

ŽILINSKÁ UNIVERZITA V ŽILINE
FAKULTA ELEKTROTECHNIKY A INFORMAČNÝCH
TECHNOLÓGIÍ

3D modelovanie objektov z RGB-D snímačov

Dizertačná práca

Štúdijný obor: 5.2.15 Telekomunikácie

Školiace pracovisko: Žilinská univerzita v Žiline

Fakulta elektrotechniky a informačných technológií

Katedra mechatroniky a elektroniky

Školiteľ: doc. Ing. Dušan Koniar, PhD.

Žilina 2020

Ing. Jozef Volak

Obsah

1	Úvod	1
2	Medicínska podstata problému	3
2.1	Aktuálny stav výskumu diagnostiky OSAS s použitím 3D dát	4
2.1.1	Diagnostika pediatrickej OSA prostredníctvom klasifikácie tváre s použitím CNN	5
2.1.2	Predikcia OSA pomocou hlbokého učenia s použitím hĺbkovej mapy tváre	7
2.1.3	Predikcia OSA pomocou 3D modelu	9
2.2	Požiadavky na systém	11
3	Technické pozadie výskumu	12
3.1	Skenovanie v reálnom čase pomocou RGB-D senzorov	12
3.1.1	Viac-pohľadové 3D snímanie a analýza pre rekonštrukciu vysoko kvalitných mračien bodov	13
3.2	Disparitná a hĺbková mapa	15
3.3	Princíp činnosti RGB-D kamier	16
3.3.1	Stereovízia	16
3.3.2	SLS snímače	17
3.3.3	ToF snímače	18
3.4	Mračno bodov a 3D povrch	21
4	Kalibrácia kamier	22
4.1	Afinné transformácie v priestore	22
4.2	Dierková kamera	25

4.3	Skreslenie šošoviek	28
4.3.1	Radiálne skreslenie	29
4.3.2	Tangenciálne skreslenie	30
4.3.3	Model reálnej kamery s korekciou skreslenia šošoviek	30
4.4	Kalibrácia	31
5	Ciele dizertačnej práce	32
6	Návrh snímacieho systému	34
6.1	Microsoft Kinect v2	34
6.2	Snímanie dynamických objektov	35
6.2.1	Jedno-kamerové snímanie dynamických objektov	35
6.2.2	Priestorové rozloženie multi-kamerového systému	36
6.2.3	Sekvenčné snímanie multi-kamerového systému	37
6.2.4	Paralelné snímanie multi-kamerového systému	37
6.2.5	Porovnanie režimov snímania dynamických objektov	37
6.3	Kalibrácia kamerového systému	39
6.3.1	Geometricka kalibrácia	39
6.3.2	Multi-kamerová kalibrácia	41
6.4	Návrh skenovacej metodiky	43
6.4.1	Nástroje použité pri vývoji algoritmu	45
6.4.2	Príprava vstupných dát	46
6.4.3	Identifikácia polohy hlavy	47
6.4.4	Filtrácia a registrácia mračna bodov	52
6.4.5	Automatizované meranie vzdialenosť tvárových črt	57
6.4.6	Segmentácia hĺbkovej mapy	60
7	Potláčanie multi-kamerovej interferencie ToF senzorov	63
7.1	Model ToF kamery	64
7.2	Model multi-kamerovej interferencie	66
7.3	Analýza multi-kamerovej interferencie	67
7.4	Filtračné metódy	69
7.4.1	SOR filter	69

7.4.2	ROR filter	69
7.4.3	PointCleanNet	70
7.4.4	Importance Map Based Median filter	70
7.5	Nastavenie experimentu	74
7.6	Výsledky experimentu	76
7.6.1	SOR filtrácia	77
7.6.2	ROR filtrácia	78
7.6.3	PointCleanNet filtrácia	79
7.6.4	IMBM filtrácia	79
7.6.5	Zhodnotenie výsledkov	81
8	Záver	84
9	Výsledky dizertácie s uvedením nových poznatkov	86
10	Publikované práce autora	88

Zoznam symbolov

ω	uhlová rýchlosť
$ * $	absolútna hodnota
μ	matica zhody
A	intenzita pixelu
$atan$	arcus tangens
B	ofset
c	rýchlosť šírenia elektromagnetickej vlny v priestore
c	stĺpcová súradnica
c_d	modulačný kontrast
$cost (*)$	účelová funkcia
$c_x; c_y; f_x; f_y$	unikátne intrinické parametre kamery
d	hlíbková súradnica
D	Hĺbka
$dist (*)$	euklidovská vzdialenosť
Δ_φ	fázový rozdiel
Δt	časový rozdiel
f	frekvencia
g	robustnosť
$g(t)$	vyžiarený (pôvodný) optický signál
φ	fáza
K	vnútorné parametre kamery
$L (*)$	stratová funkcia
$\max \{ * \}$	maximálna hodnota
$\min \{ * \}$	minimálna hodnota
N	počet snímiel/veľkosť zásobníka
P	matica parametrov kamery
$P(r; c)$	pixel na pozícii r;c
P_d	rozptyl polohy bodu

q	elektrický náboj
R	rotačná matica
r	riadková súradnica
σ	odchýlka merania hĺbky
$s(t)$	prijatý optický signál
t	čas
\mathbf{t}	translačná matica
$\mathbf{T}(\ast)$	transformačný model/ transformačná matica
T_{HIGH}	vysoká prahovacia úroveň
T_{LOW}	nízka prahovacia úroveň

Zoznam skratiek

1D	Jednorozmerný priestor (1-Dimensional)
2D	Dvojrozmerný priestor (2-Dimensional)
3D	Trojrozmerný priestor (3-Dimensional)
ADHD	Porucha aktivity a pozornosti (Attention Deficit Hyperactivity Disorder)
AHI	Apnoe - Hypopnoe Index (Apnea - Hypopnea Index)
AMCW	Amplitúdovo modulovaná kontinuálna vlna (Amplitude Modulated Continuous Wave)
BF	Bilaterálny filter (Bilateral filter)
BMI	Index telesnej hmotnosti (Body Mass Index)
CAD	Počítačom podporovaný návrh (Computer Aided Design)
CNN	Konvolučná neurónová sieť (Convolutional Neural Network)
CPU	Centrálna procesorová jednotka (Central Processing Unit)
CSA	Centrálne spánkové apnoe (Central Sleep Apnea)
CT	Počítačová tomografia (Computed Tomography)
CUDA	Architektúra grafickej karty (Compute Unified Device Architecture)
FIFO	Typ pamäťového zásobníka (First In First Out)
GPU	Grafická procesorová jednotka (Graphical Processing Unit)
HD	Hausdorffova vzdialenosť (Hausdorff Distance)
ICP	Iterative Closest Point (Iterative Closest Point)
IMBM	Výpočet mediánov založený na významnosti (Importance Map Based Median)
IR	Infračervený (Infra Red)
KC	Korelácia kernelu (Kernel Correlation)
KSS	Kamerová súradnicová sústava ()

LDA	Lineárna diskriminačná analýza (Linear Discriminant Analysis)
MCI	Interferencia viacerých kamier (Multi-Camera Interference)
MRI	Magnetická rezonancia (Magnetic Resonance Imaging)
OSA	Obštrukčné spánkové apnoe (Obstructive Sleep Apnea)
OSAS	Syndróm obštrukčného spánkového apnoe (Obstructive Sleep Apnea Syndrome)
PH	Perzistentná homológia (Persistent Homology)
PSG	Polysomnografia (Polysomnography)
pts	points (points)
rad	radius (radius)
RGB	Farebný model RGB (Red Green Blue)
RGB-D	RGB-hĺbka (Red Green Blue - Depth)
RMS	Stredná kvadratická hodnota (Root Mean Square)
ROI	Oblasť záujmu (Region Of Interest)
ROR	Radius Outlier Removal (Radius Outlier Removal)
RPM	Robustné párovanie bodov (Robust Pair Matching)
SCR	Spánkový dotazník (Sleep clinical record)
sdm	standard deviation multiplicator (standard deviation multiplicator)
SLS	Senzor založený na štruktúrovanom svetle (Structured Light Sensor)
SOR	Statistical Outlier Removal (Statistical Outlier Removal)
SSS	Svetová súradnicová sústava ()
ToF	Senzor založený na meraní doby letu (Time of Flight)
WIFA	Vážený medzisnímkový priemer (Weighted Inter-Frame Average)
WM	Vážený medián (Weighted Median)

Zoznam termínov

apnoe	prerušenie dýchania
hypertrofia	zväčšenie / nadmerný vzrast
hypoplázia	nedostatočný vývin orgánu alebo časti tela
hypopnoe	zníženie dychového objemu o viac ako 50% normálnej hodnoty trvajúce aspoň 10 sekúnd
kraniofaciálny	vyskytujúci sa na hlave a tvári
tonzila	krčná mandľa
uvula	čapík

Podákovanie

Hlavné podákovanie patrí môjmu školiteľovi doc. Ing. Dušanovi Koniarovi, PhD. za celkové vedenie práce a odborné rady pri výskume danej problematiky. Rád by som podákoval svojim spolužiakom na doktorandskom štúdiu za spoluprácu pri riešení problémov, s ktorými som sa pri výskume stretol. Prínosom boli taktiež bohaté skúsenosti kolegov z Katedry mechatroniky a elektroniky a Katedry multimédií a informačno-komunikačných technológií.

V neposlednom rade by som chcel podákovať mojej manželke, rodine a priateľom za ich prejavenú podporu pri tvorbe tejto práce.

Abstrakt

Predložená dizertačná práca prináša návrh multi-kamerového snímacieho systému na podporu diagnostiky vybraných respiračných ochorení. Vizuálny systém je založený na hĺbkových kamerách (RGB-D) s princípom merania doby letu (ToF). Práca prináša návrh typu, počtu a usporiadania obrazových senzorov, tiež návrh základných algoritmov pre kalibrovanie kamier a snímanie 3D scén. V práci je uvedená metodika na filtrovanie a opravu získaných mračien bodov (modelov) a tiež na detekciu kľúčových bodov na ľudskej hlave. Cieľom navrhovaného systému je vytvoriť nástroj na automatizovanú diagnostiku obštrukčného spánkového apnoe (OSA) v oblasti detského lekárstva.

Abstract

This dissertation thesis brings the development of multi-camera scanning system for supporting diagnostics of selected respiratory diseases. The visual system is based on depth cameras (RGB-D) working on time-of-flight (ToF) principle. The research answers the questions about the type, number or image sensors arrangement, also the basic algorithms for image sensors calibration and 3D scenes capturing. In the dissertation thesis, there is also presented the methods for filtering and restoration of point clouds (models) and detection of key points on the human head and face. The main objective of the designed system is to create a tool for automated diagnostics of obstructive sleep apnea (OSA) and apply it in the field of pediatric medicine.

1. Úvod

Syndróm obštrukčného spánkového apnoe je ochorenie, ktoré patrí medzi najčastejšie sa vyskytujúce spánkové problémy u rôznej vekovej skupiny ľudí [1]. Komplexná diagnostika tejto poruchy prebieha v špecializovaných spánkových laboratóriach pomocou polysomnografie. Takéto vyšetrenie je však časovo a finančne nákladné, čo spôsobuje dlhé čakacie doby na vyšetrenie [2].

Z tohto dôvodu sa v posledných rokoch výskum intenzívne zameriava na vývoj skríningového nástroja, ktorý by vedel klasifikovať pacientov do jednotlivých rizikových skupín. Síce ide o medicínsku problematiku, na jej riešenie sa využívajú technológie spadajúce do oblasti telekomunikácií. Z aktuálnych výskumov vyplýva, že nové diagnostické metódy využívajú k predikcii 2D snímky a 3D modely pacientov. Pre získanie týchto informácií s využívajú moderné snímacie zariadenia, ktoré poskytujú kvalitné dátá. Cena týchto zariadení je pomerne vysoká. Taktiež momentálne neexistuje komplexný systém, ktorý by sa špecializoval výlučne na detekciu obštrukčného spánkového apnoe. Jednotlivé problémy sa riešia čiastkovo a nie sú plne automatizované.

Výsledkom tejto práce je návrh cenovo dostupného snímacieho systému a vývoj príslušného algoritmu. Zariadenie pomocou RGB-D snímačov zachytáva priestorovú informáciu pacienta a umožňuje automatizované meranie špecifických kranio-faciálnych parametrov. Taktiež poskytuje možnosť vytvárania dátového setu, ktorý bude nápomocný pri vývoji klasifikačných nástrojov v neskorších fázach výskumu.

V praktickej časti práce sme navrhli metodiku snímania a filtrovania dát. Opísali sme spôsob kalibrácie multi-kamerového systému, pri ktorom sme zisťovali vnútorné a vonkajšie parametre používaných kamier. Opísali sme jav multi-kamerovej interferencie vznikajúcej pri paralelnej spolupráci ToF kamier. Navrhli sme filtračnú metódu IMBM, ktorá potláča vplyv tejto interferencie na hĺbkových mapách. Navrhli sme algoritmus umožňujúci spracovanie priatých obrazových informácií Kinect v2 v reálnom čase. Implementovali sme detektor kľúčových bodov tváre na identifikáciu pohybu objektu a pre automatizovanie merania euklidovských vzdialenosí kranio-faciálnych parametrov. Taktiež sme navrhli metodiku segmentácie tváre pomocou Mask R-CNN neurónovej siete. Tá slúži k rozdeleniu konzistentného mračna bodov vytvoreného z viacerých kamier do klasifikačných skupín (oko, ucho, tvár, ...). Takéto segmentovanie umožní vytvorenie jedinečnej databázy, ktorej dáta budú použité pri trénovaní OSA klasifikátora. Pre tento účel bol vytvorený samostatný snímací systém, ktorého cieľom je získavanie dát určených k trénovaniu segmentačnej siete. Ten je umiestnený v detskom spánkovom laboratóriu na klinike detí a dorastu JLF UK v Martine.

2. Medicínska podstata problému

Spánkové poruchy postihujú približne 30% svetovej populácie. Podľa súčasnej platnej medzinárodnej klasifikácie spánkových porúch sú respiračné spánkové poruchy druhým najčastejšie sa vyskytujúcim ochorením [3].

Respiračné spánkové apnoické ochorenia delíme na:

- **obštrukčné (OSAS):** dospelí pacienti, pediatrickí pacienti
- **centrálne (CSA):** Cheine-Stokes dýchanie, Primárne, ...
- **zmiešané:** kombinované CSA a OSAS

Syndróm obštrukčného spánkového apnoe (OSAS) je respiračná choroba, ktorá spôsobuje prerušenie oronazálnej ventilácie pri pretrvávajúcom dychovom úsilí na aspoň 10 sekúnd a opakuje sa viac ako 5 krát za hodinu [4]. Na jej vzniku sa vo významnej miere podieľa znížená priechodnosť horných dýchacích ciest a zmena mechanizmu mäkkých štruktúr horných dýchacích ciest. Najvýznamnejším rizikovo-vým faktorom vzniku OSAS je anatomické zúženie horných dýchacích ciest spôsobené predovšetkým: obezitou, kraniofaciálnymi abnormalitami, hypopláziou tváre, zväčšeným objemom mäkkých tkanív, hypertrofiou tonzíl, zväčšením uvuly a dĺžky mäkkého podnebia [5].

K presnému stanoveniu OSAS diagnózy sa vykonáva špeciálne komplexné vyšetrenie nazývané polysomnografia (PSG) [6]. Ide o celonočné monitorovanie pacienta pomocou snímačov a senzorov, ktoré dávajú lekárovi informáciu o dýchaní, ronchopátii, činnosti srdca, charaktere spánku a jeho jednotlivých cykloch, o polohe tela a o okysličovaní organizmu. PSG sa vykonáva v akreditovanom spánkovom laboratóriu, ktoré musí byť vybavené potrebnou meracou technikou. Z výsledkov sa následne určuje konkrétny spôsob liečby [7].

PSG je považovaná za štandard pre diagnostiku OSAS, avšak toto vyšetrenie je finančne nákladné a časovo náročné [8]. Navyše na Slovensku ale aj vo svete je problém s nedostatkom spánkových laboratórií, čo spôsobuje dlhé čakacie doby.

Kvôli zníženiu potreby používania PSG adolescentnou skupinou pacientov bol vytvorený validovaný dotazník nazývaný Klinický záznam o spánku (SCR). Ten pomocou kombinácie informácií z klinického vyšetrenia, skúmania subjektívnych symptomov pacienta a skúmania prítomnosti ADHD dokáže s určitou presnosťou identifikovať prítomnosť OSAS. Pozitívni pacienti sú následne uprednostňovaní na vyšetrenie PSG [9].

Vyšetrenie pomocou kontaktných meracích prístrojov je pre pediatrických pacientov častokrát stresujúce. Deti sú nepokojné a nedokážu spolupracovať s lekárom, čo predlžuje vyšetrenie a vytvára nepresnosť pri meraní parametrov tváre. Práve z tohto dôvodu vznikla požiadavka na vytvorenie meracieho systému, ktorý by dokázal zmerať požadované parametre bezkontaktne.

Jedným z možných riešení je vytvorenie skenovacieho zariadenia, ktoré dokáže s určitou precíznosťou geometricky aj textúrovo zachytiť pacienta. Takto zachytený model môže slúžiť k manuálnej diagnostike bez potreby fyzického kontaktu s pacientom alebo ako vstupné dátá pre algoritmy strojového učenia.

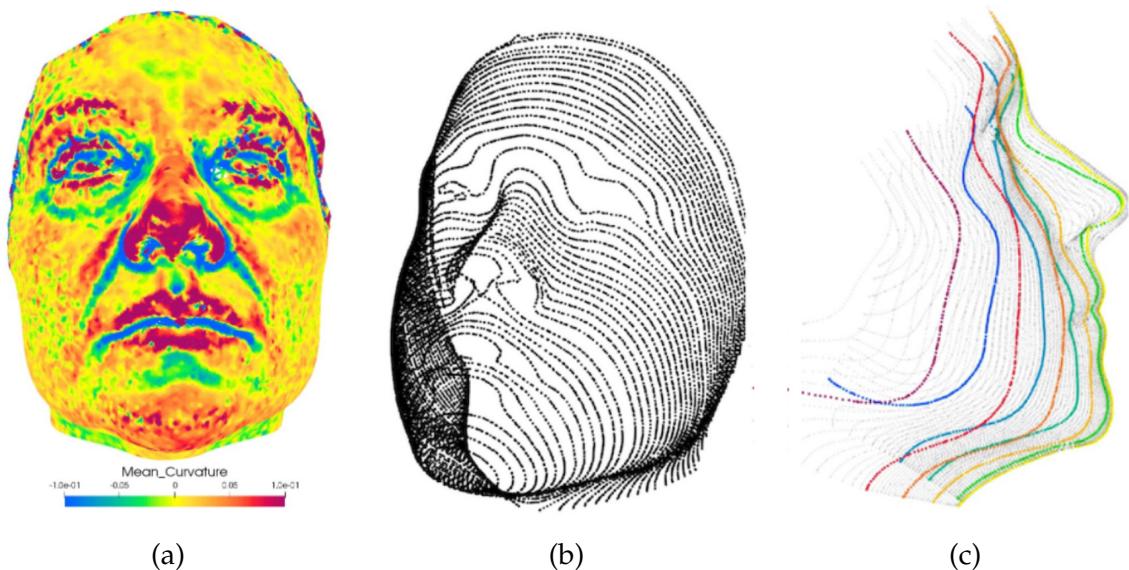
2.1 Aktuálny stav výskumu diagnostiky OSAS s použitím 3D dát

Z dôvodu finančnej a časovej nákladnosti diagnostiky OSAS sa kladie vyšší záujem o vytvorenie skríningového systému, ktorý by vedel identifikovať rizikových pacientov. Aktuálny výskum sa zameriava na predikciu ochorenia z obrazovej informácie. V minulosti a súčasnosti sa diagnostika vykonáva na CT a MRI snímkach, ktorých získanie je komplikované a limitované [10]. V posledných rokoch sa pre diagnostiku začalo využívať 3D spracovanie dát. Ich snahou je vytvoriť klasifikačný nástroj, ktorý umožní oddeliť rizikových pacientov od ostatných.

2.1.1 Diagnostika pediatrickej OSA prostredníctvom klasifikácie tváre s použitím CNN

V článku *Diagnosis of Pediatric Obstructive Sleep Apnea via Face Classification with Persistent Homology and Convolutional Neural Networks* [11] je opísaná štúdia, v ktorej sa snažia vytvoriť klasifikátor pre rozdelenie vysoko rizikových pacientov od tých menej rizikových. K analýze využívali 3D modely a 2D fotografie obsahujúce snímky týchto 3D objektov. Podľa súčasných štúdií, na ktoré sa v práci odkazujú, využili pre klasifikáciu perzistentnú homológiu (PH) [12] a konvolučné neurónové siete (CNN) [13, 14]. Pri prvom prístupe analyzujú trojuholníkové siete cez geometrické príznaky tváre, pri CNN sa pre trénovanie používajú klasifikované 2D obrazy tváre. Použitý dátový set obsahoval 172 detských pacientov. Objekty boli rozdelené lekárom do skupín podľa toho, či patrili do vyšej alebo nižšej rizikovej skupiny OSA. Každý 3D model podstúpil filtráciu (vyplnenie dier v sieti, vyhladenie), pri ktorej sa ale dbalo na to, aby nebola znehodnotená geometrická informácia.

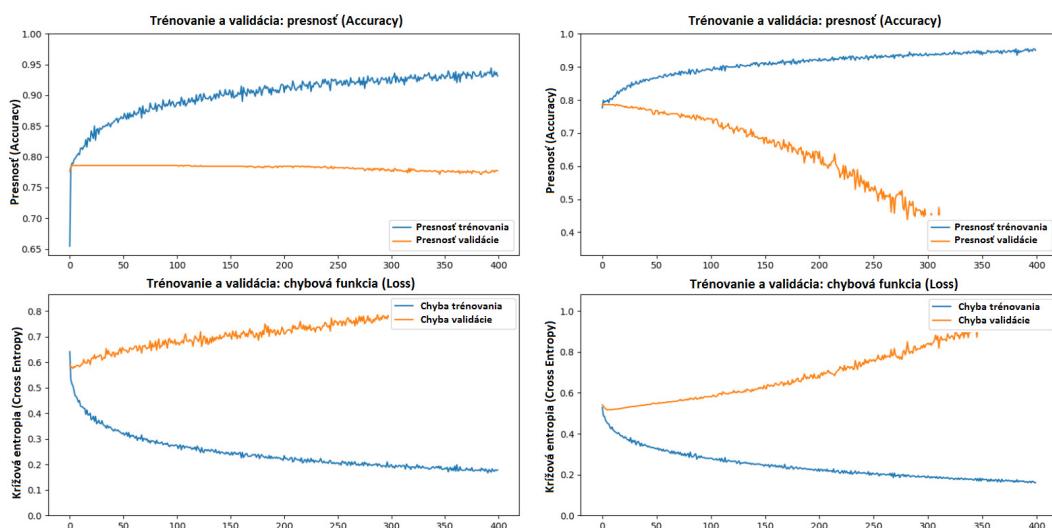
Z výsledkov PH vyplýva, že súčasné topologické analýzy dát neviedli k presnej klasifikácii. Niektoré rysy tváre súce môžu súvisieť s vysokým rizikom OSA, ale neboli postačujúce pre identifikáciu rizikových pacientov.



Obr. 2.1: Vstupné dáta pre PH klasifikáciu: (a) Priemerné hodnoty zakrivenia pre každý bod v mračne bodov. (b) Krivky tváre (c) Geodetika medzi dvoma pacientmi [11].

Pre použitie CNN metódy boli z filtrovaných modelov vytvorené séria snímok, ktoré zachytávali objekty z rôznych perspektív. Je to kvôli tomu, aby konvolučná neurónová sieť dokázala lepšie pochopiť povrch tváre. Následne boli dátá rozdelené v pomere 80% pre trénovanie a 20% pre testovanie. V článku sa používali CNN architektúry: *InceptionV3*, *InceptionResV2*, *MobileNetV2* a *NASNetLarge*. Tie boli predtrénované na ImageNet datasete.

Výsledky trénovania pre *MobileNetV2* a *NASNetLarge* sú zobrazené na obr. 2.2. Z nich vyplýva, že validácia výstupného modelu nedosahovala očakávané výsledky.



Obr. 2.2: Grafy iteračného učenia (školenia a validácie) pre *MobileNetV2* (vľavo) a *NasNetLarge* (vpravo). Spodný riadok ukazuje nákladovosť krízovej entropie a horný presnosť klasifikácie [11].

Tabuľka 2.1: Výsledky validácie predtrénovaných modelov pre použité CNN architektúry [11].

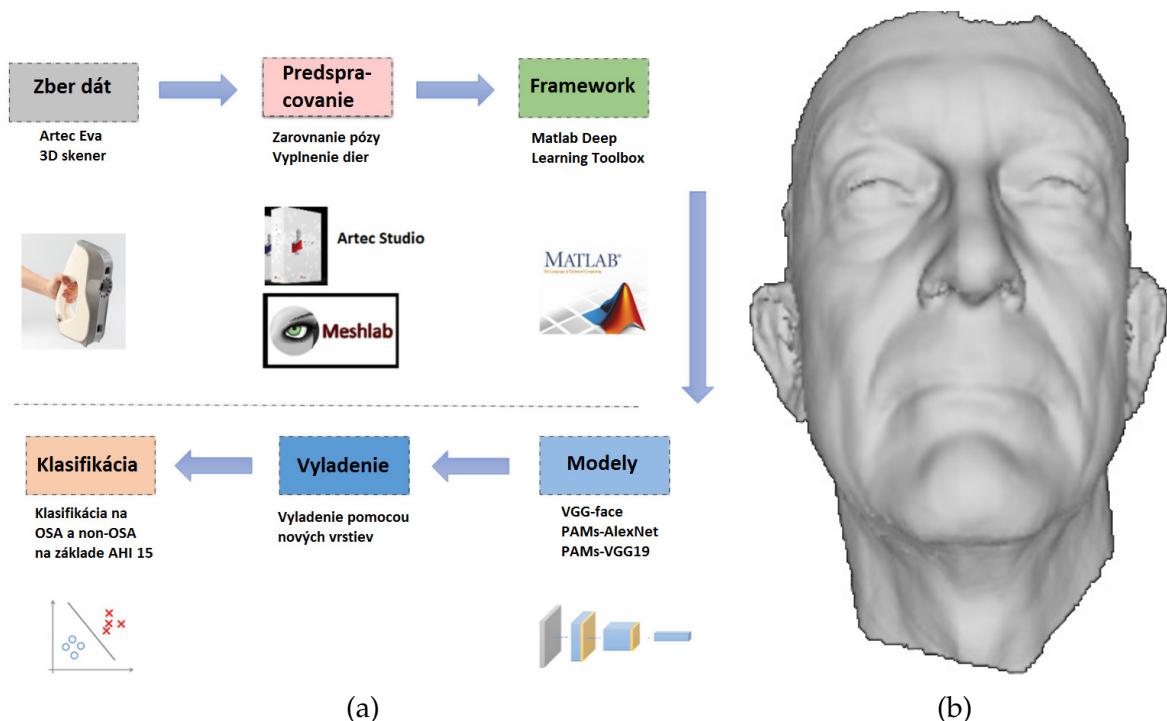
Architektúra	Trénované s	Knižnica	Validácia
InceptionV3	ImageNet	TensorFlow	70 ± 4
InceptionResV2	ImageNet	TensorFlow	71 ± 5
MobileNetV2	ImageNet	TensorFlow	64 ± 5
NASNetLarge	ImageNet	TensorFlow	72 ± 2

Možným dôvodom je, že dátové sety pediatrických OSA pacientov obsahujú malé množstvo klasifikovaných objektov. Ak by sa v budúcnosti vytvorila databáza, ktorá bude zhromažďovať potrebné dátá, je veľký predpoklad na zlepšenie presnosti klasifikovania OSA pacientov pomocou CNN.

2.1.2 Predikcia OSA pomocou hlbokého učenia s použitím hlbkovej mapy tváre

V článku *Deep Learning of Facial Depth Maps for Obstructive Sleep Apnea Prediction* autori používajú pre detekciu OSA hlbkové mapy a konvolučné neurónové siete [15]. V teoretickom úvode sa odkazujú na súčasné trendy pri diagnostike OSA. Uvádzajú, že nové diagnostické metódy smerujú k využívaniu obrazovej informácie hlavy a krku pacienta [16, 17, 18]. V ich práci sa zamerali na využitie hlbkových máp, ktoré poskytovali informáciu o geometrických parametroch tváre a krku.

3D dátá boli získané od pacientov, ktorí podstúpili PSG diagnostiku. Databáza bola tvorená z 39 mužských a 30 ženských skenov, ktoré boli vytvorené snímačom Artec Eva [19]. Tie boli následne filtrované, čím sa odstránili neželané artefakty (obr. 2.3b). Z takto upravených modelov boli následne vytvorené 2D hlbkové mapy (obr. 2.4a).

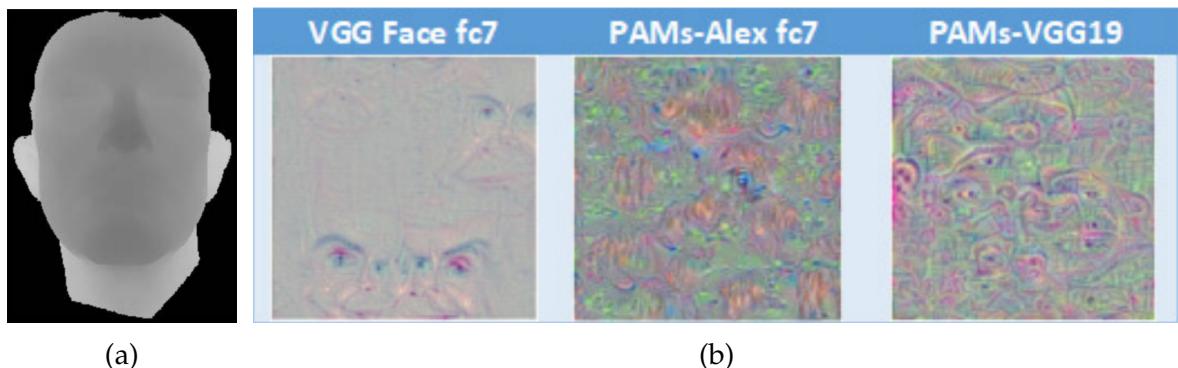


Obr. 2.3: (a) Blokový diagram prezentovaného postupu detektie OSA cez CNN. (b) Ukážka 3D objektu po filtrácii. Tento model bol získaný pomocou laserového ske- nu Artec Eva [15].

K trénovali si vybrali tri rozdielne neuronové siete, ktoré už boli predtréno- vané pre detekciu a rozpoznanie tváre. Neboli však trénovalé na hlbkových mapách.

VGG-Face bol trénovaný na 2.6 miliónovom datasete s presnosťou 98.95% [20]. Taktiež boli vybrané *Pose-Aware CNN* modely [21] *AlexNet* [22] a *VGG-19* [23]. Jemné doloďovanie modelov sa vykonávalo v *Matlab deep learning toolbox*. Z databázy hĺbkových máp bolo 70% použitých pre trénovanie (14 modelov na testovanie) a 30% pre validáciu. Celý blokový diagram je znázornený na obr. 2.3a.

Výsledné modely výstupnej vrstvy po natrénovaní siete s rôznymi vstupnými parametrami sú zobrazené na obr. 2.4b. Pri PAMs je možné vidieť rôznu aktiváciu očí a nosa na rôznych miestach modelu, pretože model sa musel vysporiadať s rozličnými pôzami a natočeniami. Výstupom *VGG Face* vrstvy je frontálny obraz tváre bez rotácie.



Obr. 2.4: (a) Hĺbková mapa získaná z 3D modelu tváre na obr. 2.3b. (b) Vlastnosti vrstvy siete [?]

Výsledné porovnanie sietí sa nachádza v tabuľke 2.2. Z výsledkov vyplýva, že *VGG-Face* dosahovala lepšie výsledky ako *PAMs-VGG19* a *PAMs-Alex*. Jej výhoda je, že bola predtrénovaná na veľkom datasate.

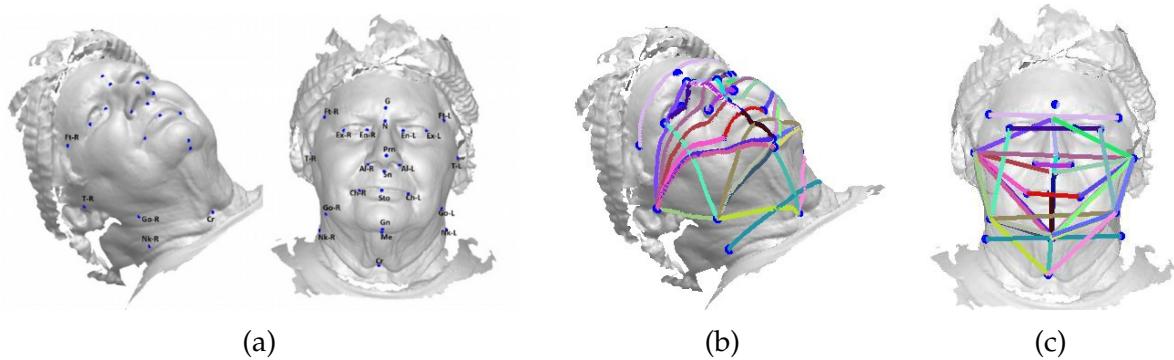
Tabuľka 2.2: Presnosť klasifikácie použitých modelov.

