

Diplomová práce



České  
vysoké  
učení technické  
v Praze

**F3**

Fakulta elektrotechnická  
Katedra řídicí techniky

## Ovládání mobilního robotu gesty

**Bc. Jiří Dolejš**

Vedoucí: Ing. Jan Chudoba  
Obor: Kybernetika a robotika  
Studijní program: Kybernetika a robotika  
Květen 2020



## I. OSOBNÍ A STUDIJNÍ ÚDAJE

Příjmení: **Dolejš** Jméno: **Jiří** Osobní číslo: **439599**  
Fakulta/ústav: **Fakulta elektrotechnická**  
Zadávající katedra/ústav: **Katedra řídicí techniky**  
Studijní program: **Kybernetika a robotika**  
Studijní obor: **Kybernetika a robotika**

## II. ÚDAJE K DIPLOMOVÉ PRÁCI

Název diplomové práce:

**Ovládání mobilního robotu gesty**

Název diplomové práce anglicky:

**Mobile robot control using gestures**

Pokyny pro vypracování:

Navrhněte metodu pro rozpoznávání gest člověka robotem, primárně pro úlohu, kdy robot tohoto člověka autonomně následuje. Sledovaný člověk by měl mít možnost ovládat robot na úrovni jednoduchých povelů, jako je zastavení, pokračování v jízdě, apod. Pro rozpoznání člověka a polohy jeho těla použijte 3D kameru Kinect. V rámci práce 1) proveďte rešerši stávajících metod rozpoznávání postav, polohy částí těla a gest a používaných aplikací využívajících rozpoznávání gest 2) zpracujte rozvalu možností využití gest při ovládání mobilního robotu 3) navrhněte a implementujte metodu pro rozpoznání zvolených gest 4) vyhodnoťte spolehlivost a použitelnost navržené metody.

Seznam doporučené literatury:

[1] Kinect – Windows app development, Online: <https://developer.microsoft.com/en-us/windows/kinect> [2] E. A. Suma, B. Lange, A. S. Rizzo, D. M. Krum and M. Bolas, 'FAAST: The Flexible Action and Articulated Skeleton Toolkit,' 2011 IEEE Virtual Reality Conference, Singapore, 2011, pp. 247-248. doi: 10.1109/VR.2011.5759491 [3] Anjum M.L., Ahmad O., Rosa S., Yin J., Bona B. (2014) Skeleton Tracking Based Complex Human Activity Recognition Using Kinect Camera. In: Beetz M., Johnston B., Williams MA. (eds) Social Robotics. ICSR 2014. Lecture Notes in Computer Science, vol 8755. Springer, Cham [4] Papadopoulos G.T., Axenopoulos A., Daras P. (2014) Real-Time Skeleton-Tracking-Based Human Action Recognition Using Kinect Data. In: Gurrin C., Hopfgartner F., Hurst W., Johansen H., Lee H., O'Connor N. (eds) MultiMedia Modeling. MMM 2014. Lecture Notes in Computer Science, vol 8325. Springer, Cham

Jméno a pracoviště vedoucí(ho) diplomové práce:

**Ing. Jan Chudoba, inteligentní a mobilní robotika CIIRC**

Jméno a pracoviště druhé(ho) vedoucí(ho) nebo konzultanta(ky) diplomové práce:

Datum zadání diplomové práce: **11.07.2019**

Termín odevzdání diplomové práce: **22.05.2020**

Platnost zadání diplomové práce: **19.02.2021**

Ing. Jan Chudoba  
podpis vedoucí(ho) práce

prof. Ing. Michael Šebek, DrSc.  
podpis vedoucí(ho) ústavu/katedry

prof. Mgr. Petr Páta, Ph.D.  
podpis děkana(ky)

## III. PŘEVZETÍ ZADÁNÍ

Diplomant bere na vědomí, že je povinen vypracovat diplomovou práci samostatně, bez cizí pomoci, s výjimkou poskytnutých konzultací. Seznam použité literatury, jiných pramenů a jmen konzultantů je třeba uvést v diplomové práci.

\_\_\_\_\_  
Datum převzetí zadání

\_\_\_\_\_  
Podpis studenta





## Poděkování

Na tomto místě bych rád poděkoval Ing. Janu Chudobovi za cenné rady a ochotu při vedení mé diplomové práce.

## Prohlášení

Prohlašuji, že jsem předloženou práci vypracoval samostatně a že jsem uvedl veškeré použité informační zdroje v souladu s Metodickým pokynem o dodržování etických principů při přípravě vysokoškolských závěrečných prací.

V Praze dne .....

.....  
podpis autora práce

## Abstrakt

Tato práce se zabývá návrhem a implementací systému pro ovládání mobilního robotu statickými i dynamickými gesty obou horních končetin. Pro sledování pohybů člověka je použit senzor Microsoft Kinect v2. V práci jsou uvedeny současné metody a aplikace sledování pohybu člověka a jeho gest. Dále je popsán návrh a implementace systému ovládání, včetně použitého hardwaru a softwaru. Nakonec je vyhodnocena spolehlivost a použitelnost implementovaného systému ovládání a jsou uvedeny možnosti jeho rozšíření a využití.

**Klíčová slova:** rozpoznávání gest, ovládání gesty, snímání pohybu, Kinect, OpenNI, sledování kostry, Python

**Vedoucí:** Ing. Jan Chudoba  
Jugoslávských partyzánů 3,  
160 00 Praha 6

## Abstract

This thesis deals with the design and implementation of a system for a mobile robot control using static and dynamic gestures of both upper limbs. For the monitoring of person's movement, the Microsoft Kinect v2 sensor is used. The thesis presents current methods and applications of monitoring the movement of a person and his gestures. The design and implementation of the control system is described, including the used hardware and software. Finally, the reliability and usability of the implemented control system is evaluated and possibilities of its expansion and usage are presented.

**Keywords:** gesture recognition, gesture control, motion tracking, Kinect, OpenNI, skeleton tracking, Python

**Title translation:** Mobile robot control using gestures

## Obsah

<b>1 Úvod</b>	<b>1</b>	<b>4 Implementace</b>	<b>31</b>
<b>2 Stávající metody a aplikace</b>	<b>3</b>	4.1 Použitý software a hardware ...	31
2.1 Stávající metody .....	3	4.1.1 Software .....	31
2.1.1 Metody pro získání příznaků .	3	4.1.2 Hardware .....	33
2.1.2 Metody pro rozpoznání gest ..	8	4.2 Detekovaná kostra .....	35
2.2 Současné aplikace .....	10	4.2.1 Úprava detekované kostry ...	36
<b>3 Návrh řešení</b>	<b>13</b>	4.3 Ovládání robotu .....	37
3.1 Návrh ovládacího systému .....	13	4.4 Bezpečnostní prvky .....	38
3.2 Detekované pozice horních končetin .....	15	4.5 Konfigurovatelné prvky .....	40
3.3 Detekovaná ovládací gesta .....	17	4.6 Grafické rozhraní .....	41
3.3.1 Statická gesta .....	17	4.6.1 Základní verze GUI .....	41
3.3.2 Dynamická gesta .....	20	4.6.2 Rozšířená verze GUI .....	43
3.4 Klasifikace .....	25	<b>5 Vyhodnocení</b>	<b>47</b>
3.4.1 Klasifikace pozice horní končetiny .....	25	5.1 Vytvořený dataset .....	47
3.4.2 Klasifikace gesta .....	29	5.2 Způsob vyhodnocení .....	48
		5.3 Dosažené výsledky .....	49
		5.3.1 Výsledky klasifikace pozic horních končetin .....	49

5.3.2 Výsledky klasifikace gest . . . .	52
5.4 Problémy . . . . .	58
5.5 Možná vylepšení . . . . .	60
5.6 Možná využití . . . . .	60
<b>6 Závěr</b>	<b>61</b>
<b>Literatura</b>	<b>63</b>
<b>A Obsah přiloženého CD</b>	<b>73</b>
<b>B Použité zkratky</b>	<b>75</b>
<b>C Konfigurační soubor</b>	<b>77</b>

## Obrázky

2.1 Obleky pro sledování pohybů celého těla.....	4	2.10 Gesto pro pořízení snímku dronem. Převzato z [51]. . . . .	11
2.2 Detekce jednotlivých částí těla prostřednictvím geodetických vzdáleností. V horní řadě jsou zobrazena hloubková data. Ve spodní řadě jsou ilustrovány geodetické vzdálenosti od těžiště, červená barva značí největší vzdálenost, modrá nejmenší. Převzato z [13]. . . . .	5	3.1 Vývojový diagram algoritmu. . . . .	14
2.3 Detekovaná kostra. Světle modré body označují nalezené význačné části těla, žlutou barvou je znázorněn získaný model kostry. Převzato z [13].	5	3.2 Detekované pozice horních končetin. . . . .	16
2.4 Příznaky definované jako těžiště trojúhelníků. Převzato z [15]. . . . .	6	3.3 Gesta pro zastavení. . . . .	18
2.5 Zpracovaná data pro sledování ruky a jednotlivých prstů . . . . .	7	3.4 Gesto pro otáčení na místě. . . . .	19
2.6 Detekce polohy prstů. Převzato z [24]. . . . .	7	3.5 Gesto pro pozastavení ovládání. . . . .	19
2.7 Obleky pro sledování pohybů celého těla. . . . .	8	3.6 Gesto pro aktivaci. . . . .	20
2.8 Konvoluční neuronová síť pro klasifikaci gest ruky. Převzato z [21].	9	3.7 Gesto pro deaktivaci. . . . .	21
2.9 Sledování polohy jednotlivých prstů prostřednictvím HTC Vive. Převzato z [42]. . . . .	11	3.8 Gesto pro změnu sledovací vzdálenosti. . . . .	22
		3.9 Gesta pro jízdu vpřed. . . . .	22
		3.10 Gesta pro jízdu vzad. . . . .	24
		3.11 Relativní kartézské souřadnice. . . . .	26
		3.12 Úhlové příznaky. . . . .	27
		4.1 Senzor Kinect v2. Převzato z [71].	33
		4.2 Čtyřkolový robot Husky se senzorem Kinect v2. . . . .	34

4.3 Souřadnicový systém. Převzato a upraveno z [71]. . . . .	35
4.4 Model kostry s vyznačenými sledovanými klouby. . . . .	35
4.5 Stavový automat ovládání. . . . .	38
4.6 Vymezený úsek pro ovládání spolu se zorným polem senzoru. . . . .	39
4.7 Hlavní menu základního uživatelského rozhraní. . . . .	41
4.8 Okno s informacemi o ovládání. . . . .	42
4.9 Okno s hloubkovým obrazem. . . . .	42
4.10 Základní menu rozšířeného uživatelského rozhraní. . . . .	43
4.11 Rozhraní pro nahrávání trénovacích dat. . . . .	44
4.12 Rozhraní pro zobrazení a mazání dat. . . . .	44
5.1 Vliv povrchu pohlcujícího infračervené záření. . . . .	58
5.2 Vliv intenzivního slunečního záření. . . . .	59

## Tabulky

3.1 Pozice horních končetin tvořící gesto <i>zastavit-do strany</i> . . . . .	18
3.2 Pozice horních končetin tvořící gesto <i>zastavit-dopředu</i> . . . . .	18
3.3 Pozice horních končetin tvořící gesto <i>otáčení na místě</i> . . . . .	19
3.4 Pozice horních končetin tvořící gesto <i>pozastavit</i> . . . . .	20
3.5 Posloupnost pozic horních končetin tvořící gesto <i>aktivovat</i> . . . . .	20
3.6 Posloupnost pozic horních končetin tvořící gesto <i>deaktivovat</i> . . . . .	21
3.7 Posloupnost pozic horních končetin tvořící gesto <i>změna sledovací vzdálenosti</i> . . . . .	22
3.8 Posloupnost pozic horních končetin tvořící gesto <i>jízda vpřed</i> . . . . .	23
3.9 Posloupnost pozic horních končetin tvořící gesto <i>jízda vpřed-na straně</i> . . . . .	23
3.10 Zjednodušené posloupnosti pozic horních končetin tvořící gesto <i>jízda vpřed</i> . . . . .	23
3.11 Posloupnost pozic horních končetin tvořící gesto <i>jízda vzad</i> . . . . .	24

3.12 Posloupnost pozic horních končetin tvořící gesto <i>jízda vzad-na straně</i> .....	24	5.8 Identifikátory gest.....	53
3.13 Zjednodušené posloupnosti pozic horních končetin tvořící gesto <i>jízda vzad</i> .....	25	5.9 Matice záměn klasifikovaných gest při použití kartézských příznaků a naivního Bayesovského klasifikátoru.	53
5.1 Identifikátory pozic horní končetiny.....	50	5.10 Vyhodnocení klasifikace gest při použití kartézských příznaků a naivního Bayesovského klasifikátoru.	53
5.2 Matice záměn klasifikovaných pozic horní končetiny při použití kartézských příznaků a naivního Bayesovského klasifikátoru. ....	50	5.11 Matice záměn klasifikovaných gest při použití úhlových příznaků a naivního Bayesovského klasifikátoru.	54
5.3 Vyhodnocení klasifikace pozic horní končetiny při použití kartézských příznaků a naivního Bayesovského klasifikátoru. ....	50	5.12 Vyhodnocení klasifikace gest při použití úhlových příznaků a naivního Bayesovského klasifikátoru. ....	54
5.4 Matice záměn klasifikovaných pozic horní končetiny při použití úhlových příznaků a naivního Bayesovského klasifikátoru. ....	51	5.13 Matice záměn klasifikovaných gest při použití kartézských příznaků a vlastního klasifikátoru. ....	54
5.5 Vyhodnocení klasifikace pozic horní končetiny při použití úhlových příznaků a naivního Bayesovského klasifikátoru. ....	51	5.14 Vyhodnocení klasifikace gest při použití kartézských příznaků a vlastního klasifikátoru. ....	55
5.6 Matice záměn klasifikovaných pozic horní končetiny při použití kartézských příznaků a vlastního klasifikátoru. ....	51	5.15 Matice záměn klasifikovaných gest při použití kartézských příznaků a naivního Bayesovského klasifikátoru s ignorací pozice <i>dopředu nahoru</i> ....	55
5.7 Vyhodnocení klasifikace pozic horní končetiny při použití kartézských příznaků a vlastního klasifikátoru.	52	5.16 Vyhodnocení klasifikace gest při použití kartézských příznaků a naivního Bayesovského klasifikátoru s ignorací pozice <i>dopředu nahoru</i> ....	56
		5.17 Matice záměn klasifikovaných gest při použití úhlových příznaků a naivního Bayesovského klasifikátoru s ignorací pozice <i>dopředu nahoru</i> ....	56

5.18 Vyhodnocení klasifikace gest při použití úhlových příznaků a naivního Bayesovského klasifikátoru s ignorací pozice <i>dopředu nahoru</i> . . . . .	56
5.19 Matice záměn klasifikovaných gest při použití kartézských příznaků a vlastního klasifikátoru s ignorací pozice <i>dopředu nahoru</i> . . . . .	57
5.20 Vyhodnocení klasifikace gest při použití kartézských příznaků a vlastního klasifikátoru s ignorací pozice <i>dopředu nahoru</i> . . . . .	57





# Kapitola 1

## Úvod

Současný rozvoj technologií přináší mimo jiné i nové způsoby interakce mezi uživatelem a zařízením. Stejně jako byla v některých oborech mechanická tlačítka nahrazena dotykovými displeji, začíná se rozšiřovat i plně bezdotykové ovládání [1]. V současnosti je již možné ovládat různá zařízení prostřednictvím hlasového projevu člověka, pohybem očí nebo pohybem končetin [2]. A právě využitím pohybu horních končetin pro ovládání mobilního robotu se zabývá tato práce.

Cílem je tedy navrhnout a implementovat systém pro ovládání mobilního robotu gesty prováděnými oběma horními končetinami. Tato gesta pak mohou být jak statická, tj. nastavení končetiny do jedné z definovaných pozic, tak dynamická, tedy pohyb mezi vybranými pozicemi, případně kombinace obou typů gest při využití obou horních končetin zároveň. Rozpoznaná gesta jsou následně využita ke spuštění vybraných akcí, např. zastavení robotu, otáčení robotu na místě, následování člověka, apod. V kombinaci s určováním relativní pozice robotu a operátora by tak měl být mobilní robot schopný člověka následovat bez dalších vnějších zásahů. Tento způsob ovládání je vhodný například v situacích, kdy se člověk nemůže neustále plně věnovat ovládání robotu nebo kdy není nutné jeho přesné řízení.

Tato práce je rozdělena do 4 hlavních kapitol. První uvádí současné metody používané pro rozpoznávání postav, poloh částí těla a gest. Dále jsou v ní zmíněny aplikace využívající tyto metody. V další kapitole je uveden návrh funkce ovládacího systému, definice detekovaných pozic horních končetin a gest a jejich klasifikace. Následující kapitola uvádí použitý software a hardware, věnuje se získání a úpravě použitých dat a popisuje vlastnosti vytvořeného ovládacího systému. V poslední kapitole je popsána vyhodnocovací metodika a jsou zde uvedeny a zhodnoceny dosažené výsledky. V závěru této kapitoly jsou kromě popisu nastalých problémů navržena možná budoucí vylepšení systému a jeho případná využití.



## Kapitola 2

### Stávající metody a aplikace

Tato kapitola popisuje stávající metody využívané pro rozpoznávání postav, poloh částí těla a gest. Dále představuje využití těchto metod v různých aplikacích.

#### 2.1 Stávající metody

Tato podkapitola se věnuje současným metodám pro nalezení a sledování poloh částí těla a celých postav a jejich využití pro rozpoznávání provedených gest. Je rozdělena do dvou hlavních částí. V té první jsou uvedeny a stručně popsány metody používané pro nalezení požadovaných objektů a následné získání vhodných příznaků, tzn. získání vhodného popisu pozic sledovaných částí těla. Druhá část se poté zabývá nejčastějšími postupy pro klasifikaci těchto příznaků.

##### 2.1.1 Metody pro získání příznaků

V následující části budou uvedeny dva základní způsoby pro získání dat vhodných k nalezení a sledování osob či částí těla. Prvním z nich je využití optického systému pro získání obrazových dat, ve kterých je poté prováděna detekce hledaných objektů. Druhým způsobem je umístění senzorů přímo na

tělo sledované osoby a využití signálů z těchto senzorů coby příznaků.

Při použití optických systémů je možné využít speciální značky usnadňující sledování pohybu osob či částí jejich těl. Může se jednat o vhodně umístěné reflexivní prvky (na Obrázku 2.1a) nebo geometrické vzory (na Obrázku 2.1b), případně kombinaci obojího.



**(a)** : Oblek s reflexivními značkami. Převzato a upraveno z [4].

**(b)** : Oblek s geometrickými značkami. Převzato a upraveno z [5].

**Obrázek 2.1:** Obleky pro sledování pohybů celého těla.

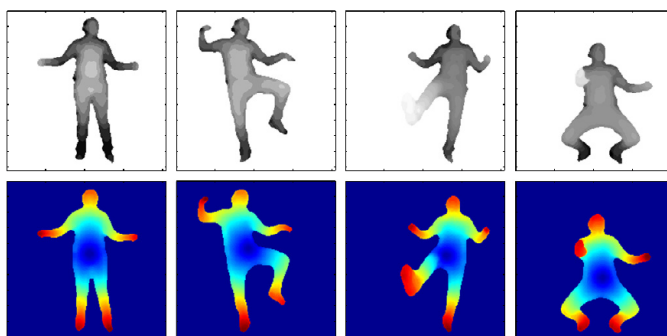
Tyto značky usnadňují sledování, protože je předem znám jejich vzhled a použité algoritmy tak mohou být uzpůsobeny přímo pro jejich detekci. Díky tomu, že je předem známá také poloha značek na těle sledované osoby, je po jejich nalezení v obraze získána množina bodů, do které je poté možné vhodným způsobem vložit model kostry a získat tak 3D reprezentaci pohybu sledované osoby [3]. Situace je obtížnější, pokud není možné či žádoucí umístit na sledovanou osobu některou z těchto značek. V takových případech je možné využít jednu z následujících metod.

Pokud je k dispozici pouze obraz ze standardní RGB kamery, používají se ve velké míře neuronové sítě. V [6] úspěšně využívají kaskádu hlubokých neuronových sítí (DNN, Deep Neural Network) pro určení polohy kloubů končetin ve dvourozměrných barevných obrazech. V [7] poté kombinují konvoluční neuronovou síť pro nalezení 2D a 3D souřadnic kloubů s následným vložením modelu kostry do těchto souřadnic, díky čemuž dokáží v reálném čase rekonstruovat 3D pohyb sledované osoby.

