

## 1. ÚVOD

Technológie sa posúvajú napred, vylepšujú sa, zdokonaľujú a nachádzajú širšie využitie v každodennom živote. Na rad prichádza aj ovládanie hlasom, presnejšie hovoreným slovom. Používa sa jednoducho, nezanepoždňuje vaše ruky, môžete ovládať zariadenie aj na diaľku - dokiaľ vás počuje. Rozšírenie sa začalo v domácnostiach používať na ovládanie elektroniky v dome, svetlá, ovládanie smartfónov alebo vyhľadávanie informácií na internete. Preto som sa rozhodol využiť ho na ovládanie modelu auta. Mój projekt nebude možné použiť na reálnej ceste, keďže je to len koncept ktorým chcem ukázať využitie ovládania hlasom, pre využitie ovládania hlasom na pohyb auta by bolo potrebných mnoho senzorov, kamier a celková monotónnosť auta.

## **2. CIELE PRÁCE**

Mojím hlavným cieľom je zostaviť model auta ktorý bude ovládaný pomocou hlasových príkazov, čo zahŕňa zostrojenie podvozku auta, zapojenie programovateľnej dosky arduino aby bolo možné podvozok ovládať, zapojenie shieldu Easy VR 3.0, ktorý slúži na rozpoznávanie reči na dosku arduino, naprogramovanie pohybu modelu pomocou dosky arduino a zakomponovanie ovládania modelu hlasom. Model bude zachytávať zvuky z okolia, a keď zaznamená hovorené slovo, ktoré je nastavené ako príkaz, spní ho.

### 3. Teoretická časť

#### 3. 1. Rozpoznávanie reči

Automatické rozpoznávanie reči (ASR) alebo speech to text (STT) umožňujú rozpoznávanie a preklad hovoreného jazyka do textu prostredníctvom počítačovej techniky. Zahŕňa vedomosti a výskum v oblasti lingvistiky, informatiky a elektrotechniky.

Niektoré systémy na rozpoznávanie reči vyžadujú "výcvik" (tiež nazývaný "zápis"), v ktorom jednotlivec hovorí do systému text alebo jednotlivé slová. Systém analyzuje konkrétny hlas osoby a používa ju na jemné vyladenie rozpoznávania reči tejto osoby, čo vedie k zvýšeniu presnosti. Systémy, ktoré nepoužívajú školenie, sa nazývajú systémy "nezávislých od hovorcov" ("speaker independent") . Systémy, ktoré používajú výcvik, sa nazývajú "závislé od hovorcov" ("speaker dependent").

Aplikácie na rozpoznávanie reči zahŕňajú hlasové používateľské rozhrania, ako napríklad hlasové vytáčanie, ovládanie spotrebiča, vyhľadávanie, jednoduché zadávanie údajov, spracovanie textu (zvyčajne nazývaný priamy hlasový vstup).

Pojem rozpoznávanie hlasu alebo identifikácia hovorca odkazuje na identifikáciu hovorca, a nie na to, čo hovorí. Rozpoznávanie hovorca môže zjednodušiť úlohu prekladu reči v systémoch, ktoré boli vycvičené na základe hlasu konkrétnej osoby, alebo ho možno použiť na overovanie totožnosti hovorca ako súčasť bezpečnostného procesu.

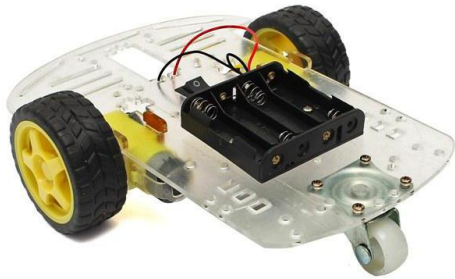
Z hľadiska technológie má rozpoznávanie reči dlhú históriu s niekoľkými vlnami významných inovácií. V poslednej dobe oblasť profitovala z pokroku v hlbokom vzdelávaní a veľkých údajoch. Pokrok je dokázaný nielen nárastom akademických publikácií publikovaných v danej oblasti, ale dôležitejšie je, že celosvetový priemysel prijal rôzne metódy hlbokého vzdelávania pri navrhovaní a zavádzaní systémov rozpoznávania reči. Tieto spoločnosti v oblasti reči zahŕňajú Google, Microsoft, IBM, Baidu, Apple, Amazon, Nuance, GoVivace, SoundHound, iFLYTEK, z ktorých mnohé zverejnili hlavnú technológiu vo svojich systémoch rozpoznávania reči ako založené na hlbokom vzdelávaní.

## 4. Využitie materiály

### 4. 1. Hardware

#### 4. 1. 1. 3-kolesový podvozok

Na túto prácu som využil 3-kolesový, 2-motorový podvozok s voľne otáčavým predným kolesom.



Obrázok 1 - podvozok

#### 4. 1. 2. Arduino

Ako základnú ovládaciu jednotku som využil programovateľnú dosku Arduino Uno, ktorá uchováva celý kód vo svojej pamäti a riadi podľa neho všetky zapojené senzory a nadstavby.



