



Západočeská univerzita v Plzni
Fakulta aplikovaných věd

BAKALÁŘSKÁ PRÁCE

Plzeň, 2023

Petr ANTAL



Západočeská univerzita v Plzni
Fakulta aplikovaných věd

BAKALÁŘSKÁ PRÁCE
Hlasová protéza ovládaná gesty ruky

Plzeň, 2023

Petr ANTAL

ZÁPADOČESKÁ UNIVERZITA V PLZNI

Fakulta aplikovaných věd
Akademický rok: 2022/2023

ZADÁNÍ BAKALÁŘSKÉ PRÁCE

(projektu, uměleckého díla, uměleckého výkonu)

Jméno a příjmení: **Petr ANTAL**
Osobní číslo: **A20B0312P**
Studijní program: **B0714A150005 Kybernetika a řídicí technika**
Specializace: **Automatické řízení a robotika**
Téma práce: **Hlasová protéza ovládaná gesty ruky**
Zadávající katedra: **Katedra kybernetiky**

Zásady pro vypracování

1. Nastudujte problematiku rozpoznávání gest a jejich následného vyhodnocení, proveďte rešerši dostupných senzorů a technologií vhodných pro rozpoznání pohybů ruky a prstů.
2. Navrhněte způsob rozpoznávání gest pomocí vybraných jednoduchých senzorů.
3. Navrhněte a realizujte převod gest do psaného textu. Psaný text s využitím syntézy řeči převedte na mluvenou řeč.

Rozsah bakalářské práce: **30-40 stránek A4**
Rozsah grafických prací:
Forma zpracování bakalářské práce: **tištěná**

Seznam doporučené literatury:

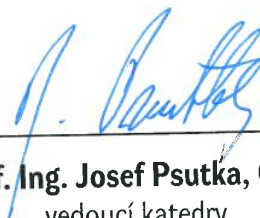
Dodá vedoucí práce.

Vedoucí bakalářské práce: **Ing. Luboš Šmídl, Ph.D.**
Katedra kybernetiky

Datum zadání bakalářské práce: **17. října 2022**
Termín odevzdání bakalářské práce: **22. května 2023**



Doc. Ing. Miloš Železný, Ph.D.
děkan



Prof. Ing. Josef Psutka, CSc.
vedoucí katedry

V Plzni dne 17. října 2022

Prohlášení

Prohlašuji, že jsem bakalářskou práci na téma:

Hlasová protéza ovládaná gesty ruky

vypracoval samostatně pod odborným dohledem vedoucího bakalářské práce za použití pramenů uvedených v přiložené bibliografii.

V Plzni dne 14. srpna 2023

.....

Petr Antal

Poděkování

Chtěl bych vyjádřit svou vděčnost vedoucímu práce, Ing. Luboši Šmídlovi, Ph.D., za cenné rady a pečlivé konzultace, které mi poskytl při vedení této bakalářské práce. A stejně tak děkuji své rodině za neocenitelnou podporu při studiu.

Abstrakt

Bakalářská práce se zaměřuje na využití nejnovějších technologií k zlepšení kvality života jedincům s komunikačním postižením. Pro tuto skupinu je problematická komunikace s lidmi, kteří neovládají znakový jazyk. V první části práce jsou diskutovány stávající technologické postupy a funkční systémy, které se snaží řešit problémy v každodenním životě němých lidí. Zaměření práce se soustředí na rukavice pro rozpoznávání gest znakového jazyka, ale jsou zde také zmiňovány jiné alternativní výzkumy, například systémy založené na počítačovém vidění nebo invazivní operace. Dále je vytvořeno srovnání těchto technologií a diskutovány možnosti dalšího vývoje v dané oblasti.

Druhá část práce se zaměřuje na návrh a konstrukci systému pro rozpoznávání gest pomocí jednoduchých senzorů a jejich převod do psaného textu. Následně je tento text transformován na mluvené slovo pomocí syntézy řeči. Na závěr jsou diskutovány možnosti dalšího vývoje a aplikace této technologie v jiných oblastech.

Klíčová slova: rozpoznávání gest, datová rukavice, znakový jazyk, senzor, rozhraní člověk-stroj, rozpoznávání vzorů, klasifikace

Abstract

The bachelor's thesis focuses on the use of the latest technologies to improve the quality of life for individuals with communication disorder. Communication with people who do not know sign language is problematic for deaf people. The first part of the thesis discusses existing technological approaches and functional systems that aim to address the daily challenges faced by deaf individuals. The focus of the thesis is on gloves for recognizing sign language gestures, but other alternative research, such as computer vision-based systems or invasive surgery, is also mentioned. Furthermore, a comparison of these technologies is created, and the possibilities for further development in the given area are discussed.

The second part of the thesis focuses on the design and construction of a gesture recognition system using simple sensors and their conversion into written text. Subsequently, this text is transformed into spoken words using speech synthesis. Finally, the possibilities for further development and application of this technology in other areas are discussed.

Keywords: gesture recognition, data glove, sign language, sensor, human-machine interface, pattern recognition, classification

Obsah

| | | |
|----------|--|-----------|
| 1 | Úvod | 13 |
| 1.1 | Naše motivace v této práci | 13 |
| 1.2 | Popis cíle a přínos práce | 13 |
| 2 | Medicínské přístupy k obnově hlasu | 14 |
| 2.1 | Ne-invazivní metody | 14 |
| 2.1.1 | Elektrolarynx | 14 |
| 2.1.2 | Pneumatický umělý larynx | 15 |
| 2.1.3 | Ezofageální řeč | 15 |
| 2.2 | Invazivní metody | 15 |
| 2.2.1 | Tracheoezofageální punkce a protéza | 15 |
| 2.2.2 | TEP a protéza u pacientů s neoesofagem | 16 |
| 2.3 | Lékařský slovník | 17 |
| 3 | Rozpoznávání gest | 18 |
| 3.1 | Technologie datové rukavice | 21 |
| 3.1.1 | Senzory na rozpoznávání ohybu prstu | 21 |
| 4 | Hlasová syntéza ovládaná datovou rukavicí | 24 |
| 5 | Návrh a implementace datové rukavice pro ovládání syntézy hlasu | 25 |
| 5.1 | Počítačový software | 26 |
| 5.2 | 1. Prototyp | 30 |
| 5.3 | 2. Prototyp | 32 |
| 6 | Práce s rukavicí a programem | 34 |
| 7 | Diskuze o rozšíření a vylepšení navržené protézy | 36 |
| 8 | Závěr a shrnutí výsledků práce | 37 |
| | Reference | 38 |

1 Úvod

1.1 Naše motivace v této práci

Na katedře kybernetiky se zabýváme syntézou řeči pěknou řádkou let a bylo by při nejmenším zajímavé ji využít z nové perspektivy. Hlasová protéza umožní lidem s omezenou schopností mluvit znovu komunikovat a byla by neocenitelným pomocníkem v každodenním životě.

Hlasová protéza může být lákavou alternativou k existujícím řešením, jako jsou invazivní operace. Pacienti by tak měli více možností volby, což může mít pozitivní vliv na jejich život. Použití těchto protéz může významně zlepšit kvalitu života lidí s postižením tím, že jim umožní nezávislost a samostatnost v každodenním životě. To může zahrnovat schopnost komunikovat s ostatními lidmi, ovládat prostředí a získávat nové zkušenosti. Tyto přístroje mohou také pomoci lidem s postižením zlepšit svou pracovní a vzdělávací výkonnost a omezit sociální izolaci.

Nicméně i když jsou hlasové protézy ovládané gesty ruky velmi slibné, stále existuje řada výzev, kterým je třeba čelit. Proto je inovace a vývoj v této oblasti velmi žádaná.

Vývoj hlasové protézy ovládané gesty ruky může být motivován také výzkumem v oblasti rozpoznávání gest a jiných pohybových aktivit. Tento typ protézy může sloužit jako testovací platforma pro nové algoritmy a metody rozpoznávání gest, které se mohou využít v různých odvětvích jako je virtuální realita či ovládání humanoidních robotických paží.

