

Západočeská univerzita v Plzni
Fakulta aplikovaných věd
Katedra informatiky a výpočetní techniky

Diplomová práce

Automatické vytvoření 3D postavy na základě jedné fotografie

Plzeň, 2021

Bc. Jan Kašák

ZÁPADOČESKÁ UNIVERZITA V PLZNI

Fakulta aplikovaných věd

Akademický rok: 2020/2021

ZADÁNÍ DIPLOMOVÉ PRÁCE

(projektu, uměleckého díla, uměleckého výkonu)

Jméno a příjmení:	Bc. Jan KAŠÁK
Osobní číslo:	A18N0047P
Studijní program:	N3902 Inženýrská informatika
Studijní obor:	Počítačová grafika
Téma práce:	Automatické vytvoření 3D postavy na základě jedné fotografie
Zadávací katedra:	Katedra informatiky a výpočetní techniky

Zásady pro vypracování

1. Seznamte se s problematikou vytváření 3D modelu postav.
2. Prozkoumejte alespoň 8 existujících metod vytváření 3D modelů postav z barevných fotografií a proveďte jejich srovnání na teoretické úrovni.
3. Po dohodě s vedoucím práce si vyberte metodu vhodnou pro automatické vytvoření 3D modelu postavy na základě jediné barevné fotografie a proveďte její implementaci.
4. Proveďte verifikaci a validaci naimplementované metody v různých scénářích (např. různé pohlaví, typ ošacení, póza, apod.).
5. Proveďte zhodnocení dosažených výsledků a nastiňte možnosti navazující práce.

Rozsah diplomové práce: **doporuč. 50 s. původního textu**
Rozsah grafických prací: **dle potřeby**
Forma zpracování diplomové práce: **tištěná**

Seznam doporučené literatury:
dodá vedoucí diplomové práce

Vedoucí diplomové práce: **Doc. Ing. Josef Kohout, Ph.D.**
Nové technologie pro informační společnost

Datum zadání diplomové práce: **11. září 2020**
Termín odevzdání diplomové práce: **20. května 2021**

L.S.

Doc. Dr. Ing. Vlasta Radová
děkanka

Doc. Ing. Přemysl Brada, Sc., Ph.D.
vedoucí katedry

V Plzni dne 24. září 2020

Prohlášení

Prohlašuji, že jsem tuto diplomovou práci vypracoval samostatně pod vedením pana Doc. Ing. Josefa Kohouta, Ph.D. a že jsem v seznamu použité literatury uvedl všechny zdroje použité při zpracování práce.

V Plzni dne 1. května 2021

Jan Kašák

Poděkování

Na tomto místě bych rád poděkoval panu Doc. Ing. Josefovi Kohoutovi, Ph.D. za odborné rady a pomoc při tvorbě diplomové práce.

Abstract

This master thesis deals with automatic creation of 3D human models based on a single photography. The main object is to compare several available methods, choose one of them and test it.

3D models are created based on photographs with persons with different kinds of clothes and with different poses. The method is tested with both a man and a woman. 3D models created by the chosen method are automatically measured on predefined body parts. There is also created a module for automatic estimation of person height in the master thesis. The estimated height is used for scaling the models. The scaled models are closer to the reality with their sizes.

Next, an internet application is created during the works on the thesis. It is used for automatic creation of 3D models from a single photography and easy interactive viewing of created models. The user can share his obtained results if he wants to.

There is an evaluation and a comparison of the results at the conclusion of the thesis. Very good results were achieved for height estimation. Estimated height of tested person corresponded with his real height very well. The proportions of created 3D models were better by the man and the female models were farther away from the reality.

Finally, there are shown options of next possible master thesis extensions.

All the calculations are performed in MS Excel 2016. Files with the source codes, photographs and calculations are accessible in the electronically included folder.

Key words

3D model, 2D to 3D, shape, silhouette, human body, photograph

Abstrakt

Diplomová práce se věnuje automatickému vytváření 3D modelů postavy na základě jedné fotografie. Cílem je porovnat dostupné metody, jednu z nich vybrat a otestovat.

3D modely jsou vytvářeny na základě fotografií osob v různých typech oblečení a s různými pózami těla. Testování se zúčastní muž a žena. Získané 3D modely vytvořené vybranou metodou jsou automaticky měřeny na předem definovaných partiích. V diplomové práci je vytvořen modul pro automatický odhad výšky postavy. Pomocí odhadnuté výšky postavy jsou modely škálovány. Naškálované modely více odpovídají reálným mírám osob.

V průběhu práce je vytvořena internetová aplikace. Ta slouží k automatickému vytváření 3D modelů z jedné fotografie a snadnému interaktivnímu prohlížení vytvořených modelů. Uživatel se může rozhodnout pro sdílení získaných výsledků.

V závěru práce jsou vyhodnoceny a porovnány výsledky testování. Bylo dosaženo velmi dobrých odhadů výšky postavy. Odhadnutá výška téměř dokonale odpovídala reálné výšce testované osoby. Rozměry získaných 3D modelů byly blíže skutečným rozměrům osoby u mužské postavy, u ženské postavy došlo k větším odchýlkám od reality.

Nakonec jsou nastíněny možnosti dalšího rozšíření diplomové práce.

Všechny výpočty jsou prováděny v MS Excel 2016. Zdrojové soubory, fotografie a výpočty jsou dostupné v elektronicky přiložené složce.

Klíčová slova

3D model, 2D do 3D, tvar, silueta, postava, fotografie

Obsah

Seznam obrázků.....	x
Seznam tabulek	xii
1 Úvod.....	- 1 -
2 Vytváření 3D modelu postavy	- 2 -
3 Metody vytváření 3D modelu postavy z barevné fotografie	- 3 -
3.1 3D Body Shapes Estimation from Dressed-Human Silhouettes	- 3 -
3.2 Body Talk Crowdshaping Realistic 3D Avatars with Words	- 4 -
3.3 End-to-end Recovery of Human Shape and Pose	- 5 -
3.4 Automatic Estimation of 3D Human Pose and Shape from a Single Image. -	6 -
3.5 Expressive Body Capture: 3D hands, Face and Body from a Single Image. -	7 -
3.6 Closing the Loop Between 3D and 2D Human Representations	- 9 -
3.7 Parametric Reshaping of Human Bodies in Images.....	- 10 -
3.8 Photo Wake-Up 3D Character Animation from a Single Photo	- 11 -
3.9 Physics-Inspired Garment Recovery from a Single-View Image	- 12 -
4 Srovnání metod	- 15 -
5 Použitá metoda.....	- 16 -
5.1 Aproximace těla kapslemi.....	- 17 -
5.2 Objektivní funkce.....	- 18 -
5.3 Optimalizace	- 20 -
6 Návrh experimentu	- 21 -
7 Realizace, implementace	- 26 -
7.1 Internetová aplikace	- 26 -
7.2 Odhad výšky osoby	- 31 -
7.3 Škálování modelu.....	- 37 -
7.4 Měření modelu	- 37 -
8 Dosažené výsledky	- 41 -

8.1	Internetová aplikace	- 41 -
8.2	Výška osoby	- 43 -
8.3	Rozměry modelu	- 53 -
8.3.1	Žena	- 53 -
8.3.2	Muž 1	- 57 -
8.3.3	Muž 2	- 62 -
8.4	Shrnutí	- 66 -
9	Možnosti rozšíření práce	- 68 -
10	Závěr	- 69 -
	Příloha	- 71 -
	Použitá literatura	- 76 -
	Pro vypracování diplomové práce byly použity následující zdroje:	- 76 -
	Obsah elektronických příloh	- 81 -

Seznam obrázků

Obrázek 3.1: Postup metody 3D Body Shapes Estimation.....	- 4 -
Obrázek 3.2: Postup metody Keep it SMPL.....	- 7 -
Obrázek 3.3: Metoda SMPLify-X s modelem SMPL-X	- 9 -
Obrázek 3.4: Postup metody Parametric Reshaping.....	- 11 -
Obrázek 3.5: Postup metody Photo Wake-Up	- 12 -
Obrázek 3.6: Postup metody Garment Recovery.....	- 14 -
Obrázek 5.1: Dosažitelné výsledky metody SMPLify.....	- 16 -
Obrázek 5.2: Aproximace těla kapslemi	- 17 -
Obrázek 6.1: Ukázka typů oblečení	- 25 -
Obrázek 6.2: Ukázka typů póz.....	- 25 -
Obrázek 7.1: Vliv rozmazání fotografie na detekci hran.....	- 32 -
Obrázek 7.2: Proces odhadu výšky – 1. část.....	- 32 -
Obrázek 7.3: Proces odhadu výšky – 2. část.....	- 36 -
Obrázek 7.4: Měření modelu – vrchní část těla	- 39 -
Obrázek 7.5: Měření modelu – spodní část těla.....	- 40 -
Obrázek 8.1: Hodnocení získaná pomocí internetové aplikace	- 42 -
Obrázek 8.2: Ukázka nevhodných fotografií	- 44 -
Obrázek 8.3: Ukázka vhodné fotografie	- 45 -
Obrázek 8.4: Histogram výšek pro kernel (1,3).....	- 46 -
Obrázek 8.5: Histogram výšek pro kernel (3,3).....	- 46 -
Obrázek 8.6: Histogram výšek pro kernel (3,4).....	- 46 -
Obrázek 8.8: Histogram výšek pro kernel (1,4).....	- 47 -
Obrázek 8.9: Histogram výšek pro kernel (2,3).....	- 47 -
Obrázek 8.10: Histogram rozměrů hrudníku – žena.....	- 55 -
Obrázek 8.11: Histogram rozměrů pasu – žena.....	- 56 -
Obrázek 8.12: Histogram rozměrů boků – žena	- 56 -
Obrázek 8.13: Histogram rozměrů hrudníku – muž 1	- 59 -
Obrázek 8.14: Histogram rozměrů pasu – muž 1	- 59 -
Obrázek 8.15: Histogram rozměrů boků – muž 1	- 59 -
Obrázek 8.16: Rozměry hlavních partií – muž 1	- 61 -
Obrázek 8.17: Histogram rozměrů hrudníku – muž 2	- 64 -
Obrázek 8.18: Histogram rozměrů pasu – muž 2	- 64 -

Obrázek 8.19: Histogram rozměrů boků – muž 2.....	- 64 -
Obrázek 8.20: Rozměry hlavních partií – muž 2.....	- 66 -
Obrázek 10.1: Internetová aplikace – úvodní stránka.....	- 75 -
Obrázek 10.2: Internetová aplikace – stránka s výsledky.....	- 75 -

Seznam tabulek

Tabulka 8.1: Srovnání výsledků odhadu výšky pro různé typy kernelů.....	- 48 -
Tabulka 8.2: Výsledky odhadu výšky pro fotografie skupiny A a B	- 49 -
Tabulka 8.3: Vliv pozice hlavy na odhadnutou výšku.....	- 50 -
Tabulka 8.4: Vliv rozlišení na dobu běhu a spotřebu paměti RAM	- 51 -
Tabulka 8.5: Vliv rozlišení na odhadnutou výšku	- 53 -
Tabulka 8.6: Vstupní rozměry – žena	- 53 -
Tabulka 8.7: Výsledné rozměry modelů – žena	- 55 -
Tabulka 8.8: Vstupní rozměry – muž 1	- 57 -
Tabulka 8.9: Výsledné rozměry modelů – muž 1	- 58 -
Tabulka 8.10: Vstupní rozměry – muž 2	- 62 -
Tabulka 8.11: Výsledné rozměry modelů – muž 2	- 63 -

1 Úvod

V dnešní době se lidé čím dál tím více setkávají s digitálními 3D lidskými avatary. Možné aplikace se neustále rozšiřují. S 3D avatary se můžeme setkat v počítačových hrách, filmovém průmyslu, zdravotnictví, virtuální realitě nebo v nákupním odvětví. Lze pozorovat, že tvoření realistických 3D avatarů je ve spojení s rozvíjejícími se technologiemi aktuálním tématem počítačové grafiky s širokými možnostmi využití.

Jedním z cílů diplomové práce je popsat dostupné metody pro vytvoření 3D modelu z jedné barevné fotografie a vzájemně je porovnat. Dalším cílem diplomové práce je jednu z uvedených metod vybrat a ověřit její kvality – především kvalitu vytvořených modelů. Pro tento účel budou provedeny a vyhodnoceny statistické testy. Pro vyhodnocení podoby osoby s modelem a interaktivní prohlížení vytvořeného modelu bude také vytvořena internetová aplikace, kde uživatelé slovně subjektivně vyhodnotí spokojenost s podobou vytvořeného modelu se zaslanou vstupní fotografií.

V diplomové práci jsou v úvodní části nejprve představeny dostupné metody pro vytvoření 3D modelu postavy z jediné barevné fotografie. Toto řešení vytvoření 3D modelu je jednodušší a mnohem levnější než použití drahých 3D skenerů. Jednotlivé uvedené metody jsou v příštích kapitolách mezi sebou na teoretické úrovni navzájem porovnány. Jedna z představených metod je vybrána, naimplementována a pomocí internetové aplikace verifikována použitím v reálném prostředí. Vytvořené modely budou také dále měřeny na předem definovaných tělesných partiích a porovnány se skutečnými tělesnými měřeními testovacích osob. Testovací osoba bude vyfocena v různých typech oblečení a s různými tělesnými pózami.

Motivací pro vybrání daného tématu je možné využití práce v praxi. Odborníci ze zdravotnictví dokáží s pomocí aplikace zjistit, zda pacient trpí poruchou příjmu potravy. Je pro to využít 3D model pacientova těla a jeho virtuální upravování. Diplomová práce může mít v budoucnosti pozitivní dopad.

Jako zdroj informací pro tuto práci jsou použity především zdrojové texty jednotlivých metod pro vytváření 3D lidských avatarů z jediné barevné fotografie.

2 Vytváření 3D modelu postavy

Tato kapitola je zpracována dle zdroje [1].

Metody pro rekonstrukci 3D modelu postavy lze rozdělit do dvou hlavních kategorií: parametrické a neparametrické. Další rozdělení může být na metody automatické a poloautomatické, které vyžadují drobný manuální zásah uživatele.

Neparametrické metody vedou k získání 3D sítí, které lze přirovnat k naskenovanému mračnu bodů. Skenování je provedeno převážně pomocí drahých 3D skenerů nebo senzorem Kinect. Na skenované osoby je kladený požadavek, aby byly co nejméně ošacené. Do této skupiny metod patří např. [2], [3] a [4].

Pro vytváření 3D modelu postav existuje velké množství parametrických metod. Tyto metody jako zdroj informací využívají jedné nebo většího množství fotografií. Fotografie jsou oproti 3D skenerům dostupnější a jejich získání je výrazně levnější a rychlejší.

Jedna z prvních metod představila statistický model sloužící pro naučení tvaru podobné pózy. Některými vědci jsou použity nové pokročilejší modely zvané SCAPE [5] nebo SMPL [6] (viz Příloha), které dovolují úspěšné modelování různých póz a deformací lidského těla. Existují též metody umožňující animaci lidského těla nebo dokonce detailní modelaci ošacení. Odlišnost v metodách lze nalézt kromě použitého modelu též v počtu potřebných zdrojových fotografií – může být jediná nebo více z různých úhlů.

Kapitola 3: Metody vytváření 3D modelu postavy z barevné fotografie podrobněji popisuje základní parametrické metody využívající jediné barevné fotografie.

3 Metody vytváření 3D modelu postavy z barevné fotografie

V následujících podkapitolách jsou uvedeny a popsány současné používané metody pro vytvoření 3D lidské postavy na základě jedné barevné fotografie.

3.1 3D Body Shapes Estimation from Dressed-Human Silhouettes

Tato kapitola byla zpracována dle zdroje [1].

Autory této metody jsou Dan Song¹, Ruofeng Tong¹, Jian Chang², Xiasong Yang², Min Tang¹ a Jian Jun Zhang² z ¹Zhejiang University, China a ²Bournemouth University, UK (2016).

V práci je představena metoda pro efektivní odhad 3D tvaru těla pod ošacením. Pro tento účel je vytvořena databáze párů 3D nahých a oblečených těl.

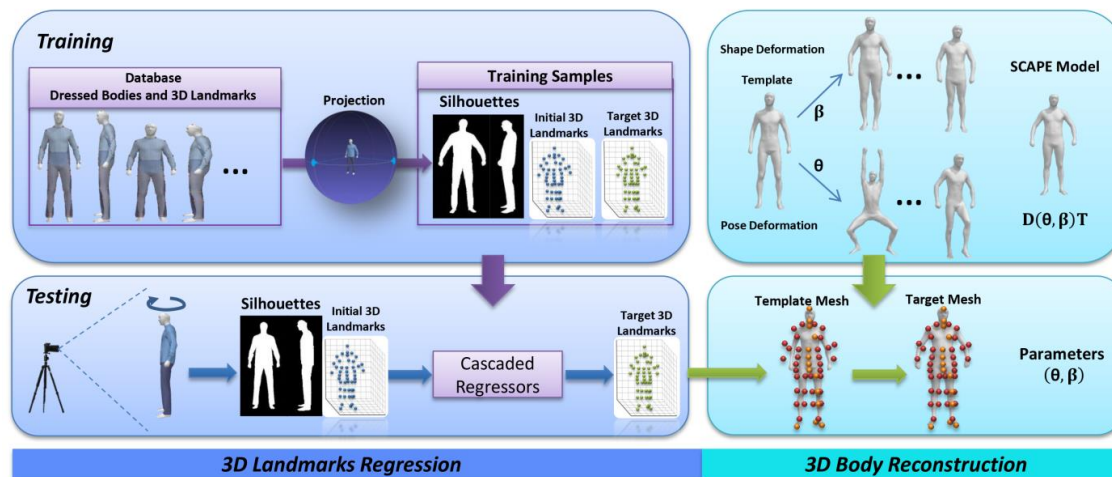
Hlavní aplikací této metody je virtuální zkoušení ošacení. Uživatelé si mohou vytvořit své vlastní věrné 3D tělo a prohlédnout nové ošacení ve virtuální zkušebně.

Pro analýzu vztahu mezi nahou 3D postavou a tvarem ošacení je vytvořena databáze 6042 párů 3D nahých a oblečených postav pro tři různé druhy ošacení. Z důvodu obtížného získání databáze s 6042 mužskými skutečnými těly ve stojící póze je databáze vytvořena syntézou 56 reálných těl z databáze MPI [7]. Těla jsou simulací oblečena třemi sety jednovrstvého ošacení – trikem s dlouhými rukávy a dlouhými kalhotami, trikem s krátkými rukávy a dlouhými kalhotami a trikem s krátkými rukávy a krátkými kalhotami. V ideálním případě je s pomocí této databáze cíleno k dosažení stejné přesnosti pro „odhad vysvětlečného těla“ a pro „simulaci ošacení“.

Na základě této databáze jsou vytvořeny trénovací vzorky obsahující siluety oblečených postav, počáteční 3D významné body a cílové 3D významné body. Siluety jsou vytvořeny projekcí modelů z databáze z předního a bočních pohledů. Významných bodů je celkem 61 po celém těle. Pro kombinaci 3D významných bodů na nahém těle s oblečenými siluetami postav je vytvořen efektivní „feature descriptor“. Regresory jsou natrénovány pro přesun významných bodů (z počátečních na cílové) podle siluet oblečených postav s trénovacími vzorky. V testovací fázi jsou jako vstup použity oblečené siluety postav a set počátečních významných bodů a na cílové 3D významné body na těle je použita regrese s pomocí výsledků trénování.

Významné body po regresi jsou použity jako omezení pro SCAPE model [5] (viz Příloha) při finální rekonstrukci těla.

Obrázek 3.1 [1] zobrazuje kompletní postup této metody v grafické podobě.



Obrázek 3.1: Postup metody 3D Body Shapes Estimation

3.2 Body Talk Crowdshaping Realistic 3D Avatars with Words

Tato kapitola byla zpracována dle zdroje [8].

Autory této metody jsou Stephan Streuber¹, M. Alejandra Quiros-Ramirez¹, Matthew Q. Hill², Carina A. Hahn², Silvia Zuffi³, Alice O'Toole² a Michael J. Black¹ z ¹Max Planck Institute for intelligent Systems, Tübingen, Germany, ²University of Texas at Dallas a ³ITC-CNR, Milan, Italy.

V metodě je využito „crowdsourcingu“ pro vytvoření hodnocení 3D těl na základě slovního popisu vlastností. Sbírkou slovního hodnocení postavy na fotografii poskytuje dostatečně silná omezení pro její 3D tvar. Tento proces hodnocení a vytváření 3D postav na základě slov je nazván „tvarování davem“ („crowdshaping“). Je ukázáno, že systém Body Talk vytváří tvary, které jsou vizuálně nerozeznatelné od těl vytvořených skenery s vysokým rozlišením a že tato přesnost je dostatečná pro mnoho úkolů. [9]

Autoři předpokládají tři hypotézy:

- Lidé si pod určitým slovním popisem představí shodný tvar.
- Kolektivní posouzení tvaru „davem“ obsahuje robustní informaci o opravdovém tvaru těla.

- Korelace mezi hodnocením tvaru a jeho vztahu k statistickým veličinám tohoto tvaru poskytuje dostatečná omezení pro vytvoření přesného 3D tvaru.

Možné aplikace této metody zahrnují např. vytváření 3D postav z knih a postav podle zavedených somatotypů nebo vyhledávání v databázích dle klíčových slov.

System Body Talk vyžaduje jedinou fotografii osoby a 15 lidí, kteří ohodnotí tvar těla na fotografii použitím maximálně 30 slov. V trénovací fázi je nejprve od skupiny lidí získán slovní popis syntetických těl a provedena lineární regrese k namapování modelu ze slovního popisu na geometrickou podobu. Body Talk využívá jako model těla SMPL [6] (viz Příloha). Tento model dovoluje vytvářet 3D digitální postavy podle slovního popisu a naopak.

Je ukázáno, že přesnost se zvyšuje, pokud postavu hodnotí více než jeden člověk. Skupinový pohled na tvar těla a pestré slovní popisy omezují tvar těla s velmi dobrou metrickou přesností.

Hodnocení tvaru těla je prováděno na pěti-stupňové škále. S použitím pouze tohoto hodnocení produkuje BodyTalk 3D modely s průměrnou Euklidovskou chybou mezi vertexy okolo 9 mm v porovnání s 3D sítěmi těl skutečně odpovídajících daným lidem. Pokud osoba na fotografii dodá svoji výšku a váhu, může být chyba snížena na přibližně 8 mm.

Autoři metody vytvořili webovou stránku (<http://bodytalk.is.tue.mpg.de/>), kde si uživatelé mohou vyzkoušet vytvořit avatary za pomoci posuvných tlačítek reprezentující hodnoty slovního popisu a porozumět vztahu mezi slovy a tvarem těla. Pro vědecké účely lze vytvořené sítě stáhnout.

3.3 End-to-end Recovery of Human Shape and Pose

Tato kapitola byla zpracována dle zdroje [10].

Autory této metody jsou Angjoo Kanazawa¹, Michael J. Black², David W. Jacobs³ a Jitendra Malik¹ z ¹University of California, Berkeley, ²MPI for Intelligent Systems, Tübingen, Germany a ³University of Maryland, College Park (2018).

V práci je popsán Human Mesh Recovery (HMR), což je „framework“ (komplexní aplikační rámec sloužící ke snadnějšímu vystavění programů) umožňující rekonstrukci

3D síť lidského těla z jedné RGB fotografie. Repräsentace je parametrizována tvarem a úhly 3D kloubů.

Metoda využívá generativní model lidského těla SMPL [6] (viz Příloha), který parametrizuje síť úhly 3D kloubů a nízko-dimenzionálním lineárním prostorem tvaru. Informace o 3D hlavě a orientaci končetin jsou zachyceny díky kinematickému stromu, do kterého jsou ukládány matice relativních 3D rotací pro každý kloub. Předpovídání rotací též zajišťuje, že jsou končetiny symetrické a validní délky. Model se implicitně učí limity úhlů kloubů z datasetů 3D modelů těl.

Autoři metody přináší nový přístup pro rekonstrukci 3D sítě. Je využito propojení mezi 2D a 3D daty. Databáze 2D fotografií obsahují klíčové popisy jejich obsahu a 3D datasey sítě se skládají z různých póz a tvarů těl. Tyto oddělené soubory dat jsou v této práci spárovány. Podle fotografie musí vytvořená síť odvodit parametry 3D sítě a kamery tak, aby 3D klíčové body odpovídaly 2D klíčovým bodům po projekci. Tyto parametry jsou poslány do diskriminační sítě, jejímž úkolem je rozhodnout, zda 3D parametry odpovídají tělům reálných lidí nebo ne. Síť se implicitně učí limity úhlů pro každý kloub a je jí bráněno ve vytváření lidí s nepřírozenými tvary těl.

3.4 Automatic Estimation of 3D Human Pose and Shape from a Single Image

Tato kapitola byla zpracována dle zdroje [11].

Autory této metody jsou Federica Bogo², Angjoo Kanazawa³, Christoph Lassner^{1,4}, Peter Gehler^{1,4}, Javier Romero¹ a Michael J. Black¹ z ¹Max Planck Institute for Intelligent Systems, Tübingen, Germany, ²Microsoft Research, ³University of Maryland a ⁴University of Tübingen (2016).

Tato metoda je plně automatická a v odhadnuté 3D síti zachycuje z 2D fotografie i pózu i tvar postavy.

Uvedený přístup využívá vysoce kvalitní 3D model lidské postavy, který je natrénován z tisíců 3D skenů, a proto zachycuje různé variace tvarů těla v populaci stejně jako deformace těla spojené se změnou pózy.

Řešení problému je rozděleno do dvou hlavních kroků.