Obrázok 2 - Arduino UNO

#### 4. 1. 3. Ultrasonický senzor

Modul pre arduino ktorý sníma vzdialenosť objektu pred sebou pomocou ultrasonických vln, pozostáva z prijímača a vysielača a meria vzdialenosť v rozsahu 3 - 400cm.



Obrázok 3 - Ultrasonický senzor

#### 4. 1. 4. EasyVR 3.0

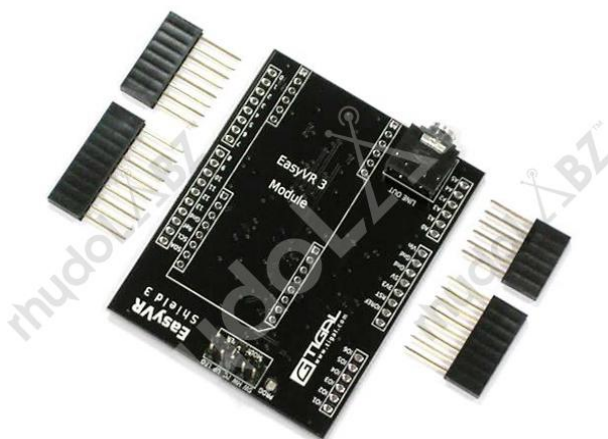
Čip EasyVR 3.0 slúži na získanie zvuku z okolia pomocou pripojeného mikrofónu a následné rozpoznanie príkazu, či už vstavaného, alebo naučeného, pretvorenie hovoreného príkazu na elektronický signál, a taktiež na vytvorenie zvukovej odozvy, ktorá je vopred nahraná na dosku. Podporuje 64 používateľsky definovaných príkazov závisiacich od hovorcov (Speaker Dependent), ktoré môžu byť vyškolené v ľubovoľnom jazyku rozdelené na maximálne 16 skupín. Taktiež ponúka 26 vstavaných príkazov nezávislých od hovorca (Speaker Independent) pre základné ovládacie prvky pripravené na prevádzku v nasledujúcich jazykoch - angličtina (US), taliančina, nemčina, francúzština, španielčina, japončina.



Obrázok 4 - EasyVR 3.0

#### 4. 1. 5. EasyVR shield

je doska určená na zjednodušenie práce s doskou EasyVR 3.0, ktorá sa napája na shield, a ten sa následne dá jednoducho napojiť priamo na Arduino Uno. Zjednodušuje celkové použitie a programovanie dosky EasyVR 3.0.



Obrázok 5 - EasyVR shield

## 4. 2. Software

### 4. 2. 1. Arduino IDE

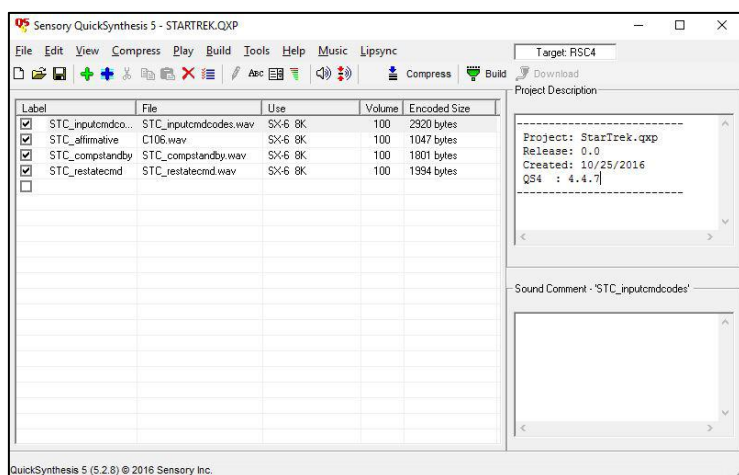
Program ktorý slúži na vytváranie kódu a komunikovanie s doskou Arduino Uno. Podporuje programovacie jazyky C a C++, dodáva softvérovú knižnicu funkcií, ktoré uľahčujú písanie najzákladnejších operácií s hardvérom. Obsahuje editor kódu s bežnými vlastnosťami ako farebné označovanie syntaxu, automatické zarovnávanie a párovanie zátvoriek. Je schopné program skompilovať a nahráť do Arduina jedným kliknutím tlačidla. Kód sa skladá z 2 častí, setup() a loop(). Časť setup() sa spúšťa len raz na začiatku programu, používa sa na nastavenie parametrov, časť loop() je periodicky spúšťaná, prebiehajú v nej procesy ovládania, snímania a iné.

### 4. 2. 2. Balabolka

Program vytvorený na prevádzanie textu na reč. Text môžete napísať v programe alebo nahráť zo súboru, je možné vybrať si z viacerých hlasov s rôznou výslovnosťou a taktiež zmeniť parametre hlasu ako sú rýchlosť a výšku hlasu.

