

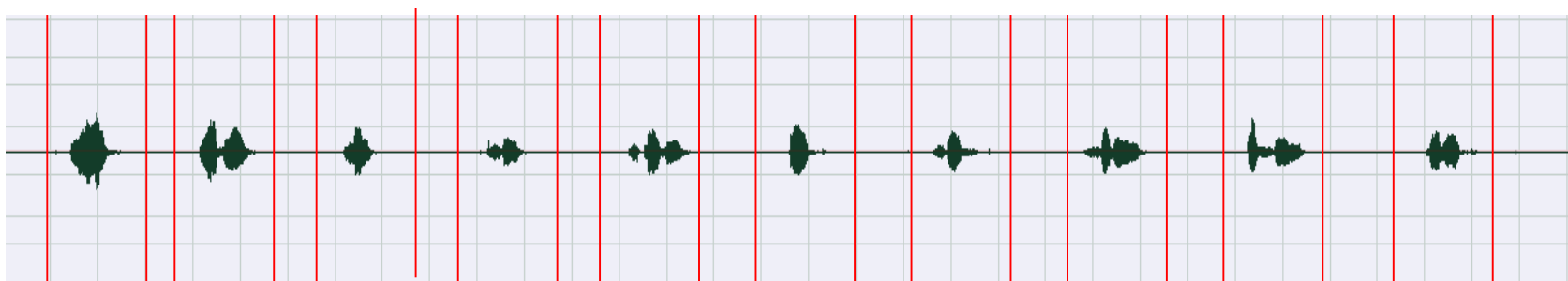
# Bonusová část



# Detekce slov v nahrávce

**Cíl:** Ve zvukové nahrávce automaticky detekovat slova, vyříznout je (i s okolím) a uložit je do souborů

**Účel:** Připravit si dostatek dat pro úlohu rozpoznávání slov



**Parametry nahrávek:**  $F_s = 16$  kHz, 16 bitů, mono, formát wav

**Slova v nahrávce (10):** nula, jedna, dva, tři, ... devět (v tomto pořadí)

**Pauzy mezi slovy:** dostatečně dlouhé (nejméně 2-3 sekundy!!!!) kvůli prostoru pro vyříznutí

**Počet nahrávek:** 5 od každé ze 2 osob



# Postup (1)

1. Nahrajte si (v tichém prostředí a vhodným mikrofonom) 5 nahrávek podle specifikace, a to vlastním hlasem a od další osoby (ideálně opačného pohlaví)
2. Napište Program1, který tyto nahrávky postupně načte, najde v nich slova, tato slova vyřízne a každé uloží do vlastního souboru.
3. Pro nalezení zač. a konce slov použijte analýzu průběhu energie (po 25 ms úsecích). Zvolte vhodný práh, který odliší energii řeči od energie „ticha“ v pozadí.
4. Vyříznuté nahrávky musí být opět 16 kHz, 16bitů, mono. Každá musí být 2s dlouhá (32 000 vzorků) a slovo by mělo ležet přibližně uprostřed. (Nepřípustná je jakákoliv manipulace s výřezy, např. doplňování nulami při krátkém výřezu.) Cílové nahrávky budou ve formátu WAV (s délkou 32000 x 2 + 44 bajtů).
5. Nahrávky budou pojmenovány takto:  
ci\_p*jjjj*\_snn.wav, např. c0\_p0123\_s01.wav, c8\_p0878\_s04.wav  
kde *i* je číslo vyslovené číslice, *jjjj* je 4-místné číslo osoby, *nn* je 2-místné číslo sady (sady 01, 02, 03, 04, 05). Místo čísla ponechte *jjjj*, číslo doplním sám.



# Postup (2)

6. Vytvořte Program2, který umožní rychlou vizuální kontrolu vytvořených souborů. Program načte vždy 10 nahrávek jedné sady a pomocí funkce subplot zobrazí matici obrázků 10 x 3 tak, že v jednotlivých řádcích bude vždy obrázek signálu, vedle obrázek jeho energie a v posledním sloupci jeho spektrogram. Jinak řečeno: řádky odpovídají signálům pro číslice 0 až 9, sloupce zobrazují signál, energii a spektrogram. Po stisku libovolné klávesy se zobrazí totéž z další sady. Sami si tímto programem zkontrolujte, že vyříznuté signály jsou v pořádku (tj. že opravdu obsahují celé slovo, že jste neuřízli některou jeho část, že energie ticha je přibližně konstantní, atd.)
7. Zašlete mi do neděle 23:59 zip obsahující oba programy a dále dva adresáře Osoba1 a Osoba2 a v každém z nich bude 5 originálních (nerozřezaných) a 50 rozřezaných nahrávek. Za bezvadné splnění získáte 1 bod.