Nejprve jsou odhadnuty pozice 2D kloubů použitím dříve uvedené konvoluční neuronové sítě (CNN) zvané DeepCut [9]. Konvoluční neuronové sítě byly v minulosti úspěšné

v odhadu 2D lidské pózy, ale ne 3D pózy a tvaru z jediné fotografie. Proto byl přidán druhý krok, kterým je odhad 3D pózy a tvaru z 2D kloubů použitím 3D generativního modelu nazvaného SMPL [6] (viz Příloha).

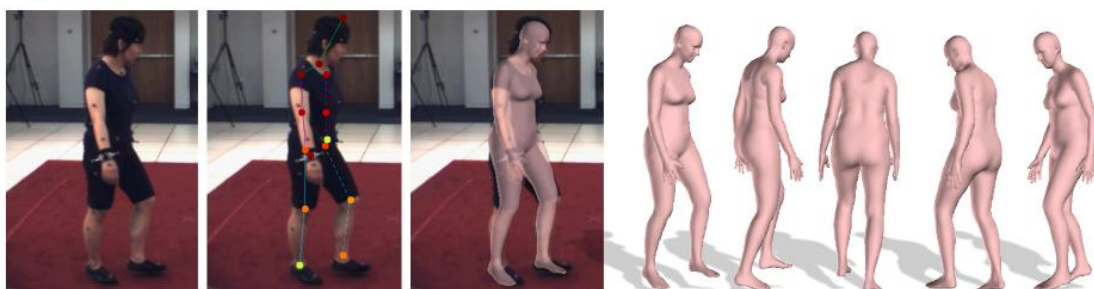
Je definována objektivní funkce a póza a tvar jsou optimalizovány přímo tak, aby se promítnuté klouby 3D modelu nacházely blízko 2D kloubů odhadnutých CNN. To je provedeno minimalizací chyby mezi 3D a 2D. Pasování pouze 2D kloubů vede k věrohodnému odhadu 3D tvaru těla.

Na daném 3D tvaru těla je pro zjednodušení definován set „kapslí“, které aproximují tvar těla. Rozměry kapslí jsou dány lineární regresí z parametrů modelu tvaru. Tato reprezentace dovoluje jednoduchý a efektivní výpočet. Je ukázáno, že tento postup též pomáhá bránit nesprávným pózám.

SMPL je závislý na pohlaví. To znamená, že rozlišuje mezi ženami a muži. Aby mohla být metoda plně automatická, je vytvořen též pohlavně neutrální model. Pokud není známo pohlaví, je pro napasování do fotografie použit právě tento model. Pokud je pohlaví známo, je pro získání lepších výsledků použit model konkrétního pohlaví.

Výsledný 3D model může být okamžitě pózován a animován. Metoda SMPLify je vyhodnocena na různých veřejně dostupných datasetech (Leeds Sports [12], HumanEva [13] a Human3.6M [14]) a je ukázáno, že poráží další state-of-the-art metody.

Obrázek 3.2 [11] zobrazuje kompletní postup této metody v grafické podobě.



Obrázek 3.2: Postup metody Keep it SMPL

Kód, 2D klouby a 3D modely uvedené ve zdrojovém textu jsou volně k dispozici pro vědecké účely.

3.5 Expressive Body Capture: 3D hands, Face and Body from a Single Image

Tato kapitola byla zpracována dle zdroje [15].

Autory této metody jsou Georgios Pavlakos^{1,2}, Vasileios Choutas¹, Nima Ghorbani¹, Timo Bolkart¹, Ahmed A. A. Osman¹, Dimitrios Tzionas¹ a Michael J. Black¹ z ¹MPI for Intelligent Systems, Tübingen, DE, ²University of Pennsylvania, PA, USA (2019).

Pro usnadnění analýzy lidských činů, interakcí a emocí je v této práci představena metoda, která umožňuje z jediné fotografie vytvoření 3D modelu lidského postavy včetně zachycení pózy těla a rukou a také výrazu obličeje. Za tímto účelem je pomocí tisíců 3D skenů natrénován nový unifikovaný 3D model lidského těla – SMPL-X (viz Příloha), který rozšiřuje původní model SMPL o ruce s plnohodnotnými klouby (dlaně a prsty) a o expresivní obličej. Odvodit parametry SMPL-X přímo z fotografií je bez spárovaných fotografií a bez odpovídajících 3D znalostí obtížné. Proto je pokračováno v přístupu dříve představené metody SMPLify, která odhaduje 2D rysy osoby a poté optimalizuje parametry modelu tak, aby model odpovídal těmto rysům.

Metoda SMPLify je vylepšena v následujících ohledech.

1. Jsou navíc detekovány 2D rysy (klouby) pro obličej, ruce a chodidla. Na tyto rysy je napasován nový model SMPL-X.
2. Z velkého MoCap datasetu je natrénována nová neuronová síť pro pasování modelu.
3. Je definována nová rychlá a přesná pokutová funkce proti vzniku samoprůniků modelů.
4. Jsou automaticky detekovány pohlaví osob a vybírány příslušné modely (mužský, ženský nebo neutrální).
5. Nová PyTorch implementace s využitím GPU dosahuje oproti staré Chumpy implementaci více než osminásobného urychlení.

Nová metoda, která se používá pro pasování SMPL-X modelu, je nazvána SMPLify-X. Dle výsledků autorů této práce je nová metoda s využitím nového modelu rychlejší a zároveň i mírně přesnější, než starší metoda SMPLify spolu s modelem SMPL. Zlepšení kvality modelu nastává hlavně v oblasti hlavy a rukou.

Obrázek 3.3 [15] zobrazuje výsledky metody SMPLify-X s modelem SMPL-X.



Obrázek 3.3: Metoda SMPLify-X s modelem SMPL-X

Modely SMPL-X, kód a data metody SMPLify-X jsou dostupné pro vědecké účely na <https://smpl-x.is.tue.mpg.de>.

3.6 Closing the Loop Between 3D and 2D Human Representations

Tato kapitola byla zpracována dle zdroje [16].

Autory této metody jsou Christoph Lassner^{1,2}, Javier Romero³, Martin Kiefel², Federica Bogo⁴, Michael J. Black² a Peter V. Gehler⁵ z ¹Bernstein Center for Computational Neuroscience, Tübingen, Germany, ²MPI for Intelligent Systems, Tübingen, Germany, ³Body Labs Inc., New York, United States, ⁴Microsoft, Cambridge, UK a ⁵University of Würzburg, Germany (2017).

Metoda Unite the People: Closing the Loop Between 3D and 2D Human Representations je vylepšením dříve představené metody SMPLify [11] uvedené v kapitole 3.4.

Prostor pro vylepšení lze nalézt v:

- nepřesných popisech,
- použití vyššího rozlišení a
- lepšímu umístění klíčových bodů z fotografie na SMPL modelu.

Data, kód a modely jsou dostupné zdarma pro vědecké účely na domovské stránce projektu: <http://up.is.tuebingen.mpg.de/>.

Pro vytvoření počátečního datasetu 3D těl je použita vylepšená verze metody SMPLify, která místo 2D klíčových bodů využívá model celého těla pózy a tvaru. Robustnější inicializace a další cíl pasování umožňuje algoritmus použít na klíčové body datasetů se standardní lidskou pózou. Lidé, kteří tvoří popisy dat, třídí napasování na dobrá a špatná. Tento postup vede k vytvoření počátečního datasetu UP-3D s kvalitními popisy.

Tento polo-automatizovaný přístup má několik výhod. Je snížen čas potřebný pro vytvoření popisů. Projekcí povrchů nebo klíčových bodů z pasování na originální fotografie jsou získány popisy, zatímco se zachovává výkon. Kvalitní reprezentace a flexibilní proces pasování umožňuje snadnou integraci datasetů s různými popisy (s různými lokacemi klíčových bodů).

Tato data mohou být použita pro natrénování modelů, které produkují výsledky s vysokými detaily. Modely predikují 31 oblastí a lokaci 91 klíčových bodů na těle.

Predikce z modelu s 91 klíčovými body zlepšuje metodu 3D pasování, která nejprve vygenerovala popisy pro natrénování klíčových bodů. Výsledky jsou ukázány na datasetech HumanEva [13] a Human3.6M [14].

3.7 Parametric Reshaping of Human Bodies in Images

Tato kapitola byla zpracována dle zdroje [17].

Autory této metody jsou Shizhe Zhou¹, Hongbo Fu², Ligang Liu¹, Daniel Cohen-Or³ a Xiaoguang Han¹ z ¹Zhejiang University, ²City University of Hong Kong a ³Tel Aviv University (2010).

Metoda využívá modelového přístupu a slouží především pro úpravy lidského těla. Pro dosažení deformací v jednotlivých částech těla je využito 3D upravovatelného modelu lidského těla. Přístup je založen na dvou klíčových pozorováních. Zaprvé, existující metody cílí na věrnou rekonstrukci 3D tvaru, zatímco cíl autorů této metody je co nejvíce napodobit 3D model tvaru na fotografii. Zadruhé, retušování profesionálních umělců naznačuje, že změny lidských těl na fotografiích jsou ve velké míře řízeny změnami 2D obrysů těla.

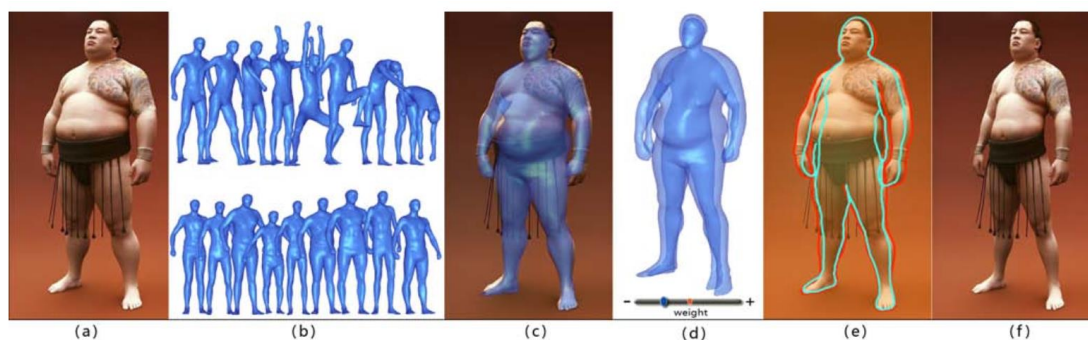
Model je nejprve přibližně napasován na fotografii s minimální interakcí uživatele. Uživatel poté může napasovaný model upravit několika atributy, jako jsou váha a výška. Nakonec je tvar na fotografii deformován tak, aby změny s ohledem na kostru (změny

převážně na obrysu postavy) co nejvíce odpovídaly změnám na 3D modelu. Díky práci v 2D prostoru fotografie je možné se vyhnout obtížnému problému přesné rekonstrukce 3D modelu a jeho textury. Technika je proto bez potíží použitelná pro zpracování postav s volným ošacením a libovolnými pózami.

Jako model, který je pasován do fotografie, je využitý SCAPE [5] (viz Příloha). Ten vyžaduje řešení parametrů pózy a tvaru těla. Při řešení je využita pomoc uživatele. Pro tento účel je vytvořeno uživatelské rozhraní, které od uživatele vyžaduje pouze malé množství počátečních vstupů a umožňuje případné vylepšování. Aby byla metoda interaktivní v reálném čase, je upřednostněna rychlost před přesností. To je provedeno zjednodušením komplexní optimalizace pasování [18].

Parametrický charakter této techniky dovoluje uživatelům jednoduše dosáhnout rozličných výsledků úpravou malého setu posuvných tlačítek odpovídajících jednotlivým vlastnostem těla. Rychlý algoritmus poskytuje odezvu na úpravy v reálném čase, což dovoluje uživateli dosáhnout požadovaných výsledků bez velké námahy. Je ukázáno, že tato technika je efektivní pro lidská těla s rozdílnými pózami, tvary a ošaceními pro běžné i profesionální fotografie.

Obrázek 3.4 [17] zobrazuje kompletní postup této metody v grafické podobě.



Obrázek 3.4: Postup metody Parametric Reshaping

3.8 Photo Wake-Up 3D Character Animation from a Single Photo

Tato kapitola byla zpracována dle zdroje [19].

Autory této metody jsou Chung-Yi Weng¹, Brian Curless¹ a Ira Kemelmacher-Shlizerman^{1,2} z ¹University of Washington a ²Facebook Inc. (2018).

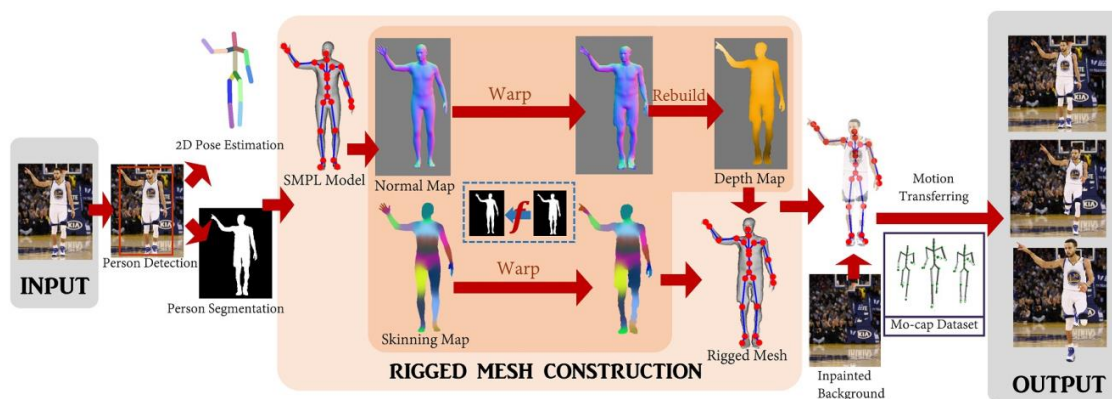
V práci Photo Wake-Up je představena metoda a aplikace pro animaci osoby na jedné fotografii. Postava může chodit, běhat, sedět nebo skákat ve 3D prostoru. Tento systém jako první umožňuje právě tuto 3D animaci oblečené postavy na základě jedné fotografie.

V této metodě jsou nejprve aplikovány state-of-the-art algoritmy pro detekci osoby, segmentaci a odhad 2D pózy. Pro detekci a segmentaci osoby je použita metoda „Mask R-CNN“ [20]. Odhad 2D pózy těla je provedený metodou [21] a segmentace osoby je vylepšena pomocí Dense CRF [22]. Jakmile je osoba vysegmentována z fotografie, je pro vyplnění míst, kde byla osoba, použitý algoritmus „PatchMatch“ [23].

Na výsledky je použita vytvořená metoda pro získání 3D sítě. Tuto síť lze již libovolným způsobem animovat.

Metoda Photo Wake-Up podle autorů funguje s velkým množstvím, převážně zepředu focených, fotografií s celým tělem. Navíc je uživateli dána možnost upravit postavu na obrázku, zobrazit si 3D rekonstrukci a prozkoumat ji v rozšířené realitě.

Obrázek 3.5 [19] zobrazuje kompletní postup této metody v grafické podobě.



Obrázek 3.5: Postup metody Photo Wake-Up

3.9 Physics-Inspired Garment Recovery from a Single-View Image

Tato kapitola byla zpracována dle zdroje [24].

Autory této metody jsou Shan Yang, Zherong Pan, Tanya Amert, Ke Wang, Licheng Yu, Tamara Berg a Ming C. Lin z University of North Carolina at Chapel Hill (2018).

Tento algoritmus nedokáže zachytit pouze celkový tvar a geometrii ošacení, umí také extrahovat malé detaily, jako jsou záhyby. Detailní ošacení je z jediné fotografie získáno využitím statistických, geometrických a fyzikálních vlastností a kombinací odhadu

parametrů, sémantickým rozbohem, znovuzískáním tvaru a fyzikálně založenou simulací ošacení.

Místo reprezentace oblečené osoby jednou sítí je pro ošacení definována vlastní síť, což dovoluje modelovat bohaté interakce mezi ošacením a lidským tělem. Takový přístup také pomáhá zachytit záhyby v ošacení způsobené různými zdroji.

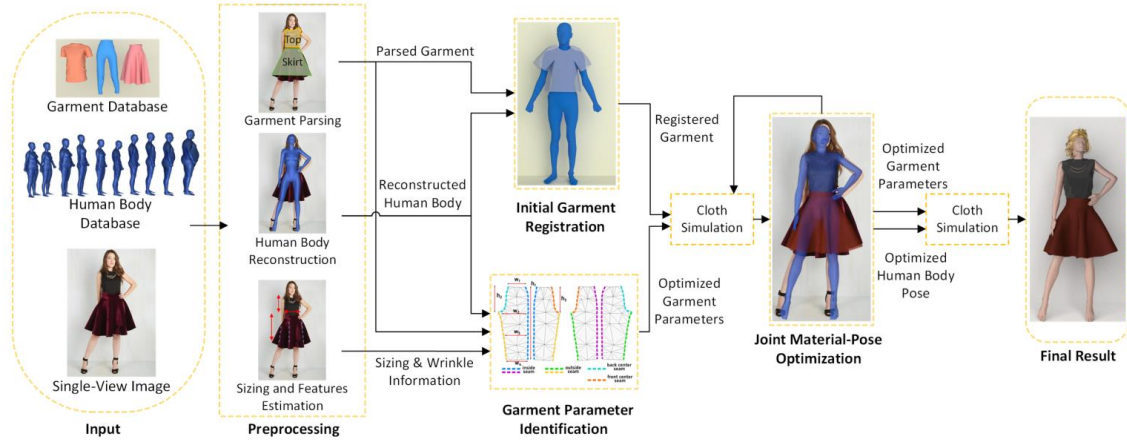
S omezeným vstupem informací z jediné fotografie je řešení problému rozloženo do tří kroků. Prvním krokem je vytvoření statistického modelu těla sestaveného ze setu nahých lidských těl. Statistický model je použitý pro extrakci tvaru a pózy těla na fotografii. Druhou důležitou částí je kolekce různých švů na běžných typech ošacení – sukně, kalhoty, krátké kalhoty, trika a šaty z databáze vzorů všech typů ošacení. Třetí částí je set všech možných konfigurací a dynamických stavů ošacení modelovaných fyzikální simulací. Simulace přidává dodatečná 3D fyzikální omezení, která chybí na 2D fotografii.

Postup metody je následující. Pro vytvoření přesného modelu těla zvolí uživatel na fotografii 14 pozic kloubů a načrtne hrubý obrys lidské siluety (tento krok může být též automatizován použitím technik pro zpracování obrazu). Z této informace je použitý statistický model k automatickému vytvoření 3D sítě lidského těla na fotografii. Pro odhad modelu ošacení je nejprve spočítán sémantický rozbor oděvu na fotografii pro identifikaci a lokalizaci jednotlivých částí ošacení. Sémantická segmentace je spočítána automaticky metodou pro rozpoznávání ošacení [25]. Výsledek je poté použitý k získání informací o velikostech – obvod pasu nebo délka sukně. Tyto informace jsou následně použity pro namapování vyobrazeného ošacení na existující vzory ošacení a úpravu vzorů švů na základě extrahovaných parametrů. Též jsou analyzovány segmentované kusy ošacení pro nalezení lokace a hustoty záhybů v rekonstruovaných kusech ošacení, což je nutné pro odhad materiálů ošacení.

Jakmile jsou získány modely těla a ošacení, je proveden proces identifikace parametrů z fotografie, který optimalizuje parametry vzorového ošacení podle rekonstruovaného lidského těla a informací z fotografie. 3D vzorová síť ošacení je napasována na lidské tělo, čímž je získáno počáteční 3D ošacení, a poté jsou společně optimalizovány parametry materiálů, tvaru těla a pózy pro získání konečného výsledku.

Obrázek 3.6 [24] zobrazuje kompletní postup metody v grafické podobě.

Metody vytváření 3D modelu postavy z barevné fotografie



Obrázek 3.6: Postup metody Garmet Recovery

4 Srovnání metod

Všechny dříve uvedené metody se shodují v jednom aspektu. Jako vstup jim postačuje jediná barevná fotografie, na které je vyobrazena postava člověka z předního pohledu. Z této 2D fotografie postavy je vytvářen 3D model těla.

Hlavní rozdíl metod je v pojetí přístupu k řešení daného problému. Metody lze rozdělit na parametrické a nparametrické (převážně vedou k naskenovanému mraku bodů; nejsou uvedeny v této práci).

Ne všechny metody lze použít pro ženské i mužské pohlaví. Metodu 3.1 lze pro ženy použít pouze po rozšíření zdrojové databáze. Metoda 3.4 umožňuje také použití modelu neutrálního pohlaví.

Některé způsoby vytváření 3D avatarů vyžadují manuální zásah uživatele. Nejsou tak plně automatické. Mezi takové metody lze zařadit metodu 3.2 (nutné ohodnocení postavy) nebo 3.7.

Rozdílem je i různá pokročilost porovnávaných metod. Pomocí metody 3.5 lze modelovat i výraz obličeje a pozici dlaní a prstů, jiné metody toto neumožňují. Metoda 3.9 umožňuje modelování oblečení osoby.

Další rozdíl je v uvolnění zdrojových kódů jednotlivých metod. Autoři některých metod dali kód své aplikace volně k dispozici (3.4, 3.5, 3.6), zatímco jiní tento krok nečinili. Metodu 3.2 si lze vyzkoušet online, avšak kód uvolněn není.

5 Použitá metoda

V této kapitole je důkladně popsána zvolená metoda, kterou je 3.4 Automatic Estimation of 3D Human Pose and Shape from a Single Image. V práci je využitý model SMPL a systém, který vytváří 3D síť z 2D fotografie, je autory nazván SMPLify¹.

Hlavní důvody pro zvolení metody jsou:

- Metoda je plně automatická.
- Použití kvalitního modelu SMPL.
- Veřejně dostupný kód aplikace.

Metoda nejprve pomocí konvoluční neuronové sítě DeepCut CNN odhadne 2D klouby J_{est} na jediné vstupní barevné fotografii. Pro každý 2D kloub i poskytuje neuronová síť míru věrohodnosti w_i . Následně je napasován 3D model lidského těla tak, aby byla minimalizována vážená chyba promítnutých kloubů.

Obrázek 5.1 [11] zobrazuje možné dosažitelné výsledky této metody (zleva: originální fotografie, napasovaný 3D model, 3D model vyrenderován z jiného pohledu).



Obrázek 5.1: Dosažitelné výsledky metody SMPLify

Model těla je definován jako funkce $M(\beta, \theta, \gamma)$, kde β je tvar, θ póza a γ translace. Výstupem funkce je trojúhelníkový povrch M s 6890 vrcholy. Parametry β jsou koeficienty nízko-dimenzionálního prostoru tvaru naučené z trénovacího setu tisíců skenů. Je použit jeden ze tří modelů tvaru (mužský, ženský a pohlavně neutrální model). SMPL uvažuje pouze mužský a ženský model. Autoři této metody natrénovali nový pohlavně neutrální model z 2000 mužských a 2000 ženských těl. Tento nový model dovoluje metodě, aby byla plně automatická. Pokud je pohlaví známé, použije se pro

¹ Pokročilá metoda SMPLify-X s modelem SMPL-X nebyla použita, jelikož byla zveřejněna až v průběhu zpracování diplomové práce.

získání lepších výsledků konkrétní odpovídající model. Model lze rozeznat podle barvy: růžový je pohlavně závislý a světle modrý je pohlavně neutrální.

Póza těla je definována kostrou s 23 klouby. Parametry pózy θ reprezentují úhly rotace mezi jednotlivými částmi. $J(\beta)$ je funkce, která předvídá umístění 3D kloubů kostry z tvaru těla. U SMPL jsou klouby sparse lineární kombinací vrcholů na povrchu nebo ekvivalentně funkcí koeficientů tvaru. Klouby mohou být umístěny do libovolných pozic užitím globální rigidní transformace.

Klouby i modelu dané pózy jsou označeny $R_\theta(J(\beta)_i)$, kde R_θ je globální rigidní transformace daná pózou θ . SMPL definuje deformace závislé na póze. Pro pohlavně nezávislý model jsou použity ženské transformace, které se ukázaly jako dostatečné. Model SMPL a kostra DeepCut mají mírně odlišné klouby. Proto jsou DeepCut klouby přiřazeny nejpodobnějším SMPL kloubům. Pro projekci SMPL kloubů na fotografii je použitý perspektivní model kamery definován parametry K .

5.1 Aproximace těla kapslemi

Předchozí metody mohou produkovat 3D pózy, které jsou fyzikálně nemožné. To je dané následkem průniků mezi jednotlivými částmi těla. Výhodou 3D modelu tvaru této metody je, že tento problém umožňuje detekovat a zabránit mu. Výpočet průniků je ovšem náročný pro komplexní nekonvexní povrchy, jako je tělo. Povrch těla je proto aproximován souborem „kapslí“. Obrázek 5.2 [11] zobrazuje tyto kapsle na dvou objektech (zleva: původní tvar, tvar aproximovaný kapslemi, změna pózy kapslí; žlutý mrak bodů reprezentuje vrcholy aproximovaného modelu). Každá kapsle má svůj poloměr a délku osy.



Obrázek 5.2: Aproximace těla kapslemi

Z parametrů tvaru modelu je natrénován regresor na parametry kapslí (délka osy a poloměr) a jsou nastaveny do pózy podle R_θ . Nejprve je napasováno 20 kapslí (jedna

pro každou část těla s výjimkou prstů na ruce a nohou) na povrch trénovacích tvarů těla bez pózy pro naučení SMPL. Kapsle jsou manuálně umístěny na klouby vzoru a je provedena na gradientu založená optimalizace jejich poloměrů a délek os pro minimalizaci obousměrné vzdálenosti mezi kapslemi a povrchem těla. Následně je naučený lineární regresor z koeficientů tvaru těla. Jakmile je regresor natrénovaný, je procedura zopakována a jsou inicializovány kapsle s výstupem regresoru.

