



ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE

**Fakulta elektrotechnická
Katedra radioelektroniky**

Vliv dynamické komprese signálu na parametr STI určovaného přímou metodou
Influence of Dynamic Signal Compression on Intelligibility Expressed by STI Using Direct Method

Diplomová práce

Studijní program: Elektronika a komunikace
Studijní obor: Audiovizuální technika a zpracování signálů

Vedoucí práce: Ing. Marek Brothánek, Ph.D.

Bc. Václav Patočka

Praha 2019

I. OSOBNÍ A STUDIJNÍ ÚDAJE

Příjmení: **Patočka** Jméno: **Václav** Osobní číslo: **434737**
Fakulta/ústav: **Fakulta elektrotechnická**
Zadávací katedra/ústav: **Katedra radioelektroniky**
Studijní program: **Elektronika a komunikace**
Studijní obor: **Audiovizuální technika a zpracování signálů**

II. ÚDAJE K DIPLOMOVÉ PRÁCI

Název diplomové práce:

Vliv dynamické komprese signálu na parametr STI určovaného přímou metodou

Název diplomové práce anglicky:

Influence of Dynamic Signal Compression on Intelligibility Expressed by STI Using Direct Method

Pokyny pro vypracování:

V práci popíšete možné způsoby měření parametru STI. Při použití přímé metody experimentálně ověřte vliv změny dynamických vlastností signálů na výslednou hodnotu STI určovaného zjednodušenou metodou STIPA. Měřicí signál modifikujte různými druhy aditivního šumu (bílý, 1/f, vážený filtrem A) a dosažené výsledky detailně zhodnoťte. Zaměřte se na různé typy uspořádání řídicí části dynamického kompresoru a použití různých detektorů (špičkový, RMS apod.). Dle specifikace vedoucího práce proveďte výpočty v prostředí MATLAB, za použití audio-pluginů, případně fyzického zařízení.

Seznam doporučené literatury:

- [1] International Electrotechnical Commission: IEC 60268-16 - Sound system equipment - Part 16: Objective rating of speech intelligibility by speech transmission index, Edition 4.0, Geneva, 2011.
- [2] Steeneken, H. J. M., Houtgast, T.: A physical method for measuring speech transmission quality, J. Acoust. Soc. Am., 67(1), 318-326, 1980.
- [3] Giannoulis, D., Massberg, M., Reiss, J. D.: Digital Dynamic Range Compressor Design - A Tutorial and Analysis, J. Audio Eng. Soc, 60(6), 399-408, 2012.

Jméno a pracoviště vedoucí(ho) diplomové práce:

Ing. Marek Brothánek, Ph.D., katedra fyziky FEL

Jméno a pracoviště druhé(ho) vedoucí(ho) nebo konzultanta(ky) diplomové práce:

Datum zadání diplomové práce: **08.02.2019**

Termín odevzdání diplomové práce: **07.01.2020**

Platnost zadání diplomové práce: **20.09.2020**

Ing. Marek Brothánek, Ph.D.
podpis vedoucí(ho) práce

doc. Ing. Josef Dobeš, CSc.
podpis vedoucí(ho) ústavu/katedry

prof. Ing. Pavel Ripka, CSc.
podpis děkana(ky)

III. PŘEVZETÍ ZADÁNÍ

Diplomant bere na vědomí, že je povinen vypracovat diplomovou práci samostatně, bez cizí pomoci, s výjimkou poskytnutých konzultací. Seznam použité literatury, jiných pramenů a jmen konzultantů je třeba uvést v diplomové práci.

Datum převzetí zadání

Podpis studenta

Abstrakt

Tato práce se zabývá stavbou dynamického kompresoru, jeho částí a parametrů. Dále je popsán parametr srozumitelnosti STI a metody jeho získání. Nakonec je studován vliv jednotlivých parametrů dynamického kompresoru na parametr STI měřeného pomocí přímé metody a vliv přidaného šumu do měřicího řetězce.

Klíčová slova

Dynamický kompresor, STI, STIPA, šum, srozumitelnost

Abstract

This thesis deals with the construction of dynamic range compressor, its parts and parameters. In the next part, the objective intelligibility parameter STI is introduced, as well as methods of obtaining it. Finally, the influence of the dynamic range compressor on the STI parameter obtained by the direct method is observed and the influence of noise on the measurement is studied.

Keywords

Dynamic range compressor, STI, STIPA, noise, intelligibility

Prohlášení

Prohlašuji, že jsem předloženou práci vypracoval samostatně a že jsem uvedl veškeré použité informační zdroje v souladu s Metodickým pokynem o dodržování etických principů při přípravě vysokoškolských závěrečných prací.

V Praze dne

.....

Podpis

Poděkování

Chtěl bych poděkovat svému vedoucímu práce, Ing. Markovi Brothánkovi, Ph.D. za trpělivost a ochotnou pomoc při řešení této závěrečné práce. Dále bych chtěl poděkovat Ing. Zbyňku Lébrovi z oddělení elektroakustiky v GZ Media za pomoc při studiu a řešení dynamického kompresoru. Nakonec bych chtěl také poděkovat své rodině za neustálou podporu.

Obsah

1. Úvod.....	12
2. Nelineární systémy.....	14
3. Dynamický kompresor	16
3.1 Detektor v kompresoru.....	16
3.1.1 Peak detektor	17
3.1.2 RMS detektor	18
3.2 Napětím řízený zesilovač	19
3.3 Dynamická komprese a její vlastnosti	19
3.4 Uspořádání kompresoru	22
3.4.1 Zpětnovazební a dopřednovazební návrh.....	22
3.4.2 Umístění detektoru ve výpočetním řetězci	23
4. Parametr STI.....	26
4.1 Historie STI.....	26
4.2 Princip výpočtu	26
4.3 Metody měření.....	27
4.3.1 Nepřímá metoda.....	27
4.3.2 Přímá metoda	27
4.3.3 Přímá metoda – zjednodušené postupy	28
4.4 Výpočet STI metodou STIPA.....	29
5. Experimentální část	34
5.1 Výroba kompresoru v prostředí Matlab	34
5.2 Vliv různých parametrů kompresoru na hodnotu STI.....	35
5.2.1 Vliv délky testovacího souboru na hodnotu STI	36
5.2.2 Vliv attack time a release time na STI	37
5.2.3 Vliv parametru threshold na hodnotu STI.....	39
5.2.4 Vliv detektoru a zalomení převodní charakteristiky.....	41
5.2.5 Vliv délky výpočetního okna efektivní hodnoty na STI.....	42
5.3 Vliv přidaného šumu v kanálu na měření	44
5.3.1 Zpracování v MATLABu	44
5.3.2 Zpracování v pluginu.....	47
6. Závěr	50
7. Seznam obrázků.....	52
8. Literatura.....	54

