

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

Fakulta informačních technologií

PŘEDMĚT - SIGNÁLY A SYSTÉMY

2020/2021

**Analýza vlivu rúšky na reč**

# Úvod

V tomto projekte sa zameriavam na analýzu vplyvu rúška na reč. Projekt je celý vypracovaný v MATLABE. Študijné zdroje sú vo všeobecnosti v tomto projekte z nasledovných zdrojov:

- 1) MATLAB fórum
- 2) Študijná etapa ku projektu
- 3) Prezentácie prednášok
- 4) Internetové články, výskumný materiál zameriavajúci sa na tematiku projektu
  - bližšie uvedenie zdrojov sú na potrebných miestach **v protokole** a v súbore riešenia **solution.m**

**Ďakujem za pozornosť a tak teda tu sú moje výsledky skúmania... :)**

## 1. Tabuľka s informáciami o nahraných tónoch

	DĚKA VIET	
NÁZVY SÚBOROV	V SEKUNDÁCH	VO VZORKOCH
maskoff_tone.wav	4.62	74000
maskon_tone.wav	4.07	65120

## 2. Tabuľka s informáciami o nahraných vetách

	DĚKA VIET	
NÁZVY SÚBOROV	V SEKUNDÁCH	VO VZORKOCH
maskoff_sentence.wav	2.29	36688
maskon_sentence.wav	2.11	33704

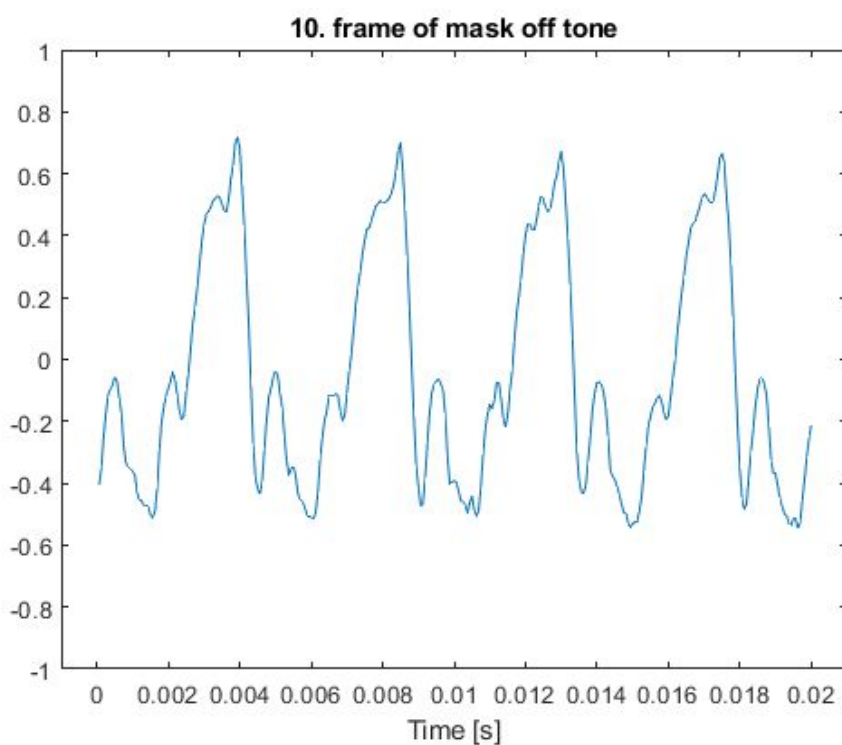
3. Vzorec pro **výpočet velikosti rámce vo vzorkách**, graf dvoch zvolených rámcov (**s rúškou, bez rúšky, prvý a druhý graf v jednom**)