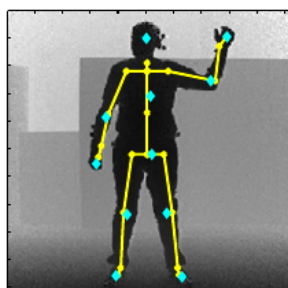
Pro sledování pohybu osob či částí těl je možné využít také hloubkové kamery. Ty dokáží určit vzdálenost objektů ve scéně a prezentovat tato data ve formě 2D obrazu. Hodnoty pixelů tohoto obrazu poté reprezentují změřené vzdálenosti od odpovídající části scény [8]. Pokud je tedy využita hloubková kamera a stačí určit polohu osoby či její siluetu, je možné použít segmentaci obrazu podle vzdálenosti od senzoru (v [9]). Bez dalších opatření však mohou

v takových případech vznikají falešné detekce, když se v záběru objeví cizí objekt blíže než detekovaná osoba. Získané výsledky tak nemusí být vždy spolehlivé. V [10] proto provádějí pokročilou segmentaci, kdy nejprve prostřednictvím 2D a 3D modelu hlavy naleznou hlavy osob v záběru a poté provedou expanzi daného prostorového regionu, dokud neobsáhnou celou oblast odpovídající detekované osobě. Díky tomuto postupu snižují možnost falešných detekcí. Nevýhodou této metody je potřeba rozpoznat hlavu osoby. Hlava tedy musí být dostatečně viditelná a neměly by se na ní nacházet zvláště tvarované pokrývky (např. velké klobouky), které by znemožňovaly detekci.

I v případě použití hloubkových dat je možné získat model kostry využitím neuronové sítě, jak je demonstrováno v [11] a [12]. Odlišný přístup zvolili v [13], kde ze získaných 3D dat nejprve vytváří graf spojující jednotlivé 3D body a následně určují polohu význačných částí těla (hlava, horní a dolní končetiny, atd.) na základě geodetické vzdálenosti vrcholů grafu od těžiště (viz Obrázek 2.2). Nakonec je podle zjištěných pozic modifikován model kostry obsahující 16 kloubů (Obrázek 2.3). Výhodou této metody je, že oproti neuronovým sítím nevyžaduje předešlé učení.

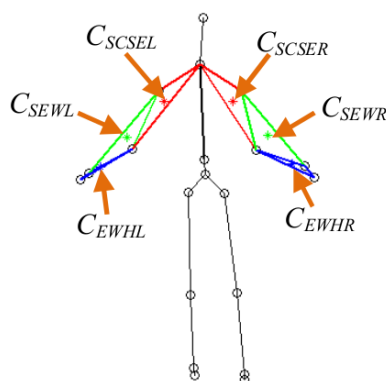


**Obrázek 2.2:** Detekce jednotlivých částí těla prostřednictvím geodetických vzdáleností. V horní řadě jsou zobrazena hloubková data. Ve spodní řadě jsou ilustrovány geodetické vzdálenosti od těžiště, červená barva značí největší vzdálenost, modrá nejmenší. Převzato z [13].



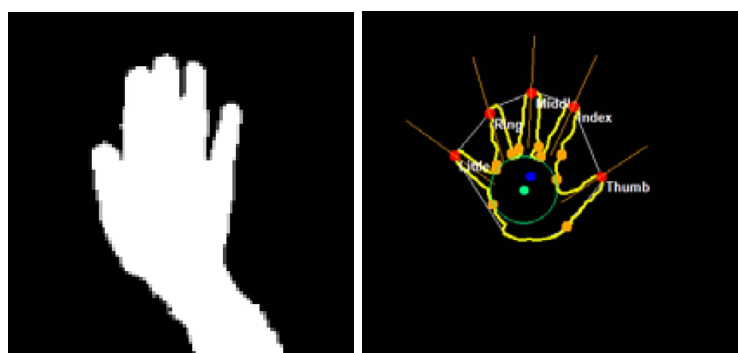
**Obrázek 2.3:** Detekovaná kostra. Světle modré body označují nalezené význačné části těla, žlutou barvou je znázorněn získaný model kostry. Převzato z [13].

Po získání odhadu kostry jsou jako příznaky často využívány prostorové souřadnice jejich jednotlivých kloubů. V [14] však namísto toho zvolili pro popis poloh horních končetin 8 vhodně nadefinovaných úhlů. V [15] pak místo souřadnic používají příznaky definované jako vzdálenosti těžišť šesti definovaných trojúhelníků od kloubu odpovídajícímu krku. Definované trojúhelníky a jejich těžiště jsou zobrazeny na Obrázku 2.4.



**Obrázek 2.4:** Příznaky definované jako těžiště trojúhelníků. Převzato z [15].

Pokud není záměrem sledovat celou postavu, ale pouze polohu dlaně a nastavení jednotlivých prstů, je možné pro získání příznaků využít segmentaci obrazu na základě barvy či textury pro oddělení obrazu ruky od pozadí [16], [17], [18]. Při použití hloubkového obrazu lze v některých případech předpokládat, že detekovaná ruka je nejbližší objekt ke kameře a segmentaci lze provést prahováním na základě vzdálenosti objektu od senzoru. Tento princip byl použit například v [19]. Z pořízených dat se po provedeném předzpracování získají příznaky například ve formě siluety ruky (v [20], viz Obrázek) či jejího binárního obrazu v kombinaci s hloubkovými daty (v [21]). Takto získaná data už mohou být použita pro klasifikaci, případně z nich lze získat odlišné typy příznaků. V [22] tak například z binárního obrazu ruky vypočítávají příznaky ve formě Huových momentů. Odlišným přístupem použitým v [23] je získání příznaků ve formě souřadnic špiček jednotlivých prstů, které jsou nalezeny v množině bodů ležících zároveň na okraji oblasti odpovídající sledované ruce a na okraji konvexního obalu této oblasti.

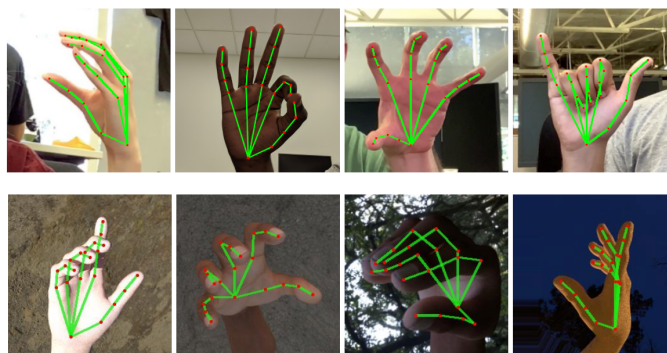


(a) : Získaná silueta ruky.  
Převzato z [20].

(b) : Nalezení pozic špiček prstů  
(označeno červenou barvou). Převzato z [23].

**Obrázek 2.5:** Zpracovaná data pro sledování ruky a jednotlivých prstů

Výše uvedené metody jsou dostatečné pro sledování ruky ve 2D. Pokud je však třeba zachytit komplexní pohyb ruky ve 3D, je vhodnější využít metody schopné modelovat konfiguraci sledované části těla. Například v [24] uplatňují strojové učení pro nalezení a sledování 21 význačných bodů ruky, kterými definují její nastavení a nastavení všech prstů (viz Obrázek 2.6). Jako příznaky poté využívají úhly mezi definovanými klouby, podle kterých určují stav jednotlivých prstů.



**Obrázek 2.6:** Detekce polohy prstů. Převzato z [24].

Odlišným postupem pro získání poloh částí těl je využití senzorů přímo na těle sledované osoby. V takovém případě nejsou potřeba žádné kamerové systémy, ale různé senzory jsou součástí rukavic (viz Obrázek 2.7b) či celých obleků (viz Obrázek 2.7a). Tento princip úspěšně využívá například [25] při umístění inerciálních měřicích senzorů do trička za účelem sledování pohybů horní poloviny těla. Podobný princip je využit v [26] pro sledování pozic jednotlivých prstů umístěním inerciálních a magnetometrických senzorů do rukavice. V těchto metodách jsou výslednými příznaky použitými pro

klasifikaci natočení a zrychlení naměřená prostřednictvím využitých senzorů.



(a) : Tričko s inerciálními měřicími senzory. Převzato z [25].

(b) : Rukavice s inerciálními měřicími senzory. Převzato z [26].

**Obrázek 2.7:** Obleky pro sledování pohybů celého těla.

Zvolení vhodné metody tedy závisí na konkrétním zamýšleném použití. Sensory umístěné přímo na těle je vhodné aplikovat v situacích, kdy není možné používat externí optické systémy, ať už z důvodu nedostatečného okolního prostoru, či naopak kvůli velké oblasti pohybu osoby. Výhodou těchto metod je, že nevyžadují pohyb osoby pouze ve vymezeném prostoru a nehrozí u nich ztráta sledování některých částí těl kvůli jejich zakrytí. Nevýhodou je poté nutnost umístění senzorů přímo na těle osoby, včetně veškeré obslužné elektroniky a napájení, což může být v některých případech nežádoucí.

V tomto ohledu je výhodnější využití sledování prostřednictvím optického systému, kde sledovaná osoba nemusí být označena vůbec, případně pouze jednoduchými pasivními značkami. Nevýhodou vizuálního sledování je nutnost pohybu pouze v prostoru vymezeném zorným polem kamery a potřeba viditelnosti sledované osoby.

### 2.1.2 Metody pro rozpoznání gest

Po získání příznaků popisujících sledované části těla je třeba tato data nějakým způsobem klasifikovat a zjistit, jaké gesto bylo provedeno. K tomu se využívá řada různých metod, zmíním zde ty nejrozšířenější.

Prvním druhem používaných metod je prosté určení provedeného gesta na základě kontroly splněných podmínek. V [23] tak ze souřadnic špiček jednotlivých prstů určí počet a druh natažených prstů a jejich vzájemné úhly. Obdobný postup je zvolen v [24], kde součtem úhlů mezi jednotlivými články prstů rozhodnou, zda je daný prst skrčený či natažený. V obou případech je poté podle stavu všech prstů určeno provedené gesto.

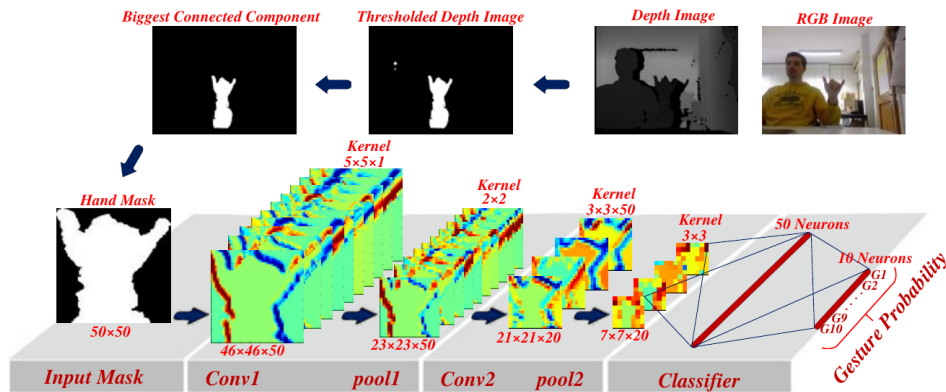
Pokud je třeba využít pokročilejší způsob klasifikace gest, může být použit naivní Bayesovský klasifikátor. Ten vychází z Bayesovy věty o podmíněné



pravděpodobnosti a provádí klasifikaci hledáním třídy s maximální aposteriorní pravděpodobností při daném pozorování [27]. Tento klasifikátor je použit například v [28] pro klasifikaci gest prováděných rozhodčím kriketu a jako jeden z klasifikátorů pozic horních končetin je využit i v této práci. Podrobnější popis jeho funkce je tak uveden v Podkapitole 3.4.1.

Další využívanou metodou je tzv. metoda podpůrných vektorů (SVM, Support Vector Machine). Tato metoda užívá rozdělicí nadrovinu pro oddělení vzorků ze dvou tříd při zachování co největšího odstupů vzorků a nadrovinu [29]. Původně je však určena pouze k binární klasifikaci. Pro klasifikaci dat do  $N$  tříd, kde  $N$  je větší než 2, je používána v tzv. režimu „one-against-all“, kdy je vytvořeno  $N$  klasifikátorů, každý pro jedno gesto. Vstupní data jsou poté klasifikována všemi klasifikátory a jako výsledná klasifikace je vybrána třída, jejíž SVM klasifikátor dosáhl nejvyšší věrohodnosti [30]. Tato metoda je úspěšně využita například v [31], [32].

Pokud je cílem zpracovávat přímo získaná binární obrazová data, je možné použít konvoluční neuronové sítě (CNN). Tato metoda je zvolena například v [21], kde byla implementována CNN pro klasifikace 10 různých gest. Tato vytvořená konvoluční neuronová síť je zobrazena na Obrázku 2.8



**Obrázek 2.8:** Konvoluční neuronová síť pro klasifikaci gest ruky. Převzato z [21].

Předchozí metody byly použity převážně pro klasifikaci gest v jednom snímku, vhodným rozšířením obslužných struktur je však možné je využívat i při detekci déle trvajících gest. Následující metody jsou pak cíleny přímo na dynamická gesta.

První zmíněnou metodou je tzv. metoda dynamického borcení času (DTW, Dynamic Time Warping) využitá v [33] a [34]. Tato metoda dokáže určit podobnost dvou časových průběhů příznaků i když mají různou dobu trvání. Při klasifikaci jsou poté naměřená data porovnávána se vzorovými průběhy definovaných gest a výstupem klasifikace je gesto, jehož průběhu jsou naměřená data dle zvolené metriky nejpodobnější.

Dalším způsobem vhodným pro klasifikaci dynamických gest je použití neuronové sítě pracující s časovou informací, např. neuronové sítě s časovým

zpožděním (TDNN, Time-Delay Neural Network). Tento druh neuronové sítě je úspěšně využit pro detekci dynamických gest v [35] a [36].

Pravděpodobně nejrozšířenějšími metodami pro klasifikaci dynamických gest jsou metody založené na skrytých Markovových modelech (HMM, Hidden Markov Model). Ty jsou v různých konfiguracích využity například v [14], [22] a [37]. Základní princip HMM spočívá ve vytvoření statistického modelu systému s nepozorovatelnými stavy [38]. Pro každé detekované gesto je během učení klasifikátoru vytvořen jeden HMM a při detekci je poté zjišťována věrohodnost tvrzení, že daný HMM mohl vygenerovat získaná vstupní data. Výstupem klasifikace je poté gesto, jehož model dosáhl největší věrohodnosti [39].

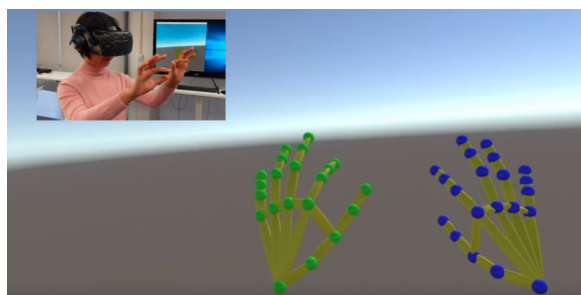
S ohledem na širokou škálu používaných příznaků tedy nelze jednoznačně určit, která metoda je pro klasifikaci gest nejlepší. Vždy záleží na konkrétním případě.

## 2.2 Současné aplikace

Využití detekce gest pro interakci a ovládání různých zařízení už nalezlo uplatnění v celé řadě oblastí [1]. V této podkapitole bude zmíněno několik případů využití tohoto způsobu ovládání v praxi.

Prvním příkladem je originální verze senzoru Kinect. Ta byla původně vyvinuta jako periferie k herní konzoli Xbox 360, kde sloužila k ovládání systému a samotných her [40]. Prováděním gest a pohybem celého těla tak mohl uživatel hrát některé hry či ovládat uživatelské rozhraní konzole.

Rozpoznání a sledování částí těla je již delší dobu využíváno v aplikacích virtuální reality. Jednalo se však převážně o sledování celkového pohybu těla a paží a až v nedávné minulosti došlo k rozšíření sledování i na jednotlivé prsty. HTC například v roce 2019 uvolnilo pro vývojáře SDK poskytující možnosti pro sledování polohy rukou a hlavně jednotlivých prstů (viz Obrázek 2.9). To umožňuje vytvoření uživatelského rozhraní, které nevyžaduje žádné ovládací zařízení a může tak být intuitivnější. V kombinaci s monitorováním pohybu celého těla je tak uživatel schopen pohybovat se ve virtuálním prostředí a provádět pokročilou interakci s okolím [41].



**Obrázek 2.9:** Sledování polohy jednotlivých prstů prostřednictvím HTC Vive. Převzato z [42].

Pokud zůstaneme u rozpoznávání gest pouze samotné ruky, můžeme nalézt uplatnění například při ovládání multimediálních systémů u nových automobilů, kde může uživatel pohybem ruky před senzorem přeladit rádiovou stanicí či změnit hlasitost audiosystému, aniž by musel spustit oči ze silnice před sebou [43]. Náročnější úlohou je například využití rozpoznání gest ruky pro překlad znakové řeči, což je úspěšně demonstrováno v pracích [44] a [45].

Dalším oborem vhodným pro použití ovládání gesty je lékařství, a to z důvodu absence fyzického kontaktu mezi uživatelem a zařízením při interakci, což snižuje riziko kontaminace [46]. Tato vlastnost byla využita v [47] při vytvoření systému s vizuální detekcí gest, prostřednictvím kterých může lékař manipulovat s daty z magnetické rezonance během operace, aniž by se musel fyzicky dotýkat ovládacích prvků. Odlišný přístup k získání dat pro detekci gest zvolili v [46], kde vytvořili zařízení ve formě náramku, který dokáže monitorovat pohyb ruky, a tím umožnit lékaři bezdotykově ovládat potřebná zařízení.

Jak již bylo zmíněno v úvodu této práce, jejím cílem je využít detekci gest pro ovládání mobilního robotu. Obdobným problémem už se různými způsoby zabývali například v [48] či [49]. Podobný způsob ovládání už je využíván i na reálných produktech. Například některé drony od společnosti DJI využívají svou kameru pro detekci gest provedených uživatelem, jejichž prostřednictvím lze ovládat fotoaparát, měnit polohu dronu či spustit režim autonomního následování [50]. Způsob použití gesta pro pořízení snímku můžeme vidět na Obrázku 2.10



**Obrázek 2.10:** Gesto pro pořízení snímku dronem. Převzato z [51].

Na uvedených příkladech je možné pozorovat, že množství aplikací využívajících gesta jako způsob interakce se neustále zvyšuje. Podle některých zdrojů leží budoucnost interakce člověka s osobními počítačovými zařízeními v přirozených uživatelských rozhraních (NUI), které budou ve velké míře využívat právě detekci gest [52].

## Kapitola 3

### Návrh řešení

V této kapitole je zpracován návrh ovládacího systému, jsou zde uvedeny jednotlivé detekované pozice horních končetin a gesta, která se z těchto pozic skládají. Součástí kapitoly je také popis klasifikátorů pozic horních končetin a klasifikátoru gest.

