nevyzařují světlo či jinou další informaci. Optický motion capture využívá řadu speciálních kamer, které pozorují scénu z různých úhlů. Stejně jako binokulární vidění umožňuje lidem vidět svět ve třech dimenzích, použití dvou nebo více kamer sledující stejný objekt nám umožňuje znovu rekonstruovat tento objekt ve 3D. Již zmíněné markery jsou umístěny na speciálním obleku. Ten těsně přiléhá k tělu, aby co nejvíce eliminoval případný posun markerů, které jsou suchým zipem připevněny na látku obleku. Obleky jsou obvykle černé, aby snížily odraz světla na minimum. Co ovšem odráží světlo velmi dobře jsou markery (vyrobeny ze stříbrného reflektivního vinylu) které snadno software v kameře rozpozná. Následné nahrávání těchto bodů během pohybu, nám umožňuje určit pozici hercova těla v jakýkoli okamžik.¹

V této práci se ovšem tímto typem motion capture zabývat nebudeme, proto je tato stránka jen velmi obecný úvod do světa tohoto odvětví. Zkráceně lze říci, že ve filmovém a

¹ OKUN, J. A., ZWERMAN, S. *The VES Handbook of Visual Effects: Industry Standard VFX Practices and Procedures. 1st Edition.* Focal Press/Elsevier, 2010. ISBN 978-0240812427, str. 335

videoherním průmyslu jde o zachycení lidského, nebo i zvířecího pohybu a jeho aplikace na digitálního dvojníka.

1.2 FACIAL MOTION CAPTURE

FaceCap je mladší a méně vyvinutý sourozenec body motion capture. Zatímco snímání pohybu těla se primárně snaží zachytit a rekonstruovat úhly kloubů složené kostry, snímání pohybů obličeje se zaměřuje na měření jemnějších deformací obličejové tkáně. Percepční pohyb hraje ve FaceCap důležitější roli, protože zachytit emoce herce a jeho artikulaci je pro správné vyjádření situace klíčové, a tyto aspekty nejsou tak snadno definovatelné jako úhly kloubů.²

Metody k zaznamenání FaceCap můžeme obecně rozdělit na dvě kategorie: technika s markery (marker based) a technika bez nich (markerless). Technika s markery využívá již zmíněných markerů aplikovaných na hercův obličej, takže capture systém může měřit jejich pohyb. Tyto markery mohou být v podobě reflexivních menších kuliček nalepených na obličej, či make-upem nanesené markery přímo na kůži. Metoda bez markerů uplatňuje jako sledování rysů hercovu obličejovou strukturu.³

Přidáním markerů na obličej lze snadno a přesně dosáhnout herecké akce, protože markery nabízejí jasně vymezený fiduciální bod (objekt umístěný v zorném poli obrazového systému, který se objevuje na vytvořeném obraze, sloužící jako referenční bod nebo měřidlo)⁴, který lze sledovat a rozlišovat z různých pohledů. Unikátní povrchová struktura je klíčová v jakémkoli systému, kde je 3D geometrie vytvářena z obrazů objektu. S ohledem na polohu fiduciálního bodu ve dvou nebo více kamerách lze jeho polohu v trojrozměrném prostředí vypočítat pomocí techniky triangulace, která ve zkratce z 2D pozic markerů jednotlivých kamer vytvoří paprsky, jejichž průsečík udává 3D souřadnici.⁵

² OKUN, J. A., ZWERMAN, S. *The VES Handbook of Visual Effects: Industry Standard VFX Practices and Procedures. 1st Edition*. Focal Press/Elsevier, 2010. str. 370

³ Tamtéž, str. 371

⁴ *Fiducial marker*. Online. In: Wikipedia. Stránka byla naposledy editována 7.11.2023 v 19:58. Dostupné z: https://en.wikipedia.org/wiki/Fiducial_marker [cit. 2024-01-27].

⁵ HARTLEY, R. I., STURM, P., *Triangulation*. Version 1. Online, dataset. In: Elsevier. 19. dubna 2002. str. 146, Dostupné z: <https://doi.org/10.1006/cviu.1997.0547>. [cit. 2025-04-23].

Jemné pohyby tváře jsou pro animátora vždy problematickou částí práce, musí animovat to, co je pro nás přirozený a nevědomý pohyb, bez kterého ale působí výraz v tváři uměle a nepřirozeně. Tento problém řeší do jisté míry právě facial motion capture, kde je ale pro změnu zapotřebí komplexní technika k výpočtu této geometrie na animaci postavy.

1.2.1 Marker-based performance capture

Aplikací markerů na obličej lze herecký výkon přesně a spolehlivě zaznamenat. Markery slouží jako jasné vymezené referenční body, které lze snadno sledovat z více kamerových úhlů a tyto body jsou zásadní při rekonstrukci 3D geometrie z obrazových dat, kdy se jejich prostorová poloha určuje například již zmíněnou triangulací. Nevýhodou markerového přístupu však je, že samotné markery mohou do určité míry omezovat herecký projev a kladou určité technické nároky na snímací prostor i samotnou produkci.

Optické markery

Optické retroreflexní markery, běžně používané ve spojení s mocap kamerami, využívající filtrové čočky (účel filtru je eliminovat veškeré světlo kromě toho, co je odraženo od markerů), vytváří v každé z nich snadno detekovatelné body. Podobně jako u optických systémů pro snímání pohybu těla, je přesnost rekonstrukce bodu v 3D prostoru přímo úměrná počtu použitých kamer a velikosti prostoru, který pokrývají.

Snímání mimiky je o něco náchylnější k nepřesnostem, protože se na obličej musí použít menší markery než na tělo. Pro kompenzaci nepřesností se musí použít více kamer v menším prostoru, což již z podstaty omezuje rozsah možného výkonu herců.

Důležitá výhoda těchto systémů ale je, že software používaný na výpočet pozic markerů je automatický a vyžaduje pouze dozor obsluhy programu. S dostatkem kamer a malým prostorem je tak možné docílit velmi přesných výsledků ovšem na úkor několika zásadním aspektům.

Pokud vynecháme již zmíněný velký počet kamer, které jsou potřeba k redukci chybovosti při určení lokace markeru ve 2D (nedokonalé zpracování obrazu, rozmazání pohybu atd.) ovlivňující zároveň i 3D rekonstrukci, přichází problém se zakrýváním některých markerů. Chyby způsobené tzv. „okluzí“ se projevují v následných mocap datech náhodným „vyskočením“ z normálu, často kvůli tomu, že je marker náhle viděn pouze

menším množstvím kamer, nebo protože ho systém chybně identifikuje mezi jednotlivými snímajícími kamerami.⁶

Dalším náročným aspektem je najít standardizované a zároveň optimální rozložení markerů. Na rozdíl od snímání pohybu celého těla, kde může jako vodítko pro umístění markerů sloužit kinematická struktura lidské kostry, neexistuje u snímání mimiky jednotné standardizované rozložení markerů z důvodu anatomické komplexnosti a velké tvarové variability mezi jednotlivci.

Z těchto důvodů jsou výzkumníci i profesionálové v oblasti vizuálních efektů často nuceni vytvářet vlastní, empiricky navržené konfigurace markerů, lišící se podle požadavků produkce či náročnosti snímku. Zároveň je důležité dodržet a mít na paměti následující obecné doporučení.

Příliš husté umístění v jedné oblasti zapříčiňuje, kromě zbytečné redundance v zaznamenaných datech, i větší pravděpodobnost selhání rozpoznávacího optického systému v rozlišení jednotlivých markerů mezi sebou. Pokud je jich zase málo v oblastech s jemnými a dynamickými pohyby obličeje, mohou být některé informace o pohybu dané části nedostatečně zastoupeny.⁷

Make-up markery

Druhým a v dnešní době mnohem více rozšířeným typem marker-based metody snímání obličeje jsou tzv. make-up markery, tečky nakreslené přímo na tvář. Ty jsou pro herce obecně pohodlnější než přilepovací retroreflexní varianta, protože neodpadávají a nijak neomezují výkon, přitom si stále zachovávají možnost rekonstrukce pohybu obličejových svalů.

Na rozdíl od retroreflexních markerů však make-up markery vyžadují záznam standartního non-interlaced (nepřekládaného, plně vykresleného) videa, což může znamenat nižší snímkovou frekvenci než už metody s optickými markery a menší rozlišení kvůli náročnosti ukládání obrazových dat.

Přechod na videozáznam však otevírá nové hardwarové možnosti, kupříkladu použití speciálně upravených helem, na kterých jsou připevněny kamery (HMC – head mounted

⁶ OKUN, J. A., ZWERMAN, S. *The VES Handbook of Visual Effects: Industry Standard VFX Practices and Procedures. 1st Edition.* Focal Press/Elsevier, 2010. str. 372

⁷ LE B. H., ZHU, M. a DENG, Z. *Marker Optimazation for Facial Motion Acquisition and Deformation.* Version 1. Online, dataset. In: IEEE. 17. května 2013. Dostupné z: <https://doi.org/10.1109/tvcg.2013.84>. [2025-4-23].

cameras) snímající fixně hercův obličej při jakékoli akci. Tento záznam je nejen hlavním pilířem rekonstrukčního algoritmu pro vytvoření realisticky se hýbajícího CGI obličeje, ale poskytuje i cennou referenci pro animátory, kteří pak následně doladují finální výstup.

Nejzásadnější výhodou HMC je ale to, že umožňuje herci volný pohyb a odstraňuje omezení malého snímacího prostoru. Kvůli tomu, že helma i s namontovanou kamerou musí být lehká, aby herce příliš neomezovala v rotaci a náklonu hlavy, vzniká tím i nucený kompromis v podobě jistého omezení na rychlost snímkové frekvence, komprese obrazu nebo rozlišení, byť se zdá, že v posledních letech se tyto kompromisy zredukovaly již na minimum.⁸

1.2.2 Markerless performance capture

Metoda zachycení mimiky obličeje bez použití jakýchkoli markerů může hrát v preprodukcí významnou roli, obzvlášť v dnešní rychlé době, kdy každá hodina líčení navíc představuje významný nárůst produkčních nákladů. Pro úvodní seznámení zvolím způsob snímání tváře pomocí hloubkových dat, jednou z mnoha možností bezmarkerového performance capture (image-based, stereoskopická atd.), která je do jisté míry napojená na metodu v praktické části a je důležité se s ní seznámit.

V červenci roku 2011 byla publikována vědecká studie čtyř tehdejších doktorandů Švýcarské federální technické vysoké školy v Lausanne (EPFL)⁹, která nejenže prokázala, že kvalitní animace mimiky může být dosažena i s využitím relativně jednoduchého a finančně dostupného hardwaru, ale zároveň definovala zcela nový a inovativní směr výzkumu v oblasti a rekonstrukce obličejových výrazů.

Technologický hardware, který Thibaut Weise a jeho tým použili, byl původně určený k hraní videoher vlastním tělem. Kinect byl samostatně se prodávající hardware k herní konzoli Xbox 360, pomocí něhož pak šlo ovládat vybrané hry vlastním tělem, kdy postava ve hře replikovala stejné pohyby, které hráč před televizí vytvářel.

Jednalo se o senzor podobný webkameře, který má ve své podstatě dvě separátní kamery: infračervenou kameru a standartní RGB kameru, jejichž data pak zpracovává

⁸ OKUN, J. A., ZWERMAN, S. *The VES Handbook of Visual Effects: Industry Standard VFX Practices and Procedures. 1st Edition.* Focal Press/Elsevier, 2010. str. 373-374

⁹ WEISE, T., BOUAZIZ, S., LI, H., PAULY, M. *Realtime Performance-Based Facial Animation.* Version 1. Online, dataset. In: ACM Transaction on Graphics. 25. července 2011. Dostupné z: <https://doi.org/10.1145/2010324.1964972>. [cit. 2025-04-23].

system na čipu (integrovaný obvod, který zahrnuje všechny součásti počítače do jediného čipu) od společnosti PrimeSense, skutečným původním tvůrcem technologie.

Způsob, kterým Kinect docílí třidimenzionálního zobrazení je realizován pomocí infračervené kamery, obsahující projektor, který vysílá mřížku o rozlišení 640x480 infračervených paprsků rychlostí 30 Hz. Tato mřížka je kuželovitě promítána do prostoru a odráží se zpět od objektů v prostředí. Kinect následně pomocí infračerveného detektoru vypočítá dobu, za kterou se jednotlivé paprsky vrátí zpět do zařízení.

Čip pak díky těmto naměřeným časovým hodnotám změří vzdálenost každého bodu od zařízení a vrací hostitelskému počítači trojrozměrný vektor, složený z datových bodů v rozmezí od 40 do 2000. Tyto datové body přímo odpovídají počtu milimetrů, o které je daná osoba nebo objekt vzdálen od zařízení.¹⁰

Performance driven animace

Místo klasického snímání pomocí reflexních markerů, využívá tedy metoda Weise a jeho týmu hloubkovou kameru Kinectu, která zaznamenává tvar obličeje jako 3D point cloud (shluk bodů v 3D prostoru) v každém nahraném snímku.

Nahraná hloubková data jsou nejdříve vyhlazena vysoko frekvenčním noise filtrem, jelikož technologie nebyla primárně zamýšlena ke snímání obličeje, tudíž nedisponuje příliš velkým rozlišením (méně než 10 % z celkové oblasti snímání), následně porovnávána s předem připraveným blendshape modelem.

Cílem systému je pak v každém snímku najít takový vektor vah blendshapu, který co nejlépe odpovídá výrazu zachycenému v hloubkových datech Kinectu.¹¹ K tomuto procesu je využita optimalizační metoda nazývaná „maximum a posteriori“ (MAP), která obecně odhaduje neznámou kvantitu na základě pozorování dosavadních výstupních dat.¹²

¹⁰ BLAIR, T. B., DAVIS, CH. E., *Innovative Engineering Outreach: A Special Applications of the Xbox 360 Kinect Sensor*. Version 1. Online, dataset. In: IEEE Frontiers in Education Conference. 19. prosince 2013. Dostupné z: <https://ieeexplore.ieee.org/document/6685036>. [cit. 2025-04-23].

¹¹ WEISE, T., BOUAZIZ, S., LI, H., PAULY, M. *Realtime Performance-Based Facial Animation*. Version 1. Online, dataset. In: ACM Transaction on Graphics. 25. července 2011. Dostupné z: <https://doi.org/10.1145/2010324.1964972>. [cit. 2025-04-23].

¹² BASSETT, R., DERIDE, J., *Maximum a Posteriori Estimators as a Limit of Bayes Estimators*. Version 2. Online, dataset. In: arXiv. 21. ledna 2018. str. 2, Dostupné z: <https://doi.org/10.48550/arXiv.1611.05917>. [cit. 2025-04-23].

Tato metoda hledá takovou konfiguraci vah, která současně: co nejpřesněji přibližuje aktuálně nasnímaný tvar obličeje a zároveň odpovídá přirozenému výrazu podle statistického modelu, který byl naučen z reálných blendshape animací.

Klíčovou vlastností dobrého pravděpodobnostního trackovacího algoritmu je množství tréninkových dat, která jsou offline nahrána před započítím performance – driven animace, pro zrekonstruování vhodných blendshapeů. Čím více těchto dat je, tím snáze se trackovaný model shoduje s obličejovými výrazy prováděnými uživatelem.¹³

¹³ WEISE, T., BOUAZIZ, S., LI, H., PAULY, M. *Realtime Performance-Based Facial Animation*. Version 1. Online, dataset. In: ACM Transaction on Graphics. 25. července 2011. Dostupné z: <https://doi.org/10.1145/2010324.1964972>. [cit. 2025-04-23].

2 SOFTWARE K ZÁZNAMU FACIAL CAPTURE

Pokud se zaměříme na software, který je schopný uspokojivého výsledku, existuje velmi mnoho nástrojů, které se v průběhu let objevili na poli nejen profesionálních produkcí, ale s rozvojem AI a výpočetního výkonu moderních počítačů, jsou jejich dostupnější deriváty stále častěji k dispozici i amatérským tvůrcům. Obecně ale stále platí, že rozhodujícím faktorem kvality výstupu animace obličeje je stále z velké části hardware, který je velice drahou komoditou i pro velká studia, natož pro nezávislé produkce, nehledě na odborné znalosti k obsluze. Opravdové „filmové“ kvality CGI tudíž dosáhne ve skutečnosti pouze silně zredukovaný okruh studií.