	VGG Face	PAMs-VGG19	PAMs-Alex
Presnosť validácie	68,75	62,5	60,15
Presnosť testovania	67,42	57,14	59,37

Táto práca predstavuje ako prvá detekciu OSA založenú na hĺbkových mapách. Presnosť klasifikácie, ktorá bola dosiahnutá, je porovnateľná s výsledkami iných prác zaobrajúcich sa detekciou OSA [24]. Autori článku uvádzajú, že problémom pri trénovaní je malý dátový set hĺbkových máp pacientov. Plánujú jeho rozšírenie a automatizáciu vybraných procesov.

2.1.3 Predikcia OSA pomocou 3D modelu

V práci *Predicting Sleep Apnea From Three-Dimensional Face Photography* sa zaoberejú predikciou OSA pomocou lineárnych a geodetických vzdialenosí medzi kľúčovými bodmi na tvári [25]. Cieľom autorov je nájsť čo najvhodnejšie kombinácie meraní, pre čo najlepšiu predikciu stupňa ochorenia OSA. Trojrozmerné snímky prekonávajú limity 2D snímania a pomáhajú určiť štruktúru tkaniva oveľa lepšie ako z 2D fotografie [26]. Dátový set bol vytvorený zo 400 3D modelov dospelých pacientov, ktorí podstúpili PSG vyšetrenie. Podľa AHI indexu boli pacienti rozdelení do 4 skupín v závislosti od štadia a závažnosti ochorenia. Taktiež bol vypočítaný BMI index a odmeraný obvod krku. 3D modely boli vytvorené kamerou značky 3dMD, ktorá meria vzdialenosť na princípe triangulácie [27, 28]. Pacienti boli pri snímaní naklonený dozadu, aby bola lepšie viditeľná hrdlová časť. Merania lineárnych vzdialenosí, uhlov a geodetických vzdialenosí boli realizované pomocou anatomických značiek z 3D snímok, z ktorých sa následne určila prítomnosť ochorenia OSA a jeho štadium. Pre detekciu OSA bolo zvolených 24 kľúčových bodov, ktoré boli identifikované v každom 3D modeli (obr. 2.5a).



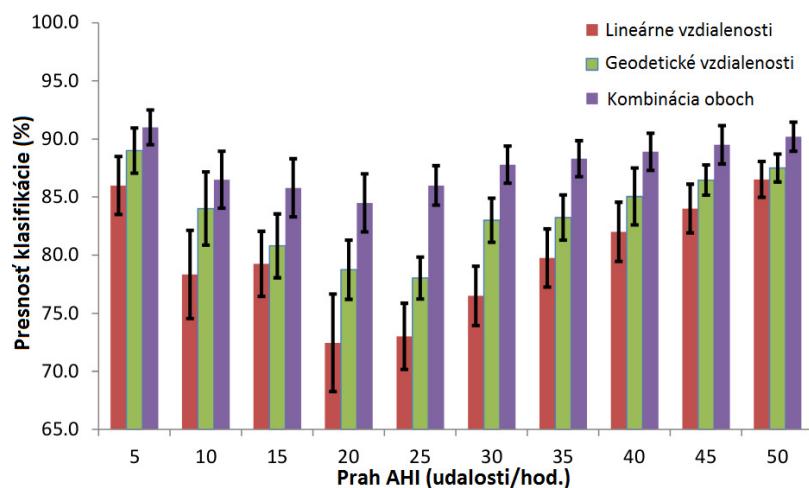
Obr. 2.5: (a) Anotované kľúčové body tváre a krku. (b) Geodetické vzdialenosť medzi vybranými párami bodov. (c) Euklidovské vzdialenosť medzi vybranými párami bodov [25].

Pre detekciu OSA bolo zvolených 24 kľúčových bodov, ktoré boli identifikované v každom 3D modeli [29, 30]. Dva z týchto bodov boli vyznačené na tvári pacienta ešte pred skenovaním. Identifikovaním týchto bodov sa získali súradnice (x, y, z), ktoré boli použité pre ďalšiu analýzu. Vďaka týmto súradniciam je možné získať až 276 rôznych vzdialenosí medzi dvojicou kľúčových bodov. Na základe predchádz-

júcich štúdií bolo vybraných 25 párov bodov, medzi ktorými boli vypočítané euklidovské a geodetické vzdialenosť. Tieto získané informácie boli použité ako príznaky v ďalšej analýze.

Na základe prahu $AHI \geq 5$ boli vstupné dátá rozdelené na kontrolné (100) a OSA (300). Pomocou tejto hodnoty bol navrhnutý a natrénovaný *Linear Discriminant Analysis* (LDA) algoritmus, ktorý využíval geodetickú vzdialenosť, euklidovskú vzdialenosť a uhol každej snímky. Úlohou algoritmu bolo klasifikovať nové prípady podľa prahu AHI s hodnotami 5, 10, 15, 20, 25, 30, 35, 40, 45 a 50.

Cieľom LDA algoritmu bolo pomocou všetkých popísaných príznakov nájsť 1D priestor, kde vzdialenosť medzi príznakmi v rámci rovnakej triedy bola minimálna a zároveň vzdialenosť medzi príznakmi dvoch odlišných tried maximálna [31]. Dátá boli týmto spôsobom rozdelené do desiatich skupín. Z každej skupiny bolo 90% dát použitých na trénovanie a zvyšných 10% pre testovanie.



Obr. 2.6: Presnosť klasifikácie pre predpovedanie OSA, keď bol algoritmus trénovaný a testovaný pre rozdielne prahové hodnoty AHI (5 až 50) s použitím euklidovskej vzdialenosť (červený stĺpec), geodetickej vzdialenosť (zelený stĺpec) a kombináciu (fialový stĺpec).

Štúdia ukazuje, že geodetická informácia výrazne zlepšuje schopnosť identifikovať rizikových pacientov. Najväčšia presnosť (91%) klasifikácie bola získaná pri kombinácii geodetickej a euklidovskej vzdialenosť. Výsledné dátá tejto štúdie nasvedčujú tomu, že 3D skenovanie má potenciálnu úlohu pri skríningu OSA, avšak nedokáže nahradiť PSG vyšetrenie. Z výsledkov taktiež vyplýva, že pomocou geodetickej vzdialenosť je možné určiť aj stupeň závažnosti OSA.

2.2 Požiadavky na systém

Z predchádzajúcej časti vyplýva, že hĺbková informácia môže byť užitočná pri predikcii syndrómu obštrukčného spánkového apnoe (OSAS). Aktuálny výskum sa uberá spracovaním priestorovej informácie, pričom sa využívajú technológie spracovania obrazu a signálov. Z aktuálnych poznatkov nie je úplne zrejme, ktorá metodika je najvhodnejšia pre diagnostiku. Aktuálny problém je aj s nedostatkom datasetu OSA pacientov, ktorý by obsahoval 3D informácie. Dôvodom je, že skenovacie zariadenia sú pomerne drahé a doposiaľ neexistuje plne automatizovaný systém zberu dát. Spánkové laboratóriá sa z týchto dôvodov nezapájajú do spoločného výskumu, ktorý by pomohol k riešeniu celosvetového problému.

Z tohto dôvodu vznikla požiadavka vytvoriť cenovo dostupné zariadenie, ktoré by umožňovalo zachytiť priestorové informácie pacientov. Dôležitým je zosnímanie tváre a krku, pričom sa kladie požiadavka na presnosť. Tá zatiaľ nie je presne špecifikovaná, taktiež nie sú pevne špecifikované príznaky. Dôležité je ale minimalizovanie chyby vzniknutej pohybom pri skenovaní. V budúcnosti môže byť špecifikované skenovať tvár pacienta v rôznych polohách a mimikách.

Pri príprave dát sa vykonávalo veľa manuálnych úkonov. Väčšinou išlo o zlepšenie kvality modelov odstraňovaním artefaktov v 3D modeloch. Taktiež sa v niektorých prácach vykonávala manuálna identifikácia kľúčových bodov tváre. Skenovací systém by mal do určitej miery umožňovať automatizáciu procesov.

3. Technické pozadie výskumu

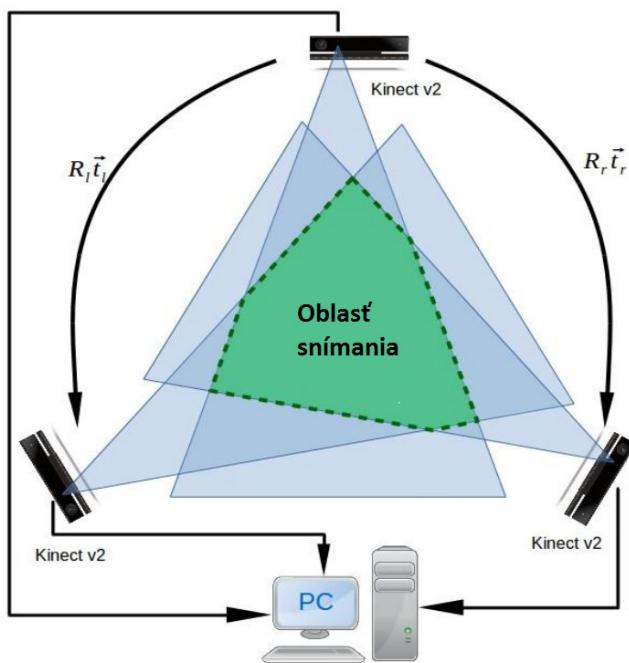
Z kapitoly 2 vyplýva, že v posledných rokoch medicínsky výskum využíva priestorové videnie a 3D informáciu na predikciu OSA. K zachyteniu 3D priestoru sa využívajú hĺbkové snímače. Tie dokážu transformovať 3D priestor do 2D obrazu [32]. Táto informácia je uchovávaná v matici, ktorá je nazývaná hĺbková mapa. Spôsob vytvárania hĺbkovej mapy kamerou závisí od použitej technológie. Medzi základné spôsoby patrí stereovízia, snímanie štrukturovaným svetlom (SLS) a meranie doby letu (ToF).

3.1 Skenovanie v reálnom čase pomocou RGB-D senzorov

Skenovanie dynamických objektov v reálnom čase je v centre záujmu už niekoľko rokov. V posledných rokoch sa hĺbkové snímače stávajú súčasťou riešenia rôznych technických problémov. Existuje viacero spôsobov, ktoré dokážu kvalitne rekonštruovať objekty v reálnom čase. Väčšinou sú to komplexné riešenia, ktoré sú cenovo nedostupné pre bežných užívateľov. Existuje však aj niekoľko cenovo dostupných RGB-D senzorov, ktoré môžu byť použité na 3D snímanie a rekonštrukciu v reálnom čase. V ďalšej časti sú opísané aktuálne riešené problematiky 3D skenovania dynamických objektov v reálnom čase.

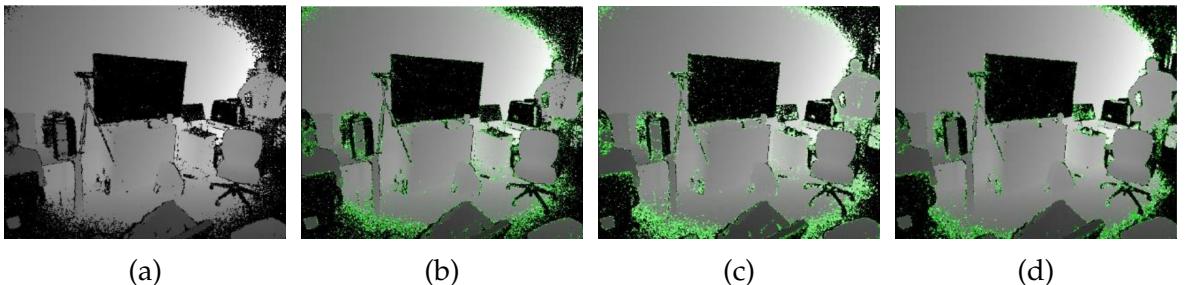
3.1.1 Viac-pohľadové 3D snímanie a analýza pre rekonštrukciu vysoko kvalitných mračien bodov

V článku *Multiview 3D Sensing and Analysis for High Quality Point Cloud Reconstruction* sa autori zaoberajú rekonštrukciou mračien bodov pomocou viacerých ToF kamier [33]. V prvom kroku riešili kalibráciu, ktorá je kritickým krokom v celom procese. Využili metódu kalibrácie pomocou planárneho kalibračného vzoru (šachovnice), kde identifikovali rotačné a translačné parametre pre jednotlivé kamery [34].



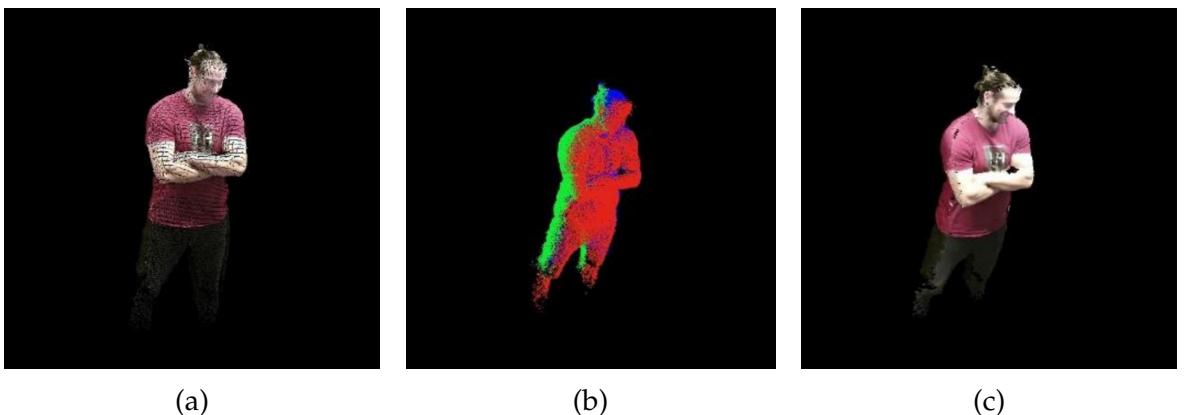
Obr. 3.1: Rekonštrukčný systém pozostávajúci z troch senzorov Kinect v2.

V práci používali trojicu kamier Kinect v2 pripojených k jednému PC čo je rozdielny prístup oproti opísaným v [35, 36]. Pre odstránenie multi-kamerovej interfeference kamery pracovali v sekvenčnom režime. Kvalita hĺbkovej mapy je kľúčovým faktorom pri tvorbe a rekonštrukcii modelu. Ak sa v nej vyskytnú chyby, tie sa potom prenášajú do mračna bodov. Túto hĺbkovú mapu filtrovali použitím bilaterálneho filtra (BF) [37], Weight Median filtra (WM) a Radius Outlier Removal filtra (ROR). Vplyv filtračných metód na hĺbkovú mapu je zobrazený na obr. 3.2. Okrem toho tiež použili Weighted Inter-Frame Average filter (WIFA), ktorý pracuje so sériou hĺbkových máp. Ten odstraňuje priestorový šum, ale výrazný pohyb robí aplikáciu WIFA nevhodnou pre vysoko dynamické scény [38].



Obr. 3.2: Výsledky filtrácie hĺbkovej mapy, kde zelená farba predstavuje rozdiely medzi prístupmi filtrácie: (a) Vstupná hĺbková mapa. (b) Filtrácia pomocou BF. (c) Filtrácia pomocou BF + WM. (d) Filtrácia pomocou BF + WM + ROR [33].

Mračná bodov zo všetkých Kinect v2 sa následne spájajú do jedného mračna, čím sa vytvorí jednotná uniformná reprezentácia snímanej scény. Algoritmus rekonštrukcie sa používa na vytvorenie vernej a podrobnej reprezentácie povrchu skenovaných tvarov. V ďalšej fáze sa generuje trojuholníková sieť zachyteného 3D objektu.



Obr. 3.3: Rekonštrukcia scény z troch snímačov Kinect v2: (a) Mrak fúzovaných bodov generovaný každým snímačom Kinect. (b) Farebné oddelenie fúzovaných mračien bodov pre každú kameru. (c) Trojuholníková sieť vytvorená zo spojených mračien bodov [33].

V článku sa zamerali aj na meranie rýchlosťi spracovania filtrácie. Pre prácu použili CUDA paralelné spracovanie na grafickej karte GeForce 780GTX. Meranie rýchlosťi spracovania jednotlivých častí algoritmu sa vykonávalo na 400 hĺbkových mapách. Priemerný čas filtrácie so sieťovou rekonštrukciou trval 14,19ms.

Autori predstavili návrh multi-kamerovej spolupráce 3 ToF snímačov značky Kinect v2. Bol opísaný spôsob kalibrácie a registrácie mračien bodov z jednotlivých kamer. Zároveň sa zamerali aj na filtráciu a trojuholníkovú rekonštrukciu snímaných objektov, pričom testovali rýchlosť spracovania.

3.2 Disparitná a hĺbková mapa

Výstupný obraz zachytávajúci vzdialenosť je závislý od použitej technológie snímania. Výstupom kamery je buď disparitná alebo hĺbková mapa [39]. Disparita zachytáva relatívnu vzdialenosť navzájom si odpovedajúcich bodov v stereo-páre obrazu. Pri vytváraní disparity sa predpokladá s nulovou vertikálnou paralaxou. K výpočtu je teda potrebné mať dvojicu obrazov v epipolárnej rovine. Ak bod x_1 v ľavom obraze na pozícii [20,0] odpovedá v pravom obraze bodu x_2 [10,0], hodnota disparity D má hodnotu 10. Pri výpočte sa jeden obraz zo stereo-páru berie ako referenčný.

$$D = x_1 - x_2 \quad (3.1)$$

V disparitnej mape každý pixel nesie informáciu o dispareite. Zobrazením vzniká obraz v odtieňoch sivej, ktorý vytvára ucelenú informáciu o priestore. Disparitná mapa je vždy vytváraná pre jeden obraz z dvojice, ktorej pozícia bodu sa berie ako referencia. Takáto mapa je už pre aplikáciu v stereovízií užitočná, pretože je z nej možné odlišiť rozloženie snímanej scény. Neposkytuje však informáciu o reálnej vzdialnosti.

Hĺbková mapa je taktiež šedo-tónový obraz, ktorý zachytáva informáciu o abso-lútnej vzdialenosťi Z snímanej scény od kamery.



Obr. 3.4: Ukážka hĺbkovej mapy povrchu tváre [40].

$$Z = \frac{bf}{x_1 - x_2} = \frac{bf}{D} \quad (3.2)$$

Pre výpočet je potrebné poznať rozostup medzi šošovkami kamier b a ohniskovú vzdialenosť f .

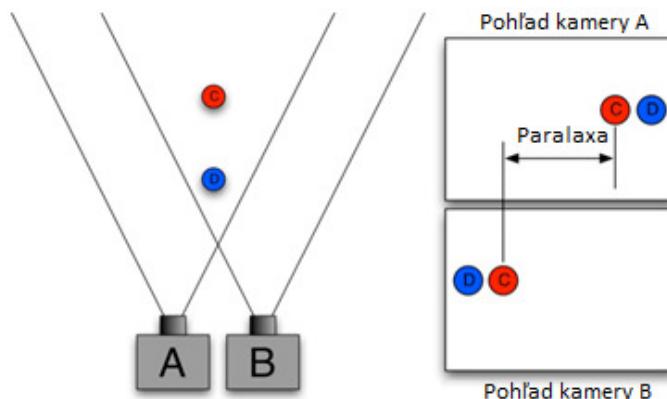
3.3 Princíp činnosti RGB-D kamier

RGB-D kamery sú optické snímače, ktoré zachytávajú hĺbkovú informáciu scény. Existuje viacero metód, ktorými tieto zariadenia pracujú. V tejto časti práce sú opísané najčastejšie používané metódy pre meranie hĺbky.

- **stereo-vízia:** ZED, Intel RealSense d415, stereo RGB
- **snímanie štrukturovaným svetlom (SLS):** Intel RealSense SR300
- **meranie času doby letu (ToF):** Microsoft Kinect v2

3.3.1 Stereovízia

Metóda stereovízie je založená na synchronizovanom snímaní scény pomocou viacerých párov kamier (RGB alebo IR), ktoré sú voči sebe vzájomne posunuté. Najčastejšie rozloženia kamier sú Toe-In a Off-Axis, ktorá je znázornená na obr. 3.5. Výstupom je farebný obraz scény zosnímanej z viacerých perspektív.

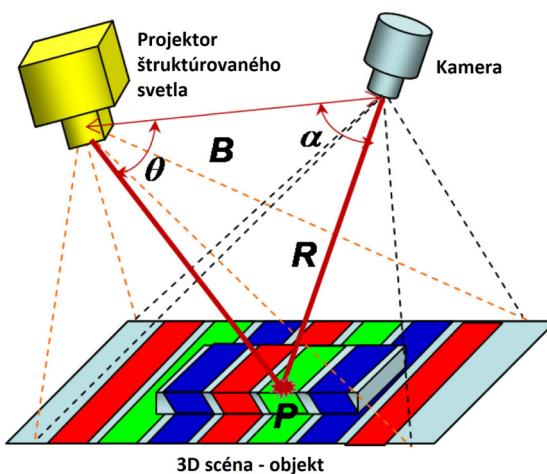


Obr. 3.5: Princíp stereovízie s Off-Axis nastavením kamier [41].

Na základe rozdielu z týchto dvoch perspektív a známych vonkajších parametrov stereo-kamerovej sústavy je možné dopočítať hĺbkový obraz [42]. Hlavnou nevýhodou je vysoká výpočtová náročnosť algoritmu a cena kvalitných snímacích RGB senzorov. Taktiež je tento systém náchylný na meniace sa svetelné podmienky. Výhodou je nulová vzájomná interferencia.

3.3.2 SLS snímače

Structured Light Sensors (SLS) snímače sú zložené z projektora štrukturovaného svetla a snímača [43]. Projektor aktívne osvecuje scénu so špeciálne navrhnutým 2D vzorom, častokrát priestorovo modulovaným svetlom. Kamera následne sníma osvetlenú scénu a získané dáta porovnáva s projektovaným vzorom. Ak je scéna plánárna, snímaný vzor sa zhoduje s referenčným vzorom. Ak sú však v scéne povrchové variácie, geometrický tvar povrchu narúša projektované štrukturované svetlo. To sa následne nebude zhodovať s projektovaným vzorom.



Obr. 3.6: Princíp SLS kamery [43].

Na obr. 3.6 je znázormený geometrický vzťah medzi projektorom, kamerou a snímaným povrchom. Tento vzťah je možné vyjadriť triangulačným princípom:

$$R = B \frac{\sin \theta}{\sin \alpha + \theta} \quad (3.3)$$

Kľúčom k 3D zobrazovaniu na báze triangulačnej techniky je správne priradiť zosnímaný bod k projekčnému bodu [43]. Na tento účel boli navrhnuté rôzne schémy, ktoré sa delia na:

- **Sekvenčnú projekciu:** binárny kód, šedý kód, fázový posun, hybrid
- **Priebežne meniacu projekciu:** dúhový kód, priebežne meniaci farebný kód
- **Pásikový index:** farebne kódované pásy, segmentované pásy, De Bruijn, ...
- **Mriežkovaný index:** pseudo-náhodné binárne body, mini-vzor ako kód, ...
- **Hybridné metódy**

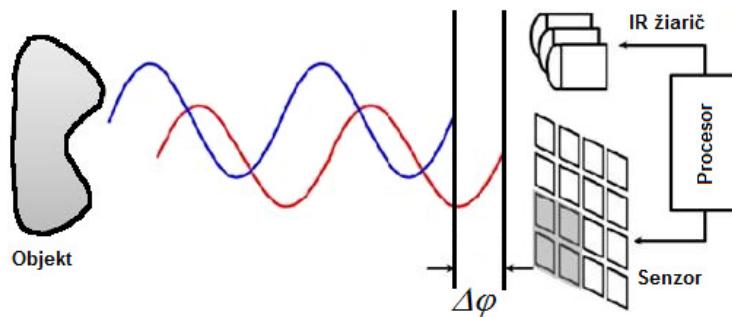
Hlavnou výhodou štruktúrovaného svetla je dosiahnutie vysokého priestorového rozlíšenia. Kamery pracujúce na tomto princípe nevyžadujú žiadnu špeciálnu úpravu na úrovni snímača. Akékoľvek rušenie je znázornené ako variácie v povrchu. Táto metóda snímania so sebou prináša problém multi-kamerovej interferencie. Tento fakt znemožňuje použitie týchto kamier v paralelnej spolupráci.

3.3.3 ToF snímače

Technológia *Time of Flight* priniesla revolúciu v 3D strojovom videní, pretože umožňuje meranie hĺbky pomocou lacného CMOS čipu a aktívneho modulovaného svetla [44]. Kamera obsahuje zdroj IR svetla, ktorý je reprezentovaný polovodičovým laserom alebo LED s vlnovou dĺžkou $\sim 850\text{nm}$. Vyžiarené svetlo sa odráža od snímaných objektov a späť dopadá do prijímača IR svetla. Pre získanie vzdialenosť d sa zisťuje fázový posun φ medzi vyžiareným a prijatým signálom. Parameter c predstavuje rýchlosť šírenia elektromagnetickej vlny v priestore.

$$d = \frac{c}{2} \frac{\Delta\varphi}{2\pi f} \quad (3.4)$$

Vyžiarené svetlo sa moduluje pulzne alebo kontinuálnou vlnou (continuous wave). Pulzná modulácia je bežnejšia kvôli ľahšej realizácii pomocou elektronických obvodov [45]. Na obr. 3.7 sa nachádza princíp detekcie hĺbky použitím ToF technológie so sínusovou moduláciou IR svetla [46].

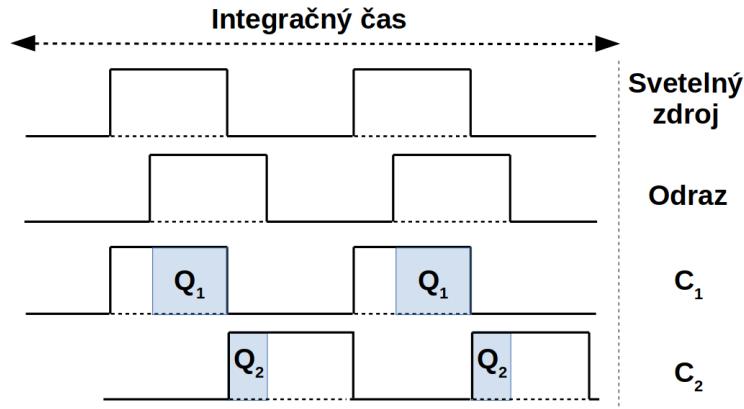


Obr. 3.7: Princíp činnosti ToF senzorov [45].

Pri pulznej modulácii sa zdroj svetla rozsvecuje na dobu Δt . Odrazená energia sa paralelne vzorkuje pomocou dvoch vzorkovacích okien C_1 a C_2 , ktoré sú fázovo

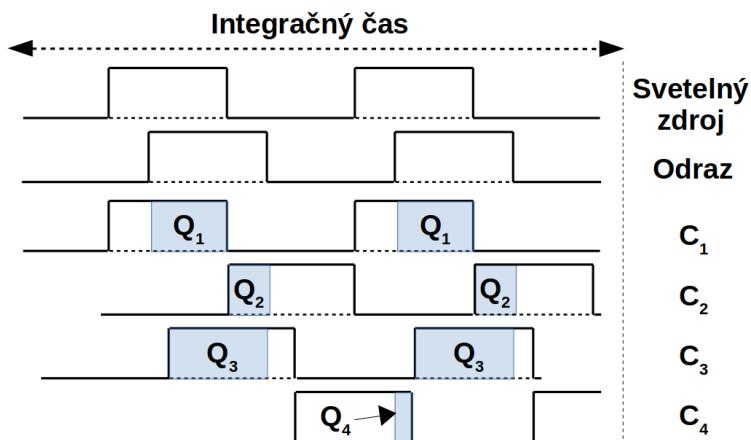
posunuté o 180° . Elektrické náboje Q_1 a Q_2 , akumulované počas doby vzorkovania, sa používajú pri výpočte vzdialenosťi d pomocou vzorca 3.5. Časový diagram pulznej modulácie sa nachádza na obr. 3.8.

$$d = \frac{c}{2} \cdot \Delta t \left(\frac{Q_2}{Q_1 + Q_2} \right) \quad (3.5)$$



Obr. 3.8: Pulzná modulácia ToF senzora [44].

Metóda s použitím kontinuálnej vlny (CW) využíva 4 vzorkovacie okná C_1 až C_4 posunuté voči sebe o 90° . Počas doby vzorkovania sa akumulujú náboje Q . Použitím tejto techniky je možné vypočítať fázový uhol φ medzi priatým a vyžiareným svetlom. Ten sa následne používa pre výpočet hĺbky.



Obr. 3.9: Modulácia ToF senzora pomocou kontinuálnej vlny [44].

$$\varphi = \arctan \left(\frac{Q_3 - Q_4}{Q_1 - Q_2} \right) \quad (3.6)$$

$$d = \frac{c}{4\pi f} \varphi \quad (3.7)$$

Pri CW modulácii sa z nameraných nábojov Q_1 až Q_4 vypočítava intenzita pixelu A a offset B .

$$A = \frac{\sqrt{(Q_1 - Q_2)^2 + (Q_3 - Q_4)^2}}{2} \quad (3.8)$$

$$B = \frac{Q_1 + Q_2 + Q_3 + Q_4}{4} \quad (3.9)$$

CW modulácia zabezpečuje zníženie chyby merania spôsobenej zmenou vnútorných parametrov elektronických komponentov. Napríklad zmena teploty kamery môže mať dopad na výsledné zosilnenie prijatých elektrických nábojov Q . V konečnom dôsledku to pri pulznej modulácii spôsobí nepresný výpočet hĺbky d v rovnici 3.5. Pomocou parametrov A a B je možné aproximovať rozptyl hĺbky σ [44].

$$\sigma = \frac{c}{4\sqrt{2\pi f}} \frac{\sqrt{A + B}}{c_d A} \quad (3.10)$$

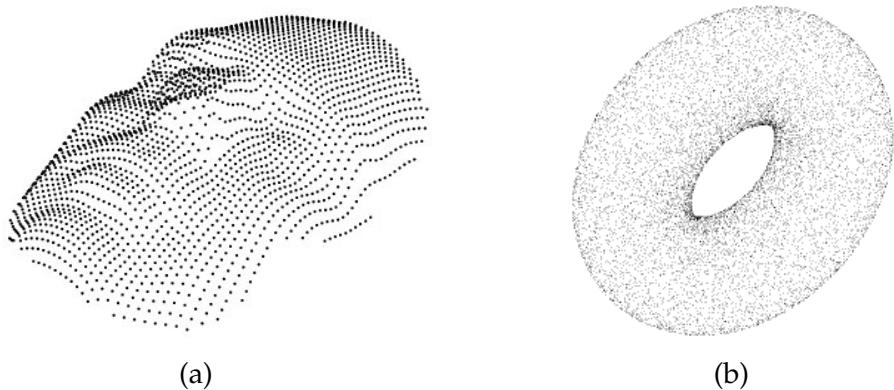
Modulačný kontrast c_d opisuje ako efektívne ToF senzor separuje a zbiera prijaté fotoelektróny. Intenzita pixelu A je funkciou optického výkonu, B predstavuje offset okolitého svetla a reziduálneho systému. Z tejto rovnice možno vyvodíť, že vysoká hodnota A , vysoká modulačná frekvencia f a vysoký modulačný kontrast zvyšujú presnosť merania. Naopak veľká úroveň ofsetu B spôsobuje saturáciu prijímača a znižuje presnosť systému [44].

Medzi výhody ToF snímačov patrí zníženie vplyvu externého osvetlenia na kvalitu merania. Je to spôsobené tým, že nedochádza k ovplyvňovaniu frekvencie vyzárovaného svetla. Naopak pri použití viac-kamerového systému s rovnakou modulačnou frekvenciou môže dochádzať k vzájomnej interferencii signálov. Tento jav je detailne opísaný v kapitole 7.

3.4 Mračno bodov a 3D povrch

Hĺbková mapa zachytáva 3D priestor v 2D rovine. Tento obraz je získavaný z optických hĺbkových snímačov. Pre spätnú transformáciu hĺbkového obrazu do 3D priestoru sa jednotlivé obrazové pixely prevádzajú do takzvaných mračien bodov (anglicky Point Cloud). Základnou jednotkou mračna bodov je dátový bod, ktorý v sebe ukladá informáciu o polohe (x,y,z). Dátový bod môže v sebe uchovávať aj iné informácie ako napríklad farbu, jas, normálový vektor a podobne [47].

Takéto mračno bodov má veľké využitie v priemyselných 3D CAD modeloch, pri metrológii a inšpekcii kvality a v iných sférach, pretože reprezentuje geometrické vlastnosti reálnych objektov.



Obr. 3.10: Mračno bodov: (a) Priestorová rekonštrukcia tváre vytvorennej z hĺbkovej mapy (Obr. 3.4) [40]. (b) Priestorová rekonštrukcia torusu [48].

Prevod mračna bodov na 3D povrch je realizovaný pomocou rekonštrukcie povrchu (*surfacer construction*). Medzi známe rekonštrukčné metódy patrí Ball-Pivoting [49] a Poissonova rekonštrukcia [50], kde sú dátové body transformované do siete. Polygónová sieť je zložená z vertexov (vertices), hrán (edges) a masiek (faces), tak tiež definuje tvar polyhedrálneho objektu v 3D grafike. Masky sa zvyčajne skladajú z trojuholníkov, štvorhranov a iných jednoduchých konvexných polygónov. Volumetrické siete explicitne reprezentujú povrch aj objem štruktúry, pričom polygónová sieť reprezentuje iba povrch. Objekty vytvorené z polygónov musia ukladať rôzne typy elementov. To zahŕňa vertexy, hrany, masky, polygóny a povrhy. Polygóny sú zvyčajne reprezentované trojuholníkmi [51].