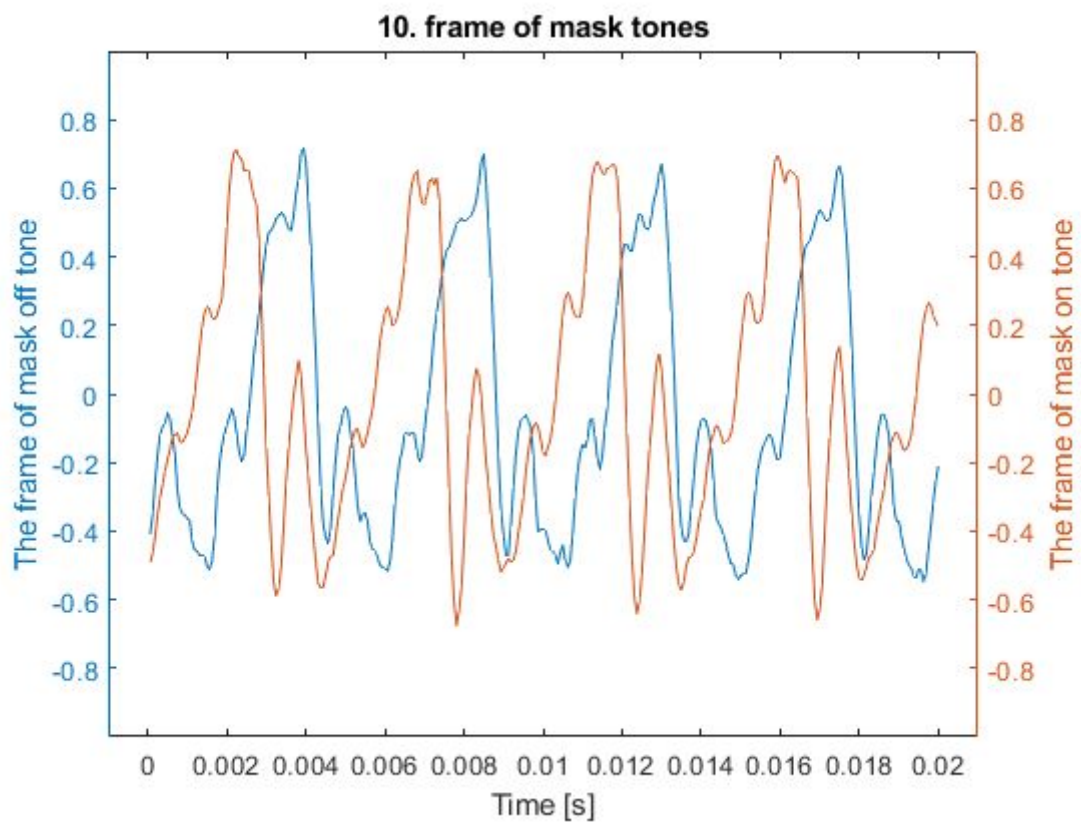
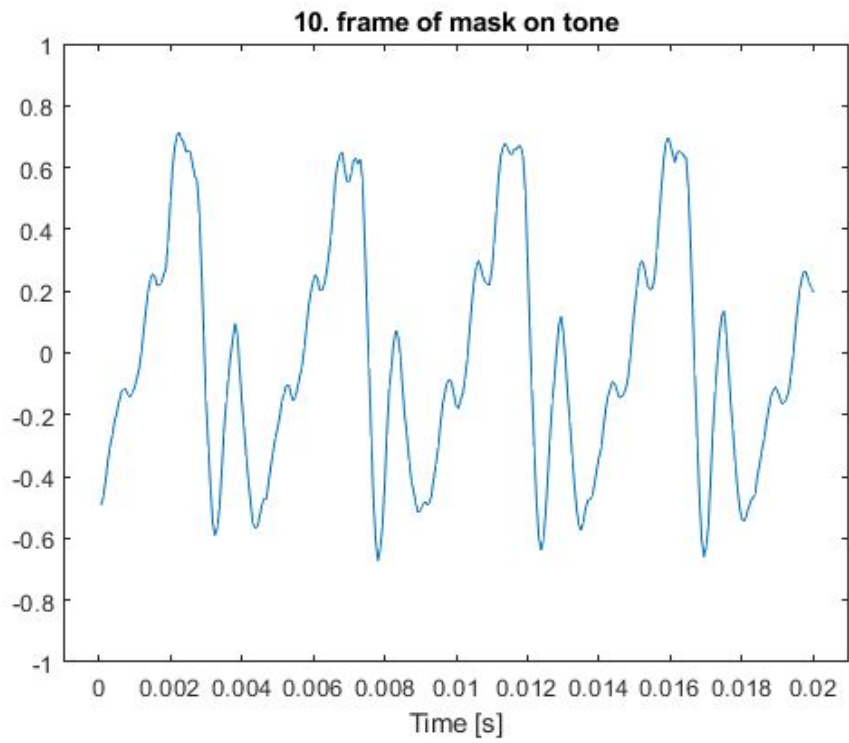
a) **Výpočet velikosti rámce vo vzorkách**

VZOREC
$\text{samples\_per\_ms} = \text{length}(s) / \text{length\_in\_ms};$ $\text{one\_frame\_length} = \text{samples\_per\_ms} * \text{howmuch};$

LEGENDA
<b>samples_per_ms</b> - vzorky prisluchajúce ku 1 milisekunde <b>length(s)</b> - dĺžka signálu <b>s</b> vo vzorkách <b>length_in_ms</b> - dĺžka signálu v milisekundách <b>one_frame_length</b> - dĺžka jedného rámca vo vzorkách <b>howmuch</b> - konštanta pre určenie veľkosti rámca v milisekundách