1. Úvod

Tato práce má za úkol čtenáře seznámit s funkcí dynamického kompresoru, parametrem srozumitelnosti STI a vlivem dynamického kompresoru na tento parametr. Zkoumána jsou různá nastavení parametrů kompresoru, jako je attack time, release time, threshold, nebo typ detektoru. Dále je porovnáván nasimulovaný kompresor v MATLABu se známou konstrukcí a neznámý kompresor z volně dostupného pluginu pro hudební procesory. Nakonec je pozorován vliv kompresoru na zašuměný signál, který má simulovat reálné podmínky měření.

V první části je lehce nastíněna problematika nelineárních soustav, jejichž zástupcem je právě i dynamický kompresor.

Další kapitola popisuje již zmíněný dynamický kompresor. Popsána je jeho funkce, postavení v hudebním i nehupebním průmyslu, jednotlivé části, zpracování signálu a nastavitelné parametry důležité k jeho funkci. Následuje popis detektorů a jejich umístění v procesu průchodu signálu soustavou.

Čtvrtá kapitola je věnována objektivnímu hodnocení srozumitelnosti a parametru STI. Nejprve se zabývá problematikou srozumitelnosti a vzniku parametru STI. Dále jsou popsány metody získání tohoto parametru.

Následující kapitola je věnována implementaci a testování kompresoru jak v prostředí MATLAB, tak v komerčním zvukovém procesoru. Je zde popsán návrh kompresoru a postup výpočtu STI přímou metodou. Dále se nachází část věnující se samotným experimentům týkajících se vlivu různých parametrů kompresoru na hodnotu STI, a dále vlivu aditivního šumu přidaného do měřicího řetězce na výpočet STI.

V závěru práce jsou shrnuty výsledky experimentů a zhodnocení vlivu dynamické komprese signálu na parametr STI.

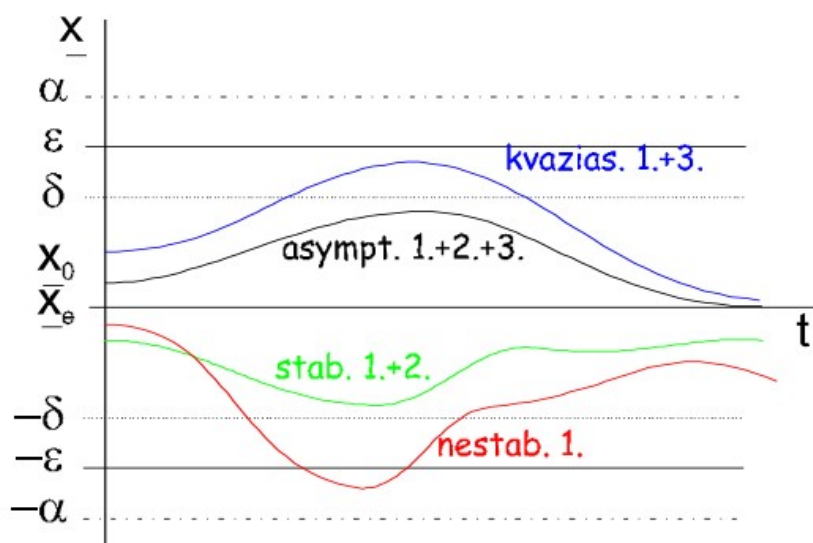
2. Nelineární systémy

V reálných podmínkách se málokterá soustava chová plně lineárně. Proto je dobré se seznámit se soustavami nelineárními. Příkladem těchto nelineárních soustav může být například saturace (nasycení) snímačů, reproduktorů, zesilovačů; snímače s pásmem necitlivosti; regulační členy se železným jádrem (hysterezní smyčka) nebo zařízení s dvupolohovou charakteristikou reléového typu.

Přirozené nelinearity jsou zapříčiněny fyzikálními vlastnostmi materiálu, ze kterého je prvek vyroben. Tyto nelinearity jsou pro nás nežádoucí a snažíme se jich zbavit. Úmyslně zaváděné nelinearity slouží buď ke kompenzaci jiné nelinearity, nebo pro náš specifický účel, například dynamická komprese zvuku.

U nelineárních systémů nemůžeme počítat s vlastnostmi systémů lineárních, tedy že odezva na výstupu tohoto systému není násobkem hodnoty na vstupu. Neplatí zde princip superpozice.

Abychom mohli v praxi využít nelineární systém, je třeba znát jeho podmínky stability. Jeden ze způsobů určení podmínek stability stanovil ruský matematik Ljapunov a zní následovně: Pro každou vstupní hodnotu x_0 , kde $|x_0 - x_e| \leq \delta$, musí platit, že vyvolá odezvu uvnitř definovaného intervalu výstupních hodnot $|x(t) - x_e| \leq \varepsilon$ a zároveň tento výstup konverguje v čase k hodnotě x_e $|\lim_{t \rightarrow \infty} x(t)| = x_e$. Názorná ukázka je na obrázku 1. Červená křivka představuje výstup, jehož průběh překračuje hodnotu ε , čímž nesplňuje první podmínku. Naopak



Obrázek 1 - Stabilita nelineárního systému. Barevné křivky ukazují splnění podmínek stability.