Pro nastínění cen profesionálního vybavení, využijí příklad americké firmy FaceWare, která se již několik let považuje za špičku v oblasti snímání obličeje a využívají ji videoherní společnosti jako RockstarGames, EA¹⁴ atd. pro nespornou kvalitu výstupu.

Cenovka celého setu plného příslušenství včetně head-mounted kamer pro dva performery a přidruženého softwaru se ovšem může vyšplhat až na 33 500 \$¹⁵, což pro velká studia není nejspíš nijak závratná částka vzhledem k funkcím a možnostem daného zařízení, nicméně je to laťka neodmyslitelně často oddělující high-end výstupy od amatérských pokusů o totožnou věc.

Na druhou stranu s postupným zlepšováním umělé inteligence je velmi pravděpodobné, že se mezera mezi těmito oddělenými tábory bude postupně zmenšovat a celkově posouvat kvalitu obou exponenciálně nahoru. Skvělým dnešním příkladem je plugin do iClone (uživatelsky přívětivý real-time 3D animační software od Reallusion)¹⁶ jménem AccuFACE, což je AI obličejový tracker poháněný grafickými kartami NVIDIA RTX, umožňující záznamu animace buď v reálném čase nebo po nahrání klasických video souborů.¹⁷ Uživatel pak může následně upravovat sílu interpretace pohybu jednotlivých částí, pro lepší kalibraci a uvěřitelnost CGI charakteru, vyhladit vysokofrekvenční šum a jiné úpravy.

Obstrukce pro tvorbu tohoto složitého typu animace jsou tedy spíš už historií a vzhledem k možnostem, které máme k dispozici společně s moderním výpočetním výkonem PC komponentů nebylo vytvořit decentní výstup animace nikdy snazší. Pro účely mé práce

¹⁴ FACEWARE. Online. [2025]. Dostupné z: <https://facewaretech.com/pricing>. [cit. 2025-04-24].

¹⁵ FACEWARE. *Pricing*. Online. [2025]. Dostupné z: <https://facewaretech.com/pricing>. [cit. 2025-04-24].

¹⁶ REALLUSION. *iClone 8, Dramatically simplify character animation*. Online. [2025]. Dostupné z: <https://www.reallusion.com/iclone/>. [cit. 2025-04-24].

¹⁷ REALLUSION. *AccuFACE, Face Tracking from Live or Recorded Video*. Online. [2025]. Dostupné z: <https://www.reallusion.com/iclone/motion-capture/accuface.html>. [cit. 2025-04-24].

jsem zvolil dva softwary: Rokoko a Unreal Engine 5, oba nabízející interní možnost vytvoření pomocí facial capture animaci tváře uživatele v kvalitě, kterou hodlám porovnat.

2.1 Rokoko

Dánská firma založená v březnu roku 2014 Jakobem Balsevem je svou nabídkou cenově dostupných gyroskopických MoCap obleků hojně využívána zejména u indie tvůrců a malých herních studií k nahrání animací pro CGI charaktery. Termínem cenově dostupný se myslí cena v porovnání s ostatními mocap systémy jako například Vicon, které jsou několikanásobně dražší a mnohem náročnější na kalibraci, prostory a obsluhu. Jako první produkt byl v roce představen tzv. Smartsuit Pro, oblek vyroben z odolného nylonového materiálu, skládající se z 19 senzorů vestavěnými přímo do obleku, měřící úhlovou rychlost nebo změny v úhlové pozici ve třech dimenzích, kterým se říká „roll, pitch a yaw”¹⁸ a které umožňují rekonstrukci pohybů člověka v obleku v reálném čase. Oblek je vybaven nastavitelnými popruhy, aby vyhovoval různým typům postav, a všechny elektronické součástky jsou chráněny gumovými ochrannými kabely integrovanými na povrch obleku, což umožňuje volný pohyb herce.¹⁹

Pro sledování pohybu v reálném čase je využíván software Smartsuit Studio, který obsahuje uživatelsky přívětivé grafické rozhraní, snadné propojení obleku přes LAN místní síť a možnost nahrát a exportovat animace do jiných 3D softwarů. Firma také vyrábí tzv. „Smartgloves“, rukavice, které se prodávají jako samostatný produkt a umožňují motion capture jednotlivých prstů.²⁰

Posledním nabízeným produktem této firmy je odvětví face capture neboli záznam výrazů obličeje, ke kterému firma nabízí několik věcí jako je „headrig“, což je speciálně zkonstruovaná helma s připevněným nosičem kamery (v tomto případě iPhonu), který je fixně nasměrovaný proti obličeji herce, zajišťující plynulé snímání tváře aktéra při otáčení hlavy či jiných akcích vyžadujících rychlý pohyb.

Rokoko navíc nabízí svou vlastní aplikaci „Rokoko Remote“, která je schopna přenášet přes LAN rozhraní real-time face capture lidské tváře. Aplikace je prozatím pouze na iOS

¹⁸ Pitch, yaw, and roll. Online. In: Wikipedia. Stránka byla naposledy editována 7.4.2023 v 5:52. Dostupné z: https://simple.wikipedia.org/wiki/Pitch,_yaw,_and_roll [cit. 2024-01-26]

¹⁹ MIHCIN, S. et al., *Investigation of Wearable Motion Capture System Towards Biomechanical Modelling*. Version 1. Online, dataset. In: IEEE Xplore. 28. červen 2019, str. 2. Licence: CC BY 4.0. Dostupné z: <https://doi.org/10.1109/MeMeA.2019.8802208> [cit. 2024-01-26]

²⁰ ROKOKO. *Capture accurate finger motions with Smartgloves*. Online. Dostupné z: <https://www.rokoko.com/products/smartgloves> [cit. 2024-01-26]

(mobily obsahující přední TrueDepth kameru, iPhone X a výše) a je založená na Apple ARkit technologii, podporující nahrávání při 60 snímcích za sekundu včetně 52 možných výrazů tváře, předdefinovaných již firmou Apple. Výstupní data jsou přenášena do Rokoko Studio, samostatné pracovní stanice, ze které mohou být dále exportována do dalších softwarů pro digitální tvorbu obsahu ve formátech FBX, BVH nebo CSV. Dostupná je i možnost živého vysílání těchto dat přes integrované pluginy do Unity nebo třeba do již zmiňovaného Unreal Engineu.²¹

2.2 Unreal Engine 5

Je nutné na začátku zmínit, že jak už název napovídá, nejedná se o nový software, nýbrž již pátou verzi v pořadí vydanou v dubnu roku 2022 společností Epic Games. Unreal Engine taktéž známý pod zkratkou UE, je nástroj primárně zaměřený na vývoj videoher. S pomocí tohoto engineu mají vývojáři možnost nejen vytvářet videohry všech možných žánrů, ale slouží také pro architektonickou a automobilovou vizualizaci, tvorbu lineárního filmového a televizního obsahu, může skrz něj být řízena produkce živých vysílání nebo pro simulace různých fyzikálních jevů.²²

Od čtvrté verze umožňuje engine uživatelům a vývojářům vytvářet interaktivní prostředí bez psaní jediného řádku kódu, pomocí „Blueprint Visual Scripting“ systému založeného na konceptu využívání nodálního rozhraní pro vytváření různých elementů přímo v Unreal Editoru. Toto byl obrovský krok vpřed v oblasti vývoje her. Úkol, který by vyžadoval spolupráci artisty, designéra a programátora, mohl nyní být proveden kýmkoliv z těchto tří lidí, všichni za použití stejného nástroje. UE4 přišel také s komunitním Unreal Marketplace, platformou, umožňující tvůrcům prodávat tzv. assety, digitální 3D modely, které byly buď zadarmo, nebo za peníze nabízeny návštěvníkům obchodu a po koupi mohli být vloženy do vlastního projektu. Toto místo je nyní velká studnice nejen assetů ale i třeba oficiálních pluginů, skriptů a mnoha dalšího.

Společnost Epic Games v posledních několika letech rozšiřuje působnost UE5 nejen na pole videoherního průmyslu, ale také na filmový trh. V podobě virtuální produkce, což je technologie, při níž jsou velké LED panely používány jako pozadí pro scénu, na kterých je

²¹ THACKER, J. Rokoko launches facial capture add-on for its mocap systems. In: *CGchannel.com* [online]. 12. července 2019. Dostupné z: <https://www.cgchannel.com/2019/07/roko-launches-facial-capture-add-on-for-its-mocap-systems/> [cit. 2024-01-26]

²² EPIC GAMES, Inc. Frequently asked questions (FAQs). In: *UnrealEngine.com* [online]. 2004-2024. Dostupné z: <https://www.unrealengine.com/en-US/faq> [cit. 2024-1-26]

následně zobrazeno video nebo počítačová grafika reagující na pohyb kamery v reálném čase, se jako první proslavil seriál *Mandalorian* režiséra Jona Favrea. ²³

Ten spolupracoval s Industrial Light & Magic na vývoji technologií pro lepší vizualizaci záběrů v rámci CGI prostoru. Výhoda této technologie spočívá v tom, že se herci, na rozdíl od natáčení před zeleným klíčovacím pozadím, mnohem lépe orientují v prostoru a vidí scénu tak jak bude ve filmu. V postprodukci je následně mnohem méně práce, a díky nativnímu světlu z panelů umožňuje technologie lépe simulovat přirozené osvětlení, odrazy a stíny ve scéně. Celý tento proces běžel na Unreal Enginu 5. ²⁴

Hlavní výhodou tohoto softwaru je nejen to, že je zdarma, je také zároveň uživatelsky přívětivý pro začátečníky. ¹⁰ To se může zdát jako marginální věc, nicméně společnost Epic Games touto strategií úspěšně dokazuje, že tyto dva atributy dokážou přivést do herního průmyslu velké množství nových uživatelů, kteří buď engine zkusí a pak od něj upustí, nebo u něj zůstanou a začnou produkovat. Jako důvodem, proč momentálně Unreal Engine patří mezi nejdiskutovanější softwary na poli 3D industry, považují inovátorské novinky, kterými nejen ulehčuje velký kus práce vývojářům, zároveň je to způsob, jak sám sebe promovat mezi širokou veřejností. Schopnost uchvátit novými technologiemi je bezesporu jednou z nedílných vlastností tohoto softwaru.

LiveLink

Účelem Live Link je poskytnout společné rozhraní pro streamování a přijímání animačních dat z externích zdrojů do Unreal Enginu. Je navržen tak, že umožňuje třetím stranám možnost vyvíjet nové funkce pomocí Unreal pluginů, bez permanentní nutnosti zasahovat do samotného zdrojového kódu UE. S nástroji jako například Maya nebo Motionbuilder, umožňuje Live Link externě upravovat animace a zároveň sledovat náhled práce uživatele v reálném čase přímo v UE. Systémy jako Shogun, vytvořené na zaznamenání motion capture dat tak mohou pomocí tohoto propojení streamovat data do UE a nahrávat tak živý performance herce přímo do enginu, kde se zobrazuje jako CGI charakter.

²³ JENSEN, K. T. 25 Years Later: The History of Unreal and an Epic Dynasty. In: *PCMag.com* [online]. 22. květen 2023, Dostupné z: <https://www.pcmag.com/news/25-years-later-the-history-of-unreal-and-an-epic-dynasty> [cit. 2024-01-26].

²⁴ DESOWITZ, B. The Mandalorian: How ILM's Innovative StageCraft Tech Created a 'Star Wars' Virtual Universe. In: *IndieWire.com* [online]. 20 února 2020 17:01. Dostupné z: <https://www.indiewire.com/features/general/the-mandalorian-ilm-stagecraft-1202212389/> [cit. 2024-01-26].

II PRAKTICKÁ ČÁST

3 ZADÁNÍ EXPERIMENTU

Vzhledem k mé praktické bakalářské práci, kde hodlám natočit krátký film za pomoci technologie motion capture, stála přede mnou už na začátku velmi stěžejní otázka a to zjistit, jakým způsobem zajistím, aby v mém filmu nebyly jen chodící těla ale opravdové postavy s tvářmi, schopnou odvyprávět nějaký příběh, zkrátka to, co běžný člověk považuje u filmu za samozřejmost. Dialog mezi herci, vyjádření nesouhlasu, vzdoru, či radosti, jednoduše vše, co nalézáme v normálních filmech bez digitálních charakterů. Když jsem v březnu minulého roku ovšem uviděl technickou demonstraci nového pluginu do Unreal Engine, který je schopen bez větších obtíží velmi kvalitně zaznamenat i jemnou mimiku obličeje pouze pomocí iPhone, okamžitě mě napadlo, kolik možností budou mít všichni tvůrci až plugin Epic Games vydají. Jak už je v poslední době pro Epic Games zvykem, obdařují nejen profesionální studia, ale hlavně malé nezávislé tvůrce a vývojáře nástroji, kterými ještě před pár lety disponovali pouze ti největší hráči na trhu. Na stranu druhou, již nějakou dobu slyším o softwaru s názvem Rokoko, který nejen že je primárně zaměřen na motion capture, dokáže ale i přes mobilní aplikaci Rokoko Remote zvládat i facial capture, kterou pak lze živě přenášet do dalších 3D softwarů, kde jí lze použít například na CGI charakter, kterého pak můžeme okamžitě vidět hýbat obličejem, a lépe si tím třeba vizualizovat celou scénu.

Cíl experimentu

Cílem zadaného experimentu je zjistit, jaký z vybraných softwarů je přesnější na měření face capture, tj. záznam výrazů tváře reálného člověka aplikovaný na digitální postavu uvnitř softwaru, zda to bude zatím experimentální plugin do Unreal Engine s názvem „Metahuman Animator“ nebo již nějakou dobu zasetý software Rokoko Studio využívající živého přenosu dat na vybraný obličej 3D postavy.

Hypotéza

Mám za to, že zcela odlišná technologie MetaHuman Animatoru, využívající mimo jiné hloubková data, bude vyspělejší a přesnější než záznam z Rokoko Remote, přenášený živě do Unreal Engine, omezený již defaultně počtem blendshapes.

3.1 Výzkumné otázky

Cílem práce je ověřit tvrzení, že streamování dat pomocí Rokoko Remote aplikace využívající klasický ARkit, je méně přesné oproti Metahuman Animatorem zpracovanému videu.

Jak moc jsou oba systémy náchylné k chybám nebo odchylkám při zachycování pohybu na tváři a jak se s těmito situacemi vyrovnávají?

Způsob měření

Rozhodl jsem se pro metodické měření.

Technika měření je tracking markerů na rtech a následné zobrazení jejich globálních souřadnic pomocí grafů. Z nich vzniknou křivky, které nám odpoví na výzkumné otázky.

Důvod měření

Měření jsem se rozhodl udělat kvůli obecně známé informaci, že hláska „m“ může mít u facial motion capture systémů vyšší chybovost, projevující se při sledování pohybů obličejových svalů a tvorby pohybů rtů v závěrečném výstupu během výslovnosti této nosové souhlásky. Při výslovnosti této konsonanty jsou totiž na malou chvíli ústa zavřená, což může mít za následek chybné zaznamenání třeba obličejových markerů, u kterých může být obtížnější je umístit tak, aby byly viditelné a poskytovaly přesné informace. Velkou roli také může hrát produkce souhlásky „m“, která vzniká přechodným vytvořením překážky neboli okluze, která chvilkově brání proudění vzduchu, zapříčiněno neprůchodným spojením horního a dolního rtu. Tento jemný a rychlý pohyb vyžaduje výkonný a citlivý systém, schopný reprodukovat pak na obličej CGI charakteru tuto subtilní souhlásku.

Postup měření

Pro správné měření odchylky musím nejdříve vytvořit můj obličej ve 3D, procesem, který může proběhnout několika způsoby. Fotogrammetrií, což, jak už název naznačuje, je technika měření tří rozměrů využívající fotografie jako základní médium pro metrologii nebo měření. Základním principem, který je ve fotogrammetrii využíván, je triangulace. Pořízením fotografií z alespoň dvou různých míst lze vytvořit tzv. „směry pohledu“ z každé kamery k bodům na objektu. Tyto směry pohledu, někdy nazývané paprsky kvůli své optické povaze, jsou matematicky průsečíkem, což vede k vytvoření trojrozměrných souřadnic bodů

zájmu. Jednotlivé obrázky pak vytvoří digitální model reálného světa.²⁵ V tomto případě mají hlavy.