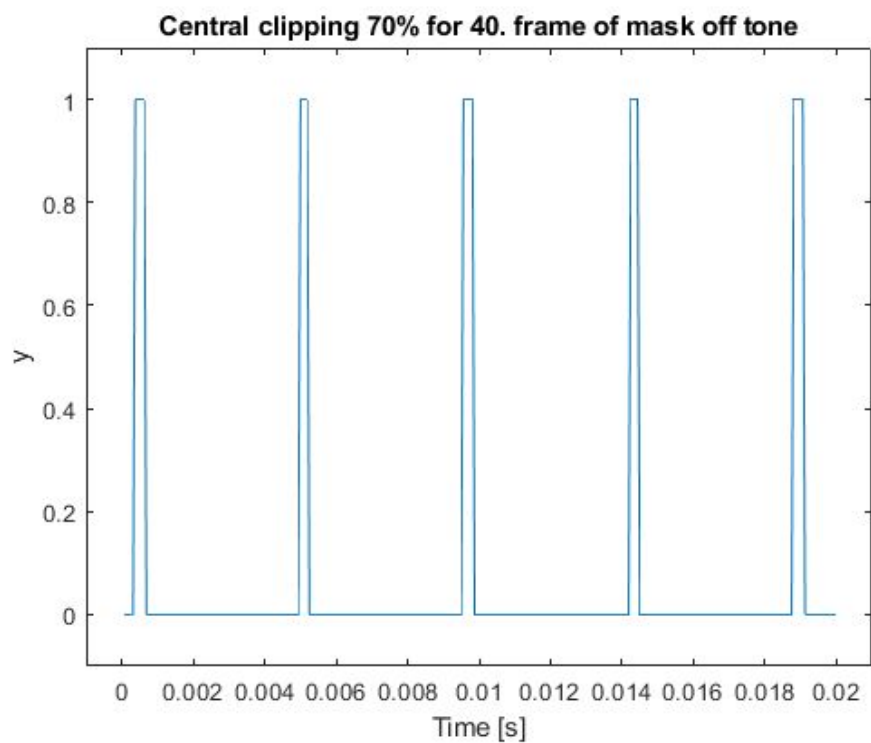
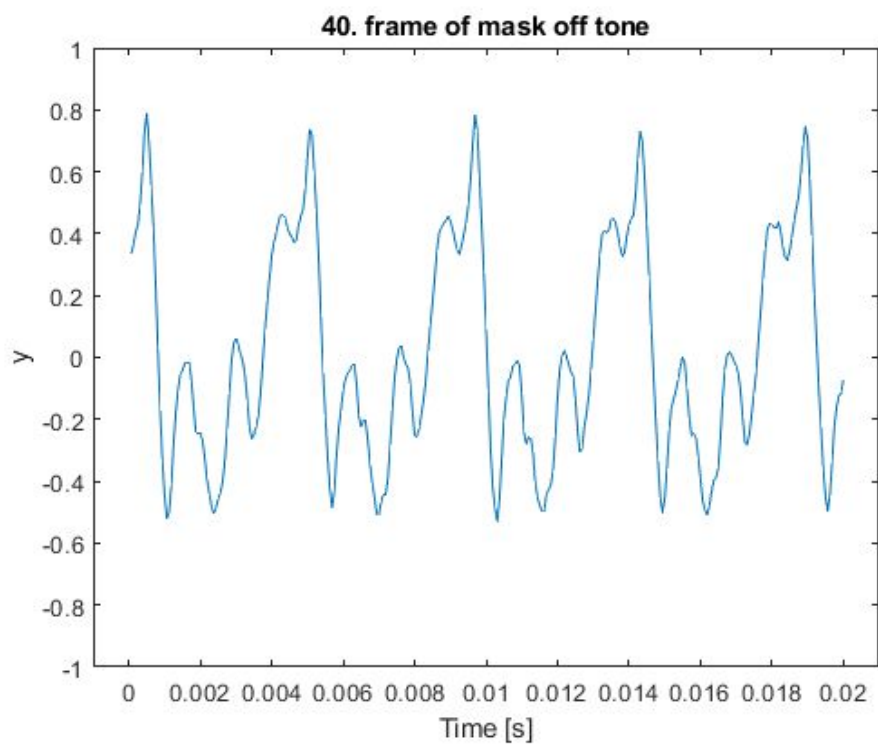
b) Graf dvoch zvolených rámcov **s rúškou, bez rúšky, prvý a druhý graf v jednom**

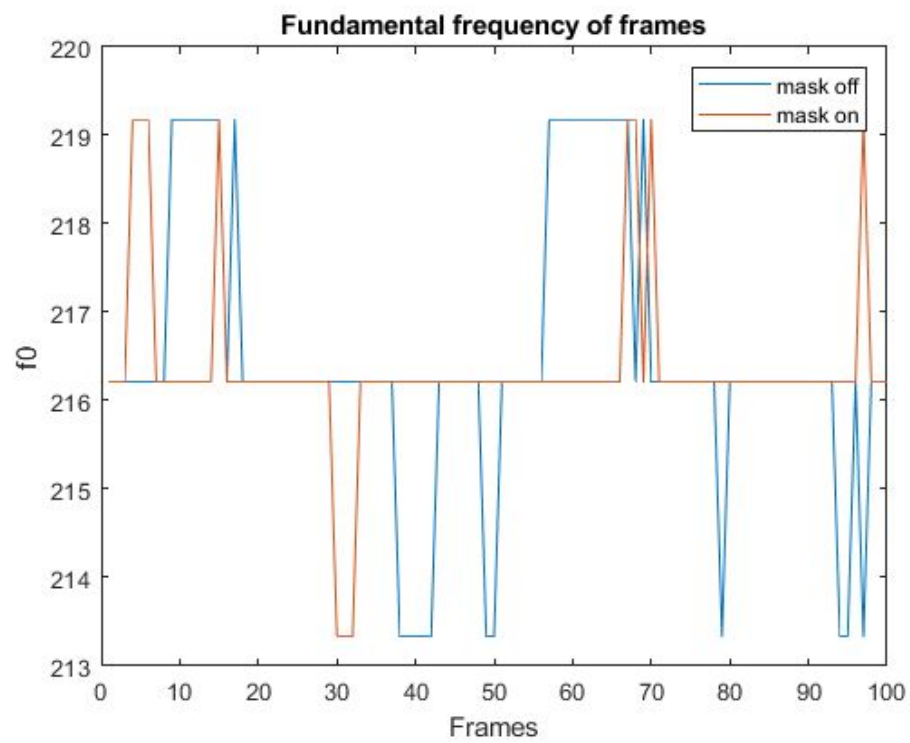
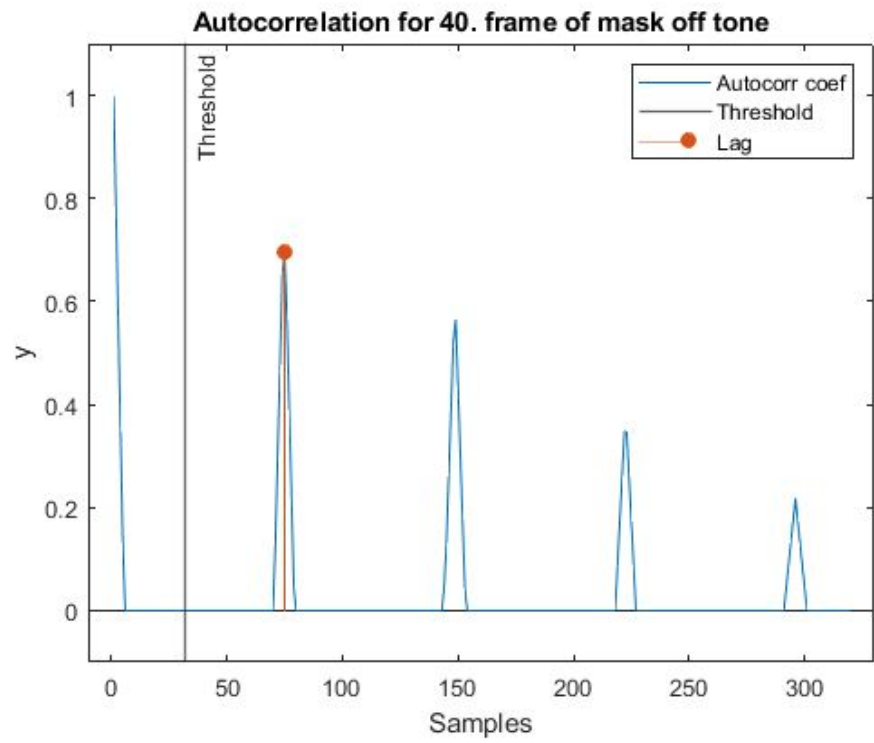