1.2 Popis cíle a přínos práce

Naším cílem je vytvořit průzkum dosavadní technologie datové rukavice, kterou využíváme jako médium pro získání dat polohy prstů a rozpoznání gest. Na základě toho si vytvoříme vlastní prototypový hardware a software, čímž se ponoříme do problematiky hlouběji. Tím si vytvoříme obraz o tom co si můžeme dovolit a jak dále rozvíjet naši myšlenku.

2 Medicínské přístupy k obnově hlasu

Hlas je jednou ze základních komunikačních technik a tvoří se na základě proudění vzduchu a tlaku, který uvádí do pohybu hlasivky. To vytváří základní zvuk v hrtanu. Tento zvuk je dále artikulován v ústech pomocí rtů, jazyka a zubů. Hlas je tvořen tedy především dýchacím ústrojím, tj. plíce, průdušky, průdušnice, hrtan, hltan a ústní a nosní dutina se podílí na tvorbě hlasu [1, 2]. Jakmile vypadne jediná z těchto komponent, tak se hlas buď nevytvoří nebo nebude dokonalý (chraplavý, špatná intonace, tichým, aj.). Ztrátou hlasu se samozřejmě zabývá medicína, která vyvinula několik různých způsobů, jak navrátit postižené osobě její hlas. Tyto metody můžeme rozdělit na neinvazivní a invazivní neboli operativní. Každá z těchto metod má své klady a zápory a stejně tak důvody, proč například zvolit jednu metodu před druhou. Ale stejně tak má poslední slovo pacient, který řekne, jestli si zvolí operativní metodu nebo jinou možnou alternativu.

Vzhledem k využití odborných medicínských termínů je na konci kapitoly uveden slovník [3, 4].

2.1 Ne-invazivní metody

2.1.1 Elektrolarynx

Je to elektrické zařízení, které při zapnutí a kontaktu s měkkou tkání (přiložením na vnější část krku) vytváří slyšitelné vibrace. Tyto vibrace může pacient využít k vytvoření slov pomocí rtů, zubů a jazyka. Tento přístroj nejčastěji využívají lidé, jenž podstoupili laryngektomii, nebo také tracheostomii. Existuje několik provedení tohoto přístroje, ale principiálně jsou stejné - přiložení elektrolarynxu k oblasti pharynxu, vytvoření vibrací nahrazuje funkci hlasivek, artikulace a tvorba hlasového projevu.

Hlavní výhodou elektrolarynxu je jeho jednoduchost. Pacient se s ním velmi rychle naučí zacházet a snadno se s ním manipuluje.

Nevýhodou je umělý robotický hlas, nelze intonovat a je společensky velmi nápadný. Zařízení je na baterie, což je značně omezující [1, 5, 6, 7].

2.1.2 Pneumatický umělý larynx

Je to pneumatické zařízení poháněné foukáním vzduchu do trubičky vedoucí do stomy, kde se následně rozvibruje o membránu a nahradí funkci larynxu. Patří to k levnějším metodám obnovy hlasu po tracheostomii. Tato metoda se často využívá v Tchaj-wanu, jelikož je v jazycích východního světa velký důraz na intonaci, kterou PUL poskytuje.

Výhodou PUL je jednoduché a levné provedení, rychlé naučení a možnost intonace. A i přestože je toto zařízení velmi účinné, tak nese velké, převážně sociální, nedostatky. Zařízení je zevní a potřebuje častou údržbu, nedá se skrýt před zraky ostatních, takže budí velkou pozornost a někteří pacienti dokonce přestanou používat PUL kvůli pocitu trapnosti při používání [1, 8].

2.1.3 Ezofageální řeč

Ezofageální nebo-li jícnová řeč je způsob komunikace, kde je zvuk vytvářen kmitáním jícnového ústí. V tomto případě není potřeba jakýkoliv vnější zásah, ale pouze cvik a péče pacienta na naučení této metody. Musíme podotknout, že tato metoda je považována za tu nejtěžší ze všech a pouze 50 - 60% pacientů je schopno se tuto metodu komunikace naučit. Ezofageální řeč využívají pacienti zejména po totální laryngektomii.

Nicméně tato metoda navzdory velké obtížnosti není závislá na bateriích nebo jiném zařízení, což konkuruje předchozím metodám [1, 7].

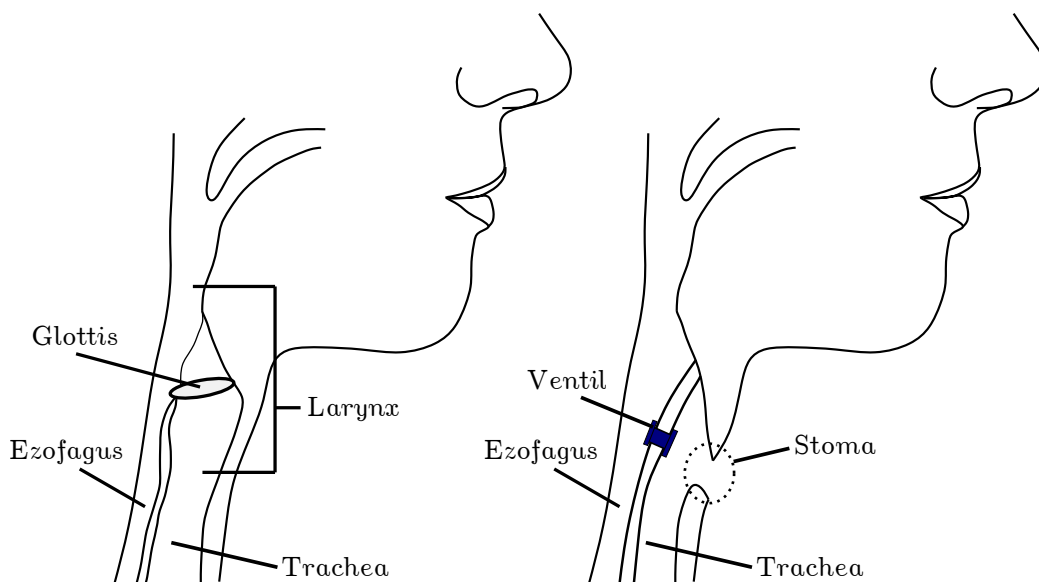
2.2 Invazivní metody

2.2.1 Tracheoezofageální punkce a protéza

Tracheoezofageální punkce (TEP) je proces, kdy se mezi tracheou a ezofagem vytvoří fistula, do které se vloží trubička s jednosměrnou záklopkou. Pro její zavedení se vytvoří tracheo stoma, která bude také fungovat jako mechanismus pro odklonění vzduchu z trachey do ezofagu. Při ucpání stomy prstem nebo umělohmotnou záklopkou se vzduch z trachey dostane přes trubičku do ezofagu a pharynxu. Stěny ezofagu se rozvibrují a začnou vytvářet zvuk. Následně se tento zvuk artikuluje za pomoci jazyka, zubů a rtů.

Jelikož má trubička jednosměrnou záklopkou, tak se nestane, že se do trachey dostane potrava či tekutiny. Procedura vyžaduje minimální trénink, má nízké riziko aspirace a nevyžaduje žádné externí zařízení. Nesmíme ale opomenout potenciální komplikace při zavedení cizího tělesa do organismu, jako

je například růst plísní a bakterií, zanesení hlenem, vdechnutí nebo spolknutí trubičky a stenóza trachey [1, 7, 9, 10].



Obrázek 1.: Tracheoesofageální punkce (TEP)

2.2.2 TEP a protéza u pacientů s neoesofagem

Procedura pro pacienty s neoesofagem se nijak zvlášť neliší od obyčejné TEP operace. Fistula je vytvořena mezi tracheou a neoesofagem, přičemž existují různé možnosti pro neoesofagus. Jedny z možností jsou gastric pull-up (doslova vytažení žaludku ke krku) nebo transplantace ilea či jejunumu [1].

2.3 Lékařský slovník

Aspirace: Vdechnutí cizího tělesa nebo tekutiny do dolních dýchacích cest.

Ezofagus: Jícen, část trávicí trubice mezi hltanem a žaludkem.