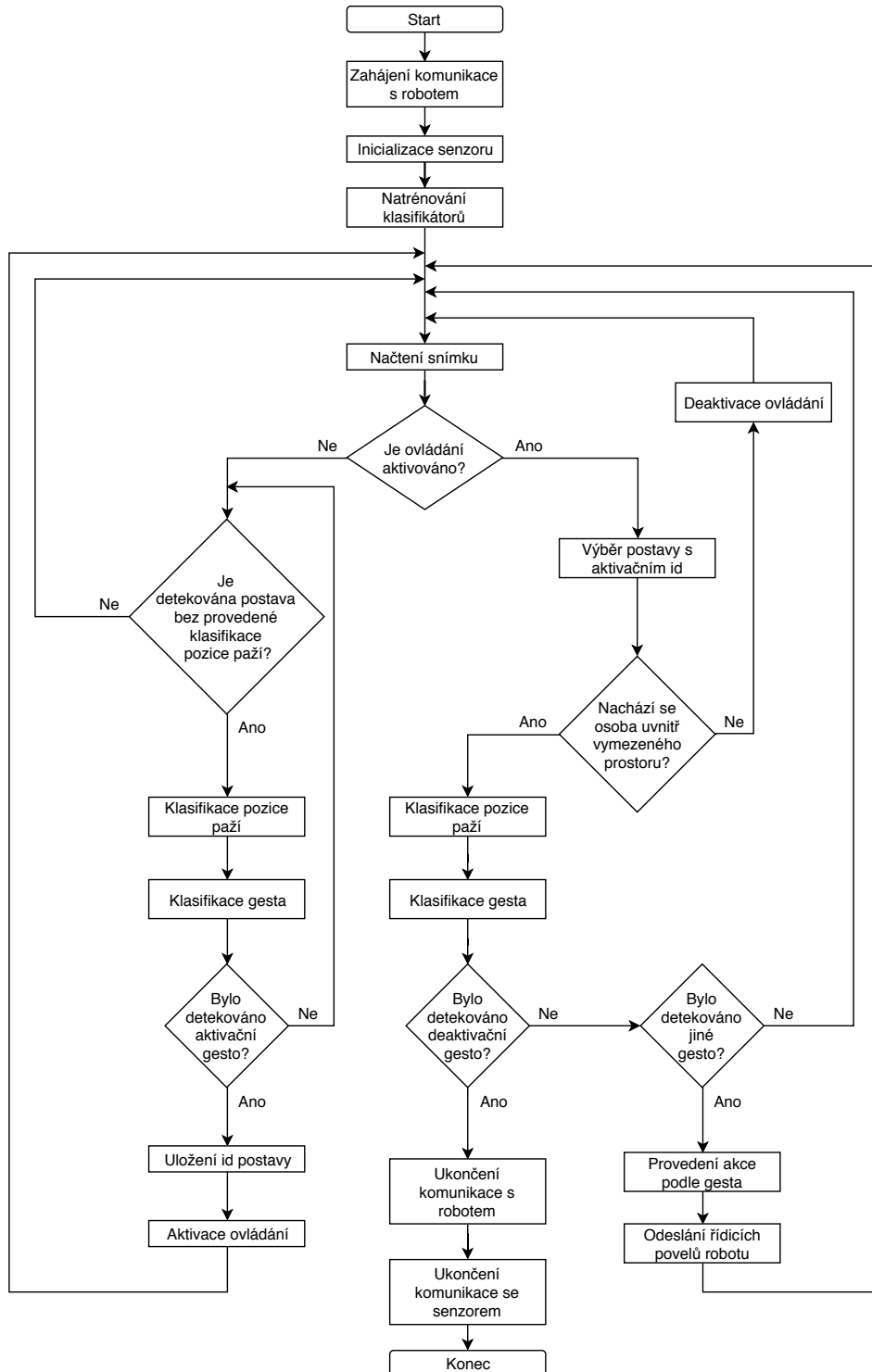
#### 3.1 Návrh ovládacího systému

Ovládací systém byl navržen pro mobilní robot schopný pohybovat se vpřed, vzad a otáčet se na místě. Ovládání je realizováno prostřednictvím jednoduchých gest prováděných horními končetinami. Způsob detekce gest je inspirován prací [28].

Základní činnost ovládacího systému spočívá v cyklickém načítání snímků z použitého senzoru. Při zpracování tohoto snímku jsou nejprve použitým softwarem detekovány postavy v záběru a následně jsou získány souřadnice jednotlivých sledovaných částí těla od každé detekované postavy. Po vhodné úpravě těchto dat dojde k jejich klasifikaci do tříd odpovídajících nadefinovaným pozicím horní končetiny (podrobně popsány v Podkapitole 3.2). Na základě klasifikovaných pozic horních končetin jsou poté detekována jednotlivá gesta (podrobně popsána v Podkapitole 3.3), která způsobují změny stavu systému. Na konci každého cyklu jsou poté vypočítány výstupní hodnoty určené pro řízení robotu. Tyto hodnoty závisí na současném stavu systému a na vzájemné pozici robotu a operátora.

Pro zamezení možnosti nežádoucích vstupů od více osob je zavedena nutnost použití aktivačního gesta pro umožnění ovládacího robotu. Po provedení

tohoto gesta již systém dále zpracovává pouze pohyby osoby, která dané gesto vykonala a činí tak, dokud osoba neprovede deaktivální gesto nebo dokud nedojde ke ztrátě vizuálního kontaktu s danou osobou.



Obrázek 3.1: Vývojový diagram algoritmu.

Na Obrázku 3.1 je zobrazen vývojový diagram popisující základní činnost ovládacího algoritmu. Jednotlivé části jsou podrobněji popsány následujících částech práce, převážně v Kapitole 4.

## 3.2 Detekované pozice horních končetin

Jako základní konstrukční bloky pro ovládací gesta bylo zvoleno celkem 10 význačných pozic horních končetin. Na Obrázku 3.2 jsou jednotlivé pozice demonstrovány na pravé horní končetině modelu. Samotné pozice byly zvoleny tak, aby je bylo možné spolehlivě detekovat i na větší vzdálenost, aby byly vzájemně dostatečně odlišné a aby pokrývaly co největší rozsah. Vzhledem ke způsobu sledování a následné klasifikaci jsou k popisu pozic horních končetin použity souřadnice ramen, loktů a zápěstí. Na určení pozice tedy nemá vliv natočení ruky v zápěstí. Zobrazené modely postavy byly vytvořeny prostřednictvím webové aplikace [53].



(a) : Dopředu dolů.



(b) : Dopředu.



(c) : Dopředu nahoru.



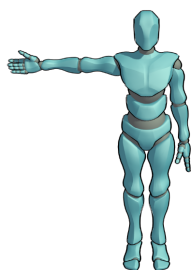
(d) : Dopředu kolmo.



(e) : Dolů.



(f) : Do strany dolů.



(g) : Do strany.



(h) : Do strany nahoru.



(i) : Nahoru.



(j) : Do strany kolmo.

**Obrázek 3.2:** Detekované pozice horních končetin.



## ■ 3.3 Detekovaná ovládací gesta

Gesta byla spolu s ovládacími operacemi navržena s ohledem na vlastnosti robotu, pro který byl systém určen a na kterém byl také testován (více informací o robotu bude uvedeno v Podkapitole 4.1.2). Vzhledem k různým definicím a odlišnostem následného zpracování se gesta dělí na dva typy, statická a dynamická. Celkem je prostřednictvím gest možné provádět 8 ovládacích operací. Některé z navržených operací je možné aktivovat dvěma různými gesty, celkový počet detekovaných gest je tedy 11. Možnost provedení některých akcí dvěma gesty byla zavedena pro zajištění větší variability a potlačení problému uvedených v Podkapitole 5.4.

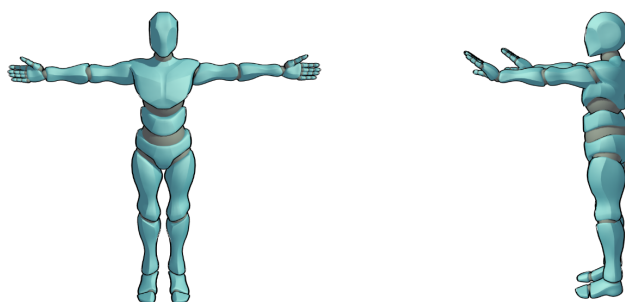
### ■ 3.3.1 Statická gesta

Statická gesta jsou tvořena nastavením obou horních končetin do požadovaných pozic. Hlavním důvodem požadavku na použití obou končetin je snížení pravděpodobnosti náhodného provedení gesta při běžném pohybu operátora.

Níže jsou uvedeny jednotlivé akce, které je možné prostřednictvím statických gest provést.

#### ■ Zastavení

Provedením jednoho ze dvou uvedených gest dojde k zastavení veškerého pohybu robotu. Obě gesta umožňující zastavení robotu jsou zobrazena na Obrázku 3.3. V Tabulkách 3.1 a 3.2 jsou poté uvedeny pozice horních končetin, kterými jsou tato gesta definována.



(a) : Zastavit-do strany.

(b) : Zastavit-dopředu.

**Obrázek 3.3:** Gesta pro zastavení.

Levá horní končetina	Pravá horní končetina
do strany	do strany

**Tabulka 3.1:** Pozice horních končetin tvořící gesto *zastavit-do strany*.

Levá horní končetina	Pravá horní končetina
dopředu	dopředu

**Tabulka 3.2:** Pozice horních končetin tvořící gesto *zastavit-dopředu*.

### ■ Otáčení na místě

Provedením gesta pro otáčení (viz Obrázek 3.4 a Tabulka 3.3) dojde k zastavení dopředného pohybu robotu, nadále je umožněno pouze otáčení na místě směrem za operátorem.



**Obrázek 3.4:** Gesto pro otáčení na místě.

Levá horní končetina	Pravá horní končetina
do strany kolmo	do strany kolmo

**Tabulka 3.3:** Pozice horních končetin tvořící gesto *otáčení na místě*.

### ■ Pozastavení ovládní

Tímto gestem (Obrázek 3.5 a Tabulka 3.4) dojde k zastavení pohybu robotu a pozastavení ovládní. Veškerá gesta mimo aktivačního jsou ignorována, avšak osoba, která provedla pozastavení, je v rámci zorného pole senzoru stále sledována a pokud nedojde ke ztrátě vizuálního kontaktu, může pouze tato osoba znovu provést aktivaci ovládní.



**Obrázek 3.5:** Gesto pro pozastavení ovládní.

Levá horní končetina	Pravá horní končetina
nahoru	nahoru

**Tabulka 3.4:** Pozice horních končetin tvořící gesto *pozastavit*.

### 3.3.2 Dynamická gesta

Dynamická gesta jsou definována jako posloupnosti několika detekovaných pozic horní končetiny. Následující navržená gesta kombinují dynamická a statická gesta, statické gesto je však vyhodnocováno až spolu s poslední pozicí dynamického gesta.

Níže jsou uvedeny akce, které je možné provést prostřednictvím dynamických gest.

#### Aktivace ovládání

Uvedené gesto (Obrázek 3.6 a Tabulka 3.5) slouží k aktivaci ovládání po prvním spuštění nebo po pozastavení ovládání.



**Obrázek 3.6:** Gesto pro aktivaci.

Krok	Levá horní končetina	Pravá horní končetina
1	–	nahoru
2	–	nahoru do strany
3	–	nahoru
4	dolů	nahoru do strany

**Tabulka 3.5:** Posloupnost pozic horních končetin tvořící gesto *aktivovat*.

### ■ Deaktivace ovládání

Tímto gestem (Obrázek 3.7 a Tabulka 3.6) dojde k deaktivaci ovládání. Robot tedy zastaví na místě a aktivaci ovládání může případně provést jiný uživatel.



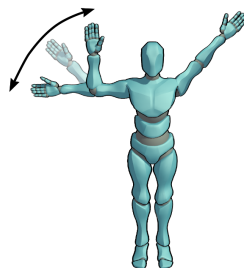
**Obrázek 3.7:** Gesto pro deaktivaci.

Krok	Levá horní končetina	Pravá horní končetina
1	–	nahoru
2	–	nahoru do strany
3	–	nahoru
4	do strany kolmo	nahoru do strany

**Tabulka 3.6:** Posloupnost pozic horních končetin tvořící gesto *deaktivovat*.

### ■ Změna sledovací vzdálenosti

Toto gesto (Obrázek 3.8 a Tabulka 3.7) slouží ke změně vzdálenosti, kterou robot udržuje během následování operátora. Při provedení gesta je aktuální vzdálenost mezi robotem a operátorem uložena a k její změně na výchozí hodnotu dojde při deaktivaci ovládání.



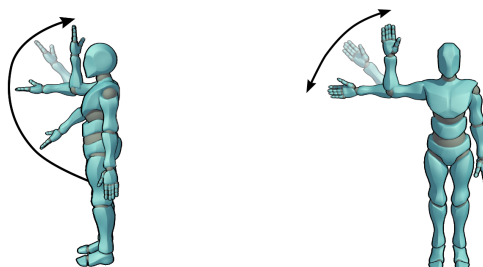
**Obrázek 3.8:** Gesto pro změnu sledovací vzdálenosti.

Krok	Levá horní končetina	Pravá horní končetina
1	–	do strany
2	–	do strany kolmo
3	–	do strany
4	do strany nahoru	do strany kolmo

**Tabulka 3.7:** Posloupnost pozic horních končetin tvořící gesto *změna sledovací vzdálenosti*.

### ■ Jízda vpřed

Po provedení některého z gest zobrazených na Obrázku 3.9 bude robot následovat operátora a snažit se při tom udržovat definovanou vzdálenost. Může se však pouze otáčet za operátorem a pojíždět vpřed. Pokud se tedy operátor přesune blíže než na definovanou vzdálenost, robot se za ním bude pouze otáčet.



(a) : Jízda vpřed.

(b) : Jízda vpřed-na straně.

**Obrázek 3.9:** Gesta pro jízdu vpřed.

Krok	Levá horní končetina	Pravá horní končetina
1	–	dolů
2	–	dopředu dolů
3	–	dopředu
4	dolů	dopředu kolmo

**Tabulka 3.8:** Posloupnost pozic horních končetin tvořící gesto *jízda vpřed*.

Krok	Levá horní končetina	Pravá horní končetina
1	–	do strany
2	–	do strany kolmo
3	–	do strany
4	dolů	do strany kolmo

**Tabulka 3.9:** Posloupnost pozic horních končetin tvořící gesto *jízda vpřed-na straně*.

K původně definovaným posloupnostem (Tabulky 3.8 a 3.9) byly z důvodu časté ztráty sledování horní končetiny při jejím rychlém pohybu směrem k senzoru přidány posloupnosti vynechávající jednu ze dvou prostředních pozic. Tyto zjednodušené definice jsou uvedeny v Tabulce 3.10.

Krok	Levá horní končetina	Pravá horní končetina
1	–	dolů
2	–	dopředu dolů
3	dolů	dopředu kolmo

(a) : Zjednodušená posloupnost vynechávající pozici *dopředu*.

Krok	Levá horní končetina	Pravá horní končetina
1	–	dolů
2	–	dopředu
3	dolů	dopředu kolmo

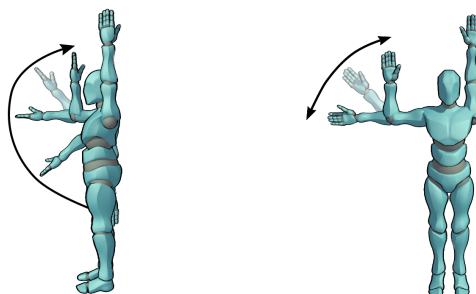
(b) : Zjednodušená posloupnost vynechávající pozici *dopředu dolů*.

**Tabulka 3.10:** Zjednodušené posloupnosti pozic horních končetin tvořící gesto *jízda vpřed*.

## ■ Jízda vzad

Po provedení jednoho z gest z Obrázku 3.10 bude robot opět udržovat definovanou vzdálenost, v tomto případě se však může pouze otáčet a pojíždět vzad.

Pokud se tedy operátor přesune blíže než na definovanou vzdálenost, robot od něho bude couvat. Jestliže se operátor naopak vzdálí robot se bude pouze otáčet. Posloupnost tvořící tato gesta jsou uvedena v Tabulkách 3.11 a 3.12.



(a) : Jízda vzad.

(b) : Jízda vzad-na straně.

**Obrázek 3.10:** Gesta pro jízdu vzad.

Krok	Levá horní končetina	Pravá horní končetina
1	–	dolů
2	–	dopředu dolů
3	–	dopředu
4	nahoru	dopředu kolmo

**Tabulka 3.11:** Posloupnost pozic horních končetin tvořící gesto *jízda vzad*.

Krok	Levá horní končetina	Pravá horní končetina
1	–	do strany
2	–	do strany kolmo
3	–	do strany
4	nahoru	do strany kolmo

**Tabulka 3.12:** Posloupnost pozic horních končetin tvořící gesto *jízda vzad-na straně*.

Stejně jako v případě předchozí akce byly i zde použity také zjednodušené posloupnosti pro definici gesta *jízda vzad*. Tyto posloupnosti jsou uvedeny v Tabulce 3.13.



Krok	Levá horní končetina	Pravá horní končetina
1	–	dolů
2	–	dopředu dolů
3	nahoru	dopředu kolmo

(a) : Zjednodušená posloupnost vynechávající pozici *dopředu*.

Krok	Levá horní končetina	Pravá horní končetina
1	–	dolů
2	–	dopředu
3	nahoru	dopředu kolmo

(b) : Zjednodušená posloupnost vynechávající pozici *dopředu dolů*.

**Tabulka 3.13:** Zjednodušené posloupnosti pozic horních končetin tvořící gesto *jízda vzad*.

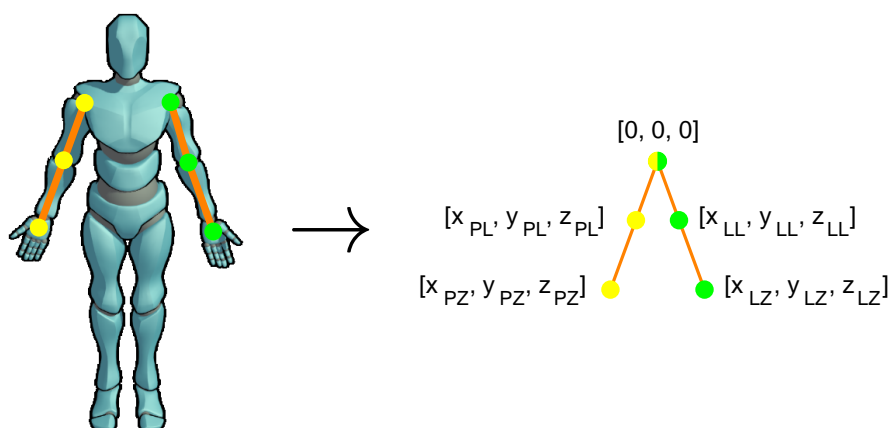
## 3.4 Klasifikace

Detekce ovládacích gest probíhá ve dvou hlavních částech. V první části jsou nejprve klasifikovány samotné pozice obou horních končetin a na základě těchto detekovaných pozic poté dochází ke klasifikaci ovládacích gest.

### 3.4.1 Klasifikace pozice horní končetiny

Klasifikace pozice horní končetiny je prováděna na základě upravených souřadnic sledovaných kloubů. Za účelem nalezení nejvhodnějšího řešení byly z těchto souřadnic získány dva odlišné druhy příznaků.

Prvním druhem byly kartézské souřadnice kloubů po úpravách, které budou popsány v podkapitole 4.2.1. Pro každou horní končetinu jsou dostupné 3 souřadnice pro 3 sledované klouby. Jako příznaky pro klasifikaci jsou použity relativní souřadnice zápěstí a lokte vůči odpovídajícímu ramenu. Každá pozice jedné sledované končetiny je tedy nakonec popsána 6 hodnotami.

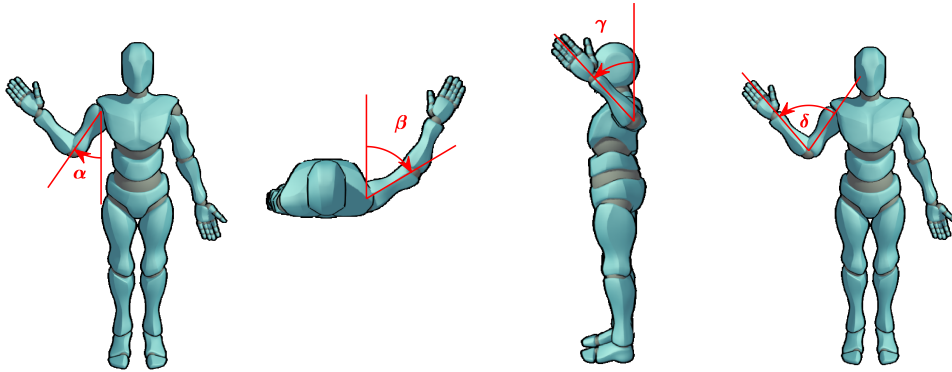


**Obrázek 3.11:** Relativní kartézské souřadnice.

Druhým typem příznaků použitým pro popis pozice horní končetiny byly 4 vhodně nadefinované úhly zobrazené na Obrázku 3.12. Úhly  $\alpha$ ,  $\beta$  a  $\gamma$  reprezentují natočení paže, úhel  $\delta$  poté odpovídá ohnutí končetiny v lokti, tedy úhlu mezi paží a předloktím. Při testování funkčnosti však bylo zjištěno, že kombinací definice úhlů a nepřesností získaných souřadnic při sledování kostry dochází v okolí některých pozic k výrazným změnám měřených hodnot. Takovou problematickou pozicí je například nastavení paže, ve kterém  $\alpha \rightarrow 0^\circ$ , resp.  $\alpha \rightarrow 180^\circ$ . V takovýchto případech způsobí i malá změna polohy horní končetiny velkou změnu měřené hodnoty úhlu  $\beta$  v plném rozsahu  $(-180^\circ, 180^\circ)$ . Stejný problém nastává pro úhel  $\gamma$  při zarovnání paže a předloktí, tj. když  $\delta \rightarrow 180^\circ$ . Z tohoto důvodu byla zavedena následující pravidla pro určení hodnot  $\beta$  a  $\gamma$ :

$$\beta = \begin{cases} \beta_{změřeno} & \text{jestliže } \alpha < 30^\circ \\ 0 & \text{jinak} \end{cases}, \quad (3.1)$$

$$\gamma = \begin{cases} \gamma_{změřeno} & \text{jestliže } \delta > 150^\circ \\ 0 & \text{jinak} \end{cases}. \quad (3.2)$$



Obrázek 3.12: Úhlové příznaky.

