govtech Polska

Politechnika Wrocławska

AI Tech
AKADEMIA INNOWACYJNYCH ZASTOSOWAŃ TECHNOLOGII CYFROWYCH

Inne zastosowania metod uczenia maszynowego w animacjach

SZTUCZNA INTELIGENCJA W ANIMACJI

Tomasz Zamojski

Fundusze Europejskie Polska Cyfrowa

Rzeczpospolita Polska

Unia Europejska Europejski Fundusz Rozwoju Regionalnego

Projekt współfinansowany ze środków Unii Europejskiej w ramach Europejskiego Funduszu Rozwoju Regionalnego Program Operacyjny Polska Cyfrowa na lata 2014-2020,
Oś Priorytetowa nr 3 "Cyfrowe kompetencje społeczeństwa" Działanie nr 3.2 "Innowacyjne rozwiązania na rzecz aktywizacji cyfrowej"
Tytuł projektu: „Akademia Innowacyjnych Zastosowań Technologii Cyfrowych (AI Tech)”
Porozumienie nr POPC.03.02.00-00-0001/20-00

Plan wykładu

1. Wprowadzenie - rys historyczny

2. Przegląd współczesnych systemów przechwytywania ruchu

- Hierarchiczny model szkieletu człowieka
- Systemy optyczne (pasywne i aktywne)
- Systemy nieoptyczne

3. Przechwytywanie ruchu z wykorzystaniem metod AI

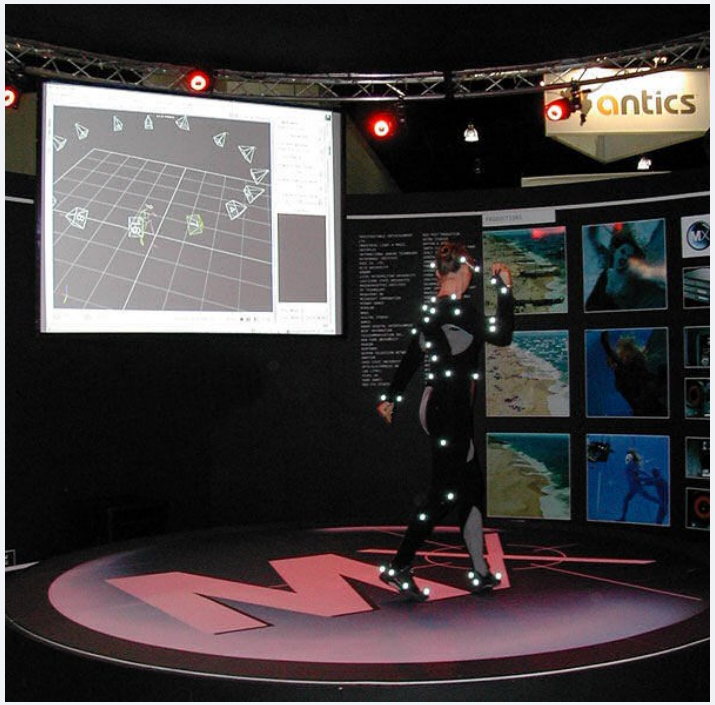
- Szkielet pozy człowieka - Open Pose (2d)
- GHUM & GHUML: generatywne modele 3D postaci człowieka i jej pozy przegubowe
- Przechwytywanie ruchu bez specjalistycznego sprzętu i danych (przykłady aplikacji)

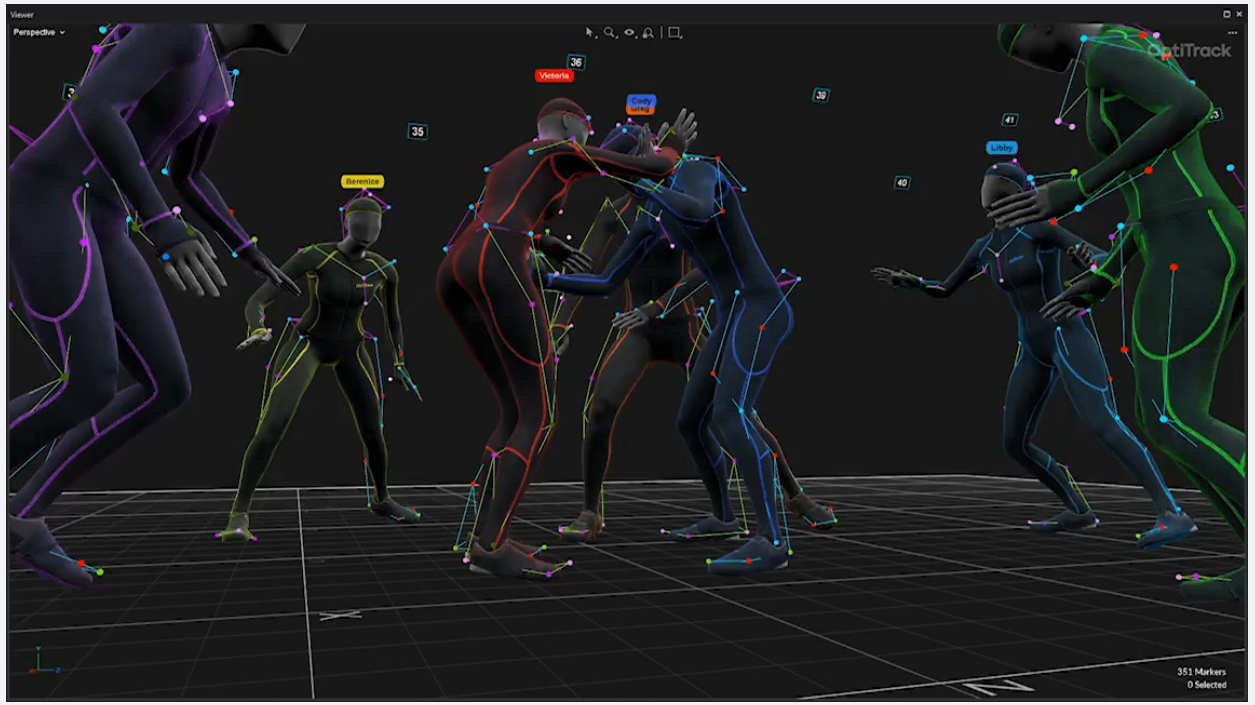
4. Literatura, bibliografia, źródła

1. Wprowadzenie, rys historyczny

1. Przechwytywanie ruchu – co to?

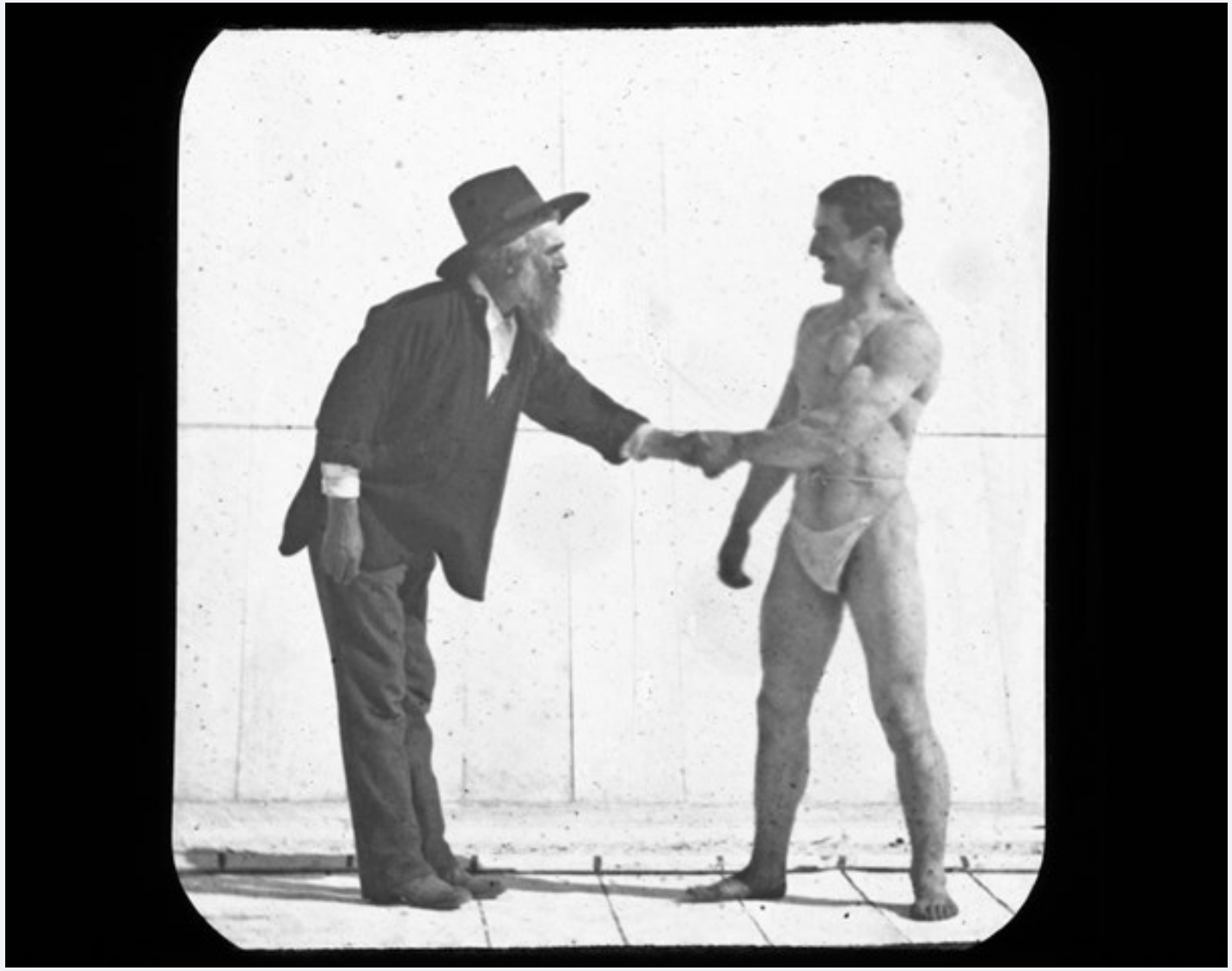
- **Przechwytywanie ruchu** (ang. *motion capture*, skrótowo *mocap*) – technika stosowana w filmach i grach komputerowych, polegająca na „przechwytywaniu" trójwymiarowych ruchów aktorów i zapisywaniu ich w komputerze.
- Dzięki niej zarejestrowane postacie poruszają się bardzo naturalnie i realistycznie, a przechwycone dane są gotowe do dalszej obróbki. Technologia ta nie zastępuje w pełni pracy animatora, wnosi jednak znaczące wsparcie do procesu animacji.
- Aby stworzyć animacje za pomocą tej techniki potrzebne są kamery oraz markery (czujniki) *(przynajmniej do niedawna)*



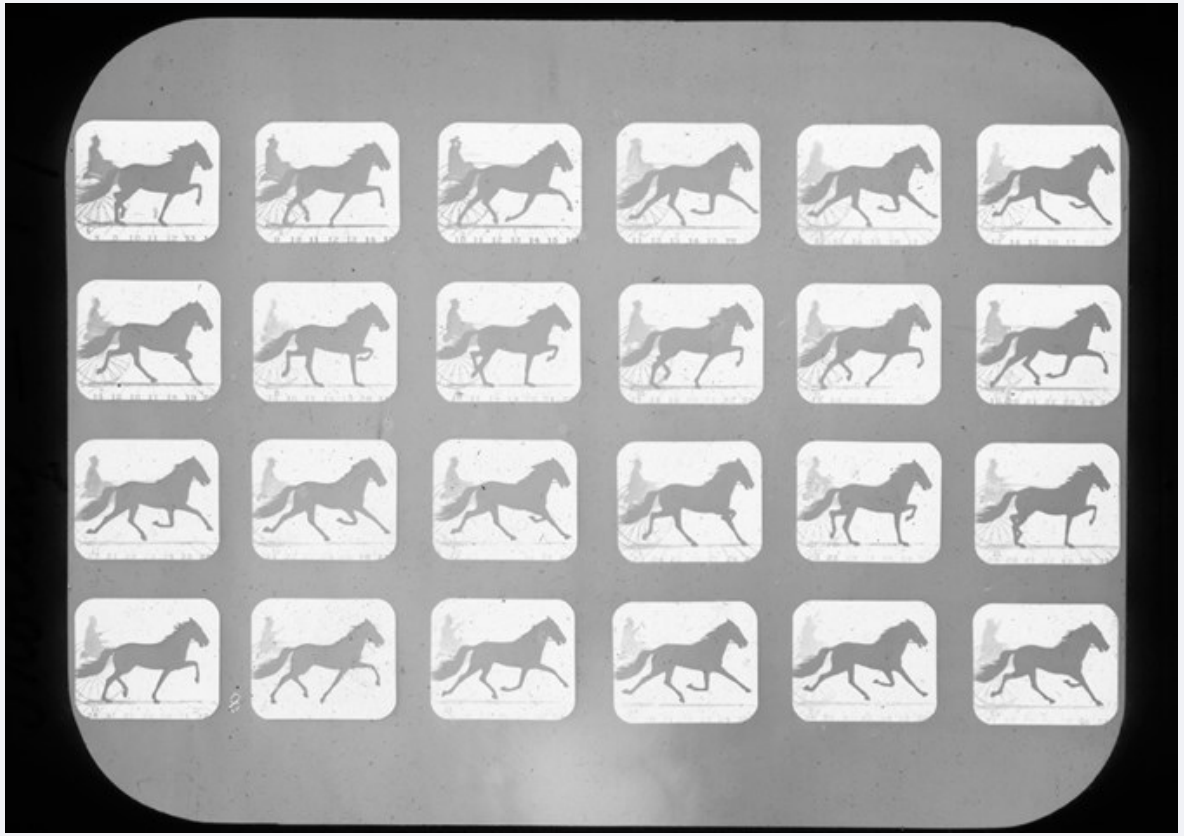


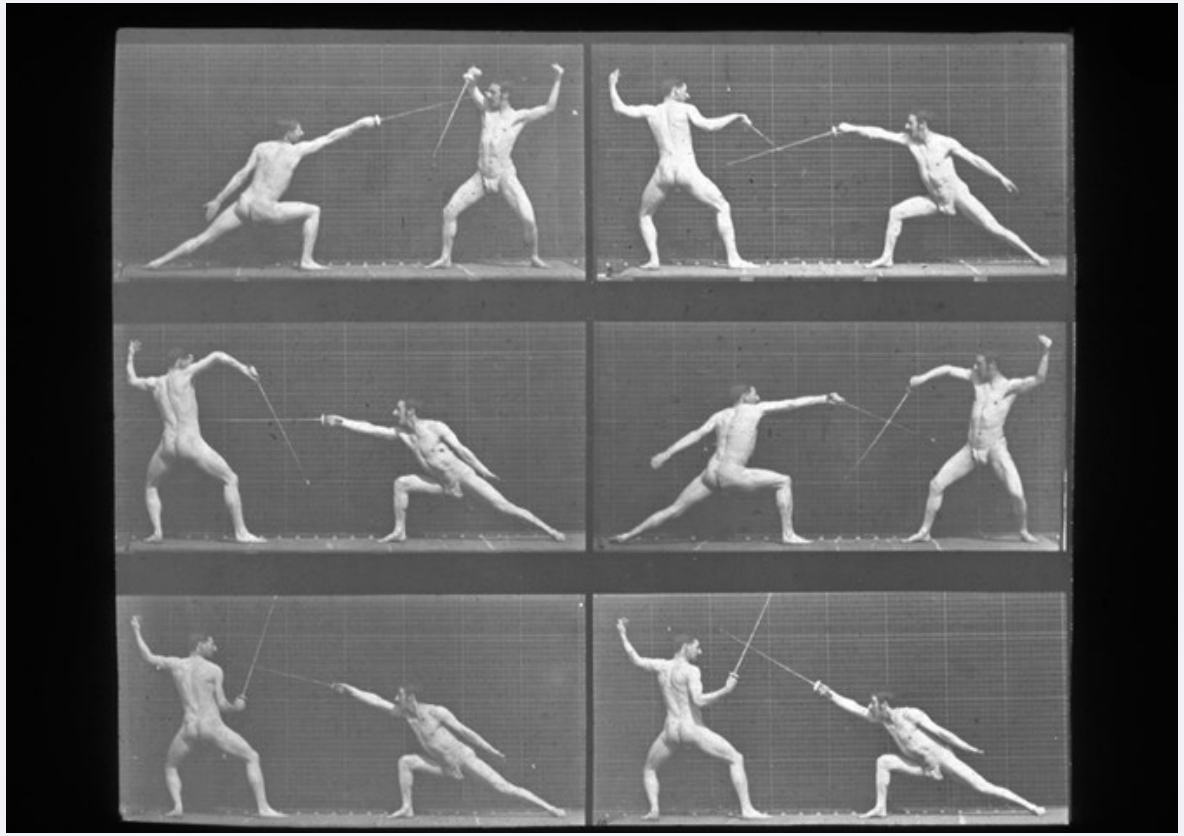
1.1. Eadweard Muybridge (1878 r.)

- Zwierzę w ruchu
- Postać ludzka w ruchu
- Zoöpraxiscope

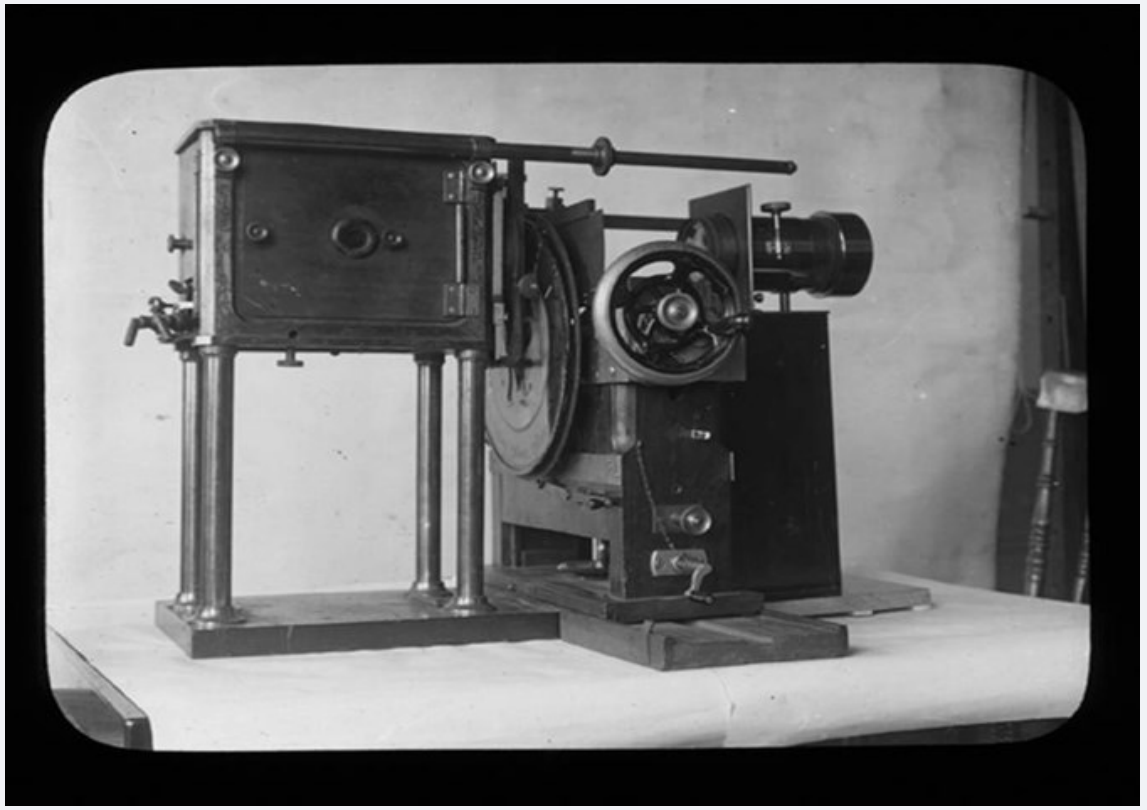


1.1. Eadweard Muybridge (1878 r.)

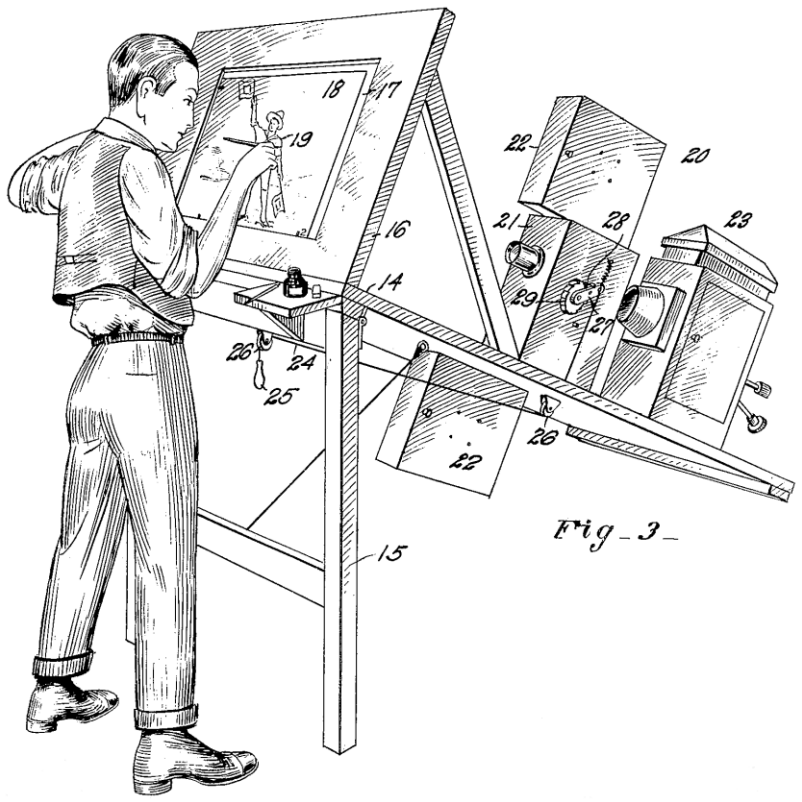




1.1. Eadweard Muybridge (1878 r.)



1.2. „Rotoskop” – bracia Maks i Dave Fleischer (1915)