Druhým způsobem je sken pomocí TrueDepth přední kamery iPhoneu za pomoci LiveLink aplikace, ze které se pak natočený materiál importuje propojením přes IP adresu společné sdílené domácí sítě LAN (Local Area Network) do Unreal Engine. TrueDepth systém umožňuje autentizaci uživatele pomocí rozpoznávání obličeje s Face ID a je implementován od generace iPhone X. Přesné registrování obličejových dat probíhá prostřednictvím projekce 30 000 infračervených bodů, které jsou zaznamenávány 60krát za sekundu infračervenou kamerou a zpracovávány neuronovým enginem čipu smartphonu. „Infrared flood illuminator“ přidává infračervené světlo umožňující měření i ve tmě. Z naměřených dat je vypočítána 3D hloubková mapa obličejové struktury, která je spolu s infračerveným obrazem převedena na matematickou reprezentaci obličejových rysů uživatele.²⁶

Pokusím se tedy o několik způsobů, ze kterých následně vyhodnotím ten nejvhodnější.

²⁵ SIEGEL, Jay A.; SAUKKO, Pekka a HOUCK, Max M. (ed.), [2013]. *Encyclopedia of forensic sciences. Second edition*. Amsterdam: Elsevier. ISBN 978-0-12-382165-2, str. 368 [cit. 2024-01-27]

²⁶ TAEGER J., BISCHOFF S., HAGEN R., RAK K. *Utilization of Smartphone Depth Mapping Cameras for App-Based Grading of Facial Movement Disorders: Development and Feasibility Study*. Version 9. Online, dataset. In: JMIR Publications 2021. Dostupné z: <https://mhealth.jmir.org/2021/1/e19346>. [3,37 MB]. [cit. 2024-01-27]

III PROJEKTOVÁ ČÁST

4 EXPERIMENT

4.1 Reality capture

Pro mé podmínky zkusím nejdříve získání mého 3D obličeje pro experiment pomocí Reality Capture. Umožní mi totiž vytvořit přesnou geometrii mého obličeje, díky mnoha kamerovým úhlům, které jsou nutné pokrýt jak předek hlavy, tak boky a týl. Reality Capture provede rekonstrukci všech nahraných fotografií mého obličeje, které byly vyfoceny po orbitální dráze mé hlavy, s pouze proměňující se výškou. Focení by proběhlo na iPhone 14 Pro s 48MP hlavní kamerou s ohniskovou vzdáleností 24 mm a clonou f/1.78. Menší clona u tohoto objektivu by znamenala malou hloubku ostrosti, a tudíž i pravděpodobné jemné rozostření pozadí, kdy by měl fokus zůstat pouze na obličej. Clona s nižším číslem propouští více světla, což bude užitečné při fotografování v místnosti, kde je absence přirozeného světla a musím se spolehnout pouze na umělé zdroje svícení.²⁷ Efekt „bokeh“ jak se také tento jev rozmazaného pozadí nazývá, může snížit vizuální rušení od pozadí a soustředit se na hlavní objekt. To může usnadnit algoritmům Reality Capture identifikaci a sledování relevantních prvků obličeje. Dalším aspektem správného skenu je světlo, které musí být rovnoměrně rozmístěno tak, aby na objektu nebyly stíny, které by pak mohli být součástí výsledné textury.

Samotné nanesení tracking markerů bude provedeno permanentkovou fixou Centropen, která není úplně šetrná k lidské kůži, na druhou stranu se marker nerozpije či jinak nezdeformuje a poskytne spolehlivou šablonu k měření.

Mobil, který používám jako kameru, je sice připevněný na stativu, ale důvodem je hlavně to, že stativ drží nejen iPhony ale i ring light (bílý světelný prstenec), který poskytuje skvělé světelné podmínky pro experiment.

Pro správné zachycení hereckého výkonu člověka, který má na obličej tyto černé tečky je potřeba vyšší snímková frekvence (48+ fps), která umožní detekovat i jemné pohyby tváře.²⁸ LiveLink aplikace, přes kterou můj herecký výkon, v podobě věty s opakující se souhláskou „m“, budeme natáčet, umožňuje snímání v 60 snímcích za sekundu, to znamená plynulé zachycení i případných detailů.

²⁷ LUBINA, I. Jak nastavit clonu. 7 tipů pro nastavení clony. In: *Fotografvna.cz* [online] 10. listopadu 2021. Dostupné z: <https://fotografvna.cz/jak-nastavit-clonu-7-tipu-pro-nastaveni-clony/> [cit. 2024-01-27]

²⁸ OKUN, J. A., ZWERMAN, S. *The VES Handbook of Visual Effects: Industry Standard VFX Practices and Procedures. 1st Edition.* Focal Press/Elsevier, 2010. str. 374

Po úspěšném skenu mojí tváře je potřeba fotografie přeměnit na mesh, neboli digitální trojrozměrný model obličejové struktury, obsahující informace o geometrii, textuře atd. K tomuto úkonu je vhodný již zmíněný software Reality Capture, původně slovenský startup společnosti Capturing Reality, který již zezáátku nabízel velkou škálu formátů, ze kterých dokáže 3D objekt generovat, včetně skenů pomocí laseru. Software je schopný zacházet velkými daty, akceptuje neomezený počet zdrojových obrázků s využitím více dostupných CPU a GPU jader ke zpracování dat²⁹. V březnu 2021 spadla společnost akvizicí pod křídla Epic Games, který ale nijak neovlivňuje existující obchodní činnost Capturing Reality ba naopak, snížil cenovou politiku pro RealityCapture tak, že je nyní přístupná i pro normální uživatele.³⁰

Po otevření programu klikneme na tlačítko „Inputs“, kde následně vybereme veškeré fotografie pořízené na sken. Důležité je, abychom fotografie importovali v surovém neupraveném stavu, bez korekcí barev či ořezu, protože kalkulace jsou kriticky založeny na originálním vztahu mezi pixely, středem objektivu atd.³¹

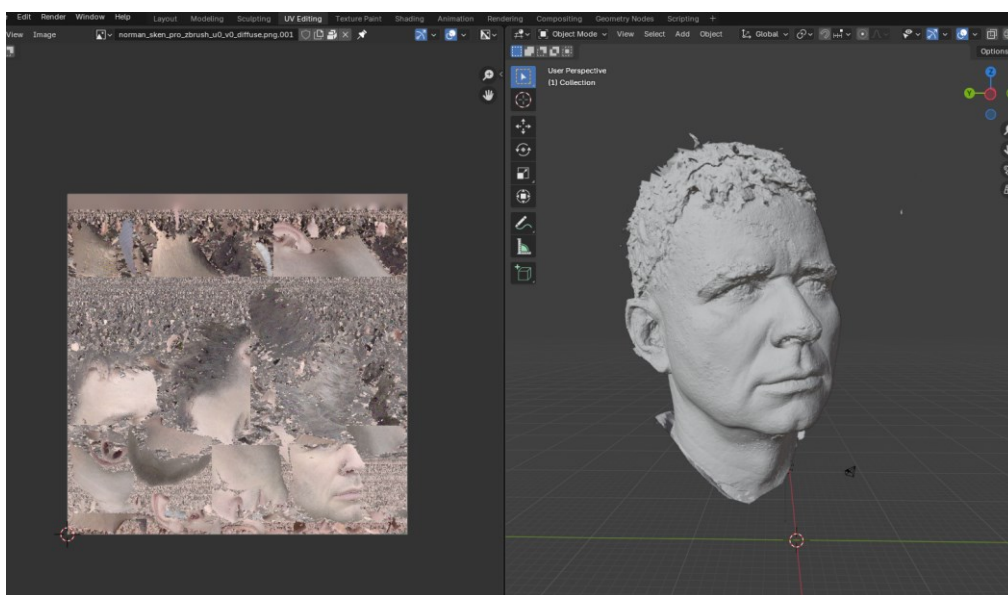
Následně pomocí tlačítka „Align Images“ software synchronizuje na pozadí tyto snímky v jednotném prostoru. Dále stačí už jen spustit proces generování 3D meshe, který trvá několik minut v závislosti na výkonu počítače, počtu fotografií a jejich kvalitě, rozlišení atd. Dalším krokem je „refining“ neboli oprava nevyžádané geometrie, vzniklé chybnou kalkulací fotografické návaznosti, či nedostatku informací v dané lokace předmětu. To se nejčastěji děje na temenu hlavy nebo v oblasti krku kde software „nemá z čeho brát“. Mým úkolem by bylo tedy označit hybné části špičaté vyčnívající geometrie a následně ji zahladit, nebo vyplnit.

²⁹ THACKER, J. Capturing Reality releases RealityCapture. In: *CGchannel.com* [online]. 4. února 2016. Dostupné z: <https://www.cgchannel.com/2019/07/rokoko-launches-facial-capture-add-on-for-its-mocap-systems/> [cit. 2024-01-26]

³⁰ KERR, C. Epic Games acquires photogrammetry software developer CapturingReality. In: *Gamedeveloper.com* [online] 9. března 2021. Dostupné z: <https://www.gamedeveloper.com/business/epic-games-acquires-photogrammetry-software-developer-capturing-reality> [cit. 2024-01-26].

³¹ Real Human to Metahuman Photogrammetry Workflow. In: Youtube [online]. 12. června 2022. Dostupné z: <https://www.youtube.com/watch?v=8JVWGb9kBng&t=798s> [cit. 2024-01-25]

Tento způsob získání finálního 3D obličeje je při správných světelných podmínkách velmi efektivní a přesný, byť pro náš výzkum postrádá jednu důležitou vlastnost a tou je textura. Ta sice je obsažena při exportu do Unreal Engine, jelikož se ale jedná o sken, a ne ručně vymodelovaný asset, je UV mapa skenu (UV mapa je dvourozměrná reprezentace textury, kde osy U a V představují horizontální a vertikální směr textury)³² hustý shluk takzvaných „vertices“. Tento termín definujeme jako vrchol bodu nebo bodů v trojrozměrném prostoru, které tvoří základní stavební prvky pro vytváření 3D modelů. Vrcholy jsou definovány svými souřadnicemi (x,y,z) a jsou spojeny hranami, které vytvářejí síť a tvoří povrch objektu. Právě tato síť je ovšem na skenovaném objektu velice hustá a nepravidelná, vyžadující mnoho úprav a zjednodušení celkové geometrie k tomu, aby textura byla reprodukovatelná i na jiné modely. Po aplikování na vytvořeného MetaHumana, jehož základní textura je specificky daná a přesně rozměrově fungující na vytvořený mesh, je textura fungující pouze na sken nefunkční a ztrácí význam.

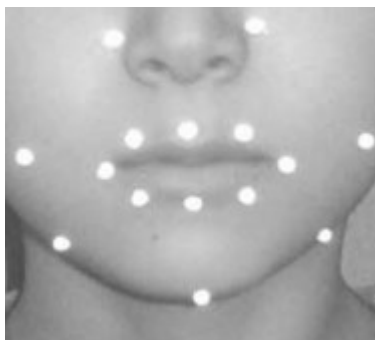


Obr. 1, Naskenovaný obličej postavy Normana z mého bakalářského filmu pomocí fotogrammetrie, vedle je UV mapa textury

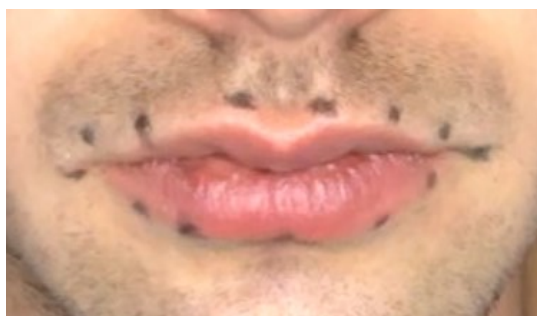
³² OKUN, J. A., ZWERMAN, S. *The VES Handbook of Visual Effects: Industry Standard VFX Practices and Procedures. 1st Edition.* Focal Press/Elsevier, 2010. ISBN 978-0240812427, str. 676

4.2 Trackovací body na rty

Aplikace trackovacích bodů neboli markerů byla provedena podle vzorového obrázku v knize *Mocap for artists*³³, kde jsem ve své variantě přidal ještě několik bodů navíc. Rád bych zdůraznil, že hlavním důvodem aplikace tracking pointů v praxi je zajištění dat pro „joints“, ovladatelné vygenerované kůstky, které jsou svázány obličejovou geometrií. Po aplikaci trackovacích bodů na stejná místa na tváři, kde jsou na digitální postavě umístěny joints, vzniká cenná informace o pohybu přesně těchto lokalit a lze tím docílit přirozeného a autentického pohybu tváře počítačové kopie herce.³⁴ Já jsem ale, kvůli zcela jiné technice snímání obličeje, o tuto metodu osvobozen (všechny použité aplikace využívají TrueDepth kameru k záznamu pohybu obličeje) tudíž černé tečky okolo rtů plní pouze funkci lokalizování vybraných bodů napříč výslednými rendery.



Obr. 2, Předloha k vytvoření mých markerů



Obr. 3, Mnou provedené omarkování rtů

V předchozí části jsem vysvětloval, že fotogrammetrický sken je bezesporu jedna z nejlepších možností, pro náš cíl je ale značně zdlouhavá a náročná, vyžadující velmi dobré světelné podmínky, ideálně solidní DLSR zrcadlovku a prostory, ve kterých je možné tuto metodu provést, nehledě na to, že i při dobrém výsledku obsahuje model nepoužitelné textury, který by bylo třeba unwrap metodou správně lokalizovat. Textura mého obličeje je pro toto měření ale stěžejní, protože právě ona nám umožní projekci mé kůže na pokusnou Metahuman postavu.

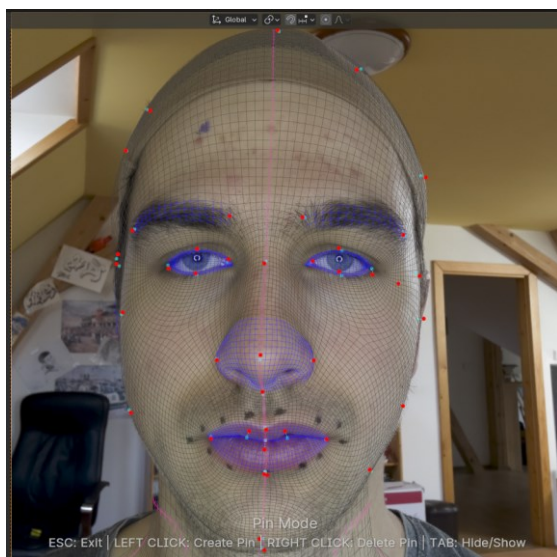
³³ KITAGAWA, Midori. *MoCap for Artists: Workflow and Techniques for Motion Capture*. Routledge, 2008. ISBN 10 0240810007. str. 160 [cit. 2024-06-12]

³⁴ Tamtéž, str. 155 [cit. 2024-06-12]

4.3 KeenTools plugin

Rozhodl jsem se tedy využít add-on do Blenderu, což je open-source 3D software, který nabízí velikou škálu možností včetně aplikování textur, modelování, renderování atd., který se jmenuje FaceBuilder for Blender a umožňuje vytvoření 3D obličeje za pomoci několika fotek³⁵. Služba je pro běžné používání placená, nabízí ale zkušební období zdarma, a tak jsem měl možnost vyzkoušet tento plugin na pár dní, kdy jsem stihl exportovat vše potřebné pro další kroky.

Principiálně funguje tento add-on na bázi takzvaného „pinování“ neboli přišpendlení základní obličejové geometrie dodané samotným addonem na obličejové rysy člověka na vložené fotografii obličeje.



Obr. 4, Funkce pluginu FaceBuilder



Obr. 5, Rotace obličeje pod jiným úhlem

Uživatel se pak po dobu procesu „výroby“ vlastní hlavy snaží co nejlépe ladit virtuálními špendlíky hrany hlavy, velikost nosu a rtů, koutky očí atd. Software je schopný také tzv. „align“ tlačítkem udělat kus práce za uživatele, přičemž trefí lépe i proporčnost hlavy a rysů, detailní úpravy však uživatel musím doladit sám. Jelikož program pracuje také s ohniskovou vzdáleností, je lepší místo přednastaveného odhadnutí této veličiny, napsat skutečnou hodnotu. V mém případě jsem fotil dvě referenční fotky na iPhone 14 Pro, který má na přední kameře ohniskovou vzdálenost objektivu 23 mm. Vyfotit referenční fotografie, které budou určeny nejen na model mé hlavy a tváře, ale i na textury by bylo samozřejmě

³⁵ KEENTOOLS.IO. *KeenTools add-ons for Blender*. Online. [2023]. Dostupné z: <https://keentools.io/download/facebuilder-for-blender> [cit. 2024-01-26]

mnohem lepší zachytit na zadní kameru, která je kvalitnější a má víc megapixelů, chtěl jsem ale zachovat stejnou ohniskovou vzdálenost, kterou používají i aplikace jako Live Link Face a Rokoko Remote, které budou více zmíněny až později.