Při použití obou typů příznaků jsou data klasifikována do jedné z 11 možných tříd. Tyto třídy odpovídají 10 definovaným pozicím rozšířeným o třídu *žádná*, do které jsou data přiřazena, pokud nebyla klasifikována jako žádná z definovaných pozic.

### ■ Použité klasifikátory

Pro klasifikaci získaných dat do tříd reprezentujících definované pozice horních končetin byly použity dva různé klasifikátory:

#### ■ Naivní Bayesovský klasifikátor

Prvním z použitých klasifikátorů byl naivní Bayesovský klasifikátor pro příznaky s předpokládaným normálním rozdělením. Tento klasifikátor předpokládá podmíněnou nezávislost jednotlivých příznaků a normální rozdělení jejich výskytu [54]. Tento předpoklad často není v reálných úlohách splněn, přesto lze s naivním Bayesovským klasifikátorem dosáhnout dostatečné úrovně klasifikace [27].

Počet příznaků  $N_x$  je při použití kartézských souřadnic  $N_x = 6$  a při použití úhlových příznaků  $N_x = 4$ . Jednotlivé příznaky budeme označovat jako  $x_i$ , kde  $i = 1, \dots, N_x$ . Celkový počet klasifikovaných tříd, tj. detekovaných pozic horní končetiny, pojmenujeme jako  $N_c$  a jednotlivé třídy tak budeme označovat jako  $c_k$ , kde  $k = 1, \dots, N_c$ . Pro všechny detekované pozice horních končetin předpokládáme stejnou apriorní pravděpodobnost,

spočteme ji tedy podle vztahu

$$P(c_k) = \frac{1}{N_c}. \quad (3.3)$$

Při učení tohoto klasifikátoru dochází k odhadu střední hodnoty  $\mu_{x_i, c_k}$  a rozptylu  $\sigma_{x_i, c_k}^2$  pro každý příznak v každé třídě. Tyto hodnoty jsou získány z trénovacích dat pro odpovídající třídu využitím následujících vztahů:

$$\mu_{x_i, c_k} = \frac{1}{N_{c_k}} \sum_{j=1}^{j=N_{c_k}} x_{i_j}, \quad (3.4)$$

$$\sigma_{x_i, c_k}^2 = \frac{1}{N_{c_k}} \sum_{j=1}^{j=N_{c_k}} (x_{i_j} - \mu_{x_i, c_k})^2, \quad (3.5)$$

kde  $N_{c_k}$  je počet trénovacích sad příznaků pro  $k$ -tou třídu a  $x_{i_j}$  je příznak  $x_i$  pro  $j$ -tou sadu trénovacích příznaků.

Během klasifikace je poté pro každou třídu vypočítána pravděpodobnost výskytu každého změřeného příznaku podle vztahu

$$P(x_i|c_k) = \frac{1}{2\pi \cdot \sigma_{x_i, c_k}^2} \exp \left\{ -\frac{(x_i - \mu_{x_i, c_k})^2}{2\sigma_{x_i, c_k}^2} \right\}. \quad (3.6)$$

Z těchto pravděpodobností je následně určena aposteriorní pravděpodobnost změřené sady příznaků pro každou třídu

$$P(x|c_k) = \prod_{i=1}^{i=N} P(x_i|c_k). \quad (3.7)$$

Poté dojde k výpočtu podmíněných pravděpodobností výskytu každé třídy při naměřených příznacích Bayesovým vztahem

$$P(c_k|x) = \frac{P(x|c_k) \cdot P(c_k)}{\sum_{m=1}^{m=N_c} P(x|c_m) \cdot P(c_m)}. \quad (3.8)$$

Standardně by výstupem klasifikace byla třída s nejvyšší podmíněnou pravděpodobností [54]. V tom případě by však byla vždy zvolena některá z definovaných pozic a nemohlo by dojít k potlačení chybné klasifikace nedefinované pozice. Z tohoto důvodu byla pro výslednou klasifikaci zavedena podmínka 3.9, teprve při jejímž splnění je vzorek dat klasifikován do třídy s největší podmíněnou pravděpodobností.

$$P(c_a|x) > g \cdot P(c_b|x) \quad (3.9)$$

$c_a$  je třída s nejvyšší podmíněnou pravděpodobností,  $c_b$  třída s druhou nejvyšší podmíněnou pravděpodobností,  $P(c_a|x)$  nejvyšší podmíněná pravděpodobnost,  $P(c_b|x)$  druhá nejvyšší podmíněná pravděpodobnost a  $g$  je nový parametr klasifikátoru.

### ■ Vlastní klasifikátor

Druhý klasifikátor byl navržen a vytvořen v rámci této práce přímo pro klasifikaci pozic horních končetin. Tento klasifikátor využívá ke klasifikaci celkem 6 příznaků. Jedná se o relativní souřadnice lokte a zápěstí vůči ramenu v kartézském souřadnicovém systému. Jediným nastavitelným parametrem tohoto klasifikátoru je násobící konstanta  $\lambda$  pro získání maximální dovolené vzdálenosti od referenční pozice kloubu, při které ještě může být klasifikovaný vzorek přiřazen do dané třídy.

Při trénování tohoto klasifikátoru jsou pro každou definovanou polohu horní končetiny načtena odpovídající data. Následně jsou z těchto dat vypočítány průměrné hodnoty jednotlivých souřadnic, čímž jsou získány referenční relativní souřadnice lokte a zápěstí vůči ramenu pro danou třídu. Poté je vypočítána střední vzdálenost každého kloubu od jeho průměrné pozice. Přenásobením této vzdálenosti násobící konstantou  $\lambda$  získáme maximální dovolenou vzdálenost od referenční pozice. Tuto vzdálenost označíme  $d_{max}$ .

Při klasifikaci je poté pro každou definovanou polohu vypočtena euklidovská vzdálenost mezi klasifikovanými polohami lokte a zápěstí a referenčními souřadnicemi těchto kloubů. Pokud je vzdálenost obou kloubů od referenčních pozic nižší než definovaná maximální vzdálenost, je rozhodnutí klasifikovat daný vzorek dat do této třídy přiřazena hodnota  $h$  podle vztahu

$$h = 1 - \frac{d_l + d_z}{2 \cdot d_{max}} \quad (3.10)$$

kde  $d_l$  je vzdálenost lokte od referenční polohy lokte,  $d_z$  vzdálenost zápěstí od referenční polohy zápěstí a  $d_{max}$  je definovaná maximální dovolená vzdálenost. Pokud je alespoň jeden kloub ve větší než definované maximální vzdálenosti od odpovídající referenční pozice, je rozhodnutí klasifikovat vzorek dat do dané třídy přiřazena nulová hodnota. Jako výstupní klasifikace je poté zvolena třída s nejvyšší hodnotou.

### ■ 3.4.2 Klasifikace gesta

Klasifikace gest je prováděna na základě detekovaných pozic horních končetin. Pro statická gesta se jedná o pouhou soustavu podmínek, při jejichž splnění je detekováno odpovídající gesto. Pokud je tedy například na levé horní končetině detekována poloha *do strany* a zároveň je i na pravé horní končetině detekována poloha *do strany*, je vyhodnoceno gesto *zastavit* a následně je provedena odpovídající akce.

Pro klasifikaci dynamických gest je v paměti udržováno posledních 10 detekovaných poloh pro každou končetinu. Aktualizace těchto struktur probíhá

v každém cyklu a k přidání nově detekované pozice dojde pouze v případě, kdy je nová pozice odlišná od té předchozí a nejedná se o pozici klasifikovanou jako *žádná*, tj. nedefinovaná. Při naplnění prostoru určeného k uložení sekvencí pozic dojde před zápisem nejnovější pozice k odstranění té nejstarší. Při samotné detekci je každé definované dynamické gesto porovnáno s nejnovějšími částmi uložených sekvencí obou horních končetin o odpovídající délce. Při nalezení shody je provedena příslušná operace a dojde k vymazání uložených sekvencí pozic. Pokud není detekováno žádné statické ani dynamické gesto, výstupem klasifikace je gesto označené jako *žádné*.



## Kapitola 4

### Implementace

V této kapitole je popsána implementace systému pro ovládání robotu gesty. Je zde uveden a popsán použitý software a hardware, způsob úpravy použitých dat, vlastnosti výsledného ovládacího systému a vytvořené grafické uživatelské rozhraní.

#### ■ 4.1 Použitý software a hardware

##### ■ 4.1.1 Software

Implementace probíhala převážně v OS Ubuntu [55] ve verzi 18.04, funkčnost byla ověřena také na verzích 19.04 a 19.10.

##### ■ Programovací jazyk

Jako programovací jazyk pro vytvoření ovládacího systému byl zvolen Python [56] ve verzi 3.7, s drobnými úpravami je možný provoz i s verzí 2.7, avšak některé moduly (např. Kivy) již brzy ukončí podporu této starší verze Pythonu. Z tohoto důvodu je doporučeno použít Python alespoň ve verzi 3.5 a novější.

Pro správný chod programu je potřeba nainstalovat následující moduly:

- `openni` – rozhraní pro OpenNI a NiTE v Pythonu [57],
- `kivy` – framework pro vytváření multiplatformních aplikací [58],
- `numpy` – knihovna pro práci s vektory a maticemi [59],
- `scikit-learn` – knihovna pro analýzu dat a strojové učení [60],
- `playsound` – knihovna pro přehrávání zvuků [61],
- `simple_pid` – knihovna implementující PID regulátor [62],
- `vg` – knihovna pro analytickou geometrii [63],
- `matplotlib` – knihovna pro vykreslování grafů [64].

#### ■ `libfreenect2`

Pro zprovoznění senzoru Kinect v OS Ubuntu byla použita knihovna `libfreenect2`. Tato open-source knihovna je vyvíjena komunitou `OpenKinect` a obsahuje ovladače pro získání RGB, infračerveného a hloubkového obrazu a registraci RGB a hloubkového obrazu [65].

#### ■ `OpenNI2`

`OpenNI2` je aplikační rozhraní poskytující přístup k datům z různých senzorů. Umožňuje uživateli inicializovat senzory a vyčítat z nich dostupná data. Poskytuje také jednotné rozhraní pro software třetích stran [66].

#### ■ `NiTE2`

`NiTE2` je uzavřený software vyvinutý společností `PrimeSense`, který je mimo jiné schopen nalézt a následně sledovat postavu v hloubkovém obraze. Vytváří při tom zjednodušený model kostry člověka (Obrázek 4.4) a umožňuje tak získání souřadnic jednotlivých sledovaných částí těla.



## 4.1.2 Hardware

### Kinect

Jako zdroj dat pro následnou detekci osob je použit senzor Kinect v2 [67] od společnosti Microsoft. Tento senzor byl původně určen jako periferie k herní konzoli Xbox One, stejně jako jeho předchozí verze však našel uplatnění v celé řadě jiných oblastí, včetně 3D skenování různých objektů, SLAM (simultánní lokalizace a mapování) a nebo právě rozpoznání gest [68], [69], [70].



**Obrázek 4.1:** Senzor Kinect v2. Převzato z [71].

Tento senzor obsahuje:

- RGB kameru s rozlišením 1920 x 1080 pixelů,
- infračervený projektor,
- infračervenou kameru s rozlišením 512 x 424 pixelů,
- mikrofónové pole se 4 mikrofóny.

Pro sledování pohybu osob a následnou detekci gest jsou využívána pouze hloubková data získána prostřednictvím infračerveného projektoru a infračervené kamery. Infračervená kamera má horizontální zorný úhel 70° a vertikální zorný úhel 60° a v kombinaci s infračerveným projektorem a příslušným softwarem je možné detekovat osobu až do vzdálenosti 4,5 m [72].

Princip pro získání hloubkových dat použitý u tohoto zařízení se nazývá Time-of-Flight. Při tomto způsobu získávání hloubkových dat dochází k vysílání infračerveného záření modulovaného obdélníkovým signálem a následnému zachycení odrazu tohoto záření infračervenou kamerou. Změřením fázového posuvu  $\Delta\varphi$  mezi vyslaným a přijatým signálem lze při známém

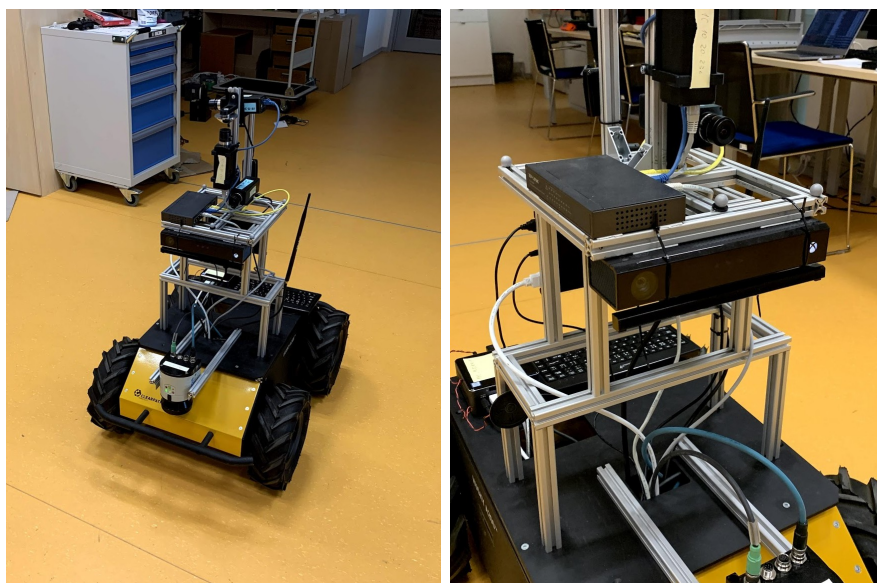
modulačním kmitočtu  $f_{mod}$  a rychlosti šíření světla  $c$  určit prostřednictvím vztahu 4.1 vzdálenost objektu  $d$ , od kterého se světlo odrazilo [73].

$$d = \frac{\Delta\varphi \cdot c}{4\pi \cdot f_{mod}} \quad (4.1)$$

Tímto způsobem je pro každý pixel infračerveného senzoru určena vzdálenost oblasti scény, od které se infračervené záření odrazilo. Tím dojde k získání hloubkového obrazu.

## ■ Robot

Pro účely testování byl senzor Kinect spolu s ovládacím systémem nainstalován na čtyřkolový robot Husky od společnosti Clearpath Robotics Inc. Základní verze tohoto robotu má hmotnost 50 kg a dokáže se pohybovat rychlostí až 1 m/s, a to po dobu až 3 hodin [74]. Kromě standardního pohybu vpřed a vzad se dokáže také otáčet na místě.



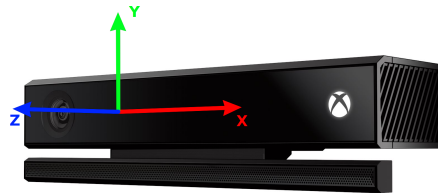
**Obrázek 4.2:** Čtyřkolový robot Husky se senzorem Kinect v2.

Na robotu byl nainstalován PC Intel NUC7i5BNH s procesorem Intel® Core™ i5-7260U a 8 GB operační paměti. Na tomto PC je spuštěn řídicí systém obstarávající pohyb robotu. Ovládání robotu je realizováno prostřednictvím dvou hodnot, které odpovídají dopředné rychlosti v metrech za sekundu a rotační rychlosti v radiánech za sekundu. Pro příjem těchto hodnot zpřístupňuje řídicí systém socket, na kterém očekává zprávu ve formátu

$move$  [dopředná rychlost v  $m/s$ ] [rotační rychlost v  $rad/s$ ], taková zpráva tedy může vypadat např. takto:  $move\ 0.5\ 0.23$ . Na základě přijatých hodnot jsou poté ovládány jednotlivé motory robotu. Tento řídicí systém byl poskytnut vedoucím práce a nebyl tedy součástí tvorby této práce.

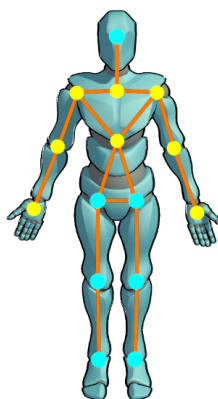
## 4.2 Detekovaná kostra

Data získaná z hloubkového obrazu prostřednictvím softwaru NiTE2 mají podobu 3D souřadnic jednotlivých kloubů detekované kostry (Obrázek 4.4) v souřadnicovém systému s počátkem v senzoru Kinect (Obrázek 4.3).



Obrázek 4.3: Souřadnicový systém. Převzato a upraveno z [71].

Z této množiny je pro další zpracování sledováno celkem 8 kloubů, a to konkrétně zápěstí, lokty, ramena, krk a střed páteře (na Obrázku 4.4 označeny žlutou barvou). Před samotnou klasifikací jsou souřadnice jednotlivých kloubů upraveny, jak je popsáno v části 4.2.1.



Obrázek 4.4: Model kostry s vyznačenými sledovanými klouby.

### ■ 4.2.1 Úprava detekované kostry

Před provedením transformací detekované kostry jsou nejprve uloženy souřadnice kloubu, který odpovídá středu páteře. Tyto souřadnice jsou později použity jako referenční bod pro určení vzájemné pozice robotu a operátora.

#### ■ Změna měřítka v ose Z

První operací je změna měřítka v ose Z. Během provedených experimentů bylo totiž zjištěno, že změna polohy v ose Z vyvolává přibližně dvojnásobnou změnu měřených souřadnic oproti pohybu v osách X a Y. Naměřené hodnoty v ose Z jsou tedy upraveny experimentálně zjištěnou konstantou tak, aby bylo výsledné měřítko všech 3 os shodné.

#### ■ Rotace před senzor

Druhou operací je rotace kostry kolem osy Y tak, aby se střed páteře nacházel uprostřed horizontálního zorného pole senzoru, tzn. aby platilo  $x_{\text{střed páteře}} = 0$ .

#### ■ Rotace do svislé polohy

Aby bylo zamezeno vlivu vzájemného natočení senzoru a ovládací osoby, například při najetí robotu na nerovnost, dochází k rotaci detekované kostry tak, aby byla spojnice kloubů odpovídajících krku a středu páteře rovnoběžná s osou Y. Nejprve je tedy provedena rotace kolem osy X pro zarovnání páteře v rovině Y-Z a poté následuje rotace kolem osy Z pro zarovnání v rovině X-Y. Uvažováno bylo také otočení kostry kolem osy Y tak, aby byla spojnice ramen rovnoběžná s osou X. Tato transformace však nakonec nebyla zahrnuta, a to z důvodu nepřesnosti získaných souřadnic, které zapříčiňovaly chybnou transformaci a zhoršenou úroveň klasifikace pozic.

## ■ Změna velikosti

Předposlední operací je normalizace velikosti kostry, díky čemuž by měla být zajištěna možnost ovládání pro osoby různého vzrůstu, za předpokladu podobných proporcí těla. Jako základní rozměr, odpovídající po transformaci jednotkové vzdálenosti, je použita vzdálenost mezi klouby kostry odpovídajícími krku a středu páteře. Touto hodnotou jsou poděleny souřadnice všech kloubů, čímž dojde k normalizaci vzdáleností jednotlivých kloubů.