Rotoskop – maszyna umożliwiająca zamianę filmu aktorskiego na film animowany poprzez ręczne, klatka po klatce, odrysowanie występujących na nim form – rotoskop wyświetlał klatki filmu na papier, gdzie były kopiowane przez rysownika. Źródło: https://pl.wikipedia.org/wiki/Rotoskop

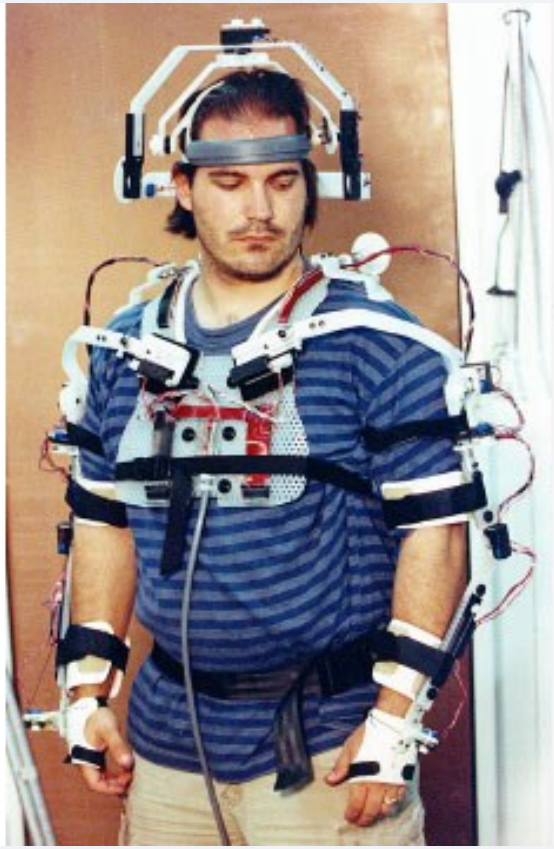
1.3. Mechaniczne przechwytywanie ruchu – lata 1950 – 1960: „Animac”

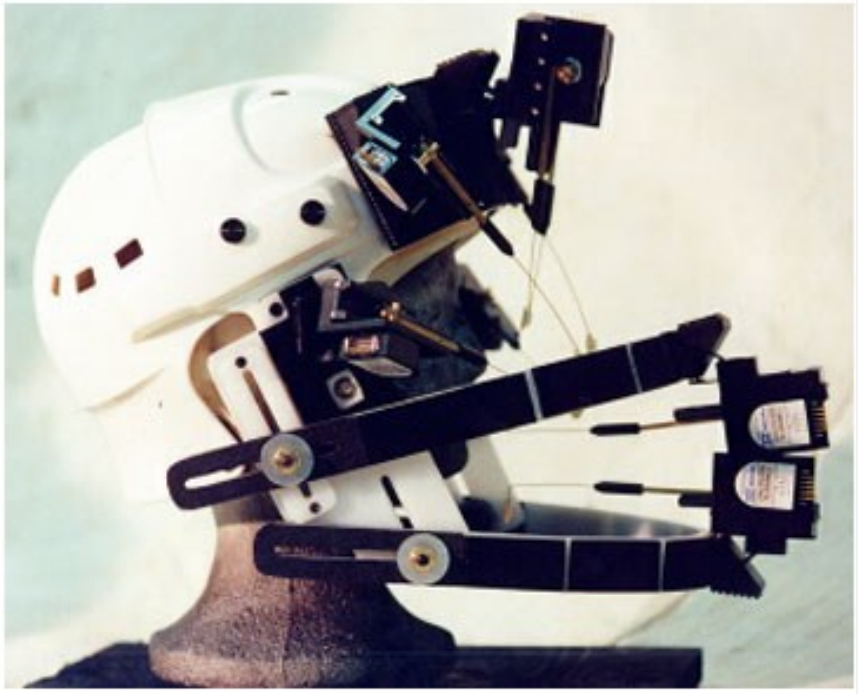
- W latach 60-tych Harrison eksperymentował z układami analogowymi oraz kineskopem (CRT) do tworzenia animowanych postaci

- System Animac wymagał ubrania aktora w **kombinezon z potencjometrami i prętami** (3), które w czasie rzeczywistym przekazywały na żywo obraz trójwymiarowy tancerza w uprzęży animacyjnej

1.4. Exoszkieletowy system przechwytywania ruchu Toma Calverta (lata 80-te XX w.)

Profesor **Tom Calvert** (kinezjolog i informatyk) – opracował exoszkielet do przechwytywania ruchu.
System składał się z mechanicznego „pancerza” oraz potencjometrów umieszczonych na każdym ze stawów, aby mogły mierzy ich zgięcie i wygięcie. Dane z wyjścia analogowych przekształcano na format cyfrowy i przekazywano do systemu animacji komputerowej.





Komentarz: W XXI wieku nadal korzystamy z systemu potencjometrów, zwłaszcza w produkcjach nisko budżetowych. (2)

1.5. Optyczne przechwytywanie ruchu (D. R. Maxwell i C. Ginsberg; MIT): Graficzna marionetka – lata 80-te

Marionetka graficzna wykorzystywała system o nazwie **Op-Eye**, opierający się na podczerwonych diodach LED podłączonych do ubrania w miejscach stawów i innych elementów kostnych.
System dwóch kamer przechwytywał ruch diod. (1)

Komentarz: *Udoskonalone systemy optyczne są powszechnie stosowane w większości współczesnych systemów optycznego do przechwytywania ruchu postaci.*

1.6. Komercyjny, mechaniczny system przechwytywania ruchu: Waldo C. Graphic (1989)(1)

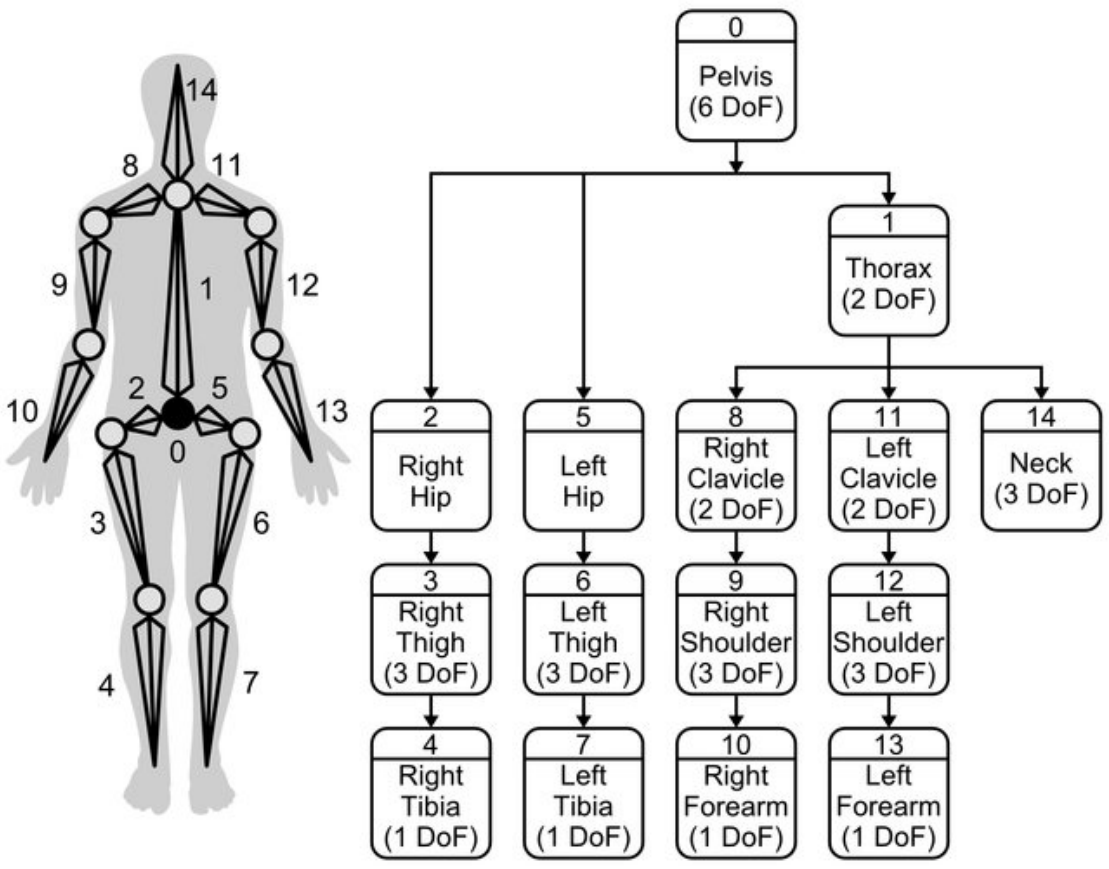
Postać ta jest uważana za pierwszą prawdziwą postać motion – capture, **stworzona dla ogólnodostępnej telewizji**. W systemie **Waldo C. Graphic,** lalkarze wchodzą w interakcję z animowaną postacią w czasie rzeczywistym.
Mechaniczna marionetka, połączona ze znacznikami i czujnikami, dzięki którym komputer rejestrował położenie ręki aktora i następnie generował animację komputerową (TV) w oparciu o trójwymiarowy model wirtualny. (2)(3)

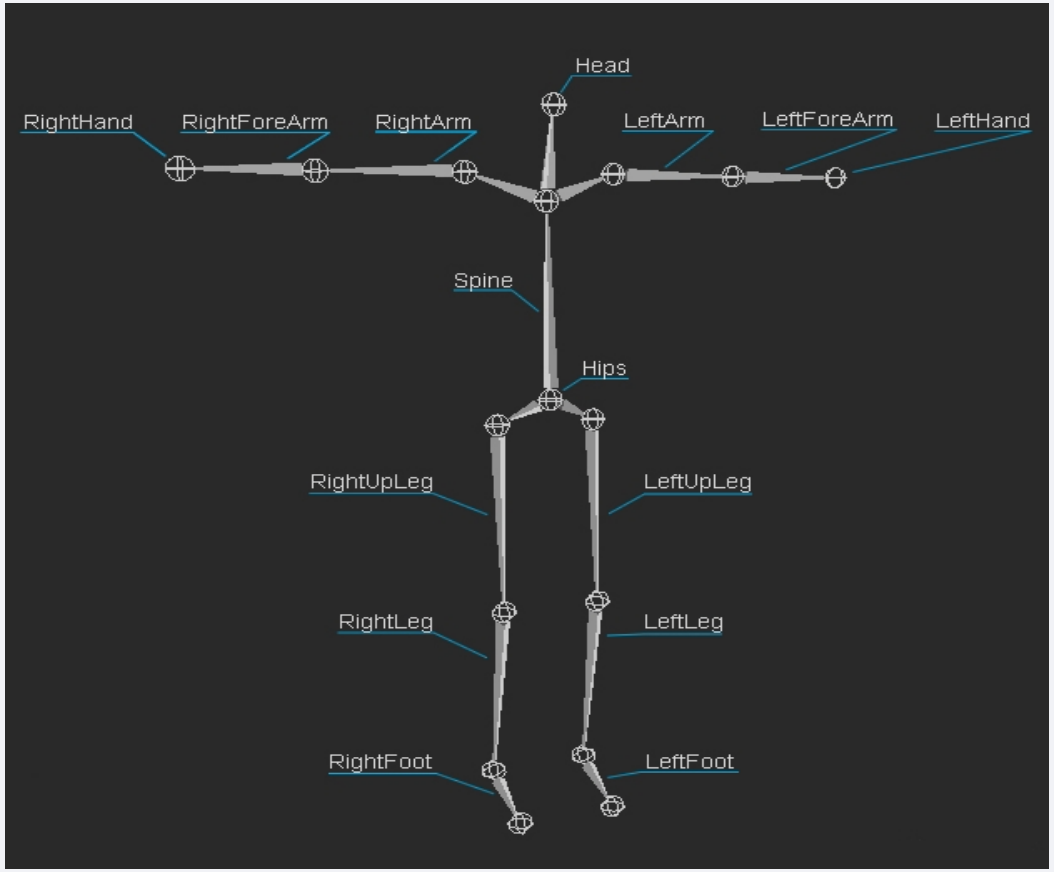
Komentarz: *Ze względu na ograniczone możliwości komputerów uzyskiwano postacie o niskiej rozdzielczości - ze względu na niską wydajność komputerów generowanie 2 min. Animacji wymagało 120 h rendering.*

2. Przegląd współczesnych systemów przechwytywania ruchu

- *Hierarchiczny model szkieletu człowieka*
- *Systemy optyczne (pasywne i aktywne)*
- *Systemy nieoptyczne*

2.1. Hierarchiczny model szkieletu człowieka (1)





Współcześnie, podstawowa wersja szkieletu składa się najczęściej z 15 stawów. Są to najważniejsze stawy, które są odpowiedzialne za ruch człowieka.

2.1. Modele szkieletu człowieka (1)

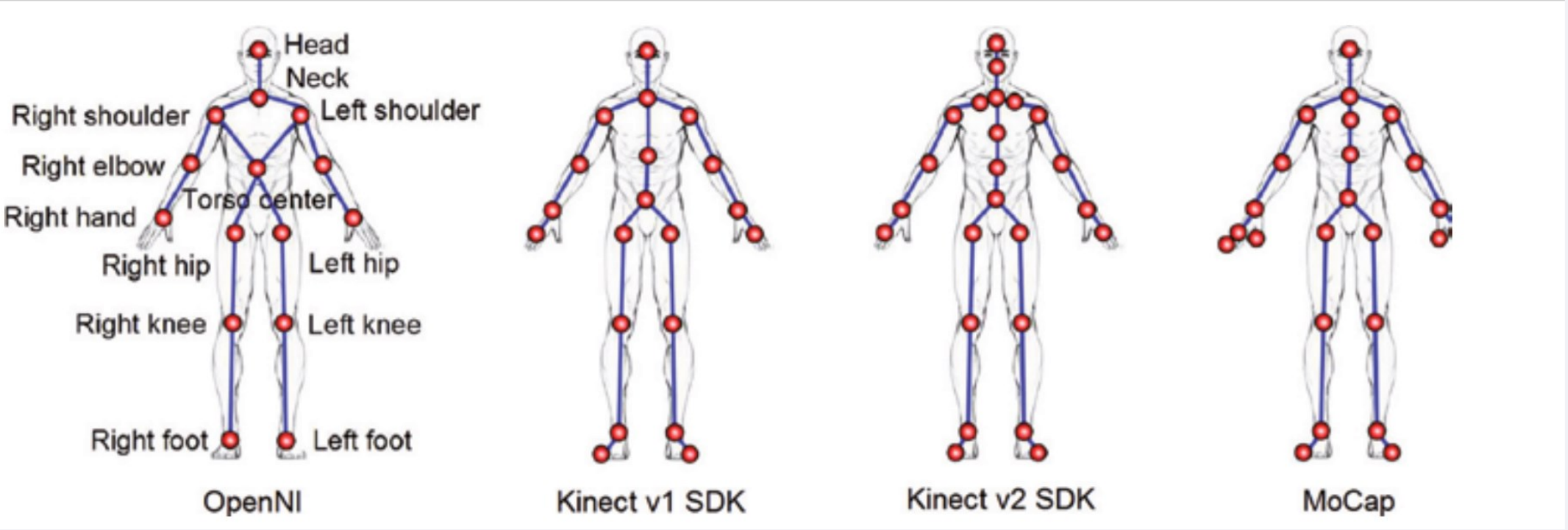
Budowane modele szkieletów człowieka opierają się o zasadę hierarchii, to znaczny wybieramy centralny punkt z którym wiążą się poszczególne elementy szkieletu, może to być miednica lub też aktywne ramię człowieka.

Szkielet wykorzystywany w przetwarzaniu danych jest uproszczoną wersją prawdziwego ludzkiego szkieletu.

Zazwyczaj jest on modelowany za pomocą łańcucha kinematycznego, który składa się ze zbioru stawów połączonych segmentami - kośćmi. Może on być postrzegany jako graf - drzewo, w którym krawędzie reprezentują kości, a wierzchołki reprezentują stawy.

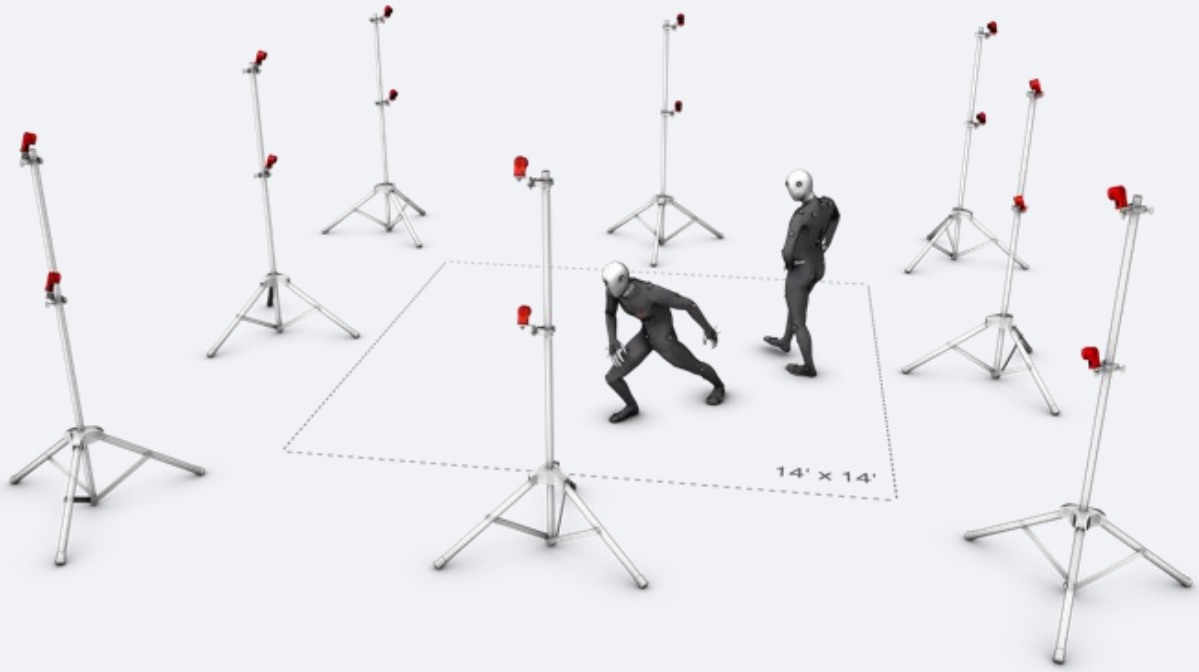
Korzeniem drzewa jest zazwyczaj staw umiejscowiony na środku miednicy. Na podstawie pomiarów ruch człowieka można sformalizować jako sekwencję postaw lub zestaw trajektorii stawów.