5.2 Objektivní funkce

Pro napasování 3D pózy a tvaru na 2D klouby nalezené pomocí CNN je minimalizována objektivní funkce, která je součtem pěti chybových výrazů: chyba týkající se kloubů, tři chyby pózy a chyba tvaru. Dostáváme výraz (1).

$$E(\beta, \theta) = E_J(\beta, \theta; K, \beta, J_{est}) + \lambda_\theta E_\theta(\theta) + \lambda_a E_a(\theta) + \lambda_{sp} E_{sp}(\theta; \beta) + \lambda_\beta E_\beta(\beta), \quad (1)$$

kde K jsou parametry kamery a λ_θ , λ_a , λ_{sp} a λ_β jsou skalární váhy.

Následující výraz (2) pokutuje vážené 2D vzdálenosti mezi odhadnutými klouby J_{est} a odpovídajícími SMPL klouby po projekci.

$$E_J(\beta, \theta; K, J_{est}) = \sum_{kloub\ i} w_i \rho(\Pi_K(R_\theta(J(\beta)_i)) - J_{est,i}), \quad (2)$$

kde Π_K je projekce z 3D do 2D daná kamerou s parametry K . Příspěvek každého kloubu je vážen mírou důvěryhodnosti jeho odhadu w_i poskytnuté CNN. Pro vypořádání se se zašumělými odhady je použita robustní diferencovatelná Geman-McClure pokutová funkce ρ [26].

Následující výraz (3) pokutuje nepřirozené ohyby loktů a kolen.

$$E_a(\theta) = \sum_i \exp(\theta_i), \quad (3)$$

kde se sčítá přes parametry pózy (rotace) odpovídající ohybu kolen a loktů. Exponenciála silně pokutuje rotace porušující přirozená omezení (hyperextenzi). Pokud kloub není ohnutý, hodnota θ_i je nula. Negativní ohyb je přirozený, a tak není příliš pokutován, zatímco pozitivní ohyb je nepřirozený, a proto je pokutován více.

Většina metod pro odhad 3D pózy využívá určité apriorní informace pro upřednostnění pravděpodobných póz před nepravděpodobnými. Jako mnoho dřívějších metod je

i metoda SMPLify natrénována pro získání apriorní informace na CMU datasetu [27]. Za předpokladu, že se pózy velmi odlišují, je důležité reprezentovat multi-modální podstatu dat a zároveň udržet zvladatelnou výpočetní složitost. Pro vytvoření apriorní informace jsou použity pózy získané pasováním SMPL na CMU data pomocí MoSh [28]. Následně je napasována směs Gaussovských funkcí na přibližně jeden milión póz sta subjektů. Použití smíšeného modelu přímo v optimalizaci je výpočetně problematické, jelikož je potřeba optimalizovat záporný logaritmus sumy. Suma Gaussovských funkcí je proto aproximována operátorem maxima v následující rovnici (4).

$$\begin{aligned} E_{\theta}(\theta) &\equiv -\log \sum_j \left(g_j \mathcal{N}(\theta; \mu_{\theta,j}, \Sigma_{\theta,j}) \right) \approx -\log \left(\max_j \left(c g_j \mathcal{N}(\theta; \mu_{\theta,j}, \Sigma_{\theta,j}) \right) \right) = \\ &= \min_j \left(-\log \left(c g_j \mathcal{N}(\theta; \mu_{\theta,j}, \Sigma_{\theta,j}) \right) \right), \end{aligned} \quad (4)$$

kde g_j jsou váhy smíšeného modelu osmi Gaussovských funkcí a c je kladná konstanta potřebná v implementaci řešiče.

Obecná Gaussovská funkce [29] v N-D prostoru je definována jako

$$G(\vec{x}, \sigma) = \frac{1}{(\sqrt{2\pi}\sigma)^N} \exp\left(-\frac{|\vec{x}|^2}{2\sigma^2}\right), \quad (5)$$

kde σ je směrodatná odchylka (šířka kernelu).

Je definován výraz využívající aproximaci kapslemi, který pokutuje vnitřní samoprůniky. Chyba je spojována s průnikem objemů „nekompatibilních“ kapslí (kapsle, které se neprotínají v přirozených pózách). Jelikož objem protínajících se kapslí není snadno vypočitatelný, jsou kapsle dále zjednodušeny na koule se středy $C(\theta, \beta)$ a poloměrem $r(\beta)$. Pro každou kouli se uvažuje 3D izotropní Gaussovská funkce s $\sigma(\beta) = \frac{r(\beta)}{3}$ a pokuta se definuje jako škálovaná verze integrálu součinu Gaussovských funkcí vztahujících se k „nekompatibilním“ částem.

$$E_{sp}(\theta; \beta) = \sum_i \sum_{j \in I(i)} \exp\left(\frac{\|C_i(\theta, \beta) - C_j(\theta, \beta)\|^2}{\sigma_i^2(\beta) + \sigma_j^2(\beta)}\right), \quad (6)$$

kde se sčítá přes všechny koule i a $I(i)$ jsou koule nekompatibilní s i . Je nutné podotknout, že výraz (6) pokutuje, ale ne přímo zabraňuje průnikům. Je avšak splněný

požadavek diferencovatelnosti s ohledem k póze a tvaru. Výraz není použitý při optimalizaci tvaru, jelikož by upřednostňoval tenké tvary těl, aby se vyhnul průnikům.

Pro tvar je použitý následující výraz (7) $E_{\beta}(\beta)$.

$$E_{\beta}(\beta) = \beta^T \Sigma_{\beta}^{-1} \beta, \quad (7)$$

kde Σ_{β}^{-1} je diagonální matice s druhými mocninami singulárních hodnot odhadnutých pomocí analýzy hlavních komponent z tvarů trénovacího setu SMPL.

5.3 Optimalizace

Je předpokládáno, že posun kamery a orientace těla jsou neznámé. Je však požadována znalost ohniskové vzdálenosti kamery nebo alespoň její hrubý odhad. Posun kamery (ekvivalentně γ) je inicializován předpokladem, že osoba stojí rovnoběžně s rovinou fotografie. Hloubka je odhadnuta poměry podobných trojúhelníků, které jsou definovány délkou trupu průměrného tvaru SMPL a odhadnutých 2D kloubů. Protože tento předpoklad není vždy správný, je odhad dále zlepšen minimalizací E_J přes samotné klouby trupu s ohledem na posun kamery a orientaci těla. Během této optimalizace je β zafixována na průměrný tvar. Ohnisková vzdálenost není optimalizována, protože problém neobsahuje dostatek omezujících podmínek pro optimalizaci společně s posunem.

Po odhadu posunu kamery je model napasován minimalizací rovnice (1). Bylo pozorováno, že začátek s vysokými hodnotami λ_{θ} a λ_{β} a jejich postupné snižování během jednotlivých fází optimalizace vede k efektivnímu způsobu vyhnutí se lokálnímu minimu.

Pokud je subjekt zachycen z bočního pohledu, vyhodnocení, v jakém směru je tělo postaveno, může být nejednoznačné. Proto jsou vyzkoušeny dvě inicializace, pokud je 2D vzdálenost mezi 2D ramenními klouby odhadnutými CNN menší než zadaný práh. Jedna verze je inicializace s orientací těla popsané výše a druhá s orientací otočenou o 180 stupňů. Nakonec je vybrána možnost s nejnižší E_J .

Rovnice (1) je minimalizována pomocí metody Powell's dogleg [29], OpenDR [30] a Chumpy [31]. Optimalizace jedné fotografie trvá méně než jednu minutu na běžném stolním počítači.

6 Návrh experimentu

Autoři metody SMPLify uvolnily její zdrojové kódy. K dispozici je metoda pro automatické vytvoření 3D modelu z jedné fotografie a k tomu i potřebný detektor kloubů osoby. Pro běžného uživatele však chybí jednoduchá možnost realizace metody a zobrazení a prohlížení výsledného 3D modelu.

V diplomové práci bude vytvořena internetová aplikace, jejímž vstupem bude fotografie postavy a jejímž výsledkem bude možnost interaktivního prohlížení získaného modelu. Fotografie bude nahrána na server, kde bude algoritmem zpracována. Získaný 3D model bude následně uživateli zobrazen v interaktivním okně (možnost prohlížení, pohybu kamery, přiblížení). Pokud bude mít uživatel zájem, může sdílet spokojenost s podobou své postavy se získaným modelem. Tyto nepovinné údaje budou uloženy do databáze.

Rozměry modelů a kvalita metody SMPLify budou dále vyhodnoceny na mužské a ženské postavě s pomocí statistických testů. Měřeno bude celkem osm tělesných partií, konkrétně obvod hrudníku, pasu, boků, zápěstí, stehna, kolena, kotníku a krku. Kromě těchto partií bude měřena i výška a rozpětí paží osoby.

Dále bude přidána funkce pro automatický odhad výšky postavy. Výška bude odhadována podle výšky dveří, u kterých se uživatel vyfotí. Zjištěná odhadnutá výška poslouží pro naškálování získaného modelu na rozměry, které by měly co nejvíce odpovídat skutečným mírám uživatele. Bylo zjištěno, že bez škálování modelu metoda SMPLify z fotografie nedokonale odhaduje výšku postavy a vytvářené modely jsou celkově ve všech ohledech výrazně větší, než jaké jsou skutečné míry vyfocené osoby.

Získaná data budou statisticky vyhodnocena v tabulkovém procesoru MS Excel 2016. Využité programovací jazyky při tvorbě diplomové práce jsou Python (verze 2 a 3), html (pro tvorbu internetové aplikace), PHP (pro ovládání a správu internetové aplikace) a skriptovací jazyk Bash interpretující příkazový řádek v Unix prostředí (pro snadné spouštění jednotlivých skriptů).

V diplomové práci jsou v kapitole 8.2 Výška osoby uvedena časová měření běhu algoritmu pro odhad výšky postavy. Pro měření času běhu je využita Python knihovna „timeit“ [32]. Na začátku skriptu je zavolána funkce „default_timer“ a hodnota uložena do proměnné, obdobný postup je proveden na konci běhu. Výsledný čas je pak dán odečtením hodnot získaných na konci a na začátku.

Pro porovnání kvalit vybrané metody SMPLify poslouží následující statistická kritéria:

- Minimum – nejmenší hodnota celého souboru
- Maximum – největší hodnota celého souboru
- Střední hodnota – Střední hodnota odpovídá hodnotě, která je v seřazeném souboru přesně uprostřed. Nezáleží tedy na hodnotě čísla, ale její pozici vůči ostatním v seřazeném souboru. Pokud je počet prvků lichý, je jako střední hodnota brán průměr dvou hodnot uprostřed tohoto souboru.
- Průměr

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i, \quad (8)$$

kde x_i jsou funkční hodnoty (jednotlivé rozměry modelu) a n je počet vzorků.

- Směrodatná odchylka

$$\sigma = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2}, \quad (9)$$

kde x_i jsou funkční hodnoty, \bar{x} je průměr a n je počet vzorků.

- Výběrová směrodatná odchylka

$$s = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2}, \quad (10)$$

kde x_i jsou funkční hodnoty, \bar{x} je průměr a n je počet vzorků.

- Rozdíl střední hodnoty a naměřené hodnoty

$$r = |E - y|, \quad (11)$$

kde E je střední hodnota získaných rozměrů a y je skutečná naměřená hodnota.

- Jednovýběrový T-test [33]

$$T = \frac{\bar{x} - \mu_0}{s} \cdot \sqrt{n}, \quad (12)$$

kde \bar{x} je průměr, μ_0 je střední hodnota, s je výběrová směrodatná odchylka a n je počet vzorků. Kritická hodnota pro hladinu α je

$$t_{1-\frac{\alpha}{2}}(n-1), \quad (13)$$

kde t je kvantil Studentova t -rozdělení a n je počet stupňů volnosti.

- Interval spolehlivosti [34]

$$\left(\bar{x} - \frac{s}{\sqrt{n}} \cdot z_{1-\frac{\alpha}{2}}, \bar{x} + \frac{s}{\sqrt{n}} \cdot z_{1-\frac{\alpha}{2}} \right), \quad (14)$$

kde \bar{x} je průměr, $z_{1-\frac{\alpha}{2}}$ je kvantil normovaného normálního rozdělení, α je % intervalu spolehlivosti, s je výběrová směrodatná odchylka a n je počet vzorků.

Interval spolehlivosti [34] je takový interval, uvnitř kterého se hodnota neznámého parametru vyskytuje s požadovanou pravděpodobností. Typickými hodnotami spolehlivosti bývají zvoleny 90–99 %. V tomto případě jsou zvoleny hodnoty hladiny $\alpha = 95\%$ a $\alpha = 99\%$. Interval se určuje na základě již získaných statistických dat a odhady jsou provedeny pro rozměry jednotlivých partií testované osoby.

Jednovýběrový T-test [33] umožňuje ověřit hypotézu, zda normální rozdělení, z něhož pochází náhodný výběr (s neznámým rozptylem), má zadanou střední hodnotu. V tomto případě je pomocí T-testu ověřováno, zda jednotlivé získané rozměry modelů z různých fotografií (tvořící náhodný výběr) odpovídají skutečným naměřeným hodnotám osoby. Kromě jednovýběrového T-testu existuje také párový T-test a dvouvýběrový T-test. Párový T-test umožňuje porovnání středních hodnot mezi prvními a druhými prvky uspořádaných dvojic. Dvouvýběrový T-test slouží k porovnání středních hodnot dvou různých skupin.

Získanou testovou statistiku (p -value, p -hodnota) porovnááme s kritickou hodnotou testu pro vybranou hladinu α . Hladina spolehlivosti pro T-test je zvolena $\alpha = 95\%$. Testovaná nulová hypotéza H_0 a alternativní hypotéza H_1 jsou definovány

$$H_0: \mu = \mu_0 \text{ a}$$

$$H_1: \mu \neq \mu_0,$$

kde μ je střední hodnota a μ_0 je konstanta. μ_0 bude při testování představovat skutečnou výšku postavy a její naměřené tělesné míry.

Pokud platí

$$|T| < t_{1-\frac{\alpha}{2}}(n-1), \quad (15)$$

nulovou hypotézu H_0 nezamítáme. V opačném případě je H_0 zamítnuta a přijata H_1 . Nezamítnutí H_0 znamená, že naměřená hodnota modelu odpovídá se zadanou hladinou spolehlivosti reálné naměřené hodnotě postavy. Zamítnutí H_0 znamená, že naměřená hodnota modelu se zadanou mírou spolehlivosti neodpovídá reálné naměřené hodnotě postavy.

Modely jsou vytvořeny dle fotografií s osobami v následujících typech oblečení:

- spodní prádlo / plavky,
- úzké kalhoty, úzké triko,
- úzké kalhoty, úzké triko + úzká mikina,
- úzké kalhoty, úzké triko + úzká mikina + lehká bunda,
- volné kalhoty, volné triko,
- volné kalhoty, volné triko + volná mikina,
- volné kalhoty, volné triko + volná mikina + zimní bunda,
- volné kalhoty, volné triko + volná mikina + zimní bunda + kapuce, rukavice

a v těchto pózách:

- ruce v bok,
- ruce překřížené přes prsa,
- ruce za hlavou,
- ruce vzhůru,
- překřížené nohy,
- chodidlo jedné nohy opřené o koleno druhé nohy, ruce v bok,
- ruce překřížené přes prsa, dlaně na ramena,
- dřep.

Pro různé oblečení je osoba vyfocena v základní přirozené póze (ruce volně podél těla). Pózy jsou vyfoceny v oblečení typu úzké kalhoty a úzké triko, stejné oblečení je použito i pro vytvoření intervalů spolehlivosti. Osoba vždy stojí přímo proti fotografovi, nenatáčí se bokem.

Obrázek 6.1 zobrazuje ukázkou typů oblečení použitých pro vytváření modelů (zleva – druhý, pátý a osmý typ). Obrázek 6.2 zobrazuje ukázkou typů póz použitých pro vytváření modelů (zleva – třetí, čtvrtá a šestá póza).



Obrázek 6.1: Ukázka typů oblečení



Obrázek 6.2: Ukázka typů póz

7 Realizace, implementace

Následující podkapitoly ukazují detaily realizace a implementace jednotlivých funkcí algoritmů a jejich nastavení a požadavků.

7.1 Internetová aplikace

Pro internetovou aplikaci je využito programovacích jazyků html a PHP.

Úvodní stránka aplikace je vertikálně rozdělena na tři třetiny.

Levá část zobrazuje jednoduchý formulář pro nahrání fotografie a zadání pohlaví. Pokud je vybráno konkrétní pohlaví, aplikace pro vytvoření modelu použije pasování na model právě s tímto pohlavím. Pokud uživatel pohlaví uvádět nechce anebo nechá příslušnou kolonku formuláře nevyplněnou, je použit zdrojový model s „neutrálním“ typem pohlaví.

Prostřední část okna zobrazuje ukázkovou fotografii, jak se má uživatel vyfotit, a krátký informační popis této fotografie. Část okna úplně vpravo pak již představuje vizualizaci vytvořeného modelu v podobě statické fotografie. Implementace se nachází v příloženém souboru „index.php“.

Po stisknutí tlačítka pro odeslání fotografie dochází ke zkontrolování souboru, zda se jedná opravdu o fotografii, zda není velikost souboru příliš velká (soubory nad 5 MB jsou zamítnuty) a zda má správnou koncovku (přípustné jsou .jpg, .jpeg a .png). Pokud je zjištěno, že vybraná fotografie se již nachází v cílové složce od dřívějšího nahrání (stejně jméno fotografie), je rovnou přistoupeno k zobrazení již dříve získaného modelu. Implementace se nachází v příloženém souboru „upload.php“.

V opačném případě je po kontrole fotografie nahrána do předem definované složky a je spuštěn hlavní spouštěcí bash skript (příložený soubor „run_smpl.sh“). Skript je spuštěn pomocí PHP funkce „shell_exec“, která vykoná zadaný příkaz pomocí shellu. Vstupními parametry skriptu jsou právě nahraná fotografie určená ke zpracování algoritmy a zvolené pohlaví pro výběr správného 3D modelu. Celý proces zahrnuje v následujícím pořadí

- zmenšení rozlišení fotografie,
- provedení algoritmu DeepCut a získání odhadu pozice kloubů a
- provedení algoritmu SMPLify a získání modelu odpovídajícího osobě na fotografii.

V tomto případě je použita neupravená verze SMPLify, to znamená bez hledání dveří, odhadu výšky a škálování modelu (popsáno dále). Rozměry modelu také nejsou měřeny. Důvodem je úspora času, aby byl získaný model co nejdříve připraven k zobrazení. Škálování modelu také není potřeba, jelikož bez srovnání s jinými předměty nelze o rozměrech modelu rozhodnout, zda jsou správné nebo ne (není poznat, zda je model vysoký např. 150 cm nebo 200 cm). Co se týká rozměrů, jde subjektivně rozhodnout pouze o správných poměrech mezi velikostmi partií. Od uživatelů využívajících aplikaci také není vyžadováno žádné měření tělesných měř.

Po dokončení procesů je zobrazena finální stránka s výsledky. Přesměrování na tuto stránku je provedeno pomocí PHP funkce „header“. Stránka s výsledky je vertikálně rozdělena na dvě poloviny.

Levá polovina okna obsahuje formulář pro odeslání spokojenosti se zobrazeným 3D modelem. Odeslání výsledků není povinné a je ponecháno na uživateli dobrovolnosti. Ukládané údaje uživatele jsou

- jméno,
- příjmení,
- věk,
- pohlaví,
- model (jeho název),
- hodnocení podobnosti s modelem a
- případné problémy s modelem nebo s aplikací.

Pohlaví není nutné opětovně vybírat, je uloženo již z výběru provedeného na úvodní stránce. Stejně tak jméno modelu je v případě odeslání formuláře uloženo automaticky bez zadávání uživatele.

Hodnocení podobnosti s modelem je rozděleno do následujících pěti podkategorií:

- „Tvar, velikost a póza modelu vypadá přesně jako já.“
- „Velikost a póza modelu vypadá jako já, ale tvar je odlišný.“
- „Tvar a póza modelu vypadá jako já, ale velikost je odlišná.“
- „Tvar a velikost modelu vypadá jako já, ale póza je odlišná.“
- „Tvar modelu vypadá jako já, ale velikost a póza je odlišná.“
- „Velikost modelu vypadá jako já, ale tvar a póza je odlišná.“

- „Póza modelu vypadá jako já, ale tvar a velikost je odlišná.“
- „Nevím.“
- „Model vůbec nevypadá jako já.“

Stránka obsahuje popis jednotlivých kritérií hodnocení, která jsou

- tvar – celková podoba s modelem, tvar postavy,
- velikost – velikostní poměry mezi jednotlivými tělesnými partiemi a
- póza – tělesná póza postavy.

Pro případné problémy s aplikací nebo další slovní hodnocení modelu je připraveno textové okno, kam uživatel může vyplnit libovolný text.

Pokud se uživatel rozhodne pro odeslání výsledků, je vyplnění všech základních údajů kromě případných problémů povinné, žádnou jinou kategorii nelze vynechat. Hodnocení je ponecháno pouze slovní a subjektivně vyjádřené, aby byl celý proces pro uživatele co nejrychlejší a nemusel se měřit. Zadávat své číselné údaje by pro některé osoby mohlo být nepříjemné a příliš osobní a mohlo by to být důvodem pro neodeslání výsledků (případně pro úplné nevyužití aplikace). Zároveň jsou všechny texty aplikace psány v anglickém jazyce, aby aplikaci mohlo v budoucnu využít co největší množství různých lidí.

Po kliknutí na tlačítko pro sdílení je spuštěn PHP úsek skriptu (stejný kód se nachází v příloženém souboru „details.php“ – kód je po stisknutí tlačítka možné spouštět buď přímo jako interní kód nebo jako kód umístěný v jiném souboru). Tento skript vytvoří spojení s připravenou databází, do které uloží hodnoty zadané ve formuláři. Připravená MariaDB databáze musí obsahovat sloupce pro ukládaná data. S každým novým odesláním dojde k automatickému zvýšení číselného id a uložení výše uvedených údajů. Jméno a příjmení jsou ukládány v textové podobě, věk jako číslo a hodnocení podoby modelu je uloženo též jako číslo v rozmezí 1 – 9. 1 značí největší spokojenost (1. výše uvedená volba) a 9 značí největší nespokojenost (9. výše uvedená volba). 8. výše uvedená volba je určena pro ty, kteří si nedokáží vybrat žádnou jinou odpověď hodnocení.

Spojení s databází je vytvořeno v souboru „dbconfig.php“, který je přiložen ve skriptu „results.php“ (popř. v „details.php“). Pokud by byla použita jiná databáze, je nutné upravit přihlašovací údaje v souboru „dbconfig.php“. Pokud jsou údaje zadané

uživatelé do databáze uloženy úspěšně, je o tom informován potvrzující hláškou, v opačném případě je vypsán důvod chyby. Údaje se ukládají následujícím sql příkazem:

```
"INSERT INTO entries (fname, lname, age, gender, model, results, problems) VALUES ('".$_POST["fname"]."', '".$_POST["lname"]."', '".$_POST["age"]."', '".$_GET["gender"]."', '".$_GET["model"]."', '".$_POST["results"]."', '".$_POST["problems"]."");"
```

Pravá polovina okna obsahuje interaktivní 3D prohlížeč s načteným získaným 3D modelem uživatele. Pro prohlížení modelů je využita knihovna Three.js [35] (verze three.js-r124). Three.js je knihovna, která využívá WebGL pro renderování 3D scén přímo v internetovém prohlížeči. Knihovna je podporována všemi dnešními moderními standardně používanými internetovými prohlížeči (Google Chrome 9+, Firefox 4+, Opera 15+, Safari 5.1+, Internet Explorer 11 a Microsoft Edge).

Nejprve jsou importovány všechny potřebné závislosti a připraveny proměnné. Zobrazování modelu se skládá ze dvou hlavních funkcí: `init` a `animate`.

Funkce `init` slouží k přípravě scény a načtení modelu. Nejprve je vytvořena kamera a nastavena její pozice ve scéně tak, aby byl viditelně zobrazen správně celý model a aby byl otočen směrem k pozorovateli. Dále jsou vytvořena základní světla (ambientní a bodové) a jsou určeny jejich parametry intenzit pro nejvhodnější osvětlení modelu. Pomocí `objLoader` a parametru odpovídající jménu souboru je načten získaný 3D model. Model je natočen do správného úhlu vzhledem ke kameře a je provedeno jeho vyhlazení pomocí funkce `setSmoothGeometry`. Dále je nastaven samotný renderer a je vytvořeno ovládání 3D scény. Je použito orbitové ovládání (`OrbitControls`) a nastavena minimální a maximální vzdálenost přiblížení. Nakonec je přidán posluchač, který reaguje na změnu velikosti okna prohlížeče. Pokud dojde ke změně, je automaticky změněna i velikost 3D scény a nastavení kamery. Výslednou 3D scénu tvoří pouze samostatný model. Model je zobrazen šedou barvou a pozadí černou kontrastní barvou.

Funkce `animate` slouží k samotnému renderování vybraného modelu. Funkce `requestAnimationFrame` vytvoří animační smyčku, která je pozastavena, pokud se uživatel přepne na jinou záložku prohlížeče. Dochází zde také k reakcím na ovládání scény uživatelem. Model je možné kolečkem myši přibližovat/oddalovat a myšička je také možné se volně pohybovat v zobrazené 3D scéně.

Testování internetové aplikace probíhalo prvně pouze lokálně prostřednictvím aplikace XAMPP [36]. XAMPP je volně dostupný softwarový balík. Používá se programátory jako testovací nástroj a je dostupný pro všechny základní platformy (Windows, Linux a macOS).

Zkratka XAMPP znamená

- X – cross–platform (multiplatformní),
- A – Apache,
- M – MariaDB,
- P – PHP a
- P – Perl.

Využité nástroje pro tvorbu a testování aplikace jsou Apache server, MariaDB databáze, phpMyAdmin pro správu databáze a PHP jako programovací jazyk. Kromě těchto dále obsahuje např. i OpenSSL, Joomla, Wordpress nebo mail server Mercury MTS a další nástroje.