Fistula: Píštěl, abnormální kanálek tvořící komunikaci mezi dutinou a jejím povrchem.

Glottis: Hlasivky, úsek hrtanu sloužící k tvorbě hlasu.

Ileum: Kyčelník, třetí a poslední část tenkého střeva, která ústí do tlustého střeva do slepého střeva.

Jejunum: Lačník, druhá část tenkého střeva.

Laryngektomie: Chirurgické odstranění hrtanu, například z důvodu zhoubného nádoru.

Larynx: Hrtan, součást horních cest dýchacích, slouží k tvorbě hlasu.

Neoezofagus: Nový nebo umělý jícen.

Pharynx: Hltan, společný oddíl trávicí a dýchací soustavy, kde se potrava smršťováním svalů posouvá do jícnu.

Stenóza trachey: Abnormální zúžení průdušnice.

Stoma: Dočasné nebo trvalé vyvedení některého dutého lidského orgánu na povrch těla.

Trachea: Průdušnice, trubice spojující hrtan s průduškami plic.

Tracheostomie: Umělé vyústění průdušnice na povrch těla.

3 Rozpoznávání gest

Chceme-li ulehčit život neslyšícím a němým, tak bychom měli uvažovat jakým způsobem žijí doposud a jak komunikují s okolím. Dosavadním a naprosto základním způsobem komunikace pro tyto lidi je znakový jazyk.

Znakový jazyk je komunikační systém založen na gestikulaci a mimice. Je to nonverbální komunikace, která má určitá pravidla a dokonce různé druhy (znakový jazyk není univerzální - v každé zemi znamenají gesta něco jiného).

Cílem výzkumu pro rozpoznání gest je analyzovat gesto, přiřadit ho ke gestu v databázi a dát mu význam [11].

Jsou tři základní přístupy pro rozpoznání gest: vizuální, senzorický a hybridní [12, 13].

Vizuální rozpoznávání gest je založeno na počítačovém vidění, tj. informace o gestech získává primárně vizuálními vstupními daty. Systém využívá buď videa nebo jen snímky pro identifikování gest. Klíčovým prvkem jsou samozřejmě kamery, které jsou v dnešní době levné a nijak neomezují fyzicky uživatele, což je jejich hlavní výhodou.

Nicméně právě s kamerou přichází problémy jako omezený zorný úhel a vysoká výpočetní zátěž. Proto je třeba pokrýt tyto problémy více kamerami z různých úhlů a decentralizovat zpracování dat pro jednotlivé kamery a výsledky nakonec spojit.

Výpočetní metody k determinaci gesta z vizuální informace lze rozdělit na tři skupiny:

- **Metody pro zpracování obrazu:**
Houghova transformace, vlnková transformace, lokální binární vzory, kombinace orientačního histogramu a statických vlastností.
- **Metody zpracování pohybu:**
Dynamické programování, Kalmanův filtr, optický tok, fuzzy logika
- **Strojové učení a neuronové sítě**

Vlnková transformace je matematická metoda pro analýzu signálů a obrazů, která umožňuje rozklad signálu na různé frekvenční složky. Tento rozklad je založen na použití vlnkových funkcí, které mají specifické vlastnosti a umožňují reprezentovat signál v různých rozlišeních. Vlnkové funkce jsou v podstatě krátké kmitající signály, které mají specifický tvar a charakteristiky. Nejprve se signál rozdělí na několik částí, které mají různé frekvenční složky. Poté se použijí vlnkové funkce k analýze každé části signálu a získání koeficientů, které popisují, jak moc se vlnková funkce shoduje s danou částí signálu. Tyto koeficienty pak mohou být použity k rekonstrukci původního signálu nebo k jeho další analýze [14].

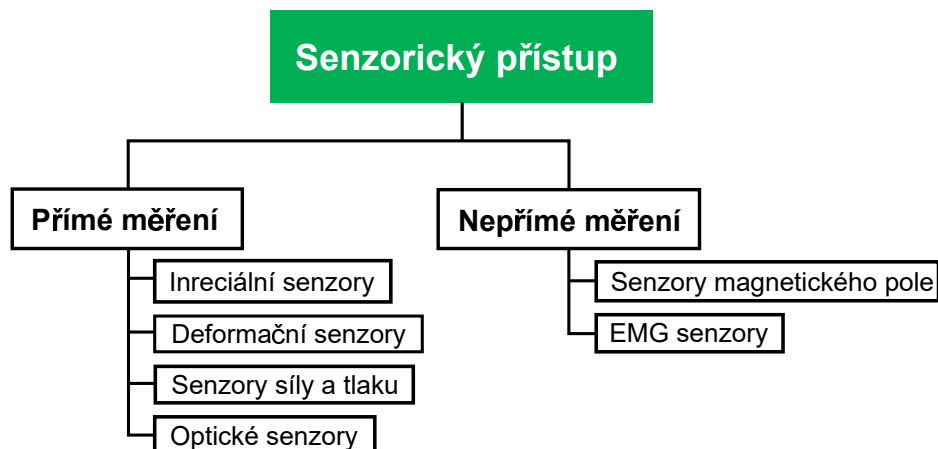
Houghova transformace je matematická technika, která se používá pro detekci geometrických tvarů v digitálním obrazu. Původně byla navržena pro detekci přímek v obrazu, ale později byla rozšířena i na detekci dalších tvarů, jako jsou kružnice, elipsy atd. Hlavním principem Houghovy transformace je transformace bodů v obrazu do parametrického prostoru, ve kterém lze hledat geometrické tvary. Například přímka v obrazu lze reprezentovat dvěma parametry, jako jsou úhel natočení a vzdálenost od středu obrazu [15].

Lokální binární vzory je metoda, která umožňuje popsat vzhled každého pixelu v obraze vzhledem k jeho okolí pomocí binárních kódů. LBV popisuje každý pixel v obraze jako binární řetězec, kde každé číslo v řetězci určuje, zda je sousední pixel větší nebo menší než středový pixel. Tyto binární řetězce jsou poté analyzovány pro získání texturových informací z obrazu. LBV se používají k detekci objektů v obraze, segmentaci obrazu, rozpoznávání obličejů a dalších úloh [16].

Kalmanův filtr je matematický algoritmus používaný pro odhadování stavů systému na základě měření, přičemž berou v úvahu šum a nejistoty v datech. Je to iterativní proces, který se skládá z predikce a aktualizace stavu [17].

Fuzzy logika je matematická teorie, která umožňuje modelovat a řešit nejistotu a neurčitost v systémech. Na rozdíl od klasické binární logiky, kde jsou pravdivostní hodnoty výroků buď pravdivé (1) nebo nepravdivé (0), fuzzy logika umožňuje pracovat s hodnotami mezi těmito krajními hodnotami, tzv. "fuzzy" hodnotami. Tyto fuzzy hodnoty mohou být vyjádřeny například jako pravděpodobnost nebo jako stupeň příslušnosti k nějakému rozsahu [18].

Senzorický přístup napovídá využití více přímočarého přístupu k získání dat o poloze prstů a ruky. Získání gesta senzory můžeme rozčlenit do několik podkategorií:



Obrázek 2.: Senzory pro rozpoznání gest

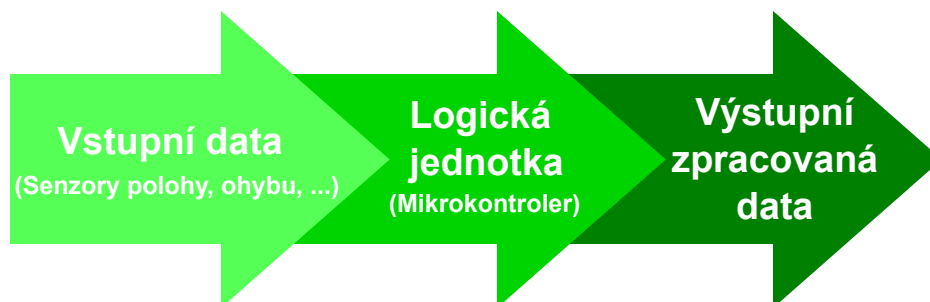
Data z těchto sensorů jsou zpracována a interpretována jako gesta. Na rozdíl od vizuálního přístupu, tak senzorický přístup není omezen zorným polem kamerového systému - systém pro rozpoznávání gest se stává mobilním. Bohužel to uživatele naopak omezuje v jeho pohyblivosti samotných prstů a zápěstí.