Hlavním důvodem, proč jsem ale ze všech možných způsobů replikace mojí tváře do virtuálního 3D prostředí zvolil právě tento, je možnost vytvoření textury, která nabízí stejnou UV mapu, jako základní šedá textura u Metahuman postavy.

Díky této vlastnosti lze pak ve Photoshopu vcelku jednoduše, pomocí změny průhlednosti, zarovnat texturu na „šablonovou“ z Metahuman. Ta jde snadno exportovat z Unreal Engine v podobě PNG formátu. Rty u našeho měření hrají tu nejzásadnější roli, proto bylo důležité je co nejpřesněji zarovnat na podkladovou vrstvu, přitom zůstat s anchor pointem (bod, kolem kterého se objekt otáčí nebo zvětšuje)³⁶ na středové ose a měnit pouze velikost se zachováním poměru stran.



*Obr. 6, Textura z pluginu FaceBuilder
aplikovaná na základní MetaHuman
texturu*

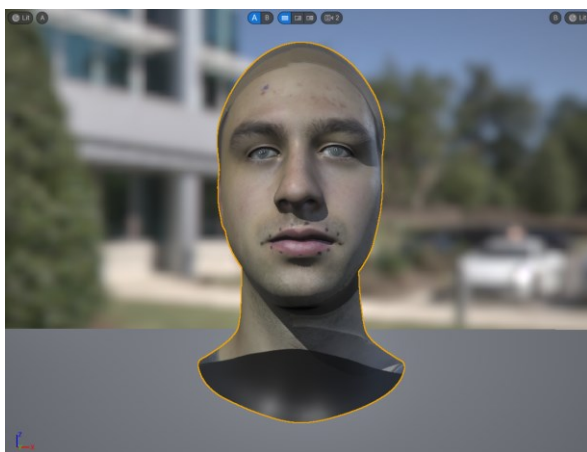
4.4 Metahuman Identity

Jádro našeho experimentu spočívá právě v těchto plně narinovaných, neboli fyziologicky obohacených, fotorealistických digitálních postavách exklusivních výhradně pro Unreal Engine. Za těmito charaktery stojí společnost 3Lateral, kterou v roce 2019 Epic

³⁶ WARD, C. How to Move the Anchor Point in After Effects. In: *Schoolofmotion.com* [online]. Dostupné z: <https://www.schoolofmotion.com/blog/anchor-point-in-after-effects>. [cit. 2024-01-26].

Games odkoupilo³⁷ a její tým pak už pod záštitou tohoto giganta začal pracovat na velmi věrohodných digitálních charakterech, kteří by šli implementovat do Unreal Enginu, zdarma, pro všechny.

Lze je upravovat, měnit rysy, postavu, barvu kůže a mnoho dalšího. Na druhou stranu uživatel do nedávna neměl zas až tak volnou ruku z pohledu modelování tváře, jelikož vždy musela být složena z již přednastavených postav, často etnický velmi rozdílných. V červnu roku 2022 vyšla ale nová funkce, která byla součástí nového Metahuman Pluginu s názvem „Mesh to Metahuman“, kde, jak už název naznačuje, lze nově vytvořit vlastní hlavu pouhým importem oskenovaného, či jinak vytvořeného nehybného meshe. Plugin sám následovně vytvoří, z jednoho zvoleného snímku definovaného uživatelem, track očí, rysů podél tváře a rtů.



Obr. 7, Mesh s aplikovanou texturou, obojí vytvořeno ve FaceBuilderu



Obr. 8, Auto-track obličejových rysů Metahuman Identity solverem

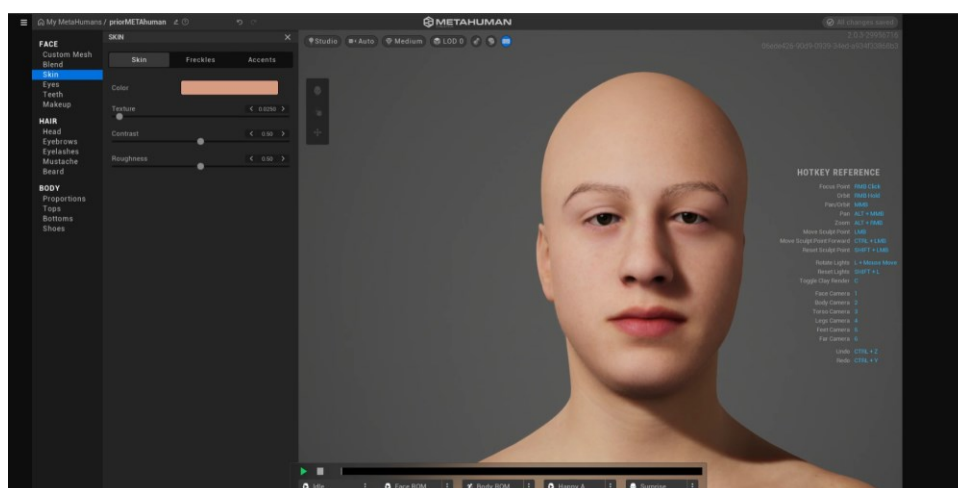
Vše lze pak manuálně dopravit a vylepšit případné neduhy, které software jednoduše nerozpozná. Pomocí pak několika dalších kliknutí na „solve Metahuman identity“ se vytvoří narigovaná a plně animovatelná kopie našeho původního meshe akorát bez textury.

Touto metodou vznikl i můj Metahuman obličej, který ale samostatně lze pouze animovat ručně. To je ale pro náš experiment nedostačující, proto jsem musel v tomto editoru udělat některé kroky k tomu, abych získal požadovaný výsledek. Jako první věc jsem zadal v kolonce „Mesh to Metahuman“ možnost vytvoření nejenom hlavy postavy, ale i celého těla. Důvodem, proč potřebujeme celou postavu i s tělem je to, že pouze u ní se pak

³⁷ MATNEY, L. Epic Games buys 3Lateral, maker of super-realistic ‘digital humans’. In: *Techcrunch.com* [online] 23. ledna 2019. Dostupné z: <https://techcrunch.com/2019/01/23/epic-games-buys-3lateral-maker-of-super-realistic-digital-humans/> [cit. 2024-01-26]

v panelu „Details“ objeví možnost propojení přes LiveLink, stěžejní funkce, o kterou se opírá celé měření. Poslední úpravou v je pak tlačítko „Prepare for Performance“, které zpustí proces adaptování Metahuman hlavy na případné použití v Metahuman Animatoru, který vyšel v červnu roku 2023³⁸ s cílem poskytnout uživatelům vysoce přesné replikace emocí lidské tváře na Metahuman postavy pouze za pomoci pluginu a mobilního telefonu.

Celý proces přípravy na performance trvá řádově několik minut s ohledem na výkon počítače, na kterém je výpočet prováděn. Po všech těchto nastaveních je možné opustit editor a importovat našeho vytvořeného Metahumana do scény. Ten ale není, jak by se očekávalo uložený v projektu, nýbrž na cloudovém úložišti Metahuman Creator, propojeného s uživatelským účtem od Epic Games. Do tohoto editoru se vstupuje pomocí webového prohlížeče a teprve tam lze Metahuman charakter upravovat.



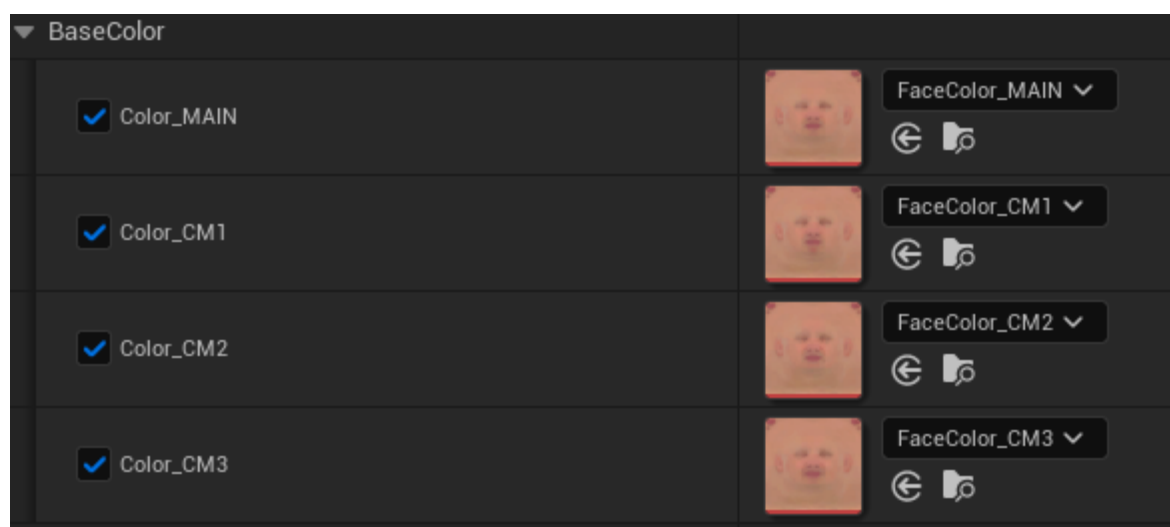
Obr. 9, Vzniklý MetaHuman charakter po aplikaci kůže v MetaHuman Creatoru

Základní postava nemá barvu kůže, vlasy ani oblečení, vše je nutné manuálně nastavit. Toto nastavení je ale prozatímní a ve výsledku nám jde jen o to, aby se postava synchronizovala s Epic Games účtem a byla v nabídce ke stažení v Unreal Engine v kolonce Quixel Bridge/My Metahumans. Z tohoto archivu je také možné stahovat high-poly 3D skeny od společnosti Quixel, exkluzivně pouze pro UE.

³⁸ 3LATERAL. Metahuman Animator Launched. In: *3Lateral.com* [online] 15.června 2023. Dostupné z: <https://www.3lateral.com/news/metahuman-animator-launched.html> [2024-01-26]

Aplikace textury

Po stažení Metahuman postavy, která disponuje uměle vygenerovaným tělem ale mojí hlavou, je nyní již možné provést experiment. Je ale před tím potřeba na hlavu aplikovat správnou texturu, kterou jsem si obstaral pomocí Face Builderu v Blenderu. To provedu rozkliknutím kolonky s materiály, kde v Blueprint Editoru zaměním všechny čtyři základní textury (Color_MAIN, Color_CM1, Color_CM2, Color_CM3) za mojí.



Obr. 10, Výšeč textur, které se musí změnit na mnou vytvořenou texturu s trackovacími body

Podnětem proč zrovna všechny čtyři je to, že po vyzkoušení nahradit pouze texturu Color_MAIN za tu mojí, se objevily prosvítající anomálie při pohybu obličeje. Těmi prosvítajícími texturami byly právě ty zbylé tři, které jsem ponechal. Ty narušovaly viditelnost trackerů, a z této zkušenosti musí proto být změněny všechny.

4.5 Nastavení telefonů

Než začnu inicializovat celý měřicí segment, musím správně připevnit mobilní telefony k sobě. K záznamu věty budu používat zařízení značky Apple a to iPhone 12 Pro a iPhone 14 Pro. Důvodem je vestavěná přední TrueDepth kamera u obou telefonů, kterou aplikace jako Rokoko Remote a LiveLink Face vyžadují ke správnému snímání tváře. Pro dobrou světelnost využiji tzv. ring light, což je obloukové měkké světlo, poskytující správné nasvícení obličeje bez nevyžádaných stínů, které by mohli zhoršit záznam dat. Ring light je připevněn na stativ, který zajišťuje potřebnou stabilitu pro těžké závaží v podobě dvou mobilů (oba telefony dohromady váží okolo 390 gramů), připevněných stahovacími gumičkami k sobě, které drží kovová nastavovací tyčka s plastovým úchopem vyrobeným ale pouze na jeden mobilní telefon. O vyrovnání se stará měkčený papír, který je umístěn mezi těly telefonů, protože jejich zadní fotoaparáty vystupují ven a narušují potřebnou rovinu. Spodní mobil, kterým se stal iPhone 12 Pro je přední kamerou dostatečně vysunut nad svého „utlačovatele“ aby mu vrchní část iPhone 14 Pro nepřekážela v zorném poli.



Obr. 11, Ukázka úchopu obou telefonů pro následující experiment



Obr. 12, Plastový držák svírající jeden z mobilů

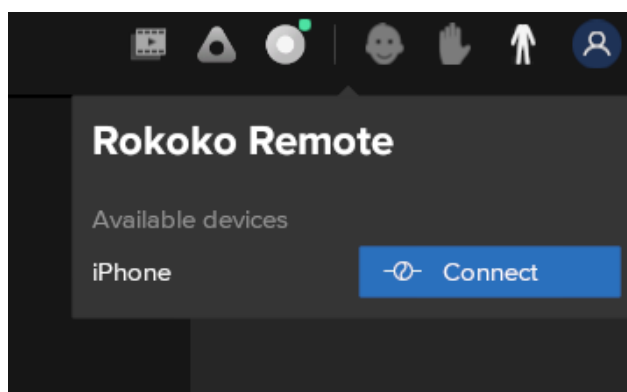
Kvůli nemožné manipulaci a ovládání spodního iPhone jsem musel všechny potřebné akce spojené s tímto zařízením, na rozdíl od druhého, na kterém displej i přes stahující gumy šel stále ovládat, spustit ještě před instalací tohoto provizorního „rigu“ na plastový úchop ring lightu.

5 DEFINICE POUŽITÝCH SOFTWAREŮ

5.1 Rokoko Studio Legacy

Existují dva softwary společnosti Rokoko, jeden s názvem Rokoko Studio a druhý s názvem Rokoko Studio Legacy. Já jsem zvolil k používání tu druhou verzi, která je profesionálnější a jde na ní hlavně spustit aplikace Rokoko Remote, kterou druhý software z nějakého důvodu nerozpozná.

Vše v Rokoku funguje na jakési lehce pofidérní bázi jednoduchosti, která ale mnohdy dělá víc problémů, než kdyby byl software složitější ale přidával více možností kalibrace a nastavení. Propojení mobilní funguje na druhou stranu ve Studio Legacy bezproblémově, stačí spustit aplikaci a být na stejné LAN síti jako počítač, na kterém Studio Legacy běží, pak jen stačí kliknout na ikonku Rokoko Remote a na mobilním zařízení vyskočí okénko, dotazující se, zda se chcete skutečně propojit přes Wi-Fi sít' s vaším počítačem.



Obr. 13, Nabídka ke spárování aplikace Rokoko Remote

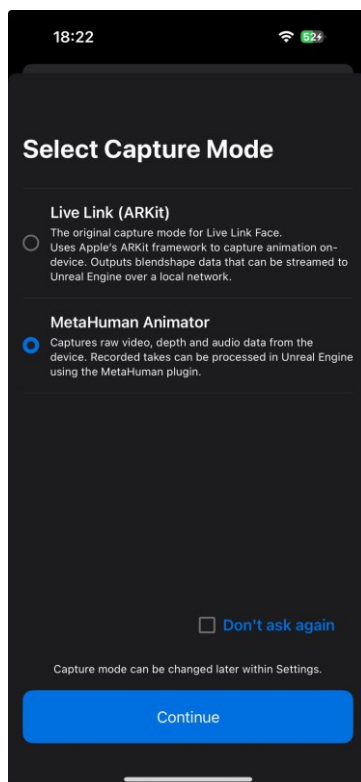


Obr. 14, Uživatelské rozhraní aplikace Rokoko Remote po spárování s PC softwarem

Po souhlasu už máte možnost ve Studiu Legacy vidět hýbající se hlavu, která v reálném čase prezentuje pohyby zaznamenané z telefonu. Pro to abyste mohli ale tyto informace přenášet na svojí Metahuman postavu, je nutné povolit streamování dat do Unrealu, což je jen z mnoha programů, které Rokoko svým živým datovým přenosem podporuje (přenos je možný i do Unity, Maya, Blenderu, MotionBuilderu atd.).