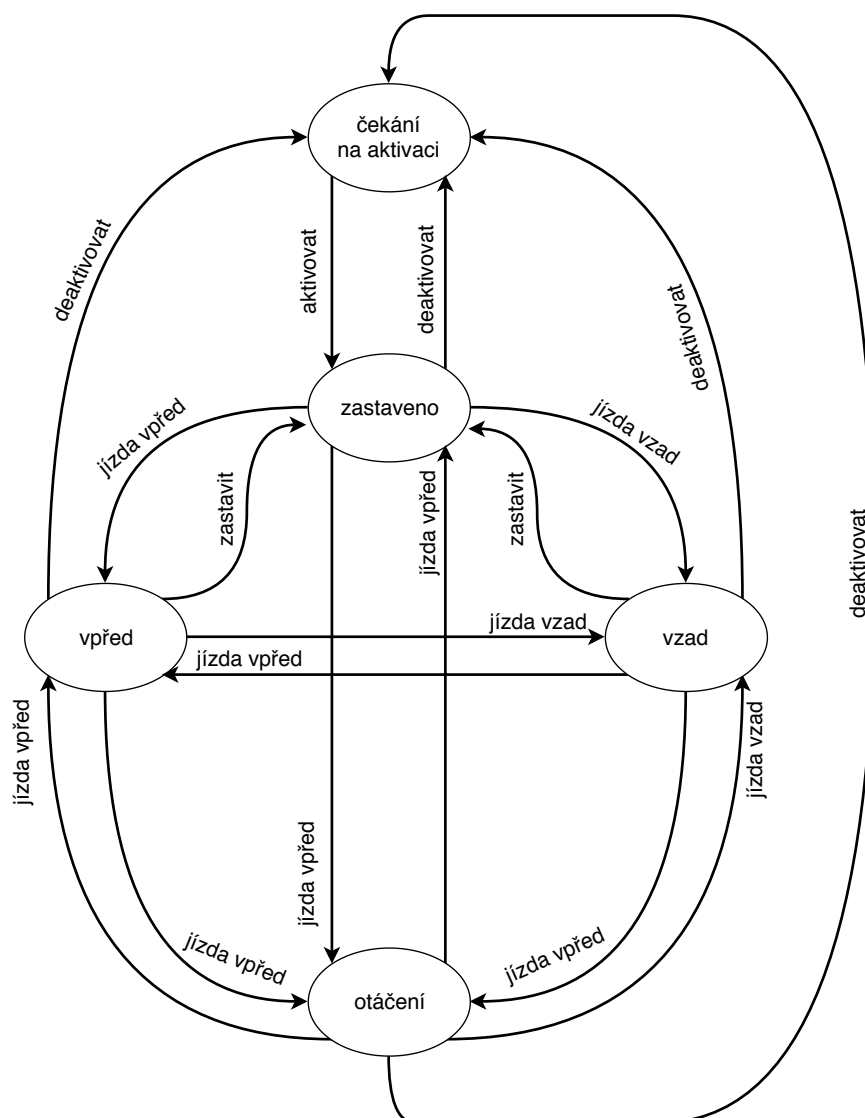
## ■ 4.3 Ovládání robotu

Samotné ovládání robotu probíhá prováděním jednotlivých řídicích gest. Na základě těchto gest se poté mění současný stav robotu a podle něj se následně určuje výpočet řídicích veličin. Na Obrázku 4.5 je zobrazen stavový automat ovládacího systému s přechody mezi jednotlivými stavy. V případě ztráty vizuálního kontaktu s operátorem nebo při jeho opuštění vymezené oblasti dojde ke stejné změně stavu jako při provedení gesta *deaktivovat*, tedy k přechodu systému do stavu *čekání na aktivaci*.

Pro řízení pohybu robotu jsou použity dvě hodnoty – dopředná a rotační rychlost. Pro získání těchto řídicích hodnot jsou použity dva obecné PID regulátory z importovaného modulu *simple\_pid*, každý je nastaven pro výpočet jedné z rychlostí. Vzhledem k výsledným rychlostem robotu se ukázalo jako dostatečné použít pouze proporcionální část regulátorů. V případě potřeby je však možné úpravou konfiguračního souboru plně využít použitou implementaci PID regulátorů vhodným nastavením všech jeho složek.

Vstupem regulátorů je vzdálenost a úhel mezi robotem a operátorem, konkrétně mezi středem senzoru a kloubem odhadované kostry, který odpovídá středu páteře. V případě, že se robot nachází ve stavu *vpřed* nebo *vzad*, dochází k výpočtu obou rychlostí. Ve stavu *otáčení* se určuje pouze potřebná rotační rychlost a ve zbývajících stavech jsou požadované rychlosti nastavené na nulovou hodnotu.

Řídicí hodnoty jsou poté v každém cyklu odesílány přes soket do řídicího systému robotu.

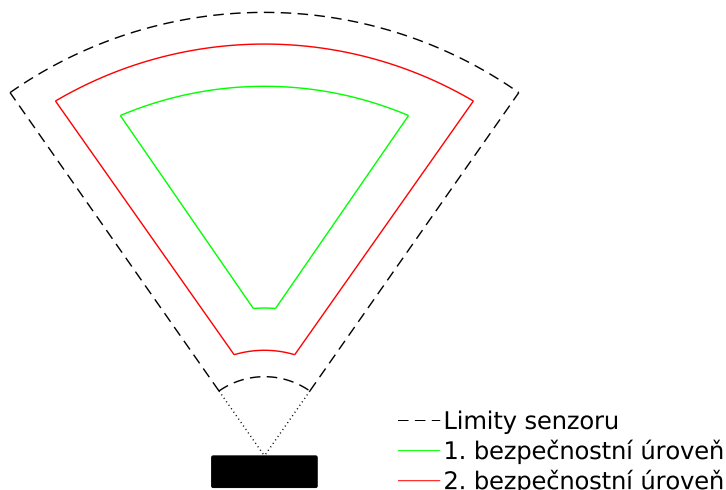


Obrázek 4.5: Stavový automat ovládní.

#### 4.4 Bezpečnostní prvky

Pro zajištění vyšší spolehlivosti ovládní bylo zavedeno několik opatření. Prvním z nich je vymezení prostoru, ve kterém se může osoba ovládající robot nacházet. Tento prostor je odvozen od parametrů použitého senzoru, konkrétně od horizontálního zorného úhlu a maximální vzdálenosti pro spolehlivý odhad pozice osoby. Výsledný bezpečný prostor byl zvolen tak, aby bylo možné v jeho hranicích spolehlivě zastavit pohyb robotu, tzn. aby bylo možné rozpažit a předvést tak gesto pro zastavení.

Vymezení prostoru pro ovládání je provedeno ve dvou úrovních. Při překročení první úrovně (na Obrázku 4.6 zelenou barvou) dojde ke spuštění akustického varování informující ovládací osobu o přiblížení se ke hranici ovládacího prostoru a hrozbě ukončení ovládání. Při překročení druhé úrovně (na Obrázku 4.6 červenou barvou) akustické varování zintenzivní a pokud se operátor během definovaného časového limitu (ve výchozím nastavení 3 sekundy) nevrátí do vymezené oblasti, dojde k zastavení pohybu robotu a ukončení ovládání.



**Obrázek 4.6:** Vymezený úsek pro ovládání spolu se zorným polem senzoru.

Dalším opatřením je prosté zastavení a ukončení ovládání při náhlé ztrátě vizuálního kontaktu s operátorem způsobené například zakrytím senzoru či pohybem ovládací osoby za překážku. Z bezpečnostních důvodů je také omezena maximální povolená rychlost pohybu robotu, a to jak při jízdě, tak při rotaci.

## 4.5 Konfigurovatelné prvky

Nastavení systému pro ovládání lze do jisté míry měnit prostřednictvím konfiguračního souboru *config.ini* umístěného v hlavním adresáři. Úpravou tohoto souboru lze nastavit:

- maximální dopřednou a rotační rychlost,
- minimální potřebné úrovně řídicích hodnot pro spuštění pohybu,
- konstanty PID regulátorů,
- vzdálenost, kterou má robot udržovat,
- bezpečnostní limity,
- časový limit při překročení bezpečnostních limitů,
- druh zvukového výstupu (slova/zvuky),
- ignoraci pozice *dopředu nahoru* při klasifikaci gest,
- kombinaci použitých příznaků a klasifikátoru.

Podrobný popis jednotlivých parametrů je uveden v Příloze C.

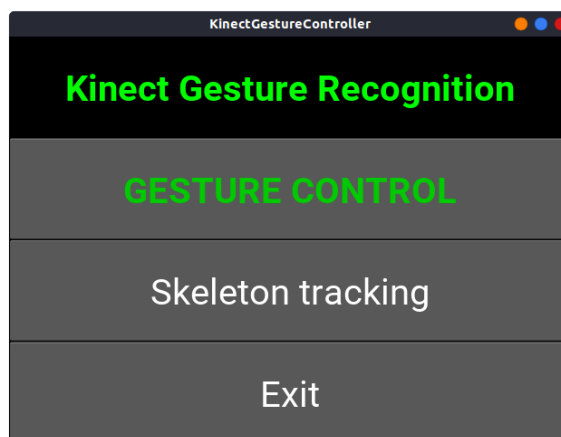


## 4.6 Grafické rozhraní

Pro snazší ovládání a kontrolu stavu systému bylo vytvořeno jednoduché grafické uživatelské rozhraní (GUI). To bylo zpracováno v základní a pokročilé verzi.

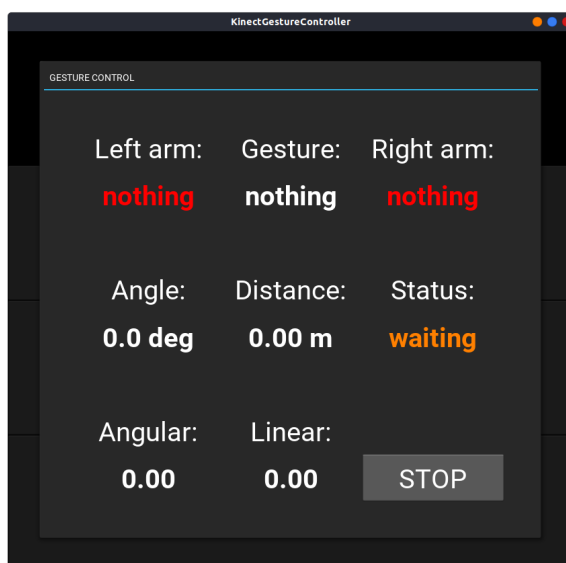
### 4.6.1 Základní verze GUI

V základní verzi GUI je možné spustit samotné ovládání a nebo zobrazit hloubkový obraz s vykreslením detekovaných koster.



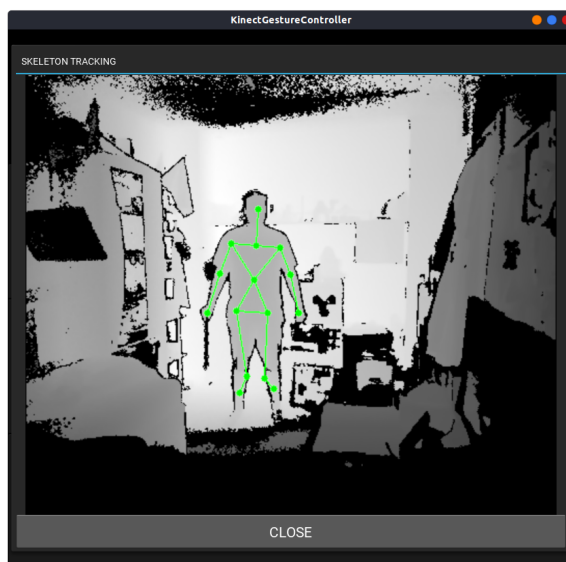
Obrázek 4.7: Hlavní menu základního uživatelského rozhraní.

Při spuštění ovládání se otevře nové okno (Obrázek 4.8), ve kterém jsou uvedeny aktuálně klasifikované pozice obou horních končetin, aktuálně detekované gesto, vzdálenost a úhel mezi robotem a operátorem, stav robotu a akční zásahy obou regulátorů. Navíc je změnou barvy textu aktuálně detekovaných poloh horních končetin indikována úroveň kvality sledování horních končetin. Červená barva označuje naprostou ztrátu sledování dané končetiny, oranžová značí velmi nízkou úroveň sledování, žlutá již znamená úroveň, na které je možné začít používat ovládání a zelená znamená maximální úroveň sledování, které je software schopný.



Obrázek 4.8: Okno s informacemi o ovládaní.

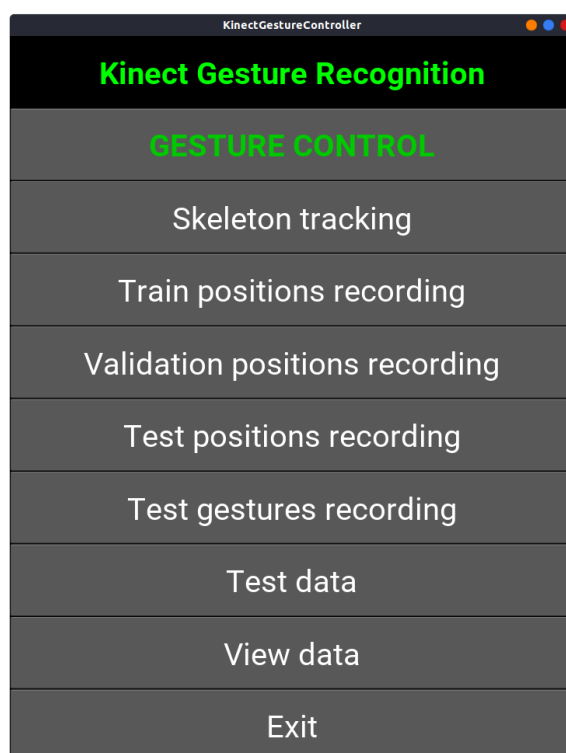
Při otevření okna se zobrazeným hloubkovým obrazem je při detekci vykreslena detekovaná kostra, jejíž barvy mají stejný význam jako barva textu uvedená v předchozím odstavci. Toto zobrazení je vhodné pro kontrolu funkčnosti detekce a zajištění vhodného náklonu senzoru pro co nejlepší pokrytí prostoru s předpokládaným pohybem sledované osoby.



Obrázek 4.9: Okno s hloubkovým obrazem.

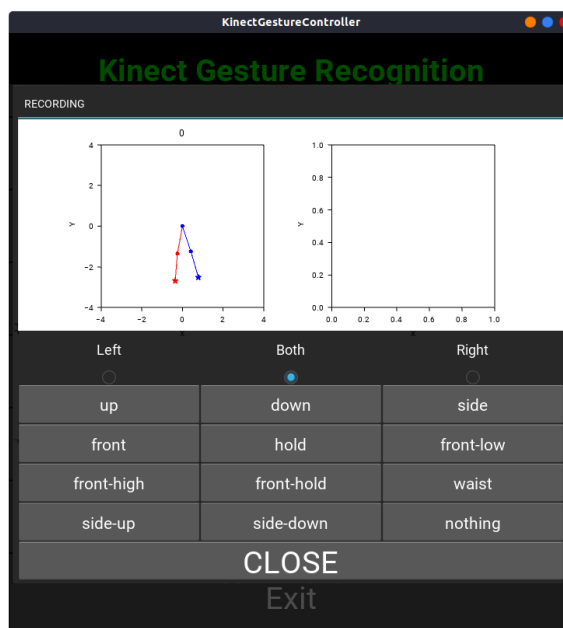
## 4.6.2 Rozšířená verze GUI

V rozšířené verzi GUI lze mimo možností dostupné v základní verzi navíc zaznamenávat data pro trénování, validaci a testování klasifikátoru, zobrazovat a případně mazat chybně zaznamenaná data nebo provést test klasifikace a zobrazit dosažené výsledky.

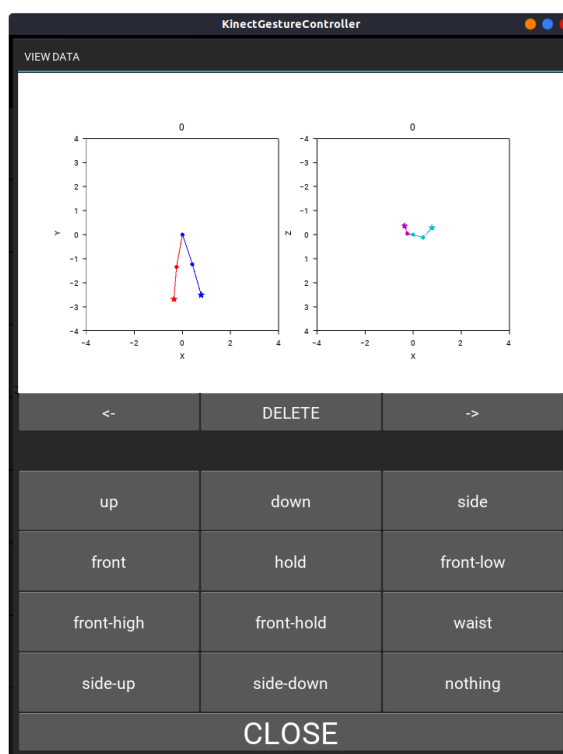


Obrázek 4.10: Základní menu rozšířeného uživatelského rozhraní.

Při nahrávání nových dat jsou v horní části okna zjednodušeně zobrazeny obě horní končetiny, a to v rovinách X-Y a X-Z. Ve spodní části okna se nacházejí tlačítka s názvy jednotlivých detekovaných pozic. Po stisku jednoho z tlačítek dojde k výpočtu příznaků ze souřadnic sledovaných kloubů a jejich uložení do odpovídajícího souboru. Vzhled okna je uveden na Obrázku 4.11.



Obrázek 4.11: Rozhraní pro nahrávání trénovacích dat.



Obrázek 4.12: Rozhraní pro zobrazení a mazání dat.

V okně pro zobrazení a mazání nahraných dat (Obrázek 4.12) jsou stejně jako v předchozím okně zobrazeny pozice obou horních končetin ve dvou rovinách a tlačítka s názvy detekovaných pozic. Po stisku některého z tlačítek se zobrazí odpovídající data. Prostřednictvím tlačítek se symboly šipek lze poté listovat mezi jednotlivými nahranými pozicemi a při nalezení chybně zpracovaných dat je lze stiskem tlačítka s nápisem *DELETE* smazat.



# Kapitola 5

## Vyhodnocení

V této kapitole je popsán způsob vyhodnocení klasifikace pozic horní končetiny a samotných gest, dále jsou zde popsány problémy, které se v rámci této práce nepodařilo dostatečně potlačit. Nakonec jsou uvedeny možná vylepšení systému a možnosti jeho využití.

### 5.1 Vytvořený dataset

Pro správnou funkci a následné vyhodnocení úrovně klasifikace bylo pro každou definovanou pozici zaznamenáno 600 nastavení obou horních končetin, to činí celkem 12000 datových vzorků pro všech 10 definovaných pozic, pokud uvažíme každou horní končetinu zvlášť. Tato data byla následně rozdělena v poměru 50:25:25 do 3 disjunktních množin. Tyto množiny byly použity jako:

- trénovací data – slouží k učení klasifikátoru,
- validační data – slouží k porovnání výsledků při různých nastaveních parametrů klasifikátoru,
- testovací data – slouží ke konečnému odhadu kvality klasifikátoru.

Pro validaci i testování klasifikátorů bylo navíc nahráno 150 nastavení horních končetin, které neodpovídaly žádné definované pozici, aby bylo možné vhodnou

úpravou parametrů klasifikátoru nastavit potlačení těchto případů. Pro ověření funkčnosti detekce gest bylo pro každé definované gesto zaznamenáno 100 sekvencí o délce 5 sekund, ve kterých bylo dané gesto provedeno. Pro ověření schopnosti klasifikátoru potlačit falešně pozitivní detekce nedefinovaných gest bylo zaznamenáno celkem 100 sekvencí, ve kterých byly provedeny pohyby horními končetinami, které neodpovídaly žádnému definovanému gestu.

## 5.2 Způsob vyhodnocení

Vyhodnocení dosažených výsledků bylo provedeno pro tři různé kombinace příznaků a klasifikátorů:

- naivní Bayesovský klasifikátor s relativními kartézskými souřadnicemi kloubů horních končetin,
- naivní Bayesovský klasifikátor s úhlovým popisem polohy horních končetin,
- vlastní klasifikátor s relativními kartézskými souřadnicemi kloubů horních končetin.

K reprezentaci získaných výsledků byly použity 3 různé metriky:

- **senzitivita** – poměr správných detekcí ku celkovému počtu výskytů daného jevu. Popisuje jak velká je pravděpodobnost detekce daného jevu při jeho výskytu [75]. Vypočítá se podle vztahu

$$\text{Senzitivita} = \frac{TP}{TP + FN}, \quad (5.1)$$

- **přesnost** – poměr správných detekcí ku celkovému počtu detekcí daného jevu. Popisuje jak věrohodná je výsledná klasifikace [75]. Vypočítá se podle vztahu

$$\text{Přesnost} = \frac{TP}{TP + FP}, \quad (5.2)$$

- **F-míra** – harmonický průměr senzitivity a přesnosti [75]. Vypočítá se podle vztahu

$$F\text{-míra} = \frac{2 \cdot \text{Přesnost} \cdot \text{Senzitivita}}{\text{Přesnost} + \text{Senzitivita}}. \quad (5.3)$$



Před samotným vyhodnocením bylo nejprve provedeno natrénování daného klasifikátoru odpovídajícími trénovacími daty. Následně byly experimentálně nastaveny dostupné parametry klasifikátoru tak, aby výsledky dosažené na validační množině dat co nejlépe odpovídaly požadavkům. To konkrétně znamenalo dosažení co největší senzitivity a zároveň co největší přesnosti, tedy dosažení co nejvyšší hodnoty F-míry.

