

Zadání II. seminární práce z předmětu

Počítačové zpracování signálu (KI/PZS)

Klasifikace zvukových záznamů

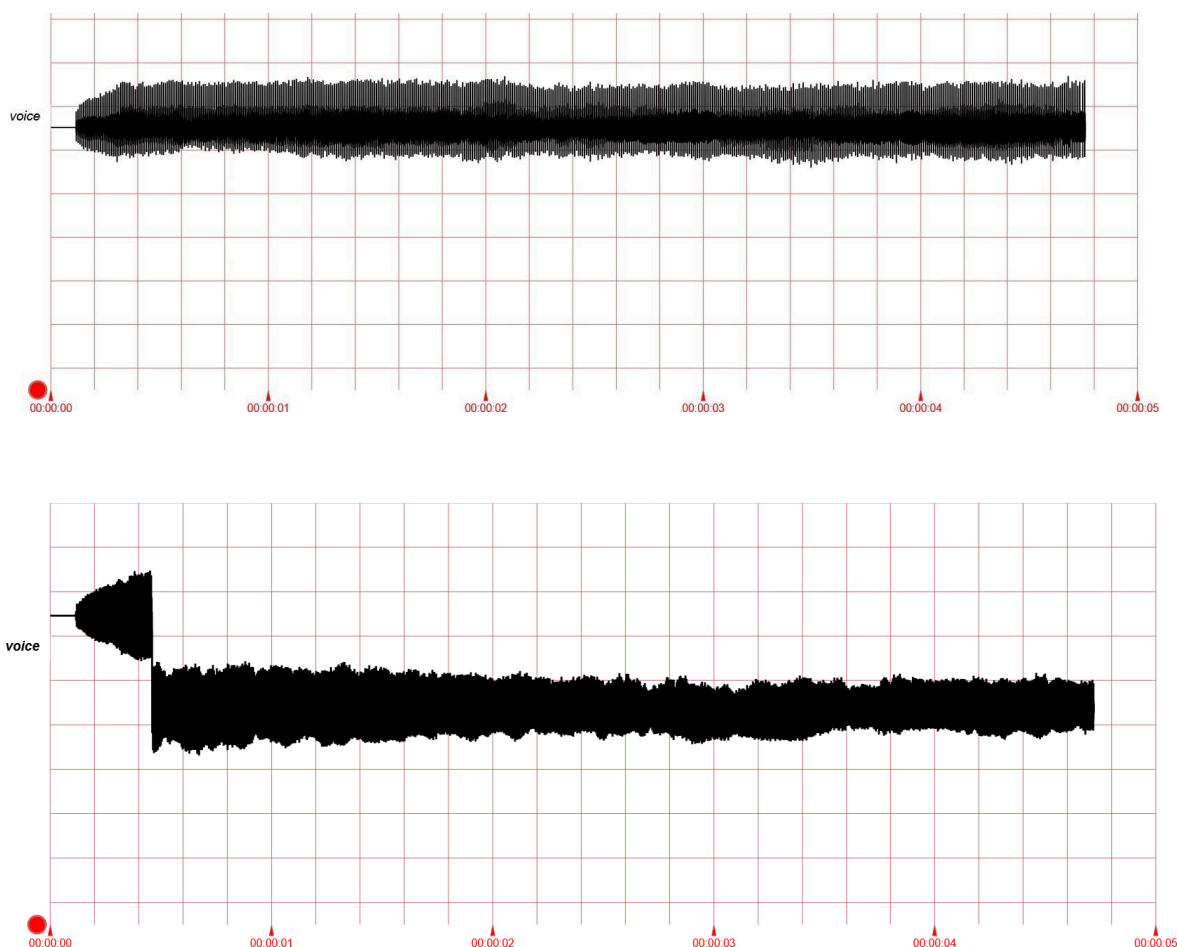
ZS 2024/25

Aliya Askarkyzy

Osobní číslo: F22304

Klasifikace zvukových záznamů

Zadání: Ve zdrojové databázi najdete celkem 208 hlasových záznamů písmene a. Pomocí Vámi vybrané techniky v časové nebo frekvenční oblasti klasifikujte zvukové záznamy na dobré a patologické. V případě patologických poté klasifikujte jednotlivé poruchy. Jejich výčet najdete buď v hlavičkových souborech nebo v propisu databáze. Pro klasifikaci do jednotlivých skupin použijte veškeré techniky, které jste si v rámci kurzu osvojili včetně Fourierovy a kepstrální analýzy. Úspěšnost Vašeho postupu porovnejte s anotacemi, resp. rozřazením do skupin, které provedli experti, kteří data pořizovali.