4. Kalibrácia kamier

Kamera je nástroj, ktorý mapuje 3D priestor do 2D roviny obrazu. Ide o technické zariadenie, ktoré je využívané širokospektrálne. Existuje viacero technologických riešení a produktov, ktoré sú špecifikované na konkrétné aplikácie. Pri potrebe merania vzdialenosť alebo zachytenia presnej hĺbkovej informácie snímanej scény je potrebné uvažovať s kalibráciou kamery [34].

Geometrická kalibrácia je nevyhnutá z dôvodu skreslenia šošovky, čo spôsobuje výrazné nepresnosti pri transformácii 3D priestoru 2D obrazovej roviny. Pri kalibrácii je dôležité poznať matematický model kamery a transformácie, ktorými sa upravuje výsledný obraz.

Pri multi-kamerovej kalibrácii hľadáme vzájomnú polohu kamier vo svetových súradničiach. Ak je aplikácia zameraná na 3D rekonštrukciu priestoru z viacerých kamier, vzájomnou kalibráciou vieme ľahšie spojiť pohľady do jedného spoločného zosúladeného priestoru.

4.1 Afinné transformácie v priestore

Afínna transformácia je metóda lineárneho mapovania, ktorá zachováva body, priame čiary a roviny [52]. Množiny rovnobežných čiar zostávajú po afinnej transformácii rovnobežné. Technika afinitnej transformácie sa zvyčajne používa na korekciu geometrických skreslení alebo deformácií, ktoré sa vyskytujú pri neideálnych uhloch kamery. Matematický zápis afinnej transformácie je $X' = A \cdot X$, kde X predstavuje pôvodnú súradnicu bodu, X' transformovanú súradnicu bodu a A danú affinú transformáciu [53]. Afínna transformácia pre 3D priestor má tvar matice o rozmere 4×4 .

Identická transformácia: Určená jednotkovou maticou, body X a X' majú totožné súradnice.

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (4.1)$$

Translačná transformácia: Posun bodu X o vektor $\vec{a} = [a_x, a_y, a_z]$, $X' = \vec{a} + X$.

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & a_x \\ 0 & 1 & 0 & a_y \\ 0 & 0 & 1 & a_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (4.2)$$

Škálovacia transformácia: Zmena bodu X o vektor $\vec{a} = [a_x, a_y, a_z]$, $X' = \vec{a}X$.

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} s_x & 0 & 0 & 0 \\ 0 & s_y & 0 & 0 \\ 0 & 0 & s_z & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (4.3)$$

Rotačná transformácia v osi x : Otočenie bodu X okolo začiatku súradnicovej sústavy o uhol α .

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \alpha & -\sin \alpha & 0 \\ 0 & \sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (4.4)$$

Rotačná transformácia v osi y: Otočenie bodu X okolo začiatku súradnicovej sústavy o uhol α .

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} \cos \alpha & 0 & \sin \alpha & 0 \\ 0 & 1 & 0 & 0 \\ -\sin \alpha & 0 & \cos \alpha & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (4.5)$$

Rotačná transformácia v osi z: Otočenie bodu X okolo začiatku súradnicovej sústavy o uhol α .

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} \cos \alpha & -\sin \alpha & 0 & 0 \\ \sin \alpha & \cos \alpha & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (4.6)$$

Súmerná transformácia v osi x: Matica súmernosti vzhľadom na os x.

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (4.7)$$

Súmerná transformácia v osi y: Matica súmernosti vzhľadom na os y.

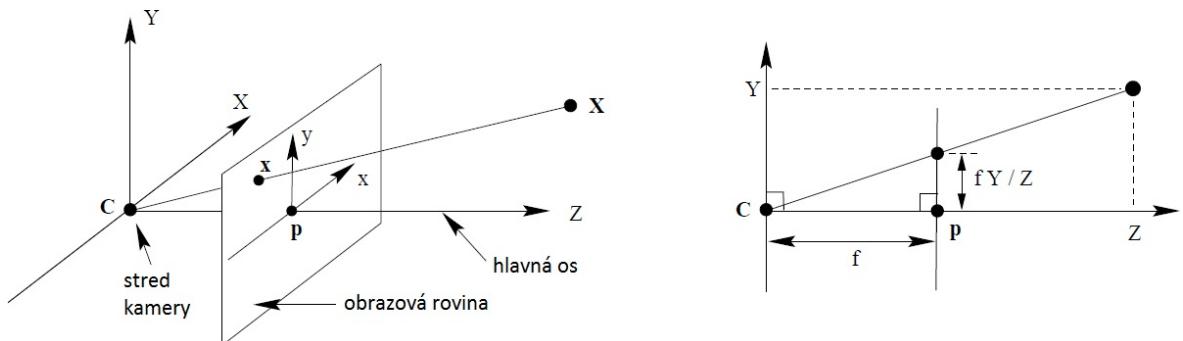
$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (4.8)$$

Súmerná transformácia v osi z: Matica súmernosti vzhľadom na os z.

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (4.9)$$

4.2 Dierková kamera

Dierková kamera (pinhole camera) predstavuje ideálnu kameru, ktorá nevykazuje žiadne skreslenia spôsobené šošovkou. Ide o najjednoduchší model kamery, ktorým svetlo namiesto šošovky prechádza cez štrbinu a prevrátené dopadá na rovinu obrazu. Dochádza tu k centrálnej projekcii bodov zo svetových súradníč do roviny obrazu [54].

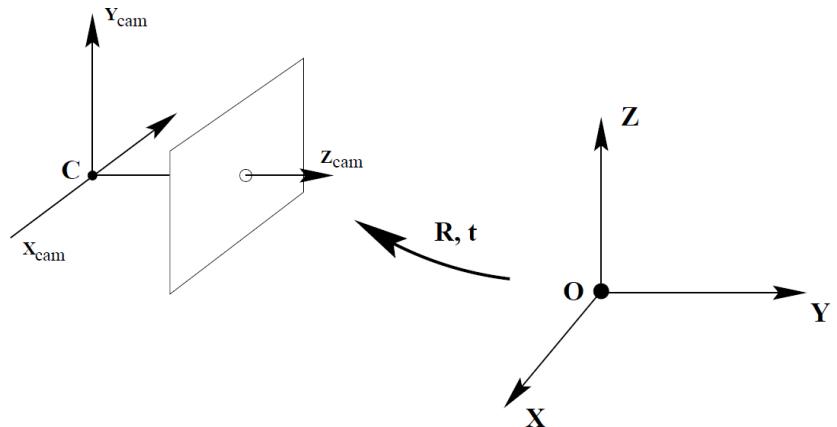


Obr. 4.1: Perspektívna projekcia štrbinovou kamerou [54].

Na obrázku 4.1 je znázornená geometria dierkovej kamery. Stredom projekcie je optické centrum. Kolmica medzi optickým centrom a obrazovou rovinou π sa nazýva hlavná os Z . Hlavným bodom p je miesto, v ktorom os Z pretína obrazovú rovinu. Rovina prechádzajúca stredom kamery rovnobežne s rovinou obrazu sa nazýva hlavná rovina kamery. C reprezentuje stred kamery (stred perspektívneho premietania). Za predpokladu, že svetové a obrazové body sú zastúpené v homogénnych súradničiach, potom sa stredná projekcia môže jednoducho vyjadriť ako lineárne mapovanie medzi ich homogénnymi súradnicami.

Prevod medzi svetovou súradnicovou sústavou (SSS) a kamerovou súradnicovou sústavou (KSS) je zobrazený na obr. 4.2 predstavený vzťahom 4.10.

$$\begin{bmatrix} X_{cam} \\ Y_{cam} \\ Z_{cam} \\ 1 \end{bmatrix} = \mathbf{R}\mathbf{T} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (4.10)$$



Obr. 4.2: Vzťah medzi svetovou súradnicovou sústavou a súradnicovou sústavou kamery [54].

Svetová súradnicová sústava je reprezentovaná trojrozmerným karteziánskym súradnicovým systémom, kde stred sústavy O (prípadne O_w) je v bode $[0,0,0]$. Vektory X, Y, Z prislúchajúce k jednotlivým súradniciam sú na seba kolmé. Definujú polohu objektu v scéne. Udávajú sa v dĺžkových jednotkách [m] alebo [mm].

Súradnicová sústava kamery je reprezentovaná trojrozmerným súradnicovým systémom v bode C (poprípade O_c), čo predstavuje stred kamery. Optická os Z_c je daná vektorom pohľadu kamery. Vzťah medzi svetovou súradnicovou sústavou a súradnicovou sústavou kamery je popísaný transláciou t a rotáciou R . Vektory sú označené ako $(X_{cam}, Y_{cam}, Z_{cam})$

Body zo súradnicovej sústavy kamery je následne potrebné premietnuť do roviny obrazu π . Nech je stred premietania pôvodom euklidovského súradnicového systému a obrazová (ohnisková) rovina $Z = f$. Podľa modelu dierkovej kamery na obr. 4.1 sa bod v priestore so súradnicami $(X, Y, Z)^T$ mapuje na bod v rovine obrazu $(f\frac{X}{Z}, f\frac{Y}{Z}, f)^T$ pomocou podobnosti trojuholníkov. Ignorovaním výsledných súradníc obrazu dostávame vzťah, ktorý popisuje mapovanie centrálnej projekcie zo svetovej súradnicovej sústavy do obrazovej súradnicovej sústavy.

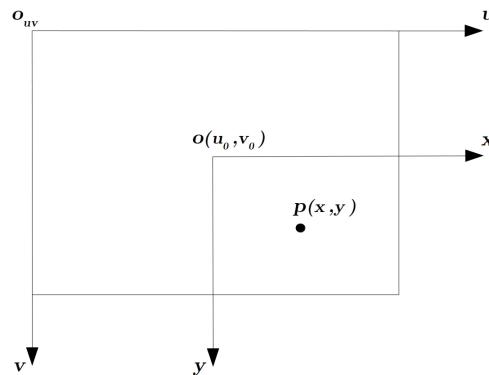
$$\frac{X_{cam}}{x} = \frac{Z_{cam}}{f}; \frac{Y_{cam}}{y} = \frac{Z_{cam}}{f} \quad (4.11)$$

$$x = f \frac{X_{cam}}{Z_{cam}}; y = f \frac{Y_{cam}}{Z_{cam}} \quad (4.12)$$

Rovnice 4.11 a 4.12 môžeme prepísať do maticového tvaru

$$\begin{bmatrix} fX_{cam} \\ fY_{cam} \\ Z_{cam} \end{bmatrix} = Z_{cam} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X_{cam} \\ Y_{cam} \\ Z_{cam} \\ 1 \end{bmatrix} \quad (4.13)$$

Rovina obrazu aj súradnicový systém obrazových bodov predstavuje dvojrozmerný priestor. Rozdielom je, že prvý spomenutý systém má počiatok v hlavnom bode a jednotka súradnicového systému je v milimetroch. Druhý spomenutý je reprezentovaný jednotkou pixel, počiatok súradnicového systému je na pozícii $(u, v) = [0, 0]$. Preto je potrebné transformovať prvý súradnicový systém do druhého.



Obr. 4.3: Vzťah roviny obrazu a súradnicového systému obrazových bodov.

Z obr. 4.3 je definovaný hlavný bod $p = o(u_0, v_0)$ (bod, ktorým prechádza hlavná os Z_{cam} v obr. 4.1). Pre prevod potrebujeme poznať hustotu pixelov na jednotku dĺžky. Tá je určená pomerom počtu pixelov a rozmerom snímacej plochy v jednotlivých smeroch. Pridaním offsetu u_0 a v_0 dostávame transformačný vzťah

$$u = \frac{x}{d_x} + u_0; v = \frac{y}{d_x} + v_0 \quad (4.14)$$

Z rovnice 4.14 získavame kalibračnú maticu K , ktorá je zložená z ohniskových vzdialenosí f_x, f_y a hlavných bodov p_x, p_y . Do úplnej matice K patrí aj koeficient skosenia s , ktorý je zvyčajne 0. Táto matica je zložená z takzvaných vnútorných pa-

rametrov kamery.

$$K = \begin{bmatrix} \frac{x}{d_x} & s & u_0 \\ 0 & \frac{y}{d_y} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} f_x & s & p_x & 0 \\ 0 & f_y & p_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad (4.15)$$

Matica kamery P sa skladá z vnútorných a vonkajších parametrov. Matica R zastupuje rotáciu a t transláciu, čo sú vonkajšie parametre (obr. 4.2). Matica K zastupuje vnútorné parametre.

$$P = K \begin{bmatrix} R & t \end{bmatrix} = \begin{bmatrix} f_x & 0 & p_x \\ 0 & f_y & p_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix} \quad (4.16)$$

Celý prevod je teda možné vyjadriť podľa rovnice 4.17

$$Z_{cam} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & p_x \\ 0 & f_y & p_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (4.17)$$

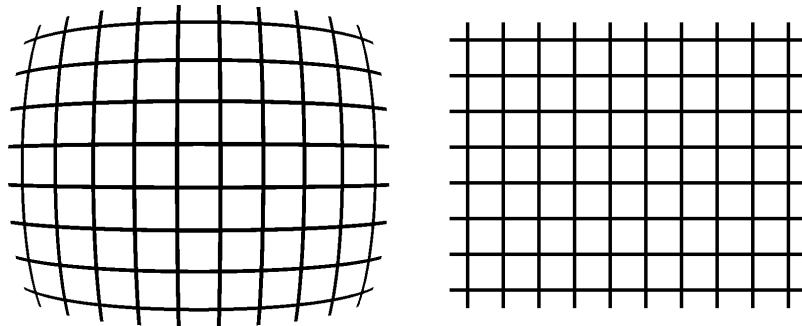
ktorá je základným matematickým opisom dierkovej kamery a prevodu svetových súradníč do súradnicového systému obrazových bodov.

4.3 Skreslenie šošoviek

V reálnych kamerách sa vyskytujú určité typy nelineárneho skreslenia, ktoré sú spôsobené pridaním šošovky do ideálnej kamery. Tieto skreslenia spôsobujú nesprávnu perspektívnu projekciu, čím sa následne stráca presná informácia o snímanej scéne. Tieto skreslenia je však možné odstrániť získaním vnútorných parametrov kamery. Medzi najčastejšie skreslenia patrí radiálne a tangenciálne skreslenie.

4.3.1 Radiálne skreslenie

Toto skreslenie vznikne, keď sa svetelné lúče ohnú viac v blízkosti okrajov šošovky ako v optickom centre. Čím je objektív menší, tým väčšie je skreslenie. Typickými radiálnymi skresleniami sú sudkovité a poduškové. Koeficienty radiálneho skreslenia modelujú tento typ skreslenia. Deformované body sú označené ako x' a y' . Prvky k_1 až k_6 predstavujú koeficienty radiálneho skreslenia.



Obr. 4.4: Radiálne skreslenie spôsobené zakrivením šošovky [55].

Neskreslené pixely označené x, y sa nachádzajú v normalizovaných súradničiach obrázka. Normalizované súradnice sú vypočítané zo súradníc pixelov preložením do optického stredu a prepočítané pomocou ohniskovej vzdialenosťi v pixeloch. Takže x a y sú bezrozmerné.

$$x' = x \frac{1 + k_1 r^2 + k_2 r^4 + k_3 r^6}{1 + k_4 r^2 + k_5 r^4 + k_6 r^6} \quad (4.18)$$

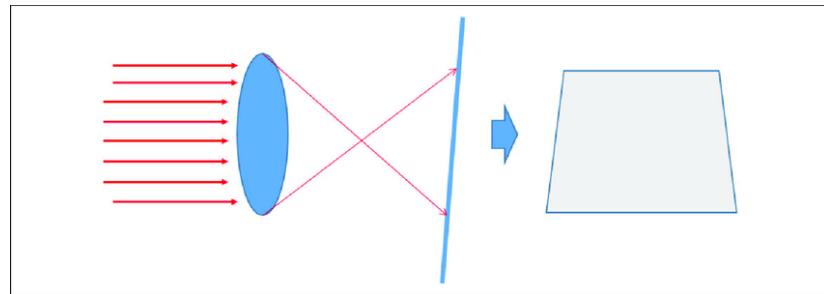
$$y' = y \frac{1 + k_1 r^2 + k_2 r^4 + k_3 r^6}{1 + k_4 r^2 + k_5 r^4 + k_6 r^6} \quad (4.19)$$

$$r^2 = x^2 + y^2 \quad (4.20)$$

Ak je $k_1 < 0$, tak je skreslenie sudkovité a ak $k_1 > 0$, tak ide o skreslenie poduškové. Koeficienty k_3 a vyššie sú pri bežných deformáciách zvyčajne rovné 0. Prejavujú sa prevažne pri širokouhlých objektívoch. r reprezentuje rádius k hlavnému bodu, kde nie je skreslenie.

4.3.2 Tangenciálne skreslenie

Tangenciálne skreslenie vzniká pri neparalelnom umiestnení senzora kamery so šošovkou. Koeficienty p_1 a p_2 predstavujú tangenciálne skreslenie šošovky:



Obr. 4.5: Tangenciálne skreslenie spôsobené nesprávnym umiestnením šošovky [56].

$$x' = 2p_1xy + p_2(r^2 + 2x^2) \quad (4.21)$$

$$y' = p_1(r^2 + 2x^2) + 2p_2xy \quad (4.22)$$

$$r^2 = x^2 + y^2 \quad (4.23)$$

4.3.3 Model reálnej kamery s korekciou skreslenia šošoviek

V reálnych podmienkach je model ideálnej (dutinkovej) kamery doplnený o rovnicu korekcie skreslenia šošoviek (tangenciálneho a radiálneho). Potreba získania koeficientov kamery je veľmi dôležitá z hľadiska správnej späťnej rekonštrukcie 2D obrazu do 3D priestoru [34].

$$\begin{bmatrix} u \\ v \\ z \end{bmatrix} = R \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + t \quad (4.24)$$

$$x' = \frac{x}{z}; y' = \frac{y}{z} \quad (4.25)$$

$$x'' = x' \frac{1 + k_1 r^2 + k_2 r^4 + k_3 r^6}{1 + k_4 r^2 + k_5 r^4 + k_6 r^6} + 2p_1 x' y' + p_2 (r^2 + 2x'^2) \quad (4.26)$$

$$y'' = y' \frac{1 + k_1 r^2 + k_2 r^4 + k_3 r^6}{1 + k_4 r^2 + k_5 r^4 + k_6 r^6} + p_1 (r^2 + 2y'^2) + 2p_2 x' y' \quad (4.27)$$

$$u = f_x x'' + c_x; v = f_y y'' + c_y \quad (4.28)$$

4.4 Kalibrácia

Kalibrácia kamery je nevyhnutný krok v 3D spracovaní obrazu, kde sa získavajú metrické informácie z 2D obrazov. Kalibráciu môžeme rozdeliť na 3 základné kategórie.

Kalibrácia pomocou korešpondencie bodov v 3D priestore

Kalibrácia kamery sa vykonáva pozorovaním kalibračného objektu, ktorého geometria v 3D priestore je presne známa. Kalibračný objekt sa obvykle skladá z dvoch alebo troch rovín kolmých na seba [57].

Autokalibrácia

Táto technika nepoužíva žiadny kalibračný objekt. Posunutím kamery v statickej scéne je možné vypočítať vnútorné parametre kamier. Ak sú snímky zosnímané tou istou kamerou s pevnými vnútornými parametrami, korešpondencia medzi tromi obrázkami je dostatočná na vypočítanie interných aj externých parametrov, ktoré nám umožňujú rekonštruovať 3D štruktúru.

Kalibrácia pomocou 2D roviny v 3D priestore

Parametre kamery sa skladajú z vnútorných (intrinických), vonkajších (extrinsic-kých) koeficientov a koeficientov skreslenia (distortion). K určeniu je potrebné mať 3D body scény (world points) a ich zodpovedajúce 2D body. Získanie týchto korešpondujúcich dát je možné extrakciou ľahko identifikateľných bodov. Jedným z najpoužívanejších kalibračných vzorov je šachovnica, kde je výrazný farebný prechod medzi hracími poľami [34].

5. Ciele dizertačnej práce

- Návrh algoritmu umožňujúceho vytvorenie precízneho priestorového modelu objektu z viacerých kamier vrátane kalibrácie vizuálneho systému.
- Návrh algoritmu pre automatizovanú detekciu kľúčových bodov na objekte a automatizované meranie vybraných geometrických priestorových parametrov.

Problematika riešená v rámci dizertačnej práce pochádza z lekárskeho klinického prostredia (Klinika detí a dorastu JLF UK, Martin). Návrhu témy predchádzala dlhorocná spolupráca medzi Žilinskou univerzitou a Jesseniovou lekárskou fakultou UK v Martine. Špecialisti z klinického prostredia aktívne vstupujú do jednotlivých fáz riešenia, čím poskytujú užitočnú spätnú väzbu v celom procese, stanovovanie limitov a referenčných hodnôt a nakoniec aj validáciu výsledkov. Splnené ciele predloženej dizertačnej práce budú implementované do komplexného nástroja automatizovanej diagnostiky vybraných ochorení.

Riešenie úlohy definovanej v prvom cieli je založené na využití multi-kamerového systému, ktorý v reálnom čase dokáže zachytiť priestorovú informáciu snímanej scény. Znižovanie času snímania objektu predstavuje z lekárskeho i technického hľadiska veľký prínos (pri zachovaní definovanej kvality). Jednotlivé priestorové informácie z kamier sa následne zlúčia do spoločného priestorového modelu, ktorý bude použitý pri riešení druhého cieľa. Podstatnou problematikou je vzájomná kalibrácia multi-kamerového systému a návrh filtračného algoritmu, ktorým sa zvýší kvalita výstupných priestorových dát.

Náplň druhého cieľa môže byť riešená viacerými spôsobmi. Podstatným krokom je identifikácia kľúčových bodov alebo oblastí v priestorovom modeli. Jedným z možných postupov je segmentácia RGB-D mapy, kde sa pomocou algoritmov klasifikujú regióny bodov. Z výsledku klasifikácie bude následne možné rekonštruovať iba vybrané regióny 3D priestorového modelu. V týchto čiastkových modeloch možno potom vykonať spoľahlivejšiu detekciu kľúčových bodov a meranie vzdialenosí, na kolko nebudú do týchto algoritmov vstupovať irelevantné časti modelov. Výsledkom bude množina mračien bodov, ktoré môžu vystupovať ako samostatné modely alebo budú reprezentovať ucelený konzistentný model.

6. Návrh snímacieho systému

Pre návrh algoritmu, ktorý umožní vytvorenie precízneho modelu dynamického objektu, je potrebné navrhnúť vhodný kamerový systém. Práve špecifická aplikácia systému môže výrazne ovplyvniť možnosti skenovania. Snímaný objekt je zastúpený pediatrickým pacientom. Jeho povrchová štruktúra nie je pevná a v čase sa môže meniť. Tento fakt môže ovplyvniť geometrickú presnosť modelu hlavne ak je dĺžka doby skenovania príliš dlhá. Cieľom tejto kapitoly je overiť a navrhnúť možnosti snímania pediatrických pacientov.

6.1 Microsoft Kinect v2

Ide o druhú generáciu hľbkových kamier od spoločnosti Microsoft. Oproti prvej generácii využíva technológiu ToF (detailnejšie v 3.3.3), pričom ide o jednu z najznámejších senzorov tohto typu na svete [58]. Rozsah snímanej scény je $0.5 - 4.5\text{m}$. Farebný obraz je dodávaný v rozlíšení $1920 \times 1080\text{p}$ a 30Hz , hľbková mapa ma rozlíšenie 512×424 .



Obr. 6.1: ToF kamera Microsoft Kinect v2 [59].

Presnosť hľbkovej mapy je oproti prvej verzii Kinectu vyššia, taktiež je znížený negatívny vplyv spôsobený slnečným svetlom.

6.2 Snímanie dynamických objektov

Pri snímaní dynamických objektov je potrebné uvažovať so vznikom pohybových artefaktov. Tie vznikajú, ak objekt v dobe skenovania mení svoje priestorové umiestnenie. Z toho dôvodu je potrebné buď stabilizovať objekt počas doby snímania alebo redukovať časovú dĺžku skenovania. Kedže tento systém má byť určený pre medicínske aplikácie, kde skenovaný objekt bude pediatrický pacient, je potrebné overiť možnosti experimentálne.

6.2.1 Jedno-kamerové snímanie dynamických objektov

Experimentálne snímanie pacientov bolo vykonávané na Klinike detí a dorastu Jesseniovej lekárskej fakulty v Martine, Laboratórium spánkovej medicíny. Do testu bolo vybraných 9 pacientov rôzneho pohlavia, vo vekovom rozmedzí od 4 do 12 rokov. Ako skenovací nástroj bola použitá kamera Kinect v2.



Obr. 6.2: Modely dynamických pediatrických pacientov [59].

V spolupráci s komerčným softvérom KScan3D boli vytvorené modely pacientov, ktoré sú zobrazené na obr. 6.2. Ako je vidieť, modely sú nekvalitné a výrazne deformované. Je to spôsobené tým, že objekty nedokázali zostať v statickej polohe počas

doby skenovania (tá presahovala 1 minútu pri každom pacientovi). Tendenciou bolo otáčať sa za kamerou, čo viackrát viedlo k nutnosti začať celý proces od znova. Takéto modely nie sú postačujúce pre diagnostikovanie OSAS. Ich miera nepresnosti je veľmi vysoká a užitočná geometria tváre sa stráca. Z experimentu vyplýva, že jedno-kamerový systém je nepoužiteľný pre skenovanie dynamických objektov ako sú pediatrickí pacienti.

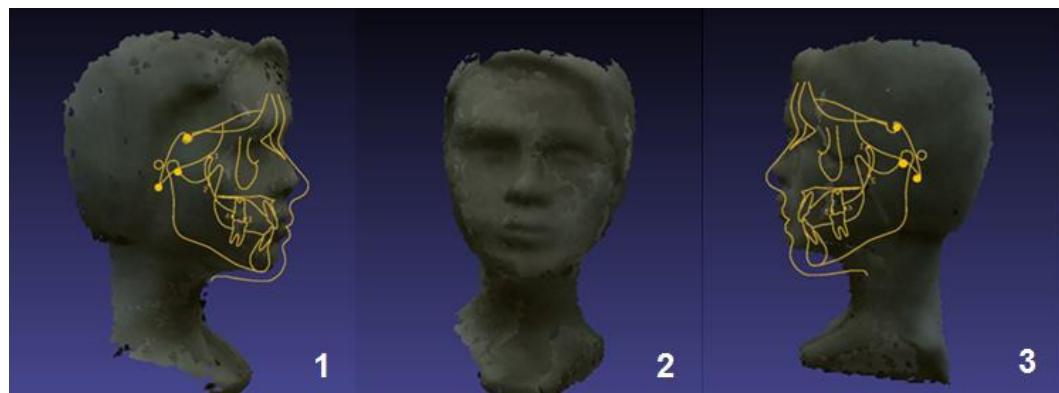
6.2.2 Priestorové rozloženie multi-kamerového systému

Pri multi-kamerovom systéme sa pre skenovanie používajú viaceré kamery, ktoré sú staticky rozložené v priestore. Ich počet závisí od detailov objektu, ktorý má byť zosnímaný.



Obr. 6.3: Rozloženie trojice snímačov Kinect v2 v skenovacej kabíne.

Medzi hlavné požiadavky na model sú: hlava musí obsahovať laterálne pohľady (umožniť prekrytie s cefalometrickou snímkou) a orientačné body mäkkého tkaniva.



Obr. 6.4: Zobrazenie modelov s vyznačenými cefalometrickými parametrami pri použití 3 snímačov.

Dôležité je aj získanie informácie o šírke krku. Na obrázku 6.3 je zobrazené rozloženie kamier v skenovacej kabíne. Ich jednotlivé 3D modely aj s verifikáciou požiadaviek sa nachádzajú na obr. 6.4.

6.2.3 Sekvenčné snímanie multi-kamerového systému

Sekvenčný mód predstavuje postupné snímanie. V jednom momente vždy sníma len jedna kamera a ostatné kamery sú vypnuté. Tento režim je pri ToF kamerách často využívaný, pretože tu nevzniká multi-kamerová interferencia. Problémom však je dlhá doba zapínania IR projektora, ktorá pri kamerách Microsoft Kinect v2 dosahuje približne $1s$. Časový odstup medzi dvoma prijatými snímkami je pri $30Hz$ okolo $33ms$.

6.2.4 Paralelné snímanie multi-kamerového systému

V paralelnom režime pracujú všetky kamery v rovnakom čase. Oproti sekvenčnému snímaniu sa inicializácia vykonáva pre každú kameru iba raz. Tým sa radikálne znižuje doba snímania objektu zo všetkých uhlov. Nevýhodou je ale vznik multi-kamerovej interferencie, ktorá dokáže negatívne ovplyvniť výstupné hĺbkové mapy. Časový rozostup medzi jednotlivými snímkami je ovplyvnený samotným spracovaním dát, maximálne však $33ms$ medzi všetkými snímkami.

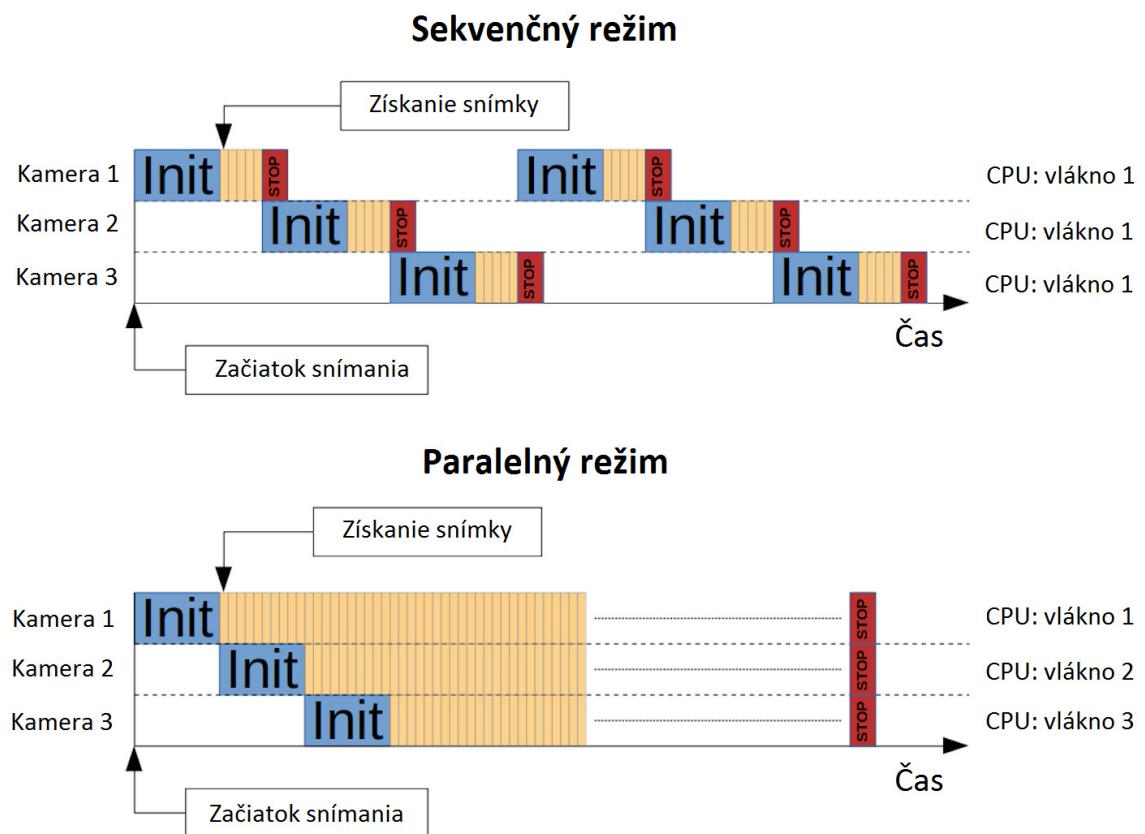
6.2.5 Porovnanie režimov snímania dynamických objektov

Pre porovnanie týchto režimov bol navrhnutý experiment, pri ktorom sa porovná veľkosť zmeny polohy objektu v závislosti na dĺžke snímania. Pre testovanie bol vyrobený merací prvok, ktorý pozostával z jednosmerného motora, ukazovateľa aktuálnej polohy a statického podstavca slúžiaceho ako uhlomer. Na hriadele motoru bol pripojený ukazovateľ, ktorý vykonával pohyb po kružnici. Rýchlosť bola nastavená regulovateľným zdrojom tak, aby otočka trvala $5s$. Obvod statického podstavca bol rozdelený na 36 dielov oddelených od seba po 10° . Kvôli identifikácii polohy podstavec obsahoval zvýraznený nultý bod a šípkou informujúcu o smere otáčania ukazovateľa.

Pri sekvenčnom režime boli snímky z kamier získavané postupne od 1. po 4. kamieru. Pri paralelnom režime je ľahké identifikovať poradie. Pre porovnanie bolo potrebné získané hľbkové mapy previesť na mračno bodov a pomocou ICP metódy ich registrovať do jedného modelu (zhodná pozícia nultého bodu).



Obr. 6.5: Porovnanie 3D mračien bodov pri rozdielnych režimoch snímania [60]: (a) Mračno bodov vytvorené pre sekvenčný režim. (b) Mračno bodov vytvorené pre paralelný režim.



Obr. 6.6: Grafické porovnanie sekvenčného a paralelného režimu snímania.