XAMPP posloužil pro otestování samotného procesu odeslání fotografie, její zpracování a interaktivní prohlížení modelu. Také pomocí něho byla ověřena funkčnost odeslání výsledků do databáze.

Návod pro úspěšné spuštění a vyzkoušení internetové aplikace se nachází v kapitole Příloha.

Obrázek 10.1 (viz Příloha) zobrazuje úvodní stránku internetové aplikace a obrázek 10.2 (viz Příloha) finální stránku s výsledky.

Internetová aplikace byla kromě lokálního testování později otestována i v reálném prostředí se zapojením dalších lidí. Pro vytvoření veřejného spojení k internetové aplikaci posloužila aplikace „ngrok“ [37] v její bezplatné verzi. Aplikace ngrok dokáže vytvořit url adresu pro připojení k lokálnímu hostu. Lokálně je potřeba mít pouze zapnutý XAMPP a jeho příslušné potřebné moduly a pomocí vygenerované url adresy se pak již kdokoliv může připojit k vytvořené internetové aplikaci.

Získané výsledky skupiny uživatelů, které byl předán odkaz, představuje kapitola 8.1 Internetová aplikace.

Vygenerovaný odkaz byl také použit pro ověření funkčnosti aplikace v prohlížeči mobilního telefonu. Při pozici mobilního telefonu na šířku byla zjištěna plná funkčnost aplikace a správné zobrazení, použití mobilního telefonu na výšku je však při současné implementaci nevhodné. Řešením může být vytvoření samostatné mobilní verze, ve které by došlo ke změnám v rozložení jednotlivých částí aplikace.

7.2 Odhad výšky osoby

Algoritmus pro odhad výšky postavy je otestován v Python 3, konkrétně ve verzi Python 3.8. Plně funkční je též pro Python 2 (testována verze 2.7.18).

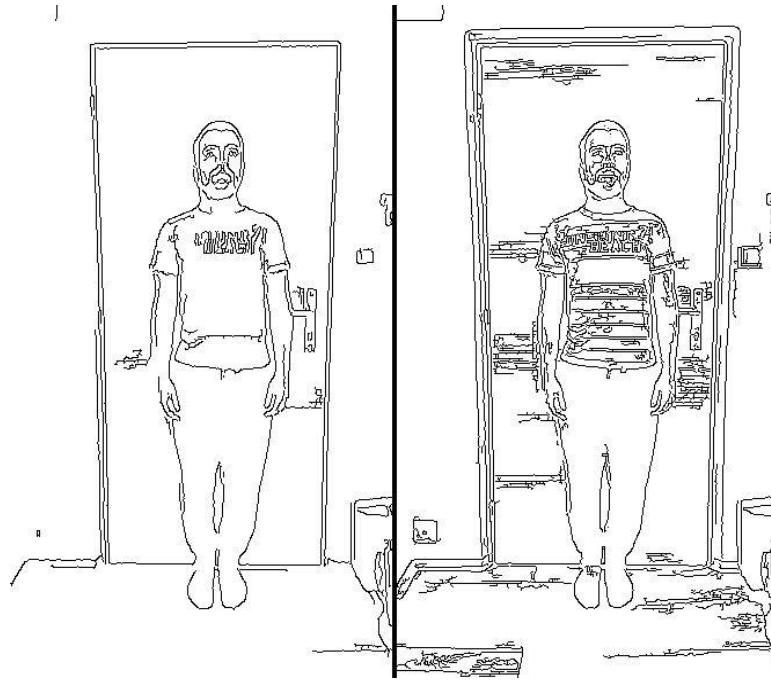
Potřebné importy metody pro odhad výšky jsou:

- OpenCV

Vstupní fotografie je nejprve převedena na černobílou fotografii pomocí OpenCV funkce „cvtColor“. Ta dokáže změnit barevný prostor fotografie. Pro správnou detekci hran v obraze je vhodné získaný černobílý obrázek rozmazat. K tomu slouží OpenCV funkce „blur“. Jejím parametrem je velikost okna, kterým se obrázek rozmazává. Rozmazání je provedeno průměrováním (tj. na středový bod okna je přiřazena průměrná hodnota okolí definovaného velikostí okna). Samotné rozmazání obrázku vede k mnohem přesnější detekci výrazných důležitých hran a vyloučení nedůležitých hran a šumu (viz obrázek 7.1 – vlevo výsledek s rozmazáním, vpravo bez rozmazání). Na získané hrany a následný odhad výšky má vliv také velikost použitého okna. Analýza různých velikostí oken se nachází v kapitole 8.2 Výška osoby. Okno (kernel) pro rozmazání okna je definováno v podobě (w, h) , kde w je šířka okna a h je výška okna. Okno velikosti $(3,3)$ tudíž zpracovává čtverec tři pixely široký a tři pixely vysoký a nová hodnota prostředního bodu okna je určena jako průměrná hodnota všech bodů uvnitř tohoto okna. Okno je postupně posouváno přes celou fotografii (konvoluce), dokud není fotografie celá zpracována.

Obrázek 7.2 zobrazuje výsledky jednotlivých kroků výše uvedeného postupu, zleva:

- vstupní fotografie,
- fotografie převedena na černobílou fotografii a
- rozmazaná fotografie.



Obrázek 7.1: Vliv rozmazání fotografie na detekci hran – vlevo s rozmazáním, vpravo bez rozmazání



Obrázek 7.2: Proces odhadu výšky – zleva: vstupní fotografie, černobílá fotografie, rozmazaná fotografie

Rozmazaný obrázek je dále převeden na obrázek hran. K tomu je použita OpenCV funkce „Canny“. Jejím parametrem je spodní a horní práh. Po testování na reálných fotografiích byla jako nejvhodnější hodnota zvolena 0 pro dolní práh a 100 pro horní práh. Canny algoritmus na vstupní obraz použije Gaussovský filtr a následně vypočte gradient. Pokud je hodnota gradientu vyšší než horní práh, je pixel vyhodnocen jako hrana. Pokud je hodnota gradientu nižší než dolní práh, je pixel zamítnut, a pokud je hodnota gradientu

mezi dolním a horním prahem, je pixel vyhodnocen jako hrana pouze tehdy, když je připojen k jinému pixelu již označenému jako hrana.

Hranový obrázek je nyní možné použít jako vstup OpenCV funkce „findContours“. Ta ve vstupním obrázku hledá kontury (souvislé křivky tvořené body se stejnou barevnou intenzitou). Kontury jsou užitečné pro tvarovou analýzu a detekci objektů. Získané pole s výsledky je již možné poslat jako argument do vlastní vytvořené funkce „find_door“.

Tato funkce ve vstupním poli prohlíží a zkoumá všechny objekty se čtyřmi hranami. Nejprve je OpenCV funkcí „approxPolyDP“ aproximován získaný tvar kontury na tvar s méně body. Jedná se o implementaci Douglas-Peucker algoritmu [38]. Dále jsou zkoumány objekty tvořené čtyřmi body (hledá se obdélník). Pomocí OpenCV funkce „boundingRect“ je vytvořen minimální ohraničující čtyřúhelník (nebere v úvahu rotaci) a jsou určeny souřadnice jeho levého horního rohu a jeho šířka a výška. Následně je vypočten jeho obsah a poměr stran. Pokud bychom chtěli brát v úvahu rotaci čtyřúhelníku, je možné použít OpenCV funkci „boxPoints“.

Obsah obecného čtyřúhelníku je vypočten jako

$$area = 2 \cdot w + 2 \cdot h, \quad (16)$$

kde w je šířka a h je výška daného čtyřúhelníku.

Poměr obsahů dvou čtyřúhelníků je vypočten jako

$$area_{ratio} = \frac{area_1}{area_2}, \quad (17)$$

kde $area_1$ je obsah největšího čtyřúhelníku a $area_2$ je obsah druhého největšího čtyřúhelníku.

Funkce ze vstupního obrázku vybírá podle obsahu vždy dva největší objekty. Pokud je poměr obsahů mezi hodnotami 0.95 a 1.05, vybírá se druhý největší čtyřúhelník (v tomto případě je největší objekt daného tvaru zárubeň dveří, druhý největší samotné dveře). Pokud je poměr obsahů jiný, anebo je nalezen pouze jeden objekt čtyřúhelníkového tvaru, je vybrán první největší čtyřúhelník.

Poměr stran čtyřúhelníku je vypočten jako

$$side_{ratio} = \frac{h}{w}, \quad (18)$$

kde w je šířka a h je výška daného čtyřúhelníku.

Následně je kontrolováno, zda má vybraný čtyřúhelník poměr stran mezi hodnotami 1.82 a 3.54 (určeno podle reálných fotografií dveří), to znamená, že se nejedná o čtverec ani o příliš úzký čtyřúhelník. Skutečný poměr stran pro dveře velikosti 197×80 cm je 2.46. Každá fotografie je však vyfocena při různých podmínkách, a tak se budou poměry stran dveří na fotografiích odlišovat. Z tohoto důvodu byly vypsány poměry stran dveří na testovacích fotografiích a pro originální ideální poměr stran 2.46 určen dolní limit 1.82 a horní limit 3.54. Tento limit zajistí, že budou správně detekovány stejné dveře na různých fotografiích, kde tyto dveře mohou vypadat užší i širší. Limit také umožňuje detekci dveří jiných rozměrů s jiným poměrem stran. Spolu s kontrolou obsahu nalezeného objektu (viz dále) by tak nemělo docházet k výběru jiného objektu, než právě požadovaných dveří (pokud se na fotografii nenachází objekt stejného tvaru a velikosti).

Dále je kontrolováno, zda má detekovaný objekt obsah větší než 1200 pixelů. Obsah je kontrolován proto, aby se nejednalo o objekt se sice správným poměrem stran, ale s nevhodnou velikostí neodpovídající dveřím (lze porovnávat s konkrétním číslem, protože je známa velikost fotografie, více viz 8.2 Výška osoby). Pokud některá z těchto podmínek neplatí, je vstupní černobílý obrázek rozmazán jiným typem kernelu a výše uvedený postup je opakován. Jiné rozmazání vede k získání mírně odlišných hran a následně k nalezení jiných objektů ve vstupním obrázku.

Sekvence použitých velikostí oken pro rozmazání je následující:

- (1,3), (2,4), (3,5), (3,6), (4,6), (5,7)

Pokud jsou dveře správně detekovány po rozmazání některým z těchto oken, další okna v sekvenci již nejsou použita a proces hledání dveří končí.

Na prvním místě sekvence jsou kromě okna velikosti (1,3) testovány také následující:

- (3,3)
- (3,4)
- (1,4)
- (2,3)

Použití různých oken na první pozici má za následek různé detekce a tím pádem vede k různým odhadnutým výškám, výsledky viz 8.2 Výška osoby.

Správnost detekce dveří byla ověřena vizuálně obarvením získaného objektu pomocí OpenCV funkce „drawContours“. Jejím vstupem je obrázek, ve kterém jsou obarveny vybrané pixely, dále aproximovaná křivka (výsledek funkce „approxPolyDP“) a zvolená barva a šířka linie pro obarvení. Stejně byla ověřena i pozice získaných vrcholů čtyřúhelníku pomocí OpenCV funkce „putText“, prostřednictvím které lze na vybranou pozici umístit do obrázku textové znaky.

Pokud dveře nejsou detekovány ani po několika různých rozmazání předem definovaných oken celé sekvence, tzn. ani v jednom případě nebyly zároveň splněny obě výše uvedené podmínky poměrů stran a obsahu, je výška osoby určena jako průměrná výška českých občanů [39]:

- muž – 181 cm,
- žena – 168 cm,
- neutrální model – 174.5 cm (průměr dvou výše uvedených výšek).

Pokud jsou dveře naopak správně detekovány, je možné přejít k samotnému odhadu výšky postavy. Je využito výsledků detektoru kloubů DeepCut, mezi jehož zjištěnými „klouby“ se nachází i pozice vrcholu hlavy. Algoritmus má nyní k dispozici znalost souřadnic vrchní hrany dveří a vrcholu hlavy, odhadnutá výška postavy je pak

$$h_{person} = h_{door} * \frac{y_1 - y_{head}}{(y_1 - y_2)}, \quad (19)$$

kde h_{door} je výška dveří v centimetrech, y_{head} je y souřadnice vrcholu hlavy zjištěná algoritmem DeepCut, y_2 je horní y souřadnice dveří a y_1 je dolní y souřadnice dveří. Rovnici lze alternativně zapsat

$$h_{person} = cm * (y_1 - y_{head}), \quad (20)$$

kde $cm = h_{door}/h$ ukazuje, kolik centimetrů ve skutečnosti má 1 pixel na fotografii, h_{door} je výška dveří v centimetrech, h je výška dveří v pixelech, y_{head} je y souřadnice vrcholu hlavy a y_1 je dolní y souřadnice dveří.

Obrázek 7.3 zobrazuje výsledky jednotlivých kroků výše uvedeného postupu, zleva:

- černobílá rozmazaná fotografie převedena na hranovou fotografii,
- nalezení pozice vrcholu hlavy detektorem DeepCut (červené zvýraznění) a
- nalezení pozice dveří (červené zvýraznění a modré souřadnice).



Obrázek 7.3: Proces odhadu výšky – zleva: hranová fotografie, detekce hlavy, detekce dveří

Využití detektoru DeepCut má výhodu v tom, že mezi vrchní hranou dveří a hlavou se může nacházet cokoli – různé vzory dřeva, prosklení dveří nebo např. různobarevné tapety nebo plakáty. Pokud by se od horní hrany dveří porovnávaly barevné hodnoty pixelů a hledala by se hlava, bylo by v případě nerovnoměrně barevných dveří mnohem obtížnější získání korektních výsledků. Postup porovnávání barev pixelů lze použít, pokud by se předpokládaly vždy pouze jednobarevné dveře (nebo co nejmenší barevné rozdíly a co nejméně šumu ve fotografii), v opačném případě je použití výsledků metody DeepCut robustnější.

Standardní šířka dveří je 60, 70, 80 a 90 cm. V algoritmu je vždy použita výška dveří 197 cm, přestože existují i jiné výšky (např. 210 cm). V době psaní diplomové práce nebyla nalezena statistika, která by ukazovala poměr mezi používáním dveří o výškách 197 cm a 210 cm, přesto je předpokládáno, že dveře výšky 210 cm jsou používány výrazně méně často. Pokud by se předem nadefinovaly konkrétní poměry stran pro rozměry určitých dveří, podle kterých by se poté vybíral předpokládaný typ dveří na fotografii, pravděpodobně by mnohem častěji docházelo k falešnému výběru dveří o výšce 210 cm. Poměr stran dveří o jedné výšce a šířce může odpovídat poměru dveří o výšce a šířce jiné. Je důležité upozornit, že ani jedny stejné dveře o standardní výšce 197 a šířce 80 cm (poměr stran 2.46) nebudou mít na různých fotografiích vždy úplně stejné poměry stran

(záleží na rotaci, výšce, odkud je foceno, atd.) a poměr stran ani nemusí odpovídat přesně číslu 2.46.

Předpokládá se použití fotografie vyfocené na výšku. Testovány byly fotografie v poměru stran 9 : 16 (šířka : výška). Osoba se postaví ideálně do středu dveří a je vhodné, aby do oblasti dveří nezasahovaly žádné jiné předměty (tj. aby zůstaly hrany viditelné co nejvíce).

V případě spuštění samotného algoritmu pro odhad výšky dojde pouze k výpisu odhadnuté výšky do konzole. Bez využití spouštěcího skriptu lze algoritmus spouštět přímo pomocí Python 2 i Python 3. Algoritmus je možné využít také ve spolupráci s algoritmem SMPLify, kde po vytvoření modelu postavy na fotografii a odhadu výšky dojde k přeškálování rozměrů modelu právě touto výškou. Návod na spuštění algoritmu pro odhad výšky a na spuštění algoritmu SMPLify včetně přeškálování vytvořeného modelu obsahuje Příloha.

7.3 Škálování modelu

Odhadnutá výška je využita pro naškálování 3D modelu získaného metodou SMPLify. Poměr odhadnuté výšky postavy a výšky 3D modelu je

$$h_{ratio} = \frac{h_{estimation}}{h_{original}}, \quad (21)$$

kde $h_{estimation}$ je odhadnutá výška vyfocené osoby a $h_{original}$ je výška modelu vytvořeného metodou SMPLify.

Tímto poměrem jsou pak vynásobeny souřadnice všech vrcholů modelu. Je získán model, který má výšku odpovídající odhadnuté výšce osoby a jehož rozměry by měly být blíže skutečným mírám osoby.

7.4 Měření modelu

Modely, které SMPLify zpracovává, zůstávají počtem vrcholů i po zpracování stále stejné, mění se jen jejich pozice. Proto je možné nadefinovat indexy vrcholů, přes které bude model měřen. Model je také vždy uveden do základní T-pózy, proto je možné rozměry modelů mezi sebou vzájemně porovnávat. Mužský, ženský i neutrální typ modelu mají všechny shodně 6890 vrcholů a liší se jen v jejich umístění v prostoru.

Euklidovská vzdálenost mezi dvěma sousedními vrcholy v 3D prostoru je definována jako

$$d_{i,j} = \sqrt{(x_j - x_i)^2 + (y_j - y_i)^2 + (z_j - z_i)^2}, \quad (22)$$

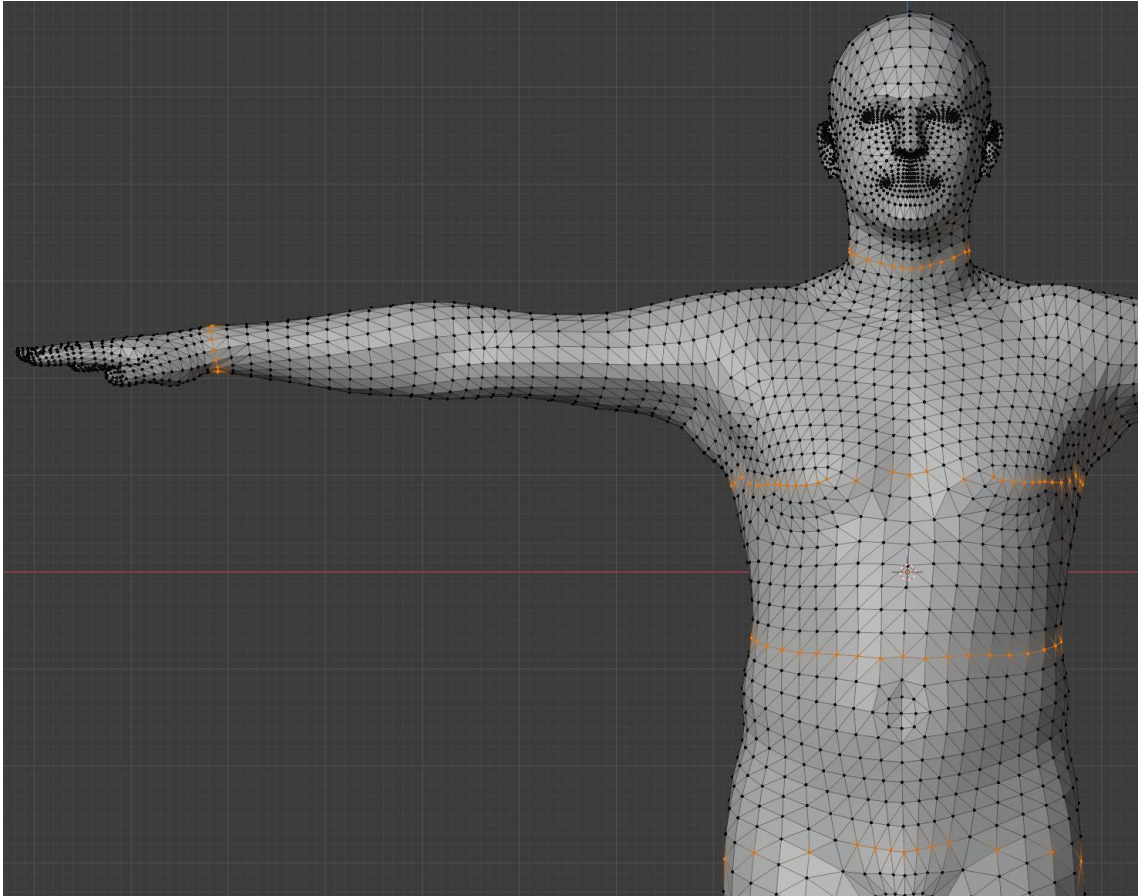
kde x, y, z jsou souřadnice vrcholů a i a j jsou indexy dvou sousedních vrcholů modelu.

Pokud je proveden součet vzdáleností mezi každými dvěma sousedními vrcholy z předem nadefinovaných polí, dostáváme obvod daných partií. Je důležité upozornit, že se jedná pouze o aproximaci. Modely neobsahují takový počet vrcholů, aby mezi nimi bylo možné vést vždy přesnou dokonalou trajektorii odpovídající reálnému měření na postavě. Žádné vrcholy do modelu pro zpřesnění trajektorie nejsou přidávány.

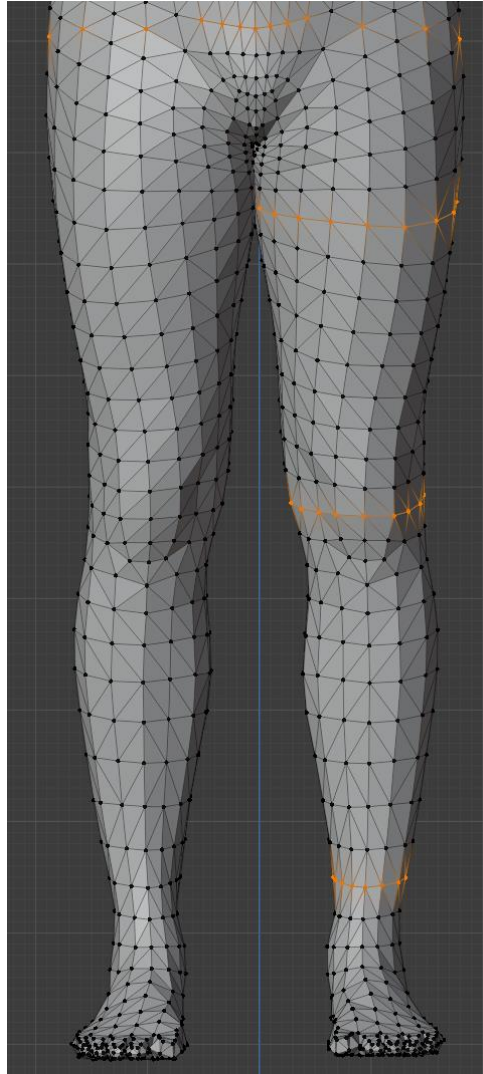
Trajektorie byly vybírány tak, aby co nejvěrněji odpovídaly reálnému měření postavy. Obrázek 7.4 zobrazuje vybrané trajektorie pro měření na vrchní části těla a obrázek 7.5 na spodní části těla. Vybraná trajektorie je obarvena oranžově. Pokud dva vrcholy v trajektorii nejsou spojeny, je i přesto počítána vzdálenost mezi těmito dvěma nejbližšími nespojenými vrcholy. Pro algoritmus počítání vzdáleností je nutné, aby vybrané body byly v poli uloženy ve správném pořadí, jak jdou za sebou.

Výška modelu a rozpětí paží jsou určeny jako vzdálenost mezi dvěma extrémními vrcholy modelu (minimální a maximální hodnota pro danou osu). Toto je možné, jelikož póza modelu je nastavena vždy stejně – vzpřímený postoj s maximálně roztaženými rukama ve tvaru písmene T.

Správnost výpočtů byla ověřena v programu Blender s nainstalovaným doplňkem MeasureIt [40]. MeasureIt je doplněk, který umožňuje výpočet a zobrazení vzdáleností ve „viewportu“ Blenderu. Výsledná měření tohoto doplňku jsou naprosto shodná s vzdálenostmi získanými vlastní vytvořenou funkcí. Měření modelů bylo ověřeno také tak, že výchozí neupravené modely téměř přesně odpovídají průměrným mírám mužů a žen ve Velké Británii [41] (to znamená, že např. nedostáváme nereálné poměry mezi rozměry jednotlivých partií).



Obrázek 7.4: Měření modelu – vrchní část těla



Obrázek 7.5: Měření modelu – spodní část těla

Návod na spuštění algoritmu SMPLify včetně měření vytvořeného modelu obsahuje Příloha. Míry jsou vypsány do konzole a také exportovány v podobě textového souboru do složky „SMPLify/smplify_public/results/measurement“. Vytvořené modely v neutrální póze jsou ukládány do složky „SMPLify/smplify_public/results/models“.

Výše uvedené platí i pro naškálované modely.

8 Dosažené výsledky

Algoritmus byl testován na následujícím PC:

- OS: Windows 10 + VirtualBox – OS Linux Mint Mate 20.1 (64bit) + nainstalované Guest Additions
- CPU: AMD Ryzen 5 2600 3.4 GHz (6C/12T) – pro virtuální systém přiřazeno 6 jader (pokud není uvedeno jinak)
- RAM: 16 GB DDR4 2800 MHz CL16 – pro virtuální systém přiřazeno 5 GB (pokud není uvedeno jinak)
- GPU: Nvidia GeForce GTX 1050Ti 4GB

Následující podkapitoly představují výsledky získané pomocí vytvořené internetové aplikace, dále získané výsledky funkce pro odhad výšky postavy a nakonec získané rozměry modelů vytvořených podle fotografií s osobami s různým oblečením a s různými pózami těl.