Po nalezení nejvhodnějších hodnot parametrů byla vyhodnocena úspěšnost klasifikace jednotlivých pozic a gest na odpovídajících testovacích datech. Dosažené výsledky jsou uvedeny ve formě matic záměn (confusion matrices). Ve sloupcích těchto matic je uvedena četnost klasifikace vzorků z jednotlivých tříd do každé výstupní třídy. V řádcích je poté možné pozorovat, ze kterých tříd jsou vzorky, které byly do dané třídy klasifikovány. Matice tedy zobrazuje četnosti výskytu jednotlivých kombinací skutečné a klasifikované třídy [76]. Dále jsou zde uvedeny matice obsahující senzitivitu, přesnost a F-míru klasifikace jednotlivých tříd a průměrné hodnoty přes všechny třídy.

## ■ 5.3 Dosažené výsledky

V této podkapitole jsou popsány výsledky klasifikací pozic a gest dosažené prostřednictvím všech tří kombinací příznaků a klasifikátorů. Výsledky každé kombinace jsou reprezentovány maticí záměn a tabulkou s vyhodnocením klasifikace podle zvolených metrik.

### ■ 5.3.1 Výsledky klasifikace pozic horních končetin

V Tabulkách 5.2–5.7 jsou uvedeny výsledky klasifikace pozic horních končetin pro všechny tři kombinace příznaků a klasifikátorů. Tabulka 5.1 obsahuje identifikátory definovaných pozic horních končetin použité v Tabulkách 5.2–5.7.

ID	Pozice
0	dolů
1	dopředu nahoru
2	dopředu kolmo
3	dopředu dolů
4	dopředu
5	do strany kolmo
6	do strany dolů
7	do strany nahoru
8	do strany
9	nahoru

**Tabulka 5.1:** Identifikátory pozic horní končetiny.

		Skutečné pozice										
		0	1	2	3	4	5	6	7	8	9	jiná
Klasifikované pozice	0	296	0	0	0	4	0	3	0	0	0	8
	1	0	240	0	0	1	0	0	0	0	0	23
	2	0	0	299	0	0	0	0	0	0	0	21
	3	0	0	0	295	0	0	0	0	0	0	50
	4	0	2	0	0	286	0	0	0	0	0	23
	5	0	0	0	0	0	294	0	0	0	1	23
	6	0	0	0	0	0	0	266	0	0	0	53
	7	0	0	0	0	0	0	0	286	0	0	0
	8	0	0	0	0	0	0	0	2	298	0	5
	9	0	0	0	0	0	0	0	1	0	297	0
žádná	4	58	1	5	9	6	31	11	2	2	94	

**Tabulka 5.2:** Matice záměn klasifikovaných pozic horní končetiny při použití kartézských příznaků a naivního Bayesovského klasifikátoru.

	Pozice											Průměrná hodnota
	0	1	2	3	4	5	6	7	8	9	jiná	
Senzitivita	0,98	0,80	0,99	0,98	0,95	0,98	0,88	0,95	0,99	0,99	0,31	0,89
Přesnost	0,95	0,90	0,93	0,85	0,91	0,92	0,83	1,00	0,97	0,99	0,42	0,87
F-míra	0,96	0,84	0,95	0,91	0,92	0,94	0,85	0,97	0,97	0,99	0,35	0,87

**Tabulka 5.3:** Vyhodnocení klasifikace pozic horní končetiny při použití kartézských příznaků a naivního Bayesovského klasifikátoru.

		Skutečné pozice										
		0	1	2	3	4	5	6	7	8	9	jiná
Klasifikované pozice	0	<b>270</b>	0	1	0	2	0	7	0	0	1	87
	1	0	<b>277</b>	0	0	3	0	0	0	1	2	34
	2	0	0	<b>292</b>	0	0	2	0	0	0	0	7
	3	8	0	0	<b>284</b>	2	0	6	0	0	0	38
	4	0	18	0	14	<b>293</b>	0	0	0	0	0	37
	5	0	0	7	0	0	<b>298</b>	0	0	0	0	85
	6	22	0	0	0	0	0	<b>285</b>	0	0	0	9
	7	0	0	0	0	0	0	0	<b>282</b>	8	1	3
	8	0	0	0	0	0	0	1	3	<b>291</b>	0	0
	9	0	5	0	0	0	0	0	15	0	<b>295</b>	0
žádná	0	0	0	2	0	0	1	0	0	1	<b>0</b>	

**Tabulka 5.4:** Matice záměn klasifikovaných pozic horní končetiny při použití úhlových příznaků a naivního Bayesovského klasifikátoru.

	Pozice											Průměrná hodnota
	0	1	2	3	4	5	6	7	8	9	jiná	
Senzitivita	0,90	0,92	0,97	0,94	0,97	0,99	0,95	0,94	0,97	0,98	0,00	0,86
Přesnost	0,73	0,87	0,97	0,84	0,80	0,76	0,90	0,95	0,98	0,93	NaN	0,79
F-míra	0,80	0,89	0,97	0,88	0,87	0,85	0,92	0,94	0,97	0,95	NaN	0,82

**Tabulka 5.5:** Vyhodnocení klasifikace pozic horní končetiny při použití úhlových příznaků a naivního Bayesovského klasifikátoru.

		Skutečné pozice										
		0	1	2	3	4	5	6	7	8	9	jiná
Klasifikované pozice	0	<b>284</b>	0	0	0	1	0	5	0	0	0	0
	1	0	<b>294</b>	0	0	2	0	0	0	0	0	10
	2	0	0	<b>281</b>	0	0	0	0	0	0	0	0
	3	1	0	0	<b>296</b>	0	0	0	0	0	0	7
	4	0	4	0	1	<b>288</b>	0	0	0	0	0	1
	5	0	0	0	0	0	<b>279</b>	0	0	0	0	0
	6	3	0	0	0	0	0	<b>288</b>	0	0	0	0
	7	0	0	0	0	0	2	0	<b>291</b>	0	0	0
	8	0	0	0	0	0	0	1	2	<b>297</b>	0	0
	9	0	0	0	0	0	0	0	3	0	<b>289</b>	0
žádná	12	2	19	3	9	19	6	4	3	11	<b>282</b>	

**Tabulka 5.6:** Matice záměn klasifikovaných pozic horní končetiny při použití kartézských příznaků a vlastního klasifikátoru.

	Pozice											Průměrná hodnota
	0	1	2	3	4	5	6	7	8	9	jiná	
<b>Senzitivita</b>	0,94	0,98	0,93	0,98	0,96	0,93	0,96	0,97	0,99	0,96	0,94	0,95
<b>Přesnost</b>	0,97	0,96	1,00	0,97	0,97	1,00	0,98	0,99	0,99	0,98	0,76	0,96
<b>F-míra</b>	0,95	0,96	0,96	0,97	0,96	0,96	0,96	0,97	0,99	0,96	0,84	0,95

**Tabulka 5.7:** Vyhodnocení klasifikace pozic horní končetiny při použití kartézských příznaků a vlastního klasifikátoru.

Z výsledků uvedených v Tabulkách 5.2–5.7 tedy vyplývá, že nejlepší klasifikace pozic bylo dosaženo při použití relativních kartézských souřadnic a vlastního klasifikátoru. Senzitivita klasifikace jednotlivých pozic byla při použití této kombinace vyrovnaná, nejvyšší senzitivity bylo dosaženo pro pozici *do strany*, kdy bylo správně klasifikováno 297 ze 300 testovaných vzorků. Naopak nejnižší senzitivity dosáhla klasifikace pozice *do strany kolmo* s hodnotou 0,93. Silnou stránkou této kombinace je také vysoká přesnost klasifikace, která pro žádnou definovanou pozici neklesla pod úroveň 0,96. Díky vysoké senzitivě i přesnosti bylo dosaženo také nejvyšší průměrné hodnoty F-míry. Velkou výhodou oproti zbývajícím kombinacím je také vysoká úspěšnost při potlačení vzorků s nedefinovanými pozicemi.

Jako druhá nejlepší možnost se ukázala být kombinace relativních kartézských souřadnic a naivního Bayesovského klasifikátoru. Při jejím použití bylo dosaženo druhé nejvyšší celkové senzitivity a mimo pozici *dopředu nahoru* můžeme opět pozorovat její vyrovnanost přes jednotlivé detekované pozice. Oproti nejlepší kombinaci klasifikátoru a příznaků je zde však výrazně nižší schopnost potlačit nedefinované pozice.

Nejhorší výsledky byly získány při použití úhlových příznaků a naivního Bayesovského klasifikátoru. Tato kombinace sice dosahuje obdobné senzitivity jako ta předchozí, oproti ní má však výrazně nižší přesnost klasifikace, obzvláště pro pozice *dolů* a *do strany kolmo*. Dalším problémem u této kombinace je potlačení nedefinovaných pozic, kdy ve všech případech došlo k chybnému přiřazení k některé z definovaných pozic.

### ■ 5.3.2 Výsledky klasifikace gest

Výsledky klasifikace definovaných gest jsou uvedeny v Tabulkách 5.9–5.20. Tabulka 5.8 obsahuje identifikátory definovaných gest horních končetin použité v Tabulkách 5.9–5.20.

ID	Gesto
0	aktivace
1	deaktivace
2	jízda vpřed
3	jízda vpřed-na straně
4	jízda vzad
5	jízda vzad-na straně
6	zastavit-na straně
7	zastavit-vpředu
8	otáčení
9	pozastavit ovládání
10	změna sledovací vzdálenosti

Tabulka 5.8: Identifikátory gest.

	Skutečná gesta												
	0	1	2	3	4	5	6	7	8	9	10	jiné	
Klasifikovaná gesta	0	97	0	0	0	0	0	0	0	0	0	0	1
	1	0	97	0	0	0	0	0	0	0	0	0	0
	2	0	0	72	0	0	0	0	0	3	3	0	4
	3	0	0	0	100	0	0	0	0	0	0	18	2
	4	0	0	0	0	77	0	0	0	0	4	0	0
	5	0	0	0	0	0	100	0	0	0	0	1	0
	6	0	1	0	0	1	0	100	0	0	1	0	3
	7	0	0	0	0	0	0	0	85	0	0	0	2
	8	0	1	0	0	0	0	0	0	97	0	0	0
	9	0	0	0	0	2	0	0	0	0	90	0	0
	10	0	0	0	0	0	0	0	0	0	0	81	0
žádné	3	1	28	0	20	0	0	15	0	2	0	88	

Tabulka 5.9: Matice záměn klasifikovaných gest při použití kartézských příznaků a naivního Bayesovského klasifikátoru.

	Gesto												Průměrná hodnota
	0	1	2	3	4	5	6	7	8	9	10	jiné	
Senzitivita	0,97	0,97	0,72	1,00	0,77	1,00	1,00	0,85	0,97	0,90	0,81	0,88	0,90
Přesnost	0,98	1,00	0,87	0,83	0,95	0,99	0,94	0,97	0,98	0,97	1,00	0,56	0,92
F-míra	0,97	0,98	0,78	0,90	0,85	0,99	0,96	0,90	0,97	0,93	0,89	0,68	0,90

Tabulka 5.10: Vyhodnocení klasifikace gest při použití kartézských příznaků a naivního Bayesovského klasifikátoru.

		Skutečná gesta											
		0	1	2	3	4	5	6	7	8	9	10	jiné
Klasifikovaná gesta	0	<b>92</b>	0	0	0	0	0	0	0	0	0	0	3
	1	0	<b>94</b>	0	0	0	0	0	0	0	0	0	0
	2	0	0	<b>57</b>	0	0	0	0	0	1	1	0	0
	3	0	0	0	<b>100</b>	0	1	0	0	0	0	3	2
	4	0	0	0	0	<b>71</b>	1	0	0	0	2	0	0
	5	0	0	0	0	0	<b>97</b>	0	0	0	0	17	0
	6	0	0	0	0	1	0	<b>90</b>	0	0	1	0	3
	7	0	0	3	0	0	0	0	<b>84</b>	0	0	0	4
	8	0	2	0	0	0	0	10	0	<b>99</b>	38	0	1
	9	0	0	0	0	0	0	0	0	0	<b>56</b>	0	0
	10	0	0	0	0	0	0	0	0	0	0	<b>79</b>	0
žádné	8	4	40	0	28	1	0	16	0	2	1	<b>87</b>	

**Tabulka 5.11:** Matice záměn klasifikovaných gest při použití úhlových příznaků a naivního Bayesovského klasifikátoru.

	Gesto												Průměrná hodnota
	0	1	2	3	4	5	6	7	8	9	10	jiné	
Senzitivita	0,92	0,94	0,56	1,00	0,71	0,97	0,90	0,84	0,99	0,56	0,79	0,87	0,83
Přesnost	0,96	1,00	0,96	0,94	0,95	0,85	0,94	0,92	0,66	1,00	1,00	0,46	0,88
F-míra	0,93	0,96	0,70	0,96	0,81	0,90	0,91	0,87	0,79	0,71	0,88	0,6	0,83

**Tabulka 5.12:** Vyhodnocení klasifikace gest při použití úhlových příznaků a naivního Bayesovského klasifikátoru.

		Skutečná gesta											
		0	1	2	3	4	5	6	7	8	9	10	jiné
Klasifikovaná gesta	0	<b>98</b>	0	0	0	0	0	0	0	0	0	0	1
	1	0	<b>97</b>	0	0	0	0	0	0	0	0	0	0
	2	0	0	<b>34</b>	0	0	0	0	0	0	0	0	0
	3	0	0	0	<b>89</b>	0	0	0	0	0	0	0	2
	4	0	0	0	0	<b>63</b>	0	0	0	0	0	0	0
	5	0	0	0	0	0	<b>100</b>	0	0	0	0	7	0
	6	0	0	0	0	0	0	<b>100</b>	0	0	1	0	3
	7	0	0	0	0	0	0	0	<b>78</b>	0	0	0	2
	8	0	0	0	0	0	0	0	0	<b>100</b>	0	0	0
	9	0	0	0	0	0	0	0	0	0	<b>97</b>	0	0
	10	0	0	0	0	0	0	0	0	0	0	<b>93</b>	0
žádné	2	3	66	11	37	0	0	22	0	2	0	<b>92</b>	

**Tabulka 5.13:** Matice záměn klasifikovaných gest při použití kartézských příznaků a vlastního klasifikátoru.

	Gesto												Průměrná hodnota
	0	1	2	3	4	5	6	7	8	9	10	jiné	
<b>Senzitivita</b>	0,98	0,97	0,34	0,89	0,63	1,00	1,00	0,78	1,00	0,97	0,93	0,92	0,86
<b>Přesnost</b>	0,98	1,00	1,00	0,97	1,00	0,93	0,96	0,97	1,00	1,00	1,00	0,39	0,93
<b>F-míra</b>	0,98	0,98	0,50	0,92	0,77	0,96	0,97	0,86	1,00	0,98	0,96	0,54	0,86

**Tabulka 5.14:** Vyhodnocení klasifikace gest při použití kartézských příznaků a vlastního klasifikátoru.

Z hodnot uvedených v Tabulkách 5.9–5.14 můžeme vypočítat nedostatečné výsledky všech tří kombinací klasifikátorů a příznaků pro gesta *jízda vpřed* a *jízda vzad*. Při bližším prozkoumání jednotlivých výsledků se ukázalo, že příčinou je falešná detekce pozice *dopředu nahoru* při rychlém pohybu horní končetiny mezi pozicemi *dopředu* a *dopředu kolmo*. Tento problém se projevuje u všech tří kombinací, nejvýrazněji však u vlastního klasifikátoru. Protože pozice *dopředu nahoru* není obsažena v žádném gestu, je možné ji při aktualizaci sekvencí pozic použitých pro detekci gesta ignorovat. Tím by mělo být zajištěno zvýšení senzitivity gest *jízda vpřed* a *jízda vzad*, ale zároveň tím může dojít ke zvýšení možnosti falešné detekce gest, jejichž definující pozice sousedí s ignorovanou pozicí. Tabulky 5.15–5.20 obsahují výsledky dosažené ignorací pozice *dopředu nahoru* při vytváření sekvencí pro detekci dynamických gest.

	Skutečná gesta												
	0	1	2	3	4	5	6	7	8	9	10	jiné	
0	<b>97</b>	0	0	0	0	0	0	0	0	0	0	2	
1	0	<b>97</b>	0	0	0	0	0	0	0	0	0	0	
2	0	0	<b>83</b>	0	0	0	0	0	3	3	0	4	
3	0	0	0	<b>100</b>	0	0	0	0	0	0	18	2	
4	0	0	0	0	<b>81</b>	0	0	0	0	4	0	0	
5	0	0	0	0	0	<b>100</b>	0	0	0	0	1	0	
6	0	1	0	0	1	0	<b>100</b>	0	0	1	0	3	
7	0	0	0	0	0	0	0	<b>85</b>	0	0	0	2	
8	0	1	0	0	0	0	0	0	<b>97</b>	0	0	0	
9	0	0	0	0	1	0	0	0	0	<b>90</b>	0	0	
10	0	0	0	0	0	0	0	0	0	0	<b>81</b>	0	
žádné	3	1	17	0	17	0	0	15	0	2	0	<b>87</b>	

**Tabulka 5.15:** Matice záměn klasifikovaných gest při použití kartézských příznaků a naivního Bayesovského klasifikátoru s ignorací pozice *dopředu nahoru*.

	Gesto												Průměrná hodnota
	0	1	2	3	4	5	6	7	8	9	10	jiné	
<b>Senzitivita</b>	0,97	0,97	0,83	1,00	0,81	1,00	1,00	0,85	0,97	0,90	0,81	0,87	0,91
<b>Přesnost</b>	0,97	1,00	0,89	0,83	0,95	0,99	0,94	0,97	0,98	0,98	1,00	0,61	0,92
<b>F-míra</b>	0,97	0,98	0,85	0,90	0,87	0,99	0,96	0,90	0,97	0,93	0,89	0,71	0,91

**Tabulka 5.16:** Vyhodnocení klasifikace gest při použití kartézských příznaků a naivního Bayesovského klasifikátoru s ignorací pozice *dopředu nahoru*.

	Skutečná gesta												
	0	1	2	3	4	5	6	7	8	9	10	jiné	
<b>Klasifikovaná gesta</b>	0	<b>92</b>	0	0	0	0	0	0	0	0	0	0	6
	1	0	<b>94</b>	0	0	0	0	0	0	0	0	0	0
	2	0	0	<b>65</b>	0	0	0	0	0	1	1	0	0
	3	0	0	0	<b>100</b>	0	1	0	0	0	0	3	2
	4	0	0	0	0	<b>72</b>	1	0	0	0	3	0	0
	5	0	0	0	0	0	<b>97</b>	0	0	0	0	17	0
	6	0	0	0	0	1	0	<b>90</b>	0	0	1	0	3
	7	0	0	3	0	0	0	0	<b>84</b>	0	0	0	4
	8	0	2	0	0	0	0	10	0	<b>99</b>	38	0	1
	9	0	0	0	0	0	0	0	0	0	<b>55</b>	0	0
	10	0	0	0	0	0	0	0	0	0	0	<b>79</b>	0
žádné	8	4	32	0	27	1	0	16	0	2	1	<b>84</b>	

**Tabulka 5.17:** Matice záměn klasifikovaných gest při použití úhlových příznaků a naivního Bayesovského klasifikátoru s ignorací pozice *dopředu nahoru*.

	Gesto												Průměrná hodnota
	0	1	2	3	4	5	6	7	8	9	10	jiné	
<b>Senzitivita</b>	0,92	0,94	0,65	1,00	0,72	0,97	0,90	0,84	0,99	0,55	0,79	0,84	0,84
<b>Přesnost</b>	0,93	1,00	0,97	0,94	0,94	0,85	0,94	0,92	0,66	1,00	1,00	0,48	0,88
<b>F-míra</b>	0,92	0,96	0,77	0,96	0,81	0,90	0,91	0,87	0,79	0,70	0,88	0,61	0,84

**Tabulka 5.18:** Vyhodnocení klasifikace gest při použití úhlových příznaků a naivního Bayesovského klasifikátoru s ignorací pozice *dopředu nahoru*.