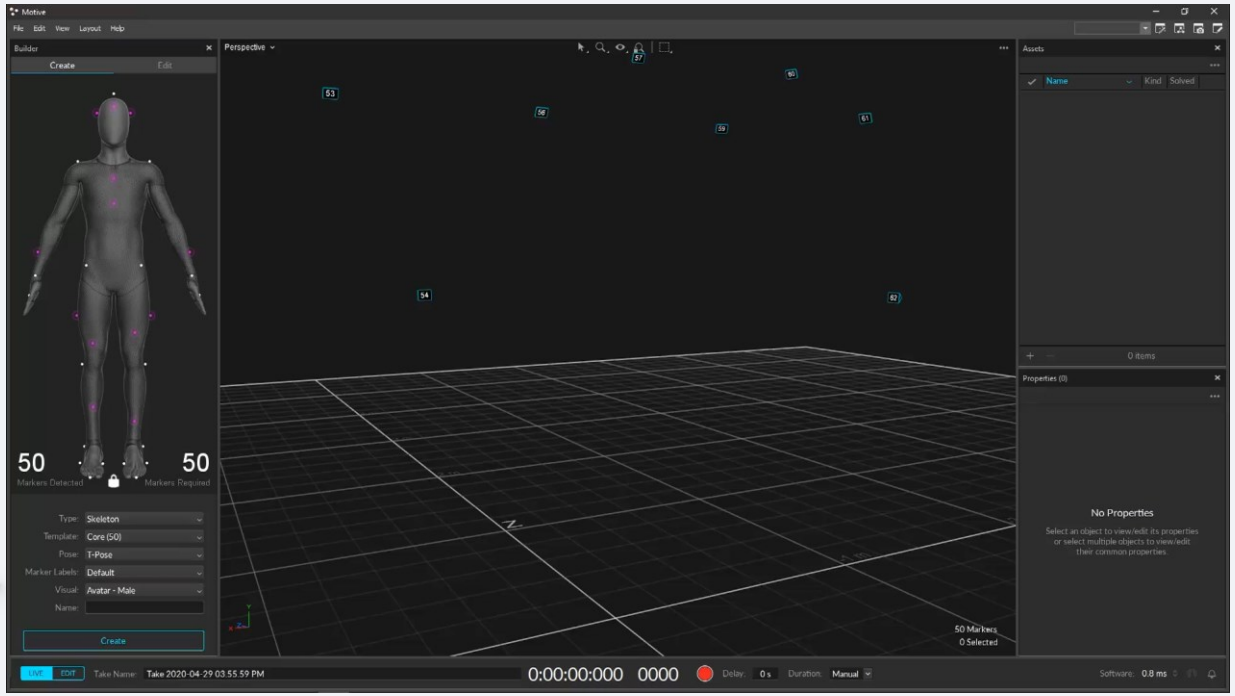
2.1. Hierarchiczny model szkieletu człowieka



Przykładowe modele o różnym stopniu szczegółowości szkieletów uzyskane z różnych urządzeń.

2.2. Optyczne systemy przechwytywania ruchu – pasywne i aktywne





Optyczna metoda motion – capture wymaga wcześniejszego ustawienia optycznych znaków na obiekcie, a następnie rejestracji przestrzennej położenia znaczników przez szybką kamerę.

Rozróżnia się znaczniki aktywne (np. diody IR Signature Pulse) i pasywne (odblaskowe).

Komentarz: *dla przechwycenia ruchu całego ciała buduję się systemy kamer (od 8 do 16)*

2.2. Pasywne optyczne systemy przechwytywania ruchu

Optyczne systemy pasywne – markery odblaskowe z możliwością regulacji natężenia odbijanego światła w celu generacji szkieletu postaci.

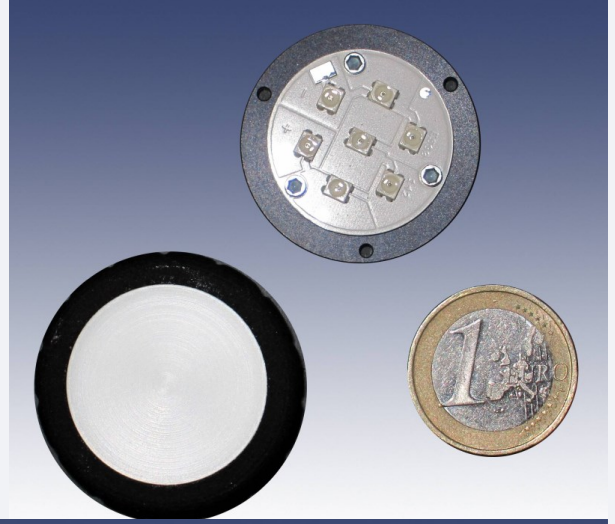
Komentarz:
Pasywne systemy nie wymagają noszenia przez aktorów przewodów i/lub sprzętu elektronicznego.

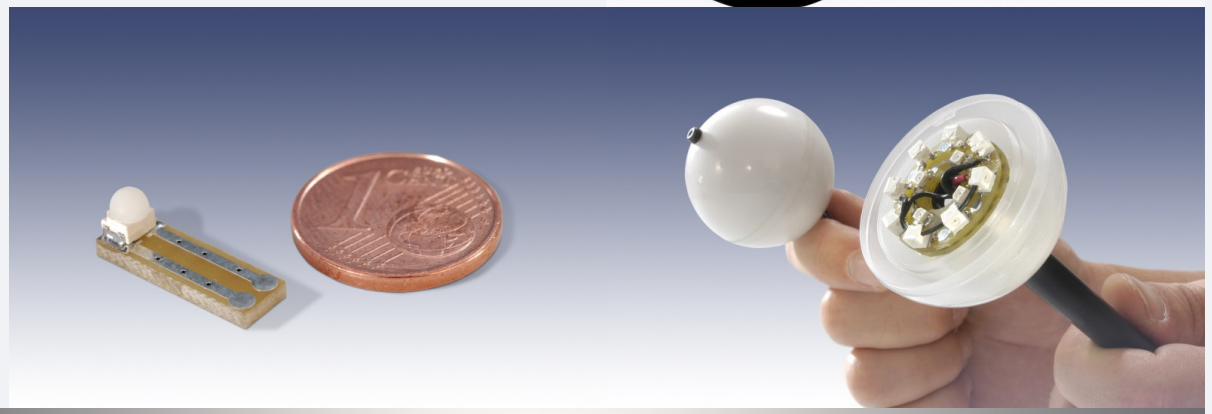
2.2. Aktywne optyczne systemy przechwytywania ruchu

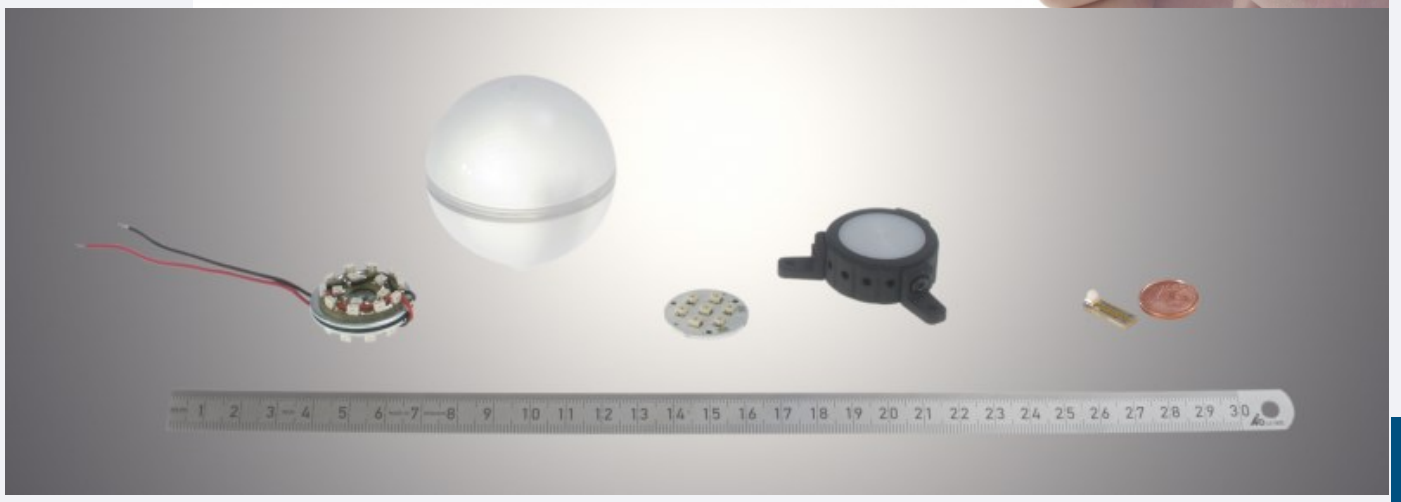
Systemy optycznej rejestracji ruchu korzystają z impulsowych diod LED, emitujących światło podczerwone (diody elektroluminescencyjne).

Diody montuję się w określonych puntach ciała.

Dzięki przypisaniu znacznikom określonych kolorów, uzyskuję się bardziej precyzyjnie nagranie ruchu postaci, bez błędów wynikających z przesłaniania markerów.







Komentarz: *Systemy aktywne oparte na impulsowych diodach LED zapewniają wyższą wydajność w czasie rzeczywistym.*

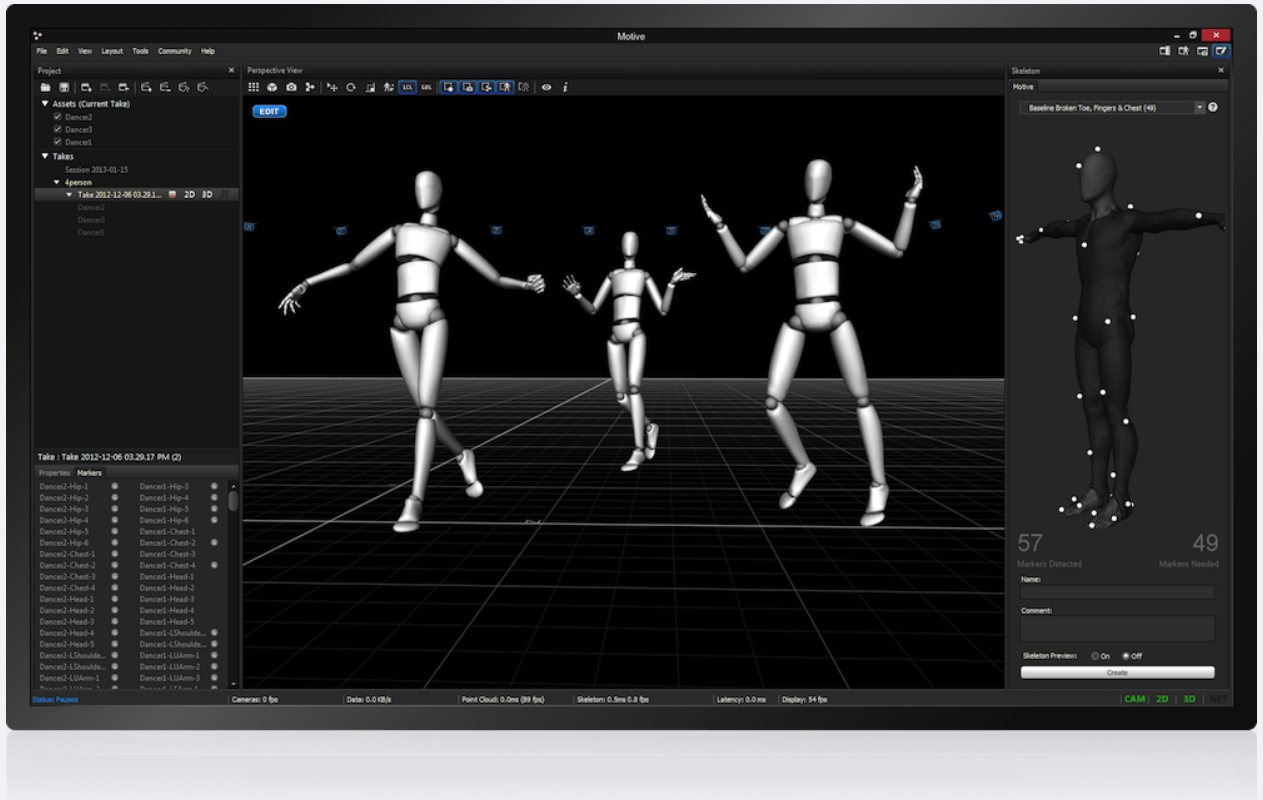
2.2. Optyczne systemy przechwytywania ruchu – wady i zalety

Zalety:

- czyste i precyzyjne systemy rejestracji ruchu
- duża swoboda w nakładaniu znaczników - możliwość rejestracji ruchu różnorodnych obiektów
- dzięki rozbudowanym systemom kamer możliwa jest rejestracja wielu obiektów (postaci) jednocześnie

Wady:

- okluzja (przesłanianie) znaczników, zwłaszcza przy dynamicznych ruchach - prowadzi do rozbudowanego i czasochłonnego „post – processingu”.
- szumy, dot. markery pasywnych - konieczność odpowiedniej kontroli oświetlenia sceny



2.3. Nieoptyczne systemy przechwytywania ruchu

Opierają się na montowaniu, tzw. **inercyjnych czujników, mierzących położenie obiektu (w tym zmiany kątów) oraz zmiany szybkości ruchu** (żyroskopy, magnetometry i akcelerometry) - pomiar prędkości obrotowych).

Są to czujniki **aktywne** przekazujące pomiary (fale elektro-magnetyczne **bezpośrednio do komputerów**. Przechwytują ruch w najtrudniejszych środowiskach, nawet z zakłóceniami magnetycznymi.

2.3. Nieoptyczne systemy przechwytywania ruchu

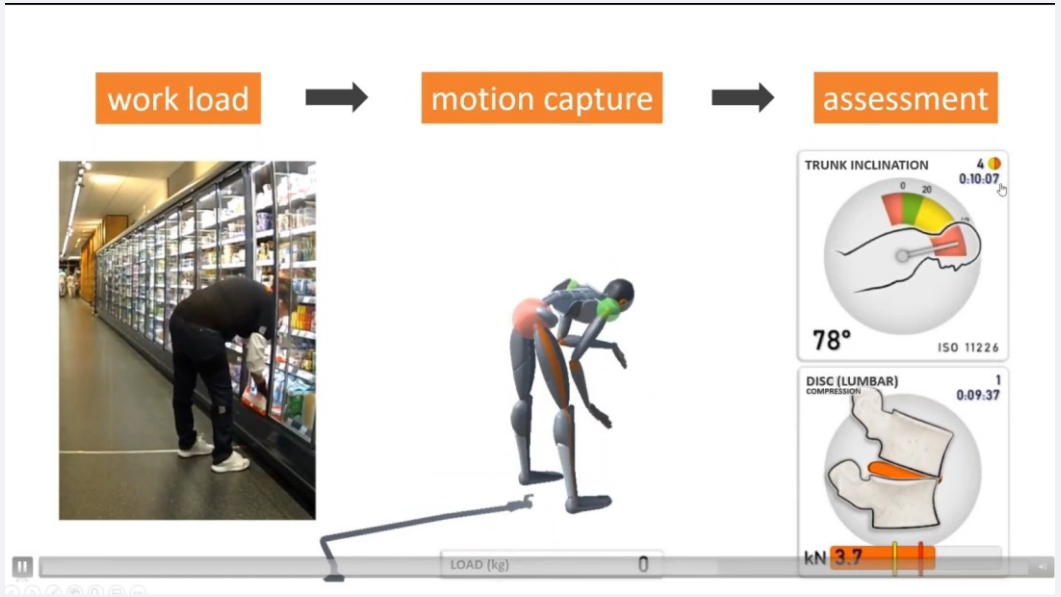
Zalety:

- przechwytywanie ruchu w różnych środowiskach: w tym w ciasnych przestrzeniach,
- przenośność i duże obszary przechwytywania,
- eliminują konieczność używania kamer

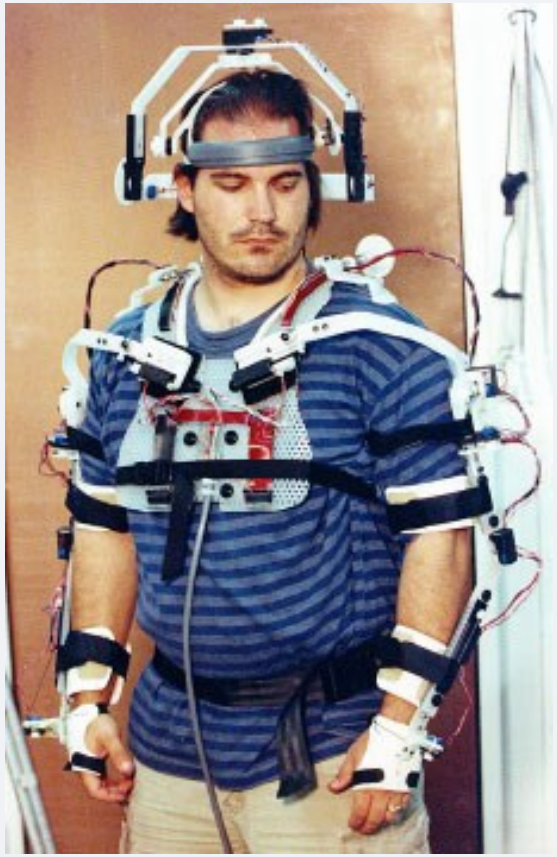
Wady:

- mniejsza dokładność pozycjonowania i dryf pozycyjny, który może się narastać z upływem czasu.