5.2 Metahuman Animator

Pro využití tohoto nástroje je zapotřebí mít v mobilním zařízení aplikaci Live Link Face, která má od června roku 2023 novou možnost nahrávání raw videa, hloubkových a zvukových dat ze zařízení³⁹. Tento červnový update je spojený s vydáním Metahuman Animatoru (dále jen MA), rozšíření stávajícího Metahuman pluginu, umožňující autentické pohyby tváře přenést s vysokou přesností a věrohodností na digitálního dvojníka. Aplikace byla z tohoto důvodu rozšířena z původní podpory pouze ARkitu o nový mód snímání, upravený na míru pro tento plugin. Rozdíl mezi těmito dvěma způsoby je ten, že ARkit nemá ve svých uložených datech depth mapu záznamu, ale používá 3D motion tracking k vypočítání správných ARkit blendshapes tj. okolo 52 přednastavených výrazů společnosti Apple.⁴⁰



Obr. 15, Úvodní nabídka aplikace Live Link Face

³⁹ Výpis historie verzí aplikace na Apple Store, 1.3.0 vyšla 14.6.2023 a přidává podporu pro MetaHuman Animator

⁴⁰ APPLE. Instance Property, blendShapes, A dictionary of named coefficients representing the detected facial expression in terms of the movement of specific facial features. In: *Developer.Apple.com* [online]. Dostupné z: <https://developer.apple.com/documentation/arkit/arfaceanchor/2928251-blendshapes> [cit. 2024-01-27]

5.2.1 Blendshapes

Termín, který zde bude ještě několikrát zmiňován, a proto by bylo dobré si ho definovat. Lze ho tedy nazvat jako podmnožina vertexů (geometrických vrcholů v 3D prostoru), které se posunuly z jejich neutrální pozice.⁴¹ Anglický výraz „blendshapes“ bychom totiž mohli volně přeložit jako „mísící se tvary“, což je přesně to, k čemu slouží.

V případě facial mocapu se jedná o meshe obličeje, obsahující geometrii i texturu, které se pomocí weight sliderů (posuvníky ovládající míru intenzity) mohou přetvořit do jiného obličejového výrazu, kde je dáno rozmezí 0 až 1, např. z neutrálního výrazu na úsměv. Vzhledem k variabilní kombinaci hodnot a výrazů, je možné i z omezené škály blend shapes docílit širokého pokrytí emocí CGI postavy.

Všechny blendshape mesh modely dané obličejové reprezentace sdílejí taktéž shodnou topologii, a proto můžeme zapsat souřadnice vrcholu V následujícím způsobem:

$$V = \sum_{i=1}^n a_i V_i$$

V je výsledná pozice jednoho vrcholu (vertexu) v aktuálním deformačním stavu obličeje. V_i specifikuje pozici téhož vrcholu v blendshapu i (např. úsměv, smutek atd.), a_i reprezentuje váhu, znázorňující, kolik se daný blendshape podílí na finální deformaci (0 až 1) a n udává počet blendshapeů.

Z rovnice pak plyne, že finální pozice jednoho vertexu v síti (mesh) obličeje se spočítá jako vážený průměr jeho pozic ve všech blendshapes, jinými slovy se každý vrchol pohne podle toho, jak moc ho jednotlivé výrazy ovlivňují.

Pro zachování stabilního chování modelu a geometrické konzistence je ovšem na váhy a_i uvaleno konvexní omezení pro všechny blendshapes:

$$a_i \geq 0$$

A zároveň všechny váhové koeficienty a_i , které určují příměs jednotlivých blendshapeů, musí dohromady dát součet 1.⁴²

⁴¹ OKUN, J. A., ZWERMAN, S. *The VES Handbook of Visual Effects: Industry Standard VFX Practices and Procedures. 1st Edition.* Focal Press/Elsevier, 2010. ISBN 978-0240812427, str. 379

⁴² JOSHI, P., TIEN, W.C., DESBRUN, M., PIGHIN, F. *Learning Controls for Blend Shape Based Realistic Facial Animation.* Version 1. Online, dataset. In: ACM SIGGRAPH 2006 Courses. 30. července 2006. Dostupné z: <https://doi.org/10.1145/1185657.1185857>. [cit. 2025-04-23].

$$\sum_{i=1}^n a_i = 1$$

V praxi toto tvrzení znamená, že připravené blendshapes lze mixovat a vytvářet tím dosud neexistující nové výrazy a nuance, pouze ale v daném rozsahu [0,1]. Není možné žádný z nich „zesílit“ (např. zvětšit úsměv) nad jeho maximální předepsaný výraz.

Z předchozí definice můžeme vyvodit, že s komplexnějšími obličejovými výrazy bude mít dříve či později tento typ záznamu problém. MA na druhou stranu využívá tzv. 4D solver kombinující depth data společně s identitou Metahuman daného performeru.⁴³

Po importování těchto nahrávek obsahující všechna potřebná data do projektu, je výsledná animace pak vytvořena lokálně pomocí GPU hardwaru během několika minut.

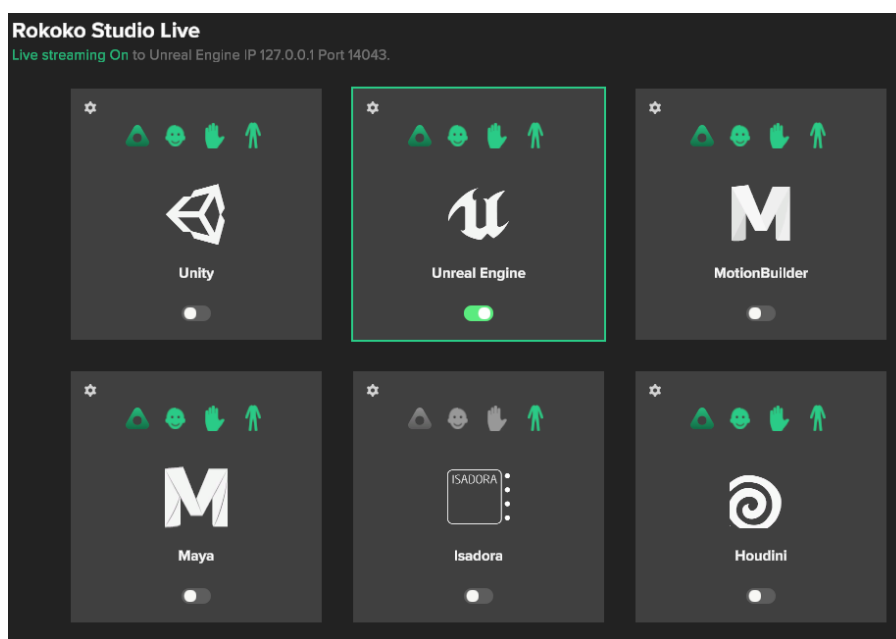
Zapnutím Rokoko Remote na iPhone 12 Pro, které spustí živé streamování dat na defaultní hlavu v Rokoko Legacy, začnu mobil připevňovat na jeho mladšího a výkonnějšího sourozence. Když jsou oba mobily již pevně zajištěny gumičkami, připevním spodní telefon k držáku ring lightu tak, aby byl festovně držel a nehrozilo vypadnutí. Sevření držáku je sice pevné, nicméně výrobce nemohl počítat s tím, že bude k upevněnému mobilu přidružen pomocí gumiček ještě další. Stabilita byla následně ujistěna ještě dodatečným druhým stativem, který podepíral iPhone 14 Pro proti případnému vyklouznutí.

⁴³ UNREALENGINE. Delivering high-quality facial animation in minutes, MetaHuman Animator is now available! In: *UnrealEngine.com* [online] 15. června 2023. Dostupné z: <https://www.unrealengine.com/en-US/blog/delivering-high-quality-facial-animation-in-minutes-metahuman-animator-is-now-available> [cit. 2024-01-27]

6 PŘENOS DAT

6.1.1 LiveLink Rokoko

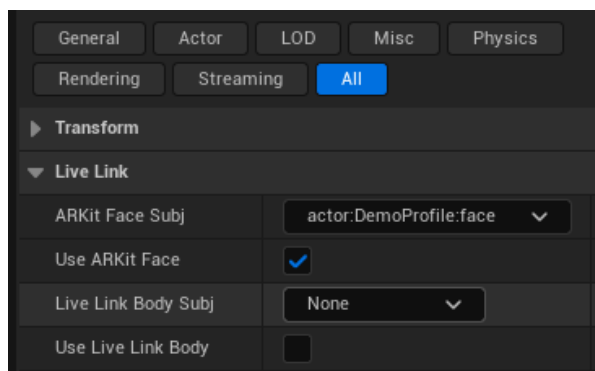
Pro přenos dat z Rokoko Studio Legacy a UE se využívá plugin Live Link přímo v enginu. V tomto případě pozor, nejedná se tentokrát o již zmíněnou mobilní aplikaci, nýbrž o samostatný plugin. Ten se musí aktivovat v nastavení, jelikož při počátečním spuštění UE k dispozici není. Když je plugin aktivovaný, lze ho najít pod kolonkou Window/Virtual production/Live Link. Po rozkliknutí se na obrazovce objeví okno, ve kterém pak můžeme přidat source neboli zdroj našeho živého vysílání. V našem případě se jedná o Rokoko Studio, kde se ale také musí nejdříve zapnout povolení streamovat data do jiného softwaru jednoduchým tlačítkem Start Live Stream. Pak už stačí vybrat jaký program chceme, a Rokoko se už interně postará o to, aby se propojení povedlo.



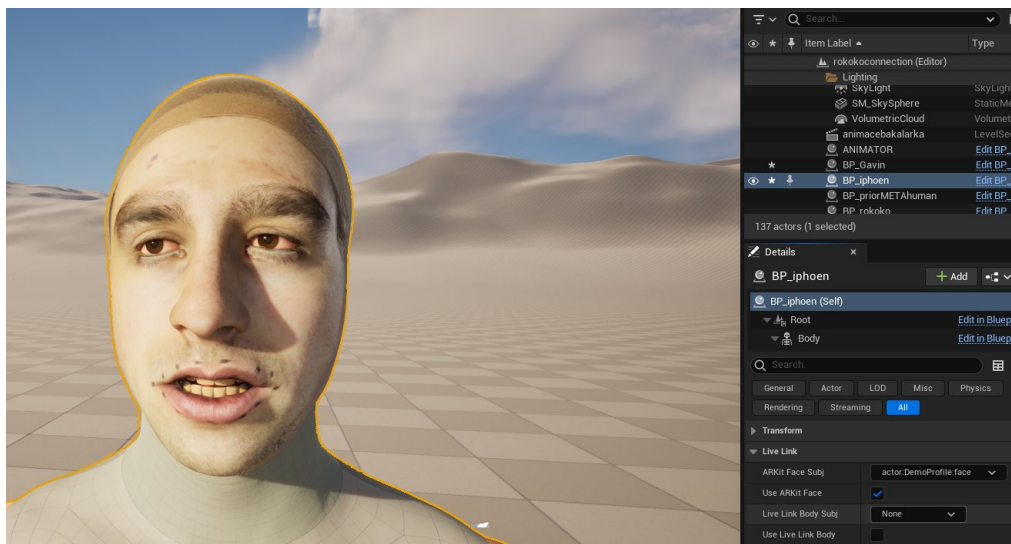
Obr. 16, Náhled propojení Rokoko Studia Legacy s UE

V Live Link manageru se pak následovně objeví možnost propojení s Rokokem, společně se dvěma zdroji, kterými jsou Rokoko suit a Rokoko face. První zdroj referuje na případný mocap oblek, jehož data mohou být, stejně tak jako obličejová, přenášena živě do UE. Druhý zdroj je to, co potřebujeme.

Zda je propojení skutečně funkční zjistíme ověřovací metodou. Klikneme na náš MH charakter a v panelu „Details“ je hned položkou s polohovými, rotačními a velikostními hodnotami sekce „Live Link“. Tam vybereme možnost „ARKit Face Subj“ což je zdroj, ze kterého bude Metahuman čerpat informace o pozicích obličejových svalů (v tomto případě Rokoko Studio Legacy). Dále stačí pouze zaškrtnout kolonku „Use ARKit Face“ a měli bychom okamžitě pozorovat, že na hlavu postavy se nyní přenáší naše obličejové výrazy.



Obr.17, Aktivace propojených softwarů

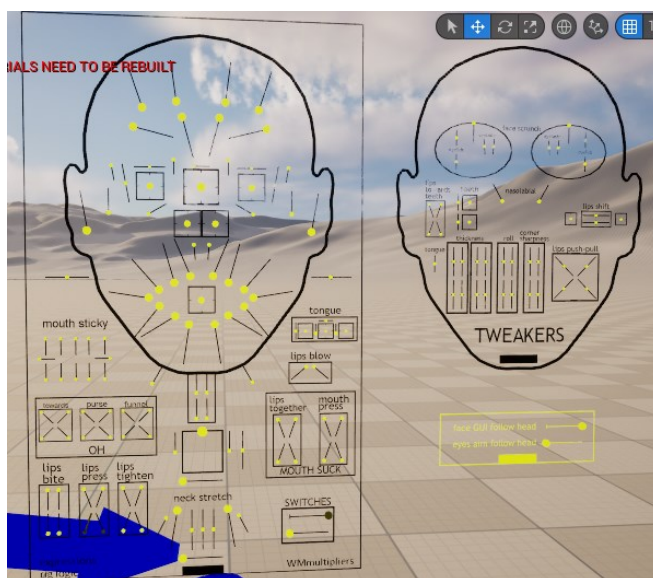


Obr. 18, Úspěšné aplikování přenosu dat na MetaHuman

6.1.2 Nastavení sekvence

Ted' k procesu zpracování měření. Takto připravenou Metahuman postavu následně vložíme do nové Level Sequence abychom mohli přenášená obličejová data z Rokoka nahrát a přeměnit na animaci. Po vložení postavy do sequenceru se objeví dva control rigy, které budeme muset smazat. Jeden je pro končetiny celého těla a druhý, který se nazývá Rig Logic

je pro obličej, na případnou ruční animaci pomocí keyframes nebo pro úpravu zapečené animace ze surových dat třeba Rokoko systému. Rig Logic závisí na univerzálním souboru pravidel, definující svalový systém lidského obličeje a patentovaný formát souboru od 3Lateral nazvaný MetaHuman DNA, který je designovaný pro ukládání kompletního popisu rigu 3D objektu a jeho geometrie.⁴⁴



Obr. 19, Podoba RigLogic

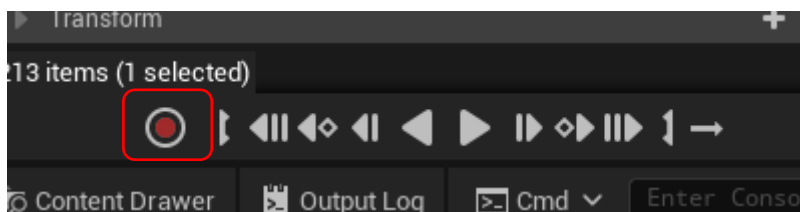
Po odstranění obličejového rigu musíme znovu aktivovat přenos dat pomocí panelu „Details“ u našeho MH charakteru. Co se týče mobilních telefonů, Rokoko Remote už je nějakou dobu aktivované a připravené. Následuje rychlá kalibrace Live Link Face aplikace, kde musíme vybrat možnost nahrávání v MetaHuman Animator módu, zajistit dostatečné světelné podmínky, nastavit mobil do vodorovné polohy tak, že nenatáčí hlavu moc ze spodu ani moc z vrchu a dodržet správnou vzdálenost pro správné zaznamenání (po kliknutí na ozubené kolečko se zobrazí nastavení aplikace, kde je k užítku zaškrtnout kolonku s názvem „Preview Depth“, kdy by měl následně po opuštění nastavení aplikace uživatel vidět šedou reprezentaci své hlavy, ideálně s minimem černých artefaktů, které indikují, že TrueDepth kamera selhává registrovat části obličeje)⁴⁵. Nyní je vše připraveno pro finální nahrání věty.

⁴⁴ RigLogic, Runtime evaluation of MetaHuman face rigs, white paper. [online] 3Lateral. Andrean Franc. Dostupné z: <https://cdn2.unrealengine.com/rig-logic-whitepaper-v2-5c9f23f7e210.pdf> [cit. 2024-01-26].

⁴⁵ UNREALENGINE. [@Unreal Engine]. Creating a MetaHuman Identity from smartphone or HMC footage. Online, video. 2023-08-30. Dostupné z YouTube, <https://www.youtube.com/watch?v=qPhn28Jk3Mo&t=461s> [cit. 2024-01-26].