A nakonec hybridní systém. **Hybridní přístup** kombinuje silné stránky senzorického a vizuálního přístupu. Obecně dojde k dělbě práce - vizuální přístup má na starosti analýzu gesta a senzorický přístup pouze pohyb na základě akcelerometru a gyroskopu. Tímto rozdělením práce se zvýší přesnost a spolehlivost. Přičemž se úměrně zvýší složitost a výpočetní zátěž.

Naším cílem bude vytvořit mobilní komunikační jednotku, takže sáhneme po senzorickém přístupu, ze kterého budeme tvořit datovou rukavici.

3.1 Technologie datové rukavice

Datová rukavice je koncept reprezentace alias digitalizace gest a dát jim význam. Esenciálně je datová rukavice vstupní zařízení, se kterým jsme schopni něco ovládat. Datovou rukavici lze logicky rozdělit na 3 části:



Obrázek 3.: Obecný koncept datové rukavice

Koncept je v mnoha výzkumech stejný, ale provedení se liší na základě využitých senzorů [12, 13, 19, 20]. Pro rozpoznání gesta potřebujeme vždy 2 věci - polohu (ohyb) prstů a orientaci (pohyb) ruky.

3.1.1 Senzory na rozpoznávání ohybu prstu

Na rozpoznání ohybu prstu lze využít flexní senzory, optické senzory, taktilní senzory a senzory pro měření intenzity magnetického pole.

Flexní senzory jsou postaveny na základě deformace (ohybu), čímž mění svůj elektrický odpor (zvyšují/snižují odpor na základě složení materiálu), tj. fungují jako potenciometry. Ze změny napětí protékajícím senzorem odhadujeme, jak moc se sensor spojený s prstem ohnul. Byť je flexibilita, hmotnost a nízká cena senzoru bezkonkurenční, tak citlivost, rozsah měření a životnost má rezervu pro zlepšení [21].

Optický senzor se pro měření ohybu prstu provádí připojením optického kanálku, ve kterém proudí světlo a změna intenzity je dále měřitelná. Jinak řečeno se při ohybu prstu ohýbá také optický kanálek, ve kterém se snižuje intenzita světla z důvodů odrazů a absorpce světla o stěny kanálku. Tato

změna je dále převedena na elektrický signál, se kterým se nadále pracuje. Výhodou těchto senzorů je vysoká přesnost měření a schopnost rozpoznat i jemné pohyby, malá velikost a snadná integrace s jinými technologiemi. Optické senzory patří spíše do dražší kategorie a vliv okolního světla může mít negativní vliv na měření. Také je nutná kalibrace pro každého uživatele, což je to časově náročné [22].

Taktilí senzory jsou vyrobeny z odolných polymerů, které při aplikaci síly mění své elektrické vlastnosti. Ohybem prstů bude na senzor působit tah a tlak způsoben připojením senzoru k rukavici. Stejně jako u flexních senzorů záleží jaký materiál byl použit a dle toho se odpor senzoru bude zvyšovat nebo snižovat působením sil. Senzory jsou víc robustní, jelikož jsou vyrobeny polymerů. Jejich úctyhodná citlivost však bohužel nese určité nevýhody. Kvůli vysoké citlivosti signál obsahuje šum a při špatné aplikaci je signál pro rozpoznání gesta dokonce až nepoužitelný [23].

Posledním typem jsou **senzory pro měření magnetické intenzity**. Mezi nejvíc využívané senzory v této oblasti jsou Hallovy sondy, které jsou schopny detekovat přítomnost magnetického pole nebo jejich intenzitu. Senzory jsou poměrně přesné a spolehlivé. Jedna z nevýhod je snadné ovlivnění měření vnějším prostředím [24].

Pro rozpoznání **pohybu a orientace ruky** se využívá kombinace akcelerometru a gyroskopu. Akcelerometr nám měří rychlost a zrychlení ruky, z čehož se dá vypočítat pohyb provedený rukou. Gyroskop nám určí úhlovou rychlost a orientaci ruky. Tyto senzory se vzájemně skvěle doplňují a jedno z možných rozšíření je přidání magnetometru, který poskytuje informaci o orientaci ruky v prostoru.

Elektrické signály ze senzorů se musí přivést do logické jednotky, která přijatá data zpracuje. Pro tuto činnost se využívají jednočipové počítače - mikrokontrolery, které jsou lehké a dostatečně výkonné. Většinou se využívají Arduina, ESP nebo STM systémy. Každý z nich má dostatečný výkon a jsou flexibilní pro jakoukoliv aplikaci.

Výstupem logické jednotky by mělo být gesto nebo data pro rekonstrukci gesta, která se dále mohou zpracovat. Tady se rozcházejí výzkumy, které se touto technologií zabývaly. Jedna část studií zpracovávala a vytvářela

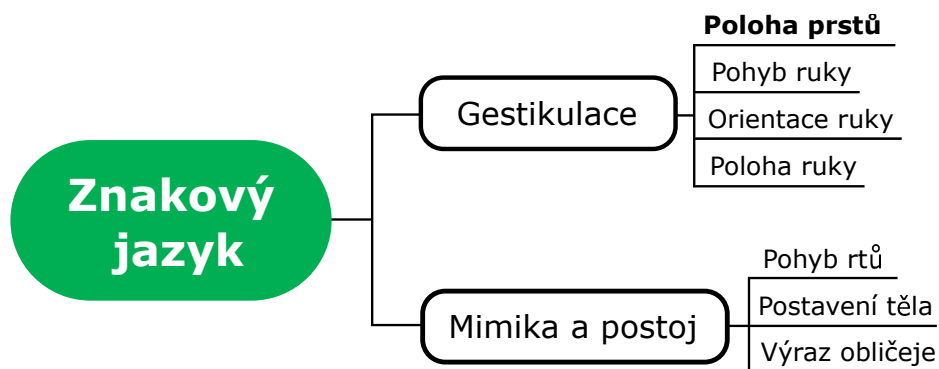
výstup pouze v mikrokontroleru, což udělalo plně funkční a samostatný celek. Druhá část odesílala data do počítače, kde se provádělo další zpracování dat. Důvodem byla konkrétní aplikace rozpoznávání gest a způsob zpracování informací.

4 Hlasová syntéza ovládaná datovou rukavicí

Využití datové rukavice pro ovládání syntetizovaného hlasu umožní uživatelům mluvit bez potřeby použití klávesnice a myši. Jelikož uživatel využívá znakový jazyk, má lepší kontrolu nad tím, co chce říct, a může si také sám definovat, co jednotlivé gesto znamená. Znakový jazyk není univerzální, ale se syntézou řeči může uživatel mluvit několika jazyky pomocí jedné znakové řeči, kterou zná. Tím se stává datová rukavice univerzálním překladatelem do všech jazyků. Samozřejmě musí existovat samotný hlasový modul v daném jazyce.

Nám byl poskytnut hlasový syntezátor katedrou kybernetiky od společnosti *Speechtech*. *Text-to-speech* (TTS) modul se využívá pro získání zvukové nahrávky požadovaného textu. Text je generován na základě vstupu datové rukavice a definice jednotlivých gest.

Vyvstává ale otázka, zda je rukavice schopna reprezentovat znakový jazyk. Znakový jazyk se skládá ze dvou podstatných částí:



Obrázek 4.: Rozdělení znakového jazyka

Manuální část jsme schopni s určitou přesností reprezentovat datovou rukavicí. Ale mimika a postoj se musí rozpoznávat dodatečnými senzory nebo vizuálně (kamerou).

My si vytvoříme unikátní sadu gest, která bude blízko znakovému jazyku, ale nebude potřeba využít mimiku. Přičemž v našem případě se zaměříme na zpracování polohy prstů.