8.1 Internetová aplikace

Přístup k vytvořené internetové aplikaci dostala skupina studentů Fakulty aplikovaných věd Západočeské univerzity v Plzni. Během určeného časového úseku s přístupem k aplikaci bylo k zpracování celkem zasláno 52 fotografií, z kterých bylo vytvořeno 44 modelů. Po zkontrolování počtů byly zasláné fotografie z důvodu anonymity smazány. Studenti dostali instrukce o zaslání fotografie celé postavy ve stoje. Dále byli informováni, že nemusí zadávat své reálné osobní údaje.

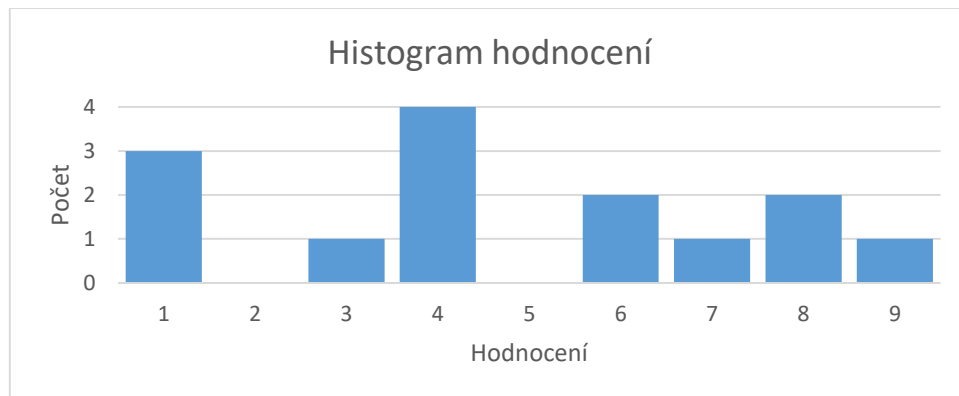
Ze 44 modelů bylo uživateli odesláno hodnocení pouze pro 14 z nich. Pro jeden z vyhodnocených modelů bylo vybráno neutrální pohlaví. Určit pohlaví modelů, pro která nebyla odeslána hodnocení, nelze jednoznačně určit, jelikož nebyla uložena.

Získaná hodnocení podobnosti modelů jsou následující:

- 3/14 – Tvar, velikost a póza modelu vypadá přesně jako já.
- 0/14 – Velikost a póza modelu vypadá jako já, ale tvar je odlišný.
- 1/14 – Tvar a póza modelu vypadá jako já, ale velikost je odlišná.
- 4/14 – Tvar a velikost modelu vypadá jako já, ale póza je odlišná.
- 0/14 – Tvar modelu vypadá jako já, ale velikost a póza je odlišná.

- 2/14 – Velikost modelu vypadá jako já, ale tvar a póza je odlišná.
- 1/14 – Póza modelu vypadá jako já, ale tvar a velikost je odlišná.
- 2/14 – Nevím.
- 1/14 – Model vůbec nevypadá jako já.

Získaná hodnocení zobrazuje též obrázek 8.1.



Obrázek 8.1: Hodnocení získaná pomocí internetové aplikace

Nejčastější odpovědí byla spokojenost s tvarem a velikostí modelu a nespokojenost s jeho pózou (čtyři zvolení). Další častou odpovědí je spokojenost se všemi kategoriemi (tři zvolení). Dvě zvolení získala volba spokojenosti s velikostí modelu a nespokojenosti s tvarem a pózou a také odpověď „nevím“.

Mezi hodnoceními se objevily i tři textové komentáře. Parafrázováno:

- Pro fotografii s pouze vrchní částí těla nebyl model vytvořen.
- Mírně odlišná pozice rukou (na fotografii držení za boky, ale model měl ruce před nimi).
- Na fotografii ruce zaťaté v pěst, ale model měl ruce (prsty) natažené.

První komentář lze vyřešit zasláním fotografie celého těla a třetí komentář by vyřešilo použití metody SMPLify-X s modelem SMPL-X.

U modelů, kde byla vybrána nespokojenost s pózou modelu, jsem i já osobně vizuálně zkontroloval vytvořený model. Dva z těchto modelů odpovídají druhému a třetímu výše uvedenému textovému komentáři – toto byly pravděpodobně jediné neshody s pózou vytvořených modelů. Třetí model má mírně pokrčené obě nohy (pravou více), pravou ruku také mírně ohnutou, levou ruku nataženou od těla a hlavu směřující dopředu před tělo. Poslední čtvrtý výběr nespokojenosti s pózou je pro model, který je v hlubokém

předklonu s rukama nataženými dozadu za tělo a s doširoka roztaženými nohama. Všechny tyto modely byly mužského pohlaví.

Zkontrolován byl také model, pro který bylo vybráno nejvíce nespokojené hodnocení „model vůbec nevypadá jako já“. Pro model bylo vybráno ženské pohlaví. Vytvořený model má jednu nohu nataženou vysoko před sebe do výšky (úroveň hlavy), stojnou nohu mírně pokrčenou, celý model je zakloněn dozadu a ruce má nepřirozeně ohnuté (opačný ohyb v lokti). Ruce navíc v místě loktů mírně pronikají do těla.

Jedno hodnocení „nevím“ bylo odesláno s výše uvedeným komentářem, že se nezobrazují modely s použitím fotografie osoby s pouze vrchní částí těla. Druhé hodnocení „nevím“ bylo použito pro ženský model s široce roztaženými nohama, jednu rukou ohnutou za zády a druhou rukou nataženou od těla.

Z hodnocení studentů lze konstatovat, že pro algoritmus SMPLify je asi nejsložitější zachycení dokonale odpovídající pózy. Je však také nutné upozornit, že některá nespokojená hodnocení byla zvolena pro modely s neobvyklými tělesnými pózami.

Ze získaných výsledků lze také pozorovat, že uživatelé jsou ochotní si internetovou aplikaci a vytvoření modelu vyzkoušet, ale jsou již méně ochotní porovnat model s osobou na fotografii a odeslat výsledky. Možným důvodem může být vyplňování osobních údajů, i když byli v tomto případě uživatelé upozorněni, že pokud nechtějí, nemusejí zadávat své reálné osobní údaje.

Pozitivní zprávou je, že uživatelé nehlásili žádné problémy se samostatným během aplikace a zobrazováním modelů.

8.2 Výška osoby

Nejprve bylo vytvořeno 16 testovacích fotografií osoby u dveří. Tyto fotografie budou dále v textu označovány jako „Skupina B“. Na fotografie byl použit vytvořený algoritmus (viz 7.2 Odhad výšky osoby) a byly analyzovány výsledné odhadnuté výšky. Podle správnosti odhadnuté výšky osoby na jednotlivých fotografiích byly detekovány vhodné a méně vhodné typy fotografií.

Skutečná výška testovací osoby je 168 cm.

Obrázek 8.2 zobrazuje ukázky nevhodně pořízených fotografií pro správný odhad výšky postavy. Odhadnutá výška pro první fotografii zleva je 162,09 cm (-7,91 cm od reálné

Dosažené výsledky

výšky), tzn. postava se jeví menší. Důvodem je vyfocení fotografie směrem „shora dolů“. Opačný případ „zdola nahoru“ by byl též nevhodný (postava by se jevila vyšší). Dveře na druhé fotografii mají špatně kontrastní pravou hranu, a tudíž nebylo možné je celé správně detekovat. Třetí fotografie je případ příliš tmavé a málo kontrastní fotografie. Jelikož nejsou dveře na druhé a třetí fotografii zleva správně detekovány, není možné pokračovat v procesu odhadu výšky z fotografie a výška je určena jako průměrná výška českých občanů [39]:

- muž – 181 cm,
- žena – 168 cm,
- neutrální model – 174.5 cm (průměr dvou výše uvedených výšek).



Obrázek 8.2: Ukázka nevhodných fotografií – zleva: fotografie pořízená „shora dolů“, špatně kontrastní pravá hrana dveří a příliš tmavá a málo kontrastní fotografie

Obrázek 8.3 zobrazuje ukázkou vhodně pořízené fotografie pro správný odhad výšky postavy. Ideální typ fotografie je pořízený zhruba z poloviny výšky dveří a s dobrým osvětlením (tzn. ani příliš tmavá ani příliš přesvětlená fotografie). Zároveň je vhodné, aby do hran dveří nezasahovaly žádné předměty. Odhadnutá výška pro tuto fotografii je 167,77 (-0,23 cm od reálné výšky).



Obrázek 8.3: Ukázka vhodné fotografie

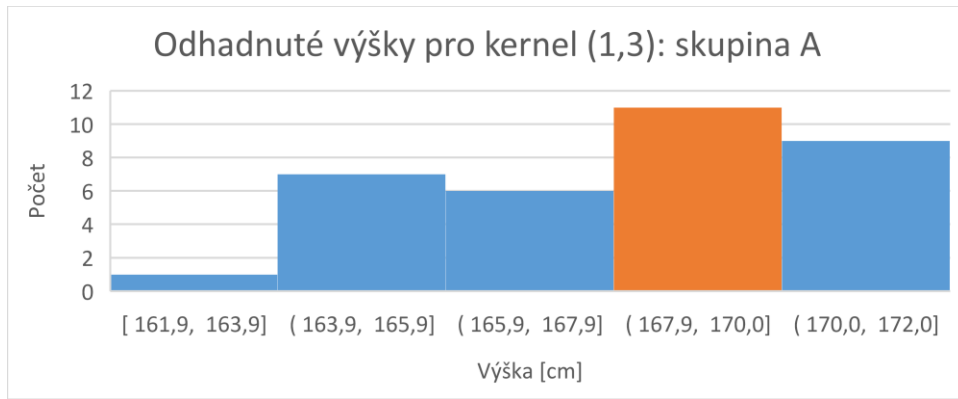
Na základě znalosti vhodných fotografií bylo vytvořeno dalších 34 fotografií pro odhad výšky s cílem co nejpřesnějšího odhadu. Tyto fotografie jsou dále v textu označovány jako „Skupina A“.

Obrázky 8.4-8.8 zobrazují histogramy odhadnutých výšek pro jednotlivé velikosti oken pro rozmazání fotografie. Zvýrazněný sloupec histogramu zobrazuje pozici reálné výšky osoby.

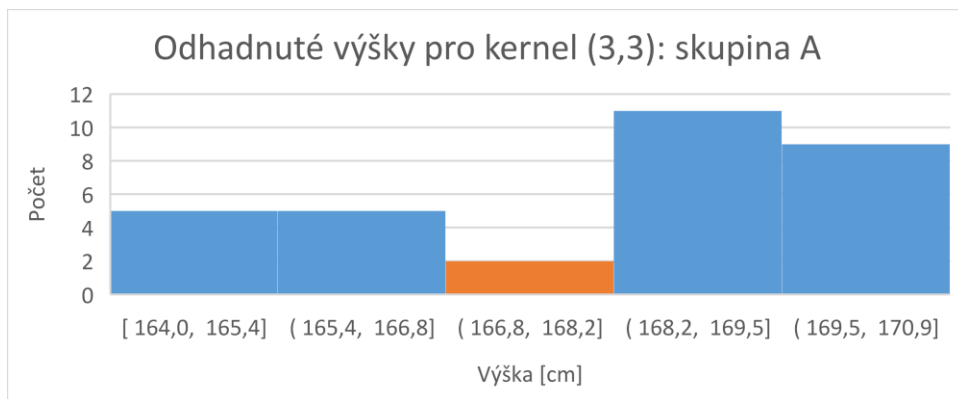
Histogram je typ grafu, který znázorňuje data v podobě stejně širokých sloupců. Šířka sloupce vyjadřuje šířku intervalů (tříd), kam data spadají, a výška sloupce vyjadřuje počet dat v daném intervalu.

V ideálním případě chceme, aby histogram tvarem odpovídal grafu hustoty normálního rozdělení pravděpodobnosti (tj. přibližně tvar „kopce“). Nejvyšší sloupec by měl ve svém intervalu obsahovat skutečnou naměřenou hodnotu (tj. být oranžově obarvený) a směrem doleva a doprava od tohoto sloupce by měly výšky sloupců klesat (tj. získáváme méně nepřesných údajů). Toto platí pro všechny dále uvedené histogramy.

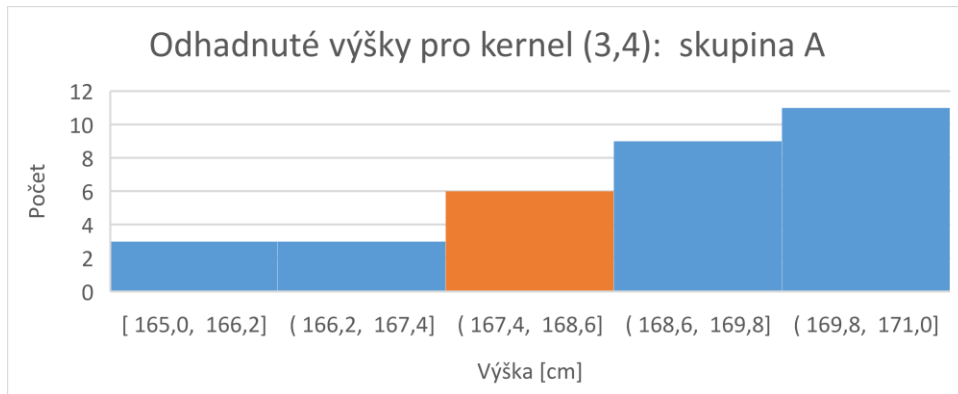
Uvedené histogramy odpovídají výškám odhadnutých na fotografiích skupiny A.



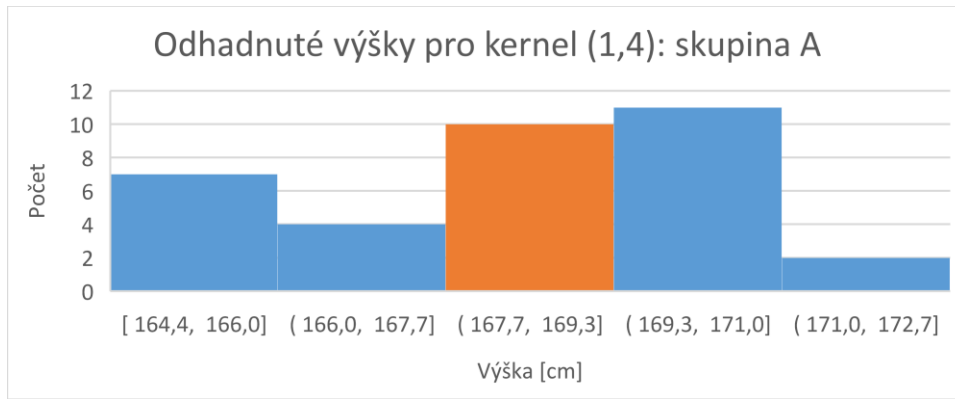
Obrázek 8.4: Histogram výšek pro kernel (1,3)



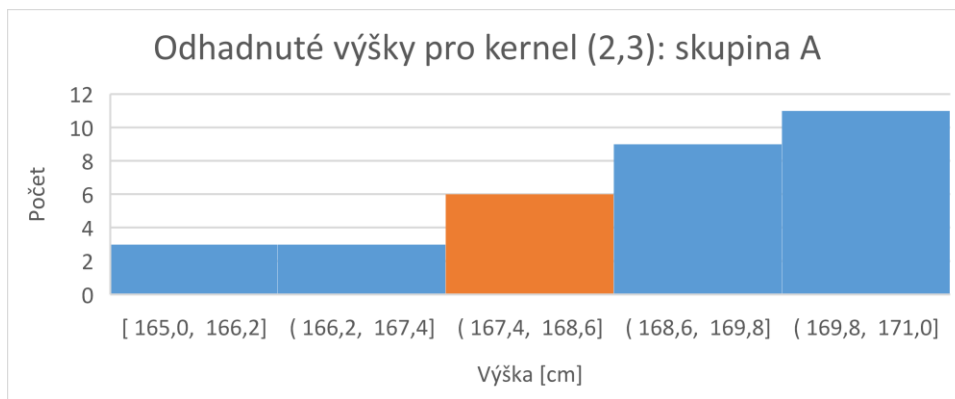
Obrázek 8.5: Histogram výšek pro kernel (3,3)



Obrázek 8.6: Histogram výšek pro kernel (3,4)



Obrázek 8.7: Histogram výšek pro kernel (1,4)



Obrázek 8.8: Histogram výšek pro kernel (2,3)

Z uvedených histogramů lze pozorovat, že nejlepších výsledků bylo dosaženo pro okno pro rozmazání fotografie velikosti (1,3). V intervalu obsahující správnou výšku se nachází největší počet odhadnutých výšek (celkem jedenáct). Interval zvýrazněného sloupce je také nejúžší z uvedených histogramů (1.1 cm). V obou směrech od nejvyššího sloupce klesají počty méně přesných odhadnutých výšek. Nejvíce nepřesné odhadnuté výšky bylo pro okno této velikosti dosaženo pouze v jednom případě.

Tabulka 8.1 zobrazuje srovnání výsledků odhadu výšky pro různé typy velikostí oken (kernelů) pro rozmazání fotografií.

Díky nejlepším získaným výsledkům odhadu výšky pro fotografie skupiny A byla velikost okna (1,3) zvolena jako finální. Výsledky byly určeny jako nejlepší na základě vyhodnocení výše uvedených histogramů a na základě číselných výsledků uvedených v tabulce 8.1. V té můžeme vidět, že pro okno velikosti (1,3) bylo pro fotografie skupiny A dosaženo nejlepší průměrné výšky a rovněž rozdíl střední hodnoty všech odhadnutých výšek a skutečné výšky je nejmenší. Nejlepší (nejnižší) hodnoty bylo dosaženo také u testovacího kritéria T-testu.

Dosažené výsledky

Dle T-testu je zvolené okna velikosti (1,3) nejvhodnější i pro fotografie skupiny B, ostatní kritéria už o tom ale jednoznačně nevyovídají. Průměrnou hodnotu a střední hodnotu lze nalézt bližší skutečné výšce u jiných typů oken.

Kernel	(3,3)	(3,4)	(1,4)	(1,3)	(2,3)
Šířka fotografie	330	330	330	330	330
Skupina A:					
Minimum	164,04	164,99	164,36	161,87	164,99
Maximum	170,91	171,00	172,65	171,98	171,00
Střední hodnota	168,81	168,94	168,72	168,44	168,94
Průměr	168,17	168,75	168,41	167,94	168,75
Směrodatná odchylka	1,98	1,60	2,07	2,27	1,60
Výběrová směrodatná odchylka	2,01	1,62	2,10	2,30	1,62
Rozdíl stř. h. a naměřené h.	0,81	0,94	0,72	0,44	0,94
T-test	0,49	2,63	1,15	-0,16	2,63
	Nezamítám H ₀	Zamítám H ₀	Nezamítám H ₀	Nezamítám H ₀	Zamítám H ₀
Skupina B:					
Minimum	151,98	156,81	151,98	151,98	156,81
Maximum	168,80	167,74	169,84	170,81	167,74
Střední hodnota	162,26	161,45	161,92	161,77	161,45
Průměr	162,73	162,10	161,84	162,25	162,10
Směrodatná odchylka	4,39	2,74	4,21	4,92	2,74
Výběrová směrodatná odchylka	4,55	2,84	4,35	5,10	2,84
Rozdíl stř. h. a naměřené h.	-5,74	-6,55	-6,09	-6,23	-6,55
T-test	-7,85	-14,23	-10,01	-7,80	-14,23
	Zamítám H ₀	Zamítám H ₀	Zamítám H ₀	Zamítám H ₀	Zamítám H ₀

Tabulka 8.1: Srovnání výsledků odhadu výšky pro různé typy kernelů

Nevýhodou vybraného okna velikosti (1,3) je jeho nevyšší směrodatná odchylka. To znamená, že může docházet k rozdílům mezi jednotlivými odhadnutými výškami. U fotografií skupiny A lze toto pozorovat u minimální odhadnuté výšky, která je u tohoto okna menší oproti jiným oknům. Dle uvedeného histogramu ale můžeme vidět, že k odhadu této nepřesné výšky došlo pouze jednou.

Tabulka 8.2 zobrazuje samostatné shrnutí statistických výsledků pro fotografie skupiny A (pouze fotografie vytvořené se znalostí, jak zachytit vhodnou fotografii) s finálním použitým kernelem pro rozmazání velikosti (1,3) a rozlišením fotografie 330 × 586 a také pro fotografie skupiny B (fotografie vytvořené bez znalosti, jak zachytit vhodnou fotografii) s finálním použitým kernelem pro rozmazání velikosti (1,3) a rozlišením fotografie 330 × 586. Použití daného rozlišení je vysvětleno dále.

Skupina	A	B
Minimum	161,87	151,98
Maximum	171,98	170,81
Střední hodnota	168,44	161,77
Průměr	167,94	162,25
Směrodatná odchylka	2,27	4,92
Výběrová směrodatná odchylka	2,30	5,10
Rozdíl stř. h. a naměřené h.	0,44	-6,23
T-test	-0,16	-7,80
	Nezamítám H0	Zamítám H0

Tabulka 8.2: Výsledky odhadu výšky pro fotografie skupiny A a B

Důležité je také upozornit, že testované typy kernelů nevedly vždy k 100% detekci dveří. Úspěšnost detekcí dveří pro testované typy kernelů při rozlišení fotografie 330 × 586 je

- (1,3) – 48/50 úspěšných detekcí,
- (3,3) – 47/50 úspěšných detekcí,
- (3,4) – 48/50 úspěšných detekcí,
- (1,4) – 50/50 úspěšných detekcí a
- (2,3) – 48/50 úspěšných detekcí.

Pokud dveře nejsou detekovány, je použita průměrná výška českých občanů [39].

Vizuální kontrolou odhadu pozice hlavy metodou DeepCut bylo zjištěno, že získaná pozice (souřadnice „y_head“) neodpovídá úplně dokonale jejímu vrcholu. Skutečný nejvyšší bod osoby na fotografii se nacházel přibližně o tři až čtyři pixely výše, než byla určená pozice. Proto byl algoritmus otestován i s nastavením pozice právě o tři a čtyři body výše vůči pozici získané metodou DeepCut. Tabulka 8.3 zobrazuje získané výsledky pro zvolený kernel velikosti (1,3).

Značení v tabulce 8.3 odpovídá těmto pozicím:

- y_head: y pozice vrcholu hlavy získaná metodou DeepCut,
- y_head - 3: y pozice vrcholu hlavy o tři pixely výše vůči souřadnici „y_head“ a
- y_head - 4: y pozice vrcholu hlavy o čtyři pixely výše vůči souřadnici „y_head“.

Kernel	(1,3)		
Pozice hlavy	y_head - 4	y_head - 3	y_head
Šířka fotografie	330	330	330
Skupina A:			
Minimum	163,54	163,12	161,87
Maximum	173,80	173,34	171,98
Střední hodnota	170,00	169,61	168,44
Průměr	169,50	169,11	167,94
Směrodatná odchylka	2,28	2,28	2,27
Výběrová směrodatná odchylka	2,31	2,31	2,30
Rozdíl stř. h. a naměřené h.	2,00	1,61	0,44
T-test	3,78	2,80	-0,16
	Zamítám H0	Zamítám H0	Nezamítám H0
Skupina B:			
Minimum	153,52	153,14	151,98
Maximum	172,38	171,98	170,81
Střední hodnota	163,44	163,02	161,77
Průměr	163,87	163,47	162,25
Směrodatná odchylka	4,94	4,93	4,92
Výběrová směrodatná odchylka	5,13	5,12	5,10
Rozdíl stř. h. a naměřené h.	-4,57	-4,98	-6,23
T-test	-5,58	-6,14	-7,80
	Zamítám H0	Zamítám H0	Zamítám H0

Tabulka 8.3: Vliv pozice hlavy na odhadnutou výšku

Zvýšení y-ové pozice hlavy vedlo ke zvýšení odhadované výšky postavy. To však pro fotografie skupiny A nepřineslo výsledky více odpovídající reálné výšce postavy. Pro fotografie skupiny B obsahující i fotografie focené bez znalosti, jak vhodně fotografii vyfotit, zvýšení y-ové pozice hlavy pomohlo k získání lepších výsledků. Tento set obsahuje i fotografie focené „shora dolů“, tudíž se postava jeví menší než ve skutečnosti, a zvýšení y-ové pozice hlavy tento problém koriguje.

Z důvodu, že jsou v diplomové práci uvedeny instrukce, jak vyfotit „vhodnou“ fotografii pro nejpřesnější odhad výšky, byla pro y pozici hlavy (tj. „y_head“) ponechána souřadnice získaná metodou DeepCut bez dalších úprav.

Během testování byl ověřen také vliv rozlišení vstupní fotografie na dobu potřebnou pro odhad výšky. Uvedený čas je v sekundách. Testování proběhlo s dvěma různými přiřazeními paměti RAM virtuálnímu operačnímu systému – 5 a 12 GB RAM. Uvedená

spotřeba paměti RAM byla určena dle „Správce úloh“ a jedná se o spotřebu maximální dosaženou během běhu procesu.

Systém	5 GB RAM			12 GB RAM		
Šířka [px]	Výška [px]	Čas [s]	RAM [GB]	Čas [s]	RAM [GB]	
330	586	2,18	1,9	2,19	1,9	
450	800	19,53	4,0	3,75	4,4	
562	1000		> 5 GB RAM	5,67	6,5	

Tabulka 8.4: Vliv rozlišení na dobu běhu a spotřebu paměti RAM

Metoda DeepCut obsahuje parametr pro nastavení, jak velké fotografie metoda zpracovává bez dělení fotografie na části. Tabulka 8.4 zobrazuje průměry ze tří časů celkové běhu programu pro odhad výšky pro vstupní fotografii o určitém rozlišení s nastavením takovým, že nedochází k dělení fotografie. Při dělení fotografie na části by mělo dojít k úspoře požadované paměti RAM. Po nastavení parametru na velikost menší, než byla velikost testovací vstupní fotografie, však docházelo během vykonávání programu k chybám a nepodařilo se získat výsledky. Konkrétně se jednalo o tuto chybu: „TypeError: slice indices must be integers or None or have an `__index__` method“. Zmíněnou chybu se podařilo opravit přetypováním proměnné zmíněné v popisu chyby na proměnnou typu „integer“. To vedlo k odstranění reportování chyby za běhu a program úspěšně došel do konce, postava však nebyla správně detekována. Všechny detekované klouby byly určeny nepravidelně v malé části fotografie nad prostorem dveří. Výše uvedená oprava tedy pravděpodobně nebyla korektní. Během tvorby diplomové práce nebylo nalezeno žádné jiné řešení daného problému.