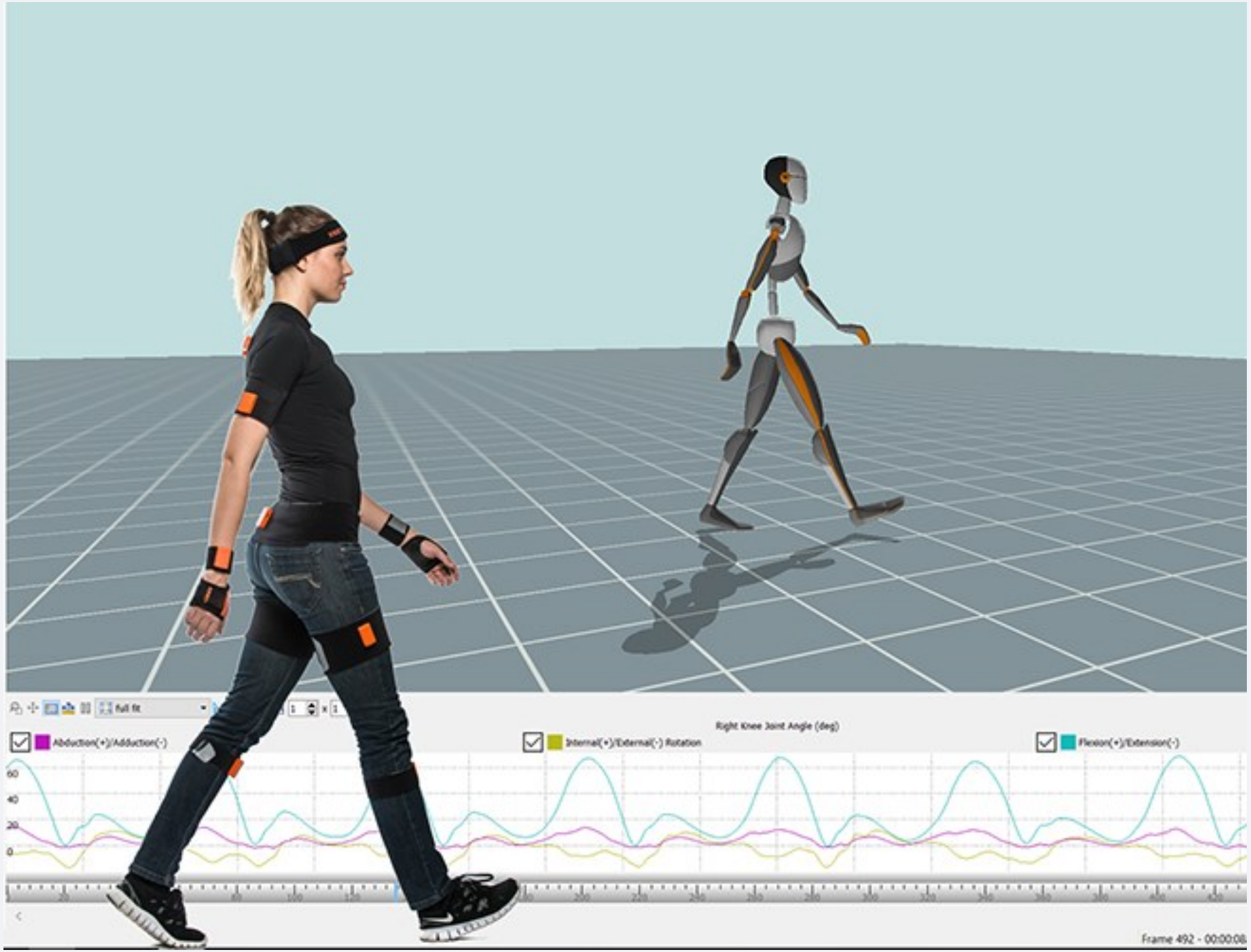
Komentarz: *Stosuję się algorytmy komunikacyjne zapewniające wysoką jakość przechwytywania ruchu, nawet środowiskach z zakłóceniami elektro-magnetycznymi.*



2.3. Współczesne egzoszkielety do przechwytywania ruchu



- Systemy elektro-mechaniczne
- Systemy magnetyczne
- Czujniki rozciągania

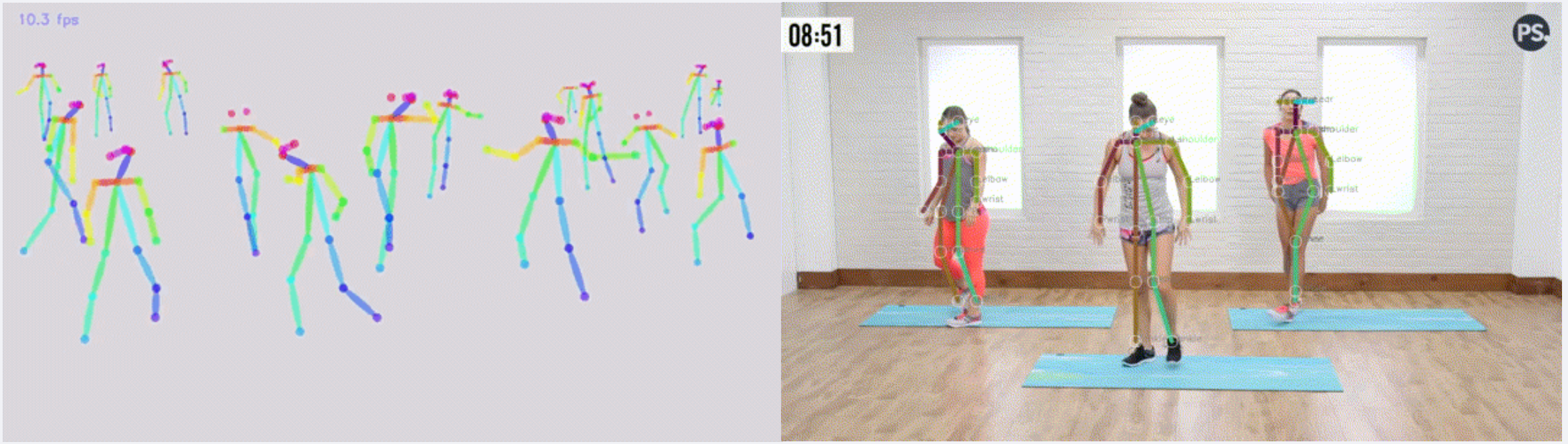


Dzięki „czujnikom" uzyskujemy komputerowy model (szkielet) poruszającego się człowieka.

3. Przechwytywanie ruchu z wykorzystaniem metod AI

- *Szkielet pozy człowieka - Open Pose*
- *GHUM & GHUML: generatywne modele 3D postaci człowieka i jej pozy przegubowe*
- *Przechwytywanie ruchu bez specjalistycznego sprzętu i danych (przykłady aplikacji)*

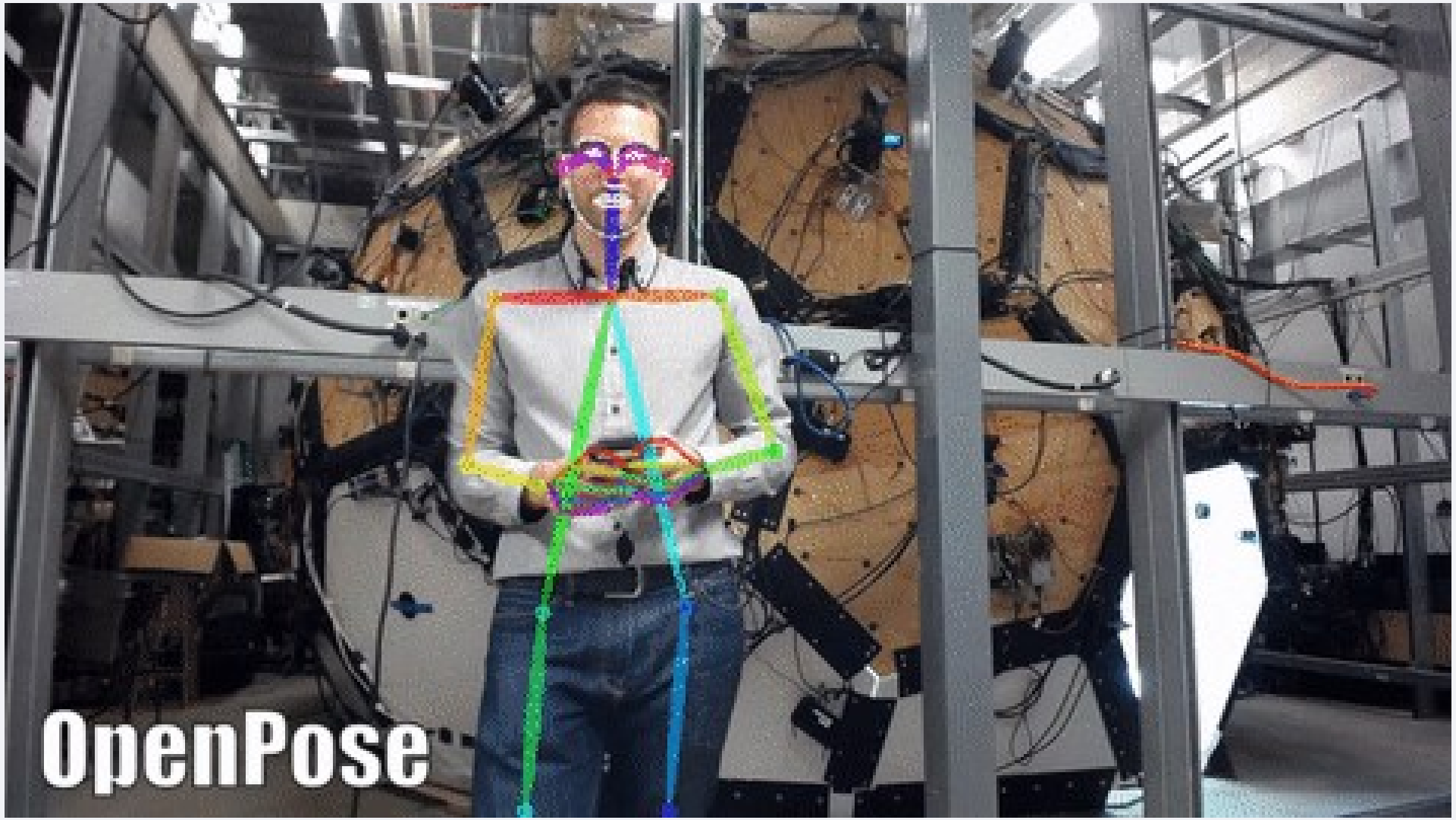
3.1. Szkielet i pozy człowieka - Open Pose 2D



Technologia Open Pose służy do generowania układów kości i stawów na płaszczyźnie 2D, w oparciu o wideo lub zdjęcia.

Technologia ta wykorzystuję przechwytywanie ruchu na płaszczyźnie (na ekranie komputera lub smartphone'a) i opiera się o wyznaczenie odpowiedniej liczby tzw. punktów kluczowych (układ kości i stawów) współpracujących ze sobą w trakcie ruchu. Opis ciała człowieka wymaga pomiędzy 15-27 punktów; rąk / dłoni - 21 punktów; twarz – 70.

System wykorzystywany jest w różnych dziedzinach, takich jak rozpoznawanie akcji, bezpieczeństwo, sport i inne.



System wykorzystywany jest w różnych dziedzinach, takich jak rozpoznawanie akcji, bezpieczeństwo, sport i inne.

Komentarz: *W bibliotece Open Pose punkty kluczowe poruszają się na jednej płaszczyźnie.*

3.2. Generatywna postać ludzka 3d i przegubowe modele naśladujące pozy.

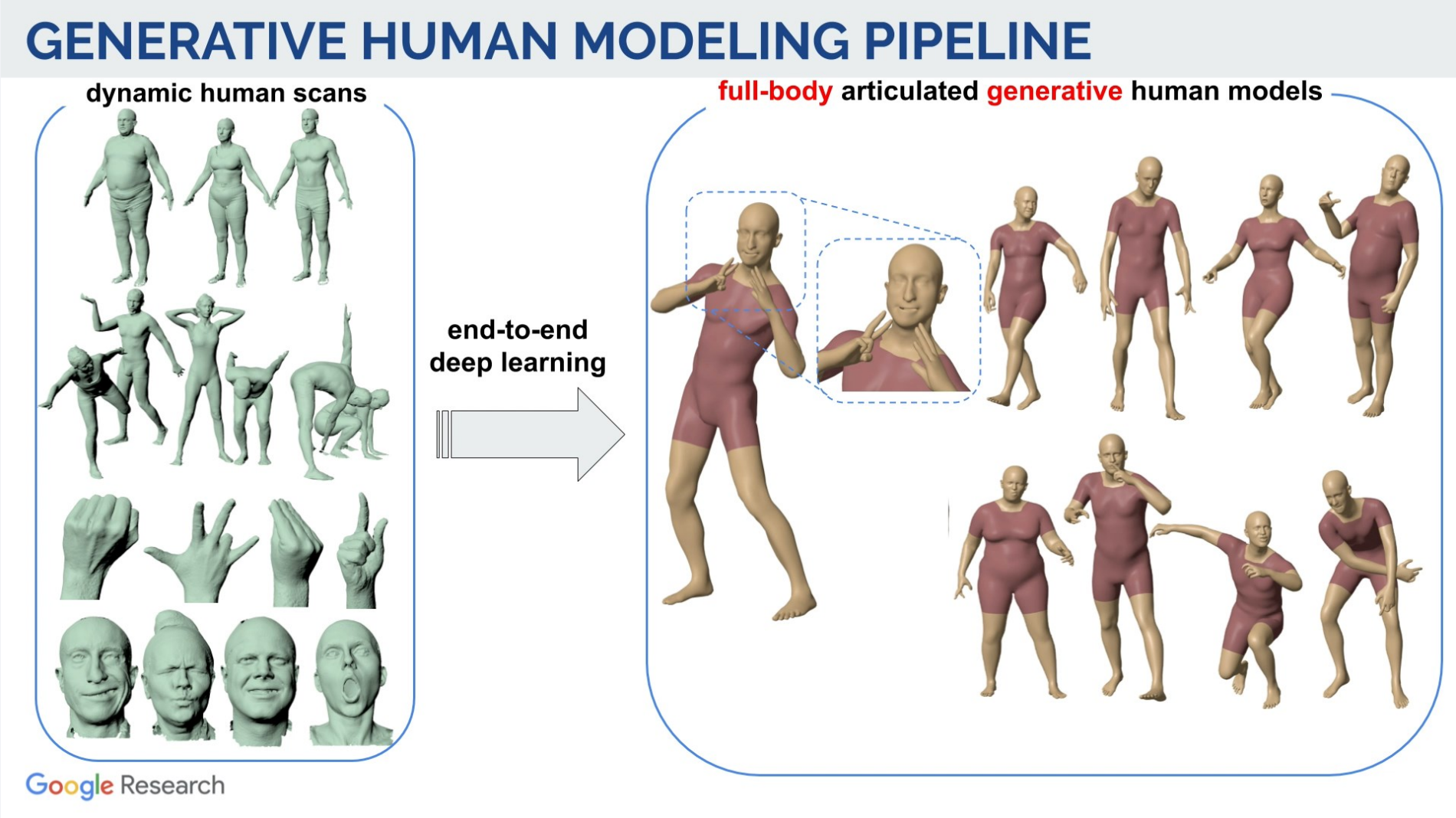
System GHUM realizujący machine-learning w Convolution Neural Networks (CNN) – Google Research

- Choć z czasem skonstruowano wiele ekspresyjnych modeli twarzy, rąk i ciał, to większość z nich - jeśli nie wszystkie, była tworzona osobno, a nie w kontekście pełnego ciała ludzkiego.

- Estymacja postawy człowieka na podstawie obrazu wideo odgrywa kluczową rolę w różnych zastosowaniach, takich jak kwantyfikacja ćwiczeń fizycznych, rozpoznawanie języka migowego i sterowanie gestami całego ciała.

3.2.1. Generatywne modelowanie człowieka (3D) – założenia, cele

Generowanie modelu 3D ludzkiej sylwetki z uwzględnieniem cech charakterystycznych wzorca (tusza, płeć itp.), jego aktualnej pozycji w realizowanej sekwencji ruchów, a także ekspresji (twarz, dłonie).



3.2.2. Generatywne modelowanie człowieka: Wcześniejsze prace, próby

- **SMPL** (modele 3d postaci człowieka)
- **FLAME** (modele 3d głowy)
- **MANO** (modele 3d ciała i rąk)

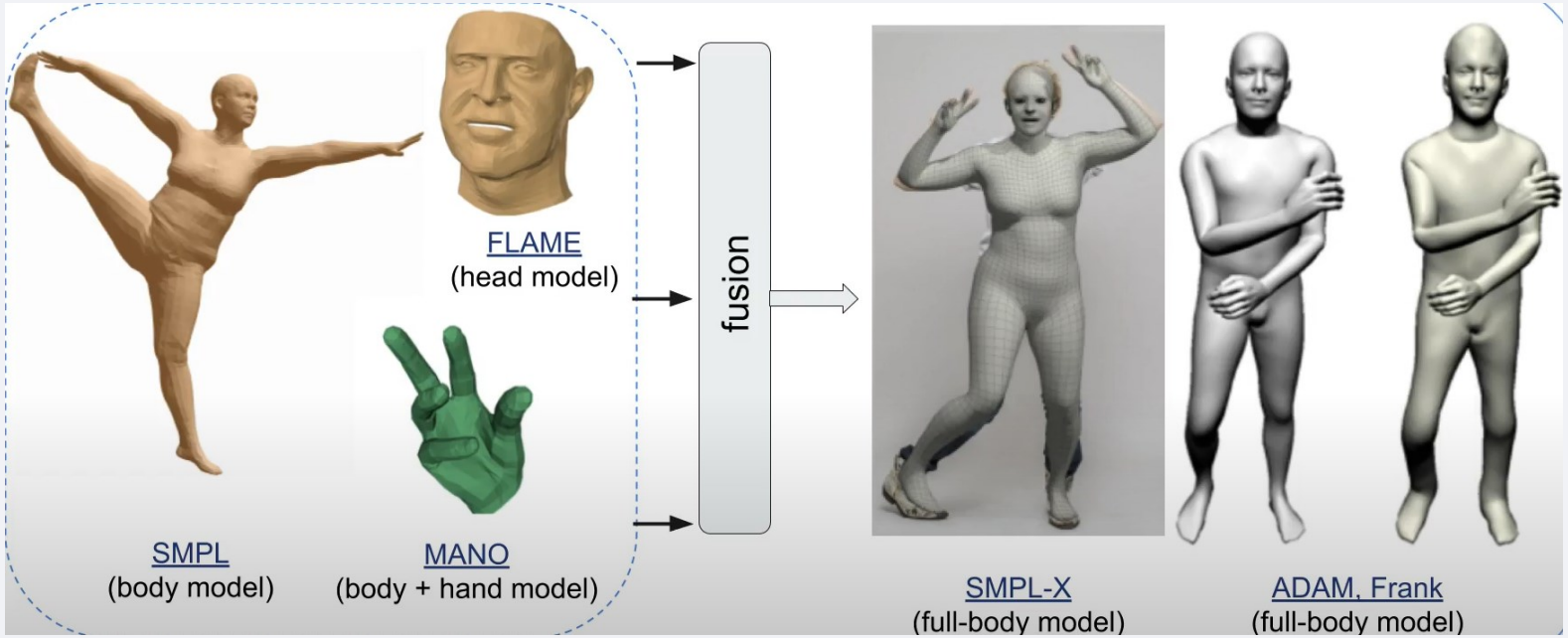
FUZJA

- **SMPL-X**

(model 3d całego ciała człowieka)

- **ADAM, Frank**

(model 3d całego ciała człowieka)

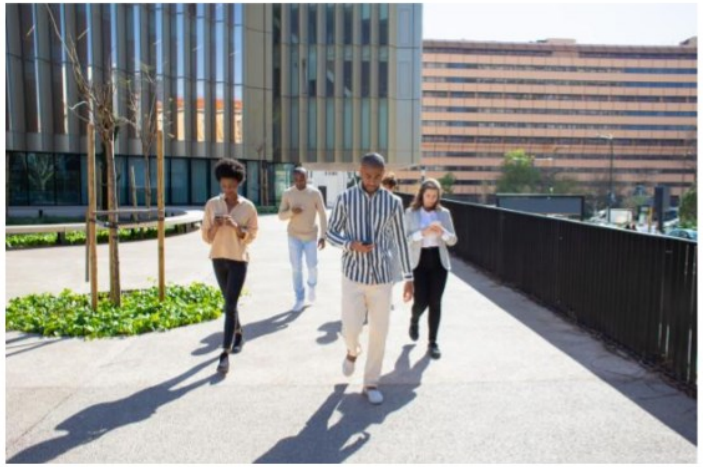


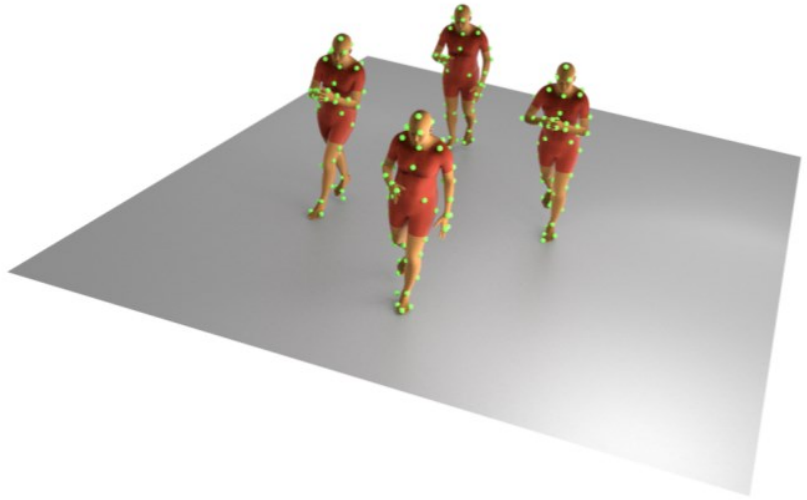
3.2.3. Generatywne modelowanie człowieka 3d : Uzasadnienie podjęcia tematu badań