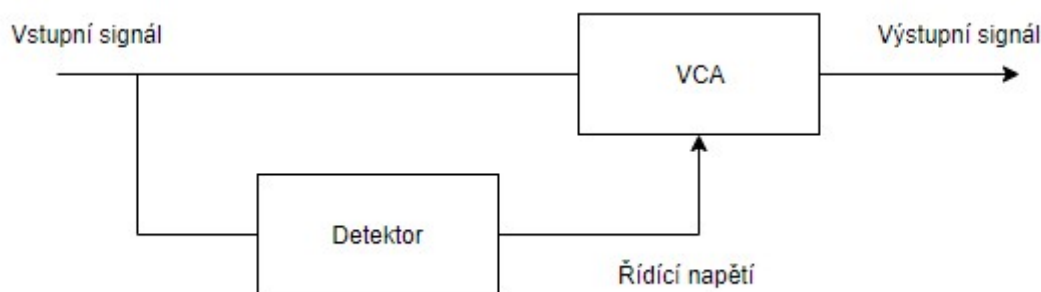
zelená křivka představuje výstup, který nekonverguje k hodnotě x_e , proto také nemůže být stabilní. Černá křivka je zástupcem plně stabilního systému. Více o nelineárních soustavách a jejich stabilitě lze nalézt v textech [4] a [5].

Dynamický kompresor je také zástupcem nelineárního systému. Jeho převodní charakteristika je nelineární, nicméně díky tomu, že vstupní signál buď projde beze změny, nebo je ztlumen, můžeme kompresor považovat za stabilní nelineární systém.

3. Dynamický kompresor

Dynamický kompresor používáme ke snížení dynamiky signálu. Využití může najít jak v hudebním průmyslu, například komprese hudby za cílem větší hlasitosti, tzv. Loudness War, nebo ochrana akustických prvků před přetížením. Dále se uplatňuje i v lékařství, kdy subjekty se sníženou schopností sluchu, například pacienti s kochleárním implantátem, potřebují snížit dynamiku veškerých signálů, aby se vešly do rozsahu poškozené kochley. Tímto problémem se zabýval například R. Plomp ve své práci [9], kde polemizuje o efektu amplitudové komprese na srozumitelnost u těchto kochleárních implantátů. V této práci se budeme zabývat vlivem tohoto dynamického kompresoru na parametr STI.

Opakem ke kompresoru je expandér, který dynamický rozsah zvětšuje. Existuje i složení obou přístrojů dohromady, nazývaný komandér, jehož funkcí je na nízkých úrovních dynamický rozsah zvyšovat a na vysokých úrovních naopak dynamický rozsah snižovat. Tyto přístroje se v praxi příliš často nevyužívají, a dominantní postavení v užitečnosti má právě kompresor. Na obrázku 2 lze vidět blokové schéma dynamického kompresoru.



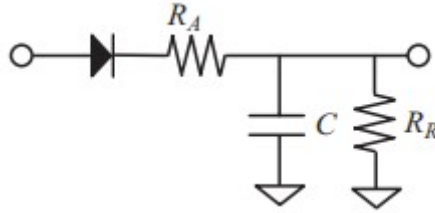
Obrázek 2 - Blokové schéma dynamického kompresoru

3.1 Detektor v kompresoru

Vstupní signál kompresoru je rozdělen na dvě části. První část je přivedena do napětím řízeného zesilovače a druhá část je použita k detekci překročení thresholdu (hodnota signálu, která určuje moment aplikace komprese, viz kapitola 3.3). Detektory úrovně mohou být dvojího typu, a to špičkový (peak) detektor, nebo detektor efektivní hodnoty (RMS detektor).

3.1.1 Peak detektor

Pokud okamžitá úroveň vstupního signálu přesáhne hodnotu threshold, v závislosti na parametru attack time se kompresor aktivuje. Z toho vyplývá, že pro peak detekci je vhodné použít krátkého attack time. Pokud nastavíme attack time větší, nedojde k úplné aplikaci zesílení. Jednoduchý analogový peak detektor je na obrázku 3.



Obrázek 3 - Jednoduchý peak detektor, sestavený z diody, dvou odporů a kondenzátoru (převzato z [1])

Když budeme považovat diodu za ideální, výstupní napětí na kondenzátoru bude rovno

$$\frac{dU_C}{dt} = \frac{\max(U_{in} - U_C, 0)}{R_A C} - \frac{U_C}{R_R C}, \quad (1)$$

kde U_C je napětí na kondenzátoru, U_{in} je vstupní napětí, R_A a R_R jsou hodnoty daných odporů. Kondenzátor C je nabíjen skrze odpor R_A při kladném napětí na diodě. Zároveň je vybíjen přes odpor R_R . Rovnice ideálního peak detektoru pak vypadá následovně:

$$y_L[n] = \alpha_R y_L[n-1] + (1 - \alpha_A) \max(x_L[n] - y_L[n-1], 0), \quad (2)$$

$y_L[n]$ je výstupní úroveň ve vzorku n , $x_L[n]$ je vstupní úroveň ve vzorku n a koeficienty α_R a α_A jsou dány velikostí použitých prvků, tedy

$$\alpha_A = e^{-1/(\tau_A f_s)} \quad \text{a} \quad (3a)$$

$$\alpha_R = e^{-1/(\tau_R f_s)}, \quad (3b)$$

kde $\tau_A = R_A C$ a $\tau_R = R_R C$.

3.1.2 RMS detektor

Detektor RMS k rozhodování využívá velikost efektivní hodnoty (RMS) vstupního signálu namísto absolutní hodnoty úrovně signálu. Toho se dá využít v případě kdy nechceme, aby kompresor reagoval na impulzní změny signálu, ale spíše na pozvolné zesílení signálu. Pro tento detektor je vhodné nastavení delšího attack i release time. Výhodou tohoto detektoru je, že RMS hodnota je více korelována s vnímanou hlasitostí. V online aplikaci je použití problematické, jelikož výpočet RMS hodnoty vyžaduje zpoždění o $M/2$ vzorků, viz následující rovnice.

$$y_L^2[n] = \frac{1}{M} \sum_{m=-\frac{M}{2}}^{\frac{M}{2}-1} x_L^2[n-m], \quad (4)$$

kde M je počet vzorků použitých k výpočtu, $y_L[n]$ je výstupní RMS úroveň ve vzorku n a $x_L[n-m]$ je vstupní úroveň signálu zpožděná o m vzorků. Proto se v praxi využívá low-pass IIR filtr, jehož výstup je poté odmocněn. Z předchozí rovnice se pak stane tato jednoduchá diferenční rovnice se zpožděním pouze jeden vzorek:

$$y_L^2[n] = \alpha y_L^2[n-1] + (1-\alpha)x_L^2[n], \quad (5)$$

kde α je koeficient filtru daný časovou konstantou τ IIR filtru ($\alpha = e^{-1/(\tau f_s)}$), f_s je vzorkovací frekvence.