#### 4. úloha

- a) Ľubovoľný rámec a na ňom aplikované **centrálné klipovanie**, **autokorelácia** a **porovnanie základných frekvencií nahrávok v grafe**.





b) Stredná hodnota a rozptyl základnej frekvencie oboch testovacích nahrávok

	STREDNÁ HODNOTA	ROZPTYL
<b>f0 bez rúšky</b>	216.491472294212	2.61916728462011
<b>f0 s rúškou</b>	216.366679007775	0.937882186288402

c) **Odpoveď na otázku:**

Použil som vedomosti z tejto prednášky z predmetu **ZRE**:

[http://www.fit.vutbr.cz/~grezl/ZRE/lectures/05\\_pitch\\_en.pdf](http://www.fit.vutbr.cz/~grezl/ZRE/lectures/05_pitch_en.pdf)

#### Clipping Level Value Estimation

As a speech signal  $s(n)$  is a nonstationary signal, the clipping level changes and it is necessary to estimate it for every frame, for which pitch is predicted. A simple method is to estimate the clipping level from the absolute maximum value in the frame:

$$c_L = k \max_{n=0 \dots N-1} |x(n)|, \quad (11)$$

where the constant  $k$  is selected between 0.6 and 0.8. Further, subdivision into several micro-frames can be done, for instance  $x_1(n)$ ,  $x_2(n)$ ,  $x_3(n)$  of one third of the original frame length. The clipping level is then given by the lowest maximum from the micro-frames:

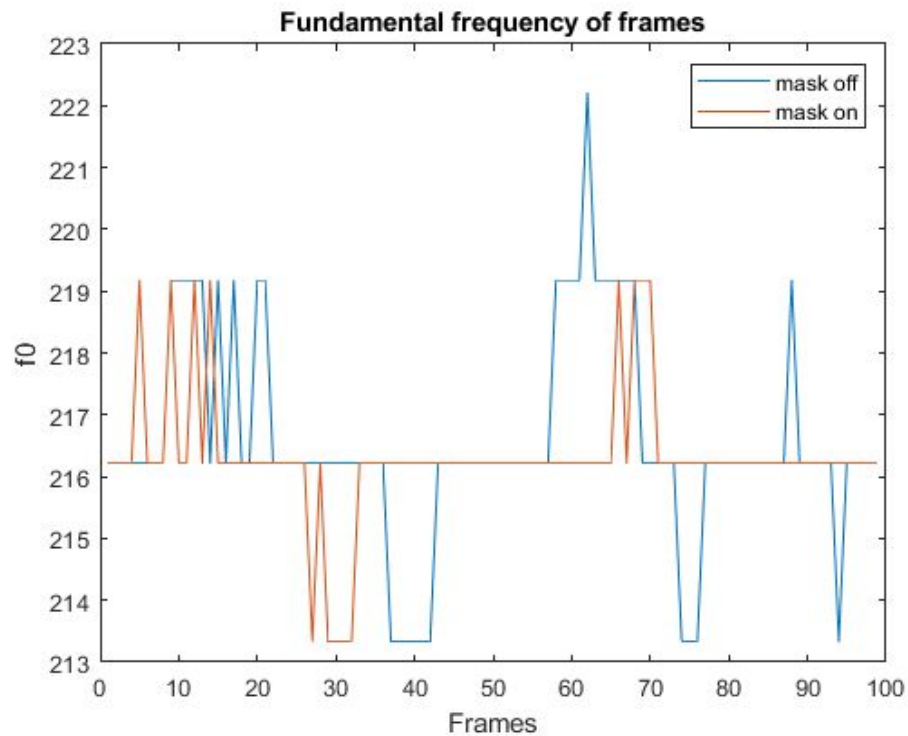
$$c_L = k \min \{ \max |x_1(n)|, \max |x_2(n)|, \max |x_3(n)| \} \quad (12)$$

**Issue:** clipping of noise in pauses, where subsequently can be detected pitch. The method therefore should be preceded by the silence level  $s_L$  estimation. In the maximum of the signal is  $< s_L$ , then the frame is not further processed.