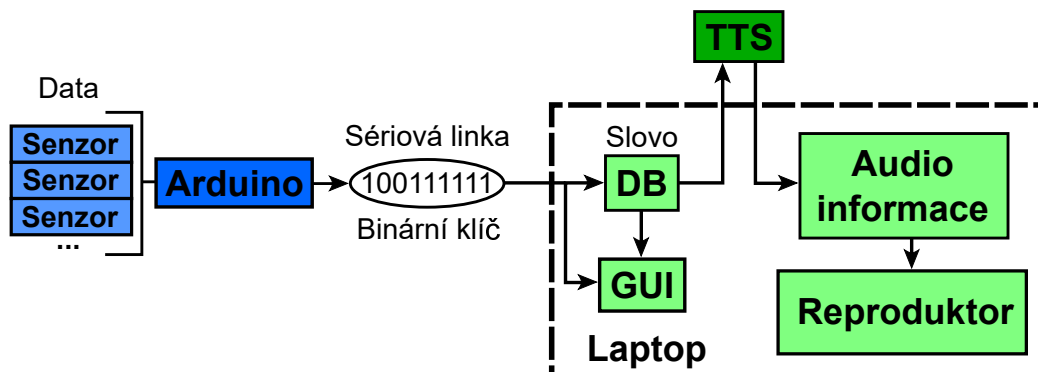
5 Návrh a implementace datové rukavice pro ovládání syntézy hlasu

Naprostým základem hlasové protézy je rozpoznávání gest. Zvolili jsme senzorický přístup a budeme se pokoušet rozeznat gesto na základě toho, zda byl určitý kloub ohnutý nebo ne. Jinak řečeno budeme řešit pouze dva stavy prstu - extenze (narovnaný prst), flexe (ohnutý prst). Pro větší přesnost budeme mít na každém prstu dva senzory pro dva články prstu. Výjimkou bude palec, který sice má vyšší stupeň volnosti než ostatní prsty, ale jeho polohu prozatím zobecňujeme jedním senzorem.

Logickou jednotkou bude Arduino Nano, se který se jednoduše pracuje a bude mít na starosti vytvářet z dat senzorů gesto. Gesto bude charakterizováno binárním klíčem. Jednička v binárním klíči bude reprezentovat flexi prstu a nula extenzi.

Zvolili jsme binární klíč s cílem přímého čtení a ověření reakce senzorů podle našich očekávání a také pro snadný přechod mezi různými prototypy datové rukavice. Dokud rukavice předává počítači stejný formát klíče, tak nemůže nastat kolize ve verzích prototypů rukavice vůči programu v počítači.

Klíč se následně pošle po sériové lince do počítače, kde se porovná s databází a najde ke gestu příslušný řetězec (slovo/větu). Nakonec se řetězec předá hlasovému modulu a přehraje se.



Obrázek 5.: Obecná implementace zapojení datové rukavice

5.1 Počítačový software

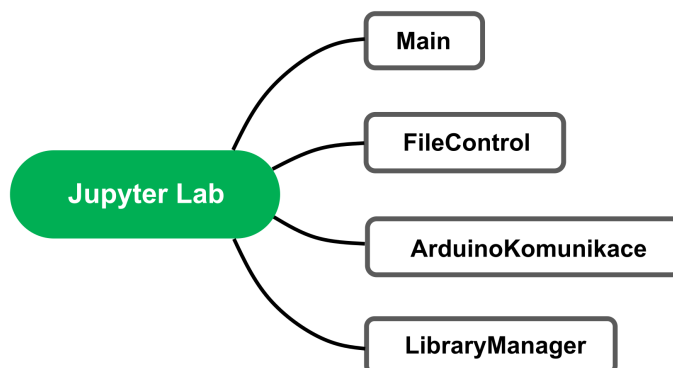
Software pro zpracování gest a ovládání syntézy řeči jsme vytvářeli v Jupyter Labu (pythonu). Software je rozdělen do 4 klíčových částí:

Main - hlavní skript, který spouští celý proces pro překlad gest na řeč

FileControl - tvorba a nastavení systémových souborů

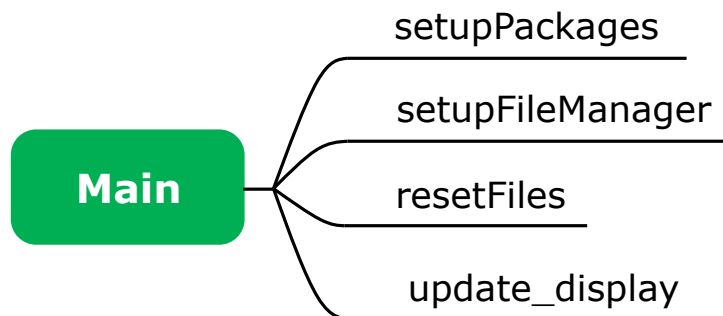
ArduinoKomunikace - vytvoření spojení (sériové linky) mezi Arduinem a počítačem

LibraryManager - tvorba, manipulace a nastavení knihoven s gesty (datábase gest a klíčů)



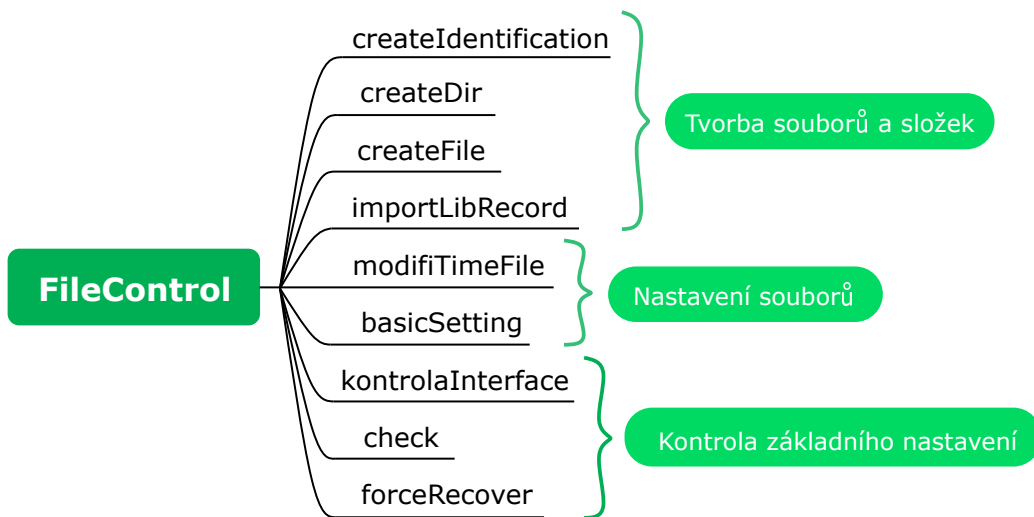
Obrázek 6.: Rozdělení hlavních souborů

Main skript je základní spouštěcí program, který zkontroluje a nastaví počítač pro práci s datovou rukavicí, vytvoří složky a základní konfiguraci souborů. Při znovu spuštění programu zkontroluje, zda jsou všechny hlavní soubory v pořádku. Najde-li chybu v souborech, tak má za úkol chybu najít a opravit. Jestli by chyba byla až moc rozsáhlá, tak spustí resetovací program, který vymaže a přepíše dosavadní soubory a složky.



Obrázek 7.: Implementované funkce pro Main skript

FileControl definuje funkce pro tvorbu a nastavení souborů a jejich kontrolu. *FileControl* je využíván *Main* skriptem pro tvorbu základního nastavení systému pro správnou funkci.