3.2 Napětím řízený zesilovač

Pro řízení zesílení je vstupní signál poslán do VCA (*voltage controlled amplifier*, napětím řízený zesilovač). Tento VCA je řízen napětím vygenerovaným v předchozím bloku, detektoru. V minulosti bylo použito několik přístupů k vytvoření VCA. Prvním z nich je například využití napěťového děliče, kde místo druhého odporu byl fotorezistor, na který svítil zdroj světla, jehož jas byl dán řídicím napětím z detektoru. S větší intenzitou světla odpor fotorezistoru klesá, tím pádem na něm klesá i napětí a tím i zesílení. Podobným způsobem lze toto vyřešit pomocí FET tranzistoru. Zesílení je určeno odporem mezi drainem a sourcem. Řídicí napětí je přivedeno na gate a tím se snižuje i odpor mezi drainem a sourcem. Ještě starší přístup byl pomocí elektronkových zesilovačů, kde řízením napětí mezi mřížkou a katodou byl řízen zisk zesilovače. Moderní kompresory využívají integrovaných VCA obvodů, které jsou více stabilní a mají lepší vlastnosti (harmonické zkreslení, dynamický rozsah).

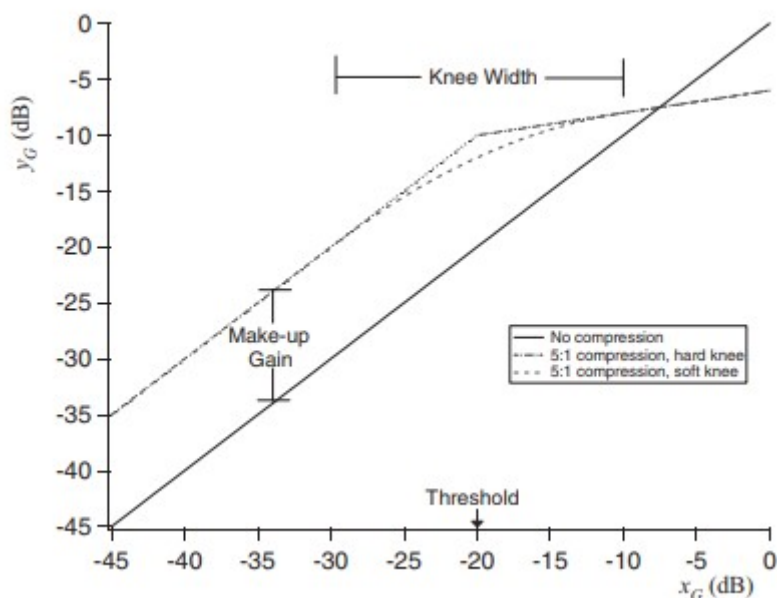
V digitální podobě je situace jednodušší. Vstupní signál můžeme rovnou vynásobit jak make-up gainem, tak řídicím napětím. V decibelové škále pak jde o pouhý součet:

$$y_{dB}[n] = x_{dB}[n] + c_{dB}[n] + M, \quad (6)$$

kde y je výstupní signál, x je vstupní signál, c je řídicí napětí a M je make-up gain.

3.3 Dynamická komprese a její vlastnosti

Následuje popis jednotlivých vlastností kompresoru. Bude se věnovat základním parametrům kompresoru – hodnotě threshold, kompresnímu poměru, make-up gainu, attack a release time, look-ahead parametru a frekvenčním kompresorům. Na obrázku 4 je uvedena převodní charakteristika dynamického kompresoru. Lze na něm také nalézt některé později uváděné vlastnosti kompresoru.



Obrázek 4 – Převodní charakteristika kompresoru (převzato z [1]). Hodnota *threshold* definuje převodní charakteristiku, se zlomem *soft knee* (čárkovaně) a *hard knee* (čerchovaně).

- *Threshold*

Hodnota *threshold* udává úroveň vstupního signálu, nad kterou se bude aplikovat komprese. Udává se v decibelech a může být vztažena k digitální nule, popřípadě hranici saturace u analogových přístrojů. *Threshold* souvisí se vstupním signálem. Pro slabé signály a vysokou hodnotu *threshold* nebude docházet ke kompresi.

- *Kompresní poměr*

Dynamická komprese rozsahu je proces, který snižuje dynamiku audio signálu. Úprava úrovně závisí na dalším nastavení kompresoru – kompresním poměru. Kompresní poměr udává sklon převodní charakteristiky za hodnotou *threshold*, viz obrázek 4. Čím větší je kompresní poměr, tím více plochá bude část křivky za hodnotou *threshold*. Nejnižší nastavení kompresního poměru je 1:1. V tomto případě kompresor nijak signál nemění, neaplikuje se žádná komprese, což ukazuje přímka na obrázku 4 pojmenovaná „No compression“. Největší nastavení je ∞ :1, kdy závislost výstupní hodnoty je rovna konstantě, tedy nezávisle na úrovni vstupního signálu bude za hodnotou *threshold* úroveň na výstupu stále stejná. Tím se z kompresoru stane tzv. limiter. V hudebním průmyslu se používá nejčastěji poměr 4:1 až 8:1. Při vyšším kompresním poměru než 8:1 již není velký slyšitelný rozdíl ve výstupním signálu vůči nižším nastavením.

- *Zalomení převodní charakteristiky*

Z obrázku 4 je také možné vidět nastavení zalomení převodní charakteristiky kompresoru. Při nastavení „hard knee“ je závislost výstupní úrovně až do poslední chvíle lineární a v bodě threshold se ostře zlomí podle zadaného kompresního poměru. Tím dosáhneme výraznější komprese. U „soft knee“ je přechod pozvolný, komprese je méně slyšitelná. Šířka zalomení (*knee width*) udává oblast postupné změny výstupní úrovně a je ovlivněna právě nastavením převodní charakteristiky kompresoru. Pro nastavení hard knee bude výstupní signál vypadat takto:

$$y_G = \begin{cases} x_G & \text{pro } x_G \leq T, \\ T + (x_G - T)/R & \text{pro } x_G > T, \end{cases} \quad (7)$$

kde y_G je výstupní úroveň, x_G je vstupní úroveň, T je threshold a R je kompresní poměr. Pro soft knee bude signál následující:

$$y_G = \begin{cases} x_G & \text{pro } 2(x_g - T) < -W, \\ x_G + \left(\frac{1}{R} - 1\right) \left(x_G - T + \frac{W}{2}\right)^2 / (2W) & \text{pro } 2|x_g - T| \leq W, \\ T + (x_G - T)/R & \text{pro } 2(x_g - T) > W, \end{cases} \quad (8)$$

kde symboly mají stejný význam jako v předchozí rovnici (7) a navíc W je šířka zalomení.