Kde po zvýšení centrálneho klipovania zo 70% (0,7) na 80% (0,8) bola viditeľne menšia veľkosť zmeny  $f_0$ .

**Zdôvodnenie:**

- I. A to z dôvodu, že pri vyššom nastavení klipovania sa aj zvýšila absolútna maximálna hodnota rámca, a z toho vyplýva, že sa aj zvýšila minimálna hranica pre odhad. Hodnoty vzoriek rámca menšie ako  **$\max(\text{abs}(\text{frame}(n)))$**  alebo väčšie ako  **$-\max(\text{abs}(\text{frame}(n)))$** . Následne neprípustné hodnoty (vzorky v klipovanom rámci) v podmienkach clippingu budú nastavené na hodnotu 0, kde práve je veľká šanca za celkom dobrých podmienok, tie vyššie skoky základnej frekvencie sa čiastočne alebo úplne eliminujú.
- II. Alebo zvoliť techniku zvýraznenú v priloženej snímke prezentácie.
- III. Prikladám ilustračný graf porovnania so základnými frekvenciami testovacích nahrávok s jedinou zmenou a to **zapnutie klipovania na 80%. Zmeny sú nepatrné. Predpokladám, že dôvodom je to, že pri nahrávaní tónu som nedržel rovnakú intenzitu tónu. Myslím, že mi skákala hlasitosť i keď som podľa ladičky sa držal v A3.**



IV. Či som sa trafil do A3 som si empiricky overil podľa dát z tohto zdroja:

<http://radkon.eu/projects/other/tones.php?lang=sk>

Na tejto webovej stránke sú ku nájdeniu frekvencie pre tóny.

A3 sa pohybuje na frekvencií **220 Hz**. Moje frekvencie v grafe sa pohybujú v intervale  **$f_0 \in <213.333; 219.178>$** . Kde sa mi potvrdzuje tvrdenie z bodu III.

Výška tónu je v medziach týchto tónov.

G#3 / Ab3	gis0 / as0	207.652348789973 <b>Hz</b>
A3	a0	220 <b>Hz</b>



## 5. DFT a spektogramy

a)

### FUNKCIA IMPLEMENTUJÚCA DFT

- výsledkom porovnania s implicitnou funkciou v matlabe **fft** bolo zistenie, že mi to počíta pomalšie v rádoch sekúnd.
- bol použitý **zero-padding**, **žeby sa to hodilo som sa dozvedel tu:**  
[https://www.fit.vutbr.cz/study/courses/ISS/public/pred/10\\_dft/dft.pdf](https://www.fit.vutbr.cz/study/courses/ISS/public/pred/10_dft/dft.pdf)

Použil som tento vzorec:

$$X[k] = \sum_{n=0}^{N-1} x[n] e^{-j \frac{2\pi}{N} kn}$$

Zdroj informácií:

- 1) [https://www.fit.vutbr.cz/study/courses/ISS/public/pred/10\\_dft/dft.pdf](https://www.fit.vutbr.cz/study/courses/ISS/public/pred/10_dft/dft.pdf)
- 2) matlab fórum