Z výsledných modelov je jasne vidieť, že paralelné spracovanie dát má vyšší zmysel pri snímaní dynamických objektov. Pri sekvenčnom režime je rozdiel polohy uka-zovateľa výrazný, čo spôsobovalo problémy aj pri ICP registrácií (obr. 6.5a). Naproti tomu je model vytvorený paralelným režimom konzistentnejší, rozdiel polohy uka-zovateľa je podstatne menší (obr. 6.5b). Ten by sa dal znížiť hardvérovým synchroni-zovaným snímaním. Pri kamerách Kinect v2 však táto možnosť chýba a je potrebná. Rozdiel medzi jednotlivými pracovnými režimami sa nachádza na obr. 6.6.

6.3 Kalibrácia kamerového systému

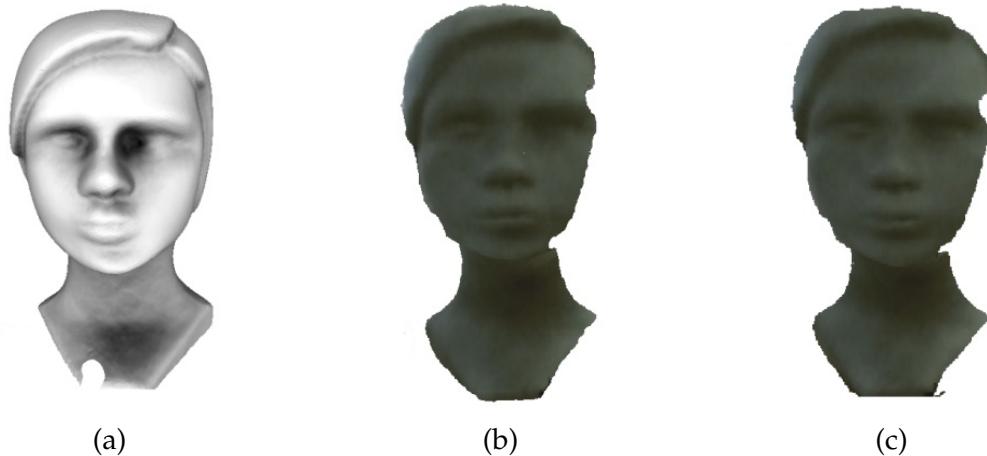
Nevyhnutným krokom pri kamerových systémoch je kalibrácia jednotlivých kamier. Pomocou nej sa odstraňuje radiálne a tangenciálne skreslenie, taktiež sa získava matica kamery P (rovnica 4.16). V tejto kapitole je opísaný postup geometrickej ka-librácie jednotlivých kamier a multi-kamerová kalibrácia.

6.3.1 Geometricka kalibrácia

Cieľom geometrickej kalibrácie je získanie vnútorných parametrov kamery spolu s korekčnými parametrami pre odstránenie tangenciálneho a radiálneho skreslenia. Tieto parametre bolo potrebné získať pre RGB aj IR senzory všetkých používaných hĺbkových kamier Kinect v2.

Pre kalibráciu RGB a IR senzora sa vytvorila séria snímok, ktoré obsahovali ka-libračný vzor snímaný z rôznych uhlov. Ten bol reprezentovaný šachovnicovým mo-tívom o rozmeroch 10×8 , pričom dĺžka hrany mala veľkosť $36mm$.

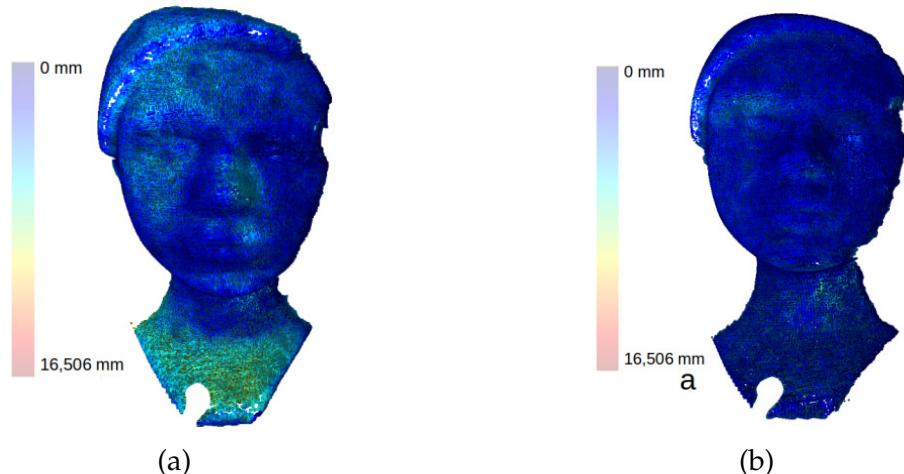
Pre overenie správnosti kalibrácie IR snímača sa z kalibrovanej a nekalibrova-nej hĺbkovej mapy vytvorili mračná bodov. Pri rovnakých podmienkach prostredia bol zosnímaný model hlavy s továrenskými nastaveniami, potom boli použité ka-libračné koeficienty získané v predchádzajúcom kroku. Tie boli transformované na trojuholníkovú sieť. Výsledné modely sa porovnávali voči referenčnému modelu. Re-ferenčný model bol vytvorený ručným laserovým skenerom. Ukážky sa nachádzajú na obr. 6.7.



Obr. 6.7: Modely trojuholníkových sietí, použité pre overenie presnosti kalibrácie: (a) Referenčný model objektu vytvorený laserovým snímačom. (b) Model vytvorený z nekalibrovanej kamery Kinect v2. (c) Model vytvorený z kalibrovanej kamery Kinect v2.

Porovnanie priestorovej rekonštrukcie hĺbkových máp bolo robené pomocou Hausdorffovej vzdialenosťi.

V nej je detailnejšie zobrazené, aký vplyv má kalibrácia na presnosť rekonštrukcie. Z modelov vzdialenosťi (obr. 6.8) je vidieť, že nekalibrovaný systém vykazoval vyššiu chybu na okrajoch modelu. To je spôsobené pozitívnym radiálnym skreslením IR senzora, ktorého intenzita je výraznejšia na okrajoch obrazu. Štatistické výsledky pre 10 modelov sa nachádzajú v tabuľke 6.1.



Obr. 6.8: Vizualizácia Hausdorffovej vzdialenosťi medzi referenčným modelom (Obr. 6.7a) a modelom získaným z kamery Kinect v2: (a) Nekalibrovaný systém (výsledok pre model z obr. 6.7c). (b) Kalibrovaný systém (výsledok pre model z obr. 6.7b).

Tabuľka 6.1: Štatistické porovnanie geometrickej kalibrácie kamier.

Model	Počet meraní [-]	Priemer [mm]	RMS [mm]
Nekalibrovaný	185187	3,0206	3,8444
Kalibrovaný	296585	2,1832	2,9506

Týmto krokom sa overil spôsob geometrickej kalibrácie hĺbkových máp. Továrenské nastavenia pre kamery Kinect v2 neposkytujú presné intrinické parametre a koeficienty skreslenia. V multi-kamerovom systéme, kde sa využíva viacero hĺbkových kamier a vyžaduje sa presnosť rekonštrukcie, je kalibrácia potrebná.

6.3.2 Multi-kamerová kalibrácia

V tomto procese sa odhadujú vzájomné pozície staticky umiestnených kamier vo svetovej súradnicovej sústave. Ich vzájomnú pozíciu určujú rotačné a translačné parametre (kombinácia affiných transformácií z kapitoly 4.1). Tie je možné získať viacerými spôsobmi. Dôležité je si určiť referenčnú kameru, ktorej matica vonkajších parametrov bude mať tvar identickej affinnej transformácie. Pri multi-kamerovej kalibrácii sa využíva podobný postup ako pri geometrickej kalibrácii. Rozdielom je však, že sa kalibračný vzor sníma súčasne z viacerých kamier. Problém nastáva, ak ich rozloženie neumožňuje súčasné snímanie. Takýto prípad je zachytený na obr. 6.3.

Riešením je separátne kalibrovanie medzi susednými kamerami (napr. 1-2, 1-3 a 2-4 alebo 3-4). V tomto procese sa vytvorí séria RGB-D párov snímok z jednotlivých dvojíc. Snímky obsahujú rovnaký kalibračný vzor so šachovnicovým motívom. Dôležité je, aby bol vzor staticky umiestnený a nevznikol polohový posun vo dvojici RGB-D mapy.

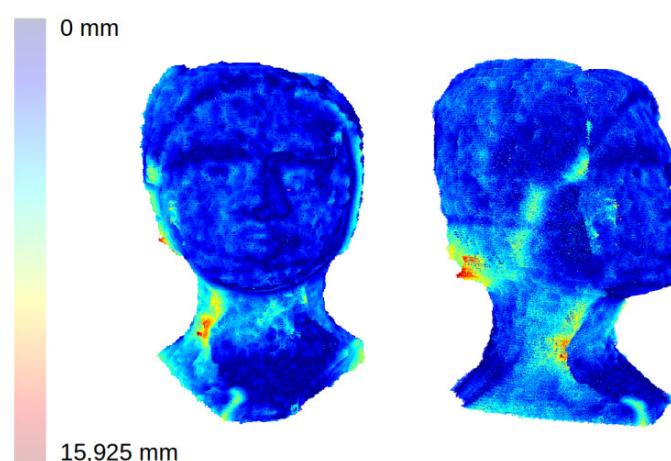


Obr. 6.9: Ukážka RGB-D páru použitého pri multi-kamerovej kalibrácii.

Efektívnejším spôsobom je hľadanie vzájomnej pozície kamier cez kľúčové body. V takomto prípade každá staticky umiestnená kamera zosníma kalibračný objekt, z ktorého vygeneruje mračno bodov. V každom mračne sa identifikujú spoločné body, ktoré slúžia na prvotné zarovnanie (*3 point picking alignment*). Následne sa pomocou ICP metódy minimalizuje suma štvorcov euklidovských vzdialenosťí bodov a vykoná sa jemnejšie zarovnanie. Výhodou statického rozmiestnenia kamier je, že multi-kamerovú kalibráciu je nutné vykonať iba raz. Získaním rotačných a translačných parametrov pre každú kameru získavame zarovnané mračná v spoločnom projektovanom priestore (6.2).



Obr. 6.10: Rekonštruované modely z jednotlivých kamier v spoločnej svetovej súradnicovej sústave: (a) Nekalibrovaná kamerová sústava. (b) Kalibrovaná kamerová sústava.



Obr. 6.11: Vizualizácia Hausdorffovej vzdialenosťí zarovnaných modelov pomocou multi-kamerovej kalibrácie.

Tabuľka 6.2: Rotačné a translačné parametre pre multi-kamerový systém (topológia systému je na obr. 6.3). Parameter K predstavuje jednotlivé kamery v tejto sústave.

K	r_{11}	r_{12}	r_{13}	r_{21}	r_{22}	r_{23}	r_{31}	r_{32}	r_{33}	t_1	t_2	t_3
1	0,2	-0,1	0,99	-0,74	0,15	0,98	0,1	-0,09	-0,99	0,08	0,04	0,62
2	1	0	0	0	1	0	0	0	1	0	0	0
3	0,09	0,16	-0,98	0,56	-0,14	0,98	0,15	-0,07	0,99	0,13	0,11	0,5

Výsledky Hausdorffovej vzdialenosť medzi referenčným modelom a zarovnaným modelom (trojuholníkové siete) sa nachádzajú v tabuľke 6.3.

Tabuľka 6.3: Štatistické porovnanie multi-kamerovej kalibrácie kamier.

Model	Počet meraní [-]	Priemer [mm]	RMS [mm]
Kalibrovaný	1071662	2,984	3,107

Z výsledkov vyplýva, že priemerná aj RMS chyba je $\sim 3\text{mm}$. Toto zarovnanie je vhodné pre systémy, ktorých rozloženie kamier je statické. Chyba však môže narásť zmenou pracovnej teploty senzora, čím sa ovplyvňuje aj presnosť merania hĺbky (pozri kapitolu 3.3.3). Z toho dôvodu je potrebné vykonávať ešte jemné zarovnanie mračien pomocou ďalších registračných metód.

6.4 Návrh skenovacej metodiky

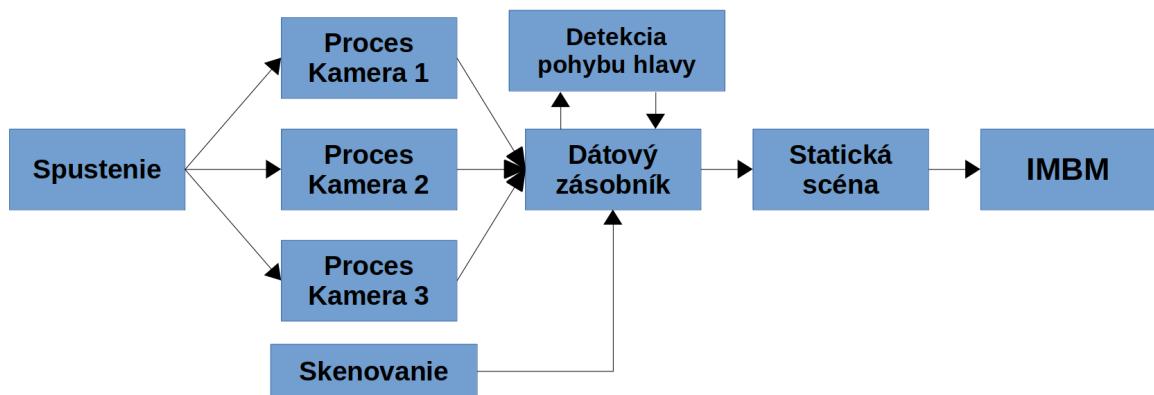
Úlohou skenovacej metodiky zloženej z prezentovaných čiastkových algoritmov je vytvoriť konzistentné mračno bodov dynamického objektu z multi-kamerového systému. Toto mračno (model) bude použité pre skríning a automatizovanú diagnostiku OSA v klinickej medicíne. Dynamický objekt je reprezentovaný hlavou pediatrického pacienta. Súčasťou skenovacieho procesu sú algoritmy na podporu automatizovaného merania vybraných kranio-faciálnych parametrov pomocou euklidovej vzdialenosť. Jadrom a vlastným vedeckým prínosom tejto časti práce je návrh metódy na potláčanie interferenčných artefaktov a vybraných degradačných procesov, na interpoláciu a opravu poškodených častí modelov. Súčasťou metodiky je tiež návrh algoritmu pre segmentáciu hĺbkových máp, ktorý identifikuje a rozdelí jednotlivé časti tváre podľa preddefinovanej klasifikácie. Z takto segmentovanej hĺbkovej

mapy bude možné vytvoriť mračná bodov jednotlivých častí tváre (oči, nos, brada, ústa,...). Takto segmentované oblasti budú v ďalšom výskume použité ako vstupné údaje do procesu automatizovanej diagnostiky.

Algoritmus môžeme rozdeliť podľa funkcionality na 4 časti:

- Zber a príprava dát v reálnom čase
- Spracovanie a filtrácia prijatých dát
- Meranie kranio-faciálnych parametrov tváre
- Segmentácia bodov podľa klasifikácie

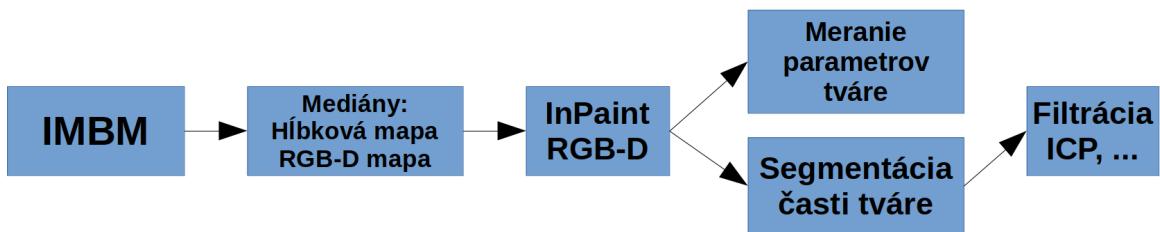
V prvej časti sa rieši paralelné spracovanie dát zo všetkých pripojených kamier. Vykonáva sa geometrická kalibrácia a základná 2D filtrácia spolu s identifikáciou pohybu tváre. Požiadavka na tento proces je, aby bol vykonávaný v reálnom čase. Ak frekvencia snímania je $30Hz$, časový rozostup medzi snímkami je $33ms$. Preto sú procesy spúšťané paralelne na CPU a GPU. Výstupom z tejto časti je obrazový zásobník, ktorý obsahuje obrazové informácie pediatrického pacienta v pokojovom stave.



Obr. 6.12: Diagram algoritmu od inicializácie kamier po IMBM filtráciu.

V druhej časti sa zo získaných obrazových dát vytvára trojrozmerná reprezentácia snímaného objektu. Prijaté dáta podstupujú filtráciu, ktorou sa zlepšuje kvalita výstupného modelu. Proces využíva štatistické filtre, ktorými sa odstraňujú chybné dáta. Aplikovaním nami navrhnutej metódy IMBM sa minimalizuje chyba multi-

kamerovej kalibrácie. Do priebehu tohto procesu vstupujú aj tretí a štvrtý proces. Tie pracujú s RGB-D obrazmi a hĺbkovými mapami, ktoré sú výstupom IMBM filtrácie (Obr. 6.13).



Obr. 6.13: Diagram spracovania dát po IMBM filtrácii.

Tretia časť sa zameriava na meranie euklidovských vzdialenosí špecifických vybraných parametrov. V RGB-D obraze sa identifikujú kľúčové body, ktoré sú následne prenesené do hĺbkovej mapy. Z týchto bodov sa vypočítajú 3D súradnice a vzdialenosí medzi jednotlivými pármí bodov. Týmto prístupom sa do určitej miery automatizuje vyšetrenie, ktoré je inak vykonávané kontaktným spôsobom. Bezkontaktný prístup je príjemnejší a menej stresujúci pre pediatrických pacientov. Taktiež skracuje dobu vyšetrenia.

V poslednej fáze sa vykonáva segmentácia výstupných RGB-D obrazov, v ktorých sa hľadajú pixely prislúchajúce jednotlivým častiam ľudskej tváre. Takoto segmentáciou v 2D obraze sa následne dokáže rozdeliť aj ucelené mračno bodov. Výstupom RGB-D segmentácie sú binárne masky, ktoré klasifikujú jednotlivé regióny v hĺbkových mapách. Pri priestorovej rekonštrukcii sa následne konkrétnym bodom nastaví identifikátor, ktorým sa body klasifikujú do tried. Takto je možné rozdeliť jedno zárovnane mračno vytvorené z viacerých hĺbkových máp.

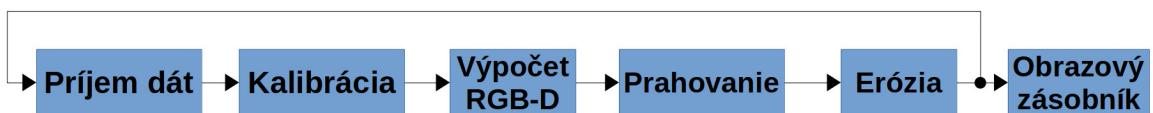
6.4.1 Nástroje použité pri vývoji algoritmu

Algoritmus a príslušný softvér bol vyvíjaný v jazyku C++17 na operačnom systéme Linux Mint 18.1. Projekt bol konfigurovaný cez CMake 3.5. Využíval pracovné rozhranie Qt 5.13.9, v ktorom bolo vyvíjané aj grafické rozhranie projektu. Kamery Kinecv2 využívali ovládač libfreenect2 [61] s podporou CudaKDE [62]. Pre spracovanie 2D obrazu bola použitá knižnica OpenCV 4. Pre prácu s 3D dátami sa využívala

knižnica *Point Cloud Library 1.9* s vizuálnou podporou knižnice *VTK 6.3*. Identifikácia kľúčových bodov využívala prostriedky *DLib 19.16*. Taktiež sa používali štandardné C++ knižnice a knižnica *Boost 1.63*. Algoritmus využíval procesorové vlákna. Softvér má predpripravené rozhranie pre spoluprácu s kamerami RealSense.

6.4.2 Príprava vstupných dát

Proces prípravy vstupných dát môžeme rozdeliť na 4 sekvencie. V prvej sa vykonáva inicializácia kamier a nastavenie konfiguračných parametrov. Medzi základné konfiguračné parametre patria kalibračné koeficienty. Po úspešnej inicializácii sa začína proces snímania. Každé snímanie je vykonávané vo vlastnom CPU vlákne, ktoré je nezávislé od hlavnej slučky programu (Obr. 6.14).



Obr. 6.14: Sekvencia spracovania prijatých dát z jednotlivých kamier vykonávaná v samostatnom vlákne.

Každá kamera tak vykonáva úlohy vo vlastnej oddelenej slučke. Samostatne sa tak prijímajú obrazové informácie z jednotlivých kamier. Získava sa 8-bitový RGB obraz, 32-bitový infračervený obraz a taktiež 32-bitová hĺbková mapa. Tieto obrazy sú zbavené skreslenia šošoviek. Mapovaním farebného obrazu na hĺbkovú mapu sa získava RGB-D obraz. Ten popisuje textúru hĺbkovej mapy. Z hĺbkovej mapy je prahovaním odstránené pozadie, pričom hranice prahu sú určené topológiou kamerového systému. Takto vznikne binárna maska objektu, ktorej eróziou zmenšíme plochu na hranách. Pomocou tej je maskovaný RGB-D obraz:

- RGB: 8-bitový 3-kanálový obraz s rozlíšením 1920×1080
- IR: 32-bitový šedo-tonový obraz s rozlíšením 512×424
- DEPTH: Prahovaný 32-bitový šedo-tónový obraz s rozlíšením 512×424
- RGB-D: Maskovaný 8-bitový 3-kanálový obraz s rozlíšením 512×424

V nasledujúcom kroku sa jednotlivé obrazové dát z kamier ukladajú do dátového zásobníka. Ten sa skladá z obrazových zásobníkov a zásobníkov mračien bodov. Jeho úlohou je spravovať sériu uložených dát, mazať staré a aktualizovať nové. Veľkosť zásobníka je prestaviteľná. Minimálnu hodnotu určuje IMBM filter, ktorý je určený na potlačenie multi-kamerovej interferencie. Jeho spravovanie beží v samostatnom vlákne, čím je znížená časová náročnosť obsluhy jednotlivých dát. Dátový zásobník poskytuje obrazové informácie ďalším procesom.

Vždy po prijatí nových obrazových dát spustí proces generovania mračna bodov, ktorý beží v samostatnom vlákne. Takto generované mračná slúžia ako 3D nadhľad v grafickom prostredí. Užívateľovi umožňujú vizuálnu kontrolu snímanej scény v reálnom čase (obr. xi.iii a xi.iv). Body mračna majú štruktúru XYZRGB. Kvôli rýchlejšiemu spracovaniu sa organizovaná štruktúra bodov (217088 bodov z hĺbkovej mapy) prevádzza na 1D pole, čím je možné odstrániť prebytočné body reprezentované 0 hodnotou pixelu v hĺbkovej mape. Takto sa počet bodov môže redukovať až o 90%, čím sa výrazne zníži výpočtová náročnosť. Mračná bodov jednotlivých kamier sa zarovnávajú do spoločného priestoru pomocou extrinzikých parametrov získaných pri multi-kamerovej kalibrácii.

Prácu s grafickým prostredím vykonáva užívateľ, ktorý si zároveň stabilizuje snímaný objekt do požadovanej polohy. Prostredie mu umožňuje vizuálne kontrolovať zmenu polohy a pohyb objektu v reálnom čase. Ak snímaná scéna vyhovuje požiadavkám, užívateľ spustí ďalší proces spracovania dát. V tomto procese sa hľadá séria snímok, kedy poloha objektu vykazuje minimálne zmeny. Táto požiadavka vyplýva z nutnosti použitia IMBM filtrácie. Nasledujúci spôsob detekcie polohy hlavy sa nachádza v ďalšej časti práce.

6.4.3 Identifikácia polohy hlavy

Z dôvodov, ktoré sú opísané v sekciu 6.4.2 je potrebné ukladať sériu snímok do dátového zásobníka. Z dát v zásobníku sa následne generuje mračno bodov, ktoré má čo najvernejšie reprezentovať snímaný objekt. Problém však nastáva, ak je objekt počas doby snímania dynamický. Pohybom môže vzniknúť problém s následným zarovnaním dát do spoločného modelu. Taktiež je veľmi pravdepodobné, že výsledný

model nebude geometricky odpovedať skutočným rozmerom. Vplyv pohybových artefaktov je znázornený na 6.2.5. Požiadavkou je získať sériu obrazových dát s minimálnou zmenou polohy snímaného objektu. Keďže kamery sú staticky umiestnené a zamerané na spoločný centrálny bod, pohyb objektu v hociktorom smere sa prejaví vo všetkých mapách. Tento fakt je rozhodujúci pri návrhu identifikačného systému. Z toho logicky vyplýva, že stačí analyzovať pohyb objektu len z dát jednej kamery. Identifikovanie kľúčových bodov, ktoré budú zhodné pre každú snímku v obrazovom zásobníku pomôže vypočítať pohyb a rozptyl polohy.

Ak je pixel P reprezentovaný konkrétnym stĺpcom r_i a riadkom c_i , tak bod záujmu P_i predstavuje jedinečnú pozíciu v obraze. Index i predstavuje konkrétny kľúčový bod detektora a N snímku v zásobníku. Bodu P_i prislúcha informácia o farbe $R_i; G_i; B_i$ (získanej z RGB-D obrazu) a o hĺbke d_i (získanej z hĺbkovej mapy).

$$P_i = P(r_i; c_i) \quad (6.1)$$

$$P_{i;N} = P(r_{i;N}; c_{i;N}) \quad (6.2)$$

$$P_i = R_i; G_i; B_i; d_i \quad (6.3)$$

Rozptyl polohy kľúčového bodu P_d sa vypočíta ako absolútna hodnota rozdielu maximálnej a minimálnej pozície pixelu v danom smere.

$$P_d(x) = |\max\{c_{i;N}\} - \min\{c_{i;N}\}| \quad (6.4)$$

$$P_d(y) = |\max\{r_{i;N}\} - \min\{r_{i;N}\}| \quad (6.5)$$

$$P_d(z) = |\max\{d_{i;N}\} - \min\{d_{i;N}\}| \quad (6.6)$$

Úlohou algoritmu je identifikovať, kedy rozptyl bodov P_d v každom smere neprekročí stanovenú hranicu. Identifikácia sa začína interakciou užívateľa systému (doktor) s grafickým prostredím programu. Snímaný objekt musí byť v pokojovom stave a nastavenej pozícii. Príklad RGB-D obrazov dynamického objektu, ktoré môžu byť použité na detekciu kľúčových bodov, sa nachádza na obr. 6.15.



Obr. 6.15: Prahované RGB-D snímky vytvorené multi-kamerovým systémom zachytávajúce dynamický objekt v rovnakom čase.

Pohybová aktivita je znížená vplyvom podporného vizualizačného systému (tabuľet premietajúci zaujímavý obsah), ktorého úlohou je udržať pohľad objektu koncentrovaný na jedno miesto.

Ideálny je frontálny pohľad, pretože tvár obsahuje veľa statických a jedinečných prvkov. V súčasnosti existuje niekoľko desiatok metód, ktoré sú určené na detekciu tváre a jej orientačných bodov. Ich úlohou je čo najpresnejšie identifikovať spomínané body. V tejto práci sa používa detektor DLib, ktorý pracuje na princípe súboru regresívnych stromov [63]. Výhodou použitia DLib detektora oproti SURF a SIFT metódam je, že dokážu ľahko identifikovať aj mimiku tváre. Takýmto spôsobom sa umožňuje detekcia pohybu hlavy pri rôznych polohách a tvárových výrazoch. Základný detektor identifikuje 68 bodov tváre, pričom vie ohraničiť tvár, oči, nos a ústa. Ukážka aplikácie detektora na RGB-D obraze sa nachádza na obrázku 6.16.



Obr. 6.16: **Vľavo:** Označenie 68 základných klúčových bodov DLib Face Landmark detektora [63]. **Vpravo:** Aplikácia na RGB-D obraze získanej z kamery Kinect v2.

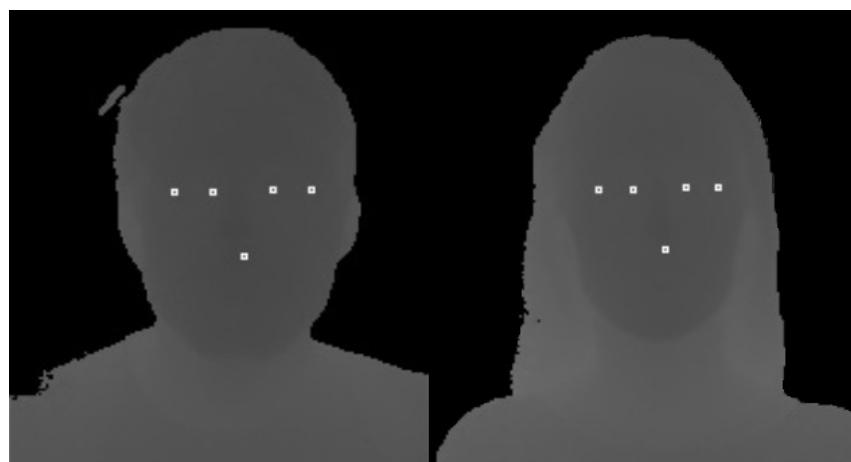
Problém nastáva pri chýbajúcich dátach v RGB-D obraze, ktoré vznikajú pri mapovaní RGB obrazu na poškodenú hĺbkovú mapu. Kvôli spracovaniu v reálnom čase je detekcia vykonávaná ešte pred filtráciou a opravou RGB-D obrazu. Preto sa k detekcii používajú ešte nefiltrované dáta. Na nich môže dochádzať k nesprávnej detekcii bodov. Problematické sú aj body 0 až 16, pretože v závislosti na fenotype tváre sa ich umiestnenie môže vyskytnúť mimo objektu.

Redukovanie počtu detekovaných kľúčových bodov na minimálny počet zabezpečí zníženie stavu, kedy bude mať bod P_i hodnotu $d_i = 0$. K nasledujúcej detekcii bol použitý detektor identifikujúci 5 kľúčových bodov.



Obr. 6.17: DLlib detektor identifikujúci 5 kľúčových bodov tváre na RGB-D obraze.

Z RGB-D obrazu (obr. 6.17) sa následne pozície kľúčových bodov r_i, c_i prenesú do hĺbkovej mapy. Získajú sa hĺbkové informácie d_i a vypočítajú sa hodnoty P_d .



Obr. 6.18: Prenesenie identifikovaných kľúčových bodov z RGB-D obrazu (Obr. 6.17) do hĺbkovej mapy.

Pre porovnanie 5 a 68 bodového detektora sa štatisticky analyzovalo, koľko detekovaných bodov malo nulovú hodnotu hĺbky $d_i = 0$. Zároveň sa zisťoval súčet hĺbkových máp, v ktorých mali detekované body 0 hodnotu $F(d_i = 0)$. V tabuľke 6.4 sú zobrazené výsledky z 87 testovacích snímok.

Tabuľka 6.4: Štatistické porovnanie 68 bodového a 5 bodového detektora na 87 hĺbkových mapách.

Model	Počet bodov	$d_i = 0$	$F(d_i = 0)$	$d_i = 0[\%]$	$F(d_i = 0)[\%]$
5 bodov	415	2	2	0,48	2,298
68 bodov	5644	41	23	0,72	26,437

Oba detektory mali veľmi nízku hodnotu prípadov, kedy identifikovanému bodu P_i chýbala hĺbková informácia. Avšak pri 68 bodovom detektore približne každý štvrtý snímok obsahoval takýto bod.

Pre porovnanie stability detekcie kľučových bodov sme vykonali meranie euklidovských vzdialenosí d_e medzi jednotlivými bodmi. Zisťovaná bola šírka ľavého a pravého oka, vonkajší a vnútorný rozostup očí. Meranie bolo vykonané na datasete, ktorý obsahoval 135 párov snímok (RGB-D, hĺbková mapa) pre oba detektory.

$$d_e(P_i, P_j) = \sqrt{(P_i(x) - P_j(x))^2 + (P_i(y) - P_j(y))^2 + (P_i(z) - P_j(z))^2} \quad (6.7)$$

Z výsledkov sa vypočítal medián, priemerná hodnota a rozdiel medzi maximálnou a minimálnou nameranou hodnotou. Údaje pre jednotlivé detektory sa nachádzajú v tabuľkách 6.5 a 6.6.

Tabuľka 6.5: Meranie vybraných faciálnych rozmerov pomocou 5 bodového detektora.

Meranie	Medián [mm]	Priemer [mm]	Rozdiel (Max-Min) [mm]
Šírka ľavého oka	28,38	28,41	5,03
Šírka pravého oka	26,21	26,17	4,39
Vnút. rozostup očí	38,43	38,22	2,32
Vonk. rozostup očí	90,06	89,89	3,22

V nich môžeme vidieť, že výsledky sú podobné. Dôležitým parametrom je rozdiel medzi maximálnou a minimálnou nameranou hodnotou. Ten hovorí o tom, ako

Tabuľka 6.6: Meranie vybraných faciálnych rozmerov pomocou 68 bodového detektora.

Meranie	Medián [mm]	Priemer [mm]	Rozdiel (Max-Min) [mm]
Šírka ľavého oka	28,46	28,52	5,40
Šírka pravého oka	27,13	27,22	4,81
Vnút. rozostup očí	37,13	37,00	3,16
Vonk. rozostup očí	85,99	89,70	4,16

stabilná je detekcia kľúčových bodov pri rôznych polohách tváre. Rozptyl detekovaných bodov je priamo úmerný veľkosti rozdielu medzi maximálnou a minimálnou hodnotou.