- *Make-Up Gain*

Make-Up Gain se používá na výstupu kompresoru a je to kompenzace snížení úrovně kompresorem. Jelikož úroveň výstupního signálu po průchodu kompresorem bude vždy menší, než úroveň vstupního signálu, aplikuje se na výstupu zesílení tak, aby měl výstupní signál stejnou úroveň jako vstupní signál.

- *Attack time a release time*

Attack time a release time jsou dva důležité parametry kompresoru. Oba dva udávají, jak rychle bude kompresor reagovat na změnu vstupního signálu. Attack time určuje, jak dlouho musí mít vstupní signál úroveň vyšší než threshold, aby se aktivovala komprese. To znamená, že samotné překročení thresholdu nestačí k tomu, aby se komprese uplatnila. V okamžik překročení thresholdu začne přecho-

dový děj s časovou konstantou o délce attack time, kdy se zesílení mění až na cílovou hodnotu.

Release time je opak attack time. Určuje čas, po který bude stále aplikována komprese v případě, že úroveň vstupního signálu spadne pod hodnotu threshold. Tato hodnota musí být pečlivě nastavena, aby nedošlo ke zkreslení signálu zesilováním krátkých tichých úseků signálu.

- *Look-Ahead*

Look-Ahead parametr je další časová konstanta kompresoru. Pokud máme offline aplikaci kompresoru, je možné tohoto parametru využít pro lepší reakci kompresoru na změnu vstupního signálu. Do paměti kompresoru se načte vzorek signálu navíc o délce look-ahead parametru a je analyzován dopředu. Pokud signál překročí threshold, bude o tom kompresor vědět dříve a může nastavit spuštění komprese s větší přesností. Většinou se nastavuje na polovinu attack time.

- *Frekvenční pásmo*

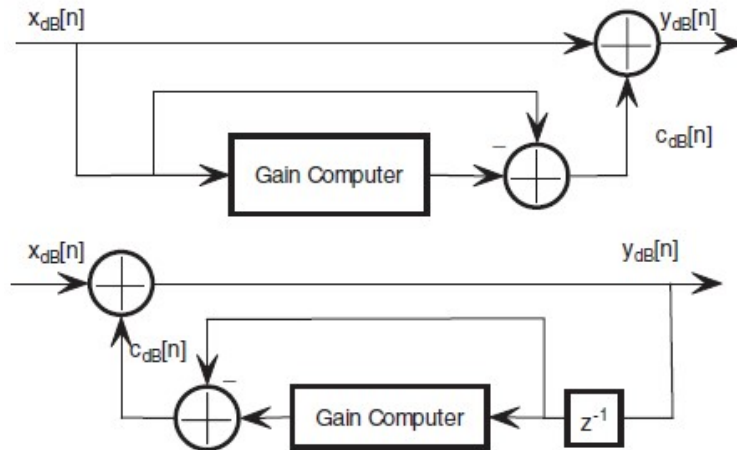
Některé aplikace vyžadují použití kompresoru pouze na určitou frekvenční oblast signálu. Proto je možné před kompresor vložit banku filtrů, která signál rozdělí na frekvenční pásma a u těch je potom možné nastavovat parametry komprese jednotlivě. Druhý způsob frekvenčního kompresoru je přidání filtru před detektor. Řídicí úroveň pro VCA je pak závislá na signálu výstupu filtru a komprimuje se podle ní celý signál procházející napětím řízeným zesilovačem – viz dále.

3.4 Uspořádání kompresoru

V této podkapitole je představen návrh kompresoru a uspořádání výpočetního řetězce.

3.4.1 Zpětnovazební a dopřednovazební návrh

Výpočetní řetězec může být uspořádán do dvou topologií, zpětnovazební a dopřednovazební. Ve zpětnovazební konfiguraci je vstupem do výpočetního řetězce již zesílený signál. V dopřednovazební situaci tomu tak není, a tak je vstup přiváděn přímo na výpočet zisku. To znamená, že se tento počítač musí vyrovnat s celým dynamickým rozsahem vstupního signálu. Většina moderních kompresorů je dopřednovazebního typu. Oba typy jsou zobrazeny na obrázku 5.



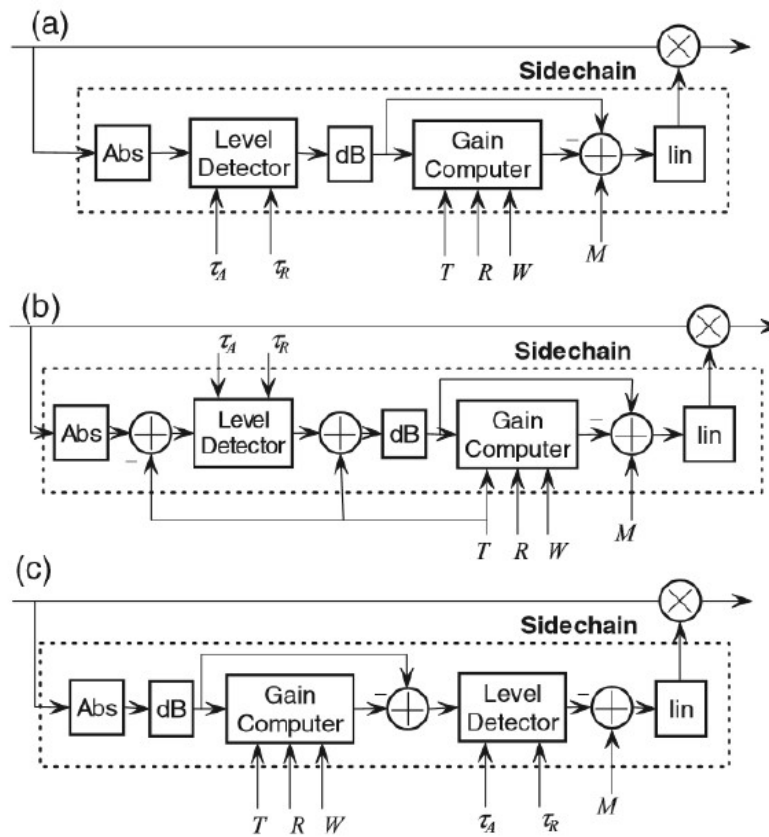
Obrázek 5 - Zpětnovazební (nahore) a dopřednovazební (dole) uspořádání výpočetního řetězce (převzato z [1])