Rozpoznawanie człowieka i jego ruchu w środowisku 3d

Możliwe zastosowania, aplikacje:

- Fitness
- Medycyna
- E-shopping
- Wirtualna rzeczywistość
- Rozszerzona Rzeczywistość
- Rozrywka
- Sport
- Analiza ruchu

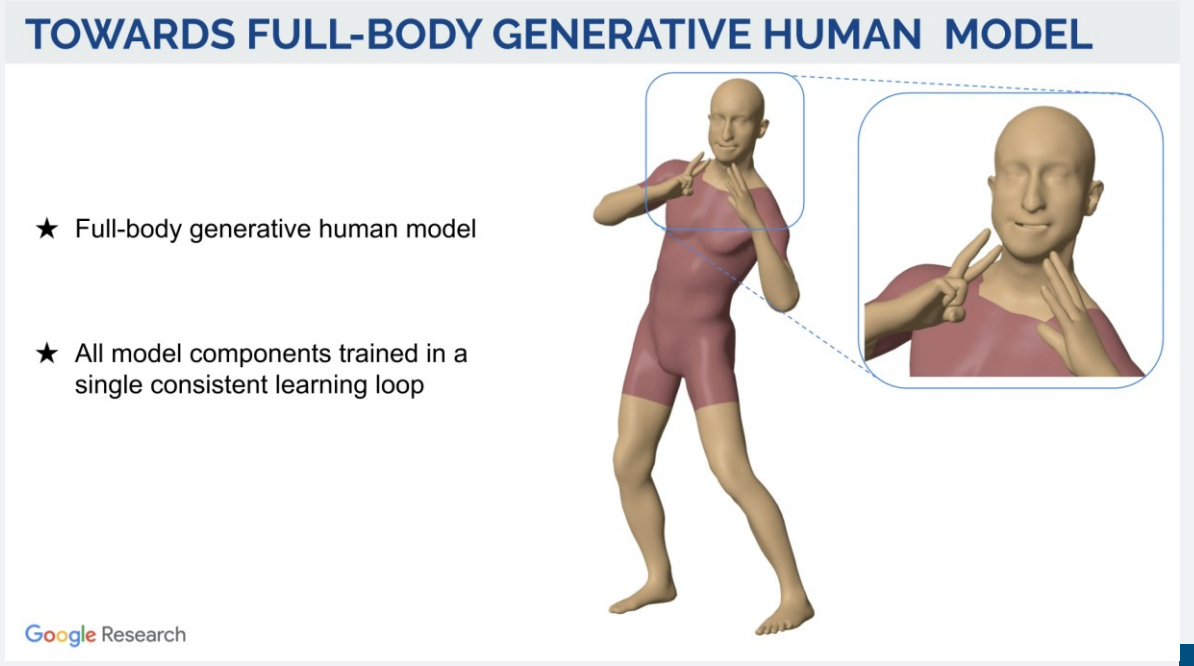




3.3. GHUM & GHUML – kolejność omawiania

Podstawowy założeniem jest generowanie modelu 3d ludzkiej postaci oraz jej ruchów.

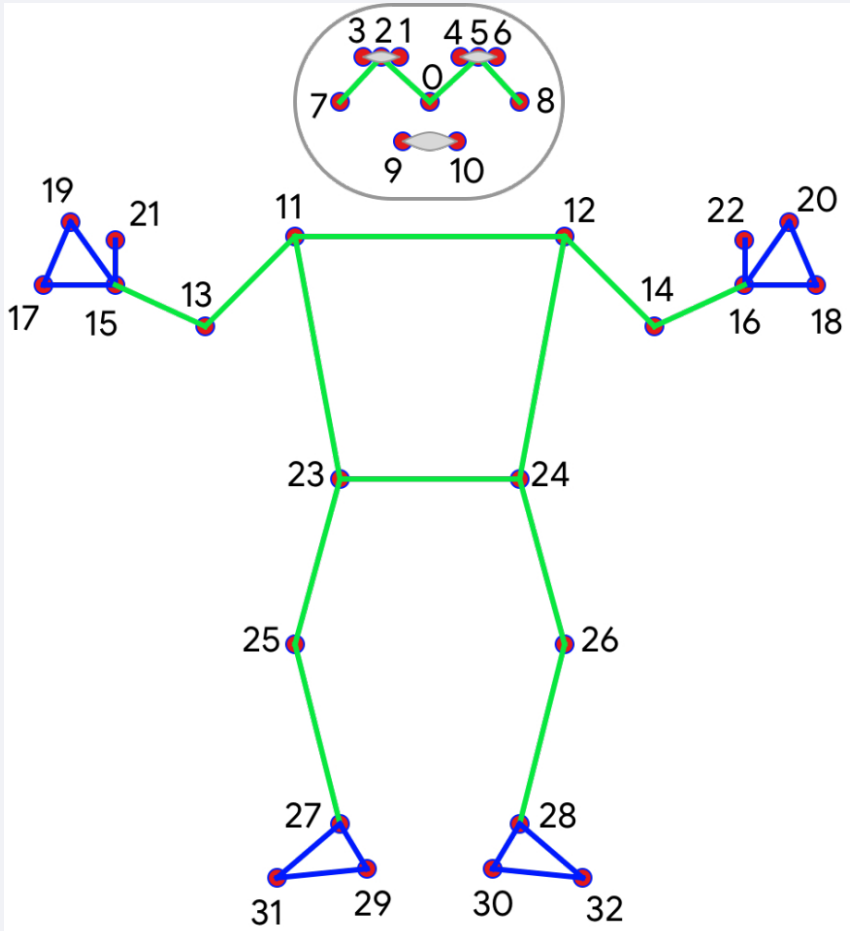
- Szkielet - podstawa modelu
- Sieć neuronowa, głębokie uczenie, bazy danych
- Możliwości systemu, np. interpolacja sylwetki
- Wersję szablonów
- Ewaluacja modelu
- Porównanie z innym oprogramowaniem
- Inne aplikacje



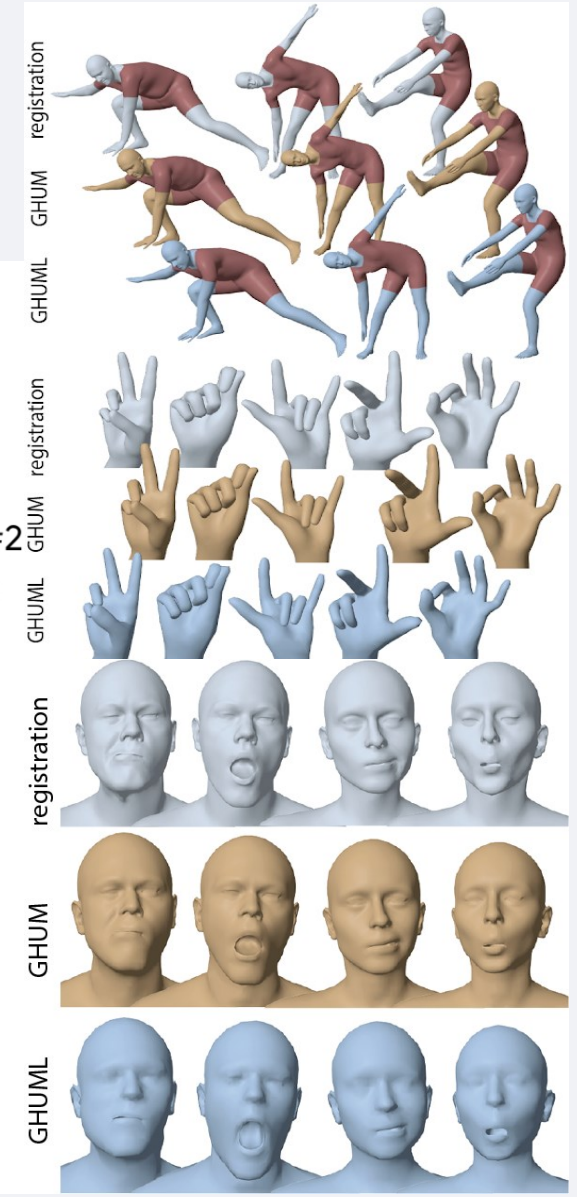
3.3.1. GHUM & GHUML - definiowanie modelu podstawowego

Wyróżnia się następujące, istotne elementy ruchu i ekspresji modelu sylwetki człowieka:

- ciało,
- szkielet,
- twarz,
- dłonie

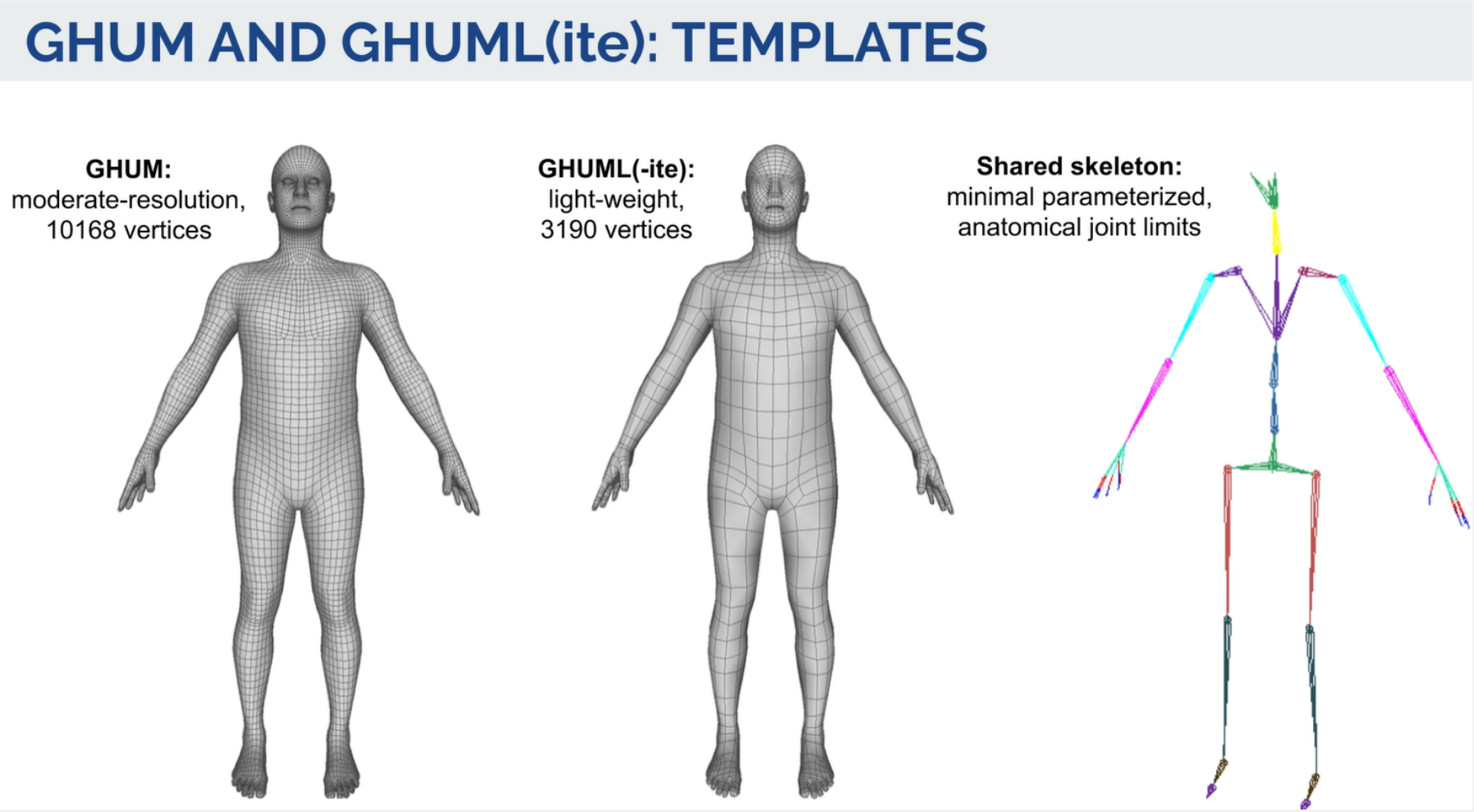


0. nose
1. right eye inner
2. right eye
3. right eye outer
4. left eye inner
5. left eye
6. left eye outer
7. right ear
8. left ear
9. mouth right
10. mouth left
11. right shoulder
12. left shoulder
13. right elbow
14. left elbow
15. right wrist
16. left wrist
17. right pinky knuckle #1
18. left pinky knuckle #1
19. right index knuclke #1
20. left index knuckle #1
21. right thumb knuckle #2
22. left thumb knuckle #2
23. right hip
24. left hip
25. right knee
26. left knee
27. right ankle
28. left ankle
29. right heel
30. left heel
31. right foot index
32. left foot index



3.3.2. Szablony: GHUM i GHUM(L). Szkielet i siatka poligonów *(mesh)*

Na szkielecie (kości i stawy) rozpina się bazową siatkę poligonową – dwie wersje modelu 3d, różniące się rozdzielczością siatki (liczbą werteksów i poligonów)



Uwaga: dokładność rekonstrukcji zależy od liczby (wierzchołków) poligonów w bazowej siatce modelu człowieka i tak:

GHUM - więcej wierzchołków ➔ lepsza jakość rekonstrukcji;

GHUML - mniej wierzchołków ➔ słabsza jakość rekonstrukcji, ale krótszy czas przetwarzania

3.3.3. GHUM – Sieć neuronowa i uczenie.

System działa w oparciu o sieci neuronowe, modelujące ruchy. Jest oparty o idee „konwolucyjnych" sieci neuronowych, w których istotą jest wykrycie elementów charakterystycznych i ich integracja.

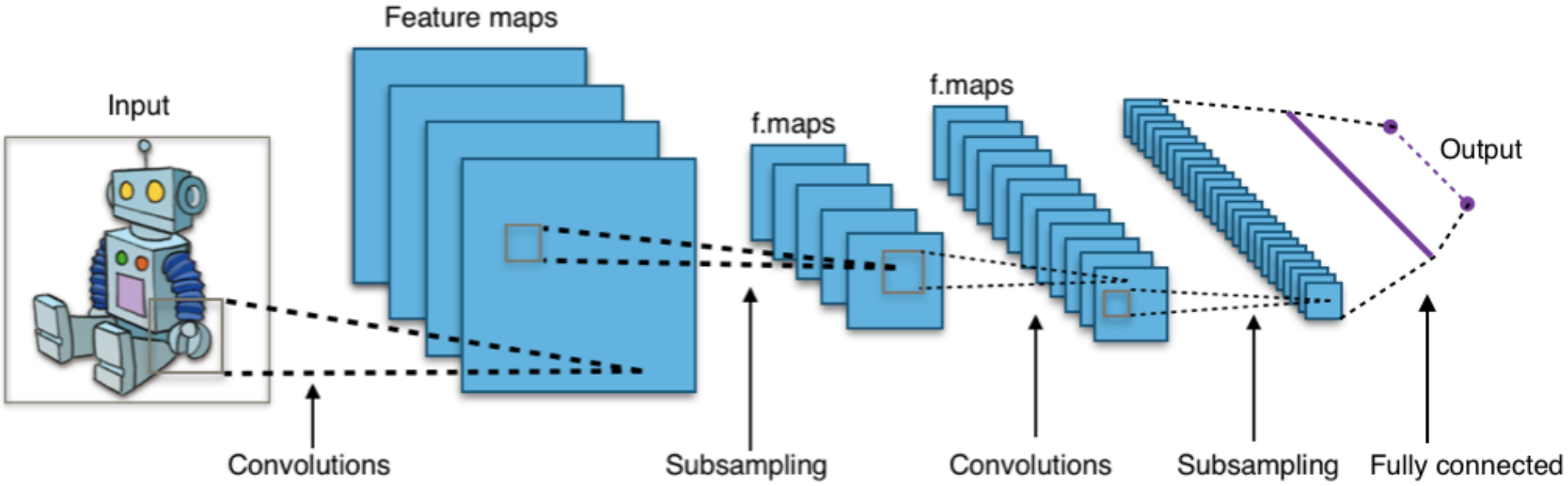
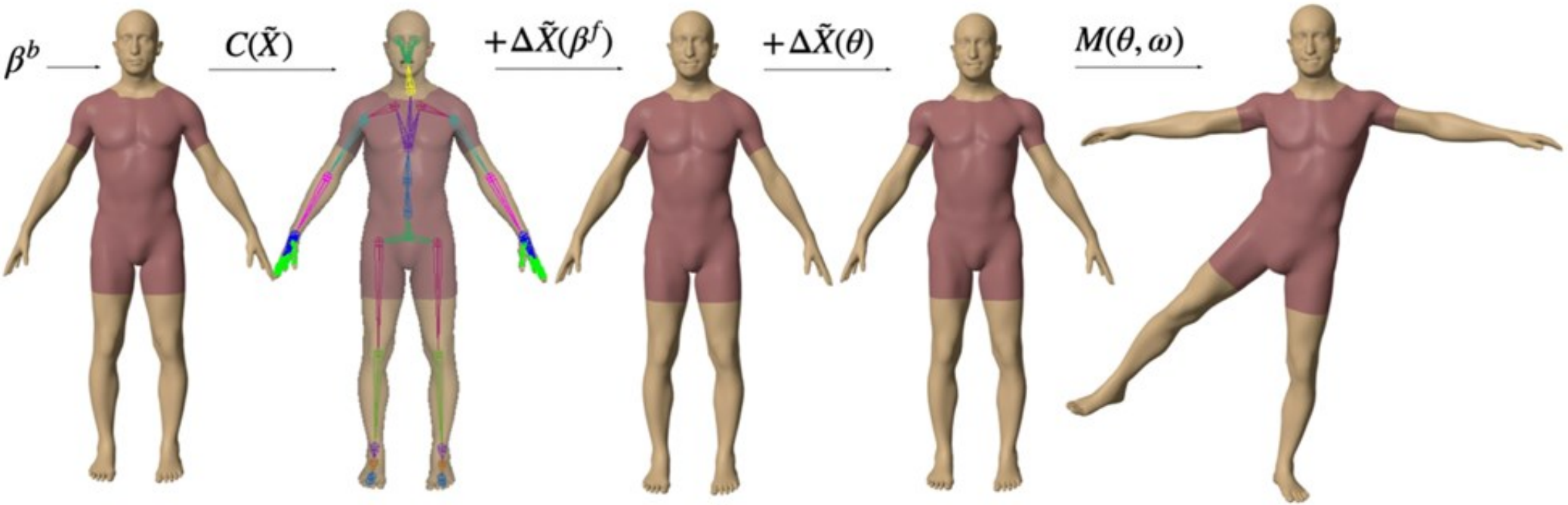


Fig. 3.1.1. Standardowa architektura sieci „konwolucyjnej" (CNN; Convolutional neural network). Źródło: https://en.wikipedia.org/wiki/Convolutional_neural_network

GHUM i GHUML(ite): Definiowanie modelu. Sieć i podsieci neuronowe dla poszczególnych elementów postaci – szkieletu, twarzy, dłoni i całości ciała – zasada uczenia