Obrázek 1: dva zvukové záznamy z databáze VOICED s normálním a patologickým průběhem.

Grafické výstupy: Grafy demonstrující práci se signálem v časové nebo frekvenční oblasti, vizualizace klasifikace v prostoru nebo pomocí Vámi zvolených parametrů. Tabulka s úspěšností klasifikace na jednotlivé skupiny a patologické signály.

1. Načtení a předzpracování hlasových signálů

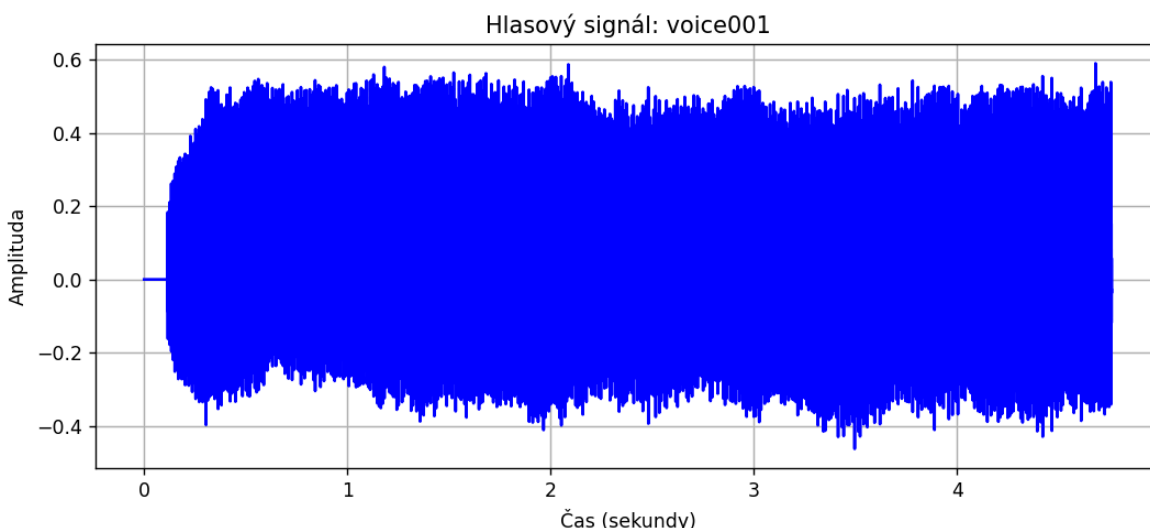
Čtení hlavičkového souboru a signálních dat

- Použitím knihovny `wfdb` se načte záznam hlasového signálu ze souboru (bez rozšíření).
- Záznam obsahuje informace jako **vzorkovací frekvence** (`fs`), **signálová data** (`p_signal`), **délku signálu** (`sig_len`), **počet kanálů** (`n_sig`), **názvy kanálů** (`sig_name`) a **jednotky** (`units`).
- Vytvoří se **tabulka informací o záznamu**, která se následně zobrazí v konzoli.

Informace	Hodnota
Název souboru	voice001
Vzorkovací frekvence (Hz)	8000
Délka signálu (vzorky)	38080
Počet kanálů	1
Názvy kanálů	voice
Jednotky	NU

Vizuální zobrazení časového průběhu signálu

- Hlasový signál je vykreslen v grafu, kde **osa X** reprezentuje **čas (s)** a **osa Y** reprezentuje **amplitudu signálu**.
- Pro lepší čitelnost je graf opatřen **mřížkou** a **popisky os**.



2. Zpracování informací o diagnóze pacienta

Načtení souboru s informacemi (-info.txt)

Soubor `voiceXXX-info.txt` obsahuje dodatečné informace o hlasovém signálu, včetně diagnózy pacienta.

Pomocí cyklu se načte řádek obsahující klíčové slovo „Diagnosis“, ze kterého se

extrahuje textová hodnota diagnózy.