3.4.2 Umístění detektoru ve výpočetním řetězci

Detektor lze v kompresoru umístit na několik míst v řetězci zpracování vstupního signálu. V prvním případě můžeme detektor umístit přímo na vstup řetězce, tedy tam, kde je signál před převodem do decibelové škály a před výpočtem zesílení. Zesílení se v tomto případě začne aplikovat až po překročení thresholdu. Problém nastává v případě poklesu signálu pod threshold, kde dojde k nespojitosti v doběhu signálu, protože řídicí napětí bude nulové. Blokové schéma je znázorněno na obrázku 6a.

Pokud odečteme hodnotu thresholdu od vstupního signálu a přičteme ji až po detekci, vyhneme se předchozímu problému – po klesnutí pod hodnotu threshold se řídicí napětí vrátí na hodnotu threshold. Toto řešení ale není vhodné pro nastavení soft knee, protože při ní není určena hodnota thresholdu jako jedno číslo. Na obrázku 6 označeno (b).

Nejlepším způsobem se ukázalo umístění detektoru do logaritmické oblasti (odkaz [1]), kde detektor reaguje na kontrolní napětí místo vstupního signálu, viz obrázek 6c. V případě nulového kontrolního napětí, tedy při neaktivované kompresi, je release time nezávislý na zvoleném parametru threshold.



Obrázek 6 - Blokové schéma uspořádání kompresoru. (převzato z [1]) (a) detekce z absolutní hodnoty, (b) detekce s parametrem threshold, (c) detekce v logaritmické oblasti

4. Parametr STI

Parametr STI (anglicky *Speech Transmission Index*), definovaný podle normy IEC 60268-16 (odkaz [2]), nám umožňuje objektivně posoudit srozumitelnost řeči v daném prostoru. To lze využít u prostorů, kde je požadována určitá míra srozumitelnosti, jako například veřejná místa s rozhlasem – stanice metra nebo vlaku, požární rozhlas na vesnici nebo ve městě, či velké sály, divadla, kongresová centra. Jedná se o objektivní metodu, z čehož vyplývá, že není potřeba velkého počtu lidí na provedení jako u subjektivního testování. Výsledky takovéto metody však nemohou být považovány za absolutní, i když je tato metoda navržena tak, aby se co nejvíce přiblížila výsledkům subjektivního testování. Tato metoda však neslouží k hodnocení maskování řeči, tedy skrývání řeči před nechtěnými posluchači, nebyla k tomu navržena.

4.1 Historie STI

Metoda STI byla vyvinuta pány Steeneken a Houtgast [6] v roce 1980, vyzkoušena byla ovšem pouze pro holandštinu. Proto byla tato metoda dále zkoumána pro použití v anglickém jazyce. Výsledky tohoto zkoumání byly pozitivní, viz [7]. Později v roce 2001 byl tento výzkum validován původními autory STI a bylo zjištěno nejen, že upravený parametr STI funguje pro anglický jazyk, ale i že hodnota STI koreluje se subjektivním měřením CVC-word score (Consonant Vowel Consonant). Dále se autoři zabývají použitím STI pro nelineární systémy, kde shledávají, že kromě jednoho typu zkreslení nemají nelinearity systému významný vliv na tento parametr. Více informací lze nalézt v článku [8].

4.2 Princip výpočtu

Metoda STI byla vytvořena jako rychlá objektivní metoda určující kvalitu přenosu řeči jedním nebo více kanály. Řečový signál obsahuje určité intenzitní obálky, a to pomalu se měnící, korespondující se slovy a hranicemi vět, a rychle se měnící, určující jednotlivé fonémy ve slovech. Bylo vyzorováno, že právě uvedené fluktuace intenzity nesou největší část informace lidské řeči. Parametr STI potom určuje, jak moc jsou tyto obálky ovlivněny průchodem komunikačního kanálu. Základním prvkem měření je Modulační přenosová funkce (MTF, *Modulation Transfer Function*), která umožňuje změřit tento vliv kanálu. Pokud se modulační spektrum změní

průchodem kanálu, a to většinou změnou hloubky modulace u jedné či více modulačních frekvencí, dojde ke snížení srozumitelnosti. Výsledkem měření STI je číslo od 0 do 1, které se vypočítá jako přínos všech frekvenčních pásem nacházejících se v řeči.

4.3 Metody měření

Obecně lze metody měření STI rozdělit na dvě skupiny: nepřímou a přímou.

4.3.1 Nepřímá metoda

Nepřímá metoda výpočtu parametru STI spočívá ve změření impulzní odezvy systému. Parametr $m_k(f_m)$ je modulační přenosová funkce s modulační frekvencí f_m v oktávovém pásmu k . Impulzní odezva eliminuje hluk pozadí, proto ho v druhé části výpočtu přičteme pomocí poměru signálu k šumu. Výpočet je následující:

$$m_k(f_m) = \frac{\left| \int_0^\infty h_k(t)^2 e^{-j2\pi f_m t} dt \right|}{\int_0^\infty h_k(t)^2 dt} \cdot \left(1 + 10^{-\frac{SNR_k}{10}} \right)^{-1}, \quad (9)$$

kde $h_k(t)$ je impulzní odezva v oktávovém pásmu k , f_m je modulační frekvence a SNR_k je poměr signálu k šumu v dB.

Nejčastějším způsobem měření impulzní odezvy je buď pseudonáhodný MLS signál (*maximum length sequence*), nebo TSP (*time stretched pulse*). Tato práce se však bude zabývat pouze přímou metodou. Více informací k nepřímé metodě lze nalézt v literatuře [3].