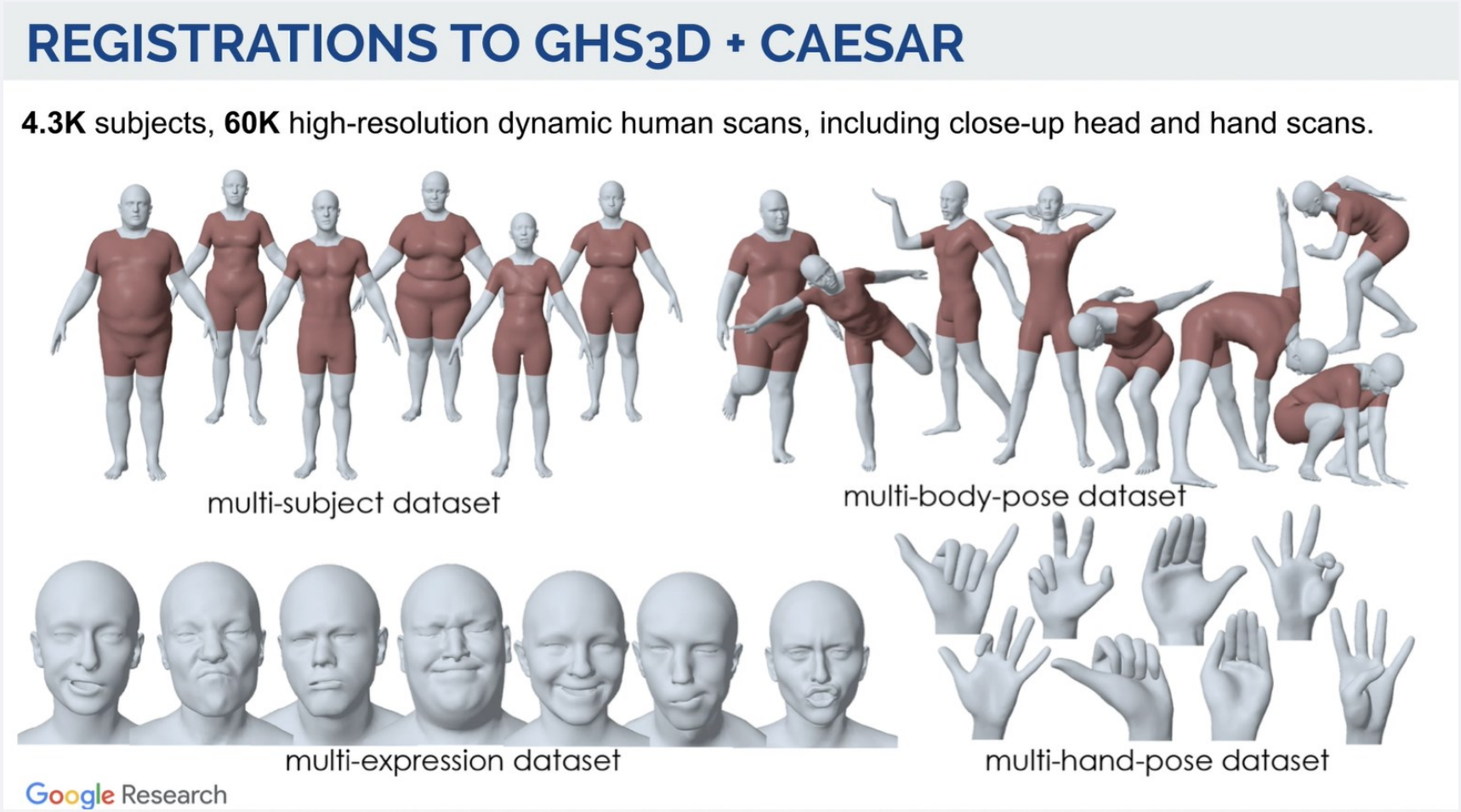
GHUM AND GHUML(ite): MODEL FORMULATION

$$X(\beta^b, \beta^f, \theta) = M(\tilde{X}(\beta^b), \Delta\tilde{X}(\beta^f), \theta, \Delta\tilde{X}(\theta), C(\tilde{X}), \omega)$$



Google Research

3.3.4. Bazy danych GHS3D + CAESAR – wzorce do naśladowania

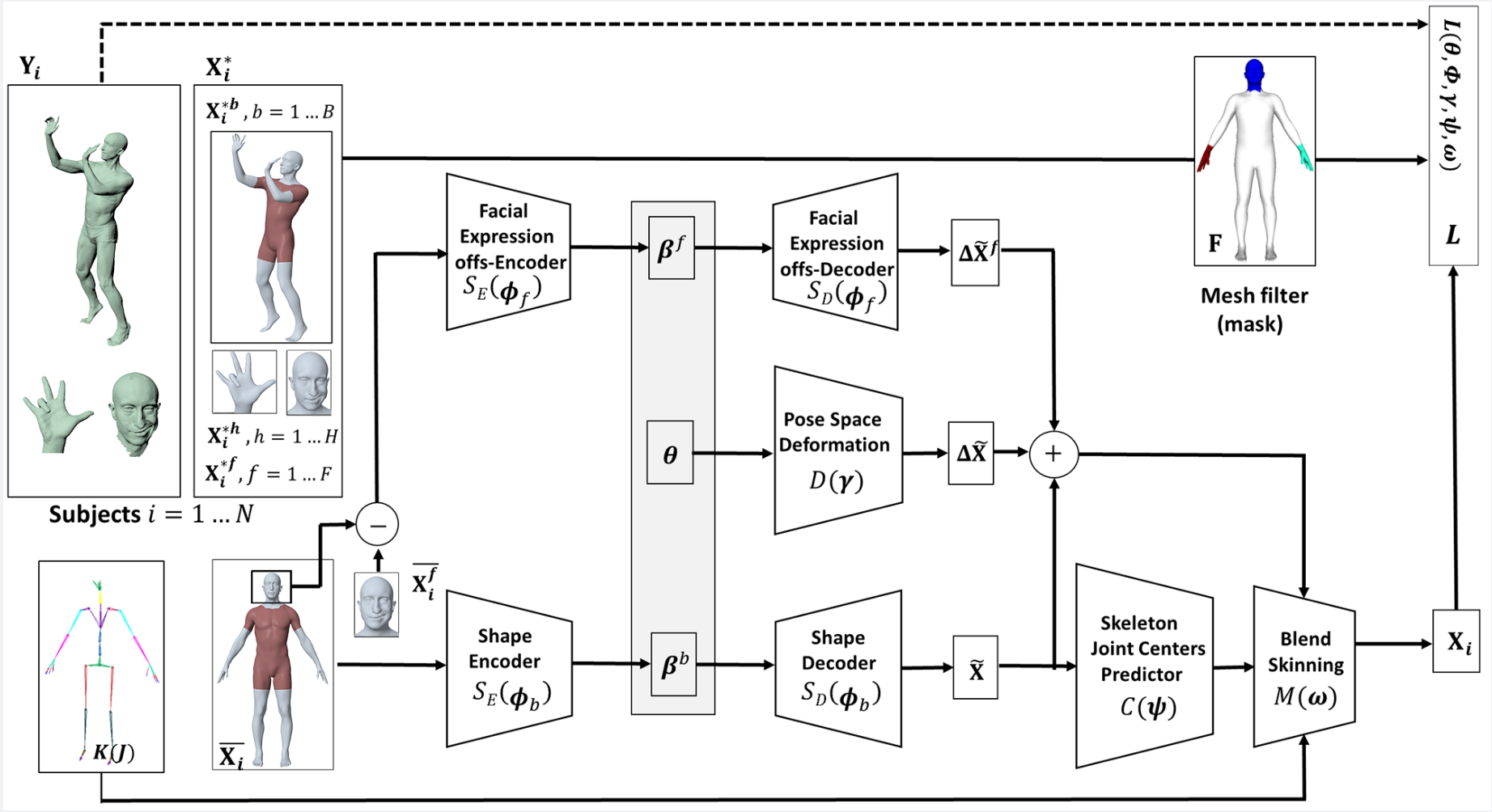


W bazach GHS3D + CAESAR zgromadzono dane dot. sylwetek, ruchów i jego faz oraz ekspresji. Dzięki tym danym w procesie deep – learningu tworzy się syntetyczne modele postaci, które następnie system adaptuję do konkretnej osoby i sytuacji ruchowej.

Deep Learning stosuję się w wielowarstwowych sieciach neuronowych (miliony neuronów i połączeń). Dzięki aktualnej mocy obliczeniowej komputerów, jesteśmy w stanie realizować w tych sieciach proces uczenia oparty o średni błąd kwadratowy oraz algorytmy wstecznej propagacji *(backpropagation)*.

3.3.5. Kompleksowy proces deep - learningu („Pipelining”)

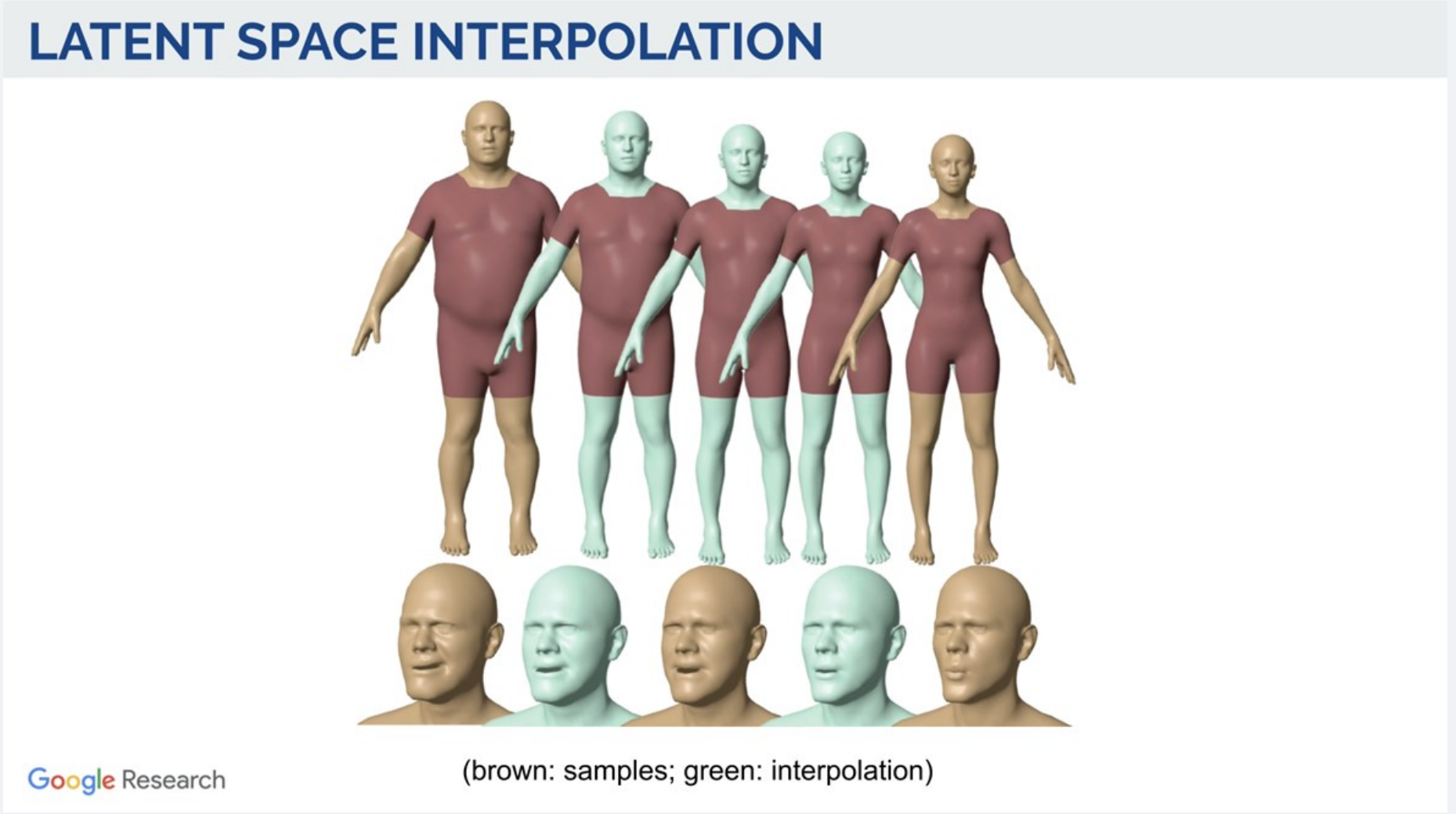
- dolny lewy róg - modelowanie sylwetki człowieka Xi na podstawie szkieletu K(J)
- lewy górny róg – na podstawie obrazów (subjects i = 1 … N) modeluje się ruchy ciała oraz ekspresję twarzy i rąk
- elementy podstawowe, są modyfikowane w kolejnych stopniach „pipelining”
- warto zauważyć trzy trapezy - enkodery których zadaniem jest deformacja odpowiednich elementów sylwetki
- Niestety autorzy nie udostępniają więcej szczegółów nt. budowy i organizacji tej sieci neuronowej oraz zasad uczenia



Schemat zasady budowy modelu człowieka oraz jego ruchów w oparciu o idee „konwolucyjnych” sieci neuronowych CNN.

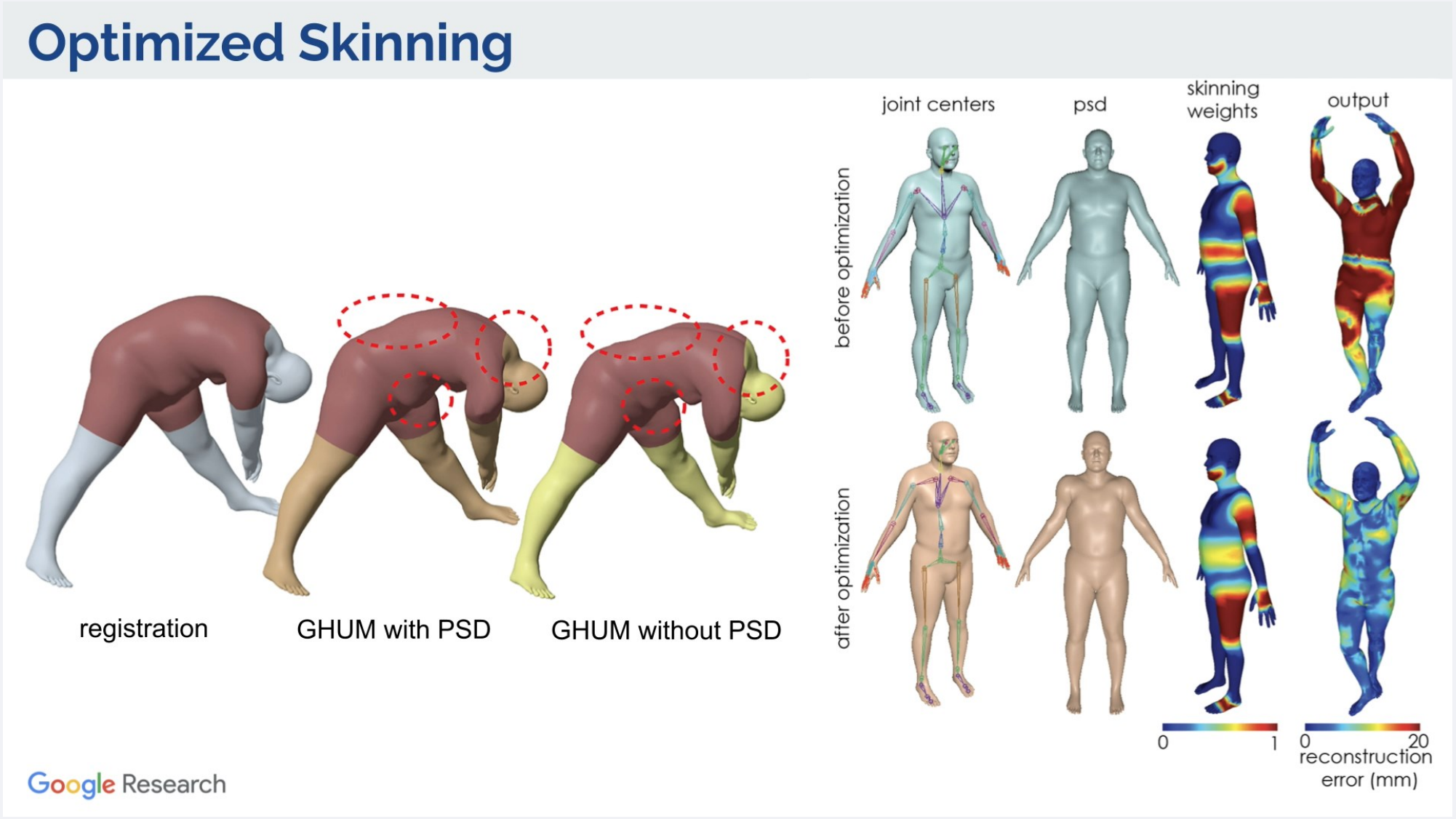
3.3.6. Możliwości systemu - formy pośrednie - interpolacja sylwetki

- System umożliwia generowanie form pośrednich ze względu na np. tuszę
- Jak również interpolację pomiędzy odmiennymi ekspresjami i mimiką twarzy



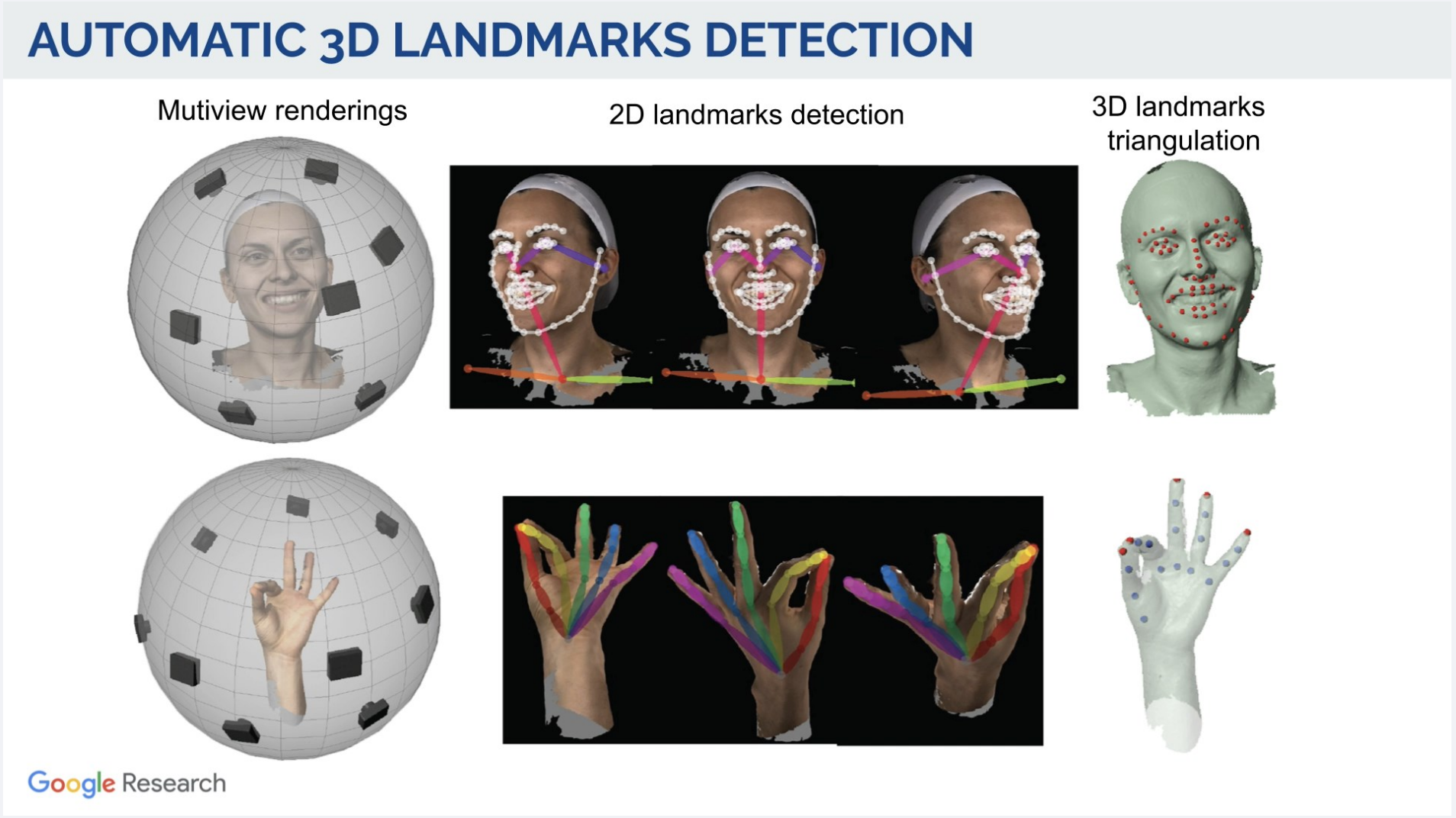
3.3.6.1. Możliwości systemu - dopasowanie skóry modelu 3d

System optymalizuje model z bazy danych pod kątem naturalnej deformacji tkanki skóry *(mesh)* sylwetki ludzkiej.