		Skutečná gesta											
		0	1	2	3	4	5	6	7	8	9	10	jiné
Klasifikovaná gesta	0	98	0	0	0	0	0	0	0	0	0	0	3
	1	0	97	0	0	0	0	0	0	0	0	0	0
	2	0	0	80	0	0	0	0	0	0	0	0	0
	3	0	0	0	89	0	0	0	0	0	0	0	2
	4	0	0	0	0	74	0	0	0	0	0	0	0
	5	0	0	0	0	0	100	0	0	0	0	7	0
	6	0	0	0	0	0	0	100	0	0	1	0	3
	7	0	0	0	0	0	0	0	78	0	0	0	2
	8	0	0	0	0	0	0	0	0	100	0	0	0
	9	0	0	0	0	0	0	0	0	0	97	0	0
	10	0	0	0	0	0	0	0	0	0	0	93	0
žádné	2	3	20	11	26	0	0	22	0	2	0	90	

**Tabulka 5.19:** Matice záměn klasifikovaných gest při použití kartézských příznaků a vlastního klasifikátoru s ignorací pozice *dopředu nahoru*.

	Gesto											Průměrná hodnota	
	0	1	2	3	4	5	6	7	8	9	10		jiné
Senzitivita	0,98	0,97	0,80	0,89	0,74	1,00	1,00	0,78	1,00	0,97	0,93	0,90	0,91
Přesnost	0,97	1,00	1,00	0,97	1,00	0,93	0,96	0,97	1,00	1,00	1,00	0,51	0,94
F-míra	0,97	0,98	0,88	0,92	0,85	0,96	0,97	0,86	1,00	0,98	0,96	0,65	0,91

**Tabulka 5.20:** Vyhodnocení klasifikace gest při použití kartézských příznaků a vlastního klasifikátoru s ignorací pozice *dopředu nahoru*.

Ve výsledcích uvedených v Tabulkách 5.15–5.20 lze skutečně pozorovat zvýšení senzitivity gest *jízda vpřed* a *jízda vzad* i snížení schopnosti potlačení falešných pozorování gesta *aktivace*. Ignorace pozice *dopředu nahoru* měla pozitivní dopad na výsledky klasifikace obzvláště při použití vlastního klasifikátoru s relativními kartézskými souřadnicemi. Na zbývající dvě kombinace klasifikátoru a příznaků měla tato změna výrazně nižší vliv.

Nejlepší klasifikace gest při ignoraci pozice *dopředu nahoru* tedy byla dosažena použitím vlastního klasifikátoru a kartézských příznaků. Pokud byly použity všechny detekované pozice, byl však lepší volbou naivní Bayesovský klasifikátor s relativními kartézskými souřadnicemi, který v tomto případě dokázal nejlépe detekovat i problémová gesta.

Nejhorších výsledků pak v obou případech dosáhla kombinace úhlových příznaků s naivním Bayesovským klasifikátorem, při jejímž použití můžeme oproti ostatním kombinacím vidět nízkou přesnost klasifikace gesta *otáčení* a *pozastavit ovládání*. Tyto zhoršené výsledky jsou nejspíše způsobeny nepřesnostmi při odhadu pozice kostry člověka, které mají za následek příliš velké chyby v měření vzájemných úhlů.

Ná základě dosažených výsledků a následného testování na reálném robotu bylo jako výchozí nastavení zvoleno použití kombinace relativních kartézských souřadnic a vlastního klasifikátoru s ignorací pozice *dopředu nahoru*.

## 5.4 Problémy

V této části jsou popsány problémy, které se nepodařilo v rámci práce dostatečně potlačit nebo které vzhledem k použitému hardwaru či softwaru ani potlačit nelze.

Prvním takovým problémem, který nemá vzhledem k použitému typu senzoru uspokojivé řešení, je pohlcování infračerveného záření některými povrchy, což může vést ke znemožnění sledování části těla či dokonce k nemožnosti detekce celé osoby. Ukázka vlivu takového materiálu na získaný hloubkový obraz je uvedena na Obrázku 5.1. Tomuto problému je tedy možné se vyvarovat pouze zamezením nošení povrchů, které výrazně pohlcují infračervené záření.

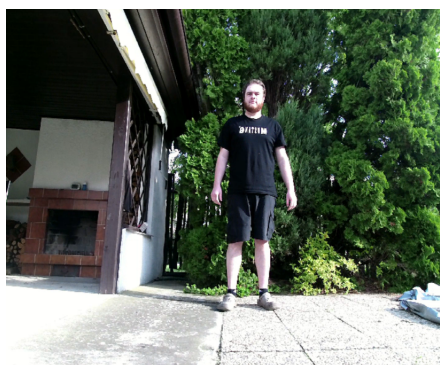


(a) : Povrch nepohlcující infračervené záření.

(b) : Povrch pohlcující infračervené záření.

**Obrázek 5.1:** Vliv povrchu pohlcujícího infračervené záření.

Vzhledem k použité infračervené technologii pro získání hloubkového obrazu je použití senzoru problematické při osvětlení operátora intenzivním slunečním zářením (viz Obrázek 5.2), které znemožňuje získání hloubkových dat a následnou detekci osoby [77]. Z tohoto důvodu je tedy téměř nemožné využití tohoto systému ve venkovních prostorech, či prostorech kde je vysoké riziko intenzivního slunečního svitu [77].



(a) : RGB obraz při nepřímém slunečním svitu.



(b) : Hloubkový obraz při nepřímém slunečním svitu.



(c) : RGB obraz při přímém slunečním svitu.



(d) : Hloubkový obraz při přímém slunečním svitu.

**Obrázek 5.2:** Vliv intenzivního slunečního záření.

Poslední problém souvisí s využitým softwarem pro detekci osoby v obraze a následné sledování. Při užívání a testování bylo totiž zjištěno, že nedokáže spolehlivě sledovat polohu končetin při jejich rychlém pohybu v případech, kdy je končetina natažena směrem k senzoru a dochází tedy k jejímu částečnému zakrytí sebou samou. To způsobuje snížení spolehlivosti detekce dopředných pozic horní končetiny při pohybu, z čehož plyne nižší spolehlivost detekce gest, která jsou těmito pozicemi definována. I z tohoto důvodu je pro provedení akcí *jízda vpřed*, *jízda vzad* a *zastavení* možné využít i gesta, která neobsahují dopředné pozice horních končetin.

## 5.5 Možná vylepšení

Mezi možná vylepšení implementovaného systému se řadí zvětšení množiny detekovaných poloh horních končetin a s tím související možnost použití většího množství ovládacích gest. Prostřednictvím těchto nových gest by mohl operátor například měnit maximální dovolenou rychlost robotu, ať už z důvodu omezení hluku či pro zvýšení bezpečnosti.

Zvýšení spolehlivosti klasifikace by mohlo být dosaženo úpravou vlastního klasifikátoru tak, aby namísto klasifikace podle vzdálenosti mezi referenční a testovanou polohou vymezil při učení prostor (např. elipsoid), ve kterém se s určitou pravděpodobností vyskytují trénovací data, a poté zjišťoval, zda klasifikovaná pozice leží ve vymezeném prostoru.

Dalším možným vylepšením rozšiřujícím možnosti systému je schopnost zapamatovat si podobu operátora pro případ ztráty vizuálního kontaktu. Tato funkce by poté umožňovala nové možnosti ovládání, například příkaz pro otočení robotu o předem definovaný úhel i s vědomím ztráty vizuálního kontaktu s operátorem.

Následující možnosti už přímo nesouvisí s ovládáním gesty, ale týkají se celkové funkčnosti. Prvním takovým vylepšením je využití Kinectu pro detekci překážek na cestě a následnou úpravu trajektorie s ohledem na detekovanou překážku. Možností je také použití mikrofonů Kinectu pro hlasové ovládání v případech, kdy je operátor mimo zorné pole senzoru nebo pokud by bylo třeba použít složitější příkazy, např. konkrétní nastavení některých parametrů systému.

## 5.6 Možná využití

Využití tohoto způsobu ovládání leží v oblastech, kde není potřeba precizní ovládání pohybu mobilního robotu a kde se člověk nemůže či nechce ovládání plně věnovat. Tento způsob je tedy vhodný například pro situaci, kdy má robotický vozík převážet náklad a může při tom následovat člověka, který ho dovede do cílové oblasti. Pokud bychom vzali v úvahu implementaci uvedených možných vylepšení, mohl by být výsledný systém nainstalován například na vozíku se zbožím, který následuje skladníka při převozu zboží v rámci skladu.

## Kapitola 6

### Závěr

V rámci této práce byla provedena rešerše stávajících metod rozpoznávání postav, poloh částí těla a gest. Na jejím základě byl navržen a implementován systém pro ovládání mobilního robotu gesty. Tento systém využívá pro detekci postav a částí jejich těl senzor Kinect v2. Systém byl navržen pro rozpoznání celkem 11 různých gest prováděných horními končetinami a prostřednictvím těchto gest je možné provádět celkem 8 operací pro ovládání mobilního robotu. Výstupem výsledné implementace jsou dvě hodnoty určené pro řízení pohybu mobilního robotu, konkrétně se jedná o dopřednou a rotační rychlost.

Pro popis jednotlivých pozic horních končetin byly zvoleny dva různé způsoby, první využívá relativní polohy zápěstí a loktů vůči ramenům, druhý používá vhodně definované úhly mezi tělem a jednotlivými částmi horní končetiny.

Pro následnou klasifikaci byly otestovány dva různé klasifikátory. Prvním byl naivní Bayesovský klasifikátor pro příznaky s předpokládaným normálním rozdělením. Druhý klasifikátor byl navržen a implementován přímo pro potřeby této úlohy. Tento klasifikátor provádí klasifikaci na základě vzdáleností mezi detekovanými polohami kloubů a jejich referenčními polohami pro každou definovanou pozici horní končetiny.

Vyhodnocení dosažených výsledků bylo provedeno určením přesnosti, senzitivity a F-míry klasifikace jednotlivých pozic horních končetin a samotných gest. Jako nejspolehlivější metodou pro klasifikaci pozic horních končetin se ukázalo být použití relativních kartézských souřadnic v kombinaci s vlastním klasifikátorem, které dosáhlo průměrné senzitivity 95 %, přesnosti 96 % a F-míry 95 %. Použití naivního Bayesovského klasifikátoru se stejným druhem příznaků oproti tomu zaznamenalo průměrnou senzitivitu 89 %, přesnost 87 % a F-míru 87 %. Nejhorší výsledky byly získány s naivním Bayesovským klasifikátorem s úhlovými příznaky, konkrétně se jednalo o průměrnou senzitivitu 86 %, přesnost 79 % a F-míru 82 %.

Pokud se zaměříme na detekci samotných gest, zde byly výsledky určeny ve dvou různých nastaveních. V prvním případě byly pro klasifikaci gest standardně použity všechny definované pozice horních končetin. Ve druhém případě byla ignorována pozice *dopředu nahoru* z důvodu jejích falešně pozitivních klasifikací. Pro první nastavení bylo dosaženo nejlepších výsledků s naivním Bayesovským klasifikátorem v kombinaci s relativními kartézskými souřadnicemi, pro který byla zaznamenána průměrná senzitivita 90 %, přesnost 92 % a F-míra 90 %. Ve druhém případě se jako nejlepší ukázalo použití vlastního klasifikátoru s relativními kartézskými souřadnicemi, které dosáhlo průměrné senzitivity 91 %, přesnosti 94 % a F-míry 91 %. V obou případech byly opět zaznamenány nejhorší výsledky s kombinací naivního Bayesovského klasifikátoru a úhlových příznaků.

Ná základě získaných výsledků a následného testování na reálném robotu byla pro detekci gest k následnému ovládní zvolena kombinace relativních kartézských souřadnic a vlastního klasifikátoru s ignorací pozice *dopředu nahoru*.



## Literatura

- [1] How will humans interface with computers in the future? [online]. HowStuffWorks [cit. 5.4.2020]. Dostupné z: <https://computer.howstuffworks.com/humans-interface-with-computers.htm>
- [2] Next Big Smart Home Trend: Gesture Control of Lights, Thermostats, and Other Devices? [online]. Electronic House [cit. 4.4.2020]. Dostupné z: <https://www.electronichouse.com/smart-home/next-big-smart-home-trend-gesture-control-of-lights-thermostats-and-other-devices/>
- [3] WEBER, Matthias, Heni Ben AMOR a Thomas ALEXANDER. Identifying Motion Capture Tracking Markers with Self-Organizing Maps. In: 2008 IEEE Virtual Reality Conference [online]. IEEE, 2008, 2008, s. 297-298 [cit. 18.5.2020]. DOI: 10.1109/VR.2008.4480809. ISBN 978-1-4244-1971-5. Dostupné z: <https://doi.org/doi:10.1109/VR.2008.4480809>
- [4] OptiTrack's newly released motion-capture suit. In: What Mocap Suit Suits You? - VFX Voice Magazine VFX Voice Magazine. [online]. Copyright © 2017 Walt Disney Pictures [cit. 13.5.2020]. Dostupné z: <https://www.vfxvoice.com/what-mocap-suit-suits-you/>
- [5] Andy Serkis in motion capture suit. In: What you need to know about 3D motion capture | Engadget. [online]. Copyright © 2020 Verizon Media. All rights reserved. [cit. 13.5.2020]. Dostupné z: <https://www.engadget.com/2014-07-14-motion-capture-explainer.html>
- [6] TOSHEV, Alexander a Christian SZEGEDY. DeepPose: Human Pose Estimation via Deep Neural Networks. In: 2014 IEEE Conference on



- Computer Vision and Pattern Recognition [online]. IEEE, 2014, 2014, s. 1653-1660 [cit. 14.5.2020]. DOI: 10.1109/CVPR.2014.214. ISBN 978-1-4799-5118-5. Dostupné z: <https://doi.org/doi:10.1109/cvpr.2014.214>
- [7] MEHTA, Dushyant, Srinath SRIDHAR, Oleksandr SOTNYCHENKO, et al. VNect. ACM Transactions on Graphics [online]. 2017, 36(4), 1-14 [cit. 14.5.2020]. DOI: 10.1145/3072959.3073596. ISSN 0730-0301. Dostupné z: <https://doi.org/doi:10.1145/3072959.3073596>
- [8] HAGGAG, Hussein, Mohammed HOSSNY, Despina FILIPPIDIS, Douglas CREIGHTON, Saeid NAHAVANDI a Vinod PURI. Measuring depth accuracy in RGBD cameras. In: 2013, 7th International Conference on Signal Processing and Communication Systems (ICSPCS) [online]. IEEE, 2013, 2013, s. 1-7 [cit. 19.5.2020]. DOI: 10.1109/ICSPCS.2013.6723971. ISBN 978-1-4799-1319-0. Dostupné z: <https://doi.org/doi:10.1109/icspcs.2013.6723971>
- [9] Human Silhouette Extraction from Depth Videos - Rivindu Madushan. [online]. Medium [cit. 20.5.2020]. Dostupné z: <https://medium.com/@rivindum.15/human-silhouette-extraction-from-depth-videos-b84d2ac77bd4>
- [10] XIA, Lu, Chia-Chih CHEN a J. K. AGGARWAL. Human detection using depth information by Kinect. In: CVPR 2011 WORKSHOPS [online]. IEEE, 2011, 2011, s. 15-22 [cit. 12.5.2020]. DOI: 10.1109/CVPRW.2011.5981811. ISBN 978-1-4577-0529-8. Dostupné z: <https://doi.org/doi:10.1109/cvprw.2011.5981811>
- [11] MARTINEZ-GONZALEZ, Angel, Michael VILLAMIZAR, Olivier CANEVET a Jean-Marc ODOBEZ. Efficient Convolutional Neural Networks for Depth-Based Multi-Person Pose Estimation. IEEE Transactions on Circuits and Systems for Video Technology [online]. 2019, 1-1 [cit. 5.5.2020]. DOI: 10.1109/TCSVT.2019.2952779. ISSN 1051-8215. Dostupné z: <https://doi.org/doi:10.1109/TCSVT.2019.2952779>
- [12] MARTIEZ-GONZALEZ, Angel, Michael VILLAMIZAR, Olivier CANEVET a Jean-Marc ODOBEZ. Real-time Convolutional Networks for Depth-based Human Pose Estimation. In: 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) [online]. IEEE, 2018, s. 41-47 [cit. 5.5.2020]. DOI: 10.1109/IROS.2018.8593383. ISBN 978-1-5386-8094-0. Dostupné z: <https://doi.org/doi:10.1109/IROS.2018.8593383>
- [13] SCHWARZ, Loren Arthur, Artashes MKHITARYAN, Diana MATEUS a Nassir NAVAB. Human skeleton tracking from depth data using geodesic distances and optical flow. Image and Vision Computing [online]. 2012, 30(3), 217-226 [cit. 11.5.2020]. DOI: 10.1016/j.imavis.2011.12.001. ISSN 02628856. Dostupné z: <https://doi.org/doi:10.1016/j.imavis.2011.12.001>



- [14] KIM, Hyesuk a Incheol KIM. Dynamic Arm Gesture Recognition Using Spherical Angle Features and Hidden Markov Models. *Advances in Human-Computer Interaction* [online]. 2015, 2015, 1-7 [cit. 11.5.2020]. DOI: 10.1155/2015/785349. ISSN 1687-5893. Dostupné z: <https://doi.org/doi:10.1155/2015/785349>
- [15] SAHA, Sriparna, Biswarup GANGULY a Amit KONAR. Gesture based improved human-computer interaction using Microsoft's Kinect sensor. In: *2016 International Conference on Microelectronics, Computing and Communications (MicroCom)* [online]. IEEE, 2016, 2016, s. 1-6 [cit. 11.5.2020]. DOI: 10.1109/MicroCom.2016.7522582. ISBN 978-1-4673-6621-2. Dostupné z: <https://doi.org/doi:10.1109/microcom.2016.7522582>
- [16] WEN, Yan, Chuanyan HU, Guanghui YU a Changbo WANG. A robust method of detecting hand gestures using depth sensors. In: *2012 IEEE International Workshop on Haptic Audio Visual Environments and Games (HAVE 2012) Proceedings* [online]. IEEE, 2012, 2012, s. 72-77 [cit. 13.5.2020]. DOI: 10.1109/HAVE.2012.6374441. ISBN 978-1-4673-1567-8. Dostupné z: <https://doi.org/doi:10.1109/have.2012.6374441>
- [17] YU, Chenglong, Xuan WANG, Hejiao HUANG, Jianping SHEN a Kun WU. Vision-Based Hand Gesture Recognition Using Combinational Features. In: *2010 Sixth International Conference on Intelligent Information Hiding and Multimedia Signal Processing* [online]. IEEE, 2010, 2010, s. 543-546 [cit. 13.5.2020]. DOI: 10.1109/IIHMSP.2010.138. ISBN 978-1-4244-8378-5. Dostupné z: <https://doi.org/doi:10.1109/iihm-sp.2010.138>
- [18] KULKARNI, Vaishali a S.D.LOKHANDE. (2010). Appearance Based Recognition of American Sign Language Using Gesture Segmentation. In: *International Journal on Computer Science and Engineering* [online]. Vol. 02, No. 03, 2010, s. 560-565 [cit. 13.5.2020]. Dostupné z: <https://pdfs.semanticscholar.org/f010/4600cb00b8be3b4a801fdde290df41d07aee.pdf>
- [19] ZHENYAO MO a U. NEUMANN. Real-time Hand Pose Recognition Using Low-Resolution Depth Images. In: *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Volume 2 (CVPR'06)* [online]. IEEE, 2006, s. 1499-1505 [cit. 2020-05-20]. DOI: 10.1109/CVPR.2006.237. ISBN 0-7695-2597-0. Dostupné z: <https://doi.org/doi:10.1109/CVPR.2006.237>
- [20] WU, Jonathan, James CHRISTIANSON, Janusz KONRAD a Prakash ISHWAR. Leveraging shape and depth in user authentication from in-air hand gestures. In: *2015 IEEE International Conference on Image Processing (ICIP)* [online]. IEEE, 2015, 2015, s. 3195-3199 [cit. 2020-05-20].