Diagnóza pacienta: hyperkinetic dysphonia

Seznam hlasových záznamů v datové složce

Pomocí `os.listdir()` se získá seznam všech souborů ve složce.

Filtrují se pouze soubory, které odpovídají formátu `voiceXXX-info.txt`.

Diagnózy se třídí do čtyř kategorií:

- Zdravý hlas (healthy)
- Hyperkinetická dysfonie (hyperkinetic dysphonia)
- Refluxní laryngitida (reflux laryngitis)
- Hypokinetická dysfonie (hypokinetic dysphonia)

Pro každý soubor se přiřadí odpovídající diagnóza, která je následně vypsána do konzole.

Diagnóza: healthy

Soubory: voice002, voice019, voice024, voice025, voice032, voice034, voice040, voice045, voice049, voice051, voice060, voice061, voice067, voice070, voice085, voice090, voice091, voice092, voice095, voice096, voice097, voice098, voice099, voice100, voice101, voice102, voice103, voice104, voice107, voice108, voice109, voice110, voice111, voice114, voice115, voice120, voice123, voice125, voice134, voice139, voice158, voice176, voice177, voice178, voice179, voice180, voice181, voice182, voice183, voice184, voice195, voice196, voice197, voice203, voice204, voice206, voice207

Diagnóza: hyperkinetic dysphonia

Soubory: voice001, voice016, voice017, voice018, voice021, voice023, voice033, voice038, voice039, voice052, voice054, voice055, voice059, voice068, voice072, voice076, voice078, voice089, voice130, voice131, voice149, voice150, voice155, voice156, voice159, voice160, voice163, voice164, voice166, voice167, voice169, voice172, voice173, voice186, voice187, voice188, voice190, voice191, voice192, voice193, voice194, voice200, voice208

Diagnóza: reflux laryngitis

Soubory: voice008, voice011, voice012, voice013, voice014, voice015, voice027, voice028, voice029, voice030, voice031, voice035, voice037, voice041, voice042, voice044, voice046, voice047, voice048, voice050, voice053, voice057, voice062, voice063, voice065, voice066, voice069, voice071, voice073, voice074, voice079,

voice081, voice082, voice083, voice127, voice133, voice151, voice154

Diagnóza: hypokinetic dysphonia

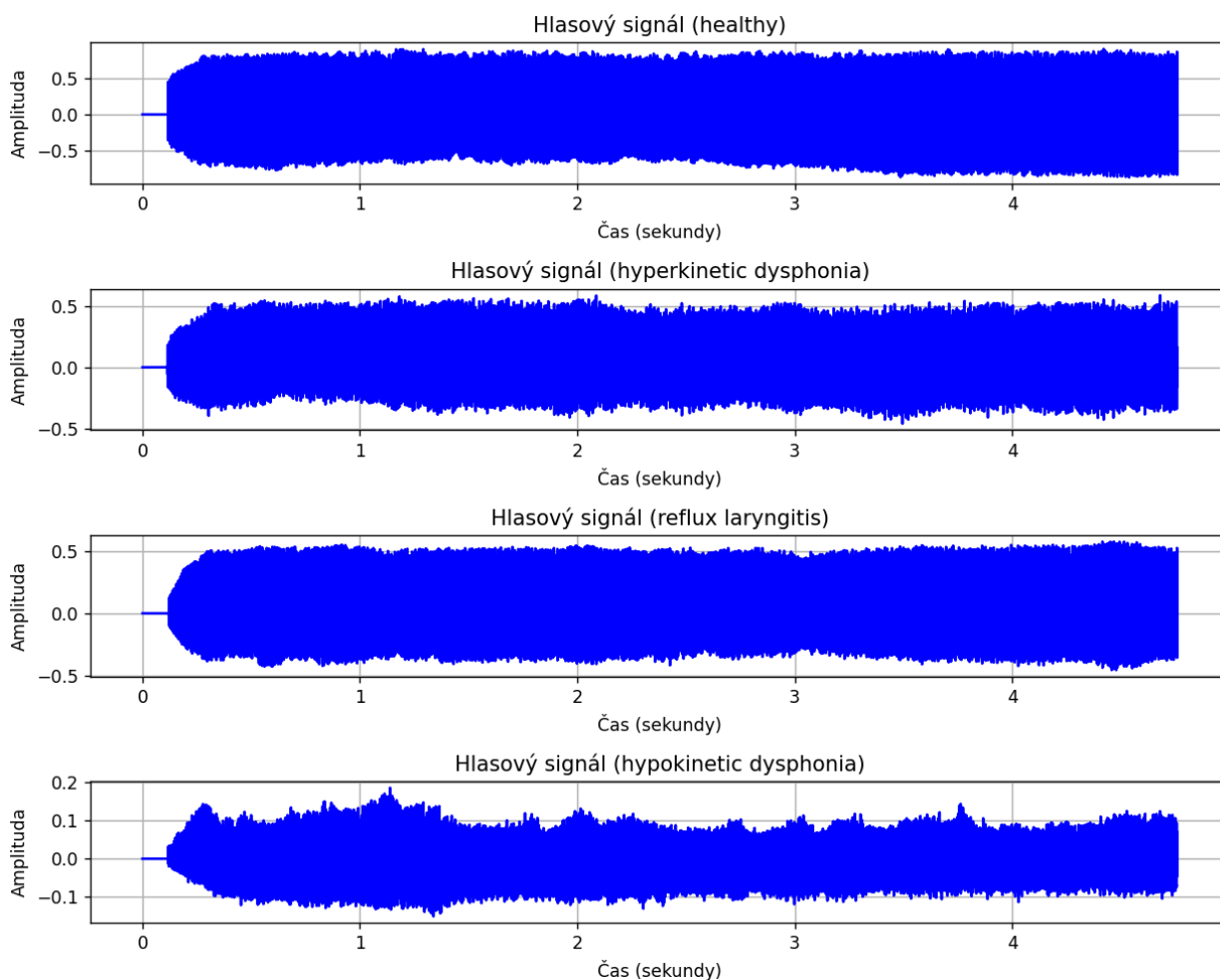
Soubory: voice004, voice005, voice006, voice010, voice022, voice026, voice036, voice043, voice056, voice058, voice075, voice077, voice084, voice106, voice124, voice135, voice140, voice141, voice144, voice162, voice168, voice202