Tabuľka 6.7: Rozdiel výsledkov medzi 68 a 5 bodovým detektorm.

Meranie	Medián [mm]	Priemer [mm]	Rozdiel (Max-Min) [mm]
Šírka ľavého oka	-0,08	-0,11	-0,38
Šírka pravého oka	-0,93	-1,05	-0,42
Vnút. rozostup očí	1,30	1,21	-0,83
Vonk. rozostup očí	0,51	0,19	-0,93

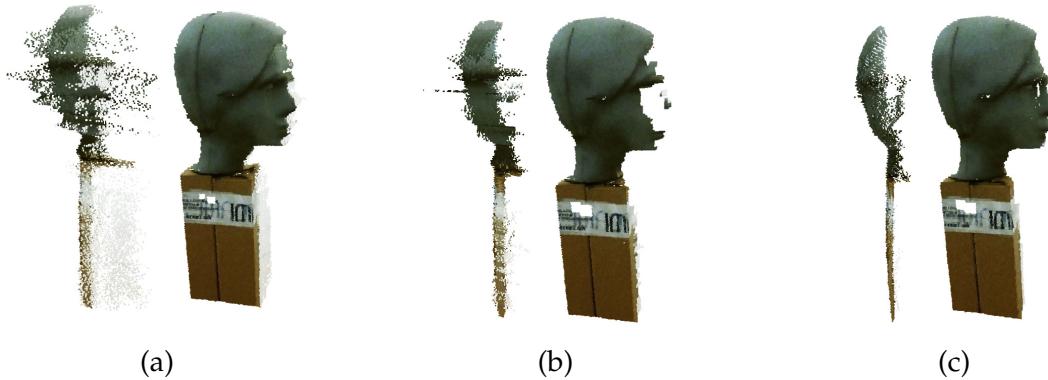
Pre lepšie porovnanie sme odčítali výsledky merania 5 bodového detektora od 68 bodového detektora. Z výsledkov vyplýva, že rozdiel medzi presnosťou jednotlivých detektorov je minimálny. Maximálna hodnota rozdielu je 1.3mm pri meraní vnútorného rozostupu očí. Z tohto experimentu vyplýva, že oba detektory sú vhodné pre meranie pohybu tváre. Každý z nich lokalizuje body pod $33ms$, čo je podmienka pri práci v reálnom čase. 68 bodový detektor má potenciálne využitie pri meraní tváre v rôznych mimikách (úsmev, ...). Tieto mimiky majú potenciál pri diagnostike OSA.

6.4.4 Filtrácia a registrácia mračna bodov

V predchádzajúcim kroku sa zabezpečilo získanie série snímok, ktoré zachytávali objekt v statickej polohe (alebo pri minimálnych pohyboch). Veľkosť série snímok určuje IMBM filter, ktorého detailný opis a optimálne nastavenia sa nachádzajú v kapitole 7. Podstatnými dátami sú hĺbkové mapy a RGB-D obrasy jednotlivých poľadov. Tieto dátá už podstúpili jednoduchú filtráciu akou bola erózia a prahovanie (pozri 6.17 a 6.18).

V ďalšom kroku sa vykonáva IMBM filtrácia, ktorej úlohou je identifikovať a odstrániť multi-kamerovú interferenciu. Multi-kamerová interferencia je v tomto prípade dôsledok využitia paralelnej spolupráce viacerých ToF kamier. Spôsobuje zašumenie hĺbkovej mapy a vznik odľahlých pixelov. IMBM filter dokáže identifikovať interferujúce miesta a odstrániť chybné pixely.

Filter pracuje s hĺbkovými mapami, výstupom z filtra je taktiež hĺbková mapa. Pre lepšiu vizualizáciu funkčnosti jednotlivých krokov IMBM sú hĺbkové mapy prevedené do mračien bodov.



Obr. 6.19: IMBM filtrácia: (a) Ukážka vplyvu interferencie. (b) Extrakcia interferova- ných miest. (c) Filtrované dátá s interpolovaním chybných miest.

Ukážka vplyvu interferencie na mračno bodov sa nachádza na obr. 6.19a. To je vytvorené zo série hĺbkových máp jednej kamery. Model obsahuje veľké množstvo chybných pixelov, ktoré sú rozptýlené okolo povrchu modelu. Navyše sa interferencia prejavuje v oblasti komplexnej geometrickej štruktúry. Na obr. 6.19b sa nachádza mračno bodov po odstránení poškodených regiónov vytvorené z jednej hĺbkovej mapy. V porovnaní s prvým modelom je vidieť, že oblasť interferencie je značná. Tento model stále obsahuje šum a interferenčné miesta, ktoré neboli identifikované IMBM filtrom. Výsledné mračno bodov (obr. 6.19c) je vytvorené zo stredných hodnôt jednotlivých filtrovaných mračien bodov. Tento model má podstatne znížený šum a počet lietajúcich pixelov bez straty interferovaných regiónov. Filtráciou však môžu v modeli vzniknúť diery. Ak sú malé a nenachádzajú sa v komplexných oblastiach, IMBM dokáže tieto hodnoty interpolovať.

Výstupný RGB-D obraz z IMBM je priemerom vstupných RGB-D obrazov. Pomocou tohto obrazu sa nanáša textúra na mračno bodov. Na obr. 6.17 je vidieť, že

vstupné obrazy obsahujú prázdne miesta (0 hodnoty pixelov) aj v oblastiach, kde hĺbka nie je nulová. Tieto miesta sú do určitej miery odstránené vo výstupnom RGB-D obraze.



Obr. 6.20: Oprava RGB-D obrazu pomocou metódy Telea : **(a)**Výstupný medián RGB-D mapy s poškodenými ROI. **(b)** Oprava farebnej informácie na miestach, kde sa v hĺbkovej mape nachádza objekt.

K úplnému odstráneniu poškodených miest používame Teleovú metódu [64]. K aplikácii použijeme hĺbkovú mapu a RGB-D obraz z IMBM filtra. Opravujeme len pixely, ktoré majú súčet R,G a B vrstiev rovný 0 a hĺbková informácia d je rôzna od 0. Výstupom je RGB-D obraz, ktorý je zbavený poškodených miest (Obr. 6.20).

Po tomto kroku máme pripravené mediány hĺbkových máp, ktoré sú použité pri meraní kranio-faciálnych parametrov tváre a pri segmentácii hĺbkovej mapy. Tieto spracovania sú opísané v samostatných kapitolách. Z týchto dát následne vytvárame mračná bodov. Pomocou nasledujúcich rovníc prebieha rekonštrukcia snímaného priestoru.

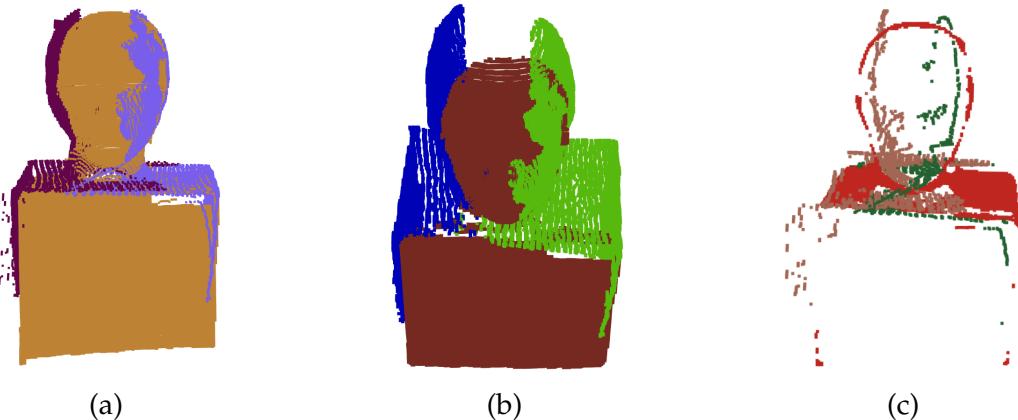
$$x = \frac{P_i(c_i) + 0.5 - c_x}{f_x} P_i(d_i) \quad (6.8)$$

$$y = \frac{P_i(r_i) + 0.5 - c_y}{f_y} P_i(d_i) \quad (6.9)$$

$$z = P_i(d_i) \quad (6.10)$$

Intrinické parametre c_x, c_y, f_x, f_y sú unikátne pre každú kameru a sú výstupom geometrickej kalibrácie kamier. Parametre $P_i(c_i), P_i(r_i)$ sa odkazujú na rovnice (6.1) až (6.3). Pomocou rotačných a translačných parametrov, ktoré boli získane pri viackamerovej kalibrácii, zarovnávame mračná bodov z jednotlivých kamier do spoloč-

nej súradnicovej sústavy. Tieto mračná sú filtrované pomocou *Radius Outlier Removal* filtra (ROR). Bližší opis tohto filtra sa nachádza v kapitole 7.6.2.

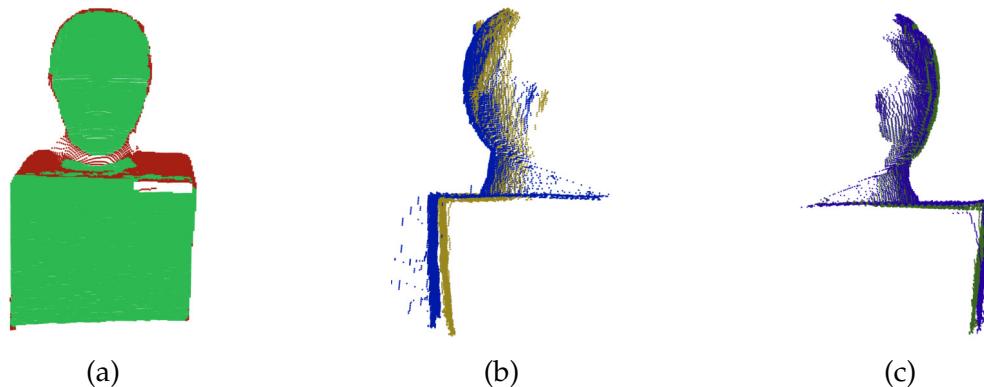


Obr. 6.21: Aplikácia ROR filtrácie s nastaveniami $rad = 0.009$ a $pts = 30$: (a) Vstupné mračno bodov vytvorené z IMBM filtrovaných hĺbkových máp. (b) Výstupné mračno bodov po ROR filtrovaní. (c) Odstránené body po ROR filtrácii.

Na obr. 6.21a, sa nachádza mračno bodov vytvorené z IMBM filtrovanej hĺbkovej mapy. Výstup ROR filtra, ktorého nastavenia boli $rad = 0.009$ a $pts = 30$, sa nachádza na obr. 6.21b. Optimálne nastavenie filtrácie je individuálne pre jednotlivé modely. Grafické prostredie programu umožňuje nastavenie parametrov a vizuálne zobrazenie filtrovaných dát.

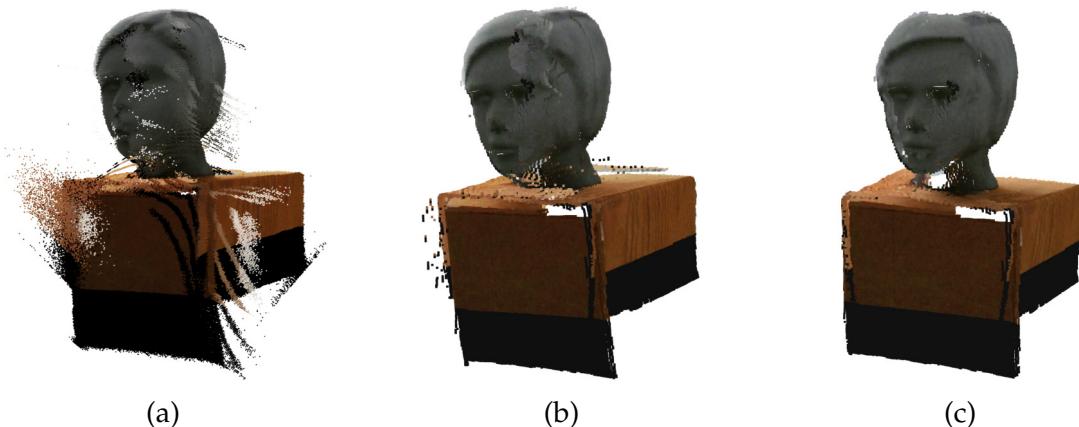
Pri multi-kamerovej kalibrácii sme získali rotačné a translačné parametre pre jednotlivé kamery. Pomocou nich dokážeme dostatočne presne zarovnať filtrované mračná do spoločného priestoru. Presnosť kalibrácie je však ovplyvnená pracovnou teplotou kamery. Jej zmena spôsobuje posun jednotlivých mračien v priestore, čím je ovplyvnená konzistentnosť modelov.

Aplikovaním ICP metódy sa zabezpečí opäťovné zarovnanie mračien bodov do spoločného priestoru. Registrácia sa vykonáva postupne medzi dvojicami podvzorkovaných mračien bodov. Predný pohľad tváre je určený ako referenčný a mračno bodov je statické. Rotačná a translačná matica má tvar identickej afinnej transformácie (rovnica 4.1). Pre ostatné mračná sú vypočítané nové rotačné a translačné parametre zložené z ostatných affinných transformácií. Vstupné nastavenia pre *Incremental registration ICP* [65] sú *MaximumIterations*= 20, *MaxCorrespondenceDistance*= 0,005 a *RansacOutlierRejectionThreshold*= 0,005. Zarovnané mračná z jednotlivých kamier sa následne spoja do jedného uceleného mračna. Kvôli vyššej hustote bodov v miestach



Obr. 6.22: Zarovnanie mračien pomocou ICP algoritmu: (a) Referenčné mračno bodov, rozdiel medzi nefiltrovaným (červená) a filtrovaným (zelená) mračnom. (b) Rozdiel medzi nezarovnaným nefiltrovaným (modrá) a zarovnaným filtrovaným (žltá) mračnom. (c) Rozdiel medzi nezarovnaným nefiltrovaným (modrá) a zarovnaným filtrovaným (zelená) mračnom.

zarovnania sa v mračne bodov tieto miesta redukujú podľa mriežky (GRID). Týmto sa získa uniformná hustota bodov v celom mračne.



Obr. 6.23: Proces filtrácie: (a) Mračna bodov získané paralelným snímaním. (b) Mračná bodov po IMBM filtrácii. (c) Mračno bodov po ROR + ICP + GRID.

Na obr. 6.23a sa nachádza mračno bodov, ktoré je vytvorené zo súrrie piatich hĺbkových máp zachytených troma ToF kamerami Kinect v2. Snímanie objektu trvalo pod $200ms$. Vplyv multi-kamerovej interferencie je však viditeľný. Na obr. 6.23b sa nachádza výstup IMBM filtra, kde je potlačený vplyv multi-kamerovej interferencie. Mračno stále obsahuje šum a lietajúce pixely. Taktiež aj zarovnanie nie je postačujúce. Následným spracovaním pomocou ROR, ICP a GRID sa odstránia chybné pixely a vytvorí sa konzistentné uniformné mračno bodov (obr. 6.23c).

6.4.5 Automatizované meranie vzdialenosť tvárových črt

Pomocou 68 bodového DLib detektora [63] je možné automatizovať meranie určitých parametrov tváre. Jednotlivé merania sú špecifikované v dotazníku využívanom pri vyšetrení na Klinike detí a dorastu Jesseniovej lekárskej fakulty v Martine, Laboratórium spánkovej medicíny (pozri prílohu xi.i). Identifikovaním špecifických kľúčových bodov tváre a prenesením týchto bodov do hĺbkovej mapy je možné pomocou euklidovskej vzdialenosť (rovnica 6.7) získať geometrické vlastnosti tvárových parametrov. Pre meranie OSA bolo vybraných týchto 19 párov kľúčových bodov:

Dĺžka nosa: Koreň nosa (28) - Špička nosa (34)

Šírka nosa: Ľavá hranica nosa (36) - Pravá hranica nosa (32)

Výška špičky nosa: Špička nosa (34) - Horná hranica pery (52)

Šírka ľavého oka: Vonkajší kútik oka (46) - Vnútorný kútik oka (43)

Šírka pravého oka: Vonkajší kútik oka (37) - Vnútorný kútik oka (40)

Výška pravého oka - 1: Vonk. horná kontúra oka (38) - Vonk. dolná kontúra oka (42)

Výška pravého oka - 2: Vnut. horná kontúra oka (39) - Vnut. dolná kontúra oka (41)

Výška ľavého oka - 1: Vonk. horná kontúra oka (45) - Vonk. dolná kontúra oka (47)

Výška ľavého oka - 2: Vnut. horná kontúra oka (44) - Vnut. dolná kontúra oka (48)

Vnútorný rozostup očí: Vnút. kútik ľavého oka (43) - Vnút. kútik pravého oka (40)

Vonkajší rozostup očí: Vonk. kútik ľavého oka (46) - Vonk. kútik pravého oka (37)

Výška pier: Horná hranica pery (52) - Dolná hranica pery (58)

Šírka pier: Ľavá hranica pery (55) - Pravá hranica pery (49)

Dĺžka tváre: Koreň nosa (28) - Spodná hranica brady (9)

Výška brady: Spodná hranica brady (9) - Dolná hranica pery (55)

Dĺžka ľavého spánku: Stred ľavého ucha (16) - Vonk. kútik ľavého oka (46)

Dĺžka pravého spánku: Stred pravého ucha (2) - Vonk. kútik pravého oka (37)

Dĺžka ľavého líca: Stred ľavého ucha (16) - Ľavá hranica pery (55)

Dĺžka pravého líca: Stred pravého ucha (2) - Pravá hranica pery (49)

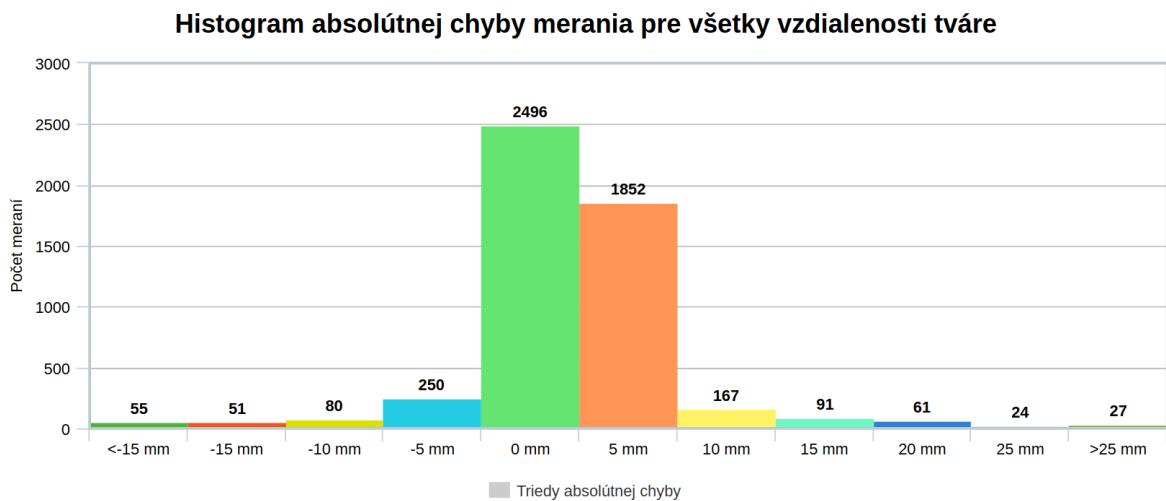
Meranie bolo vykonané na datasete 270 párov obrazov (RGB-D + hĺbková mapa).

Na detekciu kľúčových bodov bol použitý DLib model *shape-predictor-68-face-landmarks*.

Každý objekt, ktorý bol použitý pri automatizovanom meraní, mal priradené refe-

referenčné hodnoty vzdialenosť pre jednotlivé tvárové parametre.

Vypočítané hodnoty boli následne porovnané s referenčnými, čím sa získal histogram absolútnej chyby pre všetky merania (Obr. 6.24). Z celkového počtu 5130 meraní malo 4598 absolútne chybu menšiu ako 5mm. Zvyšných 532 vykazovalo vyššiu chybu merania nad 10mm.



Obr. 6.24: Histogram absolútnej chyby merania tvárových parametrov vytvorený z 5130 meraní euklidovskej vzdialenosť.

Vyšší rozdiel medzi referenčnými a nameranými hodnotami môže byť spôsobený zmenou geometrie tváre (zmena mimiky tváre). Takáto chyba sa prejavila napríklad pri výške pier a brady. Problematickými sa javí aj meranie rozmerov spánkov a lícnych častí.



Obr. 6.25: Ukážka mračna bodov s vyznačenými kľúčovými bodmi: (a) Všetky detekované body boli pod 5mm hranicou absolútnej chyby. (b) Červený bod označuje nepresnú detekciu kľúčového bodu.

Ich príslušné kľúčové parametre sa pri frontálnom pohľade nachádzajú na hranici hĺbkovej mapy. Preto aj pri minimálnej nepresnosti určenia pozície stredov pravého (2) a ľavého ucha (16) môže dôjsť k výraznejšej zmene hĺbky (Obr. 6.25).

Stredné a priemerné hodnoty sú vypočítané z rozdielov výsledných vzdialenosí ($d_e - d_e(\text{ref})$). Mínusové hodnoty znamenajú, že automaticky vypočítaná vzdialenosť je menšia ako referenčná.

Tabuľka 6.8: Meranie vybraných faciálnych rozmerov pomocou 68 bodového detektora.

Meranie	Medián [mm]	Average [mm]	Diff (Max-Min) [mm]
Dĺžka nosa	0.19	0.45	9.52
Šírka nosa	0.12	-0.27	10
Výška špičky nosa	-0.12	-0.08	6.55
Šírka ľavého oka	0.03	0.00	7.26
Šírka pravého oka	0.03	0.09	6.90
Výška pravého oka - 1	0.07	0.20	12.28
Výška pravého oka - 2	-0.10	0.08	9.56
Výška ľavého oka - 1	0.13	0.15	8.52
Výška ľavého oka - 2	-0.13	0.08	9.90
Vnútorný rozostup očí	-0.24	-0.28	4.34
Vonkajší rozostup očí	-0.89	-0.88	7.73
Výška pier	0.77	0.42	11.47
Šírka pier	-0.19	0.04	9.07
Dĺžka tváre	-1.82	-1.81	9.28
Výška brady	-3.84	-3.90	11.71
Dĺžka ľavého spánku	1.47	3.06	29.30
Dĺžka pravého spánku	-10.14	-7.35	90.66
Dĺžka ľavého líca	-2.07	-0.80	31.88
Dĺžka pravého líca	5.40	2.36	69.38

Meranie vzdialenosí tvárových parametrov, ktorých kľúčové body sa nenachádzajú na hraniciach hĺbkovej mapy, vykazujú malú absolútну chybu voči referencii. Problémom je získanie všetkých parametrov z frontálneho pohľadu. Tento problém je však možné vyriešiť dvoma meraniami, kedy poloha tváre pacienta bude vždy mierne natočená. Takéto euklidovské meranie však ukazuje svoje limity pri diagnostike OSA (pozri 2.1.3). V ďalšej časti sa preto opisuje návrh, ktorý je založený na nových modernejších metódach a môže priniesť lepšie výsledky.

6.4.6 Segmentácia hĺbkovej mapy

Pre segmentáciu tvárových častí sa používa hĺbková mapa a RGB-D obraz získaný z IMBM filtra. Pomocou *Mask R-CNN* neurónovej siete dokážeme v RGB-D obraze ohraničiť príslušné tvárové časti [66]. Následným prenesením identifikovaných regiónov do hĺbkovej mapy vieme konkrétnym pixelom nastaviť klasifikátor, ktorý ich priraduje do preddefinovaných skupín.

Pre trénovanie tejto siete sme vytvorili dátový set, ktorý pozostával z 65 RGB-D obrazov obsahujúcich 6 rôznych osôb v rozličných polohách. 50 snímok 4 osôb boli použité pre trénovanie a zvyšných 15 obrazov 2 osôb pre testovanie. Vo všetkých obrazoch sa označili regióny prislúchajúce jednotlivým klasifikačným triedam.

Klasifikované boli tieto oblasti: tvár, nos, ústa, brada, krk, ucho, oko, obočie.

Vyznačené boli základné časti tváre pomocou nástroja *Via Image Annotator* [67]. Ukážka RGB-D obrazu s označenými oblasťami sa nachádza v prílohe xi.v. Snahou bolo zachytiť tieto oblasti z rôznych uhlov tak, aby boli rozpoznateľné pri použití multi-kamerovej topológie a pri rôznych pozíciach natočenia hlavy.

Vychádzali sme z modelu, ktorý už podstúpil trénovanie na ImageNet obrazovom datasete. Tento model bol dotrénovaný na našich dátach. Trénovanie prebiehalo v 60 epochách s 1000 iteráciami. Bola použitá *resnet101* sieťová architektúra. Výsledky trénovania a chyby konkrétnych vrstiev siete sa nachádzajú v tabuľke 6.9.

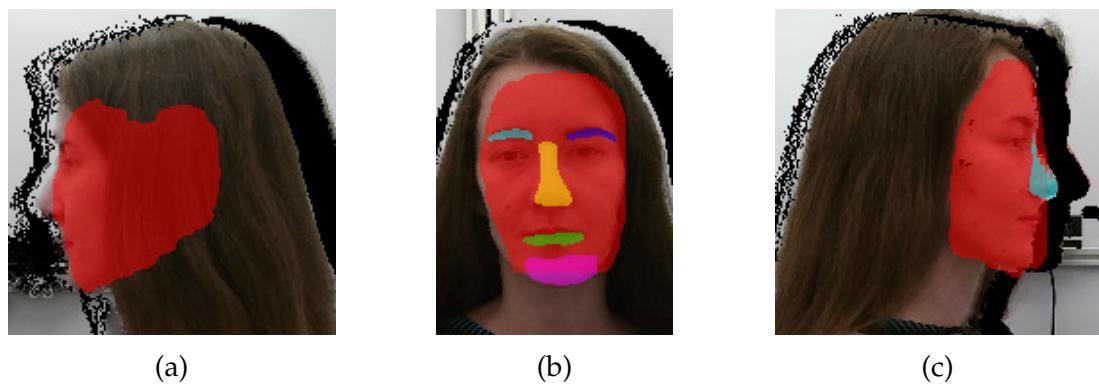
Tabuľka 6.9: Vyhadnotenie trénovania Mask R-CNN siete.

Strata	30 epocha	60 epocha	validácia
loss	5,4736	4,579	4,698
rpn class loss	0,6676	0,435	0,205
rpn bbox loss	2,4274	2,144	2,596
mrcnn class loss	0,6545	0,394	0,264
mrcnn bbox loss	1,0124	0,894	0,939
mrcnn mask loss	0,7118	0,712	0,692

Celková stratová funkcia L (*loss*) predstavuje súčet chyby klasifikácie L_{cls} (*class loss*), lokalizácie L_{box} (*bbox loss*) a segmentácie L_{mask} (*mask loss*).

$$L = L_{cls} + L_{box} + L_{mask} \quad (6.11)$$

Pre overenie funkčnosti sme klasifikátor použili na nových objektoch. Výsledky segmentácie sa nachádzajú na obr. 6.26. Z výsledkov segmentácie vidíme, že nie všetky oblasti boli klasifikované. Niektoré ROI masky boli relatívne dobre segmentované (tvár, nos, ústa). Určité oblasti však chýbali úplne (oči, krk). Predpokladom je, že rozšírením trénovacieho datasetu o nové RGB-D obrazy sa môže výrazne zvýšiť presnosť klasifikácie a segmentácie. Prenos segmentovaných častí na mračno bodov je znázornené na obr. 6.27.



Obr. 6.26: Ukážka segmentácie RGB-D obrazu použitím Mask R-CNN siete: (a) Ľavý pohľad, identifikovaná tvár. (b) Predný pohľad, identifikované oboče, tvár, nos, ústa, brada. (c) Ľavý pohľad, identifikovaná tvár a nos.

Výhodou tohto prístupu je, že dokážeme spojiť jednotlivé klasifikované regióny z viacerých pohľadov. Tieto spojené oblasti si navyše uchovávajú aj svoje geometrické vlastnosti (napr. rozostup uší, šírka brady). Pričom pri segmentácii z jednotlivých pohľadov bez spájania do spoločného priestoru by sa táto podstatná informácia strácalá.



Obr. 6.27: Mračno bodov dynamického objektu so zvýraznenými klasifikovanými oblasťami.

Zlepšenie segmentácie je cieľom budúcej práce. Pre tieto účely bol navrhnutý snímací systém pracujúci s novšími kamerami Intel RealSense d415, ktorý je umiestnený v detskom spánkovom laboratóriu na klinike detí a dorastu v Martine. Ten zabezpečuje zber hĺbkových a RGB-D obrazov pediatrických pacientov z troch perspektív. Spolu s týmito dátami sú ukladané aj informácie o klinickom vyšetrení snímaného pediatrického pacienta (obr. xi.vii.). Ukážka snímacieho zariadenia sa nachádza na obr xi.ix. Výstupné dáta budú použité k diagnostike OSA pomocou PointNet2 neurónovej siete [68].

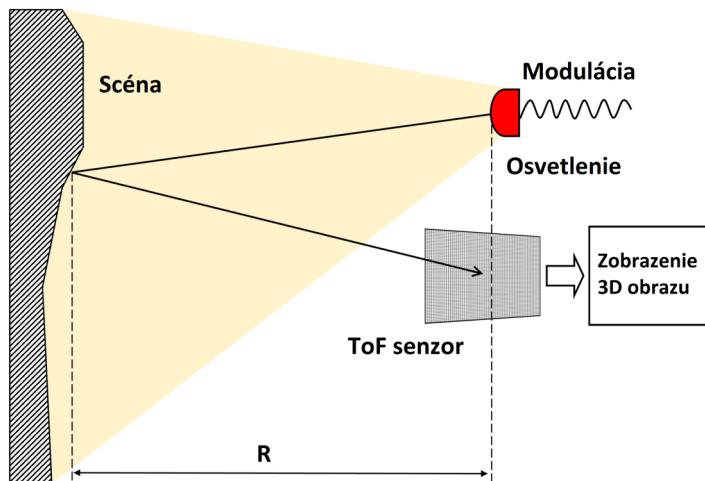
7. Potláčanie multi-kamerovej interferencie ToF senzorov

Jedným z nevýhod použitia ToF senzorov je ich multi-kamerová interferencia (MCI). Ide o chybu, ktorá ma fyzikálnu podstatu a má veľmi negatívny dopad na výstupné dátu. Ak pri snímaní scény nastane táto interferencia, skenovací systém sa stáva nepoužiteľným.

Jednou z možnosti riešenia problému je použitie kamier, ktoré pracujú na rozdielnych modulačných frekvenciach emitovaného svetla. Mnoho nedávnych prác zaoberejúcich sa problémom multi-kamerovej interferencie uvádzajú, že ToF senzory sa nesmú prevádzkovať s rovnakou modulačnou frekvenciou [69]. Spoliehajú sa na ortogonálne funkcie, ako sú sínusoidy s rôznou frekvenčnou moduláciou. Toto riešenie je však obmedzujúce, pretože vzniká problém softvérovej kompatibility s pripojením rôznych typov kamier v jednej aplikácii [70, 71, 72]. Pomocou metódy LSENS je možné obnoviť informácie o hĺbke a amplitúde. Táto metóda je založená na analýze štatistických vlastností interferencie signálu [73]. Interferenčný signál však musí mať kladné a záporné pulzovanie hĺbkovej mapy. V situáciách, keď sú kamery namierené proti sebe, môže byť kolísanie signálu extrémne vysoké. V takom prípade nebude metóda LSENS účinná. Z analýzy výstupných dát je však možné určiť podmienky potlačenia interferencie a podľa nich navrhnúť filtračný algoritmus. Úlohou filtra je odstrániť čo najviac poškodených dát pri čo najmenšej strate relevantných dát.

7.1 Model ToF kamery

ToF senzory sú optické snímače, ktoré poskytujú informácie o hĺbke scény. Obsahujú aktívny svetelný zdroj, ktorý generuje amplitúdovo modulovaný signál. Signál môže mať spojity alebo impulzný charakter. Väčšina ToF kamier vyžaruje amplitúdovo modulovanú kontinuálnu vlnu (AMCW) s frekvenciou blízkou IR na osvetlenie scény. Meranie hĺbky je založené na meraní amplitúdy fázového posunu vysielaného a prijímaného modulovaného signálu (Obr. 7.1). Informácie o hĺbke pre každý pixel sa môžu vypočítať pomocou synchronného demodulovania prijatého modulovaného svetla v detektore. Demodulovanie sa môže uskutočniť preložením s pôvodným modulovaným signálom. Tento proces sa nazýva krížová korelácia. Všeobecne je korelačná funkcia definovaná rovnicou 7.1.