7 NAHRÁNÍ DAT

Simultánně stiskneme jak v Live Link Face aplikaci, tak dole v Level Sequenceru, červené tlačítko představující snad ve všech filmových a hudebních odvětvích funkci nahrávání. V aplikaci se záznam začne nahrávat okamžitě, v sequenceru má 3 sekundy zpoždění, kvůli přípravě na nahrávání (poskytuje hercům a performerům chvilku čas se připravit a vžít do role).

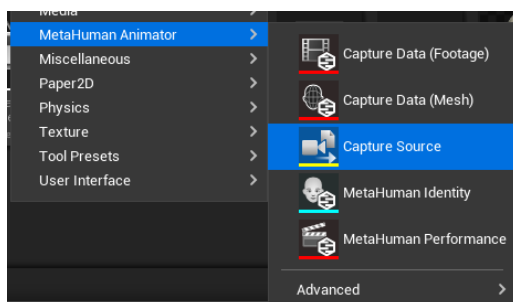


Obr. 20, Tlačítko nahrávání v Level Sequenceru

Během nahrání věty jsem na začátku sekvence udělal grimasu s otevřenou pusou, která bude sloužit jako provizorní klapka, aby mohly být oba záznamy vloženy do jedné timeline a mohla proběhnout synchronizace, jelikož jemné shody mimických svalů se v tomto případě hledají velmi těžko.

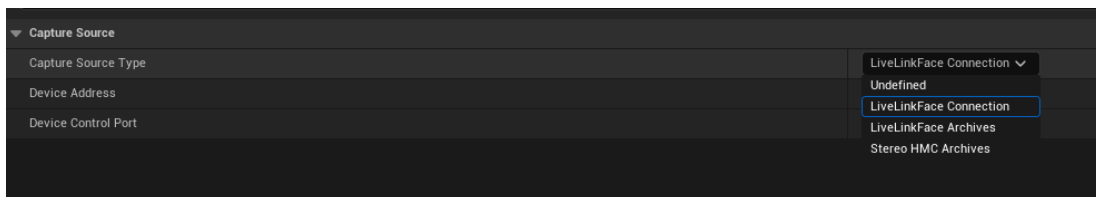
Zatímco záznam Rokoka byl úspěšně nahrán, materiál z Live Link Face aplikace je stále v mobilním zařízení. K tomu, abychom ho přesunuly do projektu jsou na výběr dvě možnosti.

Možnost číslo jedna je přesun dat do projektu přes místní síť. Pomocí pravého kliknutí v Content Browseru (široké okno umístěné dole, které slouží k vytváření, importování, organizování a náhledu obsahu každého Unreal projektu) otevřeme nabídku k tvorbě assetů. V kolonce MetaHuman Animator vybereme možnost „Capture Source“.



Obr. 21, Vytvoření Capture Source assetu

Po kliknutí na tento nově vytvořený asset se nám zobrazí nabídka výběru typu přenosu záznamu. Možnost „LiveLink Connection“ vyžaduje zadání IP adresy, kterou zjistíme buď z Live Link Face v sekci nastavení pod názvem „OSC Server“, nebo pomocí příkazového řádku každého počítače, kde pak uživatel napíše příkaz „ipconfig/all“. Následně stačí najít řádek IPv4 a adresu zkopírovat. Zprostředkovatel tohoto celého procesu se nazývá „Capture Manager“, který se stará o správné importování daného materiálu do enginu.



Obr. 22, Možnosti importu MetaHuman Animator nahrávek do UE

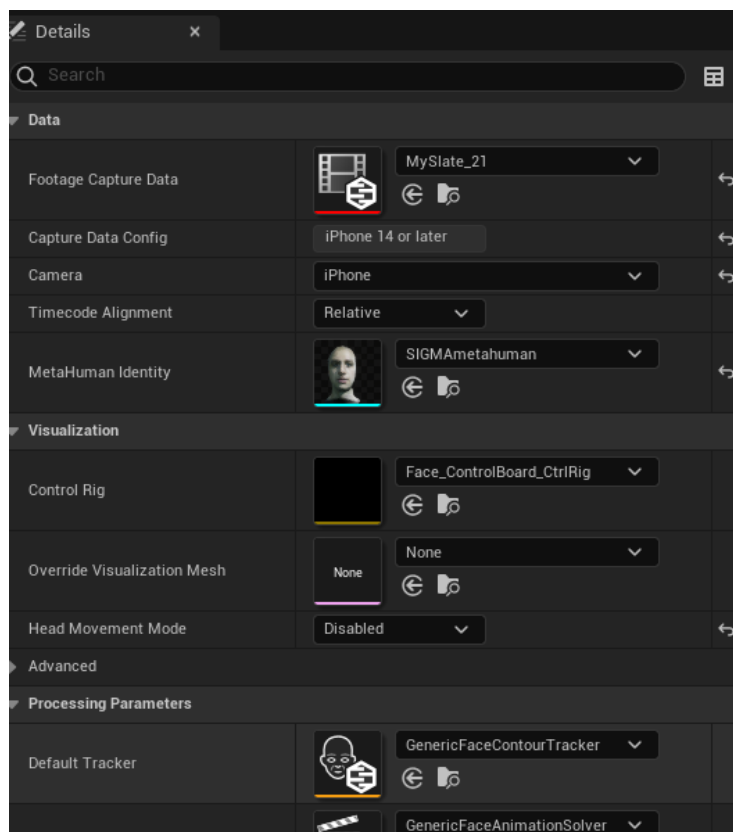
Tento ovladač se připojí na mobilní aplikaci a pomocí Wi-Fi z ní stáhne vybraná videa. Druhým způsobem je možnost záznam zazipovat v iPhonu a odeslat na cloudové úložiště, ze kterého ho můžeme stáhnout do počítače. To pak znamená v „Capture Source“ zvolit možnost „LiveLink Archives“ a vybrat složku, ze které je pak „Capture Manager“ bude importovat do projektu.

Metahuman Animator data byly tedy úspěšně importovány do Unreal Engine projektu. K jejich použití je ale nutné vybrat původní MH hlavu bez těla, která byla vytvořena pomocí Metahuman Identity procesu (z této hlavy pak vznikl celý náš Metahuman ve webovém editoru). Ta byla procesována přes příkaz „Prepare for Performance“ a je možné ji k aplikaci animace použít.

7.1 Vytvoření animace v Metahuman Animatoru

Vytvoříme si nový asset v Content Browseru, který se jmenuje „Metahuman Performance“. Jedná se o kompletně nové uživatelské rozhraní pluginu, ve kterém se raw data z iPhonu 14 Pro změní v animovanou sekvenci aplikovanou na moji hlavu, vytvořenou z importovaného meshe. Do kolonky „Footage Capture Data“ vložíme záznam z mobilní aplikace Live Link Face a jako „MetaHuman Identity“ zvolíme vygenerovanou hlavu MH Identity z původního meshe vytvořeného v Blenderu. V mém případě jsem ještě v nabídce deaktivoval pohyb hlavy, protože ani Rokoko ji nemá. Následně klikneme na tlačítko „Process“ a začne renderování animace v reálném čase. Po pár desítkách vteřin je animace

hotová a buď je možnost jí exportovat jako animaci aplikovatelnou na kteréhokoli MH nebo jí exportovat jako Level Sequence, která koresponduje se sekvencí v tomto editoru. Tuto možnost zvolíme pro naše měření, protože nejenže exportuje animovanou MetaHuman identitu, zároveň s ní také přenese kameru a původní video, na kterém tuto větu říkám. Toto video bude zároveň sloužit jako vzorový template, který umožní porovnání přesnosti jednotlivých softwarů.

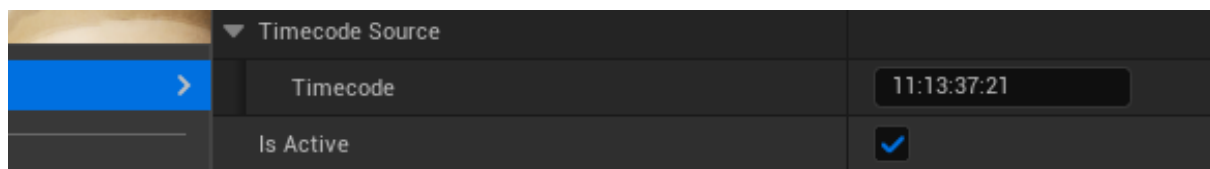


Obr. 23, Nabídka možností před spuštěním MetaHuman Animator sekvence

Manipulace a nastavení

V Level Sequence začíná snímková synchronizace těchto výstupů. Při manipulaci přiloženého videa, sloužícího jako šablona docházelo k velkým zásekům programu a nebylo proto fyzicky možné s ním jakkoli posouvat, proto bylo odstraněno a bude přidáno až v Adobe After Effects.

Pro synchronizaci obou zbylých záznamů byl přidán na záznam obličeje z Rokoko Remote aplikace timecode z Metahuman Animator sekvence. To mělo za následek přesné seřazení obou záznamů na snímek přesně.



Obr. 24, Timecode animace

Postup je teď následující, a to vyrenderovat nejdříve MA sekvenci, za použití vygenerované kamery, která nejlépe simuluje reálnou ohniskovou vzdálenost, úhel a rotaci iPhone, než kdybychom se snažili o replikování těchto hodnot manuálně. Důležité je také aby výsledný obraz byl ostrý bez minimálního motion bluru, který by zamezoval správnému budoucímu trackingu v After Effects. Po nastavení všech signifikantních vlastností kamery můžeme spustit render. Po najetí na ikonku klapky v Level Sequenceru se nám otevrou možnosti nastavení výsledného klipu, jako je druh formátu, počet snímků za sekundu, zda vyžadujeme i audio nebo do jaké složky v počítači chceme, aby se soubor uložil. Jako formát výsledného videa jsem zvolil PNG, který je známý pro svou bezztrátovou kompresi. Jako hodnotu počtu snímků za sekundu jsem zvolil 60, jelikož při trackingu se hodí mít více FPS, pro lepší chytání markerů.

Abychom docílili použití stejné virtuální kamery jako u sekvence z MetaHuman Animatoru, jednoduše umístíme „Rokoko MetaHumana“ na stejnou globální pozici jako byla hlava MH Identity a tím zajistíme stejné vlastnosti kamery i pro Rokoko záznam. Doba renderu u obou sekvencí činila za použití grafické karty NVIDIA RTX 3090 byla zhruba 13 sekund.

7.2 Tracking

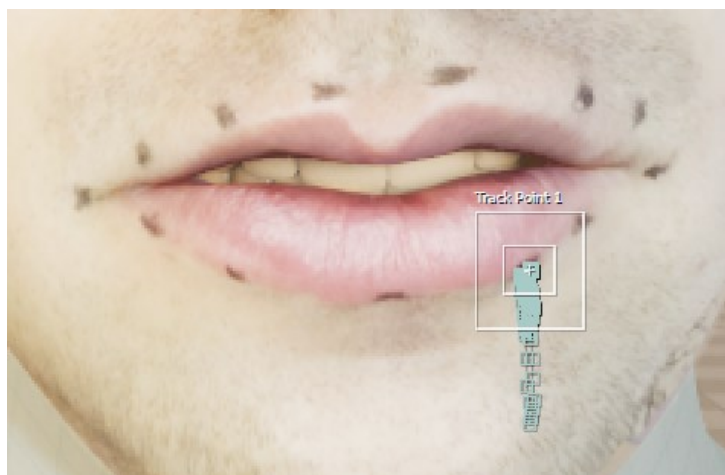
After Effects

Nyní se můžeme přesunout na další část měření a tou je tracking markerů ze třech výstupních zdrojů. Všechny vzniklé PNG sekvence importujeme do AE. Po vložení těchto sekvencí na timeline přidáme i třetí šablonové video reálného záznamu mého obličeje. Následuje postupný tracking markerů na rtech, které pomocí „Track Motion“ funkce

označíme a spustíme automatický tracking, který probíhal ve většině případů dobře, někdy bylo ale nutné automatický proces pozastavit a opravit vzniklé chyby.

Zvolil jsem několik markerů, které budou mít nosnou vypovídající hodnotu, a to dva vrchní pravé a dva spodní, z toho jeden je levý a jeden pravý. Tyto čtyři markery budou vytrackovány napříč vzniklými rendery a šablonovým videem.

Po úspěšném trackingu se musí výsledná data aplikovat na jednotlivé null objekty, které se jmenují podle pozice markeru. Každé video má tedy své čtyři null objekty zaznamenávající informace o pozici trackeru na jednotlivém snímku.



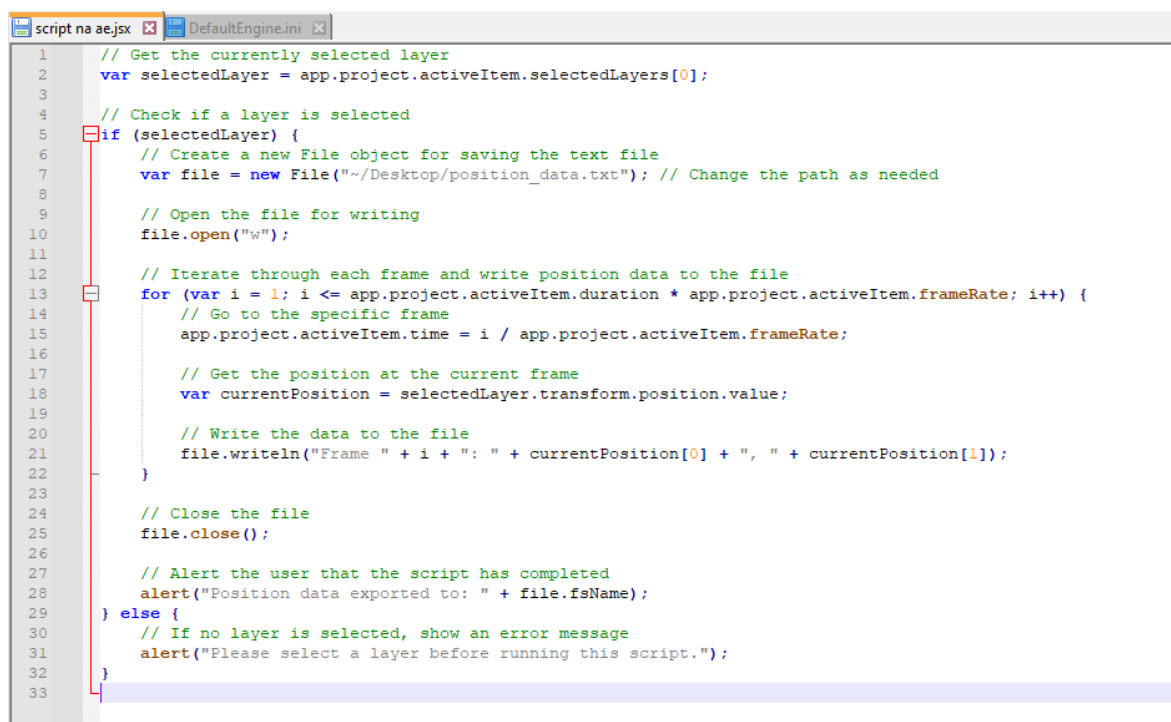
Obr. 25, Snímek obrazovky právě dokončeného tracku jednoho z markerů

8 MĚŘENÍ NASBÍRANÝCH DAT

8.1 Skript

K tomu abychom mohli ale tyto data vůbec analyzovat, je potřeba je mít v číselné podobě někde v tabulkách. Null object ale nenabízí možnost zkopírovat tyto hodnoty někam jinam než zase na některou z komponent v AE timeline. S tímto vědomím je nutné vymyslet metodu, jak data přesunout v číselné podobě do alespoň textového souboru. Program Adobe AE disponuje ovšem možností vložení skriptového příkazu, který vyvolá sled daných naprogramovaných událostí. Jelikož ale nejsem programátor a neumím programovací jazyky, použil jsem k tomuto úkonu umělou inteligenci Chat.GPT, aby pro mě vymyslela skript, který po označení null objektu spustí proces vypsání všech snímků a k nim příslušných globálních souřadnic X a Y trackeru, který sledoval daný marker do textového souboru.