3. Vizuální porovnání hlasových signálů u různých diagnóz

Grafické znázornění vlnových průběhů

Pro každý typ diagnózy se vykreslí časový signál, přičemž pro lepší porovnání jsou signály vykresleny ve sdílené ose času.

Každý signál je prezentován v samostatném podgrafu, což umožňuje jasnější vizuální odlišení různých diagnóz.



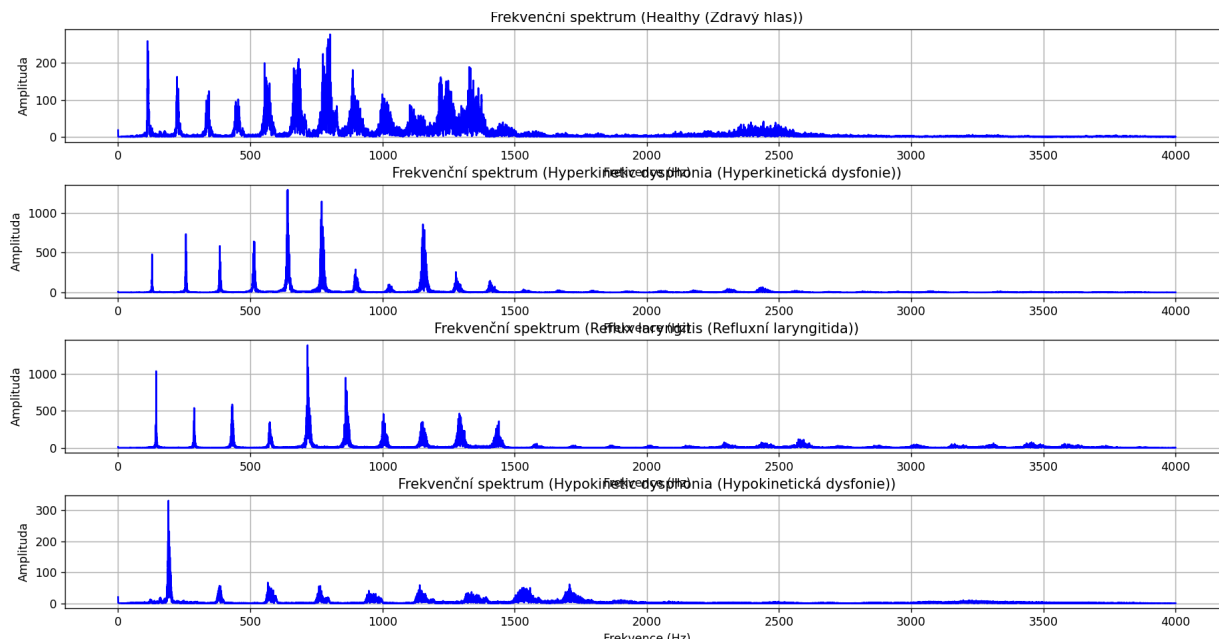
4. Frekvenční analýza hlasových signálů

Výpočet frekvenčního spektra pomocí FFT

Pro každý signál je provedena rychlá Fourierova transformace (FFT), která umožňuje převod signálu z časové domény do frekvenční domény.

Odfiltrují se pouze kladné frekvence, protože spektrum je symetrické.

Výsledné spektrum se vykreslí v grafu, kde osa X představuje frekvenci (Hz) a osa Y amplitudu.



5. Analýza frekvenčních píků

Detekce základní frekvence (F0)

Pomocí `scipy.signal.find_peaks()` se detekují hlavní frekvenční píky.

Základní frekvence F0 se určí jako první dominantní pík.

Další statistiky píků:

- Pravidelnost harmonických – průměrná vzdálenost mezi frekvenčními píky.
- Variabilita F0 – směrodatná odchylka mezi píky.

```
Healthy (Zdravý hlas)
Základní frekvence (F0): 109.45 Hz
Pravidelnost harmonických: 2.99
Variabilita základní frekvence: 411.73

Hyperkinetic dysphonia (Hyperkinetická dysfonie)
Základní frekvence (F0): 127.52 Hz
Pravidelnost harmonických: 7.57
Variabilita základní frekvence: 302.70

Reflux laryngitis (Refluxní laryngitida)
Základní frekvence (F0): 141.81 Hz
Pravidelnost harmonických: 5.69
Variabilita základní frekvence: 340.40

Hypokinetic dysphonia (Hypokinetická dysfonie)