Obr. 7.1: Princíp činnosti merania fázového posunu pri ToF kamerách [73].

$$c(\tau) = s(t) \otimes g(t) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} s(t) \cdot g(t + \tau) dt \quad (7.1)$$

kde $s(t)$ je prijatý optický signál a $g(t)$ je vyžiarený (pôvodný) signál. Použitím špecifických funkcií pre kamery ToF dostaneme rovnice

$$g(t) = \cos \omega t \quad (7.2)$$

$$s(t) = 1 + a \cdot \cos(\omega t - \varphi) \quad (7.3)$$

$$c(\tau) = \varphi_{sg}(\tau) = \frac{a}{2} \cdot \cos(\varphi + \omega\tau) \quad (7.4)$$

kde a je amplitúda modulácie a φ je fázový posun. Táto funkcia je vypočítaná pre štyri rôzne ωt argumenty, ktoré sú posunuté z 0 o 90° . Prijatý signál je väčšinou navrstvený na pozadí obrazu, čo vyžaduje pridanie offsetu b do korelačnej funkcie:

$$C(\tau) = c(\tau) + b \quad (7.5)$$

$$\begin{aligned} C(\tau_0) &= c(\tau_0) + b = \frac{a}{2} \cdot \cos(\varphi) + b \\ C(\tau_1) &= c(\tau_1) + b = -\frac{a}{2} \cdot \sin(\varphi) + b \\ C(\tau_2) &= c(\tau_2) + b = -\frac{a}{2} \cdot \cos(\varphi) + b \\ C(\tau_3) &= c(\tau_3) + b = \frac{a}{2} \cdot \sin(\varphi) + b \end{aligned} \quad (7.6)$$

S týmito štyrmi vybranými bodmi je možné vypočítať korelačnú funkciu a určiť fázu φ a amplitúdu a prijatého signálu $s(t)$:

$$\varphi = \text{atan} \left[\frac{C(\tau_3) - C(\tau_1)}{C(\tau_0) - C(\tau_2)} \right] \quad (7.7)$$

$$a = \frac{\sqrt{[C(\tau_3) - C(\tau_1)]^2 + [C(\tau_0) - C(\tau_2)]^2}}{2} \quad (7.8)$$

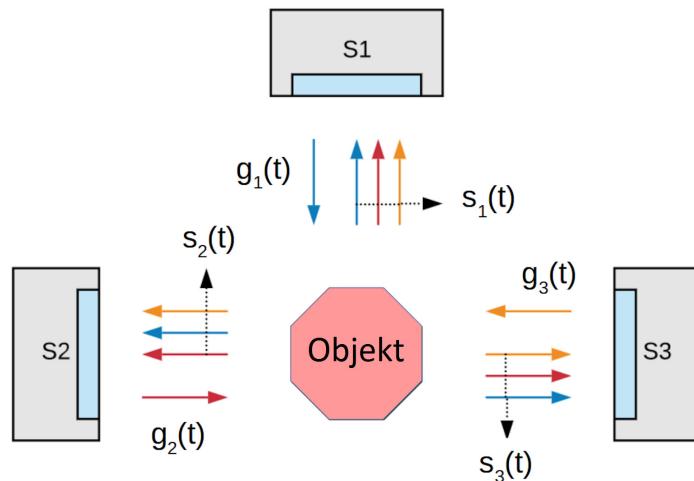
Hĺbka d sa vypočíta podľa nasledujúcej rovnice:

$$d = \frac{c \cdot \varphi}{2 \cdot 2\pi f} \quad (7.9)$$

kde c je rýchlosť svetla a f je frekvencia modulácie IR [74].

7.2 Model multi-kamerovej interferencie

V multi-kamerovom režime každý ToF senzor používa rovnakú modulačnú frekvenciu a IR vlnovú dĺžku, takže prijaté signály interagujú navzájom. V experimentálnej topológií senzorov používame tri ToF senzory S1 - S3, ktoré sú statické a umiestnené v skenovacej kabíne. Tento systém je znázornený na obrázku 7.2, ktorý popisuje vzájomné vzťahy medzi jednotlivými kamerami a objektom.



Obr. 7.2: Model multi-kamerového systému s interferenciou IR signálu.

Každá kamera generuje signál $g(t)$ a prijíma odrazený signál $s(t)$, ktorý predstavuje spojenie generovaných čiastkových signálov zo všetkých kamier S . Tieto signály možno opísť takto:

$$\begin{aligned} s_1(t) &= s_{11}(t) + s_{12}(t) + s_{13}(t) \\ s_2(t) &= s_{21}(t) + s_{22}(t) + s_{23}(t) \\ s_3(t) &= s_{31}(t) + s_{32}(t) + s_{33}(t) \end{aligned} \quad (7.10)$$

Každý čiastkový signál možno zapísť ako:

$$s_{xy}(t) = b_{xy} + a_{xy} \cdot \cos(\omega t - \varphi_{xy}) \quad (7.11)$$

Index x predstavuje cieľovú kameru a index y kameru pôvodného signálu. Nasledujúce rovnice slúžia ako príklad výpočtu modelu pre jeden signál s_1 . Pre ostatné kamery sa matematický model odvodzuje rovnakým spôsobom:

$$\begin{aligned} s_1(t) &= b_{11} + a_{11} \cdot \cos(\omega t - \varphi_{11}) + b_{12} + a_{12} \cdot \cos(\omega t - \varphi_{12}) \\ &+ b_{13} + a_{13} \cdot \cos(\omega t - \varphi_{13}) = \tilde{b}_1 + \tilde{a}_1 \cdot \cos(\omega t - \tilde{\varphi}_1) \end{aligned} \quad (7.12)$$

Po 4-fázovej korelácií rušivých signálov sa rovnice 7.7 a 7.8 transformujú do nasledujúcej podoby:

$$\tilde{\varphi}_1 = \text{atan} \left[\frac{a_{11} \sin(\tau_{11}) + a_{12} \sin(\tau_{12}) + a_{13} \sin(\tau_{13})}{a_{11} \cos(\tau_{11}) + a_{12} \cos(\tau_{12}) + a_{13} \cos(\tau_{13})} \right] \quad (7.13)$$

$$\tilde{a}_1 = \sqrt{\frac{a_{11}^2 + a_{12}^2 + a_{13}^2 + 2\bar{a}}{2}} \quad (7.14)$$

kde

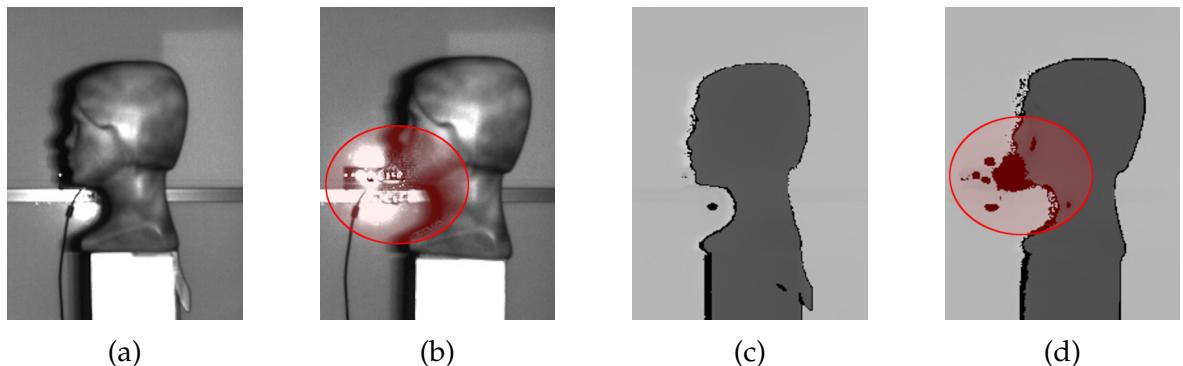
$$\begin{aligned} \bar{a} &= a_{11}a_{12}[\sin(\varphi_{11} + \varphi_{12}) + \cos(\varphi_{11} + \varphi_{12})] + \\ &a_{11}a_{13}[\sin(\varphi_{11} + \varphi_{13}) + \cos(\varphi_{11} + \varphi_{13})] + \\ &a_{12}a_{13}[\sin(\varphi_{12} + \varphi_{13}) + \cos(\varphi_{12} + \varphi_{13})] \end{aligned} \quad (7.15)$$

Toto zmiešanie signálov z rôznych kamier spôsobuje značné chyby merania a preto mapa výstupnej hĺbky obsahuje artefakty [73].

7.3 Analýza multi-kamerovej interferencie

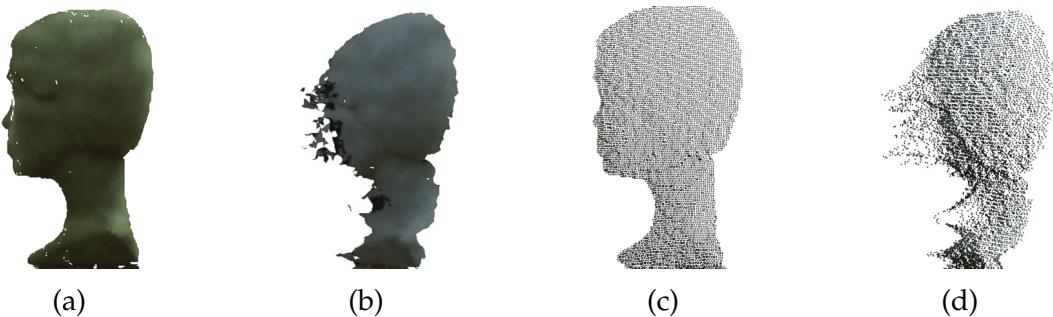
Rovnice 7.13 a 7.14 sú platné, ak každý prijatý signál $s_x(t)$ je kombináciou čiastkových signálov všetkých kamier S1 - S3, kde x sa pohybuje od 1 do 3. Interakcia signálov úzko súvisí aj so snímanou scénou. Ak si predstavíme priestorové rozloženie kamier, ktoré je znázornené na obr. 7.2, môžeme popísať určité situácie nastávajúce pri skenovaní. Ak sa v scéne nenachádza objekt, je veľmi veľká pravdepodobnosť že k interferencii dôjde medzi kamerami S2 a S3. Taktiež ak je veľkosť objektu malá a vyžarovanému IR svetlu nebude v ceste stáť žiadna prekážka. Naopak veľký objekt môže zabrániť interferencii medzi kamerami S2 a S3. Veľmi rizikové sú aj predmety a povrchy, ktoré reflektujú IR svetlo. Pri skenovaní dynamických objektov sa môže vplyv rušenia meniť, takže filtračný algoritmus musí reagovať na zmeny v čo najkratšom čase.

Ako už bolo spomenuté, výstupom ToF kamier je IR a hĺbkový obraz. Hĺbkový obraz obsahuje 3D informácie v 2D rovine obrazu. Hodnota jednotlivých pixelov (32-bit float) predstavuje absolútну vzdialenosť. Na základe týchto informácií je možné znova preniesť naskenovanú scénu, kde musia byť známe interné a externé parametre IR kamery (pozri kapitolu 4).



Obr. 7.3: Obrazy zo senzora Microsoft Kinect v2: (a) IR obraz bez interferencie. (b) IR obraz s interferenciou. (c) Hĺbková mapa IR obrazu (a). (d) Hĺbková mapa IR obrazu (b). Miesto interferencie je zvýraznené červenou farbou.

Interferencia sa prejavuje na IR obraze, čo má priamy dopad na výsledný hĺbkový obraz (pozri obr. 7.3). V hĺbkovom obraze sa v poškodenom mieste objavia pixely, ktoré naberajú extrémne odlišné hodnoty od tých skutočných. Tie sú buď negatívne alebo pozitívne, pričom sa zvyknú pulzovaním striedať.



Obr. 7.4: Rekonštruovaný 3D model statického objektu: (a) 3D trojuholníková sieť objektu bez interferencie. (b) 3D trojuholníková sieť objektu s interferenciou. (c) Mračno bodov objektu bez interferencie. (d) Mračno bodov objektu s interferenciami.

Pri pozitívnej interferencii môže ísť hĺbka do maximálnych možných hodnôt (32-bit float), pri negatívnej môže ísť do 0. Tieto miesta sú zároveň obkolesené deformovanými pixelmi, ktoré nenaberajú extrémne hodnoty, ale stále výrazne zhoršujú kvalitu výstupného modelu. Grafické znázornenie v 3D priestore je možné vidieť na

obrázkoch 7.4a a 7.4c. Práve tieto miesta sú veľmi kritické, pretože ich identifikácia je problematická.

7.4 Filtračné metódy

Existuje niekoľko prehľadových článkov, ktoré sa zaobrajú technikami filtrácie a obnovy používanými pri 3D zobrazovaní a skenovaní. Zo všetkých z nich sa práca [75] javí ako veľmi užitočná, pretože obsahuje prehľad približne 40 metód a rozdeľuje ich do niekolkých tried. Mnohé z týchto metód sú použiteľné na 3D mračná bodov, avšak žiadny známy nie je priamo určený pre potláčanie interferencie. Na tento účel bol navrhnutý IMBM filter, ktorý sa môže zaradiť medzi štatistické filtre. Pre porovnanie sú v tejto práci opísané ďalšie filtračné metódy založené na štatistike (SOR, ROR) alebo na princípoch umelej inteligencie (PointCleanNet).

7.4.1 SOR filter

Filtračná metóda SOR (Statistical Outlier Removal) je založená na štatistickej analýze okolia každého bodu. Odľahlá hodnota sa dá klasifikovať ako stratený alebo izolovaný bod, poprípade ako súbor bodov v mračne. V SOR metóde sa určuje počet bodov, ktoré sa budú považovať za susedov (definovaný počet k najbližších susedov [76]) a potom sa vypočíta priemerná vzdialenosť každého bodu od jeho susedov. Body, ktorých stredné vzdialosti nespĺňajú kritériá a sú mimo interval sa považujú za odľahlé. Interval je definovaný strednou hodnotou stredných vzdialenosťí a štandardnou odchýlkou [77]. SOR je široko používaná a efektívna metóda, aj keď vo veľkých množinách 3D dát je časovo neefektívna [78].

7.4.2 ROR filter

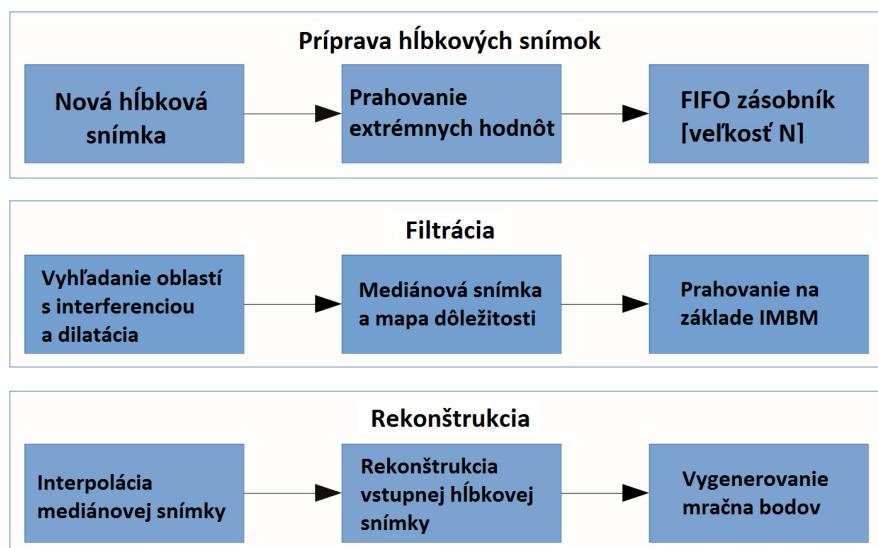
Metóda ROR (Radius Outlier Removal) predstavuje pomerne jednoduchý štatistický filter. Do filtra sa zadávajú parametre sférického polomeru rad a minimálneho počtu susedov pts . Filter odstráni body, ktoré majú v oblasti polomeru rad menej susedných bodov ako bolo nastavené parametrom pts .

7.4.3 PointCleanNet

PointCleanNet je dvojfázový algoritmus hĺbkového učenia, ktorý odstraňuje odľahlé hodnoty a znižuje šum v neusporiadaných mračnách bodov. Architektúra je založená na PCPNet (prístup založený na hlbokom učení na odhad lokálnych 3D vlastností v mračnach bodov [79]) s niekoľkými úpravami. V prvom kroku sa odstraňujú odľahlé body, v druhom sa odhadujú korekčné vektory pre zostávajúce body. Trénovanie tejto siete je časovo náročné, taktiež je vyžadované použitie CUDA.

7.4.4 Importance Map Based Median filter

Nami navrhnutý filtračný algoritmus je založený na identifikácii, extrakcii a interpolácií interferenčnej oblasti zo sekvencie hĺbkových máp získaných z ToF senzorov. Cieľom nie je poškodené miesta úplne odstrániť, ale čo najviac minimalizovať vplyv interferencie bez veľkej straty užitočných dát. Primárne využitie filtra je pre aplikáciu skenovania ľudských hláv, ale nie je obmedzené iba na túto činnosť. Základnou požiadavkou je získať dostatočný počet hĺbkových máp do obrazového zásobníka, pričom snímaná scéna nebola výrazne zmenená. Táto podmienka je splnená, ak je snímaný objekt statický alebo v pokoji (pacient nie je v strese). V prípade dynamickej scény je žiadúca minimalizácia veľkosti zásobníka, inač by mohlo dôjsť k vytvoreniu pohybových artefaktov v hĺbkovej mape.



Obr. 7.5: Blokový diagram navrhnutého IMBM filtra.

V prvom kroku je potrebné naplniť obrazový zásobník, ktorého veľkosť je špecifikovaná podľa danej snímanej scény a intenzity interferencie. Plnenie funguje na princípe FIFO. Najprv sa zásobník naplní hĺbkovými mapami a po naplnení zásobníka každý nový snímok vylúči ten najstarší, čím je zabezpečená stála veľkosť zásobníka. Všetky vstupujúce hĺbkové mapy sú prahované podľa prahových hodnôt určených topológiou a vlastnosťami systému.

Ak $d(x; y)$ predstavuje jednotlivý pixel hĺbkovej mapy, tak je prahovanie vykonávané pomocou rovnice 7.16.

$$d'(x; y) = \begin{cases} d(x; y), & \text{if } d(x; y) \in [T_{LOW}; T_{HIGH}] \\ 0, & \text{v inom prípade} \end{cases} \quad (7.16)$$

kde $d(x; y)$ je 32-bitové float číslo a T_{LOW} , T_{HIGH} sú dolné a horné prahové hodnoty. V našom prípade sú nastavené na 80 a 1300 (tieto čísla predstavujú skutočnú vzdialenosť v mm). Skenovacia kabína je znázornená na obrázku 7.9d a rozmer spodnej časti je 1520×1520 mm.

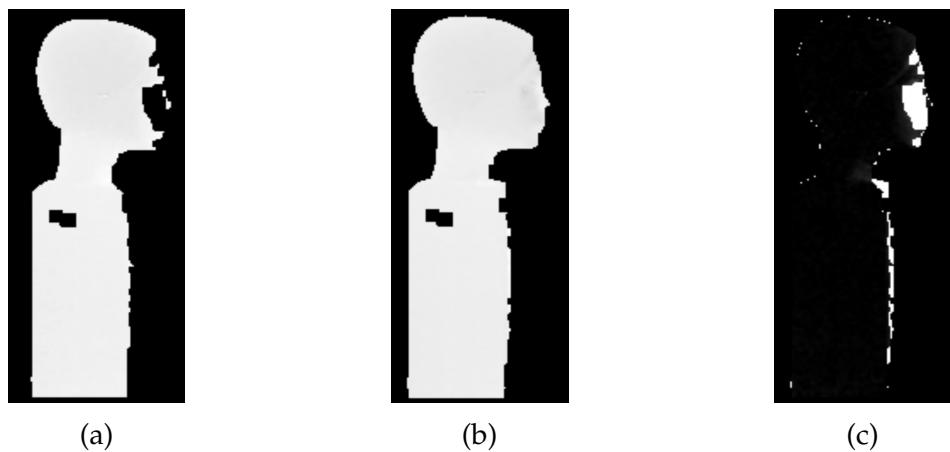


Obr. 7.6: Vplyv prahovania na interferovanú hĺbkovú mapu (a) Hĺbková mapa s interferovanými regiónymi. Pixely nadobúdajú extrémne hodnoty. (b) Hĺbková mapa bez extrémnej interferencie.

Po prahovaní hĺbky sa pozadie odstráni a obrázky obsahujú iba predmet záujmu (Obr. 7.6). Tento jednoduchý krok je dôležitý a účinný, pretože odstraňuje oblasti so spomínanou extrémnou interferenciou a pozadia.

V ďalšom kroku sa z obrazového zásobníka vyberie ako referenčný obraz filtrovaná hĺbková mapa s minimálnym počtom nenulových hodnôt. Je to preto, že tento

obrázok s najväčšou pravdepodobnosťou obsahuje minimálny počet rušivých pixelov. Z tejto referencie je vytvorená binárna mapa a každá hĺbková mapa je s ňou maskovaná. V tomto kroku sú pixely vstupných máp rozdelené do troch oblastí: pozadie (vrátane extrémnych interferenčných oblastí), pixely interferenčnej oblasti (tiež vrátane niektorých pixelov objektu) a oblasť objektov (vrátane interferenčných pixelov).



Obr. 7.7: Identifikácia interferovaných regiónov v hĺbkových mapách. (a) Referenčná mapa s maximálnym počtom 0 pixelov.(b) Vstupná mapa s interferenciou. (c) Mapa zobrazujúca možné interfereované regióny obrazu b.

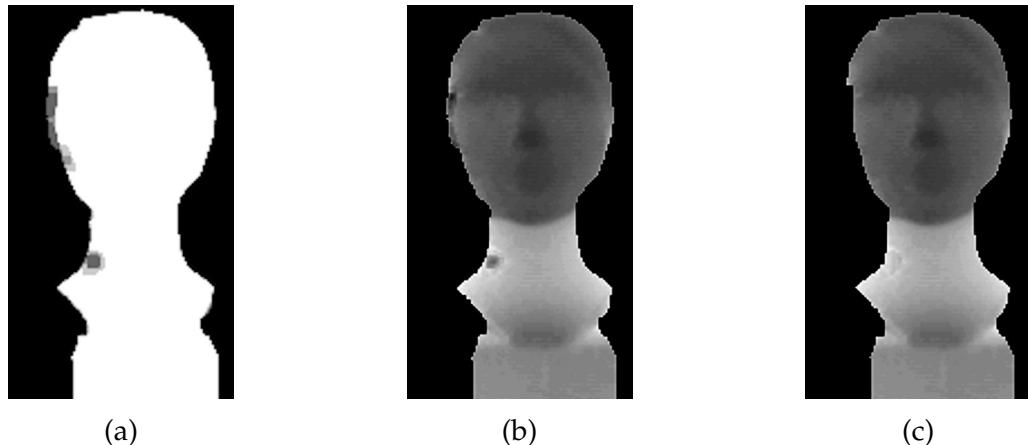
Výsledná rozdielová maska je ešte dilatovaná prvkom o štvorcovej štruktúre 1×1 , čím sú odstraňované oblasti s nízkou intenzitou interferencie. Týmto spôsobom sa zníži celkový súčet interferenčných pixelov, ale zároveň znižujeme aj súčet správnych obrazových pixelov skenovaného objektu.

Tieto morfologicky spracované snímky sa používajú na výpočet mediánu hĺbkovej mapy a na vytvorenie mapy dôležitosti pixelov.

Pixel mediánu hĺbkovej mapy je vypočítavaný ako medián všetkých konkrétnych nenulových pixelov (rovnaká pozícia riadku a stĺpca v obraze) hĺbkových máp v zásobníku.

Mapa dôležitosti pixelov nám hovorí, z kolkých hodnôt boli vypočítané konkrétné pixely v mediáne hĺbkovej mapy. Malá hodnota dôležitosti daného pixelu znamená vysokú pravdepodobnosť interferencie vyskytujúcej sa v určitom bode. Z tohto

dôvodu sa zdá byť užitočné odstrániť také oblasti nastavením prahu dôležitosti.



Obr. 7.8: Filtrácia a interpolácia hĺbkových máp. (a) Mapa dôležitosti. (b) Medián. (c) Interpolovaná mediánová hĺbková mapa.

Pixely v mediáne (Obr. 7.8b) s vybranou dôležitosťou vyššou ako $k \cdot N$ sú zachované, ostatné pixely sú nastavené na nulu. K predstavuje prahovú hodnotu dôležitosti mapy a N je veľkosť obrazového zásobníka. Koeficient k môže mať rozsah hodnôt od 0 do 1. Tento krok je posledný, znižuje vplyv interferencie na hĺbkové mapy. Ostatné časti filtračného algoritmu zabezpečujú rekonštrukciu filtrovaných dát, ale nie sú nevyhnutné.

Interpolácia filtrovaných miest je nasledujúcim krokom IMBM. Rekonštrukcia dát využíva bilineárnu interpoláciu [80], pričom tá je vykonávaná iba na špecifických filtrovaných miestach spĺňajúcich definované podmienky. K obmedzeniu interpolácie dochádza na neuzavretých dierach, pri veľkých dierach a pri vysokých rozdieloch medzi interpolačnými pixelmi. Bilineárna interpolácia môže byť nahradená inou metódou interpolácie (napr. Bikubická, biquadratická ...). Z hľadiska výpočtu je bilineárna transformácia dobrým kompromisom medzi metódou najbližšieho suseda a inými zložitejšími metódami interpolácie.

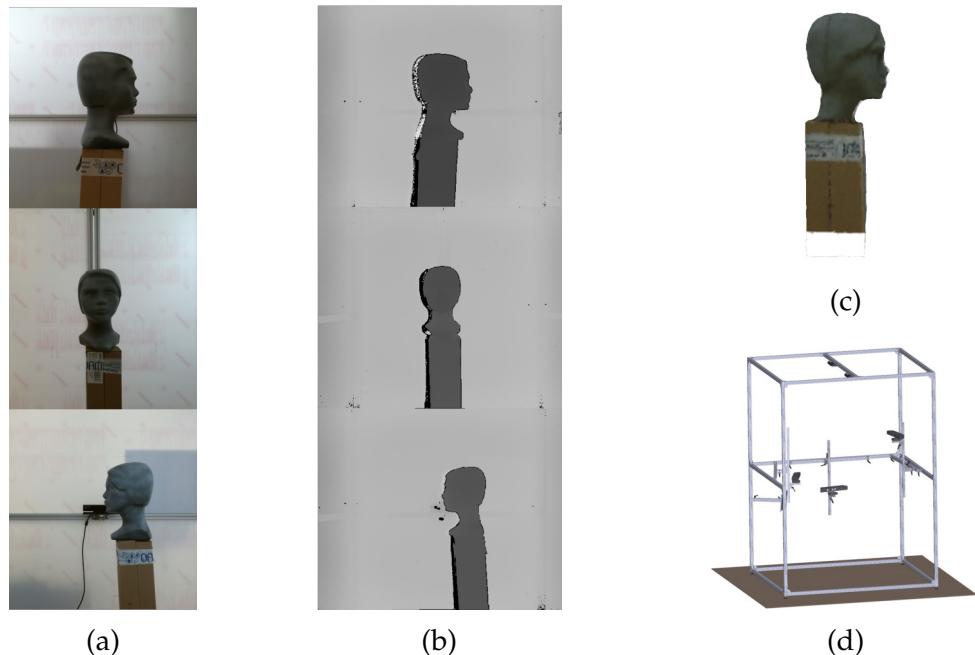
Mediánom môžeme opraviť vstupné hĺbkové mapy. Tým opravíme nie len interferované miesta ale aj znížime vplyv šumu kamier. V procese opravy sa pixely hĺbkových máp porovnávajú s pixelmi mediánu. Ak je absolútна hodnota rozdielu väčšia ako zvolená šumová prahová hodnota (táto premenná je diskutovaná v experimentálnej časti 7.6.4), pixel v pôvodnej hĺbkovej mape sa nahradí mediánovou hodnotou (alebo nulovou hodnotou, v závislosti od režimu metódy IMBM).

Hodnota šumového prahu závisí od intenzity šumu kamery pre definovanú vzdialenosť. Ak je táto hodnota príliš nízka, veľký počet neinterferujúcich pixelov bude nahradený strednou hodnotou. Na druhej strane by filter mohol zachovať veľké množstvo rušenia v pôvodnom obraze.

Všeobecne možno povedať, že použitie navrhovaného algoritmu zlepšuje hĺbkový obraz (Obr. 7.8c) znížením lietajúcich pixelov, šumových procesov a vyplnením dier vo výslednom 3D modeli.

7.5 Nastavenie experimentu

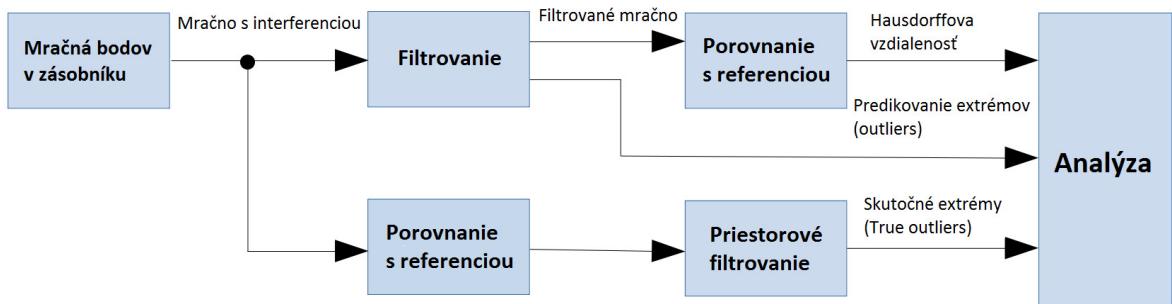
Na overenie funkčnosti vybraných metód filtrácie sme vykonali experiment v reálnych podmienkach na statickom objekte. V experimente boli použité triToF kamery typu Microsoft Kinect v2. Tie boli umiestnené do skenovacej kabíny, ktorej návrh je zobrazený na obrázku 7.9d. V strede kabíny je modelový objekt umiestnený v rozsahu od 60 cm do 80 cm od senzorov. Predmetom bola plastová hlava s pomerne komplexným tvarom. Veľkosť objektu zodpovedá normálnym rozmerom dieťaťa vo veku 5-10 rokov.



Obr. 7.9: Referenčné obrazy multi-kamerovej sústavy: (a) RGB snímky. (b) Hĺbkové mapy. (c) 3D model vytvorený z hĺbkových máp. (d) Konštrukcia skenovacej kabíny s ukážkou rozloženia kamier.

RGB obrázky prijaté z každej kamery sú znázornené na obrázku 7.9a, hĺbkové snímky snímanej scény bez rušenia (postupne pracujúce kamery) pre jednotlivé kamery sú uvedené na obrázku 7.9b. Tieto hĺbkové mapy sa použili ako referenčné údaje bez interferenčných artefaktov. Obrázok 7.9c je rekonštruovaný 3D model statického objektu.

Aby bolo možné objektívne vyhodnotiť výkon potlačenia rušenia, musí sa identifikovať povaha každého bodu v mračne bodov. Zisťuje či ide alebo nejde o skutočné rušenie. Dáta skreslené interferenciou sa získali v režime paralelného snímania. Pre získanie tejto informácie sme použili priestorové porovnávanie interferova- ných dát voči referenčným. Referenčná sieť, znázornená na obrázku 7.9c, bola zís- kaná z mračna referenčných bodov. Celkový proces vyhodnotenia metód je uvedený na obrázku 7.10.



Obr. 7.10: Blokový diagram vyhodnocovacieho procesu.

Kroky v tomto procese sú opísané nasledovne:

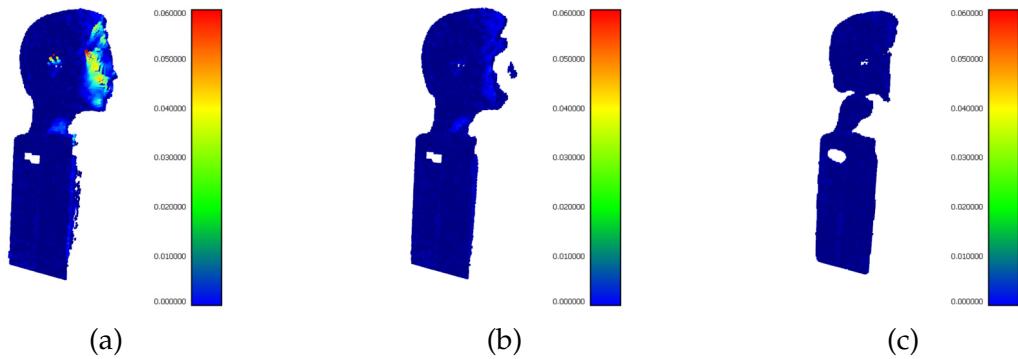
- **Zásobník** - záchytenie niekoľkých mračien bodov z konkrétneho pohľadu a zlúčenie do jedného mračna.
- **Filtrovanie** - použitie jednej z metód filtrovania opísaných v podkapitole 7.4.
- **Porovnanie s referenciou** - výpočet Hausdorffových vzdialenosí voči referenčnej sieti pre konkrétny pohľad.
- **Priestorové filtrovanie** - štatistické rozhodnutie, či ide o rušenie alebo nie na základe Hausdorffovej vzdialenosí. Za rušenie sa považujú štatistické odľahlé hodnoty.
- **Analýza** - vyhodnotenie výsledkov pomocou známych metrík z ROC analýzy.

V experimentálnej časti používame na hodnotenie výkonnosti konkrétnych metód nasledujúce metriky.

- **Accuracy** odráža celkový výkon algoritmu, takže zohľadňuje aj skutočné negatívne predpovede (bod nie je interferovaný). Presnosť je vysoká, keď je počet skutočných pozitívnych a pravdivých negatívnych predpovedí vysoký.
- **Recall** predstavuje počet odstránených bodov, ktoré boli interferované. Nezohľadňuje ale počet falošne pozitívnych odstránených bodov. Recall je vysoké, keď je počet skutočných predpovedí (filtrovaných interferenčných bodov) veľký, bez ohľadu na to, koľko užitočných bodov stratíme.
- **Precision** odráža počet skutočných predpovedí s ohľadom na počet falošne pozitívnych predpovedí.
- **F1 score** je hodnota rovnováhy medzi Recall a Precision. To znamená, že skóre F1 nie je skreslené veľkým počtom skutočne negatívnych predpovedí a môže sa považovať za rozhodujúce.
- **Hausdorffova vzdialenosť** meria euklidovskú vzdialenosť dvoch vzájomne porovnávaných bodov. Používa sa stredná hodnota vzdialnosti od bodu v mračne k najbližšiemu povrchu v referenčnej sieti .