Pokud byl parametr nastaven tak, aby fotografie zůstala v jednom kuse, potřeba paměti RAM pro zpracování fotografie s nejvyšším testovaným rozlišením (šířka × výška = 562 × 1000) již přesahovala 5 GB a proces byl operačním systémem automaticky ukončen. Zvýšení šířky fotografie z 330 pixelů na 450 pixelů vedlo u systému s nižším přiřazením paměti RAM k přibližně devítinásobnému zvýšení času běhu. Řešením bylo virtuálnímu systému přidat více paměti. V takovém případě proces úspěšně došel i pro největší testovanou fotografii, maximální spotřeba paměti byla okolo 6.5 GB RAM a zároveň došlo k poměrně rychlému zpracování. Výrazné urychlení běhu algoritmu bylo měřeno i u fotografie velikosti 450 × 800.

Zajímavostí je, že při nejvyšším rozlišení 562 × 1000 algoritmus DeepCut ve třech testovaných případech z pěti špatně detekoval vrchol hlavy. Všechny ostatní klouby byly

nalezeny správně. Pozice vrcholu hlavy byla určena přibližně v místě rtů, což vedlo k výrazně menší odhadnuté výšce. Mezi kompletními přiloženými výsledky lze i u nižších rozlišení fotografií skupiny B možné nalézt několik odhadnutých výšek v hodnotách mezi přibližně 151 a 153 cm. Je možné předpokládat, že i v těchto případech došlo k stejné špatné detekci hlavy. Z důvodu velkého množství fotografií, různých typů oken pro rozmazání fotografií, různě posouvané souřadnice y hlavy (tj. „ y_{head} “) a různých rozlišení fotografií však nebyly všechny fotografie vizuálně kontrolovány, nelze tedy s naprostou jistotou konstatovat, zda se v těchto případech jednalo o špatnou detekci hlavy, anebo zda šlo o detekci zárubně dveří.

Dále je důležité upozornit, že největší část běhu metody pro odhad výšky zabírá získání výsledků metody DeepCut, odhad výšky je následně již okamžitý. Jelikož fotografie s rozlišením 330 : 586 vede s vybraným typem okna pro rozmazání fotografií k nejpřesnějším výsledkům a jejich rychlému získání z testovacích fotografií i u systému s menší pamětí RAM, bylo toto rozlišení zvoleno jako finální. Metoda DeepCut je nastavena tak, aby fotografie tohoto rozlišení zpracovávala vcelku bez dělení.

Testovaná rozlišení byla určena poměrem z originální fotografie o velikosti 2248 × 4000:

- 15 % z této fotografie: 337 × 600, dále upraveno na 330 × 586 (cca 14.66 %)
- 20 % z této fotografie: 450 × 800
- 25 % z této fotografie: 562 × 1000

Tabulka 8.5 zobrazuje vliv rozlišení fotografie na odhadnutou výšku osoby pro vybrané velikosti oken (kernely) pro rozmazání fotografie. Bylo zjištěno, že zvýšení rozlišení nevede ve všech případech k lepšímu odhadu výšky. Někdy jsou odhadnuté výšky blíže skutečnosti u fotografií s nižším rozlišením a někdy s vyšším rozlišením. Nejlepší výsledky byly získány u fotografií s nižším rozlišením.

Kernel	(3,3)		(3,4)		(1,4)	
Šířka fotografie	330	450	330	450	330	450
Skupina A:						
Minimum	164,04	164,73	164,99	164,86	164,36	161,83
Maximum	170,91	172,78	171,00	171,54	172,65	174,59
Střední hodnota	168,81	169,20	168,94	169,19	168,72	169,65
Průměr	168,17	168,86	168,75	168,62	168,41	169,20
Směrodatná odchylka	1,98	2,36	1,60	2,02	2,07	2,66
Výběrová směrodatná odchylka	2,01	2,39	1,62	2,05	2,10	2,70
Rozdíl stř. h. a naměřené h.	0,81	1,19	0,94	1,19	0,72	1,65
T-test	0,49	2,10	2,63	1,76	1,15	2,60
	Nezamítám H0	Zamítám H0	Zamítám H0	Zamítám H0	Nezamítám H0	Zamítám H0
Skupina B:						
Minimum	151,98	159,88	156,81	160,29	151,98	159,53
Maximum	168,80	168,02	167,74	168,18	169,84	170,41
Střední hodnota	162,26	160,90	161,45	161,40	161,92	161,95
Průměr	162,73	162,13	162,10	162,72	161,84	163,51
Směrodatná odchylka	4,39	2,87	2,74	2,82	4,21	3,47
Výběrová směrodatná odchylka	4,55	2,97	2,84	2,93	4,35	3,59
Rozdíl stř. h. a naměřené h.	-5,74	-7,10	-6,55	-6,60	-6,09	-6,06
T-test	-7,85	-13,84	-14,23	-12,48	-10,01	-8,85
	Zamítám H0	Zamítám H0	Zamítám H0	Zamítám H0	Zamítám H0	Zamítám H0

Tabulka 8.5: Vliv rozlišení na odhadnutou výšku

Kompletní výsledky metody pro odhad výšky postavy včetně odhadnutých výšek pro jednotlivé fotografie a histogramů odhadnutých výšek pro fotografie skupiny A a B lze nalézt v příloženém souboru „Measurements.xlsx“ v záložce „Výška“.

8.3 Rozměry modelu

Následující podkapitoly zobrazují výsledky měření modelů pro jednotlivé testovací osoby.

8.3.1 Žena

Tabulka 8.6 zobrazuje výchozí rozměry použitého ženského modelu a naměřené míry dané osoby. Osoba byla ochotná poskytnout pro účely diplomové práce všechny potřebné fotografie kromě fotografie v oblečení prvního typu. Žena je hubené sportovní postavy.

	Výška	Rozpětí paží	Hrudník	Pas	Boky	Zápěstí	Stehno	Koleno	Kotník	Krk
Výchozí hodnoty modelu:	166,00	167,00	98,01	83,12	103,66	15,93	59,74	37,73	21,24	33,24
Naměřené hodnoty osoby:	172,00	175,00	86,00	65,00	82,00	14,00	51,00	34,00	26,00	32,00

Tabulka 8.6: Vstupní rozměry – žena

Výsledky testování hypotéz s ženským modelem jsou následující:

- různé oblečení + různé pózy

- bez škálování – 1/10 nezamítnutí H_0
- škálování výškou – 1/9 nezamítnutí H_0
- škálování rozpětím paží – 0/9 nezamítnutí H_0
- různé oblečení
 - bez škálování – 0/10 nezamítnutí H_0
 - škálování výškou – 1/9 nezamítnutí H_0
 - škálování rozpětím paží – 0/9 nezamítnutí H_0
- různé pózy
 - bez škálování – 3/10 nezamítnutí H_0
 - škálování výškou – 1/9 nezamítnutí H_0
 - škálování rozpětím paží – 1/9 nezamítnutí H_0

Bez škálování je celkových hypotéz deset (osm měřených tělesných partií, výška a rozpětí paží). Při škálování, ať už výškou nebo rozpětím pažím, je díky přeškálování tento parametr rovný přesně údají testované osoby, a proto pro tento parametr není testována hypotéza H_0 – z tohoto důvodu je při škálování pouze devět celkových hypotéz. Toto platí i pro další testovací osoby uvedené v dalších podkapitolách (8.3.2 a 8.3.3). Uvedené platí obdobně i u dále uvedených intervalů spolehlivosti.

Celkem bylo nezamítnuto 8/84 nulových hypotéz H_0 . Pokud bychom chtěli tvrdit, že metoda za všech podmínek vytváří modely odpovídající skutečným postavám, bylo by potřeba, aby nezamítnutých hypotéz byla většina. V tomto případě nám malé množství nezamítnutých hypotéz říká, že vytvořené modely příliš neodpovídají postavě testované osoby.

Modely bez škálování byly vyšší zhruba o 10 až 20 cm. Extrémním případem je póza dřepu. Z této fotografie vytvořila metoda model vysoký 2.6 m, čemuž odpovídaly i další rozměry tohoto modelu.

Tabulka 8.7 zobrazuje konkrétní statistické výsledky pro modely vytvořené dle fotografií s ženskou osobou s různým oblečením a různými pózami a škálováním modelů výškou.

Dosažené výsledky

	Rozpětí paží	Hrudník	Pas	Boky	Zápěstí	Stehno	Koleno	Kotník	Krk
Naměřené hodnoty osoby:	175,00	86,00	65,00	82,00	14,00	51,00	34,00	26,00	32,00
Přepočtené hodnoty poměrem výšky:									
Minimum	169,08	85,63	69,69	96,60	15,12	55,56	36,23	20,32	30,27
Maximum	179,28	99,55	85,17	105,90	16,35	60,79	38,51	21,76	33,87
Střední hodnota	172,00	95,04	80,60	103,05	15,81	58,92	37,76	21,18	32,33
Průměr	172,23	93,98	79,61	102,72	15,81	58,76	37,65	21,18	32,32
Směrodatná odchylka	2,37	3,58	4,22	2,69	0,33	1,41	0,61	0,35	0,98
Výběrová směrodatná odchylka	2,45	3,71	4,37	2,78	0,34	1,46	0,63	0,37	1,01
Rozdíl stř. h. a naměřené h.	3,00	9,04	15,60	21,05	1,81	7,92	3,76	4,82	0,33
T-test	-4,37	8,33	12,95	28,82	20,55	20,52	22,42	-51,07	1,21
	Zamítám H0	Zamítám H0	Zamítám H0	Zamítám H0	Zamítám H0	Zamítám H0	Zamítám H0	Zamítám H0	Nezamítám H0

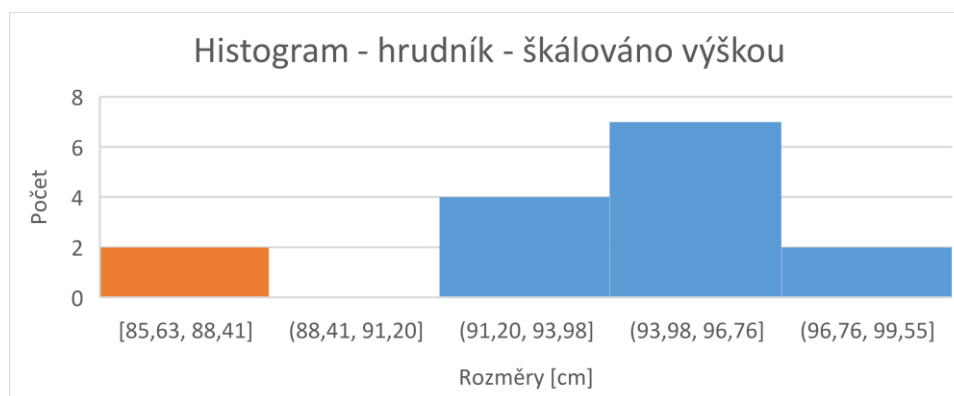
Tabulka 8.7: Výsledné rozměry modelů – žena

Výsledné rozměry modelů jsou větší oproti reálným mírám, hlavně v oblasti pasu a boků. Je nutné upozornit, že tato osoba neznala přesné trajektorie měření na modelu a byly jí zaslány pouze textové instrukce pro měření tělesných měř. To mohlo způsobit drobné ovlivnění výsledků.

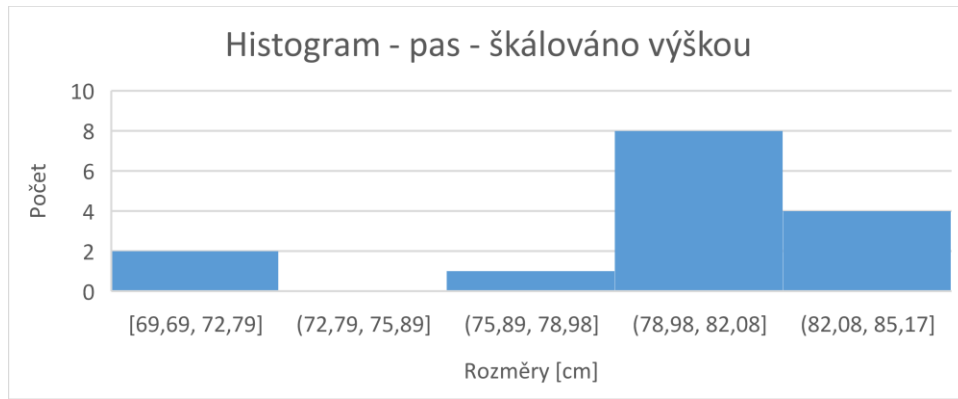
Největší rozdíl mezi reálnou hodnotou a střední hodnotou velikostí modelů byl naměřen v oblasti boků. Největší směrodatná odchylka byla naopak měřena u pasu, což znamená, že byl naměřen největší rozdíl mezi největší a nejmenší hodnotou pasu.

Obrázky 8.9-8.11 zobrazují histogramy hlavních tělesných partií (hrudník, pas a boky) pro modely vytvořené dle fotografií s ženskou osobou s různým oblečením a různými pózami a škálováním modelů výškou.

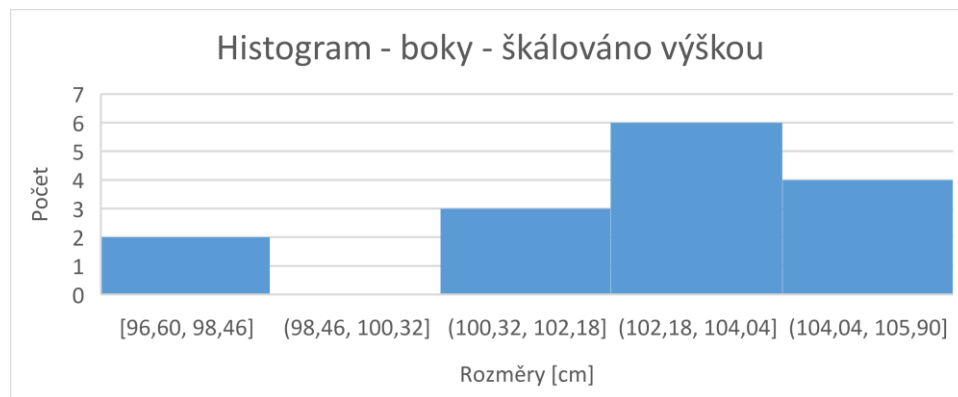
Zvýrazněný sloupec histogramu zobrazuje pozici reálné naměřené tělesné míry osoby. Pokud není zvýrazněný žádný sloupec, reálná naměřená hodnota se v intervalech histogramu vůbec nenachází, anebo nebyl vytvořen žádný model, který by tomuto rozměru odpovídal. Toto platí u všech následujících histogramů.



Obrázek 8.9: Histogram rozměrů hrudníku – žena



Obrázek 8.10: Histogram rozměrů pasu – žena



Obrázek 8.11: Histogram rozměrů boků – žena

Z histogramů lze pozorovat, že algoritmus i po naškálování vytvářel modely spíše větší, než je reálná postava. Zároveň nebyl vytvořen ani jeden model, který by svými rozměry odpovídal rozměrům pasu a boků testované osoby.

Kompletní výsledky měření pro ženský typ modelu lze nalézt v příloženém souboru „Measurements.xlsx“ v záložce „Žena“.

U této osoby bylo také ověřeno, jaké budou výsledky, pokud se pro vytváření modelu vybere model s neutrálním typem pohlaví. Výsledky testování hypotéz s neutrálním modelem jsou následující:

- různé oblečení + různé pózy
 - bez škálování – 0/10 nezamítnutí H_0
 - škálování výškou – 1/9 nezamítnutí H_0
 - škálování rozpětím paží – 1/9 nezamítnutí H_0
- různé oblečení
 - bez škálování – 0/10 nezamítnutí H_0
 - škálování výškou – 1/9 nezamítnutí H_0

- škálování rozpětím paží – 1/9 nezamítnutí H_0
- různé pózy
 - bez škálování – 3/10 nezamítnutí H_0
 - škálování výškou – 1/9 nezamítnutí H_0
 - škálování rozpětím paží – 1/9 nezamítnutí H_0

Celkem bylo nezamítnuto 9/84 nulových hypotéz H_0 . To je téměř stejné množství jako při použití ženského modelu. Získané rozměry neutrálních modelů však byly v tomto případě u setu fotografií s různým oblečením a různými pózami a škálováním modelu výškou většinou lepší oproti modelům ženským, konkrétně byly

- rozměry lepší – rozpětí paží, hrudník, boky, stehno a koleno a
- rozměry horší – pas, zápěstí, kotník a krk.

Maximální rozdíl mezi ženským a neutrálním modelem je v tomto případě téměř 10 cm (boky) ve prospěch neutrálního modelu. Lze předpokládat, že použití ženského modelu je vhodné především pro ženy s výraznými hlavními ženskými partiemi. Pro neutrálnější typ postavy může být vhodnější použití neutrálního typu modelu. Tuto hypotézu by bylo nutné ověřit zúčastněním testování většího počtu žen s rozmanitými typy postav.

Kompletní výsledky měření pro neutrální typ modelu lze nalézt v příloženém souboru „Measurements.xlsx“ v záložce „Žena (N)“.

Celkové výsledky pro tuto osobu nejsou příliš dobré, odchylky od reálných rozměrů jsou s použitím ženského modelu i okolo 15–20 cm. Zlepšení bylo dosaženo použitím neutrálního modelu, přesto rozměry stále neodpovídaly realitě.

U této osoby nebyly prováděny intervalové odhady.

8.3.2 Muž 1

Tabulka 8.8 zobrazuje výchozí rozměry použitého mužského modelu a naměřené míry dané osoby. Osoba pro účely diplomové práce poskytla všechny potřebné fotografie.

	Výška	Rozpětí paží	Hrudník	Pas	Boky	Zápěstí	Stehno	Koleno	Kotník	Krk
Výchozí hodnoty modelu:	179,00	184,00	103,38	94,08	102,44	17,76	60,54	38,72	22,47	39,11
Naměřené hodnoty osoby:	168,00	169,00	87,00	78,00	89,00	17,00	50,00	34,00	20,00	38,00

Tabulka 8.8: Vstupní rozměry – muž 1

Výsledky testování hypotéz s mužským modelem jsou následující:

- různé oblečení + různé pózy
 - bez škálování – 0/10 nezamítnutí H_0
 - škálování výškou – 1/9 nezamítnutí H_0
 - škálování rozpětím paží – 2/9 nezamítnutí H_0
- různé oblečení
 - bez škálování – 0/10 nezamítnutí H_0
 - škálování výškou – 1/9 nezamítnutí H_0
 - škálování rozpětím paží – 2/9 nezamítnutí H_0
- různé pózy
 - bez škálování – 1/10 nezamítnutí H_0
 - škálování výškou – 5/9 nezamítnutí H_0
 - škálování rozpětím paží – 7/9 nezamítnutí H_0

Celkem bylo nezamítnuto 19/84 nulových hypotéz H_0 . Oproti testované ženské osobě je toto větší počet nezamítnutí nulových hypotéz H_0 . Rozměry všech vytvořených modelů by tudíž měly být blíže realitě. Zajímavostí je, že většího počtu nezamítnutí bylo dosaženo se škálováním pomocí rozpětí paží. Rozdíl mezi reálnou naměřenou výškou osoby a rozpětím paží je 1 cm.

Modely bez škálování byly velmi vzdálené od reality. Výška těchto modelů se lišila i o až 50 cm. Extrémním případem byla póza ve dřepu. Z této fotografie vytvořila metoda model vysoký téměř 3 m, čemuž odpovídaly i další rozměry tohoto modelu. Po naškálování jsou však modely realitě již mnohem blíže.

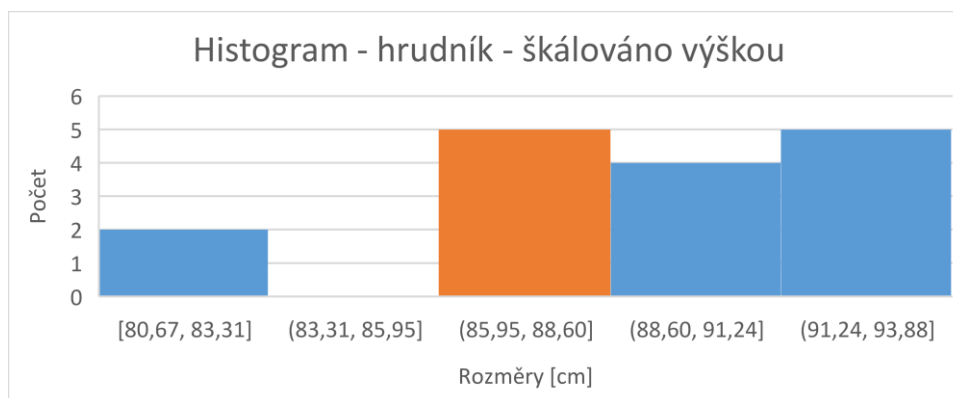
Tabulka 8.9 zobrazuje konkrétní statistické výsledky pro modely vytvořené dle fotografií s mužskou osobou s různým oblečením a různými pózami a škálováním modelu výškou.

	Rozpětí paží	Hrudník	Pas	Boky	Zápěstí	Stehno	Koleno	Kotník	Krk
Výchozí hodnoty modelu:	184,00	103,38	94,08	102,44	17,76	60,54	38,72	22,47	39,11
Naměřené hodnoty osoby:	169,00	87,00	78,00	89,00	17,00	50,00	34,00	20,00	38,00
Přečtené hodnoty poměrem výšky:									
Minimum	166,34	80,67	72,13	85,56	15,08	48,83	33,39	18,70	29,76
Maximum	176,57	93,88	87,93	94,49	16,52	54,66	35,85	20,76	35,52
Střední hodnota	170,47	89,51	82,73	91,03	15,96	50,84	35,00	19,90	33,60
Průměr	170,59	89,04	82,42	90,77	15,89	51,42	34,84	19,84	33,30
Směrodatná odchylka	2,29	3,67	4,26	2,22	0,45	1,47	0,63	0,51	1,46
Výběrová směrodatná odchylka	2,36	3,79	4,40	2,29	0,46	1,52	0,65	0,53	1,50
Rozdíl stf. h. a naměřené h.	1,47	2,51	4,73	2,03	1,04	0,84	1,00	0,10	4,40
T-test	2,70	2,15	4,02	3,09	-9,55	3,73	5,15	-1,23	-12,50
	Zamítám H_0	Zamítám H_0	Zamítám H_0	Zamítám H_0	Zamítám H_0	Zamítám H_0	Zamítám H_0	Zamítám H_0	Zamítám H_0

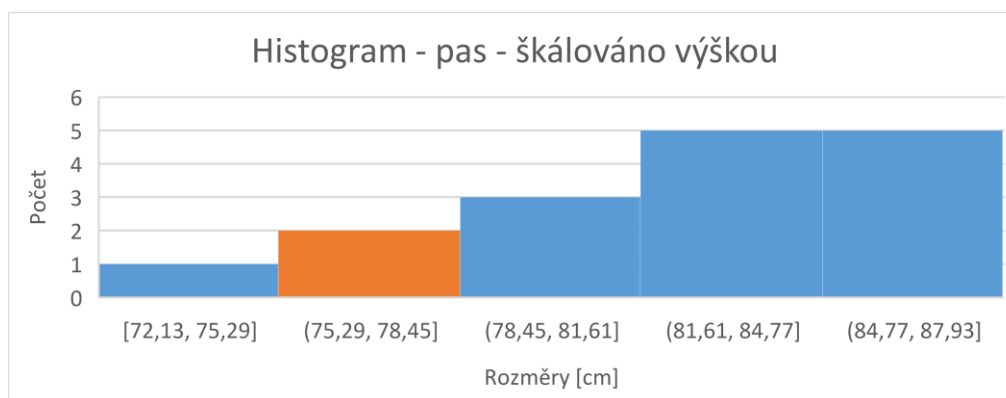
Tabulka 8.9: Výsledné rozměry modelů – muž 1

Největší rozdíl mezi reálnou hodnotou a střední hodnotou velikostí modelů byl naměřen v oblasti pasu. Dále lze pozorovat, že dolní končetiny velmi dobře odpovídají reálným naměřeným hodnotám. Vrchní část těla se liší také pouze v jednotkách centimetrů.

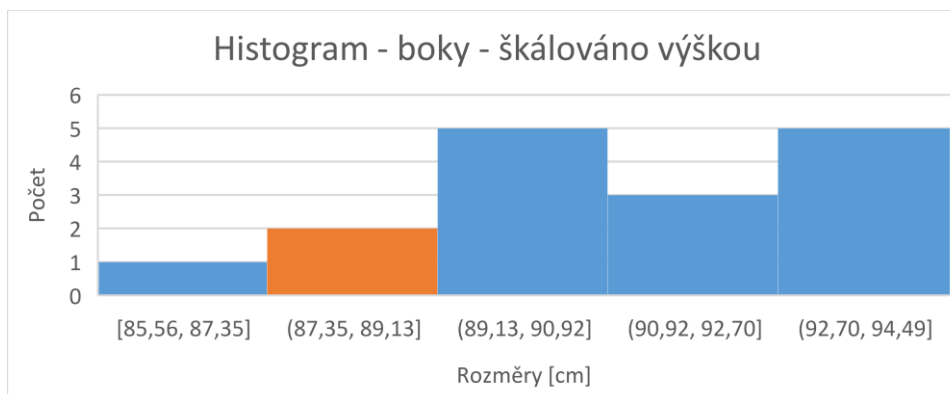
Obrázky 8.12-8.14 zobrazují histogramy hlavních tělesných partií (hrudník, pas a boky) pro modely vytvořené dle fotografií s mužskou osobou s různým oblečením a různými pózami a škálováním modelů výškou.