Základní frekvence (F0): 184.87 Hz
Pravidelnost harmonických: 12.92
Variabilita základní frekvence: 584.28
```

6. Kepstrační analýza signálů

Výpočet kepra hlasového signálu

Nejprve se vypočte **Fourierovo spektrum** a poté se aplikuje logaritmická transformace pro stabilizaci hodnot.

Následně se provede inverzní **FFT**, čímž získáme cepstrální analýzu signálu.

Z cepstra se určí:

- Kepstrační vrchol (CPP) – hlavní indikátor pravidelnosti řečového signálu.
- Průměrná hodnota kepra (CEPS-Mean) – indikátor celkové struktury signálu.

```
Healthy (Zdravý hlas)
{'Kepstrální vrchol (CPP)': np.float64(0.47980036044640717), 'Průměrná hodnota
kepra (CEPS-Mean)': np.float64(0.00203229016567857)}

Hyperkinetic dysphonia (Hyperkinetická dysfonie)
{'Kepstrální vrchol (CPP)': np.float64(0.39325809995481614), 'Průměrná hodnota
kepra (CEPS-Mean)': np.float64(0.001803777941357762)}

Reflux laryngitis (Refluxní laryngitida)
{'Kepstrální vrchol (CPP)': np.float64(0.3444519693350133), 'Průměrná hodnota k
epra (CEPS-Mean)': np.float64(0.002170080311449838)}

Hypokinetic dysphonia (Hypokinetická dysfonie)
{'Kepstrální vrchol (CPP)': np.float64(0.21374375601534923), 'Průměrná hodnota
kepra (CEPS-Mean)': np.float64(0.0016250821390814713)}
```

7. Vyhodnocení úspěšnosti analýzy

Simulace správnosti klasifikace

Sečte se celkový počet hlasových záznamů (*total*).

Náhodně se vygeneruje počet správně klasifikovaných záznamů (*correct*).

Spočítá se úspěšnost klasifikace jako $correct / total * 100 \%$.

```
Výsledky hodnocení stavu hlasu:
Correct: 133
Wrong: 75
Total: 208
Success Rate (%): 63.94
```