4.3.2 Přímá metoda

Testovací signál metody Full STI se skládá ze sedmi nosných půloktávových šumových signálů se středními frekvencemi od 125 Hz do 8 kHz. V tabulce 1 jsou uvedeny hladiny těchto půloktávových pásem, normalizované k vážené hladině 0 dB podle křivky A. Efektivní hodnota každého modulovaného šumového pásma odpovídá průměrné úrovni dlouhodobého spektra řeči. Každé pásmo je pak modulováno jednou frekvencí z třetinooktávových intervalů, začínajících na 0,63 Hz až do 12,5 Hz. Celkem tedy získáme 98 měření, odpovídajících 14 modulačním frekvencím a 7 modulovaným pásmům.

Tabulka 1 – Hladiny oktávných pásem měřícího signálu STI (převzato z [2])

Oktávné pásmo (Hz)	125	250	500	1000	2000	4000	8000
Muži (dB)	2,9	2,9	-0,8	-6,8	-12,8	-18,8	-24,8
Ženy (dB)	-	5,3	-1,9	-9,1	-15,8	-16,7	-18,0

4.3.3 Přímá metoda – zjednodušené postupy

Jelikož měření metodou Full STI je časově náročné (98 měření x 10 sekund, celkově tedy cca 15 minut), byly vyvinuty dvě další metody vycházející z Full STI, STITEL a STIPA. Vyvinuta byla i metoda RASTI, ta je ale v současné době považována za zastaralou. STI signál může být zjednodušen, pokud se z měření vyřadí modulace potřebné k přesné identifikaci nelineárních zkreslení. To nám umožní paralelní zpracování všech frekvenčních pásem najednou. Tím získáme krátký čas měření, ale ztratíme možnost přesnějšího měření. Měření se zjednoduší na jeden signál, ve kterém je více modulačních frekvencí najednou.

- RASTI

Room Acoustics STI. Skládá se ze dvou modulovaných šumových pásem se střední frekvencí 500 Hz a 2000 Hz. Celkově se v něm nachází 9 modulačních frekvencí (4 pro pásmo 500 Hz a 5 pro pásmo 2000 Hz), aplikovaných najednou na jednotlivá šumová pásma. Jedno měření trvá cca 30 sekund.

- STITEL

Použití STITEL je zaměřeno téměř výhradně pro telekomunikační kanály. Signál je složený ze sedmi modulovaných pásem, každé modulované jednou modulační frekvencí. Jedno měření trvá cca 12 sekund.

- STIPA

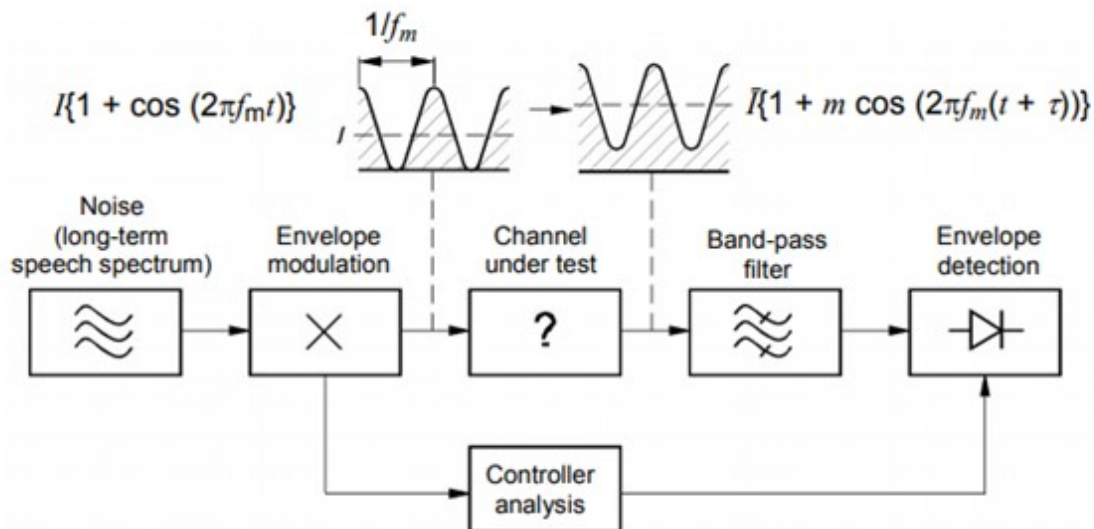
Speech Transmission Index for Public Address Systems. Signál se skládá ze sedmi modulovaných pásem a dvou modulačních frekvencí pro každé pásmo, viz tabulka 2. Poměr mezi první a druhou modulační frekvencí je 5. Hloubka modulace pro každou modulační frekvenci je 0,55 a mezi první a druhou modulační frekvencí je fázový rozdíl 180°. Takto vygenerovaný signál je založen na mužském řečovém spektru. Doba trvání jednoho měření je mezi 10 a 15 sekundami.

Tabulka 2 - Modulovaná pásma a jejich modulační frekvence pro metodu STIPA (převzato z [3])

Oktávové pásmo (Hz)	125	250	500	1000	2000	4000	8000
První modulační frekvence (Hz)	1,60	1,00	0,63	2,00	1,25	0,80	2,50
Druhá modulační frekvence (Hz)	8,00	5,00	3,15	10,00	6,25	4,00	12,50

4.4 Výpočet STI metodou STIPA

Měřicí řetězec STIPA začíná širokospektrálním růžovým šumem, který je pomocí banky filtrů rozdělen na jednotlivá oktávová pásma. Tato pásma jsou dále modulována modulačními frekvencemi podle tabulky 2. Všechna oktávová pásma jsou pak složena do jednoho signálu, který je poslán testovaným kanálem. Výstupní signál je pak analyzován a je vypočtena hodnota STI, viz dále. Na obrázku 7 lze vidět celý měřicí řetězec znázorněný blokovým schématem.



Obrázek 7 - Měřicí systém STI (převzato z [2]). Blokové schéma měření metodou STIPA.

Vstupními parametry pro výpočet parametru STI přímou metodou jsou čistý STIPA signál a STIPA signál po průchodu kanálem, tedy v našem případě signál po kompresi. Nejprve je potřeba si navrhnout banku filtrů shodnou s tou, která byla použita k výrobě STIPA signálu. Každé modulované pásmo je vhodné zkoumat samostatně, aby zde nedocházelo k různým rušením z ostatních pásem. Navíc tímto způsobem lze vidět ve spektru pouze ty modulační frekvence, které jsou použity pro toto pásmo.