### 4. 2. 3. Sensory QuickSynthetys 5

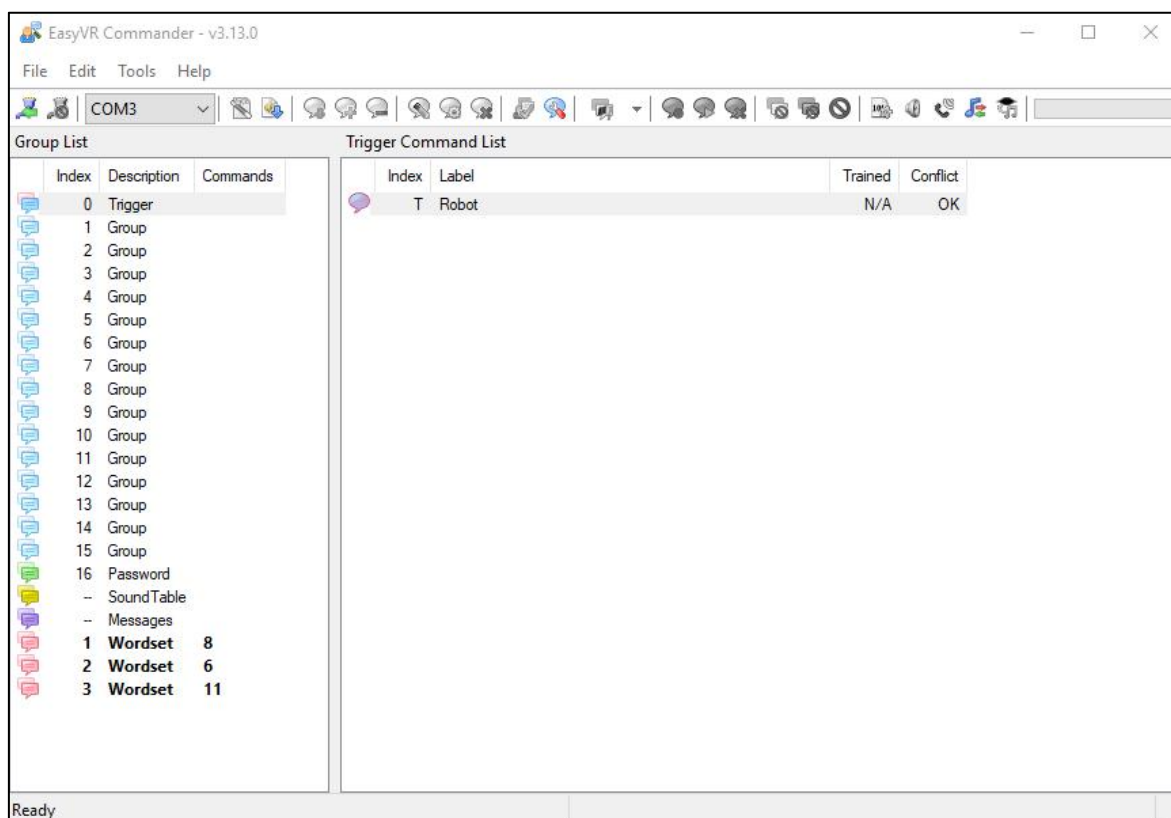
Vytvára zložku audiosúborov, ktorú je možné najhrať na dosku EasyVR a použiť ako výstup. Audio súbory, obsahujúce zvuk alebo hovorené slová/vety, s koncovkou .wav sa načítajú do programu, ktorý ich komprimuje a uloží do súboru použiteľného doskou EasyVR.



Obrázok 6 - Sensory QuickSynthetys 5

#### 4. 2. 4. EasyVR commander

Softvér slúži na komunikáciu s doskou EasyVR, nahrávanie audio súborov (vytvorené programom Sensory QuickSynthetys), vytváranie a učenie dosky hovorové príkazy a generovanie základného kódu pre arduino, ktorým dokáže arduino komunikovať a ovládať dosku EasyVR. V kóde sú už zahrnuté naučené príkazy a audio súbory ktoré je možné použiť ako výstup (audio odpoveď, výzvu na zadanie príkazu atď.). V programe je niekoľko zložiek, prvou je “Trigger”, v ktorej je už nahratý príkaz Robot, a jeden ďalší príkaz sa dá nahrať, tieto slúžia ako spúšťače príkazy ktorými arduino začne prijímať ďalšie príkazy. Ďalej sú tu zložky “Group” 1 - 15, do ktorých sa nahrávajú vlastné príkazy, v programe sa využívajú jednotlivo. Zložka “password” je určená pre ukladanie hesiel. V zložkách “Wordset” 1 - 3 sú už vopred nahrané príkazy a v zložke SoundTable sú uložené všetky zvukové odozvy vytvorené programom Sensory QuickSynthetys 5 a jeden vstavaný zvuk - pípnutie.



Obrázok 7 - EasyVR Commander

## **Zhrnutie**

V tejto práci som využil svoje znalosti v elektronike a programovaní, a taktiež som nadobudol nové, na zistenie informácií som využil najmä internet. Všetky problémy ktoré nastali som vyriešil, a preto túto prácu považujem za úspešnú, taktiež všetky ciele práce som splnil úspešne.