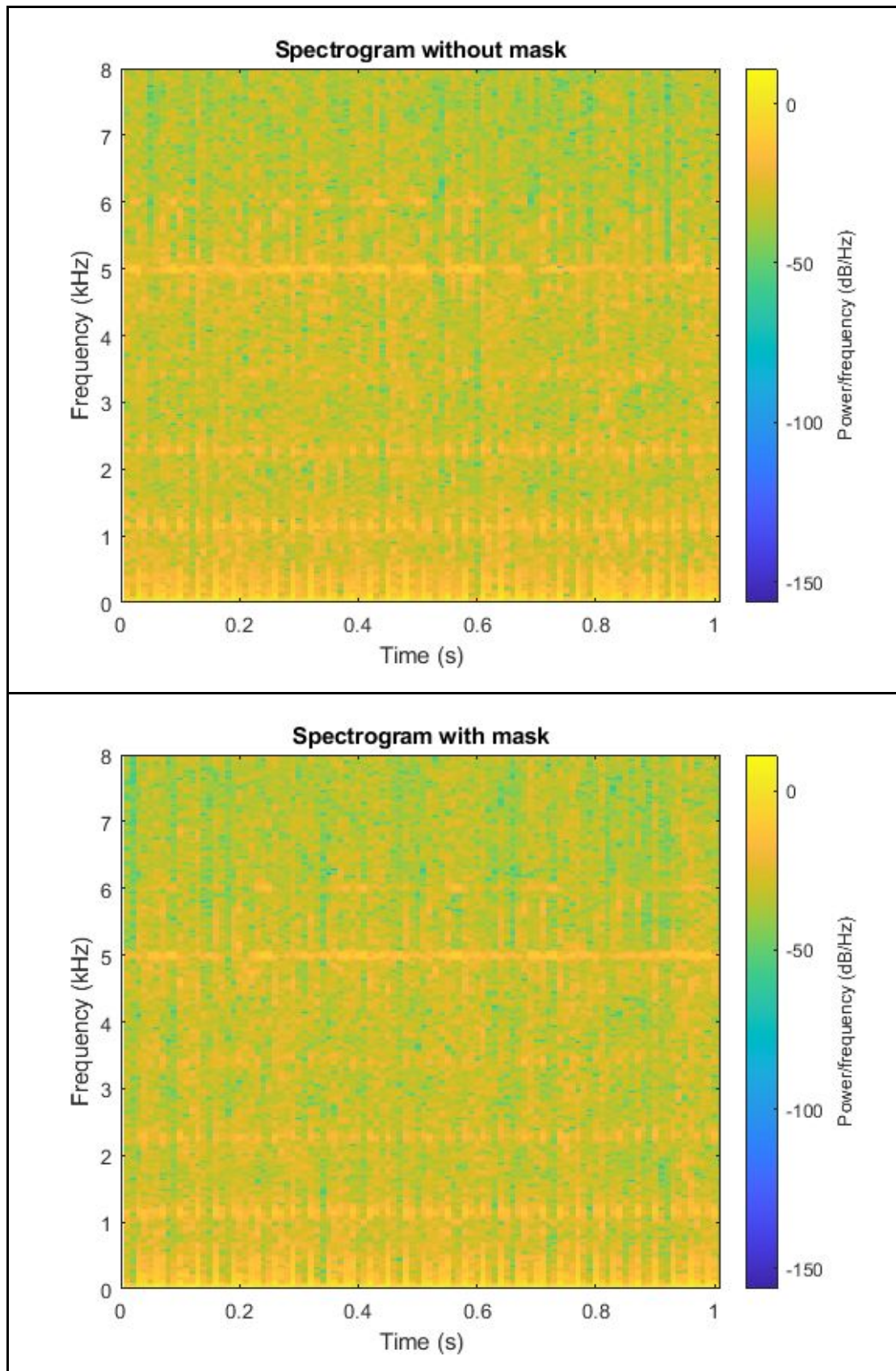
```
function dft = do_dft(input_frame,N)
    % alokujem si miesto pre vektory
    align = zeros(N,1);
    dft = zeros(N,1);
    % dft_frame_normalized = zeros(N,1);

    % vyplnime zarovnaný nulami vektor o dĺžke N datami rámca o dĺžke
    for frame_val = input_frame(:,1)
        align(1:320,1) = frame_val;
    end

    % 1. x~[n] = x[ mod N(n)]. periodizujeme
    % dft_frame_normalized(:,1) = mod(align(:,1), N);

    % aplikacia dft na data s ruskou a bez nej
    for k = 0 : N - 1
        sum = 0;
        for n = 0 : N - 1 % aplikacia vzorca pre dft na vstupne data
            sum = sum + times(align(n + 1,1), exp(-j*2*pi*k*n/N));
        end
        dft(k + 1,1) = sum; %zapisem po kazdej iteracii dopocitanu sumu
    end
    % zapisem vysledok dft
    dft = dft(1:N,1);
end
```

- b) SPEKTROGRAMY (používal som pri vykreslení spektrogramu **hammingovu okienkovú funkciu** ako jeden z parametrov pre funkciu **spectrogram**)



## 6. úloha

a) Vzťah pre výpočet  $H(e^{j\omega})$

$$H(e^{j\omega}) = |Y(e^{j\omega})| / |X(e^{j\omega})|$$

$H(e^{j\omega})$  - frekvenčná charakteristika rúška

$Y(e^{j\omega})$  - DFT dáta s rúškom

$X(e^{j\omega})$  - DFT dáta bez rúška

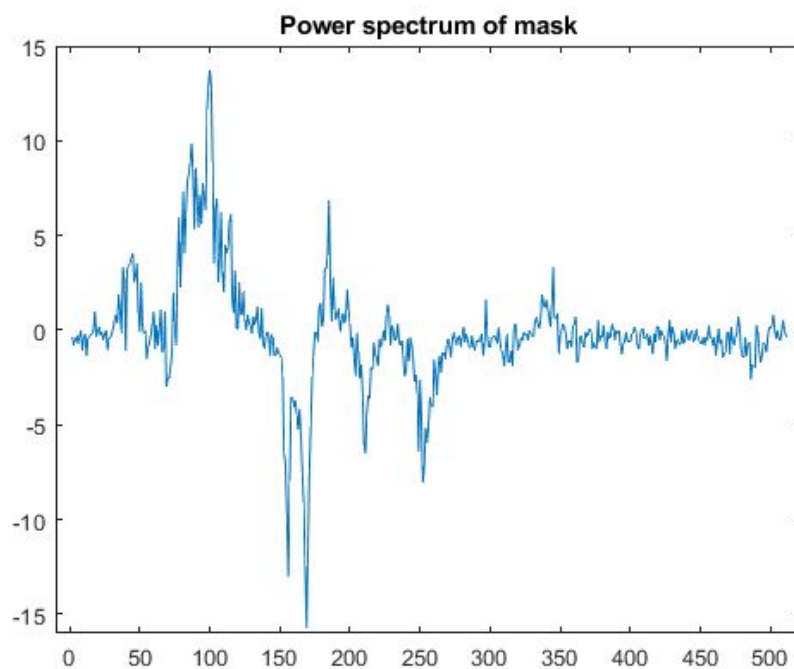
**Zdroje:**

<https://www.researchgate.net/post/What-formula-should-I-use-to-calculate-the-power-spectrum-density-of-a-FFT>