7.6 Výsledky experimentu

V prvom kroku testovania sa hľadali optimálne nastavenia vstupných parametrov pre jednotlivé filtre. Automatickou rekonfiguráciou sa získali rôzne kombinácie nastavení, z ktorých boli následne vybrané najoptimálnejšie parametre (najlepšie F1 skóre). Vysoká hodnota Accuracy môže byť v tomto prípade zavádzajúci parameter, pretože veľké množstvo TN predikcií skresľuje výsledok. Taktiež nízka Hausdorffova vzdialenosť nemusí automaticky znamenať vysoký výkon filtra. Pri použití Hausdorffovej vzdialnosti nevidíme, koľko užitočných bodov bolo zbytočne odstránených. Príklad rozdielu v hodnotení pomocou Hausdorffovej vzdialnosti a F1 skóre je vidieť na obr. 7.11.

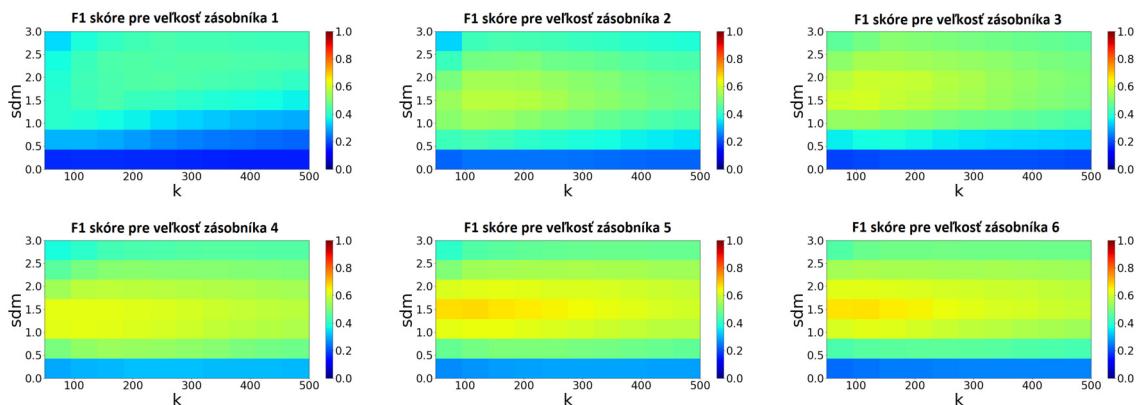


Obr. 7.11: Rozdiel v hodnotení výkonnosti filtrácie pomocou Hausdorffovej vzdialenosť a F1 skóre. (a) Mračno bodov s interferenciou. Priemerná Hausdorffova vzdialenosť je veľká. (b) Filtrované mračno bodov. Priemerná Hausdorffova vzdialenosť je nízka a skóre F1 je veľké. (c) Filtrované mračno bodov s veľkou stratou užitočných bodov. Priemerná Hausdorffova vzdialenosť je nízka, skóre F1 je nízke.

Meranie doby spracovania dát sa uskutočňovalo na operačnom systéme Ubuntu s procesorom Intel® Core™ i5-6440HQ 2,60 GHz a 16 GB operačnej pamäte.

7.6.1 SOR filtrácia

Filter SOR nám umožňuje upraviť sdm -parameter (prahová hodnota multiplikátora štandardnej odchýlky) a k -parameter (počet bodov, ktoré sa majú použiť na odhad priemernej vzdialenosť). Nasledujúce farebné mapy (Obr. 7.12) ukazujú, ako sa mení výkon filtrácie vyhodnotený ako skóre F1 v závislosti od výberu parametrov a veľkosti vyrovnávacej pamäte.



Obr. 7.12: Farebná mapa závislosti F1 skóre od vstupných parametrov SOR filtra pre každú veľkosť obrazového zásobníka.

Na základe analýzy na obrázku 7.12 sa získalo najlepšie nastavenie pre filter SOR. Zdá sa, že výkon má stúpajúci trend až do veľkosti zásobníka 5. Ideálna kombinácia

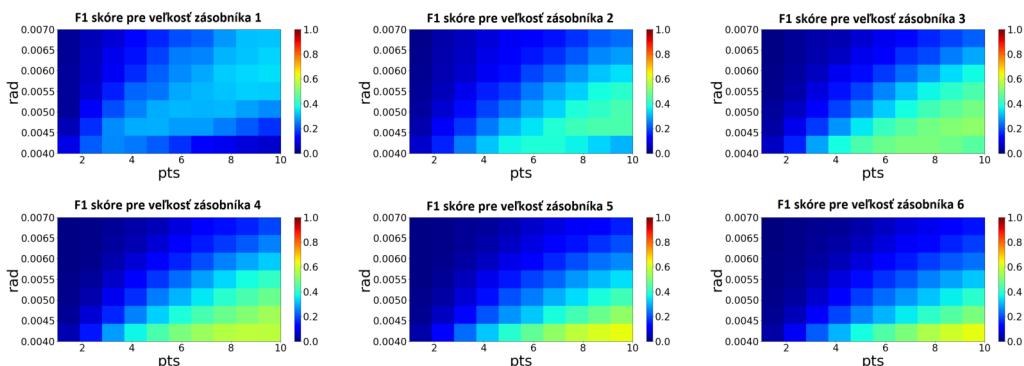
Tabuľka 7.1: Najlepšie výsledky SOR filtrácie pre veľkosť obrazového zásobníka 1-6.

Zásobník	Accuracy	Precision	Recall	F1 Score	Priem. HD [m]
1	0,964001	0,390533	0,538043	0,452571	0,001101
2	0,959534	0,474334	0,719212	0,571652	0,001073
3	0,964936	0,472317	0,853734	0,608171	0,001025
4	0,951624	0,501123	0,839413	0,627583	0,001041
5	0,968504	0,615551	0,754316	0,677905	0,001062
6	0,970065	0,598705	0,76028	0,669887	0,001053

parametrov pre veľkosť obrazového zásobníka 5 je sdm 1,5 a k 100. V tomto nastavení algoritmus dosahuje najlepšie skóre F1 0,68, ktoré je vidieť v tabuľke 7.1.

7.6.2 ROR filtrácia

Filter ROR nám umožňuje nastaviť polomer (priestorový rádius použitý na určenie najbližších susedov pts) a pts -parameter (minimálny počet susedov, ktoré musí mať bod v danom polomere vyhľadávania, aby neboli odfiltrované). Nasledujúce farebné mapy znázorňujú, ako sa mení skóre F1 v závislosti od výberu parametrov a veľkosti obrazového zásobníka.



Obr. 7.13: Farebná mapa závislosti F1 skóre od vstupných parametrov ROR filtra pre každú veľkosť obrazového zásobníka.

Podobne ako v predchádzajúcim prípade analýza na obrázku 7.13 ukazuje najlepšie nastavenie pre filter ROR. Výkonnosť filtrácie je najvyššia pre veľkosť zásobníka 5. Ideálna kombinácia parametrov je rad 0,004 a pts 10. V tomto nastavení dosahuje filter najlepšie skóre F1 0,63, ako je vidieť v tabuľke 7.2.

Tabuľka 7.2: Najlepšie výsledky ROR filtrácie pre veľkosti obrazového zásobníka 1-6.

Zásobník	Accuracy	Precision	Recall	F1 Score	Priem. HD [m]
1	0,956411	0,296154	0,418478	0,346847	0,001116
2	0,944072	0,352522	0,585222	0,44	0,001104
3	0,963341	0,446866	0,631255	0,523293	0,001064
4	0,949669	0,487503	0,711546	0,578593	0,001092
5	0,962245	0,553508	0,728088	0,628907	0,001068
6	0,968228	0,590323	0,6689	0,62716	0,001074

7.6.3 PointCleanNet filtrácia

Pri používaní pred-trénovanej siete PointCleanNet nebolo potrebné zadávať žiadne parametre pre odstraňovanie odľahlých hodnôt. V tomto prípade sa porovnávajú výsledky pre rozdielne veľkosti zásobníka. Spracovanie dát trvalo rádovo v minútach, pričom bola využívaná grafická karta s CUDA podporou.

Tabuľka 7.3: Najlepšie výsledky PointCleanNet filtrácie pre veľkosti obrazového zásobníka 1-6.

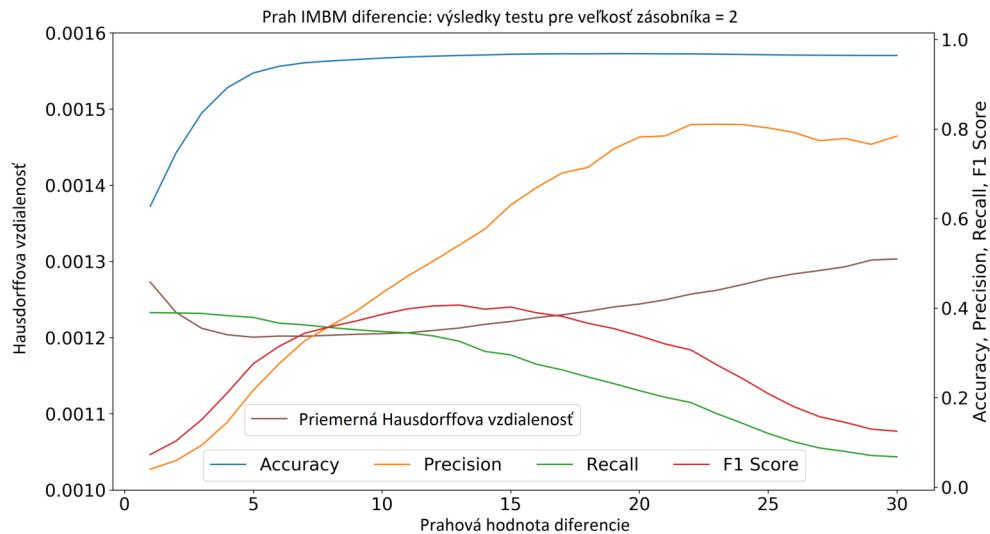
Zásobník	Accuracy	Precision	Recall	F1 Score	Priem. HD [m]
1	0,840448	0,084714	0,486413	0,144297	0,00112
2	0,835066	0,07334	0,291626	0,117205	0,001226
3	0,830422	0,06211	0,30639	0,103283	0,001181
4	0,819131	0,096918	0,327567	0,149579	0,001387
5	0,819513	0,094805	0,363546	0,150391	0,001324
6	0,823568	0,099486	0,424307	0,16118	0,001263

Výkon filtra má stúpajúci trend so zvyšujúcou veľkosťou zásobníka. Ako je však vidieť z tabuľky 7.3, najlepší výsledok získaný na základe skóre F1 je iba 0,16 pre veľkosť vyrovnávacej pamäte 6.

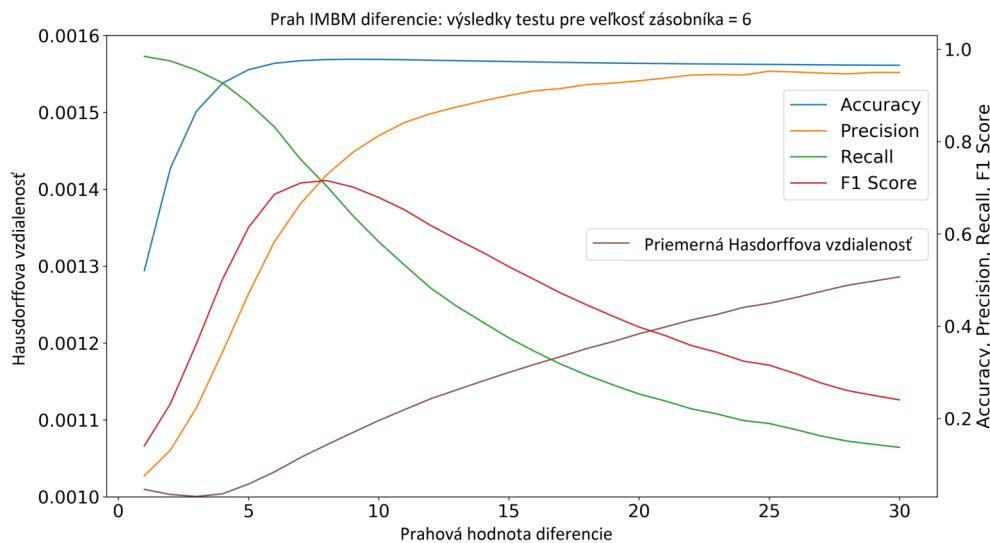
7.6.4 IMBM filtrácia

IMBM filter umožňuje nastaviť prahovú hodnotu hĺbky (odstránenie pozadia a extréme interferujúcich pixelov), prahovú hodnotu dôležitosti mapy (odstránenie oblastí s nízkou významnosťou pixelov v mediáne hĺbkovej mapy) a šumovú prahovú hodnotu (nastavenie prahovej hodnoty šumu kamery). Optimálna hodnota posledného spomenutého prahu pre veľkosť zásobníka 2 je uvedené na obrázku 7.14 a

pre veľkosť zásobníka 6 na obrázku 7.15. Z analýzy týchto obrázkov je najlepším šumovým prahom pre veľkosť zásobníka 2 hodnota 13 a pre veľkosť zásobníka 6 hodnota 8. Tento parameter bol v ďalšej časti experimentu staticky nastavený na 10.



Obr. 7.14: Výber ideálneho prahu rozdielu IMBM filtrácie pre veľkosť obrazového zásobníka 2.



Obr. 7.15: Výber ideálneho prahu rozdielu IMBM filtrácie pre veľkosť obrazového zásobníka 6.

Navrhovaný spôsob filtrovania dosahuje najvyššie F1 skóre pre veľkosť zásobníka 5. V tomto okamihu je zlepšenie priemernej Hausdorffovej vzdialenosť približne o 12 % (stredná vzdialenosť pre vstupné rušené mračná je 0,001255 m).

Tabuľka 7.4: Najlepšie výsledky IMBM filtrácie pre veľkosť obrazového zásobníka 2-6.

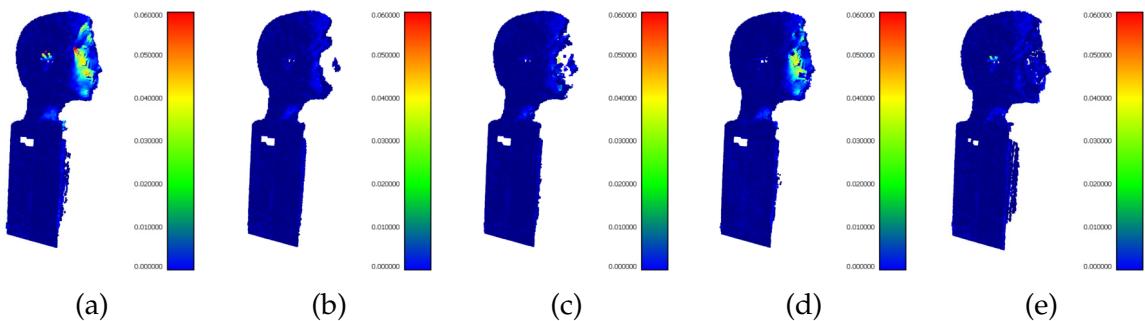
Zásobník	Accuracy	Precision	Recall	F1 Score	Priemerná HD [m]
2	0,958461	0,433661	0,347783	0,386003	0,001205
3	0,978235	0,724401	0,511932	0,59991	0,001119
4	0,950948	0,493889	0,410305	0,448233	0,001229
5	0,97733	0,815311	0,62583	0,708114	0,001114
6	0,977306	0,807993	0,566555	0,66607	0,001107

7.6.5 Zhodnotenie výsledkov

Po získaní najlepších nastavení pre konkrétné metódy filtrovania je potrebné komplexne porovnať ich výkonnéosti. Obrázok 7.16 ukazuje vizualizáciu Hausdorffových vzdialenosí výstupov z jednotlivých filtrov pre veľkosť zásobníka 5. Toto nastavenie bolo vybrané na základe najlepších výsledkov z tabuľky 7.5 (s výnimkou PointCleanNet, ktorý dosahuje najlepšie výsledky pri veľkosti 6).

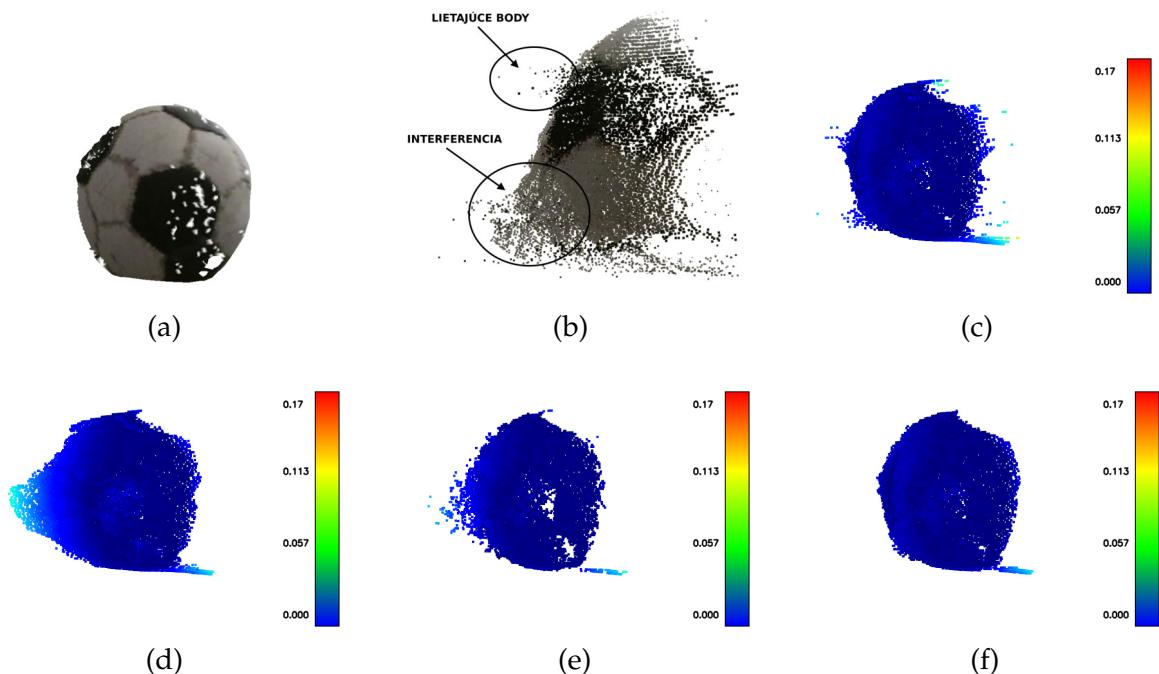
Tabuľka 7.5: Najlepšie skóre F1 pre všetky metódy filtrovania.

Zásobník	SOR	ROR	PointCleanNet	IMBM
1	0,452571	0,346847	0,144297	-
2	0,571652	0,44	0,117205	0,386003
3	0,608171	0,523293	0,103283	0,59991
4	0,627583	0,578593	0,149579	0,448233
5	0,677905	0,628907	0,150391	0,708114
6	0,669887	0,62716	0,16118	0,66607



Obr. 7.16: Komplexné porovnanie metód filtrácie pomocou vizualizácie vzdialenosí medzi filtrovanými mračnami bodov a referenčnou sieťou. (a) Vstupné mračno bodov bez potlačenia rušenia. (b) Najlepší výsledok SOR filtra s parametrami sdm 1.5 a k 100. (c) Najlepší výsledok pri použití ROR filtra s parametrami rad 0.004 and pts 10. (d) Najlepší výsledok pre PointCleanNet filtráciu s veľkosťou obrazového zásobníka 6. (e) Najlepší výsledok pre IMBM filter.

Porovnanie výstupných mračien bodov pre jednotlivé filtračné metódy je zobrazené na obr. 7.16. Interferovaná časť objektu obsahovala relatívne komplexné detaľy (oči, brada, nos). Filtre pracujúce na princípe štatistiky (SOR, ROR, IMBM) dokázali lepšie odstrániť interferujúce miesta ako filter založený na princípe neurónových sietí (PointCleanNet). Ten odstránil prevažne silno interferované body a mierne potlačil povrchovú interferenciu. Sieť však nebola trénovaná na odstránenie takéhoto typu šumu. Dotrénovaním siete na interferovaných mračnách môže tento filter vykazovať najlepšie výsledky. Je však potrebné vytvoriť dataset interferovaných mračien bodov. SOR a ROR odstránili väčšinu poškodených pixelov. Problém však je, že boli odstránené aj užitočné body, a to práve v miestach interferencie. Takouto stratou nie je možné späť rekonštruovať poškodené regióny.



Obr. 7.17: Vizualizácia Hausdorffovej vzdialenosť pre IMBM, SOR a ROR filtre. Všetky výsledky sú pre veľkosť obrazového zásobníka 5. (a) Referencia (lopta). (b) Rozdiel medzi interferenciou a lietajúcimi bodmi. (c) Potlačenie interferencie IMBM filtrom, lietajúce doby ostali. (d) SOR filter odstránil lietajúce body ale interferencia ostala. (e) ROR filter odstránil lietajúce body ale interferencia ostala. (f) Najlepší výsledok aplikovaním IMBM a SOR filtra spoločne.

Problematickými časťami filtrov SOR a ROR sú rohy a hrany, ktoré sa filtrovaním zaobľujú. Ako je uvedené v štúdii [76], filter SOR vníma body na hranach ako odľahlé body. Tieto filtre tiež odstraňujú malé ostrovy bodov (zhluky bodov, ktoré nie

sú bodmi rušenia) [81]. Ako je uvedené v [78], filtrovanie SOR nie je úplne vhodné pre výkon v reálnom čase. IMBM filter dokázal potlačiť interferenciu bez väčšej straty plochy objektu, čo je znázornené aj na obr. 7.16e. IMBM sa nezameriava iba na odstraňovanie rušenia, ale poskytuje aj interpoláciu chýbajúcich bodov, avšak výrazne predlžuje dobu spracovania dát. Nedokázal ani odstrániť všetky chybné body. Je to spôsobené tým, že tieto body boli podobné vo všetkých hĺbkových mapách v obrazovom zásobníku. Pre elimináciu chybných dát a zníženie Hausdorffovej vzdialenosťi je možné kombinovať IMBM filter s inými štatistickými filtrovmi.

Na záverečné vyhodnotenie metód filtrovania sa použil dataset zložený z 23 objektov. Tento nami vytvorený dataset obsahuje 23 skenov predmetov (guľa, plastová fľaša ...) zachytených z rôznych pohľadov. Nasledujúca tabuľka 7.6 zobrazuje medíany každej vyhodnocovanej metriky v celom datasete. Všetky výsledky sú pre veľkosť obrazového zásobníka s kapacitou 5.

Na základe výsledkov tohto prieskumu je odporúčané používať metódy filtrovania pre konkrétné aplikácie podľa ich funkcií opísaných v tabuľke 7.7.

Tabuľka 7.6: Medián všetkých použitých metrík v celom datasete.

Metóda	Accuracy	Precision	Recall	F1 score	Priemerná HD [m]
SOR	0.936164	0.980778	0.393617	0.555951	0.002393
ROR	0.927161	0.639983	0.734787	0.647839	0.001697
PCN	0.862285	0.371242	0.472615	0.409729	0.00221
IMBM	0.908833	0.673608	0.442942	0.477348	0.002563
IMBM + ROR	0.918448	0.59806	0.820666	0.653113	0.001656
IMBM + SOR	0.943387	0.704448	0.836253	0.741811	0.001548

Tabuľka 7.7: Porovnanie vlastností filtračných metód

Metóda	Potláč. interf.	Rýchlosť	Filtrovanie hrán/rohov	Obnova dát
SOR	Stredná	Stredná	Stredná	Nie
ROR	Stredná	Vysoká	Stredná	Nie
PCN	Nízka	Nízka	Stredná	Ano
IMBM	Vysoká	Vysoká	Nízka	Ano

Ideálnou voľbou je použitie IMBM filtra pre potlačenie interferencie spolu s ROR alebo SOR filtrom pre odstránenie zvyšných chybných bodov. IMBM a SOR filtre boli aplikované v kapitole 6.4.4.

8. Záver

Táto dizertačná práca bola zameraná na prácu s 2D obrazovými a 3D priestorovými dátami. Cieľom bolo navrhnúť snímací systém a algoritmus, ktorý by slúžil ako podpora pre medicínske aplikácie. V posledných rokoch sa medicínsky výskum intenzívnejšie zaobráva diagnostikou OSA pomocou 3D informácií pacientov. Síce tento problém je medicínsky, pre jeho riešenie a ďalší rozvoj je taktiež potrebný výskum v technickej oblasti.

Navrhnutý algoritmus mal umožniť vytvorenie precízneho priestorového modelu objektu s použitím viacerých RGB-D kamier. Algoritmus mal taktiež umožniť detekciu vybraných kľúčových bodov na snímanom objekte a automatizovať meranie geometrických parametrov.

Systém bol optimalizovaný pre snímanie pediatrických pacientov, ktorí svojimi vlastnosťami špecifikovali určité kritériá a požiadavky. Základnou požiadavkou bolo minimalizovanie doby skenovania a bezkontaktné meranie euklidovej vzdialenosť kranio-faciálnych parametrov.

V prvej časti sme navrhovali multi-kamerový systém a určovali sme pracovné podmienky. Experimentálne sme overili, že pre skenovanie detských pacientov je potrebné pracovať s viac-kamerovým systémom. Taktiež sme poukázali na nutnosť použitia paralelnej spolupráce kamier. Detailne sme opísali spôsob geometrickej a multi-kamerovej kalibrácie, ktoré sú využívané v procese snímania.

V druhej časti opisujeme algoritmus, ktorým je vykonávaný proces snímania. Našim návrhom sme vytvorili systém pre zber dát z viacerých RGB-D kamier v reálnom čase. Tým sme minimalizovali vplyv pohybu na výslednú presnosť 3D modelu. Navrhli sme systém, ktorý využíva detekciu kľúčových bodov na tvári k meraniu pohybu objektu počas doby skenovania. K zvýšeniu kvality zosnímaných dát sme implementovali známe filtračné a regisračné metódy. K potlačeniu multi-kamerovej interferencie, ktorá je dôsledkom paralelnej spolupráce viacerých ToF kamier, sme navrhli vlastnú filtračnú metódu s názvom IMBM. Tá dokáže odstrániť interferované regióny v hĺbkovej mape a do určitej miery rekonštruovať tieto poškodené miesta. Pomocou 68 bodového detektora kľúčových bodov sme navrhli metódu merania euklidovských vzdialenosťí vybraných kranio-faciálnych parametrov. Overili sme presnosť merania na živých objektoch a zároveň ako rozšírenie práce sme navrhli algoritmus, ktorý umožňuje segmentáciu ľudskej tváre. Takéto segmentovanie má potenciálne využitie pri procese prípravy datasetu, ktorý bude zložený z mračien bodov a bude použitý k vytvoreniu OSA klasifikátora.

Z dôsledku potrieb pri návrhu algoritmu bol taktiež vytvorený systém pre zber obrazových dát, ktorý je umiestnený v spánkovom laboratóriu. Tento systém zaznamenáva hĺbkové a RGB-D informácie pediatrických pacientov spolu s dokumentáciou o vyšetrení.

V ďalšom vývoji a výskume sa zameriame na implementáciu novších cenovo dostupných hĺbkových snímačov. Cieľom bude zvýšiť presnosť snímania pediatrických pacientov. Taktiež sa budeme venovať zvyšovaniu presnosti segmentácie tváre.

9. Výsledky dizertácie s uvedením nových poznatkov

Dizertačná práca rieši návrh skenovacieho systému, ktorý bude použitý ako skríniový nástroj pri diagnostike spánkového obstrukčného apnoe. Navrhnuté algoritmy sú univerzálne použiteľné vo viacerých oblastiach, ale pozornosť je sústredená na skenovanie pediatrických pacientov. Dosiahnuté výsledky ukazujú možnosti využitia počítačového videnia a číslicového spracovania signálov v oblasti medicínskej problematiky, konkrétnie využitia hĺbkových snímačov pre priestorovú rekonštrukciu pacientov.

Cieľom dizertácie bol návrh vizuálneho systému a algoritmu, ktorým sa zautomatizujú v súčasnosti manuálne vykonávané úkony vyšetrenia. Vytvorenie tohto systému môže viesť k rozvíjaniu a zlepšovaniu diagnostiky OSA pomocou obrazových informácií. Výhodou tohto systému oproti komerčným riešeniam je kombinácia cenovej dostupnosti a rýchlosťi snímania. Daná skutočnosť má v budúcnosti viesť k vytvoreniu robustnej databázy OSA pacientov, ktorá bude použitá pri vytváraní klasifikačného systému v procese sofistikovanej diagnostiky. Kľúčovými nástrojmi boli algoritmy číslicového spracovania signálu a analýzy obrazu, čo predstavuje prienik vedného odboru Telekomunikácie do oblasti medicíny. Algoritmy a postupy boli testované a odladené na referenčných modeloch a následne overené na reálnych dynamických objektoch.

Podľa názoru autora je pôvodným vedeckým prínosom dizertačnej práce:

1. Návrh multi-kamerového systému a algoritmov umožňujúcich rýchle vytvorenie 3D modelov objektov.
2. Návrh filtračného algoritmu, ktorý potláča vplyv multi-kamerovej interferencie vznikajúcej pri paralelnej spolupráci ToF kamier.
3. Prvotný návrh algoritmu, ktorým sa vykonáva segmentovanie mračien bodov obsahujúcich tváre. Pre zlepšenie výkonnosti segmentácie bol vytvorený systém zabezpečujúci zber potrebných obrazových dát.

Výhodou navrhnutých postupov, algoritmov a programových nástrojov je ich transparentnosť, otvorenosť a možnosť ďalšej modifikácie pre konkrétnu potrebu a aplikácie.

10. Publikované práce autora

PUBLIKÁCIE SÚVISIACE S PRÁCOU

AFC Volak, J., Koniar, D., Jabloncik, F., Hargas, L., Janisova, S.: Interference artifacts suppression in systems with multiple depth cameras, *42nd International Conference on Telecommunications and Signal Processing*, TSP 2019, July 2019, Pages 472-476 (WoS, Scopus)

AFC Volák, J., Koniar, D., Jablončík, F., Hargaš, L.: A Study on OSAS Diagnostics Supported by RGB-D Imaging, *41st International Conference on Telecommunications and Signal Processing*, TSP 2018, Greece, 4 July 2018 through 6 July 2018 (WoS, Scopus)

AFC Volak, J., Koniar, D., Hargas, L., Jabloncik, F., Sekel'ova, N., Durdik, P.: RGB-D imaging used for OSAS diagnostics, *12th International ELEKTRO Conference*, ELEKTRO 2018, Mikulov, Czech Republic; 21 May 2018 through 23 May 2018 (WoS, Scopus)

AFC Koniar D., Volák J., Šindler P., Hargaš L., Jablončík F., Janišová S.: Fast parallel 3D scanning system based on non-expensive devices, *Modelling of mechanical and mechatronic systems 2019 - conference proceedings*, s. 112-119, 2019, ISBN 978-80-553-3368-7.

AED Volák J., Koniar D., Hargaš L., Jablončík F., Ďurdík P., Šutvajová D., Duchoň F., Beňo P.: Technická podpora diagnostiky obstrukčného spánkového apnoe, *Diagnosika a terapia v pediatrii*, s. 52-55, 2019, ISBN 978-80-89797-49-3.

AFC Volák J.: Stereovision in robotic systems, *Poster 2018 - proceedings of the International Student Scientific Conference Poster*, Praha, s. 1-5, 2018, ISBN 978-80-01-06428-3.

AFC Koniar D., Volák J., Hargaš L., Janišová S., Bajzík J.: Depth cameras calibration for obstructive sleep apnea diagnostic support, *ELEKTRO 2020 – 13th International Conference*, Taormina, Italy, 2020 (čaká sa na zaradenie do databáz WoS a Scopus).

AFC Bajzík J., Koniar D., Hargaš L., Volák J., Janišová S.: Depth sensor selection for specific application, *ELEKTRO 2020 – 13th International Conference*, Taormina, Italy, 2020 (čaká sa na zaradenie do databáz WoS a Scopus).

Publikácie čakajúce na zverejnenie

ADC Volák J., Bajzík J., Janišová S., Koniar D., Hargaš L.: Real-Time interference artifacts suppression in array of ToF sensors, *MDPI Sensors*, Issue 20, 2020, ISSN: 1424-8220

Poznámka: článok bol akceptovaný 29. 6. 2020 a bude zaradený do databázy CC-WoS (Q1, IF=3,031)

Publikácie v recenznom konaní

ADD Volák J., Koniar D., Hargaš L., Bajzík J., Janišová S.: Importance of depth camera calibration in obstructive sleep apnea screening system, *Measurement Science Review*, ISSN: 1335-8871 (WoS/CCC/Q2 , IF=1,122)

AGJ Paralelný viac senzorový 3D skenovací systém s bežnými kamerami - Volák J., Koniar D., Hargaš L., Bajzík J., Janišová S.