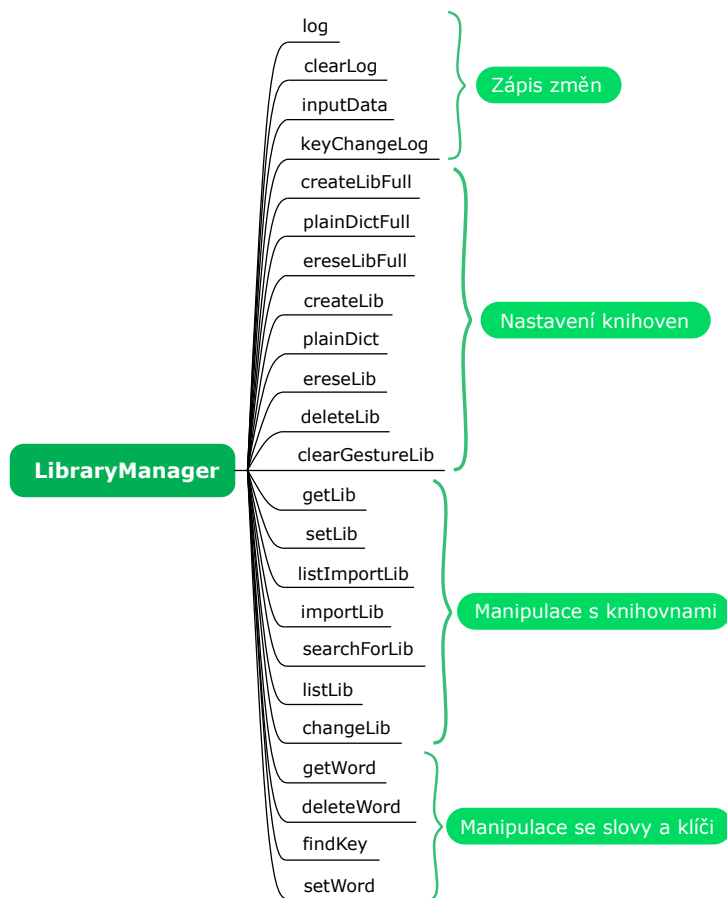
Obrázek 8.: Implementované funkce pro FileControl

ArduinoKomunikace definuje nalezení připojeného Arduina a započítí samotné komunikace s Arduinem.



Obrázek 9.: Implementované funkce pro *ArduinoKomunikace*

LibraryManager je sada funkcí, které spravují knihovny (databáze) s gesty a klíči. Základními oblastmi jsou manipulace se slovy a klíči, manipulace s knihovnami, nastavení a tvorba knihoven, zápis změn provedených s knihovnamí, klíči a slovy.



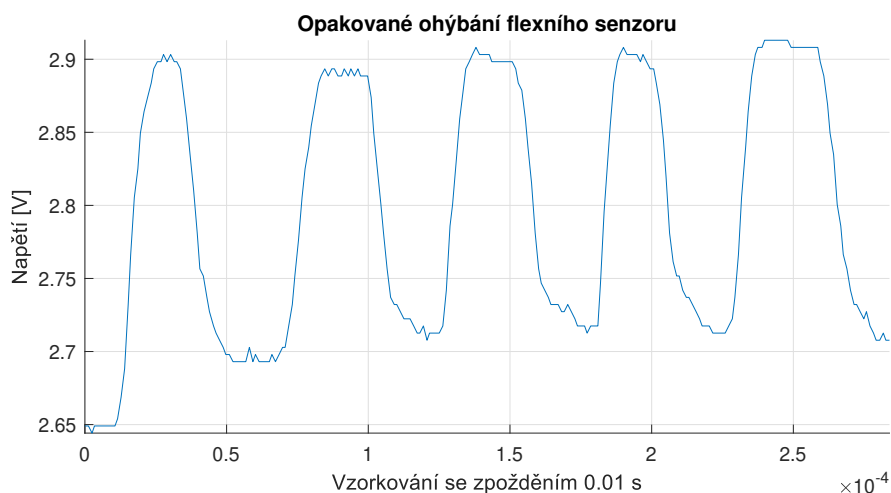
Obrázek 10.: Implementované funkce pro *LibraryManager*

V *Main* skriptu se pak uskuteční spojení s rukavicí a započne nekonečná smyčka načítání klíčů ze sériové linky. Klíče se porovnají pomocí funkcí z *LibraryManager* s knihovnami a dostáváme zpět řetězec (slovo/větu). Řetězec se předá TTS modulu a přehraje se. Aby měl výstup štábní kulturu a my měli povědomí co se na pozadí děje, tak se klíč, řetězec, stav připojení a ukončovací tlačítko vykreslí pod spuštěným skriptem. Tím můžeme kdykoliv ukončit program a vylepšovat ho. Nesmíme ale opomenout skutečnost, že nám běží nekonečná smyčka pro čtení ze sériové linky. Aktivní proces jako je nekonečná smyčka utlumí jakékoliv přerušení, které bychom chtěli udělat, jako je například aktivace ukončovacího tlačítka. Řešením je vytvořit nové vlákno, ve kterém se nekonečná smyčka spustí a v "hlavním" vlákně bude tlačítko vyčkávat na aktivaci. Po jeho aktivaci se do vlákna se smyčkou propíše aktualizace ukončovací proměnné a tím se smyčka ukončí.

Po prvním připojení prototypu jsme zjistili nemilou věc a to vysokou latenci. Ta byla způsobena pomalým čtením ze sériové linky, buffrováním a samotným během aktualizací programu. Po změně čtecí frekvence se vyřešil problém pomalého čtení a buffrování. Aktualizační algoritmus se musel optimalizovat, aby latence byla co nejmenší. Vyřešením toho jsme dostávali klíče prakticky okamžitě a stejně tak i výstupní řetězec.

5.2 1. Prototyp

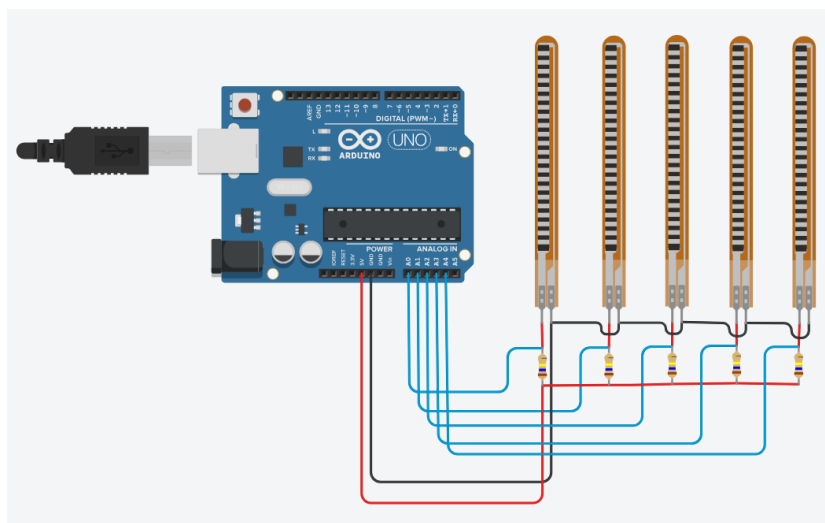
V prvním prototypu jsme využili flexních sensorů. Důvodem bylo jejich velké zastoupení v jiných výzkumech [25] a jednoduchost zapojení s Arduinem. Jak bylo řečeno, tak budeme chtít znát ohnutí prstů na dvou místech, takže využijeme kratších sensorů, aby se vešli na rukavici. Pro zapojení sensorů je nutné vytvořit napěťový dělič, abychom dostali rozumné hodnoty. Ten jsme nadimenzovali podle doporučení výrobce na $22\text{k}\Omega$.



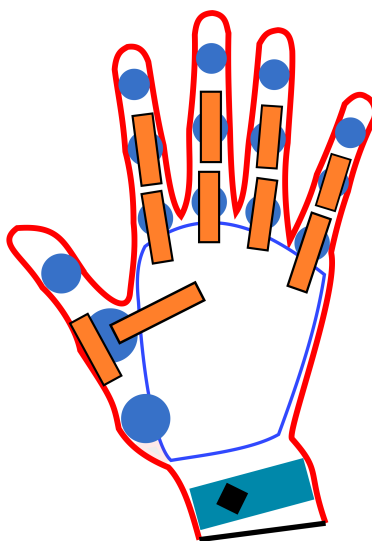
Obrázek 11.: Testování flexního senzoru

Na grafu můžete vidět průběh opakovaného ohýbání senzoru. I když napěťový rozsah není velký, pro zjištění, zda byl prst ohnutý nebo nikoliv, je to dostačující.