3.3.6.2. Możliwości systemu - rozpoznawanie cech (znaków charakterystycznych) twarzy i rąk w środowisku 2d i 3d

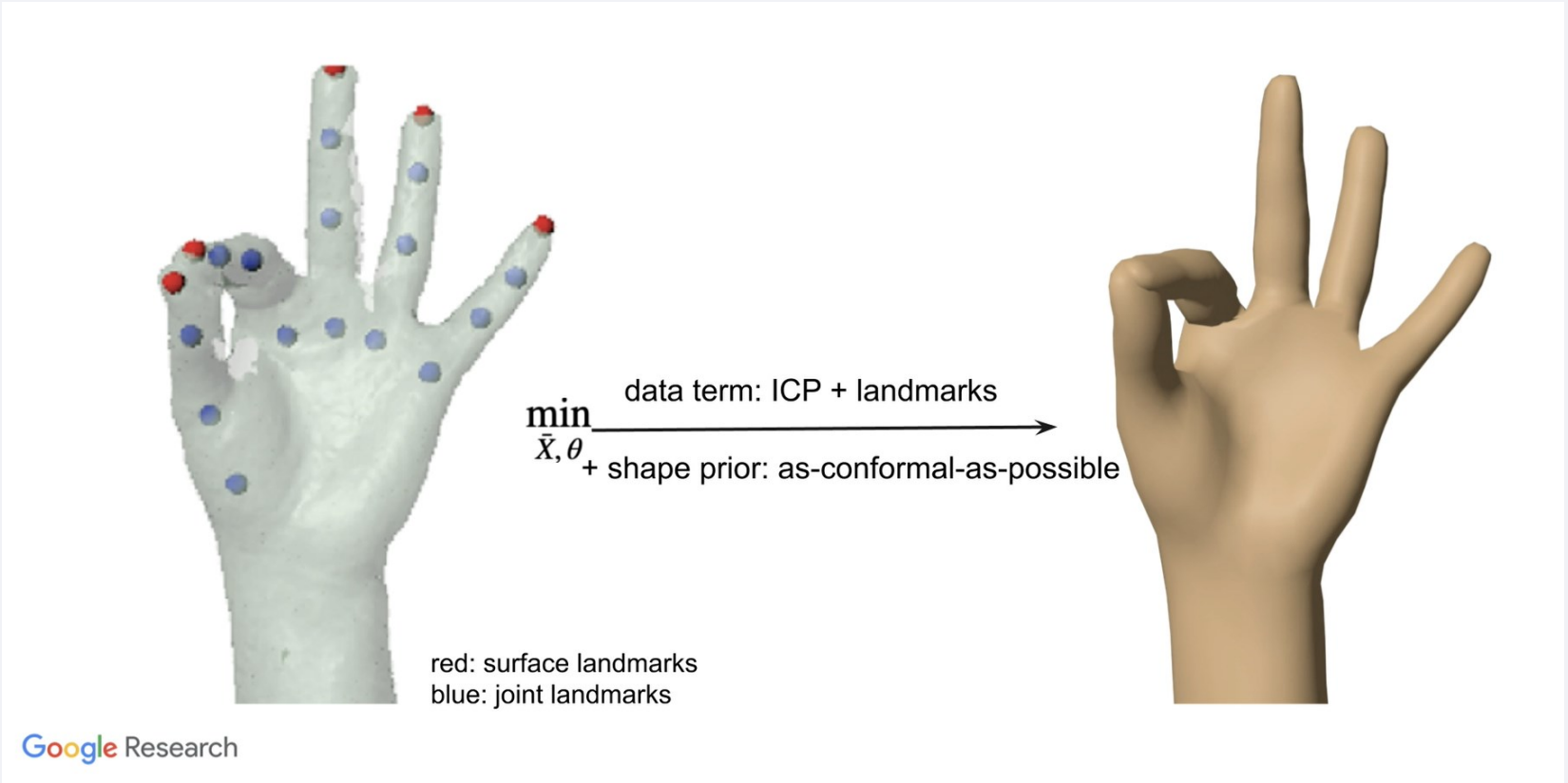
- Wielo – widokowe renderingi twarzy i rąk
- Rozpoznawanie cech szczególnych twarzy i rąk na płaszczyźnie 2D
- Triangulacja znaków szczególnych twarzy i rąk w przestrzeni 3D



Metoda skanowania 3d umożliwia wykrycie i lokalizację punktów charakterystycznych twarzy (obrazujących ekspresję) i rąk na płaszczyźnie 2d oraz w przestrzeni 3d, a także zarysowanie i powiązanie ich ze szkieletem i stawami.

3.3.6.3. Możliwości systemu - wierność dopasowania przegubowych modeli 3d

Optymalizacja sylwetki ludzkiej w pozie spoczynkowej (A-pose) i jej dopasowanie do określonej pozy odbywa się poprzez korelację znaków punktów charakterystycznych skóry *(mesh, surface) oraz szkieletu (stawy i kości)*

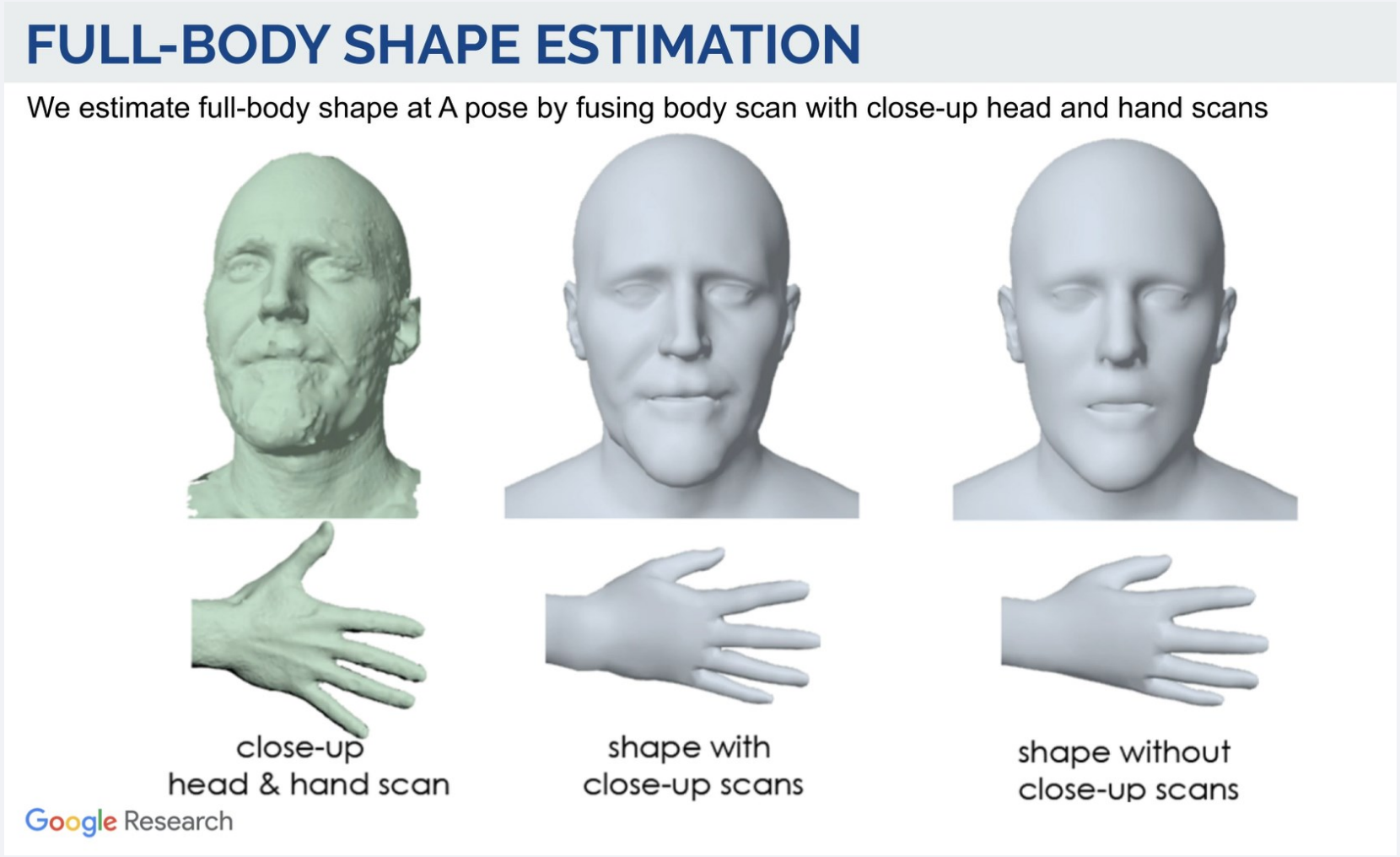


3.3.6.4. Możliwości systemu - dopasowanie formy i kształtu sylwetki ludzkiej

(1) Oryginalne skany 3d (zbliżenia głowy i rąk)

(2) Forma modelu 3d przy użyciu zbliżeń skanów rąk i głowy

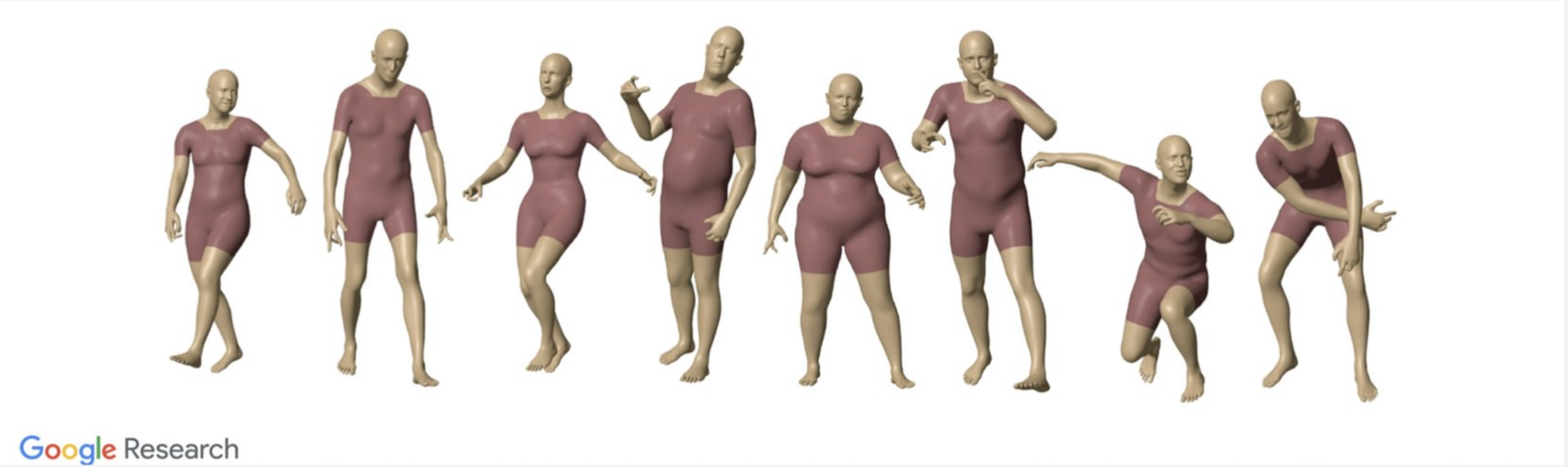
(3) Forma modelu 3d bez użycia zbliżeń skanów 3d



(1) (2) (3)

3.3.6.5. Możliwości systemu - adaptacja wygenerowanego modelu

- dostosowanie kubatury modelu do wzorca
- pozy sylwetki
- ekspresja ruchu (twarz, dłonie, nachylenie (układ) szkieletu)



3.3.7. Wyniki, Aplikacje, Zastosowania

(1) generowanie sylwetki podobnej do wskazanego wzorca

(2) modelowanie ruchu z uwzględnieniem elementów otoczenia – okluzja; np. (AR/VR)

(3) przewidywanie układu ciała na podstawie dotychczasowej fazy ruchu

(4) adaptację sylwetki do wskazanych parametrów np. chudy, gruby

(5) modelowania ekspresji rąk w połączeniu (w fuzji) z całością ciała ludzkiego

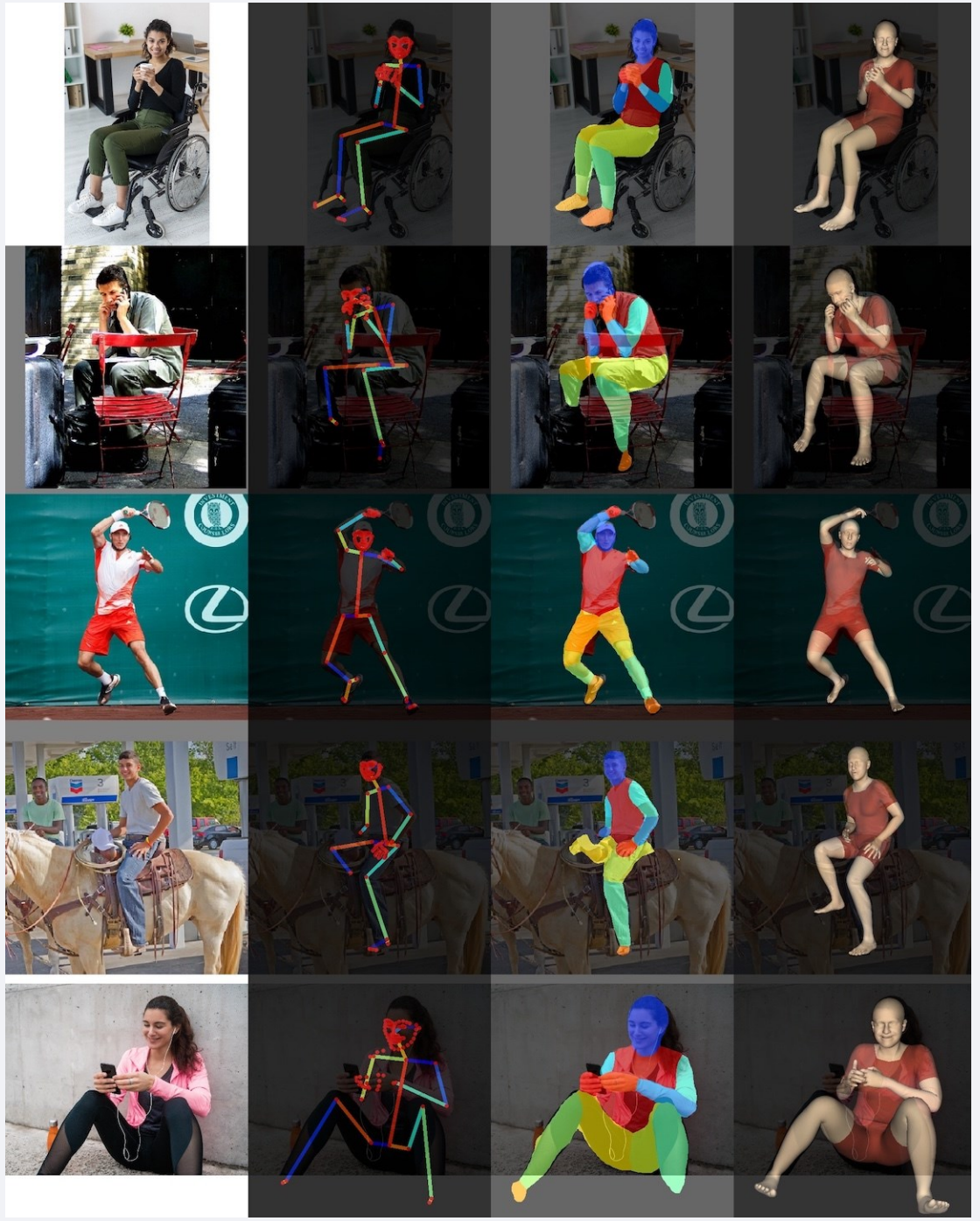
3.3.8. Problemy adaptacji modelu

System generuję model 3D pozy osoby w określonej sytuacji ruchowej w oparciu o:

- rozpoznawanie układu szkieletu (lokalizacja stawów i kości)
- segmentację sylwetki człowieka rozumianej jako deformacje siatki poligonowej poszczególnych części ciała

Uwaga:

System umożliwia **przewidywanie** stanu modelu w przypadku okluzji (przesłaniania) oraz w kolejnych jego faz ruchowych.

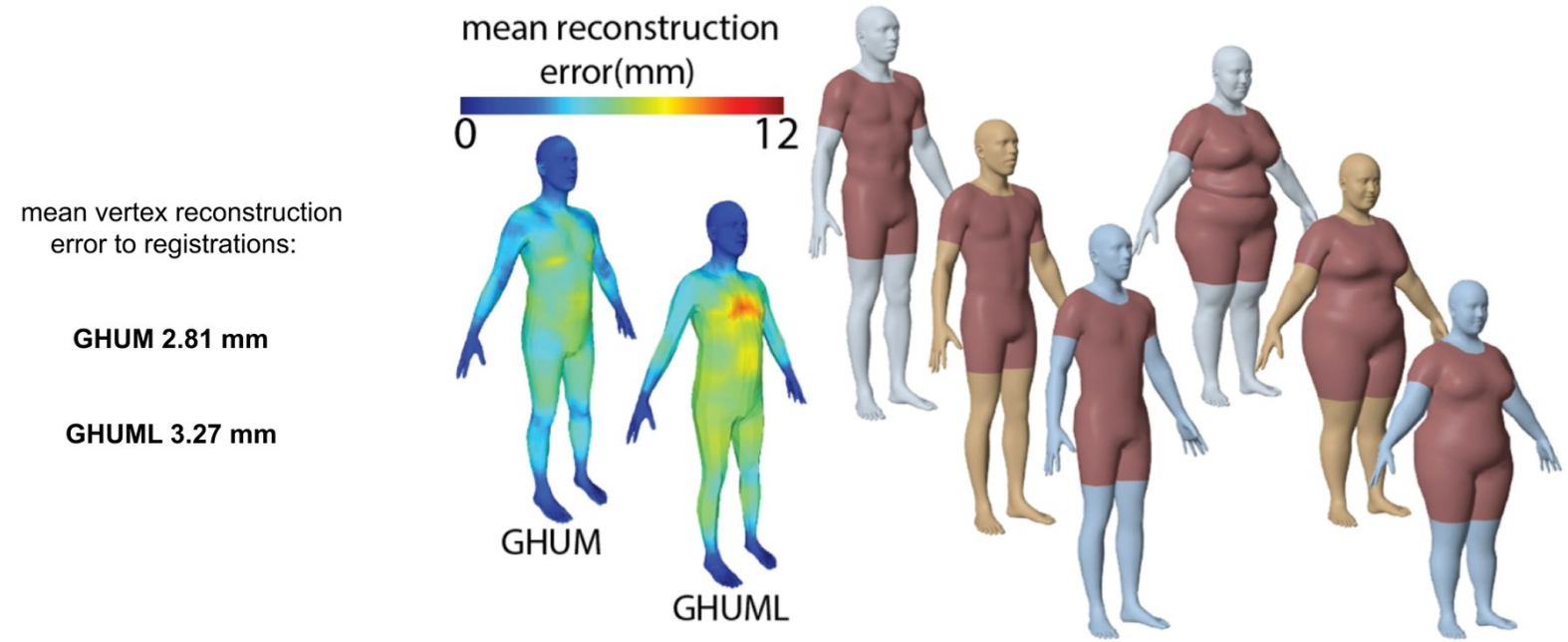


3.3.9. EWALUACJA modelu na podstawie baz danych CAESAR & GHS3D

W potokowym modelu uczenia wydzielono 4 moduły odpowiadające: sylwetce, wyrazowi twarzy, gestykulacji dłoni oraz pozycjom szkieletu.