PUBLIKÁCIE NESÚVISIACE S PRÁCOU

AFC Jabloncik, F., Hargas, L., Koniar, D., Volak, J.: Detection of static objects in an image using texture analysis, *Transportation Research Procedia*, Volume 40, 2019, Pages 265-270, ISSN:2352-1457 (Scopus, SJR=0,48)

AFC Jabloncik, F., Hargas, L., Volak, J., Koniar, D.: Detection of Static Objects in an Image Based on Texture Analysis, Lecture Notes in Computer Science (including sub-series Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), Volume 11466 LNBI, 2019, Pages 445-457, ISSN:0302-9743 (Scopus / Q2, SJR=0,43)

AFC Hargas, L., Loncova, Z., Koniar, D., Jabloncik, F., Volak, J.: Artefacts Recognition and Elimination in Video Sequences with Ciliary Respiratory Epithelium, *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, Volume 11466 LNBI, 2019, Pages 458-468, ISSN:0302-9743 (Scopus / Q2, SJR=0,43)

AFC Hargas, L., Jabloncik, F., Koniar, D., Volak, J., Loncova, Z.: Identification of static objects in the respiratory epithelium, *12th International ELEKTRO Conference*, ELEKTRO 2018, Mikulov, Czech Republic; 21 May 2018 through 23 May 2018 (WoS, Scopus)

AFC Taraba, M., Volak, J., Jabloncik, F., Koniar, D., Hargas, L.: Conditioning of light microscope for advanced methods of ciliary dyskinesia diagnostics, *12th International ELEKTRO Conference*, ELEKTRO 2018, Mikulov, Czech Republic; 21 May 2018 through 23 May 2018 (WoS, Scopus)

AFC Jabloncik, F., Hargas, L., Koniar, D., Volak, J., Loncova, Z.: Dynamic objects detection of the respiratory epithelium based on image analysis, *12th International ELEKTRO Conference*, ELEKTRO 2018, Mikulov, Czech Republic; 21 May 2018 through 23 May 2018 (WoS, Scopus)

AFC Hargaš, L., Loncová, Z., Koniar, D., Jablončík, F., Volák, J.: Detection Methods of Static Microscopic Objects, *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, Volume 10814 LNBI, 2018, Pages 163-175, ISSN:0302-9743 (Scopus / Q2, SJR=0,43)

AFC Hargaš L., Bulava J., Koniar D., Šindler P., Jablončík F., Volák J.: Microscopic mechatronic objects movement measurement by stroboscopic effect, *Modelling of mechanical and mechatronic systems 2019* - conference proceedings, s. 63-68, 2019, ISBN 978-80-553-3368-7.

AED Jablončík F., Volák J., Hargaš L., Koniar D., Ďurdík P., Kvaššayová J., Bánovčin P.: Návrh algoritmu pre automatickú segmentáciu statických cílií respiračného epitelu v obraze, *Diagnostika a terapia v pediatrii*, s. 14-17, 2019, ISBN 978-80-89797-49-3.

AFC Jablončík F., Hargaš L., Koniar D., Volák J.: Contrast enhancement methods for images from the light microscope, *Bioinformatics and biomedical engineering*, s. 85-88, 2018, ISBN 978-84-17293-36-9.

AFC Bulava J., Hargaš L., Koniar D., Šindler P., Jablončík, F., Volák J.: Stroboscopic measurement of kinematics parameters using highspeed camera, ELEKTRO 2020 – 13th International Conference, Taormina, Italy, 2020

Literatúra

- [1] Jamie CM Lam, SK Sharma, Bing Lam, et al. Obstructive sleep apnoea: definitions, epidemiology & natural history. *Indian Journal of Medical Research*, 131(2):165, 2010.
- [2] B Lam, MSM Ip, E Tench, and CF Ryan. Craniofacial profile in asian and white subjects with obstructive sleep apnoea. *Thorax*, 60(6):504–510, 2005.
- [3] Michael J Sateia. International classification of sleep disorders-third edition. *Chest*, 146(5):1387–1394, Nov 2014.
- [4] Rachel Jen, Fernanda R Almeida, Penelope Brasher, Mary M Doyle-Waters, Jim Salzman, and John Fleetham. Oral appliances for obstructive sleep apnoea. *Cochrane Database of Systematic Reviews*, Sep 2018.
- [5] Terry Young, Paul E. Peppard, and Daniel J. Gottlieb. Epidemiology of obstructive sleep apnea. *American Journal of Respiratory and Critical Care Medicine*, 165(9):1217–1239, May 2002.
- [6] Kirk Kee, Matthew T Naughton, et al. Sleep apnoea-a general practice approach. *Australian family physician*, 38(5):284, 2009.
- [7] Deepak Shrivastava, Syung Jung, Mohsen Saadat, Roopa Sirohi, and Keri Crewson. How to interpret the results of a sleep study. *Journal of Community Hospital Internal Medicine Perspectives*, 4(5):24983, Jan 2014.
- [8] David R Hillman and Leon C Lack. Public health implications of sleep loss: the community burden. *Medical Journal of Australia*, 199:S7–S10, 2013.

- [9] Maria Pia Villa, Maria Chiara Paolino, Rosa Castaldo, Nicola Vanacore, Alessandra Rizzoli, Silvia Miano, Marco Del Pozzo, and Marilisa Montesano. Sleep clinical record: an aid to rapid and accurate diagnosis of paediatric sleep disordered breathing. *European Respiratory Journal*, 41(6):1355–1361, Sep 2012.
- [10] Andy Whyte and Daren Gibson. Imaging of adult obstructive sleep apnoea. *European Journal of Radiology*, 102:176–187, May 2018.
- [11] Milad Kiaee, Adam B. Kashlak, Jisu Kim, and Giseon Heo. Diagnosis of pediatric obstructive sleep apnea via face classification with persistent homology and convolutional neural networks. *ArXiv*, abs/1911.05628, 2019.
- [12] Herbert Edelsbrunner and John Harer. Persistent homology-a survey. *Contemporary mathematics*, 453:257–282, 2008.
- [13] Christian Szegedy, Sergey Ioffe, Vincent Vanhoucke, and Alexander A Alemi. Inception-v4, inception-resnet and the impact of residual connections on learning. In *Thirty-first AAAI conference on artificial intelligence*, 2017.
- [14] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jon Shlens, and Zbigniew Wojna. Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2818–2826, 2016.
- [15] S. M. S. Islam, H. Mahmood, A. A. Al-Jumaily, and S. Claxton. Deep learning of facial depth maps for obstructive sleep apnea prediction. In *2018 International Conference on Machine Learning and Data Engineering (iCMLDE)*, pages 154–157, 2018.
- [16] Richard WW Lee, Andrew SL Chan, Ronald R Grunstein, and Peter A Cistulli. Craniofacial phenotyping in obstructive sleep apnea—a novel quantitative photographic approach. *Sleep*, 32(1):37–45, 2009.
- [17] Eung-Kwon Pae, Alan A Lowe, and John A Fleetham. Shape of the face and tongue in obstructive sleep apnea patients—statistical analysis of coordinate data. *Clinical orthodontics and research*, 2(1):10–18, 1999.

- [18] Richard WW Lee, Peter Petocz, Tania Prvan, Andrew SL Chan, Ronald R Grunstein, and Peter A Cistulli. Prediction of obstructive sleep apnea with craniofacial photographic analysis. *Sleep*, 32(1):46–52, 2009.
- [19] Artec 3D. Artec eva. <https://www.artec3d.com/portable-3d-scanners/artec-eva>. Online: 2020-04-20.
- [20] Omkar M Parkhi, Andrea Vedaldi, and Andrew Zisserman. Deep face recognition. 2015.
- [21] Iacopo Masi, Stephen Rawls, Gérard Medioni, and Prem Natarajan. Pose-aware face recognition in the wild. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4838–4846, 2016.
- [22] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105, 2012.
- [23] Ken Chatfield, Karen Simonyan, Andrea Vedaldi, and Andrew Zisserman. Return of the devil in the details: Delving deep into convolutional nets. *arXiv preprint arXiv:1405.3531*, 2014.
- [24] Asghar Tabatabaei Balaei, Kate Sutherland, Peter A Cistulli, and Philip de Chazal. Automatic detection of obstructive sleep apnea using facial images. In *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*, pages 215–218. IEEE, 2017.
- [25] Peter Eastwood, Syed Zulqarnain Gilani, Nigel McArdle, David Hillman, Jennifer Walsh, Kathleen Maddison, Mithran Goonewardene, and Ajmal Mian. Predicting sleep apnea from three-dimensional face photography. *Journal of Clinical Sleep Medicine*, 16(4):493–502, 2020.
- [26] Shih-Wei Lin, Kate Sutherland, Yu-Fang Liao, Peter A Cistulli, Li-Pang Chuang, Yu-Ting Chou, Chih-Hao Chang, Chung-Shu Lee, Li-Fu Li, and Ning-Hung Chen. Three-dimensional photography for the evaluation of facial profiles in obstructive sleep apnoea. *Respirology*, 23(6):618–625, 2018.

- [27] Seth M Weinberg and John C Kolar. Three-dimensional surface imaging: limitations and considerations from the anthropometric perspective. *Journal of Craniofacial Surgery*, 16(5):847–851, 2005.
- [28] Seth M Weinberg, Sybill Naidoo, Daniel P Govier, Rick A Martin, Alex A Kane, and Mary L Marazita. Anthropometric precision and accuracy of digital three-dimensional photogrammetry: comparing the genex and 3dmd imaging systems with one another and with direct anthropometry. *Journal of Craniofacial Surgery*, 17(3):477–483, 2006.
- [29] Kate Sutherland, Richard WW Lee, Peter Petocz, Tat On Chan, Susanna Ng, David S Hui, and Peter A Cistulli. Craniofacial phenotyping for prediction of obstructive sleep apnoea in a chinese population. *Respirology*, 21(6):1118–1125, 2016.
- [30] Richard WW Lee, Kate Sutherland, Andrew SL Chan, Biao Zeng, Ronald R Grunstein, M Ali Darendeliler, Richard J Schwab, and Peter A Cistulli. Relationship between surface facial dimensions and upper airway structures in obstructive sleep apnea. *Sleep*, 33(9):1249–1254, 2010.
- [31] Suresh Balakrishnama and Aravind Ganapathiraju. Linear discriminant analysis-a brief tutorial. In *Institute for Signal and information Processing*, volume 18, pages 1–8, 1998.
- [32] Aamir Saeed Malik and Tae-Sun Choi. A novel algorithm for estimation of depth map using image focus for 3d shape recovery in the presence of noise. *Pattern Recognition*, 41(7):2200–2225, Jul 2008.
- [33] Andrej Satnik, Ebroul Izquierdo, and Richard Orjesek. Multiview 3d sensing and analysis for high quality point cloud reconstruction. In *Tenth International Conference on Machine Vision (ICMV 2017)*, volume 10696, page 106962K. International Society for Optics and Photonics, 2018.
- [34] Zhengyou Zhang. A flexible new technique for camera calibration. *IEEE Transactions on pattern analysis and machine intelligence*, 22(11):1330–1334, 2000.

- [35] M. Kowalski, J. Naruniec, and M. Daniluk. Livescan3d: A fast and inexpensive 3d data acquisition system for multiple kinect v2 sensors. In *2015 International Conference on 3D Vision*, pages 318–325, 2015.
- [36] R. S. Yang, Y. H. Chan, R. Gong, M. Nguyen, A. G. Strozzi, P. Delmas, G. Gimel’farb, and R. Ababou. Multi-kinect scene reconstruction: Calibration and depth inconsistencies. In *2013 28th International Conference on Image and Vision Computing New Zealand (IVCNZ 2013)*, pages 47–52, 2013.
- [37] J. Fu, S. Wang, Y. Lu, S. Li, and W. Zeng. Kinect-like depth denoising. In *2012 IEEE International Symposium on Circuits and Systems (ISCAS)*, pages 512–515, 2012.
- [38] T. Mallick, P. P. Das, and A. K. Majumdar. Characterizations of noise in kinect depth images: A review. *IEEE Sensors Journal*, 14(6):1731–1740, 2014.
- [39] E Roy Davies. *Machine vision: theory, algorithms, practicalities*. Elsevier, 2004.
- [40] Thomas Fabry, Dirk Smeets, and Dirk Vandermeulen. Surface representations for 3d face recognition. *Face Recognition*, pages 273–294, 2010.
- [41] D. L. Lau. The science behind kinect. <http://lau.engr.uky.edu/news/science-behind-kinect>. Online: 2020-05-24.
- [42] Rahul Kala. *On-road intelligent vehicles: Motion planning for intelligent transportation systems*. Butterworth-Heinemann, 2016.
- [43] Jason Geng. Structured-light 3d surface imaging: a tutorial. *Adv. Opt. Photon.*, 3(2):128–160, Jun 2011.
- [44] Larry Li. Time-of-flight camera—an introduction. *Technical white paper*, (SLOA190B), 2014.
- [45] Miles Hansard, Seungkyu Lee, Ouk Choi, and Radu Patrice Horaud. *Time-of-flight cameras: principles, methods and applications*. Springer Science & Business Media, 2012.

- [46] D Van Nieuwenhove, W Van der Tempel, R Grootjans, and M Kuijk. Time-of-flight optical ranging sensor based on a current assisted photonic demodulator. In *Annual Symposium of the IEEE Photonics Benelux Chapter, Ghent, Belgium*, pages 209–212, 2006.
- [47] Chee Kai Chua, Chee How Wong, and Wai Yee Yeong. *Standards, quality control, and measurement sciences in 3D printing and additive manufacturing*. Academic Press, 2017.
- [48] L. Bamber. Intel realsense technology and the point cloud. <https://software.intel.com/content/www/us/en/develop/articles/intel-realsense-technology-and-the-point-cloud.html>, 2015. Online: 2020-05-11.
- [49] Fausto Bernardini, Joshua Mittleman, Holly Rushmeier, Cláudio Silva, and Gabriel Taubin. The ball-pivoting algorithm for surface reconstruction. *IEEE transactions on visualization and computer graphics*, 5(4):349–359, 1999.
- [50] Michael Kazhdan, Matthew Bolitho, and Hugues Hoppe. Poisson surface reconstruction. In *Proceedings of the fourth Eurographics symposium on Geometry processing*, volume 7, 2006.
- [51] Colin Smith. *On vertex-vertex systems and their use in geometric and biological modelling*. University of Calgary, 2006.
- [52] J. Chalmovianská. Afinné transformácie. http://www.sccg.sk/~pilnikova/pg/afinne_transformacie.pdf. Online: 2020-04-17.
- [53] M. Bátorová. Geometrické transformácie a ich triedy. http://www.http://fractal.dam.fmph.uniba.sk/~batorova/UPG/2_Transformacie. Online: 2020-04-17.
- [54] Richard Hartley Andrew Zisserman. Multiple view geometry in computer vision. 2004.

- [55] F Bellavia, G Gagliano, D Tegolo, and C Valenti. Underwater archaeological mosaicing. In *10th WSEAS International Conference on COMPUTERS. Vouliagmeni, Athens, Greece*, pages 1141–1146, 2006.
- [56] Hyunchul Lee and Okkyung Choi. An efficient parameter update method of 360-degree vr image model. *International Journal of Engineering Business Management*, 11:1847979019835993, 2019.
- [57] Olivier Faugeras and OLIVIER AUTOR FAUGERAS. *Three-dimensional computer vision: a geometric viewpoint*. MIT press, 1993.
- [58] Wikipedia contributors. Kinect — Wikipedia, the free encyclopedia. <https://en.wikipedia.org/w/index.php?title=Kinect&oldid=965084159>, 2020. [Online; accessed 2-July-2020].
- [59] Nikola Sekelová. Využitie rgb-d snímačov v biomedicíne, 2018.
- [60] Silvia Janišová. Využitie 3d snímania pri diagnostike obštrukčného spánkového apnoe, 2019.
- [61] J Blake, F Echtler, and C Kerl. libfreenect2: Open source drivers for the kinect for windows v2 device, 2015.
- [62] Felix Järemo Lawin, Per-Erik Forssén, and Hannes Ovrén. Efficient multi-frequency phase unwrapping using kernel density estimation. In *European Conference on Computer Vision*, pages 170–185. Springer, 2016.
- [63] Davis E King. Dlib-ml: A machine learning toolkit. *The Journal of Machine Learning Research*, 10:1755–1758, 2009.
- [64] Alexandru Telea. An image inpainting technique based on the fast marching method. *Journal of graphics tools*, 9(1):23–34, 2004.
- [65] Michael Goerner. Incremental registration. https://pointclouds.org/documentation/incremental_registration_8hpp_source.html. Online: 2020-02-03.

- [66] Kaiming He, Georgia Gkioxari, Piotr Dollár, and Ross Girshick. Mask r-cnn. In *Proceedings of the IEEE international conference on computer vision*, pages 2961–2969, 2017.
- [67] Abhishek Dutta and Andrew Zisserman. The VIA annotation software for images, audio and video. In *Proceedings of the 27th ACM International Conference on Multimedia*, MM ’19, New York, NY, USA, 2019. ACM.
- [68] Charles R Qi, Li Yi, Hao Su, and Leonidas J Guibas. Pointnet++: Deep hierarchical feature learning on point sets in a metric space. *arXiv preprint arXiv:1706.02413*, 2017.
- [69] Young Min Kim, Derek Chan, Christian Theobalt, and Sebastian Thrun. Design and calibration of a multi-view tof sensor fusion system. In *2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, pages 1–7. IEEE, 2008.
- [70] Bernhard Buttgen, Felix Lustenberger, Peter Seitz, et al. Pseudonoise optical modulation for real-time 3-d imaging with minimum interference. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 54(10):2109–2119, 2007.
- [71] Refael Z Whyte, Andrew D Payne, Adrian A Dorrington, and Michael J Cree. Multiple range imaging camera operation with minimal performance impact. In *Image Processing: Machine Vision Applications III*, volume 7538, page 75380I. International Society for Optics and Photonics, 2010.
- [72] Bernhard Buttgen and Peter Seitz. Robust optical time-of-flight range imaging based on smart pixel structures. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 55(6):1512–1525, 2008.
- [73] Lianhua Li, Sen Xiang, You Yang, and Li Yu. Multi-camera interference cancellation of time-of-flight (tof) cameras. In *2015 IEEE International Conference on Image Processing (ICIP)*, pages 556–560. IEEE, 2015.
- [74] Robert Lange. 3d time-of-flight distance measurement with custom solid-state image sensors in cmos/ccd-technology. 2000.

- [75] Xian-Feng Han, Jesse S Jin, Ming-Jie Wang, Wei Jiang, Lei Gao, and Liping Xiao. A review of algorithms for filtering the 3d point cloud. *Signal Processing: Image Communication*, 57:103–112, 2017.
- [76] Francesco Pirotti, Roberta Ravanelli, Francesca Fissore, and Andrea Masiero. Implementation and assessment of two density-based outlier detection methods over large spatial point clouds. *Open Geospatial Data, Software and Standards*, 3(1):1–12, 2018.
- [77] Juan Manuel Corso Sarmiento, M Pilar Garcia-Almirall, Daniel López Jiménez, and Jordi Casals Fernandez. Disclosure of temporary exposures as permanent website applications through the patrimonial survey. In *IOP Conference Series: Materials Science and Engineering (Online)*, number 245, pages 1–11. Institute of Physics (IOP), 2017.
- [78] Haris Balta, Jasmin Velagic, Walter Bosschaerts, Geert De Cubber, and Bruno Siciliano. Fast statistical outlier removal based method for large 3d point clouds of outdoor environments. *IFAC-PapersOnLine*, 51(22):348–353, 2018.
- [79] Paul Guerrero, Yanir Kleiman, Maks Ovsjanikov, and Niloy J Mitra. Pcpnet learning local shape properties from raw point clouds. In *Computer Graphics Forum*, volume 37, pages 75–85. Wiley Online Library, 2018.
- [80] Jozef Volak, Dusan Koniar, Frantisek Jabloncik, Libor Hargas, and Silvia Janisova. Interference artifacts suppression in systems with multiple depth cameras. In *2019 42nd International Conference on Telecommunications and Signal Processing (TSP)*, pages 472–476. IEEE, 2019.
- [81] Christian Schaller, Andre Adelt, Jochen Penne, and Joachim Hornegger. Time-of-flight sensor for patient positioning. In *Medical Imaging 2009: Visualization, Image-Guided Procedures, and Modeling*, volume 7261, page 726110. International Society for Optics and Photonics, 2009.

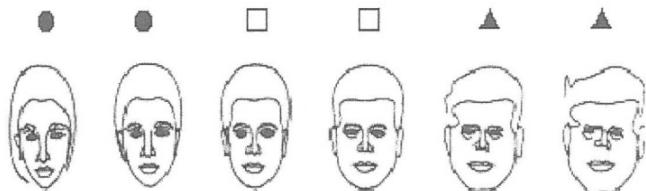
xi. Prílohy

xi.i Dotazník OSA

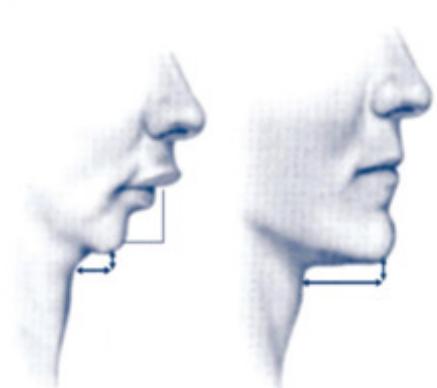
V tejto prílohe sa nachádzajú časti dotazníka, ktorý sa využíva pri vyšetrení OSAS.

FENOTYP

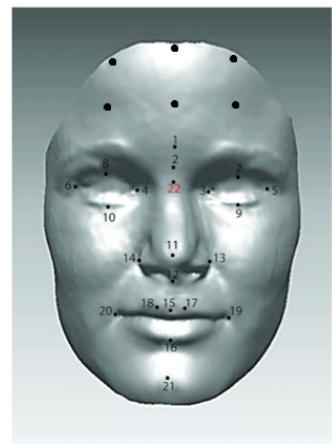
- Adenoid fenotyp
- Normálny fenotyp
- ▲ Fenotyp dospelého človeka



Obr. xi.i: Identifikácia fenotypu tváre na základe geometrických rozmerov.



(a)

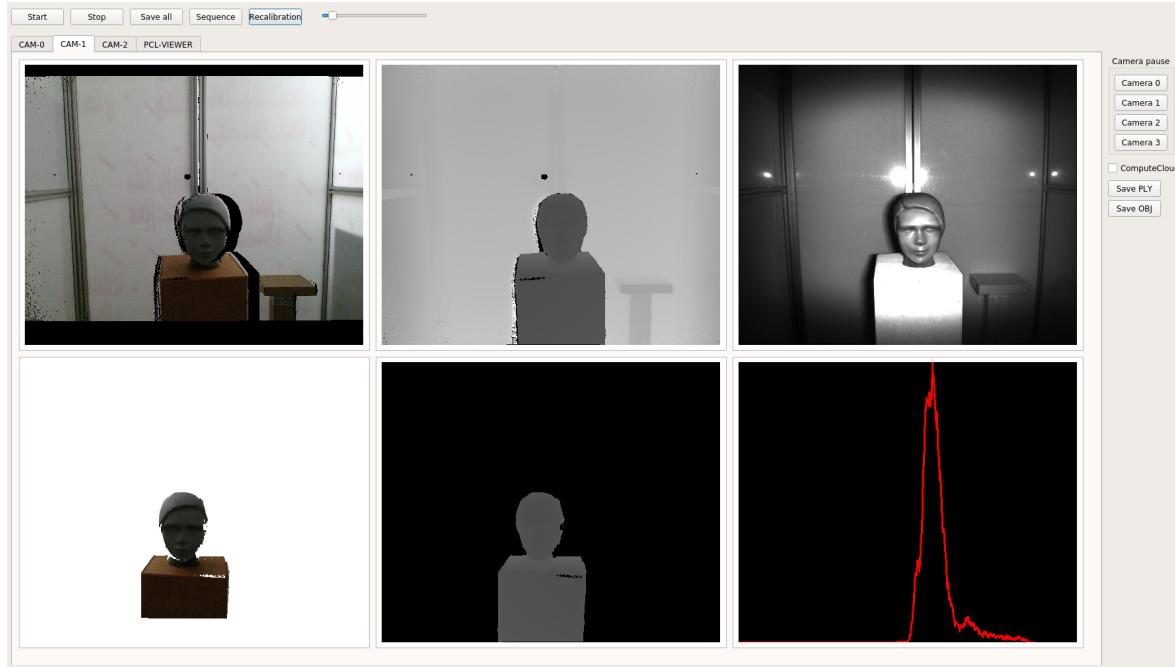


(b)

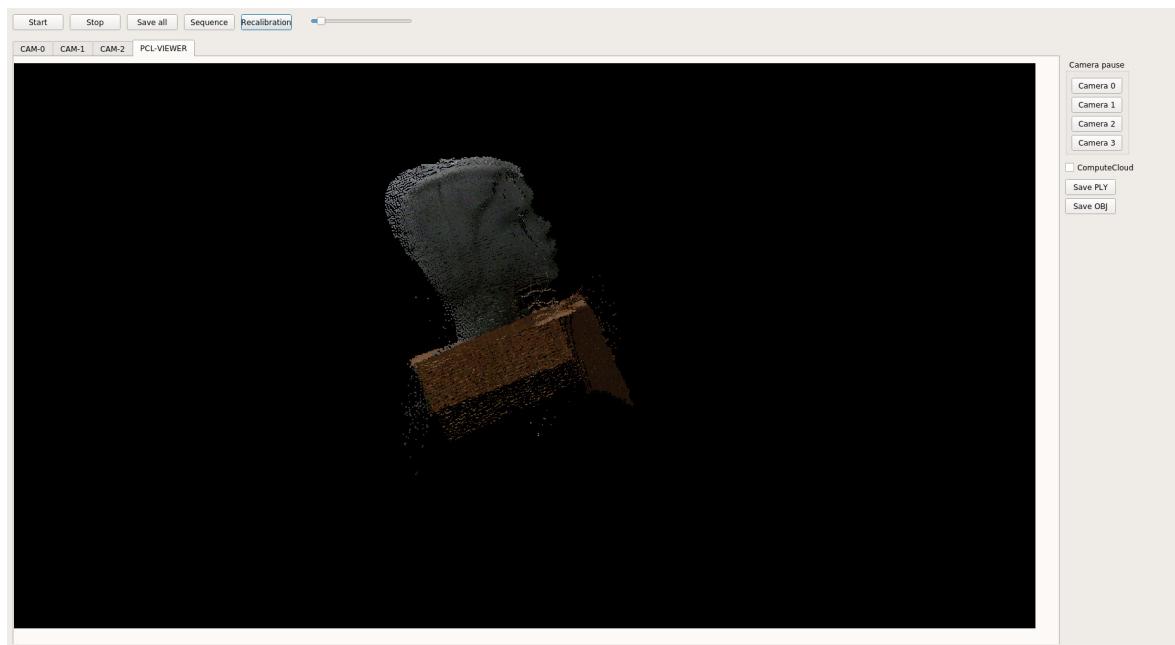
Obr. xi.ii: Zobrazenie príkladov vyšetrenia: (a) Dento-faciálne abnormality. (b) Ukážka klúčových bodov tváre.

xi.ii Grafické prostredie

Zobrazenie grafického užívateľského rozhrania, ktorý sa využíva pri vyšetrení OSA.

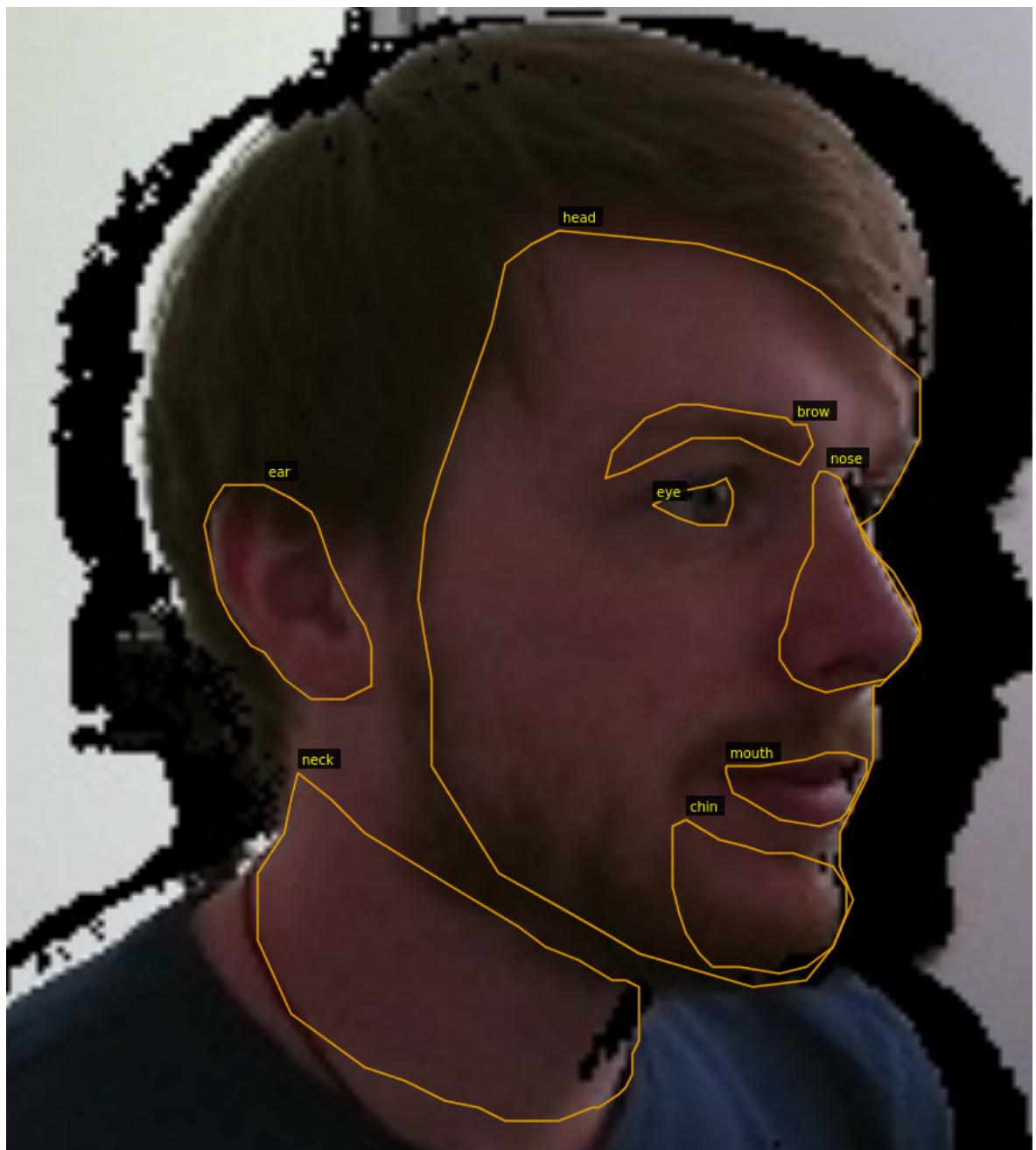


Obr. xi.iii: Zobrazenie nadhľadov spracovávaných snímok z jednotlivých kamier.



Obr. xi.iv: Zobrazenie mračna bodov v reálnom čase.

xi.iii Ukážka dátového setu pre Mask R-CNN

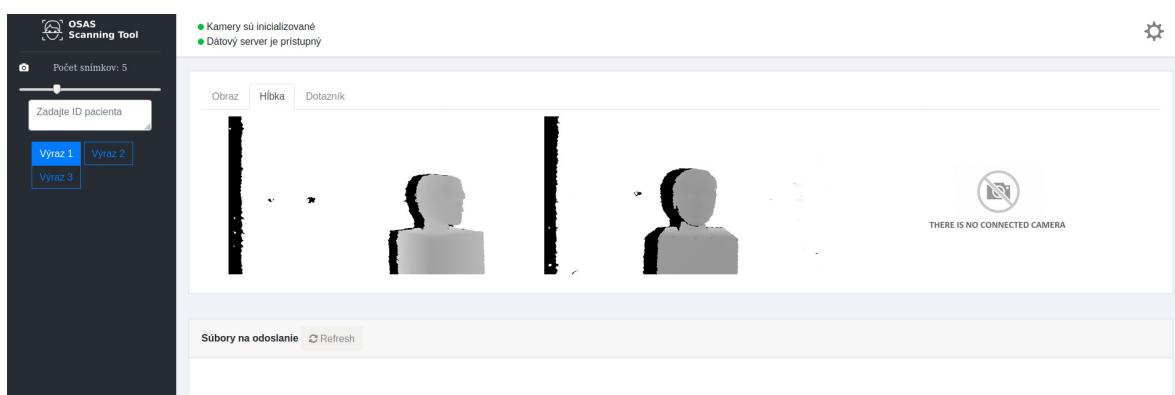


Obr. xi.v: RGB-D obraz použitý pri trénovaní Mask R-CNN .

xi.iv Grafické prostredie pre zber dát



Obr. xi.vi: Zobrazenie webového rozhrania pre zber dát, RGB obraz.



Obr. xi.vii: Zobrazenie webového rozhrania pre zber dát, hĺbkový obraz.

A screenshot of the OSAS Scanning Tool web interface showing a questionnaire form. The form consists of several input fields and dropdown menus. The fields include: Vek: 1, Výška: 60 cm, Hmotnosť: 0 kg; Obvod krku (v strede krku): 0 cm, BMI, Deviácia nosovej priehradky; Chronická obstrukcia nosa, Tonzilárna hypertrofia, Pozícia podnebia podľa Friedmana; Neurologické symptómy, Anomálie polohy čefu proti lebkej spodine, Anomálie vzájomnej polohy zubných oblúkov; Protrúzia, Úzke podnebie, Fenotyp. The sidebar on the left shows the same camera and server status as previous screenshots.

Obr. xi.viii: Zobrazenie webového rozhrania pre zber dát, dotazník



Obr. xi.ix: Konštrukcia snímacieho zariadenia použitého pre zber dát.