Celé zapojení jsme si namodelovali v prostředí **Autodesk Tinkercad**, kde je možno si napsat i kód pro testování zapojení. Hodnoty ze simulovaného prostředí a způsob zapojení byl v pořádku a byli jsme s ním spokojeni. Problémem bylo, že jsme potřebovali zapojit víc analogových signálů než má Arduino portů (Arduino Nano disponuje 8 analogovými porty). Vyřešili jsme to zapojením analogového multiplexeru, který využívá 4 digitální a 1 analogový port a rozšiřuje připojení na 16 analogových portů.



Obrázek 12.: Zapojení v Tinkercad



Obrázek 13.: Design 1. Prototypu

Při prvních testech senzorů se rukavice jevila slibně a to i přes neohrabaný design, ale po zhruba týdnu se začaly projevovat velké nedostatky. Senzory nebyly dostatečně odolné pro tento druh aplikace a docházelo k přerušení elektrického obvodu v senzorech.

Přestože se tento prototyp neosvědčil, tak nám poskytl cenné informace o nedostacích v komunikačním softwaru (již jednou zmiňovaná latence). Z těchto důvodů jsme opustili flexní senzory a přešli na nepřímé měření pomocí magnetické intenzity.

5.3 2. Prototyp

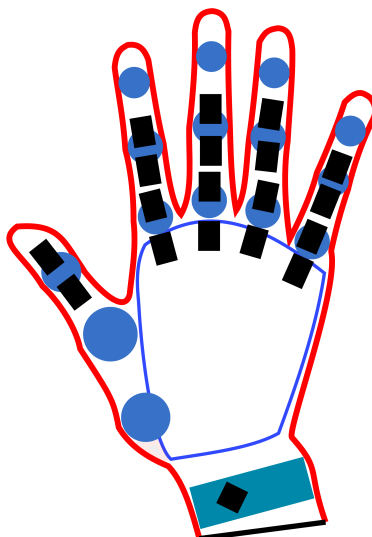
Druhý prototyp jsme navrhli na základě nepřímého měření za pomoci Hallovy sondy, které měří intenzitu magnetického pole. Pro rozmístění magnetů jsme zvolili decentralizovanou variantu a na každý prst připevnili dva malé magnety. Senzory a magnety jsou upevněny pomocí plastových držáčků vytištěných na 3D tiskárně.



Obrázek 14.: Testování Hallovy sondy

Na grafu můžete znovu vidět průběh opakovaného přiblížení senzoru k magnetu. Citlivost senzoru je o poznání větší a průběh je daleko hladší než u flexního senzoru.

Prototyp byl příjemnější na nošení a neomezoval jako první prototyp. Data získávaná ze senzorů byli uspokojivější a měli lepší rozsah (průměrně 2.4 - 3.9 Voltu). Vzhledem k tomu, že mezi senzorem a magnetem byla na rukavici obvykle malá mezera, tak to mírně snížilo napěťový rozsah, což však nebylo významné. Nevýhodou bylo připojení dalšího drátu, takže rukavice byla o něco víc obalená, ale to je spíše estetická vada. Jak již bylo řečeno dříve, tak jsme se snažili vytvořit modulární kód, jak pro počítač, tak pro Arduino. V operačním kódu pro Arduino se musel změnit pouze výpočet změny ohybu, jelikož se rozšířil monitorovaný rozsah.

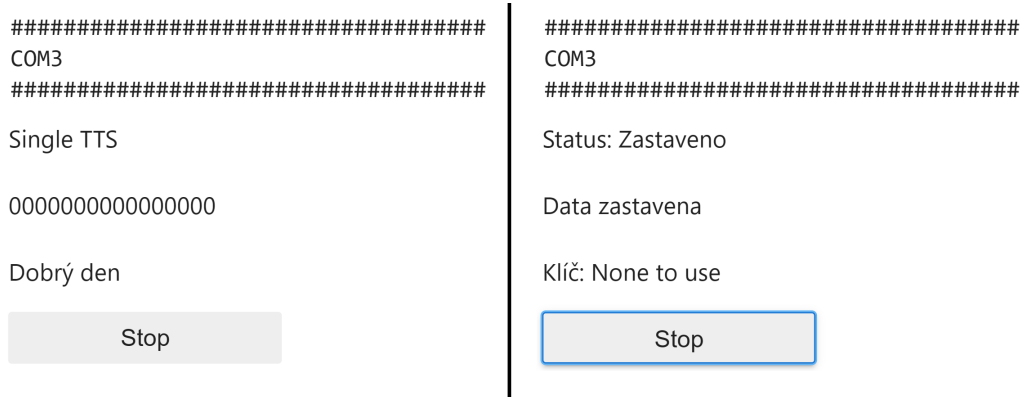


Obrázek 15.: Design 2. Prototypu

Po prvních testech byl znatelný rozdíl ve větší stabilitě získávaných dat a rozpoznávání polohy prstů. Jedna z našich obav byla možná interference magnetů s okolními senzory (senzory na vedlejších prstech). Prokázalo se však, že jelikož nejsou magnety k vedlejším sensorům kolineární, tak magnetická intenzita postranních senzorů se blížila k nule.

6 Práce s rukavicí a programem

Nyní se detailněji seznámíme s tím, jak pracovat s rukavicí a programem. Rukavice se prozatím připojuje k počítači USB kabelem, který slouží pro vytvoření datového spojení a napájení. Při připojení rukavice se v počítači musí spustit náš program pro vytvoření sériové komunikace. Není nutné definovat, ke kterému portu počítače se připojujeme; to zajišťuje program sám. Do výstupní konzole programu se zapisuje:



Obrázek 16.: Vzhled výstupní konzole (připojené | odpojené)

- Přístupový port - Jaký port se používá k připojení rukavice.
- Mód - V jakém pracovním stavu se rukavice nachází:
 1. Zastaveno - Rukavice není připojená k programu/počítači.
 2. Singel TTS - Po každém přeloženém gestu se přehraje zvuková nahrávka.
 3. Batch TTS - Gesta, která byla přeložena, se postupně shromažďují, dokud uživatel nezadá speciální gesto, které spustí přehrávání všech dosud zaznamenaných překladů gest najednou.
- Klíč - Binární klíč posílaný Arduinem.
- Řetězec - Řetězec předaný z databáze, pakliže nějaký existuje. Když ne, tak obsahuje hodnotu 'None'.
- Stop - Zastavovací tlačítko.

Po připojení rukavice k programu se provede základní kalibrace senzorů. Tato kalibrace slouží k zjištění mezních hodnot, kdy je prst v režimu flexe nebo extenze. Samotná kalibrace trvá necelých 15 sekund, z toho 6 sekund jsou věnovány udržení všech prstů ve stavu extenze a následně 6 sekund ve stavu flexe. Dokončení kalibrace spustí v programu nekonečnou smyčku pro nahrávání gest z Arduina a hledání shod v databázi.

V databázi však nemusí být obsaženy definice jakéhokoliv gesta. Definice se mohou provádět v odpojeném i připojeném stavu. Na následující ukázce můžete vidět, jak probíhá vytvoření knihovny a definice gest.

```
1 # EXAMPLE PRÁCE VYTVOŘENÍ KNIHOVNY A DEFINICE GESTA
2 import LibraryManager as LM
3 LM.createLib('HelloWorldDB.txt') # vytvoření databáze
4 LM.setLib('HelloWorldDB.txt') # nastavení programu tuto databázi pro překlad gest
5
6 LM.setWord('00000000', 'Ahoj') # definování významu gesta
7 LM.setWord('001111110', 'Světě')
8 LM.setWord('111111000', 'Toto je hlasová protéza')
```

```
Byl vytvořen slovník - HelloWorldDB.txt
Aktuální knihovna je - HelloWorldDB.txt
Přidáno nové slovo.
00000000      Ahoj
Přidáno nové slovo.
001111110     Světě
Přidáno nové slovo.
111111000     Toto je hlasová protéza
```

Obrázek 17.: Definování knihovny a významu gesta

Spuštěním těchto příkazů se provedou okamžité změny a překlady jsou pro rukavici ihned k dispozici.