Po několika iteracích a ladění promptů, které jsem psal pro lepší pochopení, co přesně po AI požaduji, se nakonec skript povedl a úspěšně pracuje tak, jak jsem původně chtěl. Textový dokument uloží na plochu, kde ho hned přejmenuji na původní název null objektu a přesunu do pro něj určené složky. Takto vytřídím všechny požadovaná data markerů.



```
1 // Get the currently selected layer
2 var selectedLayer = app.project.activeItem.selectedLayers[0];
3
4 // Check if a layer is selected
5 if (selectedLayer) {
6     // Create a new File object for saving the text file
7     var file = new File("~/Desktop/position_data.txt"); // Change the path as needed
8
9     // Open the file for writing
10    file.open("w");
11
12    // Iterate through each frame and write position data to the file
13    for (var i = 1; i <= app.project.activeItem.duration * app.project.activeItem.frameRate; i++) {
14        // Go to the specific frame
15        app.project.activeItem.time = i / app.project.activeItem.frameRate;
16
17        // Get the position at the current frame
18        var currentPosition = selectedLayer.transform.position.value;
19
20        // Write the data to the file
21        file.writeln("Frame " + i + ": " + currentPosition[0] + ", " + currentPosition[1]);
22    }
23
24    // Close the file
25    file.close();
26
27    // Alert the user that the script has completed
28    alert("Position data exported to: " + file.fsName);
29 } else {
30     // If no layer is selected, show an error message
31     alert("Please select a layer before running this script.");
32 }
33
```

Obr. 26, Zdrojový kód skriptu napsaný umělou inteligencí Chat.GPT

8.2 Přesunutí hodnot do tabulek

Po kontrole dat v textových souborech jsem zjistil, že snímků je nadbytečně moc, protože velká část sekvence je pro mé měření zbytečná, z důvodu absence jakýkoliv slov. Díky tomuto poznatku jsem zredukoval počet snímků na 77, ve kterých se nejvíce promítne daný pohyb. Po přesunutí hodnot do Excelu se mi nedařilo zrealizovat graf, který by dostatečně reprezentoval změnu polohy markerů, proto jsem využil online grafickou aplikaci na vytváření prezentací a grafů Canva, kde se mi již povedlo vyjádřit diferenci polohy markerů vyrenderovaných sekvencí z UE oproti markerům v šablonovém videu.

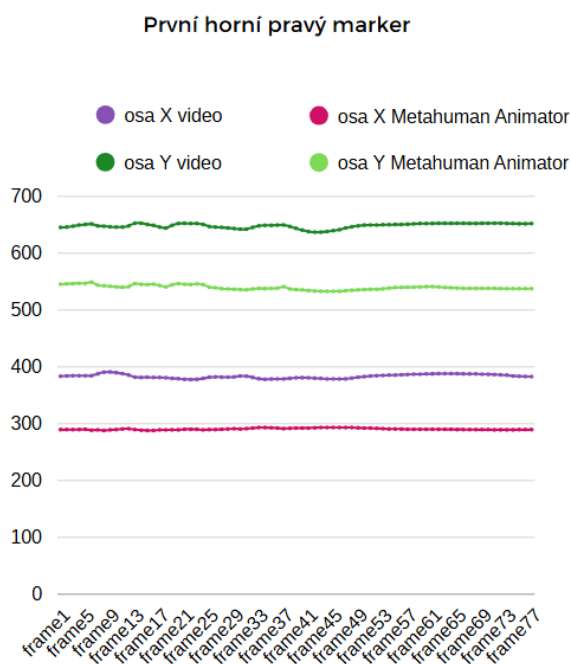
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA
1		source	trčpni spodni pravý				metahuman				rolko						rolko	drůhy spodni levý		source							
2		058 X	058 Y				058 X	058 Y	058 X	058 Y						058 X	058 Y	058 X	058 Y		058 X	058 Y					
3	Frame1	431.82978125	685.0078125				434.4009375	684.82109375	442.4846875	681.8200125						434.75	678.90625	614.048875	702.11028125	536							
4	Frame2	431.80046875	686.5625				434.4015625	685.25	442.453125	681.8575						434.75	679.90625	614.048875	702.13971875	516.5							
5	Frame3	431.0390625	688.7265625				434.4009375	685.5078125	442.40625	681.2285625						434.75	677.2734375	614.34375	704.8515625	517							
6	Frame4	430.8796875	690.44875				434.23828125	686.0315625	442.3140625	682.4821875						430.0803125	677.90625	614.8280625	706.84140625	517.5							
7	Frame5	430.8796875	692.3346875				434.24009375	687.0020125	442.266875	682.62109375						430.25	678.20625	614.859375	710.27734375	518							
8	Frame6	430.2109375	700.625				435.0	694.27734375	442.2421875	683.1171875						430.6875	679.3125	614.9766625	721.08864375	517.4769625							
9	Frame7	431.45703125	718.4009375				431.18015625	705.67978125	441.978125	687.30328125						438.7878125	684.8540625	617.70703125	742.2109375	517.51893125							
10	Frame8	433.7421875	728.703125				433.0889375	708.46796875	441.109375	671.578125						438.75	690.12109375	617.7090625	751.754375	517.01893125							
11	Frame9	433.1171875	732.18921875				433.87109375	710.75	442.88203125	681.75						438.75	705.12109375	615.70703125	784.10546875	518.2109375							
12	Frame10	430.80234375	733.8303125				434.0889375	708.75	444.0889375	690.3046875						438.75	716.2890625	613.625	756.8828125	521.62109375							
13	Frame11	428.015625	726.625				434.02134375	704.75	444.015625	682.75						438.95703125	721.62109375	611.65625	752	523.01893125							
14	Frame12	425.2390625	715.7205625				436.20703125	697.75	444.0889375	683.75						439.59875	724.200125	610.625	758.038625	524.778125							
15	Frame13	427.009375	701.125				436.92734375	688.72285625	442.4546875	682.4933125						430.4785625	711.83984375	615.7578125	716.0889375	524.2530625							
16	Frame14	427.76171875	696.86328125				434.44875	688.25	444.2734375	672.15234375						430.99809375	695.12109375	617.23828125	712.2578125	523.0234375							
17	Frame15	428.4784375	695.57109375				433.88471875	688.1171875	443.15234375	686.75						431.015625	682.7890625	616.875	710.21875	522.95078125							
18	Frame16	428.3046875	697.85546875				434.703125	689.44140625	442.796875	687.10546875						431.28125	681.71875	614.3828125	722.1484375	517.90625							
19	Frame17	427.40234375	705.125				435.125	696.75	442.80625	687.0625						431.46875	681.41796875	613.38671875	728.30889375	519.684875							
20	Frame18	427.3671875	708.73046875				434.94140625	698.75	442.34375	686.63671875						430.94140625	687.62109375	613.74609375	735.90234375	519.13125							
21	Frame19	427	701.86015625				434.8203125	692.8396875	442.1209375	671.75						430.9339375	693.51625	613.59375	722.54286875	520.0625							
22	Frame20	426.9785625	694.8621875				435.5	688.5640625	442.7421875	670.75						430.25	691.84785625	612.2734375	711.8328125	519.8804875							
23	Frame21	426.796875	693.2205625				435.33984375	687.6484375	441.9339375	686.94821875						439.984375	686.58203125	610.375	708.31640625	520.50078125							
24	Frame22	426.66015625	693.2734375				435.339375	687.390625	442.046875	687.6228125						439.25	681.4875	609.8640625	708.74464375	520.99375							
25	Frame23	425.96484375	694.46875				437.62109375	688.76171875	442.3671875	687.08894375						439.25	680.426875	610	715.34785625	519.8125							
26	Frame24	425.88828125	703.2265625				436.82140625	692.75	442.0234375	686.5546875						439.3515625	680.660625	609.90625	728.78515625	519.38671875							
27	Frame25	427.26296875	707.65234375				437.03125	694.75	442.87109375	686.91796875						439.62109375	685.5589375	609.75390625	736.03859375	518.56640625							
28	Frame26	428.8176875	707.3546875				435.8046875	693.140625	442.8007125	687.53125						439.92734375	688.72285625	611.70906875	735.11328125	517.890625							
29	Frame27	431.14453125	705.80546875				433.8359375	690.75	444.90234375	687.96875						439.25	691.76171875	613.375	729.76953125	514.59375							
30	Frame28	431.7734375	700.8125				434.1015625	689.71484375	445.20703125	687.60015625						438.98046875	691.74809375	615.60546875	729.72265625	518.94621875							
31	Frame29	432.0975625	700.1280625				435.00125	689.62109375	445.83303125	684.796875						438.49809375	687.62109375	614.5	723.953125	514.48921875							
32	Frame30	431.9785625	700.46875				437.823125	688.75	445.91015625	685.875						438.25	685.90234375	612.2890625	724.7734375	514.94821875							
33	Frame31	430.8396875	700.125				440.0406875	687.75	445.9339375	683.9785625						438.25	686.34785625	608.9785625	733.0889375	518.703125							
34	Frame32	428.84140625	697.8453125				440.80328125	683.75	445.875	686.4921875						438.25	686.4140625	608.8089375	716.09375	520.201125							
35	Frame33	427.703125	694.0125				440.821875	678.2734375	445.8515625	684.45125						438.25	685.8028125	608.6171875	706.4446875	520.1238125							
36	Frame34	427.717875	693.78515625				436.38125	677.95703125	444.78125	682.4785625						438.4140625	680.8089375	603.97265625	703.8205125	521.84464375							
37	Frame35	427.77734375	694.83984375				439.984375	678.11328125	444.07421875	687.0889375						431.328125	677.71484375	604.52734375	704.82421875	521.1640625							
38	Frame36	427.9785625	697.5				440.0078125	690.4140625	444.4021875	687.1280625						432.249625	675.71875	606.453125	711	520.0090625							
39	Frame37	426.546875	706.015625				432.1484375	689.06640625	444.4375	681.9339375						432.515625	676.3828125	607.71875	728.64403125	521.4296875							
40	Frame38	426.5234375	713.62109375				436.1171875	694.74009375	443.44140625	683.0193125						432.203125	684.5426875	608.45703125	738.1171875	518.16796875							
41	Frame39	426.8671875	707.455125				436.13671875	689.75	444.0889375	686.40625						430.88453125	693.20703125	609.84785625	730.11328125	519.66015625							
42	Frame40	427.01171875	698.64453125				437.0140625	694.703125	444.7890625	671.0078125						439.36328125	696.42109375	610.74121875	720.8515625	521.04296875							
43	Frame41	425.82978125	694.0193125				438.1875	681.75	445.22465625	687.6796875						438.10546875	691.421875	610.5625	732.8539375	521.86796875							
44	Frame42	426.03515625	692.28515625				437.2285625	679.6015625	445.33203125	685.75						439.02734375	688.3546875	610.46484375	711.96484375	520.7734375							
45	Frame43	427.2734375	691.34609375				437.8859375	678.75	445.3303125	683.91796875						438.62109375	685.5589375	612.4785625	709.9609375	520.05125							
46	Frame44	427.02978125	692.1893125				427.2284375	678.74009375	445.15625	682						439.1953125	684.3671875	613.71875	709.90234375	519.125							
47	Frame45	428.09785625	693.1890625				437.0078125	676.5193125	444.6875	682.6484375						438.98046875	682.8828125	614.625	710.12890625	519.625							
48	Frame46	428.21875	684.91796875				439.89375	679.75	444.3303125	682.5234375						438.12109375	682.4009375	615.71484375	712.046875	519.8089375							
49	Frame47	427.01171875	698.214175				436.16140625	681.4015625	444.8426875	682.45703125						438.14615625	6										

9 ANALÝZA MĚŘENÍ

Analýza proběhne vizuálním porovnáním vzniklých křivek všech měřených subjektů. Jako příkladový a ideální výsledek je vždy zdrojové video. Data jsou často snížena o 100 či 200 bodů, aby nebyly vzniklé křivky naskládány v sobě a tím zamezily vizuálnímu porovnání, jelikož by jejich tvar mezi sebou zanikl. Pro měření bylo vzato 77 snímků, které jsou reprezentovány vždy pod grafem v podobě slova „frame“ jako určovatel časové jednotky. Při trackování byla snímková frekvence všech záznamů v AE upravena na 30 FPS. Videá v kompozici byla zkrácena pomocí funkce Time Stretch o polovinu. Důvodem bylo to, že oproti originálnímu záznamu byly vyrenderované sekvence přesně o polovinu pomalejší. Nashromážděná data jsou k dispozici pod tímto odstavcem.

Při důkladnější analýze si můžeme povšimnout, že ve většině případech je osa Y mnohem přesnější a vypovídající než osa X. Důvodem může být pravděpodobně ne zcela přesný úhel kamery, lehce posunutá textura a obecně známý fakt, že když vyslovíme hlásku „m“, hlavní pohyb vytváří dolní čelist pohybem shora dolů, tudíž po ose Y.

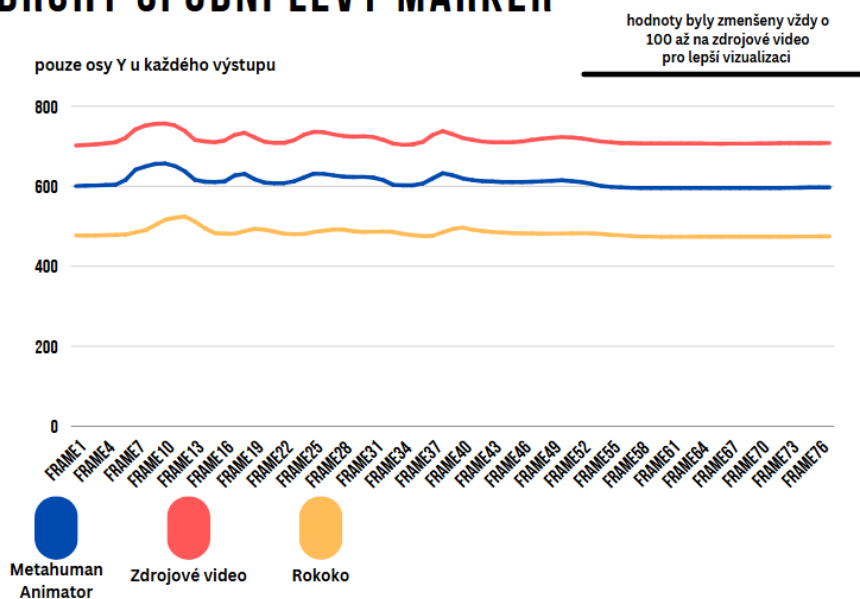
Zde je zobrazen graf č. 1, ukazující pohyb markeru prvního horního pravého rtu. Osa Y je u MetaHuman postavy až na výjimky velmi podobná svému vzoru, kdežto u osy X je podoba minimální.



Graf 1

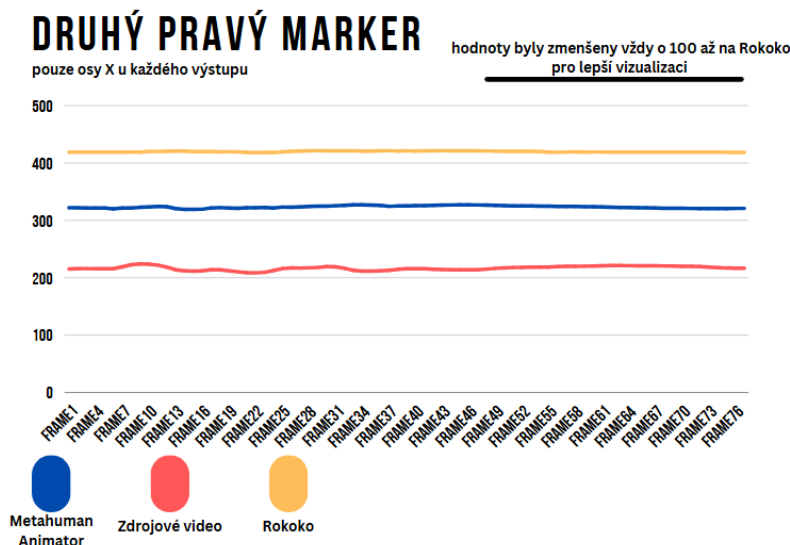
Na grafu č. 2 lze vidět dvě téměř identické elipsy zdrojového videa a MetaHuman Animatoru. Rokoko už je na tom o něco hůře, protože zcela ignorovalo jemné nuance pohybů rtů, signifikantní části však zaznamenalo také. Tento graf je z celkového měření ten nejprůkaznější a nejlépe vystihuje chování jednotlivých softwarů.