Obrázek 8.12: Histogram rozměrů hrudníku – muž 1



Obrázek 8.13: Histogram rozměrů pasu – muž 1



Obrázek 8.14: Histogram rozměrů boků – muž 1

Z histogramů lze pozorovat, že algoritmus i po naškálování vytvářel modely spíše větší, než je reálná postava. Nejvýrazněji to lze pozorovat u měření pasu. Zvýrazněný sloupec histogramu boků je velmi blízko početnějšího sloupce, který také není daleko od reality, jelikož skutečný rozměr boků je 89 cm. Rozměry modelů lépe odpovídají naměřeným mírám testovací osoby než při testování ženy, odchylky jsou v tomto případě menší.

S mužským typem modelu byly vytvořeny také intervaly spolehlivosti. Bylo pro to použito 34 fotografií ve stejném typu oblečení. Jednotlivé vytvořené modely splňují většinu limitů intervalů pro dané partie, nikdy však ne pro zcela všechny (myšleno pro vybrané modely). Některé modely tudíž splňují intervaly pro téměř všechny partie, naopak některé modely splňují intervaly pouze pro některé partie. Celkem bylo u 34 vytvořených modelů splněno následující množství intervalů s 95% spolehlivostí:

- bez škálování – 87/340 splnění 95% intervalů (25.6% úspěšnost); 1 model kompletní uvnitř všech 99% intervalů
- škálování výškou – 67/306 splnění 95% intervalů (21.9% úspěšnost)
- škálování rozpětím paží – 71/306 splnění 95% intervalů (23.2% úspěšnost)

U všech modelů bylo dosaženo zhruba 25% celkové úspěšnosti splnění. Podobné množství splnění bylo dosaženo i u modelů bez škálování. To je dáno tím, že jsou intervaly v tomto případě širší oproti intervalům u modelů se škálováním, a tak se do nich vejde podobné množství měřených rozměrů. Testování také neukazuje, jak blízko jsou modely realitě. V intervalech, které jsou splněny, vůbec nemusí ležet reálná naměřená hodnota.

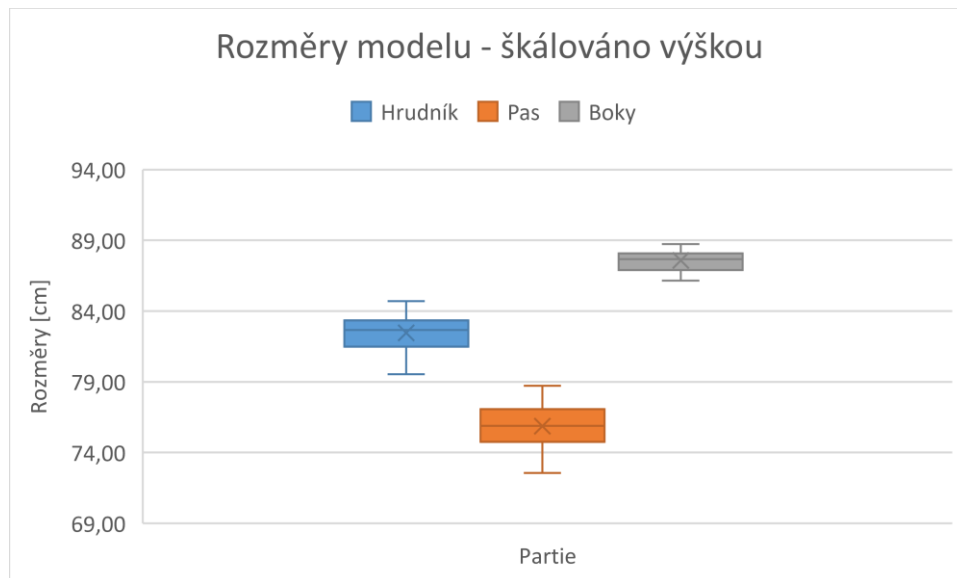
Pokud by byla naměřena úspěšnost blízká 100 %, znamenalo by to, že metoda za velmi podobných podmínek vytváří modely vždy velmi podobné a že mezi nimi nedochází k výraznějším odchýlkám.

Testovány byly i intervaly s 99% spolehlivostí, splnění těchto intervalů jsou obdobná, horní ani dolní mez se již příliš neliší.

Obrázek 8.15 zobrazuje krabicový graf hlavních tělesných měř (hrudník, pas a boky) pro modely vytvořené dle fotografií použitých pro vytvoření intervalových odhadů. Krabicový graf je typ grafu, který zobrazuje data pomocí kvartilů. Uvnitř obdélníkové části grafu je linií vyznačen medián a křížkem průměr. Horní hrana tohoto obdélníku představuje 3. kvartil (75%, oddělí 75 % nejmenších dat od 25 % nejvyšších) a dolní

hrana představuje 1. kvartil (25%, oddělí 25 % nejmenších dat od 75 % nejvyšších). Linie vycházející z tohoto obdélníku nahoru a dolů vyjadřuje celkovou variabilitu dat (maximum a minimum).

V ideálním případě požadujeme, aby variabilita byla co nejmenší a aby medián (a průměr) co nej přesněji odpovídal reálné naměřené hodnotě.



Obrázek 8.15: Rozměry hlavních partií – muž 1 (reálné rozměry: hrudník – 87 cm, pas – 78 cm, boky – 89 cm)

Nejmenší variability bylo u hlavních partií dosaženo u boků, ty byly také nejbližší reálným bokům osoby. U rozměrů hrudníku a pasu lze pozorovat již větší variabilitu a odchylku od reálných rozměrů. U dalších menších partií je variabilita již menší, než u zde uvedených partií. To znamená, že menší partie zůstávají podobně velké, i když se mění rozměry hlavních partií. Stejný závěr lze tvrdit i na základě vypočtených směrodatných odchylek.

Kompletní výsledky měření pro mužský typ modelu lze nalézt v příloženém souboru „Measurements.xlsx“ v záložce „Muž“ a výsledky intervalových odhadů včetně grafů pro další partie v záložce „Interval“.

U této osoby bylo také ověřeno, jaké budou výsledky, pokud se pro vytváření modelu vybere model s neutrálním typem pohlaví. Výsledky testování hypotéz s neutrálním modelem jsou následující:

- různé oblečení + různé pózy
 - bez škálování – 1/10 nezamítnutí H_0
 - škálování výškou – 2/9 nezamítnutí H_0

- škálování rozpětím paží – 5/9 nezamítnutí H_0
- různé oblečení
 - bez škálování – 0/10 nezamítnutí H_0
 - škálování výškou – 1/9 nezamítnutí H_0
 - škálování rozpětím paží – 2/9 nezamítnutí H_0
- různé pózy
 - bez škálování – 7/10 nezamítnutí H_0
 - škálování výškou – 7/9 nezamítnutí H_0
 - škálování rozpětím paží – 7/9 nezamítnutí H_0

Celkem bylo nezamítnuto 32/84 nulových hypotéz H_0 . To je o 13 více oproti mužskému typu modelu, získané rozměry jsou však téměř totožné. Maximální rozdíl u setu fotografií s různými typy oblečení a různými typy póz a škálování modelu výškou je zhruba 1.5 cm. Bez škálování byly modely obdobně vysoké a velké jako při použití mužského typu modelu.

Kompletní výsledky měření pro neutrální typ modelu lze nalézt v příloženém souboru „Measurements.xlsx“ v záložce „Muž (N)“.

Vytvořené modely pro tuto osobu (mužské i neutrální) považují subjektivně za dobré s přihlédnutím k tomu, že se jedná o automatickou metodu využívající pouze jedinou fotografii.

8.3.3 Muž 2

Tabulka 8.10 zobrazuje výchozí rozměry použitého mužského modelu a naměřené míry dané osoby. Jedná se o stejnou osobu jako v předchozí kapitole 8.3.2, došlo pouze k simulaci mírně nadprůměrné postavy vycpáním břicha. Osoba pro účely diplomové práce poskytla všechny potřebné fotografie kromě oblečení prvního typu, které kvůli vycpání břicha nebylo realizovatelné.

	Výška	Rozpětí paží	Hrudník	Pas	Boky	Zápěstí	Stehno	Koleno	Kotník	Krk
Výchozí hodnoty modelu:	179,00	184,00	103,38	94,08	102,44	17,76	60,54	38,72	22,47	39,11
Naměřené hodnoty osoby:	168,00	169,00	87,00	96,00	89,00	17,00	50,00	34,00	20,00	38,00

Tabulka 8.10: Vstupní rozměry – muž 2

Výsledky testování hypotéz s mužským modelem jsou následující:

- různé oblečení + různé pózy

- bez škálování – 1/10 nezamítnutí H_0
- škálování výškou – 1/9 nezamítnutí H_0
- škálování rozpětím paží – 1/9 nezamítnutí H_0
- různé oblečení
 - bez škálování – 0/10 nezamítnutí H_0
 - škálování výškou – 1/9 nezamítnutí H_0
 - škálování rozpětím paží – 0/9 nezamítnutí H_0
- různé pózy
 - bez škálování – 2/10 nezamítnutí H_0
 - škálování výškou – 2/9 nezamítnutí H_0
 - škálování rozpětím paží – 4/9 nezamítnutí H_0

Celkem bylo nezamítnuto 12/84 nulových hypotéz H_0 . To je méně než v případě stejné osoby bez vycpání břicha a použití stejného mužského modelu. Může to být způsobeno tím, že vycpání na fotografiích není příliš poznat, došlo však ke změně čísla použitého u výpočtu testovacího kritéria.

Tabulka 8.11 zobrazuje konkrétní statistické výsledky pro modely vytvořené dle fotografií s mužskou osobou s různým oblečením a různými pózami a škálováním modelu výškou.

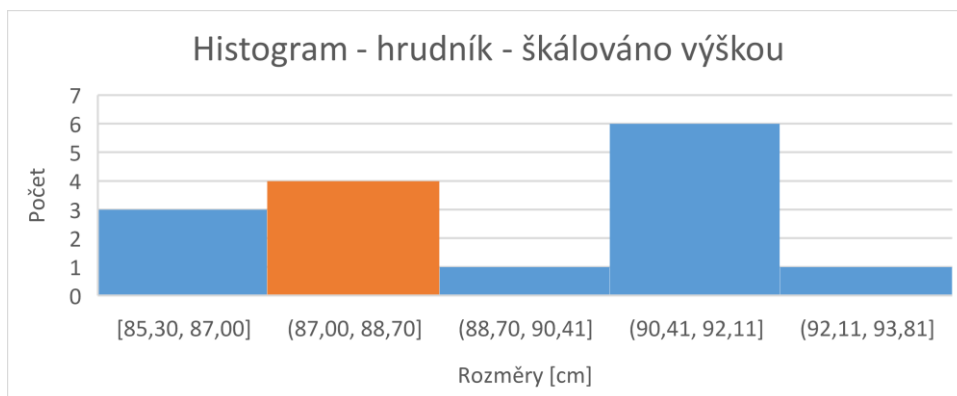
	Rozpětí paží	Hrudník	Pas	Boky	Zápěstí	Stehno	Koleno	Kotník	Krk
Naměřené hodnoty osoby:	169,00	87,00	96,00	89,00	17,00	50,00	34,00	20,00	38,00
Přepočtené hodnoty poměrem výšky:									
Minimum	168,00	85,30	77,85	87,95	15,05	47,22	34,26	18,42	29,09
Maximum	174,33	93,81	87,17	94,58	16,35	55,12	35,84	20,74	35,53
Střední hodnota	170,50	90,35	82,90	91,06	16,05	52,07	34,91	20,02	33,98
Průměr	170,29	89,37	82,86	91,12	15,93	51,96	34,93	19,95	33,61
Směrodatná odchylka	1,79	2,33	2,51	1,78	0,30	1,82	0,47	0,51	1,39
Výběrová směrodatná odchyl	1,85	2,41	2,59	1,84	0,31	1,89	0,49	0,53	1,44
Rozdíl stř. h. a naměřené h.	1,50	3,35	13,10	2,06	0,95	2,07	0,91	0,02	4,02
T-test	2,71	3,81	-19,62	4,45	-13,46	4,02	7,33	-0,35	-11,80
	Zamítám H_0	Zamítám H_0	Zamítám H_0	Zamítám H_0	Zamítám H_0	Zamítám H_0	Zamítám H_0	Nezamítám H_0	Zamítám H_0

Tabulka 8.11: Výsledné rozměry modelů – muž 2

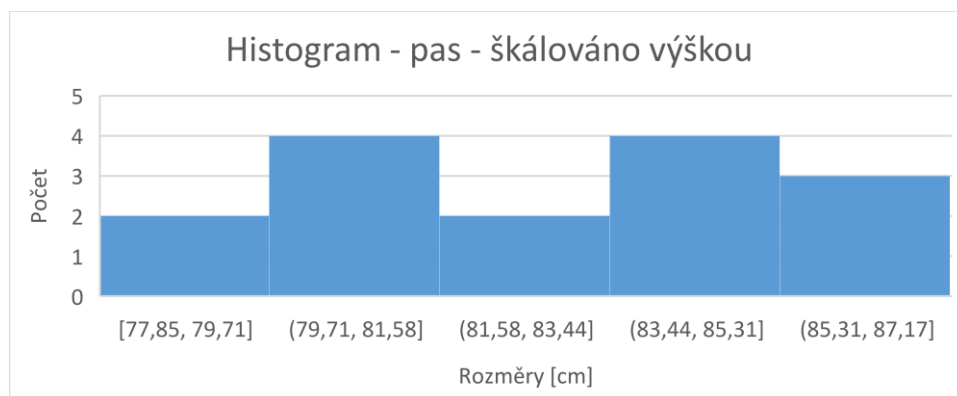
Největší rozdíl mezi reálnou hodnotou a střední hodnotou velikostí modelů byl naměřen v oblasti pasu (místo vycpání). Dále lze pozorovat, že rozměry modelů v oblasti pasu jsou velmi podobné jako u osoby bez vycpání břicha, což potvrzuje, že vycpání nebylo na fotografiích příliš poznat. Pokud by však i reálná osoba byla stejného typu postavy (tzn. rozměrnější směrem dopředu než do šířky), bylo by to pro metodu opět těžké detekovat a opět by pravděpodobně bylo dosaženo špatných výsledků v oblasti této partie.

Bez naškálování byly modely ve všech ohledech opět až příliš velké a paradoxně k nezamítnutí hypotézy H_0 došlo právě u vycpané oblasti pasu. Všechny ostatní rozměry však byly vzdálené od reality.

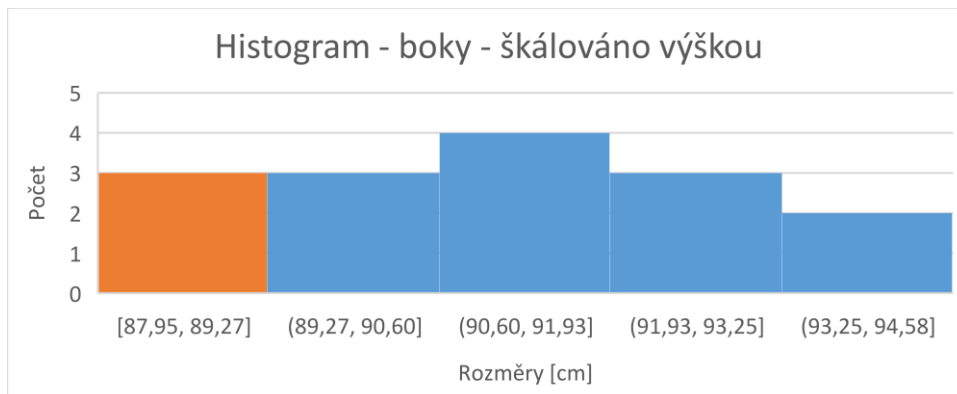
Obrázky 8.16-8.18 zobrazují histogramy hlavních tělesných partií (hrudník, pas a boky) pro modely vytvořené dle fotografií s mužskou osobou (s vycpáním břicha) s různým oblečením a různými pózami a škálováním modelů výškou.



Obrázek 8.16: Histogram rozměrů hrudníku – muž 2



Obrázek 8.17: Histogram rozměrů pasu – muž 2



Obrázek 8.18: Histogram rozměrů boků – muž 2

Z histogramů lze pozorovat, že algoritmus i po naškálování vytvářel modely spíše větší, než je reálná postava. Také můžeme vidět, že ani jednou nebyl vytvořen model se správnou velikostí pasu.

Jelikož je typ této postavy pouze simulován a tělo není zcela reálné, byl testován pouze s mužským modelem.

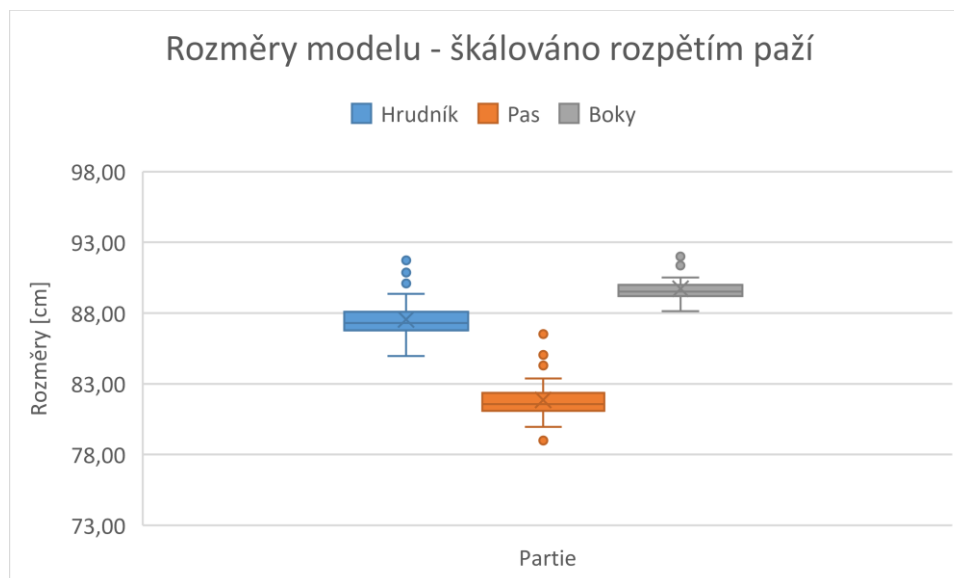
S mužským typem modelu byly vytvořeny také intervaly spolehlivosti. Bylo pro to použito 34 fotografií ve stejném typu oblečení. Jednotlivé vytvořené modely splňují většinu limitů intervalů pro dané partie, nikdy však ne pro zcela všechny. Oproti stejné postavě bez vycpání břišní části těla je však splnění limitů o něco více. Celkem bylo u 34 vytvořených modelů splněno následující množství intervalů:

- bez škálování – 97/340 splnění 95% intervalů (28.5% úspěšnost); 1 model kompletní uvnitř všech 95% intervalů a 4 uvnitř všech 99% intervalů
- škálování výškou – 87/306 splnění 95% intervalů (28.4% úspěšnost)
- škálování rozpětím paží – 99/306 splnění 95% intervalů (32.4% úspěšnost)

U všech modelů bylo dosaženo okolo 30% celkové úspěšnosti splnění. Opět platí výše uvedené, u modelů bez škálování docházelo ke splnění limitů, jelikož jsou tyto limity širší oproti limitům se škálováním a vytvářené modely jsou podobně velké.

Testovány byly i intervaly s 99% spolehlivostí, splnění těchto intervalů jsou obdobná, horní ani dolní mez se již příliš neliší.

Obrázek 8.19 zobrazuje krabicový graf hlavních tělesných měr (hrudník, pas a boky) pro modely vytvořené dle fotografií použitých pro vytvoření intervalových odhadů.



Obrázek 8.19: Rozměry hlavních partií – muž 2 (reálné rozměry: hrudník – 87 cm, pas – 96 cm, boky – 89 cm)

Nejmenší variability bylo u hlavních partií dosaženo u boků, ty byly také nejbližší reálným bokům osoby. Medián také velmi dobře odpovídá u rozměrů hrudníku, přesto je zde pozorována větší variabilita oproti bokům. U rozměrů pasu lze pozorovat také již větší variabilitu a odchylku od reálných rozměrů – medián se liší téměř o 15 cm. U dalších menších partií je variabilita menší, než u zde uvedených partií.

Kruhové body v grafu vyznačují ojedinělé extrémní hodnoty. Linie vycházející z obdélníku nahoru a dolů v tomto případě vyjadřují maximum a minimum bez zahrnutí těchto extrémních případů.

Kompletní výsledky měření pro mužský typ modelu lze najít v příloženém souboru „Measurements.xlsx“ v záložce „Muž (2)“ a výsledky intervalových odhadů včetně grafů pro další partie v záložce „Interval (2)“.

8.4 Shrnutí

Výsledky odhadu výšky pro finální nastavení parametrů algoritmu byly velmi dobré. Pro skupinu fotografií A byla průměrná odhadnutá výška pouze o zanedbatelných 0.06 cm menší než reálná výška osoby a střední hodnota o 0.44 cm větší. Pokud byly pro vyhodnocení použity fotografie skupiny B focené bez znalosti, jak udělat vhodnou fotografii, byla průměrná odhadnutá výška o 5.75 cm menší a střední hodnota o 6.23 cm menší oproti reálné výšce. To považuji za stále velmi dobré výsledky – naškálováním touto výškou by vytvořené modely byly stále blíže realitě než bez naškálování. Současně bylo dosaženo 96% úspěšnosti detekce dveří. S jiným nastavením se lze u kompletního

setu testovacích fotografií dostat až na 100 %. Hodnoty odhadnutých výšek jsou však již mírně horší (v rámci několika desetin cm).

Pokud se podíváme na statistické výsledky testování rozměrů modelů, může se zdát, že získané modely nejsou příliš kvalitní. Já osobně však i tyto výsledky považuji za dobré – toto platí převážně u mužského modelu. Pokud se vytvoří alespoň několik modelů, lze mezi nimi najít takový, který velmi dobře rozměry odpovídá reálným mírám osoby. Toto platí po naškálování modelů, bez škálování byly vytvářené modely ve všech ohledech výrazně větší. Výsledky ženských modelů oproti mužskému byly horší, výsledné modely byly oproti reálným mírám i po škálování většinou větší. U ženy se v tomto případě vyplatilo použít neutrální typ modelu, což může být dáno typem její postavy. Ženský model je pravděpodobně vhodné používat pro osoby s výraznými ženskými partiemi. Toto tvrzení by ale muselo být ověřeno otestováním většího množství osob. U muže vedlo použití mužského a neutrálního modelu k obdobným rozměrům modelů.

Mezi typem oblečení a rozměry vytvářených rozměrů nebyla nalezena přímá spojitost. Lze např. pozorovat, že některé fotografie s osobou méně oblečenou vedly k modelům větším než s fotografií s osobou více oblečenou a naopak. Toto lze ověřit porovnáním jednotlivých rozměrů modelů pro každou fotografii uvedených v příloženém souboru „Measurements.xlsx“. Naopak největší vliv na velikost modelu má póza „dřep“, při které metoda bez naškálování vytváří jednoznačně největší modely.

Z uvedených výsledků lze dále pozorovat, že algoritmus se lépe vypořádává s různými typy póz v úzkém oblečení oproti stejné základní póze a různým typům oblečení. Za zmínku také stojí výsledky modelů bez škálování, u kterých se může dle statistických testů zdát, že mají obdobnou kvalitu jako modely naškálované. To je způsobeno většími hodnotami výběrových směrodatných odchylek, které hrají roli ve výpočtech T-testů a intervalů spolehlivosti. Největší variability u vytvářených modelů dochází u hlavních partií (hrudník, pas, boky), menší vedlejší partie se pak již příliš nemění. Toto ukázaly výsledky intervalových odhadů a také směrodatné odchylky. Největší směrodatná odchylka byla ve všech třech uvedených případech testovacích osob měřena v oblasti pasu – je to partie, která se u vytvářených modelů mezi sebou navzájem nejvíce odlišuje.

Subjektivní hodnocení podoby s modely získané vytvořenou internetovou aplikací ukázalo největší nespokojenost s pózou modelů, druhou nejčastější odpovědí naopak byla spokojenost se všemi dotazovanými údaji.

9 Možnosti rozšíření práce

Jedním z možných budoucích rozšíření diplomové práce je otestování algoritmu SMPLify na větším počtu osob s rozmanitějšími typy postav. Nyní byl algoritmus vyzkoušen pouze na dvou osobách s průměrnými tělesnými mírami. Postava třetí osoby byla pouze simulována vycpáním oblečení, což ne dokonale odpovídá reálné postavě a je tak pouze obtížné hodnotit kvality metody pro takový typ postavy. Vycpání také postavu rozšířilo hlavně směrem dopředu, což na fotografiích není tolik poznat. Důvodem současného neotestování více osob byla převážně časová náročnost tohoto experimentu (focení setu fotografií v různých typech oblečení, měření tělesných měř) a s tím spojená ne příliš velká ochota se testování zúčastnit a poskytnout fotografie a údaje.

Pro sběr podrobnějších údajů včetně tělesných rozměrů by bylo možné po rozšíření využít i stávající internetovou aplikaci. Dobrovolníci by měli možnost zadat své tělesné míry a výsledkem by kromě zobrazení 3D modelů mohla být i statistika, jak moc se vytvořené modely liší od jejich naměřených údajů. Měřit by se také mohly i další tělesné partie, např. na více místech rozměry paží.