My si v rámci vývoje definovali 5 základní gest a jejich definice.

```
LM.setWord('0000000000000000', 'Dobrý den')
LM.setWord('1010111100000000', 'Jmenuju se Petr Antal')
LM.setWord('1000011110000000', 'Toto je bakalářská práce na téma')
LM.setWord('1000000000000000', 'Hlasová protéza')
LM.setWord('1111111100000000', 'Hotovo')
```

Obrázek 18.: Testovací gesta

Tyto gesta jsme využívali především k testování softwaru a ladění.

7 Diskuze o rozšíření a vylepšení navržené protézy

Pro zlepšení rozpoznávání gest datovou rukavicí lze provést několik vylepšení.

V hardwarové části můžeme zlepšit konstrukci a upevnění senzorů k ruce pro stabilnější zisk dat. Obohatit data o orientaci a dynamický pohyb ruky, čímž se o kousek blíž přiblížíme k rozpoznání znakového jazyka. Můžeme také experimentovat s možností nasazení EMG senzorů pro čtení svalové činnosti na předloktí a estimovat pohyb a polohu prstů.

Software se s přidáním dynamiky gest musí také zlepšit například využitím pokročilejších algoritmů strojového učení pro robustnější interpretaci gest. Komunikaci rukavice-počítač převést na komunikaci rukavice-telefon pro dosažení úplné mobility. A pro rozšíření možné aplikace datové rukavice se může vytvořit počítačový model, tj. digitalizace gest pro jiné využití jako je virtuální realita aj.

Je zapotřebí se také nechat poučit od odborníků na znakový jazyk, jaké aspekty znakového jazyka nezanedbat a vytvořit databázi pro přibližný až přesný překlad. A stejně tak udělat uživatelské testy pro zjištění aplikovatelnosti našeho řešení a vylepšení překladu ze znakového jazyka na mluvený projev.

8 Závěr a shrnutí výsledků práce

V této práci jsme prezentovali výsledky návrhů prototypů hlasových protéz. Přesněji dva prototypy datových rukavic, které se snaží přiblížit překladu ze znakového jazyka na mluvený projev.

Průběžně jsme analyzovali a testovali oba prototypy, abychom ověřili jejich funkčnost a uživatelskou přívětivost. Výsledky ukázaly, že technologie datových rukavic má potenciál přeložit znakový jazyk do mluveného projevu a má smysl se tomuto tématu více věnovat.

Během práce jsme také diskutovali o možnostech rozšíření a vylepšení protézy pomocí pokročilých algoritmů strojového učení, integrace dalších senzorů a zlepšení ergonomie designu. Byly učiněny návrhy, jak tyto aspekty zohlednit v budoucích iteracích vývoje prototypů, a potenciální směry dalšího výzkumu.

Jsme potěšeni z dosažených výsledků a pokroku, kterého jsme v tomto projektu dosáhli. Navzdory úspěchům si uvědomujeme, že stále zbývá mnoho práce a výzkumu, než budeme moci plně uspokojit potřeby uživatelů s omezenou schopností komunikace. Věříme, že naše navržená hlasová protéza má potenciál přispět k zlepšení kvality života uživatelů a otevřít nové možnosti pro návrat hlasu. Jsme vděční za příležitost pracovat na tomto projektu a doufáme, že naše úsilí pomůže otevřít cestu pro další výzkum a vývoj v této oblasti.

Reference

- [1] Hung-Chi Chen, Karen Kim Evans, Christopher Salgado, and Samir Mardini. Methods of voice reconstruction. *Seminars in Plastic Surgery*, 24(02):227–232, 2010.
- [2] Jan G. Švec. Studium mechanicko-akustických vlastností zdroje lidského hlasu. Master’s thesis, Přírodovědecká fakulta, Univerzita Palackého v Olomouci, 1996.
- [3] Velký lékařský slovník. <https://lekarske.slovniky.cz/>, 2023. [Online].
- [4] Wikiskripta. <https://www.wikiskripta.eu/>, 2023. [Online].
- [5] Hanjun Liu and Manwa L. Ng. Electrolarynx in voice rehabilitation. *Auris Nasus Larynx*, 34(3):327–332, 2007.
- [6] Rachel Kaye, Christopher G Tang, and Catherine F Sinclair. The electrolarynx: voice restoration after total laryngectomy. *Medical Devices: Evidence and Research*, Volume 10:133–140, 2017.
- [7] Eliška Venclová. Využití náhradních hlasových mechanismů při rehabilitaci osob po totální laryngektomii. Master’s thesis, Hradec Králové: Pedagogická fakulta Univerzity Hradec Králové, 2018.
- [8] Farzaneh Ahmadi, Farzad Noorian, Daniel Novakovic, and André van Schaik. A pneumatic bionic voice prosthesis—pre-clinical trials of controlling the voice onset and offset. *PLOS ONE*, 13(2):e0192257, 2018.
- [9] Itzhak Brook and Joseph F Goodman. Tracheoesophageal voice prosthesis use and maintenance in laryngectomees. *International Archives of Otorhinolaryngology*, 24(04):e535–e538, 2020.
- [10] Burkhard Kramp and Steffen Dommerich. Tracheostomy cannulas and voice prosthesis. *GMS Current Topics in Otorhinolaryngology - Head and Neck Surgery; 8:Doc05; ISSN 1865-1011*, 2009.
- [11] Zuzana Pavliňáková and Pavlína Zarubová. Základy znakového jazyka. https://repozitar.cz/repo/39430/Zaklady_znakoveho_jazyka.pdf?info, 2019. [Online].
- [12] Mohamed Aktham Ahmed, Bilal Bahaa Zaidan, Aws Alaa Zaidan, Mahmood Maher Salih, and Muhammad Modi bin Lakulu. A review on systems-based sensory gloves for sign language recognition state of the art between 2007 and 2017. *Sensors*, 18(7):2208, 2018.

- [13] Zinah Raad Saeed, Zurinahni Binti Zainol, B. B. Zaidan, and A. H. Alamoodi. A systematic review on systems-based sensory gloves for sign language pattern recognition: An update from 2017 to 2022. *IEEE Access*, 10:123358–123377, 2022.
- [14] Gheyath Othman and Diyar Qader Zeebaree. The applications of discrete wavelet transform in image processing: A review. *Journal of Soft Computing and Data Mining*, 1(2):31–43, 2020.
- [15] Richard O. Duda and Peter E. Hart. Use of the hough transformation to detect lines and curves in pictures. *Commun. ACM*, 15(1):11–15, jan 1972.
- [16] Matti PietikÄÖinen. Local binary patterns. *Scholarpedia*, 5(3):9775, 2010.
- [17] Alex Becker. Kalman filter. <https://www.kalmanfilter.net/default.aspx>, 2023. [Online].
- [18] Lotfi Zadeh. Fuzzy logic. *Scholarpedia*, 3(3):1766, 2008.
- [19] Adnan Rashid and Osman Hasan. Wearable technologies for hand joints monitoring for rehabilitation: A survey. *Microelectronics Journal*, 88:173–183, 2019.
- [20] Anna Findejsova. Rukavice se senzory ohybu. Bachelor’s thesis, Brno, Vysoke ucenı technicke v Brne, Fakulta elektrotechniky a komunikacnıch technologiı, 2017.
- [21] Giovanni Saggio, Francesco Riillo, Laura Sbernini, and Lucia Rita Quitadamo. Resistive flex sensors: a survey. *Smart Materials and Structures*, 25(1):013001, 2015.
- [22] Naseer Sabri, S A Aljunid, M S Salim, R B Ahmad, and R Kamaruddin. Toward optical sensors: Review and applications. *Journal of Physics: Conference Series*, 423(1):012064, apr 2013.
- [23] Uriel Martinez-Hernandez. Tactile sensors. *Scholarpedia*, 10(4):32398, 2015.
- [24] J.E. Lenz. A review of magnetic sensors. *Proceedings of the IEEE*, 78(6):973–989, June 1990.
- [25] P Weber, E Rueckert, R Calandra, J Peters, and P Beckerle. A low-cost sensor glove with vibrotactile feedback and multiple finger joint and hand motion sensing for human-robot interaction. 2016.