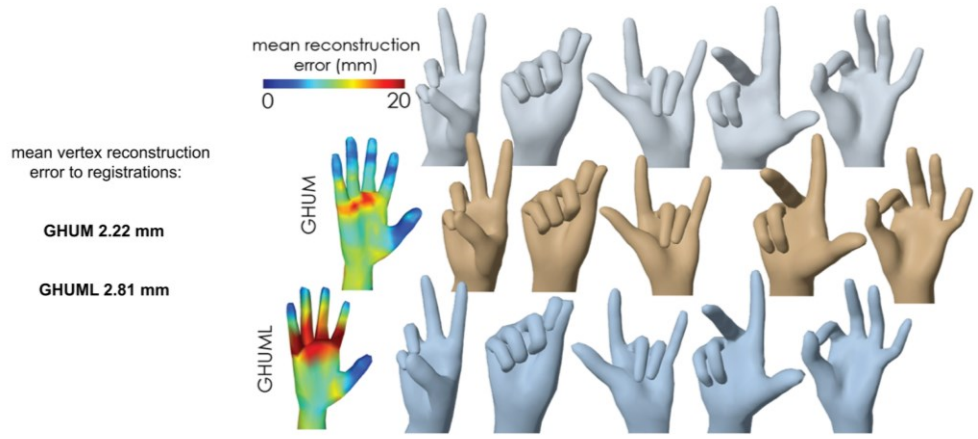
Ewaluację jakości modelowania oceniono w oparciu o średni błąd kwadratowy, a jakość modelu poszczególnych segmentów ciała, oceniono w oparciu o średni błąd rekonstrukcji (wyrażony w mm)

EVALUATION: CAESAR RECONSTRUCTION



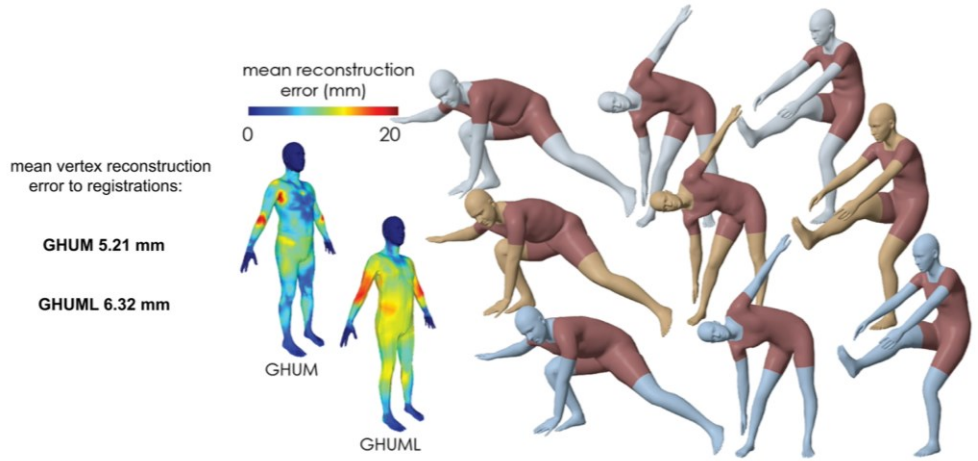
Google Research

EVALUATION: GHS3D HAND RECONSTRUCTION



Google Research

EVALUATION: GHS3D BODY RECONSTRUCTION



Google Research

3.3.10. GHUM vs. SMPL

Porównano modele GHUM i SMPL

(1) GHUM jest dokładniejszy, a może nawet trochę lepszy w wizualnym oddaniu sylwetki człowieka niż SMPL

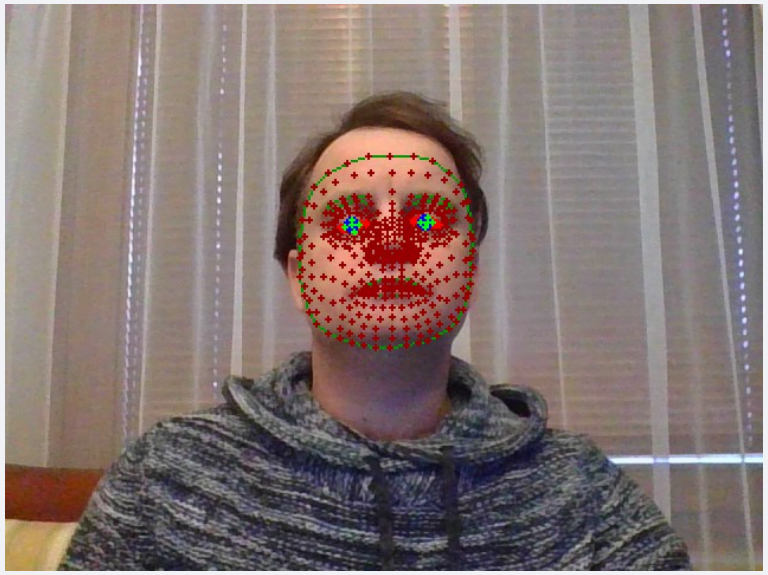
(2) GHUM uzyskuje mniejszy błąd rekonstrukcji całości ciała niż SMPL

- GHUM is close (or slightly better) to SMPL in skinning visual quality
- Vertex point-to-plane error (body-only): ***GHUM: 4.23 mm*** *vs* ***SMPL: 4.96 mm***

GHUM Registration SMPL

3.4. Aplikacje do animacji z wykorzystaniem metod AI (przykłady)

Przechwytywanie ruchu bez specjalistycznego sprzętu i danych



Przykład zastosowania SI do przechwytywania ruchu i ekspresji twarzy z wykorzystaniem biblioteki Media Pipe i Blender 3d oraz kamery w laptopie.

Więcej o bibliotece bazującej na modelu GHUM, o który omawiałem we wcześniejszej części wykładu. Mogą Państwo dowiedzieć się pod następującymi adresami www. https://mediapipe.dev/; https://google.github.io/mediapipe/

3.4. Aplikacje do animacji z wykorzystaniem metod AI (przykłady)

Przechwytywanie ruchu bez specjalistycznego sprzętu i danych

Deep motion; https://www.deepmotion.com/. Przykład zastosowania SI do generacji szkieletu i modelu 3d z nagranej sekwencji ruchów telefonem.

Deep Motion

https://www.deepmotion.com/

NVIDIA

https://developer.nvidia.com/blog/nvidia-ai-generating-motion-capture-animation-without-hardware-or-motion-data/

PLASK

https://plask.ai/

PIXCAP

https://www.pixcap.com/

RADiCAL

https://getrad.co/

4. Literatura, bibliografia, źródła:

1. Wprowadzenie, rys historyczny

2. Przegląd współczesnych systemów przechwytywania ruchu

3. Przechwytywanie ruchu z wykorzystaniem metod AI

Projekt współfinansowany ze środków Unii Europejskiej w ramach Europejskiego Funduszu Rozwoju Regionalnego Program Operacyjny Polska Cyfrowa na lata 2014-2020,
Oś Priorytetowa nr 3 "Cyfrowe kompetencje społeczeństwa" Działanie nr 3.2 "Innowacyjne rozwiązania na rzecz aktywizacji cyfrowej"
Tytuł projektu: „Akademia Innowacyjnych Zastosowań Technologii Cyfrowych (AI Tech)"
Porozumienie nr POPC.03.02.00-00-0001/20-00

1. Wprowadzenie, rys historyczny

[1.1] Eadweard Muybridge: Źródło: https://www.eadweardmuybridge.co.uk/
[1.2.1] Rotoskop: Źródło: https://www.engadget.com/2018-05-25-motion-capture-history-video-vicon-siren.html;
[1.2.2] Rotoskop: Źródło: https://pl.wikipedia.org/wiki/Rotoskop;
[1.3.1] Animac: Taryn Baker, The History of Motion Capture Within The Entertainment Industry, Metropolia University of Applied Sciences, Bachelor Thesis, Dostępne on-line: 09.04.2022
[1.4.1] Egzoszkielet: Taryn Baker, The History of Motion Capture Within The Entertainment Industry, Metropolia University of Applied Sciences, Bachelor Thesis, Date 05.05.2020. Dostępne on-line: 09.04.2022
[1.4.2] Egzoszkielet: T. W. Calvert, J. Chapman and A. Patla, "Aspects of the kinematic simulation of human movement," IEEE Computer Graphics and Applications, Vol. 2, No. 9, November 1982, pp. 41-50. Dostępne on-line: 09.04.2022
[1.5.1] Marionetka Graficzna: Taryn Baker, The History of Motion Capture Within The Entertainment Industry, Metropolia University of Applied Sciences, Bachelor Thesis, Date 05.05.2020. Dostępne on-line: 09.04.2022;
[1.5.2] Marionetka Graficzna: Źródło: https://vimeo.com/43287234?embedded=true&source=vimeo_logo&owner=2108367; Dostępne on-line: 09.04.2022
[1.6.1] Waldo C. Graphics: Źródło: https://beforesandafters.com/2019/08/24/re-visiting-waldo/;
[1.6.2] Waldo C. Graphics: Źródło: https://www.henson.com/jimsredbook/2014/04/4-1988/;
[1.6.3] Waldo C. Graphics: Źródło: https://www.youtube.com/watch?v=dP6TUB7KQc4

Projekt współfinansowany ze środków Unii Europejskiej w ramach Europejskiego Funduszu Rozwoju Regionalnego Program Operacyjny Polska Cyfrowa na lata 2014-2020,
Oś Priorytetowa nr 3 "Cyfrowe kompetencje społeczeństwa" Działanie nr 3.2 "Innowacyjne rozwiązania na rzecz aktywizacji cyfrowej"
Tytuł projektu: „Akademia Innowacyjnych Zastosowań Technologii Cyfrowych (AI Tech)"
Porozumienie nr POPC.03.02.00-00-0001/20-00

2. Przegląd współczesnych systemów przechwytywania ruchu

[2.1.1] Modele szkieletu człowieka: MODELOWANIE KSZTAŁTU I RUCHU ANIMOWANYCH POSTACI - Tomasz Kucharski (praca dyplomowa); promotor: dr inż. Ryszard Klempous (2005).
[2.1.2] Modele szkieletu człowieka: Reconstruction of 3D human motion in real-time using particle swarm optimization with GPU-accelerated fitness function; Bogdan Kwolek, Boguslaw Rymut. AGH Kraków.
[2.1.3] Modele szkieletu człowieka: Źródło: https://mocappys.com/10-step-guide-to-building-a-skeleton-for-motionbuilder/#.YlcX2ehBxPY
[2.1.4] Modele szkieletu człowieka: J. Valcik, Similarity Models for Human Motion Data. PhD thesis, Masaryk University, Brno, 2016.
[2.1.5] Modele szkieletu człowieka: Modele i algorytmy detekcji znaczników w problemach analizy ruchu obiektów. (Models and algorithms for Tag detection in object motion analysis problems). Konrad Kluwak (rozprawa doktorska) Promotor: dr hab. inż. Ryszard Klempous, prof. Uczelni Promotor pomocniczy: dr inż. Jan Nikodem; Politechnika Wrocławska, W4, 2021
[2.2.1] Optyczne systemy przechwytywania ruchu – pasywne i aktywne: Źródło: https://metamotion.com/motion-capture/optical-motion-capture-1.htm
[2.2.2] Optyczne systemy przechwytywania ruchu – pasywne i aktywne: Źródło: https://en.wikipedia.org/wiki/Motion_capture
[2.2.3] Optyczne systemy przechwytywania ruchu – pasywne i aktywne: Źródło: https://optitrack.com/cameras/flex-13/indepth.html
[2.2.4] Optyczne systemy przechwytywania ruchu – pasywne i aktywne: Rozmieszczenie markerów pasywnych w charakterystycznych miejscach na kostiumie aktora mo-cap i generacja szkieletu w oprogramowaniu komputerowym. (video); Źródło: https://optitrack.com/software/motive/

Projekt współfinansowany ze środków Unii Europejskiej w ramach Europejskiego Funduszu Rozwoju Regionalnego Program Operacyjny Polska Cyfrowa na lata 2014-2020,
Oś Priorytetowa nr 3 "Cyfrowe kompetencje społeczeństwa" Działanie nr 3.2 "Innowacyjne rozwiązania na rzecz aktywizacji cyfrowej"
Tytuł projektu: „Akademia Innowacyjnych Zastosowań Technologii Cyfrowych (AI Tech)"
Porozumienie nr POPC.03.02.00-00-0001/20-00

[2.2.5] Pasywne, optyczne systemy przechwytywania ruchu: Źródło: https://metamotion.com/motion-capture/optical-motion-capture-1.htm

[2.2.6] Pasywne, optyczne systemy przechwytywania ruchu: Rozmieszczenie markerów pasywnych w charakterystycznych miejscach na kostiumie aktora mo-cap (a) System optyczny – pasywny (Zestaw Kamer Optitrack, Flex13) (b) Ubieranie aktora w markery w charakterystycznych miejscach. (c) przygotowania do przechwytywania ruchu, zestaw markerów i kostium - aktor wraz z przedmiotem. Wydział Informatyki i Telekomunikacji. Politechnika Wrocławska. Fotografie z archiwum własnego autora. Kwiecień 2022.

[2.2.7] Pasywne, optyczne systemy przechwytywania ruchu: Video. System optyczny pasywny. Nagrania sekwencji ruchu na przykładzie japońskiej sztuki walki z użyciem dodatkowego przedmiotu (katana). Wydział Informatyki i Telekomunikacji. Politechnika Wrocławska. Fotografie z archiwum własnego autora. Kwiecień 2022.

[2.2.8] Aktywne optyczne systemy przechwytywania ruchu: Rodzaje optycznych markerów aktywnych. Źródło: http://www.cgchannel.com/2017/02/optitrack-launches-its-first-active-optical-tracking-system; https://ar-tracking.com/en/product-program/markers

[2.2.9] Optyczne systemy przechwytywania ruchu – wady i zalety: Źródła: https://optitrack.com/software/motive/;

[2.2.10] Nieoptyczne systemy przechwytywania ruchu: Źródła: https://www.xsens.com/motion-capture; https://en.wikipedia.org/wiki/Xsens; https://www.xsens.com/news/making-of-the-chemical-brothers-wide-open-ft-beck;

[2.2.11] Współczesne egzoszkielety do przechwytywania ruchu: Źródła: https://www.pccpolska.pl/motion-capture-co-powinienes-wiedziec-o-systemach-mocap/;

Projekt współfinansowany ze środków Unii Europejskiej w ramach Europejskiego Funduszu Rozwoju Regionalnego Program Operacyjny Polska Cyfrowa na lata 2014-2020,
Oś Priorytetowa nr 3 "Cyfrowe kompetencje społeczeństwa" Działanie nr 3.2 "Innowacyjne rozwiązania na rzecz aktywizacji cyfrowej"
Tytuł projektu: „Akademia Innowacyjnych Zastosowań Technologii Cyfrowych (AI Tech)"
Porozumienie nr POPC.03.02.00-00-0001/20-00

3. Przechwytywanie ruchu z wykorzystaniem metod AI

[3.1.1] Szkielet pozy człowieka - Open Pose: Źródła: https://en.wikipedia.org/wiki/Convolutional_neural_network; https://viso.ai/deep-learning/openpose/; https://viso.ai/deep-learning/pose-estimation-ultimate-overview/; https://www.youtube.com/watch?v=T1vvokFnsbU;
[3.1.2] Szkielet pozy człowieka - Open Pose: Źródło: https://viso.ai/deep-learning/openpose/;
[3.1.3] Szkielet pozy człowieka - Open Pose: Źródło: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields; Zhe Cao, Tomas Simon, Shih-En Wei, Yaser Sheikh
[3.1.4] Szkielet pozy człowieka - Open Pose: Źródło: https://github.com/CMU-Perceptual-Computing-Lab/openpose;
[3.2.1] Generatywna postać ludzka 3d i przegubowe modele naśladujące pozy: Źródło: https://google.github.io/mediapipe/solutions/pose.html;
[3.2.2] Generatywna postać ludzka 3d i przegubowe modele naśladujące pozy: Źródło: BlazePose: On-device Real-time Body Pose tracking; Valentin Bazarevsky, Ivan Grishchenko, Karthik Raveendran, Tyler Zhu, Fan Zhang, Matthias Grundmann
[3.2.3] Generatywna postać ludzka 3d i przegubowe modele naśladujące pozy: Źródło: GHUM & GHUML: Generative 3D Human Shape and Articulated Pose Models; Hongyi Xu, Eduard Gabriel Bazavan, Andrei Zanfir, William T. Freeman, Rahul Sukthankar, Cristian Sminchisescu; Google Research, 2021
[3.2.4] Generatywna postać ludzka 3d i przegubowe modele naśladujące pozy: GHUM & GHUML: Generative 3D Human Shape and Articulated Pose Models; Źródło: https://github.com/google-research/google-research/tree/master/ghum;

Projekt współfinansowany ze środków Unii Europejskiej w ramach Europejskiego Funduszu Rozwoju Regionalnego Program Operacyjny Polska Cyfrowa na lata 2014-2020,
Oś Priorytetowa nr 3 "Cyfrowe kompetencje społeczeństwa" Działanie nr 3.2 "Innowacyjne rozwiązania na rzecz aktywizacji cyfrowej"
Tytuł projektu: „Akademia Innowacyjnych Zastosowań Technologii Cyfrowych (AI Tech)"
Porozumienie nr POPC.03.02.00-00-0001/20-00

[3.2.5] Generatywna postać ludzka 3d i przegubowe modele naśladujące pozy: HUND: Neural Descent for Visual 3D Human Pose and Shape; Źródło: https://github.com/google-research/google-research/tree/master/ghum/inference/hund;
[3.2.6] Generatywna postać ludzka 3d i przegubowe modele naśladujące pozy: THUNDR: Transformer-based 3D HUmaN Reconstruction with Markers; Źródło: https://github.com/google-research/google-research/tree/master/ghum/inference/thundr;
[3.3.1] Dostępne aplikacje do animacji z wykorzystaniem metod AI: Deep Motion: Źródło: https://www.deepmotion.com;
[3.3.2] Dostępne aplikacje do animacji z wykorzystaniem metod AI: NVIDIA: Źródło: https://developer.nvidia.com/blog/nvidia-ai-generating-motion-capture-animation-without-hardware-or-motion-data/;
[3.3.3] Dostępne aplikacje do animacji z wykorzystaniem metod AI: PLASK: Źródło: https://plask.ai/;
[3.3.4] Dostępne aplikacje do animacji z wykorzystaniem metod AI: PIXCAP: Źródło: https://www.pixcap.com/
[3.3.5] Dostępne aplikacje do animacji z wykorzystaniem metod AI: RADiCAL: Źródło: https://getrad.co/;

Projekt współfinansowany ze środków Unii Europejskiej w ramach Europejskiego Funduszu Rozwoju Regionalnego Program Operacyjny Polska Cyfrowa na lata 2014-2020,
Oś Priorytetowa nr 3 "Cyfrowe kompetencje społeczeństwa" Działanie nr 3.2 "Innowacyjne rozwiązania na rzecz aktywizacji cyfrowej"
Tytuł projektu: „Akademia Innowacyjnych Zastosowań Technologii Cyfrowych (AI Tech)"
Porozumienie nr POPC.03.02.00-00-0001/20-00

govtech Polska

Politechnika Wrocławska

AI Tech
AKADEMIA INNOWACYJNYCH ZASTOSOWAŃ TECHNOLOGII CYFROWYCH

Dziękuję za uwagę! (-:

Fundusze Europejskie Polska Cyfrowa

Rzeczpospolita Polska

Unia Europejska
Europejski Fundusz Rozwoju Regionalnego

Projekt współfinansowany ze środków Unii Europejskiej w ramach Europejskiego Funduszu Rozwoju Regionalnego Program Operacyjny Polska Cyfrowa na lata 2014-2020,
Oś Priorytetowa nr 3 "Cyfrowe kompetencje społeczeństwa" Działanie nr 3.2 "Innowacyjne rozwiązania na rzecz aktywizacji cyfrowej"
Tytuł projektu: „Akademia Innowacyjnych Zastosowań Technologii Cyfrowych (AI Tech)”
Porozumienie nr POPC.03.02.00-00-0001/20-00