Dalším možným vylepšením již hotových algoritmů by mohla být automatická detekce pohlaví osoby na fotografii. Řešením by mohlo být využití neuronových sítí, podobně jako jsou pomocí nich na fotografiích odhadovány pozice kloubů osoby algoritmem DeepCut. Při dostatečné úspěšnosti odhadu pohlaví by pak byl uživatel zbaven nutnosti vybírat pohlaví manuálně.²

Nakonec by bylo zajímavé vyzkoušet jiný typ metody pro automatické vytváření 3D modelů a pro stejné vstupní fotografie mezi sebou metody navzájem statisticky porovnat. Rychlejší metodu by bylo vhodné použít pro internetovou aplikaci a pro získání modelů co nejvíce odpovídajících realitě přesnější metodu. Dle autorů metody SMPLify-X je jejich nová metoda spolu s využitím modelu SMPL-X zároveň rychlejší i přesnější.

² Později bylo zjištěno, že funkci automatické detekce pohlaví již realizovali autoři metody SMPLify-X.

10 Závěr

V diplomové práci byly popsány existující algoritmy pro automatické vytvoření 3D modelu z jedné fotografie. Algoritmy byly mezi sebou na teoretické úrovni porovnány a následně jeden z nich vybrán a otestován.

Pro prohlížení 3D modelů a vyzkoušení vybrané metody byla vytvořena internetová aplikace. Po výběru pohlaví a nahrání své fotografie je uživateli zobrazen 3D model jeho postavy s možností interaktivního prohlížení. Pokud má uživatel zájem, může sdílet své osobní údaje a zhodnocení podoby své postavy a modelu. Funkčnost aplikace byla otestována vybranou skupinou studentů. Tito studenti zároveň vyhodnotili podobu vytvořeného modelu se zaslanou fotografií. Nejvíce nespokojení byli s pózou modelu.

Kvalita vybrané metody byla dále zhodnocena na základě rozměrů získávaných 3D modelů. Testování bylo provedeno na mužské a ženské postavě s průměrnými tvary těl a dále byla simulována i mírně nadprůměrná mužská postava. Osoby byly vyfoceny v různých typech oblečení a při různých pózách a vyhodnocení proběhlo na základě získaných základních statistických údajů. Model byl měřen na osmi tělesných partiích a dále byla měřena výška osoby a rozpětí paží.

Jelikož bylo zjištěno, že metoda špatně odhaduje výšku postavy a modely jsou pak ve všech rozměrech zbytečně velké a neodpovídají realitě, byla vytvořena funkce pro automatický odhad výšky postavy podle výšky dveří. Ve vstupní fotografii jsou vyhledávány objekty obdélníkového tvaru. Podle objektu, o kterém je rozhodnuto, že s největší pravděpodobností odpovídá dveřím, je pak podle počtu pixelů na výšku odhadnuta výška osoby vyfocené u těchto dveří. Odhadnutou výškou je pak model přeškálován, čímž je zajištěno vytvoření modelů s rozměry více odpovídajícími skutečným mírám osob. Bylo vyzkoušeno i škálování rozpětím paží, což vedlo k obdobným výsledkům.

Po naškálování modelů jejich rozměry poměrně dobře odpovídají reálným naměřeným mírám testovacích osob – toto platí hlavně pro mužskou osobu. Rozměry získaných mužských modelů byly lepší než rozměry ženských modelů. Vycpáním oblečení mužské postavy v oblasti břicha byla simulována i mírně nadprůměrná postava, výsledky však byly v místě vycpání již horší. Není jasné, zda jsou tyto výsledky způsobeny snahou o zvětšení postavy, anebo zda se jedná o nedokonalost metody. Aby bylo možné tvrdit,

že testovaná metoda vždy za všech podmínek vytváří modely dokonale odpovídající realitě, bylo by nutné, aby vyhodnocované statistické testy dopadly ještě výrazně lépe.

Možným budoucím rozšířením diplomové práce je otestování metody s více osobami s rozmanitými typy postav. Pro další sběr podrobnějších dat by bylo možné po případném rozšíření využít stávající internetovou aplikaci. Vhodné by také bylo vyzkoušet pokročilou metodu SMPLify-X s novým expresivním modelem SMPL-X. Tato metoda by dle autorů měla být oproti stávající metodě SMPL rychlejší a zároveň i mírně přesnější.

Příloha

SMPL (2015) je vertexový 3D model lidského těla, který je naučený z tisíců 3D skenů těl. Přesně reprezentuje velké množství tvarů těl a póz v nízko-dimenzionálním parametrickém prostoru. SMPL používá analýzu hlavních komponent (PCA) k naučení na póze nezávislé Euklidovské reprezentaci tvaru těla. Model je dle autorů přesnější než dříve představené modely. Jeho síť obsahuje 6980 vrcholů. Je zdarma veřejně dostupný pro nekomerční výzkumné vědecké účely a je kompatibilní s existujícími renderovacími enginy. [6], [8]

SCAPE (Shape Completion and Animation for People, 2005) je metoda pro vytváření modelu lidského těla. Model pokrývá množství tvarů a póz subjektu. SCAPE model je generován ze setu skenů a je rozdělen na deformační model pózy a deformační model tvaru těla. Toto rozdělení usnadňuje matematické formulace, zlepšuje rozpoznatelnost modelu z dat a umožňuje použití efektivnějších učících algoritmů. Jednotlivé modely mohou být zkombinovány pro vytvoření postavy s tělem v libovolné póze a s realistickou deformací svalů. [5]

SMPL-X (SMPL eXpressive = expresivní, 2019) model je založený na modelu SMPL. Model si zachovává výhody svého předchůdce: kompatibilita s grafickými programy, jednoduchá parametrizace, malá velikost, efektivita, diferencovatelnost atd. SMPL-X model je složen z SMPL modelu s modelem hlavy FLAME a s modelem rukou MANO. Výsledný model je registrován na 5586 3D skenů. Naučením modelu z dat jsou zachyceny přirozené spojitosti mezi tvary těl, obličejů a rukou a výsledný model je bez artefaktů, jaké lze nalézt např. u modelu FRANK. SMPL-X je volně dostupný pro vědecké účely. [15]

Spouštění algoritmů

Pro spouštění algoritmů byly vytvořeny skripty v programovacím jazyce bash. Skripty jsou použity pro zpracování vstupních parametrů (jméno fotografie, pohlaví, popř. výška) a pro jednoduchou možnost sekvenčního spouštění Python skriptů.

- Změna rozlišení + odhad výšky:

```
./run_height.sh $file,
```

kde \$file je jméno vstupní fotografie (včetně přípony souboru).

- Změna rozlišení + DeepCut + SMPLify:

```
./run_smpl.sh $file $gender_input,
```

kde \$file je jméno vstupní fotografie (včetně přípony souboru) a \$gender_input je typ zvoleného pohlaví.

- Změna rozlišení + DeepCut + SMPLify + měření modelu:

```
./run_smpl_and_measure.sh $file $gender_input,
```

kde \$file je jméno vstupní fotografie (včetně přípony souboru) a \$gender_input je typ zvoleného pohlaví.

- Změna rozlišení + DeepCut + SMPLify + odhad výšky, škálování a měření modelu:

```
./run_smpl_and_rescale.sh $file $gender_input $height,
```

kde \$file je jméno vstupní fotografie (včetně přípony souboru), \$gender_input je typ zvoleného pohlaví a \$height je výška postavy.

U výše uvedených skriptů je povinné vyplnění pouze prvního parametru (název fotografie ke zpracování). Pokud není vyplněn typ pohlaví, je ke zpracování použitý model s neutrálním typem pohlaví. Typ zadávaného pohlaví se předpokládá v podobě „male“, „female“ nebo „neutral“. Pokud uživatel zadá svoji výšku, je získaný model škálován právě touto výškou. V opačném případě dochází k automatickému odhadu výšky.

Pokud je zadán vyšší počet argumentů, než se předpokládá, dojde k výpisu upozornění a algoritmy nejsou spuštěny.

Algoritmy, které nejsou spouštěny v kombinaci s jinými, lze spouštět následovně:

- DeepCut:

```
python3 deepcut_smpl.py $file,
```

kde \$file je jméno vstupní fotografie (včetně přípony souboru).

- SMPLify:

```
python2 fit_3d.py --file_name=$file --joints_name=$file_without_extension --gender=$gender_input,
```

kde \$file je jméno vstupní fotografie (včetně přípony souboru), \$file_without_extension je jméno vstupní fotografie (bez přípony souboru) a \$gender_input je zvolené pohlaví.

- Změna rozlišení:

```
python3 resize_image.py $file,
```

kde \$file je jméno vstupní fotografie (včetně přípony souboru).

- Nalezení dveří a odhad výšky:

```
python3 estimate_height.py $file,
```

kde \$file je jméno vstupní fotografie (včetně přípony souboru). Jelikož tento skript nevyžaduje vyplnění pohlaví, v případě neúspěšného nalezení dveří je výška určena jako 174.5 (výška definována pro neutrální typ modelu).

U výše uvedených skriptů je povinné vyplnit všechny vstupní parametry. V tomto případě nedochází ke kontrole počtu argumentů, je tedy nutné vyplnit všechny argumenty správně.

Fotografie určené ke zpracování se pro všechny typy skriptů musí nacházet v příložené složce „SMPLify/smplify_public/images/“. Skripty je nutné spouštět z příložené složky „SMPLify/smplify_public/code/“.

Metoda DeepCut je optimalizována pro Python 3, naopak metoda SMPLify byla vytvořena pro Python 2.

Potřebné importy metody DeepCut jsou

- HDF5 1.8,
- CMake,
- C++ 11,
- CUDA ≥ 7.5 (pouze pokud jsou vyžadovány výpočty na GPU) a
- Caffe (vyžaduje další importy jako např. OpenCV, BLAS, Boost a další, více viz [42]).

Potřebné importy metody SMPLify jsou

- OpenCV,
- Numpy $\geq 1.11.0$,

- Scipy>=0.17.1,
- Chumpy,
- OpenDR a
- Matplotlib.

Spuštění internetové aplikace lokálně pomocí XAMPP

Na počítač s Linuxovým typem operačního systému je nejprve nutné nainstalovat aplikaci XAMPP. Nainstalovaná verze použitá pro testování byla „XAMPP 8.0.0-2“. Pro správnou funkčnost aplikace v internetovém prohlížeči je dále nutné mít složku „SMPLify_web“ se zdrojovými soubory ve složce „htdocs“ umístěné v cestě instalace XAMPP. Odpovídající složka musí mít také nastavena linuxová práva pro zápis pro uživatele „Ostatní“. U spouštěcího souboru „run_smpl_web.sh“ je pro správnou funkčnost skriptu nutné upravit cesty vedoucí k souborům umístěným ve složce „/SMPLify/smplify_public/code/“. Stejně je nutné upravit cesty i v příslušných souborech „resize_image_web.py“ (proměnná „img_dir“), „deepcut_smpl_web“ (proměnné „image_name_path“, „model_def“, „model_bin“ a „out_name“) a „fit_3D_web.py“ (proměnné „img_dir“, „data_dir“ a „out_dir“). Důvodem je, že XAMPP potřebuje uvedené celé absolutní cesty a nejde použít relativní cesty.

Přesná cesta záleží na zvoleném umístění složek konkrétního uživatele. Stejně je nutné upravit i proměnnou „\$target_dir“ v souboru „upload.php“. Programu XAMPP pro funkční využití nainstalovaných Python knihoven nestačilo ani přidání cest do „Pythonpath“ v souboru „bashrc“. Proto musely být cesty ke knihovnám přidány přímo do spouštěných souborů pomocí modulu „sys“ a jeho funkce „path.append“ (parametrem je připojená cesta). Pro plnohodnotné vyzkoušení internetové aplikace je proto nutné upravit také cesty k těmto instalovaným knihovnám. Pro samotné spuštění skriptů bez použití internetové aplikace spuštěné s XAMPP není nutné tyto cesty upravovat.

Internetová aplikace se spustí zadáním příkazu „localhost/“ a názvu hlavního spouštěcího souboru do vyhledávací lišty prohlížeče, v tomto případě zadáním „localhost/SMPLify_web/index.php“. Stejně tak je nutné, aby se zobrazované modely nacházely ve stejné hlavní složce „htdocs“, jsou tedy ukládány do složky „htdocs/SMPLify_web/models“. Ukázkové fotografie z hlavní úvodní stránky se nacházejí uvnitř složky „htdocs/SMPLify_web/examples“.

Prohlížení a správa databáze s výsledky jsou možné zadáním příkazu „localhost/phpmyadmin“ do vyhledávací lišty prohlížeče. Po celou dobu práce s aplikací je nutné mít zapnutý program XAMPP a v něm běžící moduly Apache a MySQL.

Welcome to photo to 3D model converter

Instructions:
Select your gender, upload your photo and wait for the process to finish. Finally, you will get a 3D model of your body. You can interactively examine the created model.

Upload Photo

Select your gender:

Don't want to say (neutral)
 Male
 Female


If you select your gender, the appropriate model will be used. Otherwise, the neutral model will be used.

Select image to upload:

No file selected.


The process will take a few minutes to finish. Please, wait.

This is a type of photo you should upload:



(front view, whole body)

This will be your result:



Obrázek 10.1: Internetová aplikace – úvodní stránka

You can share your results with us.

Share details

First Name :

Last Name :

Age :

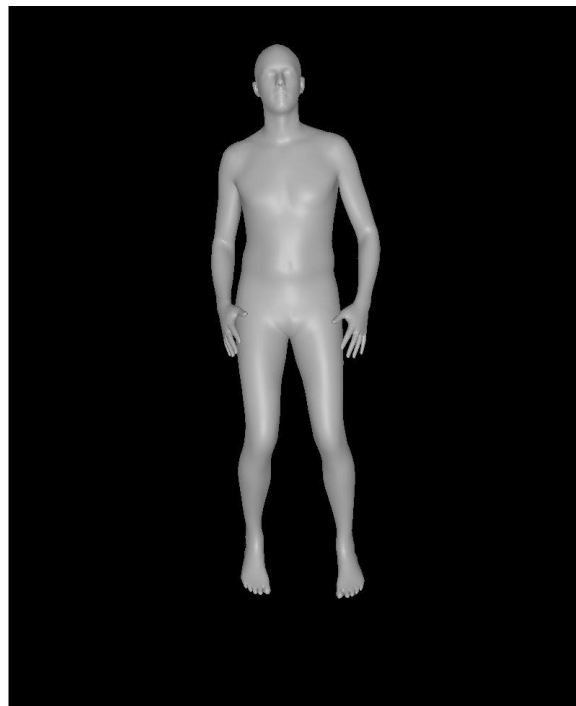
Select your results:

Please, evaluate the similarity between the model and shape, size and pose of your body.

- The shape, size and pose of the model looks exactly like me.
- The size and pose of the model looks like me, but the shape is different.
- The shape and pose of the model looks like me, but the size is different.
- The shape and size of the model looks like me, but the pose is different.
- The shape of the model looks like me, but the size and pose is different.
- The size of the model looks like me, but the shape and pose is different.
- The pose of the model looks like me, but the shape and size is different.
- I don't know.
- The model doesn't look like me at all.

Shape - the overall look and form of the model
Size - sizes of body parts in relation to each other
Pose - the body pose of the model

If something bad happened, please let us know:



Obrázek 10.2: Internetová aplikace – stránka s výsledky

Použitá literatura

Pro vypracování diplomové práce byly použity následující zdroje:

- [1] D. Song, R. Tong, J. Chang, X. Yang, M. Tang and J. J. Zhang, "3D Body Shapes Estimation from Dressed-Human Silhouettes," *Computer Graphics Forum*, vol. 35, no. 7, pp. 147-156, oct 2016.
- [2] G. K. L. Tam, Z.-Q. Cheng, Y.-K. Lai, F. C. Langbein, Y. Liu, D. Marshall, R. R. Martin, X.-F. Sun and P. L. Rosin, "Registration of 3D Point Clouds and Meshes: A Survey from Rigid to Nonrigid," *IEEE Transactions on Visualization and Computer Graphics*, vol. 19, no. 7, pp. 1199-1217, 26 nov 2012.
- [3] H. Li, E. Vouga, A. Gudym, L. Luo, J. T. Barron and G. Gusev, "3D self-portraits," *ACM Transactions on Graphics (TOG)*, vol. 32, no. 187, nov 2013.
- [4] J. Tong, J. Zhou, L. Liu, Z. Pan and H. Yan, "Scanning 3D Full Human Bodies Using Kinects," *IEEE Transactions on Visualization and Computer Graphics*, vol. 18, no. 4, pp. 643 - 650, 09 mar 2012.
- [5] D. Anguelov, P. Srinivasan, D. Koller, S. Thrun, J. Rodgers a J. Davis, „SCAPE: shape completion and animation of people,“ *ACM Transactions on Graphics (TOG)*, sv. 24, pp. 408-416, jul 2005.
- [6] M. Loper, N. Mahmood, J. Romero, G. Pons-Moll and M. J. Black, "SMPL: A Skinned Multi-Person Linear Model," *ACM Trans. Graphics (Proc. SIGGRAPH Asia)*, no. 6, pp. 248:1--248:16, oct 2015.
- [7] N. Hasler, C. Stoll, M. Sunkel, B. Rosenhahn and H. P. Seidel, "A Statistical Model of Human Pose and Body Shape," *Computer Graphic Forum*, no. 28, pp. 337-346, feb 2009.
- [8] S. Streuber, M. A. Quirós-Ramírez, M. Q. Hill, C. A. Hahn, S. Zuffi, A. O'Toole and M. J. Black, "Body Talk: Crowdshaping Realistic 3D Avatars with Words," *ACM Trans. Graph. (Proc. SIGGRAPH)*, vol. 35, no. 4, jul 2016.
- [9] L. Pischulin, E. Insafutdinov, S. Tang, B. Andres, M. Andriluka, P. Gehler a B. Schiele, „DeepCut: Joint Subset Partition and Labeling for Multi Person Pose

- Estimation," *IEEE Conference on Computer Vision and Pattern Recognition, CVPR*, pp. 4929-4937, 2016.
- [10] A. Kanazawa, M. J. Black, D. W. Jacobs and J. Malik, "End-to-end Recovery of Human Shape and Pose," *Computer Vision and Pattern Recognition (CVPR)*, 2018.
- [11] F. Bogo, A. Kanazawa, C. Lassner, P. Gehler, J. Romera and M. J. Black, "Keep it SMPL: Automatic Estimation of 3D Human Pose and Shape," *Computer Vision -- ECCV 2016*, oct 2016.
- [12] S. Johnson and M. Everingham, "Clustered Pose and Nonlinear Appearance Models for Human Pose Estimation," *Proceedings of the British Machine Vision Conference*, 2010.
- [13] L. Sigal, A. Balan and M. J. Black, "HumanEva: Synchronized Video and Motion Capture Dataset and Baseline Algorithm for Evaluation of Articulated Human Motion," *International Journal of Computer Vision*, vol. 87, pp. 1-2, 2010.
- [14] C. Ionescu, D. Papava, V. Olaru and C. Sminchisescu, "Human3.6M: Large Scale Datasets and Predictive Methods for 3D Human Sensing," *Natural Environments, IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 7, jul 2014.
- [15] G. Pavlakos, V. Choutas, N. Ghorbani, T. Bolkart, A. A. A. Osman, D. Tzionas and M. J. Black, "Expressive Body Capture: 3D Hands, Face, and Body from a Single Image," *Proceedings IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, p. 10975 – 10985, 2019.
- [16] C. Lassner, J. Romero, M. Kiefel, F. Bogo, M. J. Black and P. V. Gehler, "Unite the People: Closing the Loop Between 3D and 2D Human Representations," *arXiv:1701.02468v3*, 25 jul 2017.
- [17] S. Zhou, H. Fu, L. Liu, D. Cohen-Or and X. Han, "Parametric reshaping of human bodies in images," *ACM Transactions on Graphics (TOG)*, vol. 29, no. 126, jul 2010.

- [18] P. Guan, A. Weiss, A. Balan and M. J. Black, "Estimating Human Shape and Pose from a Single Image," *IEEE International Conference on Computer Vision*, sep 2009.
- [19] C.-Y. Weng, B. Curless and I. Kemelmacher-Shlizerman, "Photo Wake-Up: 3D Character Animation from a Single Photo," *arXiv:1812.02246v1*, 5 dec 2018.
- [20] K. He, G. Gkioxari, P. Dollár and R. Girshick, "Mask R-CNN," *arXiv:1703.06870*, 20 mar 2017.
- [21] S.-E. Wei, V. Ramakrishna, T. Kanade and Y. Sheikh, "Convolutional Pose Machines," *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4724-4732, 2016.
- [22] P. Krähenbühl and V. Koltun, "Efficient Inference in Fully Connected CRFs with Gaussian Edge Potentials," *Advances in Neural Information Processing Systems*, pp. 109-117, 2011.
- [23] C. Barnes, E. Shechtman, A. Finkelstein and D. B. Goldman, "PatchMatch: A Randomized Correspondence Algorithm for Structural Image Editing," *ACM Transactions on Graphics (Proc. SIGGRAPH)*, vol. 28, no. 3, aug 2009.
- [24] S. Yang, Z. Pan, T. Amert, K. Wang, L. Yu, T. Berg and M. C. Lin, "Physics-Inspired Garment Recovery from a Single-View Image," *ACM Transactions on Graphics (TOG)*, vol. 37, no. 170, nov 2018.
- [25] K. Yamaguchi, M. H. Kiapor and T. L. Berg, "Paper Doll Parsing: Retrieving Similar Styles to Parse Clothing Items," *Proceedings of the 2013 IEEE International Conference on Computer Vision*, pp. 3519-3526, 2013.
- [26] S. Geman and D. McClure, "Statistical methods for tomographic image reconstruction," *Bulletin of the International Statistical Institute*, vol. 52, no. 4, pp. 5-21, 1987.
- [27] Carnegie Mellon University, „CMU Graphics Lab Motion Capture Database,“ 28 10 2019. [Online]. Available: <http://mocap.cs.cmu.edu/>. [Přístup získán 28 10 2019].

- [28] M. Loper, N. Mahmood and M. J. Black, "MoSh: motion and shape capture from sparse markers," *ACM Transactions on Graphics (TOG)*, vol. 33, no. 220, nov 2014.
- [29] M. Chung, „The Gaussian Kernel,“ University of Wisconsin-Madison, [Online]. Available: <http://pages.stat.wisc.edu/~mchung/teaching/MIA/reading/diffusion.gaussian.kernel.pdf.pdf>. [Přístup získán 09 05 2021].
- [30] J. Nocedal and S. Wright, Numerical optimization, Springer, 2016.
- [31] M. M. Loper and M. J. Black, "OpenDR: An Approximate Differentiable Renderer," *European Conference on Computer Vision, ECCV*, pp. 154-169, 2014.
- [32] „chumpy,“ 28 10 2019. [Online]. Available: <https://github.com/mattloper/chumpy>. [Přístup získán 28 10 2019].
- [33] Python, „timeit — Measure execution time of small code snippets,“ [Online]. Available: <https://docs.python.org/3/library/timeit.html>. [Přístup získán 11 04 2021].
- [34] WikiSkripta, „Studentův t-test,“ 30 11 2016. [Online]. Available: https://www.wikiskripta.eu/w/Student%C5%AFv_t-test. [Přístup získán 11 04 2021].
- [35] M. Turčičová, „Intervaly spolehlivosti,“ [Online]. Available: https://www2.karlin.mff.cuni.cz/~turcic/Konfidencni_intervaly.pdf. [Přístup získán 11 04 2021].
- [36] „three.js,“ [Online]. Available: <https://threejs.org/>. [Přístup získán 11 04 2021].
- [37] Apache Friends, „XAMPP Apache + MariaDB + PHP + Perl,“ 2021. [Online]. Available: <https://www.apachefriends.org/index.html>. [Přístup získán 11 04 2021].
- [38] ngrok, „ngrok,“ [Online]. Available: <https://ngrok.com/>. [Přístup získán 28 04 2021].

-
- [39] Wikipedia, „Ramer–Douglas–Peucker algorithm,“ 30 11 2020. [Online]. Available:
https://en.wikipedia.org/wiki/Ramer%E2%80%93Douglas%E2%80%93Peucker_algorithm. [Přístup získán 11 04 2021].
- [40] WorldData, „Body size by country,“ [Online]. Available:
<https://www.worlddata.info/average-bodyheight.php>. [Přístup získán 11 04 2021].
- [41] Blender, „MeasureIt,“ 2021. [Online]. Available:
https://docs.blender.org/manual/en/latest/addons/3d_view/measureit.html.
[Přístup získán 11 04 2021].
- [42] Humanics, „SizeUK: average measurements for MEN & WOMEN,“ [Online]. Available: https://www.humanics-es.com/sizeuk_analysis.pdf. [Přístup získán 11 04 2021].
- [43] Berkeley AI Research, „Installation,“ [Online]. Available:
<http://caffe.berkeleyvision.org/installation.html>. [Přístup získán 11 04 2021].

Obsah elektronických příloh

Příložený adresář elektronických příloh obsahuje složky a soubory v této struktuře:

- Text_prace – textové soubory diplomové práce ve formátu .docx a .pdf
- Poster – poster ve formátu .pub a .pdf
- Aplikace_a_knihovny
 - SMPLify – zdrojové soubory metody DeepCut a metody SMPLify včetně stažených modelů SMPL
 - SMPLify_web – zdrojové soubory internetové aplikace a knihovna three.js-r124
- Vstupni_data – fotografie použité pro testování aplikace
 - Dvere – testovací fotografie pro odhad výšky
 - Muz – testovací fotografie pro měření modelu muže
 - Muz_(2) – testovací fotografie pro měření modelu muže (2)
 - Zena – testovací fotografie pro měření modelu ženy
- Vysledky
 - Excel – zpracování výsledků ve formě excelových tabulek a grafů ve formátu .xlsx
 - Modely – exportované vytvořené modely a textové soubory s výsledky měření partií
- Readme.txt – popis aktuální adresářové struktury