- Conference on Robotics and Biomimetics (ROBIO) [online]. IEEE, 2013, 2013, s. 798-803 [cit. 13.4.2020]. DOI: 10.1109/ROBIO.2013.6739560. ISBN 978-1-4799-2744-9. Dostupné z: <https://doi.org/doi:10.1109/robio.2013.6739560>
- [29] HEARST, M.A., S.T. DUMAIS, E. OSUNA, J. PLATT a B. SCHOLKOPF. Support vector machines. IEEE Intelligent Systems and their Applications [online]. 1998, 13(4), 18-28 [cit. 25.4.2020]. DOI: 10.1109/5254.708428. ISSN 1094-7167. Dostupné z: <https://doi.org/doi:10.1109/5254.708428>
- [30] YI LIU a Y.F. ZHENG. One-against-all multi-class SVM classification using reliability measures. In: Proceedings. 2005 IEEE International Joint Conference on Neural Networks, 2005 [online]. IEEE, 2005, s. 849-854 [cit. 25.4.2020]. DOI: 10.1109/IJCNN.2005.1555963. ISBN 0-7803-9048-2. Dostupné z: <https://doi.org/doi:10.1109/ijcnn.2005.1555963>
- [31] PRISACARIU, Victor Adrian a Ian REID. 3D hand tracking for human computer interaction. Image and Vision Computing [online]. 2012, 30(3), 236-250 [cit. 25.4.2020]. DOI: 10.1016/j.imavis.2012.01.003. ISSN 02628856. Dostupné z: <https://doi.org/doi:10.1016/j.imavis.2012.01.003>
- [32] BISWAS, K. K. a Saurav Kumar BASU. Gesture recognition using Microsoft Kinect®. In: The 5th International Conference on Automation, Robotics and Applications [online]. IEEE, 2011, 2011, s. 100-103 [cit. 25.4.2020]. DOI: 10.1109/ICARA.2011.6144864. ISBN 978-1-4577-0330-0. Dostupné z: <https://doi.org/doi:10.1109/icara.2011.6144864>
- [33] BAUTISTA, Miguel Ángel, Antonio HERNÁNDEZ-VELA, Victor PONCE, Xavier PEREZ-SALA, Xavier BARÓ, Oriol PUJOL, Cecilio ANGULO a Sergio ESCALERA. Probability-Based Dynamic Time Warping for Gesture Recognition on RGB-D Data. JIANG, Xiaoyi, Olga Regina Pereira BELLON, Dmitry GOLDGOF a Takeshi OISHI, ed. Advances in Depth Image Analysis and Applications [online]. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, 2013, s. 126-135 [cit. 15.5.2020]. Lecture Notes in Computer Science. DOI: 10.1007/978-3-642-40303-3\_14. ISBN 978-3-642-40302-6. Dostupné z: [https://doi.org/doi:10.1007/978-3-642-40303-3\\_14](https://doi.org/doi:10.1007/978-3-642-40303-3_14)
- [34] RIOFRIO, Santiago, David POZO, Jorge ROSERO a Juan VASQUEZ. Gesture Recognition Using Dynamic Time Warping and Kinect: A Practical Approach. In: 2017 International Conference on Information Systems and Computer Science (INCISCOS) [online]. IEEE, 2017, 2017, s. 302-308 [cit. 15.5.2020]. DOI: 10.1109/INCISCOS.2017.36. ISBN 978-1-5386-2644-3. Dostupné z: <https://doi.org/doi:10.1109/inciscos.2017.36>



- [44] CHANSRI, Chana a Jakkree SRINONCHAT. Hand Gesture Recognition for Thai Sign Language in Complex Background Using Fusion of Depth and Color Video. *Procedia Computer Science* [online]. 2016, 86, 257-260 [cit. 4.1.2020]. DOI: 10.1016/j.procs.2016.05.113. ISSN 18770509. Dostupné z: <https://doi.org/doi:10.1016/j.procs.2016.05.113>
- [45] SHARATH KUMAR, Y. H. a V. VINUTHA. Hand Gesture Recognition for Sign Language: A Skeleton Approach. DAS, Swagatam, Tandra PAL, Samarjit KAR, Suresh Chandra SATAPATHY a Jyotsna Kumar MANDAL, ed. *Proceedings of the 4th International Conference on Frontiers in Intelligent Computing: Theory and Applications (FICTA) 2015* [online]. New Delhi: Springer India, 2016, 2016-10-25, s. 611-623 [cit. 4.1.2020]. *Advances in Intelligent Systems and Computing*. DOI: 10.1007/978-81-322-2695-6\_52. ISBN 978-81-322-2693-2. Dostupné z: [https://doi.org/doi:10.1007/978-81-322-2695-6\\_52](https://doi.org/doi:10.1007/978-81-322-2695-6_52)
- [46] ZHAO, Hongyang, Shuangquan WANG, Gang ZHOU a Daqing ZHANG. Gesture-Enabled Remote Control for Healthcare. In: *2017 IEEE/ACM International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE)* [online]. IEEE, 2017, s. 392-401 [cit. 16.12.2019]. DOI: 10.1109/CHASE.2017.123. ISBN 978-1-5090-4722-2. Dostupné z: <https://doi.org/doi:10.1109/chase.2017.123>
- [47] WACHS, J. P., H. I. STERN, Y. EDAN, M. GILLAM, J. HANDLER, C. FEIED a M. SMITH. A Gesture-based Tool for Sterile Browsing of Radiology Images. *Journal of the American Medical Informatics Association* [online]. 2008, 15(3), s. 321-323 [cit. 5.1.2020]. DOI: 10.1197/jamia.M241. ISSN 1067-5027. Dostupné z: <https://doi.org/doi:10.1197/jamia.M241>
- [48] ZHOU, Dajun, Minghui SHI, Fei CHAO, Chih-Min LIN, Longzhi YANG, Changjing SHANG a Changle ZHOU. Use of human gestures for controlling a mobile robot via adaptive CMAC network and fuzzy logic controller. *Neurocomputing* [online]. 2018, 282, 218-231 [cit. 29.4.2020]. DOI: 10.1016/j.neucom.2017.12.016. ISSN 09252312. Dostupné z: <https://doi.org/doi:10.1016/j.neucom.2017.12.016>
- [49] BARCAJ, Adam. Návrh dálkového ovládní mobilního robotu pomocí Microsoft Kinect [online]. Brno, 2012 [cit. 17.5.2020]. Diplomová práce. Vysoké učení technické v Brně. Fakulta strojního inženýrství. Dostupné z: <http://hdl.handle.net/11012/9284>
- [50] 10 Ways to Use SmartCapture with the Mavic Air - DJI Guides [online]. Copyright © 2020 DJI ALL RIGHTS RESERVED [cit. 16.5.2020]. Dostupné z: <https://store.dji.com/guides/use-smartcapture-with-the-mavic-air/>
- [51] DJI Spark – Gesture Control – Selfie. In: *DJI Spark Gesture Control - How to use it?* [online]. el Producente [cit. 16.5.2020]. Dostupné z:

<http://elproducente.com/dji-spark-gesture-control-how-to-use-it-review-manual/>

- [52] Brave NUI World: Rise of touch-less gesture control [online]. UX Planet [cit. 16.5.2020]. Dostupné z: <https://uxplanet.org/brave-nui-world-rise-of-touch-less-gesture-control-882be077cdfa>
- [53] JustSketchMe [online]. JustSketchMe [cit. 4.1.2020]. Dostupné z: <https://justsketchme.web.app/>
- [54] Naive Bayes Classifiers [online]. GeeksforGeeks [cit. 2.1.2020]. Dostupné z: <https://www.geeksforgeeks.org/naive-bayes-classifiers/>
- [55] Desktop | Ubuntu CZ/SK [online]. Copyright © Canonical Ltd. Ubuntu a Canonical jsou registrovanými známkami společnosti Canonical Ltd. [cit. 3.1.2020]. Dostupné z: <https://www.ubuntu.cz/desktop/>
- [56] Welcome to Python.org [online]. Copyright ©2001 [cit. 3.1.2020]. Dostupné z: <https://www.python.org/>
- [57] GitHub - severin-lemaignan/openni-python: Python bindings for OpenNI2 and NiTE2. [online]. GitHub, Inc. [cit. 2.12.2019]. Dostupné z: <https://github.com/severin-lemaignan/openni-python>
- [58] Kivy: Cross-platform Python Framework for NUI Development [online]. [cit. 3.12.2019]. Dostupné z: <https://kivy.org/>
- [59] NumPy [online]. Copyright © Copyright 2019 NumPy developers. [cit. 4.1.2020]. Dostupné z: <https://numpy.org/>
- [60] F. PEDREGOSA, G. VAROQUAUX, A. GRAMFORT, V. Michel, B. THIRION, O. GRISEL, M. BLONDEL, P. PRETTENHOFER, R. WEISS, V. DUBOURG, J. VANDERPLAS, A. PASSOS, D. COURNAPEAU, M. BRUCHER, M. PERROT, a E. DUCHESNAY, Scikit-learn: Machine learning in Python. In: Journal of Machine Learning Research [online]. 12 (2011), s. 2825–2830. [cit. 4.1.2020]. Dostupné z: <http://www.jmlr.org/papers/volume12/pedregosa11a/pedregosa11a.pdf>
- [61] GitHub - TaylorSMarks/playsound: Pure Python, cross platform, single function module with no dependencies for playing sounds [online]. GitHub, Inc. [cit. 4.1.2020]. Dostupné z: <https://github.com/TaylorSMarks/playsound>
- [62] GitHub - m-lundberg/simple-pid: A simple and easy to use PID controller in Python [online]. GitHub, Inc. [cit. 4.1.2020]. Dostupné z: <https://github.com/m-lundberg/simple-pid>
- [63] GitHub - lace/vg: Linear algebra for humans. [online]. GitHub, Inc. [cit. 2.12.2019]. Dostupné z: <https://github.com/lace/vg>



- [64] CASWELL, Thomas A, Michael DROETTBOOM, John HUNTER, Eric FIRING, Antony LEE, Jody KLYMAK, David STANSBY, Elliott Sales De ANDRADE, Jens Hedegaard NIELSEN, Nelle VAROQUAUX, Benjamin ROOT, Tim HOFFMANN, Phil ELSON, Ryan MAY, Darren DALE, JAE-JOON LEE, Jouni K. SEPPÄNEN, Damon MCDOUGALL, Andrew STRAW, Paul HOBSON, Christoph GOHLKE, Tony S YU, Eric MA, Adrien F. VINCENT, Steven SILVESTER, Charlie MOAD, Jan KATINS, Nikita KNIAZEV, Federico ARIZA a Elan ERNEST. matplotlib v2.2.4 [online]. 2019, Zenodo. [cit. 9.5.2020]. Dostupné z: <https://doi.org/doi:10.5281/ZENODO.2669103>
- [65] LINGZHU XIANG, Florian ECHTLER, Christian KERL, Thiemo WIEDEMMEYER, , LARS, HANYAZOU, Ryan GORDON, Francisco FACIONI, LABORER2008, Rich WAREHAM, Matthias GOLDHORN, ALBERTH, GABORPAPP, Steffen FUCHS, JMTATSCH, Joshua BLAKE, , FEDERICO, Henning JUNGKURTH, Yuan MINGZE, VI-NOUZ, Dave COLEMAN, Brendan BURNS, Rahul RAWAT, Serguei MOKHOV, Paul REYNOLDS, P.E. VIAU, Matthieu FRAISSINET-TACHET, LUDIQUE, James BILLINGHAM a , ALISTAIR. Libfree-nect2: Release 0.2 [online]. 2016, Zenodo. [cit. 9.5.2020]. Dostupné z: <https://doi.org/10.5281/ZENODO.50641>
- [66] OpenNI Programmer's Guide. [online]. [cit. 3.12.2019]. Dostupné z: [https://s3.amazonaws.com/com.occipital.openni/OpenNI\\_Programmers\\_Guide.pdf](https://s3.amazonaws.com/com.occipital.openni/OpenNI_Programmers_Guide.pdf)
- [67] Kinect v2 | Kinect Support [online]. Copyright © 2019 [cit. 2.12.2019]. Dostupné z: <https://support.xbox.com/en-US/xbox-on-windows/accessories/kinect-for-windows-v2-info>
- [68] CHANG, M. a Z. KANG. AN INDOOR SLAM METHOD BASED ON KINECT AND MULTI-FEATURE EXTENDED INFORMATION FILTER. ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences [online]. 2017, XLII-2/W7, s. 327-332 [cit. 29.12.2019]. DOI: 10.5194/isprs-archives-XLII-2-W7-327-2017. ISSN 2194-9034. Dostupné z: <https://doi.org/doi:10.5194/isprs-archives-XLII-2-W7-327-2017>
- [69] KHAWALDEH, Saed, Tajwar Abrar ALEEF, Usama PERVAIZ, Vu Hoang MINH a Yeman Brhane HAGOS. Complete End-to-End Low Cost Solution to a 3D Scanning System with Integrated Turntable. International Journal of Computer Science and Information Technology [online]. 2017, 9(4), s. 39-55 [cit. 29.12.2019]. DOI: 10.5121/ijcsit.2017.9404. ISSN 09754660. Dostupné z: <https://doi.org/doi:10.5121/ijcsit.2017.9404>
- [70] LEKOVA, A., D. RYAN a R. DAVIDRAJUH. Fingers and Gesture Recognition with Kinect v2 Sensor. Information Technologies and Control [online]. 2016, 14(3), s. 24-30 [cit. 4.1.2020]. DOI: 10.1515/itc-

- 2017-0009. ISSN 1312-2622. Dostupné z: <https://doi.org/doi:10.1515/itc-2017-0009>
- [71] XBOX ONE - Senzor Kinect v2. In: ATC Market - XBOX ONE - Senzor Kinect v2. ATC Market [online]. Copyright © AT Computers a.s. 1998 [cit. 9.5.2020]. Dostupné z: <https://www.atcmarket.cz/detail/XBOX-ONE-Senzor-Kinect-v2/438784>
- [72] WASENMÜLLER, Oliver a Didier STRICKER. Comparison of Kinect V1 and V2 Depth Images in Terms of Accuracy and Precision. CHEN, Chu-Song, Jiwen LU a Kai-Kuang MA, ed. Computer Vision – ACCV 2016 Workshops [online]. Cham: Springer International Publishing, 2017, 2017-03-16, s. 34-45 [cit. 17.12.2019]. Lecture Notes in Computer Science. DOI: 10.1007/978-3-319-54427-4\_3. ISBN 978-3-319-54426-7. Dostupné z: [https://doi.org/doi:10.1007/978-3-319-54427-4\\_3](https://doi.org/doi:10.1007/978-3-319-54427-4_3)
- [73] SELL, John a Patrick O'CONNOR. The Xbox One System on a Chip and Kinect Sensor. IEEE Micro [online]. 2014, 34(2), s. 44-53 [cit. 5.12.2019]. DOI: 10.1109/MM.2014.9. ISSN 0272-1732. Dostupné z: <https://doi.org/doi:10.1109/MM.2014.9>
- [74] Husky UGV - Outdoor Field Research Robot by Clearpath [online]. Copyright © Clearpath Robotics Inc. All rights reserved. [cit. 4.1.2020]. Dostupné z: <https://clearpathrobotics.com/husky-unmanned-ground-vehicle-robot/>
- [75] POWERS, David. Evaluation: From precision, recall and F-measure to ROC, informedness, markedness & correlation. Journal of Machine Learning Technologies [online] 2011, 2. 2229-3981 [cit. 9.5.2020]. DOI: 10.9735/2229-3981. ISSN 22293981. Dostupné z: <https://doi.org/doi:10.9735/2229-3981>
- [76] FAWCETT, Tom. An introduction to ROC analysis. Pattern Recognition Letters [online]. 2006, 27(8), 861-874 [cit. 9.5.2020]. DOI: 10.1016/j.patrec.2005.10.010. ISSN 01678655. Dostupné z: <https://doi.org/doi:10.1016/j.patrec.2005.10.010>
- [77] SUAREZ, Jesus a Robin R. MURPHY. Using the Kinect for search and rescue robotics. In: 2012 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR) [online]. IEEE, 2012, 2012, s. 1-2 [cit. 9.5.2020]. DOI: 10.1109/SSRR.2012.6523918. ISBN 978-1-4799-0165-4. Dostupné z: <https://doi.org/doi:10.1109/ssrr.2012.6523918>



## Příloha A

### Obsah přiloženého CD

Na přiloženém CD se kromě elektronické verze této práce nachází zdrojový kód vypracovaného ovládacího systému v programovacím jazyce Python, vytvořená data pro učení a testování klasifikátorů, zdrojové soubory této práce ve formátu  $\text{\LaTeX}$ a videozáznam s demonstrací funkce ovládacího systému.

Obsah CD je rozdělen do následujících adresářů a podadresářů:

- **latex** – zdrojové soubory textu této práce ve formátu,  $\text{\LaTeX}$
- **média** – fotografie použitého robotu s nainstalovaným senzorem Kinect v2, videozáznam zachycující ovládání robotu pomocí gest,
- **python** – zdrojové soubory programu v jazyce Python, vytvořený dataset, externí knihovny NiTE2 a OpenNI2,
  - **NiTE2** – software pro detekci kostry v hloubkových datech,
  - **OpenNI2** – software pro přístup k datům ze senzoru,
  - **position\_data** – datové soubory s pozicemi a gesty,
    - **test\_gestures** – testovací gesta ve formátu NumPy matic,
    - **test\_positions** – testovací pozice ve formátu NumPy matic,
    - **train\_positions** – trénovací pozice ve formátu NumPy matic,
    - **validation\_positions** – validační pozice ve formátu NumPy matic,
  - **sounds** – zvukové soubory pro slovní oznámení.



## Příloha B

### Použité zkratky

Zkratka	Význam
CNN	konvoluční neuronová síť (Convolutional Neural Network)
DNN	hluboká neuronová síť (Deep Neural Network)
DTW	metoda dynamického borcení času (Dynamic Time Warping)
FN	falešně negativní pozorování (False Negative)
FP	falešně pozitivní pozorování (False Positive)
GUI	grafické uživatelské rozhraní (Graphical User Interface)
HMM	skrytý Markovův model (Hidden Markov Model)
LL	levý loket
LZ	levé zápěstí
NUI	přirozené uživatelské rozhraní (Natural User Interface)
OS	operační systém
PL	pravý loket
PZ	pravé zápěstí
RGB	červená-zelená-modrá (Red-Green-Blue)
SDK	sada vývojových nástrojů (Software Development Kit)
SLAM	simultánní lokalizace a mapování (Simultaneous Localization and Mapping)
SVM	metoda podpůrných vektorů (Support Vector Machine)
TDNN	neuronová síť s časovým zpožděním (Time Delay Neural Network)
TP	skutečně pozitivní pozorování (True Positive)





## **Příloha C**

### **Konfigurační soubor**

Tabulka s popisem jednotlivých parametrů konfiguračního souboru se nachází na následující straně.

Parametr	Popis	Obor hodnot
classification_mode	Kombinace příznaků a klasifikátoru.	$\{\text{bayer-cartesian, bayer-angles, sphere-distance}\}$
ignore_front_high_position	Ignore pozice <i>dopředu nahoru</i> při klasifikaci gest.	$\{0, 1\}$
soft_margin_threshold	Boční ohraničení 1. bezpečností úrovně.	$\mathbb{R}^+$
hard_margin_threshold	Boční ohraničení 2. bezpečností úrovně.	$\mathbb{R}^+$
depth_near_soft_threshold	Minimální nutná vzdálenost operátora pro 1. bezpečností úroveň.	$\mathbb{R}^+$
depth_near_hard_threshold	Minimální nutná vzdálenost operátora pro 2. bezpečností úroveň.	$\mathbb{R}^+$
depth_far_soft_threshold	Maximální dovolená vzdálenost operátora pro 1. bezpečností úroveň.	$\mathbb{R}^+$
depth_far_hard_threshold	Maximální dovolená vzdálenost operátora pro 2. bezpečností úroveň.	$\mathbb{R}^+$
safety_timer	Maximální dovolený čas překročení 2. bezpečností úrovně.	$\mathbb{R}^+$
angular_limit	Maximální dovolená rotační rychlost.	$\mathbb{R}^+$
angular_control_threshold	Minimální požadovaná hodnota rotační rychlosti spuštění pohybu.	$\mathbb{R}^+$
angular_P	Proporcionální složka PID regulátoru pro rotační rychlost.	$\mathbb{R}$
angular_I	Integrační složka PID regulátoru pro rotační rychlost.	$\mathbb{R}$
angular_D	Derivační složka PID regulátoru pro rotační rychlost.	$\mathbb{R}$
linear_setpoint	Požadovaná vzdálenost robotu od operátora.	$\mathbb{R}^+$
linear_limit	Maximální dovolená dopředná rychlost.	$\mathbb{R}^+$
linear_control_threshold	Minimální požadovaná hodnota dopředné rychlosti spuštění pohybu.	$\mathbb{R}^+$
linear_P	Proporcionální složka PID regulátoru pro dopřednou rychlost.	$\mathbb{R}$
linear_I	Integrační složka PID regulátoru pro dopřednou rychlost.	$\mathbb{R}$
linear_D	Derivační složka PID regulátoru pro dopřednou rychlost.	$\mathbb{R}$
only_beeping	Použití pípání místo slovních oznámení.	$\{0, 1\}$