DRUHÝ SPODNÍ LEVÝ MARKER



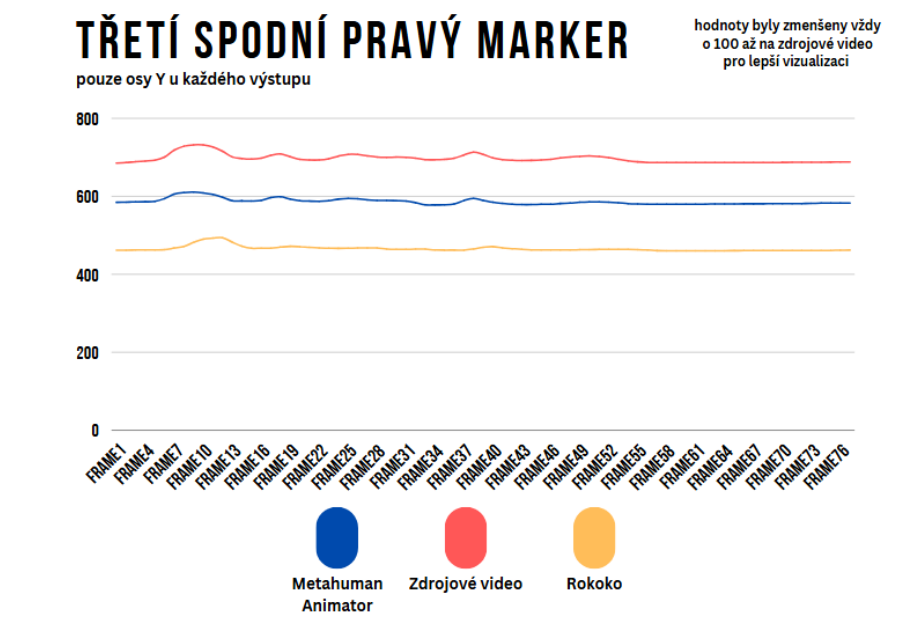
Graf 3

DRUHÝ PRAVÝ MARKER



Graf 2

Graf č. 3 ukazuje, že problematická osa X u Rokoka nebyla téměř registrována, u MetaHuman Animator sekvence je vidět jemné zakřivení zhruba odpovídající zdrojovému videu.



Graf 4

Stejný případ jako u předchozího výsledku měření osy Y u testovaných subjektů. Metahuman velmi přesvědčivě zvládá zachytit i jemné pohyby rtů, na rozdíl od Rokoko systému, kterému se znovu podařilo zaznamenat pouze ty největší a nejznatelnější pohyby.

10 ZÁVĚR

Na základě provedeného experimentu měření přesnosti Rokoko Studia a Metahuman Animatoru via jejich vlastní aplikace zprostředkovávající data na vybrané digitální MetaHuman charaktery lze konstatovat, že z dostupných informací v podobě grafického srovnání výsledného trackingu jednotlivých markerů je blíže šablonovému videu plugin MetaHuman Animator. Tento fakt potvrzuje hypotézu, která byla nakloněna spíše k MA, byť jde stále o experimentální plugin, jsou jeho výsledky velmi dobré. U systému Rokoko lze pozorovat křivky, které ale bohužel v některých případech jen velmi tlumeně kopírují šablonové video.

Původní hypotéza tedy nakonec vešla v platnost a byla potvrzena tímto měřením. Během experimentu jsem naskenoval svoji tvář omarkrovanou okolo rtů černými tečkami, použil sken jako podklad pro vytvoření narigovaného obličej v UE, nahradil naskenovanou texturu za původní texturu MH, nastavil snímací aparaturu telefonů, nahrál simultánně oba výstupy, natracoval markery všech výstupů včetně zdrojového referenčního videa pomocí After Effects, pomocí skriptu převedl null objekty do souřadnic aplikovatelných do grafů a změřil přesnost obou snímaných sekvencí porovnáním se zdrojovou nahrávkou pořízenou přední kamerou.

Z pohledu využití obou softwarů v praxi lze vyvodit, že Rokoko bude sloužit spíše jako slušný nástroj pro vizualizaci obličejových emocí, už jen díky možnosti živého streamingu dat přímo na postavu v reálném čase, než software, přes který by se nahrávaly cinematics nebo cutscenes do videoher. Pro kvalitní výstup, použitelný právě do těchto segmentů je tou správnou volbou MetaHuman Animator.

SEZNAM POUŽITÉ LITERATURY

- [1] OKUN, J. A., ZWERMAN, S. *The VES Handbook of Visual Effects: Industry Standard VFX Practices and Procedures. 1st Edition.* Focal Press/Elsevier, 2010. ISBN 978-0240812427, str. 335, [cit. 2024-01-27].
- [2] SIEGEL, Jay A.; SAUKKO, Pekka a HOUCK, Max M. (ed.), [2013]. *Encyclopedia of forensic sciences. Second edition.* Amsterdam: Elsevier. ISBN 978-0-12-382165-2, str. 368, [cit. 2024-01-27].
- [3] KITAGAWA, Midori. *MoCap for Artists: Workflow and Techniques for Motion Capture.* Routledge, 2008. ISBN 10 0240810007. str. 160, [cit. 2024-06-12].

SEZNAM INTERNETOVÝCH SBORNÍKŮ

- [4] HARTLEY, R. I., STURM, P., *Triangulation.* Version 1. Online, dataset. In: Elsevier. 19. dubna 2002. str. 146, Dostupné z: <https://doi.org/10.1006/cviu.1997.0547>. [cit. 2025-04-23].
- [5] LE B. H., ZHU, M. a DENG, Z. *Marker Optimazation for Facial Motion Acquisition and Deformation.* Version 1. Online, dataset. In: IEEE. 17. května 2013. Dostupné z: <https://doi.org/10.1109/tvcg.2013.84>. [2025-4-23].
- [6] WEISE, T., BOUAZIZ, S., LI, H., PAULY, M. *Realtime Performance-Based Facial Animation.* Version 1. Online, dataset. In: ACM Transaction on Graphics. 25. července 2011. Dostupné z: <https://doi.org/10.1145/2010324.1964972>. [cit. 2025-04-23].
- [7] BLAIR, T. B., DAVIS, CH. E., *Innovative Engineering Outreach: A Special Applications of the Xbox 360 Kinect Sensor.* Version 1. Online, dataset. In: IEEE Frontiers in Education Conference. 19. Prosince 2013. Dostupné z: <https://ieeexplore.ieee.org/document/6685036>. [cit. 2025-04-23].
- [8] BASSETT, R., DERIDE, J., *Maximum a Posteriori Estimators as a Limit of Bayes Estimators.* Version 2. Online, dataset. In: arXiv. 21. ledna 2018. str. 2, Dostupné z: <https://doi.org/10.48550/arXiv.1611.05917>. [cit. 2025-04-23].
- [9] MIHCIN, S. *et al., Investigation of Wearable Motion Capture System Towards Biomechanical Modelling.* Version 1. Online, dataset. In: IEEE Xplore. 28. červen 2019, str. 2. Licence: CC BY 4.0. Dostupné z: <https://doi.org/10.1109/MeMeA.2019.8802208> [cit. 2024-01-26].
- [10] TAEGER J., BISCHOFF S., HAGEN R., RAK K. *Utilization of Smartphone Depth Mapping Cameras for App-Based Grading of Facial Movement Disorders: Development and Feasibility Study.* Version 9. Online, dataset. In: JMIR Publications 2021. Dostupné z: <https://mhealth.jmir.org/2021/1/e19346>. [3,37 MB]. [cit. 2024-01-27].
- [11] JOSHI, P., TIEN, W.C., DESBRUN, M., PIGHIN, F. *Learning Controls for Blend Shape Based Realistic Facial Animation.* Version 1. Online, dataset. In: ACM SIGGRAPH

2006 Courses. 30. července 2006. Dostupné z: <https://doi.org/10.1145/1185657.1185857>. [cit. 2025-04-23].

SEZNAM INTERNETOVÝCH ZDROJŮ

[12] *Fiducial marker*. Online. In: Wikipedia. Stránka byla naposledy editována 7.11.2023 v 19:58. Dostupné z: https://en.wikipedia.org/wiki/Fiducial_marker [cit. 2024-01-27].

[13] FACEWARE. Online. [2025]. Dostupné z: <https://facewaretech.com/pricing>. [cit. 2025-04-24].

[14] FACEWARE. *Pricing*. Online. [2025]. Dostupné z: <https://facewaretech.com/pricing>. [cit. 2025-04-24].

[15] REALLUSION. *iClone 8, Dramatically simplify character animation*. Online. [2025]. Dostupné z: <https://www.reallusion.com/iclone/>. [cit. 2025-04-24].

[16] REALLUSION. *AccuFACE, Face Tracking from Live or Recorded Video*. Online. [2025]. Dostupné z: <https://www.reallusion.com/iclone/motion-capture/accuface.html>. [cit. 2025-04-24].

[17] *Pitch, yaw, and roll*. Online. In: Wikipedia. Stránka byla naposledy editována 7.4.2023 v 5:52. Dostupné z: https://simple.wikipedia.org/wiki/Pitch,_yaw,_and_roll [cit. 2024-01-26].

[18] ROKOKO. *Capture accurate finger motions with Smartgloves*. Online. Dostupné z: <https://www.rokoko.com/products/smartgloves> [cit. 2024-01-26].

[19] EPIC GAMES, Inc. Frequently asked questions (FAQs). In: *UnrealEngine.com* [online]. 2004-2024. Dostupné z: <https://www.unrealengine.com/en-US/faq> [cit. 2024-1-26].

[20] JENSEN, K. T. 25 Years Later: The History of Unreal and an Epic Dynasty. In: *PCMag.com* [online]. 22. květen 2023, Dostupné z: <https://www.pcmag.com/news/25-years-later-the-history-of-unreal-and-an-epic-dynasty> [cit. 2024-01-26].

[21] DESOWITZ, B. The Mandalorian: How ILM's Innovative StageCraft Tech Created a 'Star Wars' Virtual Universe. In: *IndieWire.com* [online]. 20 února 2020 17:01. Dostupné z: <https://www.indiewire.com/features/general/the-mandalorian-ilm-stagecraft-1202212389/> [cit. 2024-01-26].

[22] LUBINA, I. Jak nastavit clonu. 7 tipů pro nastavení clony. In: *Fotografvna.cz* [online]. 10. listopadu 2021. Dostupné z: <https://fotografvna.cz/jak-nastavit-clonu-7-tipu-pro-nastaveni-clony/>. [cit. 2024-01-27].

[23] Real Human to Metahuman Photogrammetry Workflow. In: Youtube [online]. 12. června 2022. Dostupné z: <https://www.youtube.com/watch?v=8JVGb9kBNg&t=798s> [cit. 2024-01-25].

- [24] KERR, C. Epic Games acquires photogrammetry software developer CapturingReality. In: *Gamedeveloper.com* [online] 9. března 2021. Dostupné z: <https://www.gamedeveloper.com/business/epic-games-acquires-photogrammetry-software-developer-capturing-reality>. [cit. 2024-01-26].
- [25] KEENTOOLS.IO. *KeenTools add-ons for Blender*. Online. [2023]. Dostupné z: <https://keentools.io/download/facebuilder-for-blender> [cit. 2024-01-26].
- [26] WARD, C. How to Move the Anchor Point in After Effects. In: *Schoolofmotion.com* [online]. Dostupné z: <https://www.schoolofmotion.com/blog/anchor-point-in-after-effects>. [cit. 2024-01-26].
- [27] MATNEY, L. Epic Games buys 3Lateral, maker of super-realistic ‘digital humans’. In: *Techcrunch.com* [online] 23. ledna 2019. Dostupné z: <https://techcrunch.com/2019/01/23/epic-games-buys-3lateral-maker-of-super-realistic-digital-humans/> [cit. 2024-01-26].
- [28] 3LATERAL. Metahuman Animator Launched. In: *3Lateral.com* [online] 15. června 2023. Dostupné z: <https://www.3lateral.com/news/metahuman-animator-launched.html> [2024-01-26].
- [29] APPLE. Instance Property, blendShapes, A dictionary of named coefficients representing the detected facial expression in terms of the movement of specific facial features. In: *Developer.Apple.com* [online]. Dostupné z: <https://developer.apple.com/documentation/arkit/arfaceanchor/2928251-blendshapes>. [cit. 2024-01-27].
- [30] UNREALENGINE. Delivering high-quality facial animation in minutes, MetaHuman Animator is now available! In: *UnrealEngine.com* [online] 15. června 2023. Dostupné z: <https://www.unrealengine.com/en-US/blog/delivering-high-quality-facial-animation-in-minutes-metahuman-animator-is-now-available>. [cit. 2024-01-27].
- [31] *RigLogic, Runtime evaluation of MetaHuman face rigs, white paper*. [online] 3Lateral. Andrean Franc. Dostupné z: <https://cdn2.unrealengine.com/rig-logic-whitepaper-v2-5c9f23f7e210.pdf>. [cit. 2024-01-26].
- [32] UNREALENGINE. [@Unreal Engine]. *Creating a MetaHuman Identity from smartphone or HMC footage*. Online, video. 2023-08-30. Dostupné z YouTube, <https://www.youtube.com/watch?v=qPhn28Jk3Mo&t=461s>. [cit. 2024-01-26].

SEZNAM OBRÁZKŮ

Obr. 1. Naskenovaný obličej postavy Normana z mého bakalářského filmu pomocí fotogrammetrie, vedle je UV mapa textury

Obr. 2. Předloha k vytvoření mých markerů

Obr. 3. Mnou provedené omarkrování rtů

Obr. 4. Funkce pluginu FaceBuilder

Obr. 5. Rotace obličeje pod jiným úhlem

Obr. 6. Textura z pluginu FaceBuilder aplikovaná na základní

Obr. 7. Mesh s aplikovanou texturou, obojí vytvořeno ve FaceBuilderu

Obr. 8. Auto-track obličejových rysů MetaHuman Identity solverem

Obr. 9. Vzniklý MetaHuman charakter po aplikaci kůže v MetaHuman Creatoru

Obr. 10. Výšeč textur, které se musí změnit na mnou vytvořenou texturu s trackovacími body

Obr. 11. Ukázka úchopu obou telefonů pro následující experiment

Obr. 12. Plastový držák svírající jeden z mobilů

Obr. 13. Nabídka ke spárování aplikace Rokoko Remote

Obr. 14. Uživatelské rozhraní aplikace Rokoko Remote po spárování s PC softwarem

Obr. 15. Úvodní nabídka aplikace Live Link Face

Obr. 16. Náhled propojení Rokoko Studio Legacy s UE

Obr. 17. Registrování Rokoko Studio pluginem Live Link

Obr. 18. Úspěšné aplikování přenosu dat na MetaHuman

Obr. 19. Podoba RigLogic

Obr. 20. Tlačítko nahrávání v Level Sequenceru

Obr. 21. Vytvoření Capture Source assetu

Obr. 22. Možnosti importu MetaHuman Animator materiálu do UE

Obr. 23. Nabídka možností před spuštěním MetaHuman Animator sekvence

Obr. 24. Timecode animace

Obr. 26. Zdrojový kód skriptu napsaný umělou inteligencí Chat.GPT

Obr. 27. Tabulka, zaznamenávající jedny z mnoha dat o souřadnicích os X a Y

SEZNAM TABULEK

Graf 1	53
Graf 3	54
Graf 2	54
Graf 4	55

SEZNAM POUŽITÝCH SYMBOLŮ A ZKRATEK

MA	MetaHuman Animator
MH	MetaHuman
AE	Adobe After Effects
ARkit	vývojový nástroj umožňující tvorbu rozšířené reality na zařízeních iOS prostřednictvím integrovaných funkcí kamery a senzorů.
RIG	proces přidělování kostí a ovládacích prvků modelu, umožňující animátorům pohybovat a deformovat digitální postavy či objekty
RENDER	proces vytváření vizuálního výstupu z 3D modelu nebo scény, kde počítač generuje obrázky či animace
ASSET	digitální objekt nebo soubor obsahující 3D model, textury, a další prvky
TRACKING	proces identifikace a sledování specifických bodů v obrazu nebo videu za účelem určení jejich pohybu a pozice v čase
GPU	grafická jednotka počítače
CPU	procesorová jednotka počítače
MESH	sít' propojených vrcholů, hran a stěn, která reprezentuje geometrii trojrozměrného objektu v počítačové grafice
UE	Unreal Engine
BLENDSHAPE	metoda animace, při které se výraz tvoří interpolací mezi předem vytvořenými tvarovými variantami (např. úsměv, mračení)
CUTSCENE	naskriptovaná filmová sekvence ve hře, která slouží k vyprávění děje mimo hráčovu kontrolu