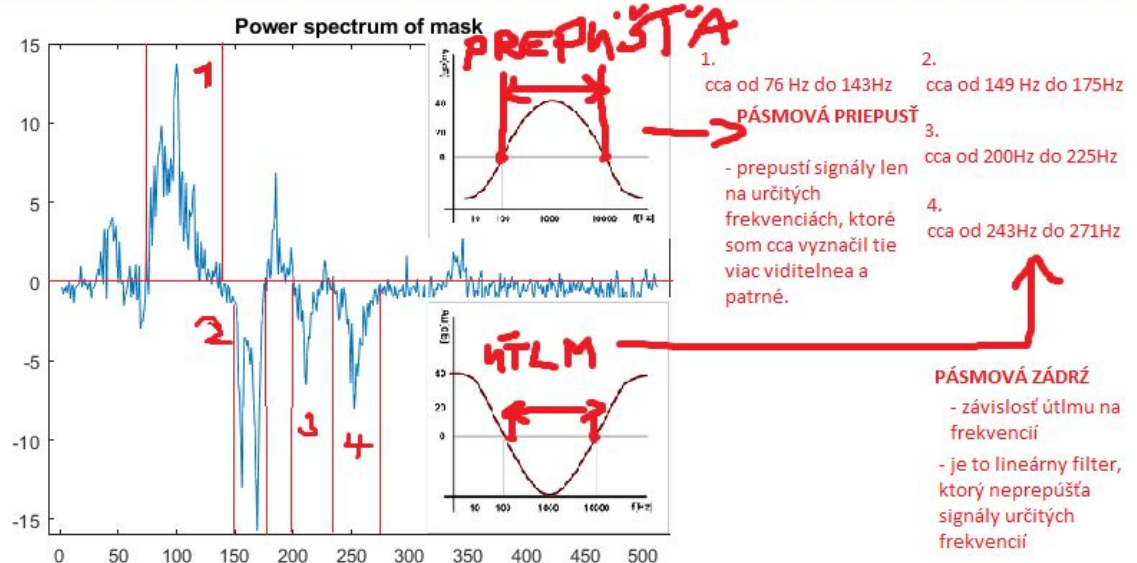
**Postup:** Pre dáta s rúškom a bez neho som spočítal DFT a uložil do matice. Odsekol časť, ktorá bola už časťou symetrickou.

Výsledná veľkosť matice je **[N/2+1 x 100]**. Kde N je hodnota pre DFT 1024 z predošlej úlohy. Ďalej aby som nedelil a nepriemeroval komplexné zložky dát som pre obe matice spočítal absolútne hodnoty. Aby som dostal frekvenčnú charakteristiku pre každý rámec, tak som matice dát navzájom vydelen a podiel spriemeroval cez všetky rámce aby som získal jednu frekvenčnú charakteristiku.

b) Frekvenčná charakteristika rúšky



c) Krátky komentár ku filteru pomocou znázornenia na frekvenčnej charakteristike rúška



7. úloha

a) Vlastná implementácia IDFT

#### FUNKCIA IMPLEMENTUJÚCA IDFT

Použil som tento vzorec:

zpětná transformace IDFT :

$$x[n] = \frac{1}{N} \sum_{k=0}^{N-1} X_{DFT}[k] e^{j2\pi k n / N} \quad n = 0, 1, \dots, N-1$$

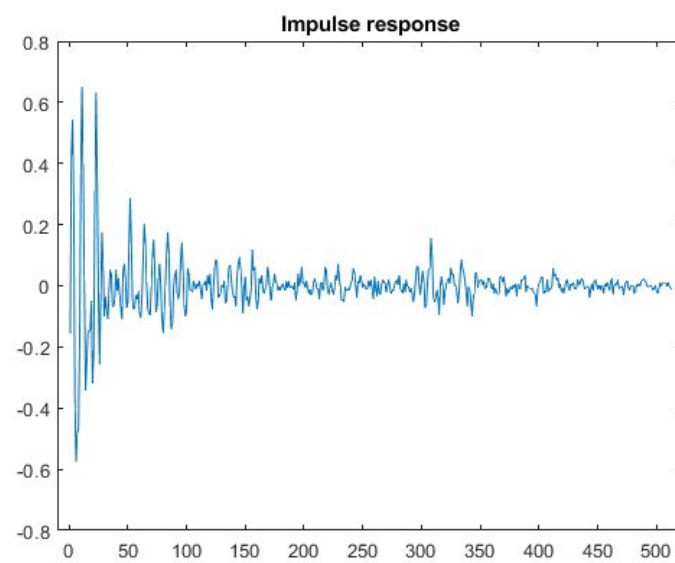
Zdroj informácií:

1) [http://www.kiv.zcu.cz/~mautner/Azs/Azs5\\_Fourierova\\_transformace.pdf](http://www.kiv.zcu.cz/~mautner/Azs/Azs5_Fourierova_transformace.pdf)

```
function idft = do_idft(input_data,N)
%alokujem si miesto pre vektor
idft = zeros(1,N);

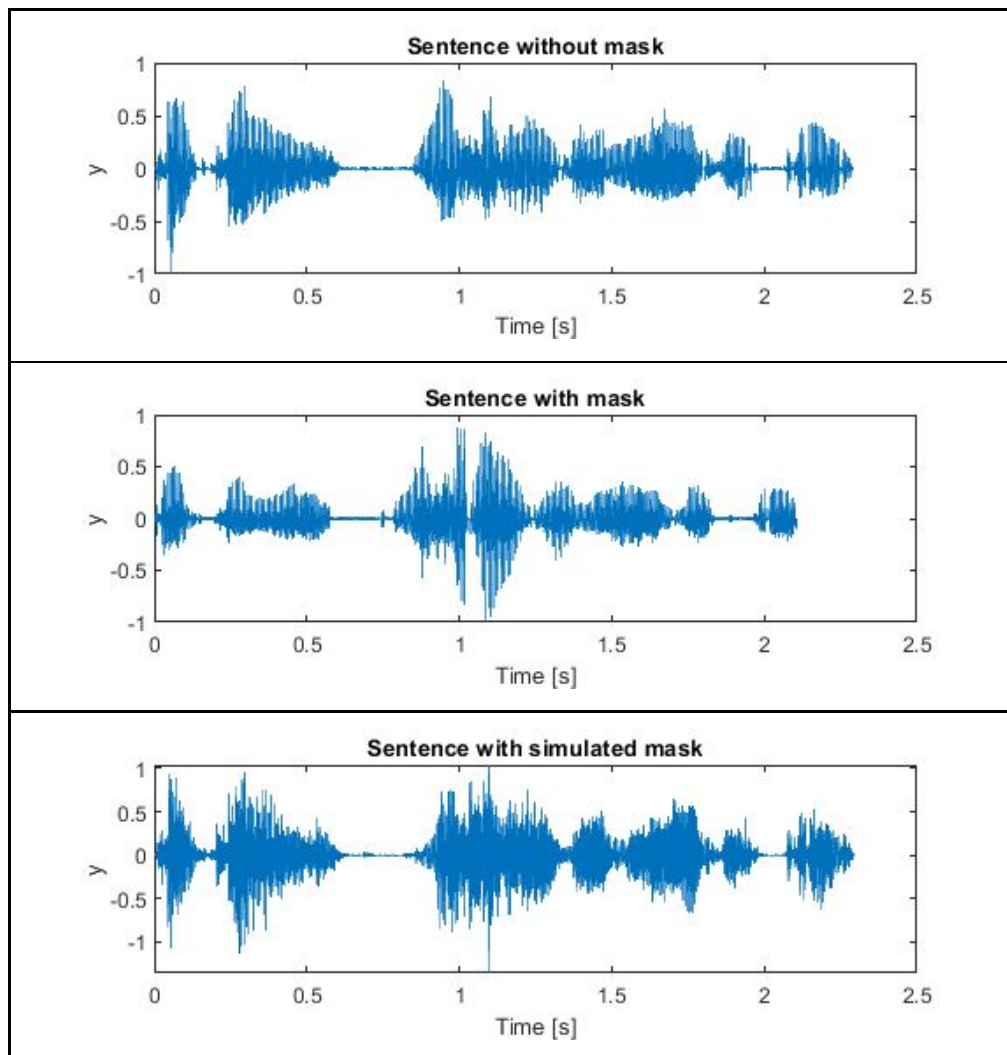
% aplikacia dft na data s ruskou a bez nej
for n = 0 : N - 1
    sum = 0;
    for k = 0 : N - 1 % aplikacia vzorca pre dft na vstupne data
        sum = sum + times(input_data(1,k + 1), exp(j*2*pi*k*n/N));
    end
    % zapisem po kazdej iteracii dopocitanu sumu
    idft(1,n + 1) = times((1/N), sum);
end
%zapisem vysledok dft
idft = idft(1,1:N);
end
```

b) Graf impulznej odozvy rúška



## 8. úloha

- a) Graf nahranej vety bez rúška, s rúškou a so simulovaným rúškom



- b) Odpoveď na otázky

Signál so simulovaným rúškom oproti signálu so skutočným rúškom sa javí zosilnený ako i pri posluchu, tak i pri vizuálnej kontrole signálu. A to v kladnej i zápornej časti podľa osi Y. Predpokladám, že príčinou tohto efektu "zosilnenia" je pásmová priepusť v kladnej časti osi y a pásmová zádrž v zápornej časti osi Y filtra pre simulovanie rúška. Signály sú najviac podobné v stredných hodnotách signálu v okolí bodu 0 na osi Y. Najviac sa líšia po okrajoch signálu, čo najďalej od stredy.

## 9. Záver

Na hlavné črty filtra som poukázal v úlohách 6c) a 8b), kde by som dodal, že pásmová zádrž zadržiava a nepúšťa práve tie frekvencie, v ktorých som nahrával tón a cca  $\pm 50$  Hz nad a pod touto frekvenciou, takže ju tlmí a to bol žiadaný efekt simulácie rúšky podľa mňa. No bolo by to žiadaným efektom, keby to pásmová priepusť nepokazila a neprepustila nižšie frekvencie s minimálnym útlmom. Keďže sú prítomné v jednom filtry obe priepuste, tak nakoniec je konečný filtrovaný signál zosilnený. A som za to, že to nebol žiadaný efekt. Na druhú stranu nahrávky po prefiltrovaní sú pekne jasne počuteľné a slová sú zrozumiteľné. Jemné praskanie v nahrávke vety je spôsobené nekvalitnou nahrávkou z dôvodu HW závady. Takže informácia v signále nebola stratená. Myslím, že riešenie by bolo úspešné ak by bola prítomná iba jedna z priepustí a to pásmová zádrž.

Na úplný záver chcem poďakovať všetkým, čo toto zadanie vymysleli. Veľa som si počas riešenia toho zopakoval a naučil sa. Napríklad práca v Matlabe bola pre mňa španielskou dedinou a veľa som sa v ňom naučil a je to cenná skúsenosť do budúcnosti.