Dalším krokem je normalizace rozsahu vstupního a výstupního signálu. Jelikož kompresor bez použití make-up gainu signál obecně zeslabuje, je nutné toto vykompenzovat normalizací hlasitosti pomocí efektivní hodnoty signálu. Toto zajistí zjednodušení výpočtu – pokud mají oba signály stejnou střední hodnotu, nemusí se v kroku výpočtu poměru modulačního indexu počítat hloubka modulace dané modulační frekvence, ale stačí hodnota amplitudy těchto frekvencí.

Funkce modulačního přenosu (MTF, *Modulation transmission function*) je soubor poměrů modulačních indexů všech modulačních frekvencí pro všechna půloktávová pásma. Poměr modulačních indexů se vypočítá jako podíl

$$m_{k,f_m} = m_{k,f_m,i} / m_{k,f_m,o},$$

kde $m_{k,f_m,i}$ je modulační index pro půloktávové pásmo k a modulační frekvenci f_m před průchodem kanálu a $m_{k,f_m,o}$ je ten samý modulační index, akorát pro signál po průchodu kanálu. Následuje implementace výpočtu MTF jednotlivých modulací v tomto kompresoru. Signál filtrovaný na dané pásmo z předchozího kroku se převede do frekvenční oblasti pomocí FFT. Poté se naleznou dvě modulační frekvence

dané pro zkoumané pásmo a vypočte se poměr jejich amplitud, komprimovaný signál k čistému signálu. Provedením tohoto kroku pro každé modulované pásmo dostaneme celou modulační přenosovou funkci.

Nyní přijde na řadu jednoduchý převod MTF na STI. Hodnoty MTF se převedou na efektivní hodnotu SNR podle následujícího vzorce:

$$SNR_{eff\ k, f_m} = 10 \cdot \lg \left(\frac{m_{k, f_m}}{1 - m_{k, f_m}} \right), \quad (10)$$

kde m_{k, f_m} je hodnota MTF pro oktávové pásmo k a modulační frekvenci f_m . Jelikož tato hodnota může vyjít nekonečno, je škála této veličiny omezena na interval od -15 dB do +15 dB. Dále se tato veličina upraví tak, aby její škála byla od 0 do 1. Tím vznikne veličina transmission index (TI):

$$TI_{k, f_m} = \frac{SNR_{eff\ k, f_m} + 15}{30}, \quad (11)$$

kde $SNR_{eff\ k, f_m}$ je efektivní poměr signálu k šumu pro každé oktávové pásmo k a modulační frekvenci f_m , vypočítaný podle vzorce (10).

Hodnoty TI jsou dále průměrovány přes všechny modulační frekvence v daném modulovaném pásmu. Takto vznikne modulation transfer index (MTI_k) pro oktávové pásmo k .

$$MTI_k = \frac{1}{n} \sum_{m=1}^n TI_{k, f_m}, \quad (12)$$

kde TI_{k, f_m} je transmission index pro každé oktávové pásmo k a modulační frekvenci f_m , m je index modulační frekvence a n je počet modulačních frekvencí na oktávové pásmo (v případě STIPA vždy 2).

Posledním krokem je aplikace koeficientů α a β . Parametr α je koeficient pro MTI z daného půloktávového pásma a β je koeficient redundance mezi daným pásmem a pásmem následujícím. Jednotlivé váhovací parametry jsou uvedeny v tabulce 3. Následuje konečně výpočet STI:

$$STI = \sum_{k=1}^7 \alpha_k \cdot MTI_k - \sum_{k=1}^6 \beta_k \cdot \sqrt{MTI_k \cdot MTI_{k+1}}, \quad (13)$$

kde MTI_k je modulation transfer index pro oktávové pásmo k , α_k a β_k jsou váhovací a redundantní parametry podle tabulky 3.

Tabulka 3 - Váhovací (α) a redundantní (β) parametry MTI (převzato z [2])

Oktávové pásmo Hz		125	250	500	1 000	2 000	4 000	8 000
Muži	α	0,085	0,127	0,230	0,233	0,309	0,224	0,173
	β	0,085	0,078	0,065	0,011	0,047	0,095	-
Ženy	α	-	0,117	0,223	0,216	0,328	0,250	0,194
	β	-	0,099	0,066	0,062	0,025	0,076	-

5. Experimentální část

Tato kapitola se bude zabývat implementací kompresoru a měřicího systému STI v prostředí MATLAB. V další části se nachází výsledky různých měření vlivu kompresoru na hodnotu STI měřenou přímou metodou STIPA. Na konci se také zkoumá vliv aditivního šumu při cestě kanálem na výsledky měření STIPA.

5.1 Výroba kompresoru v prostředí Matlab

Implementace kompresoru byla poměrně jednoduchá. Prvním krokem zpracování signálu je jeho převod do decibelové škály. Tento krok je proveden v souladu s příručkou k výrobě kompresoru dle [1].

Okamžitá hodnota signálu se nikdy neudává v decibelové škále. Threshold by také mohl být uveden v absolutních jednotkách, nicméně tato hodnota není tolik uživatelsky přívětivá, jako hodnota v decibelech. Převedení do logaritmické oblasti se provádí pouze pro detekci překročení tresholdu, samotné zeslabení signálu se aplikuje přímo na vstupní signál v lineární oblasti.

Další na řadu přichází detekce překročení tresholdu, která určí gain pro daný vzorek, a to v závislosti na aktuální hodnotě (peak detektor), nebo na efektivní hodnotě (RMS detektor). Vypočtený gain také závisí na typu zalomení převodní charakteristiky. Výpočet proběhne dle vzorce (7) pro hard knee kompresor, nebo podle vzorce (8) pro soft knee kompresor (viz kapitola 3.3).

Velmi malé hodnoty signálu je vhodné upravit. Pokud by se amplituda signálu v některém vzorku rovnala nule, v logaritmické oblasti by se objevila hodnota mínus nekonečno, což pro další výpočty není vhodné. Proto je v tomto kompresoru nastaven pevný limit hladiny vstupního signálu, který je nastavený na -90 dB. To znamená, že všechny hodnoty menší, než -90 dB budou nastaveny na tuto hodnotu. Hodnota -90 dB byla zvolena podle úrovně šumu v nejvyšším oktávovém pásmu signálu STIPA (8000 Hz).

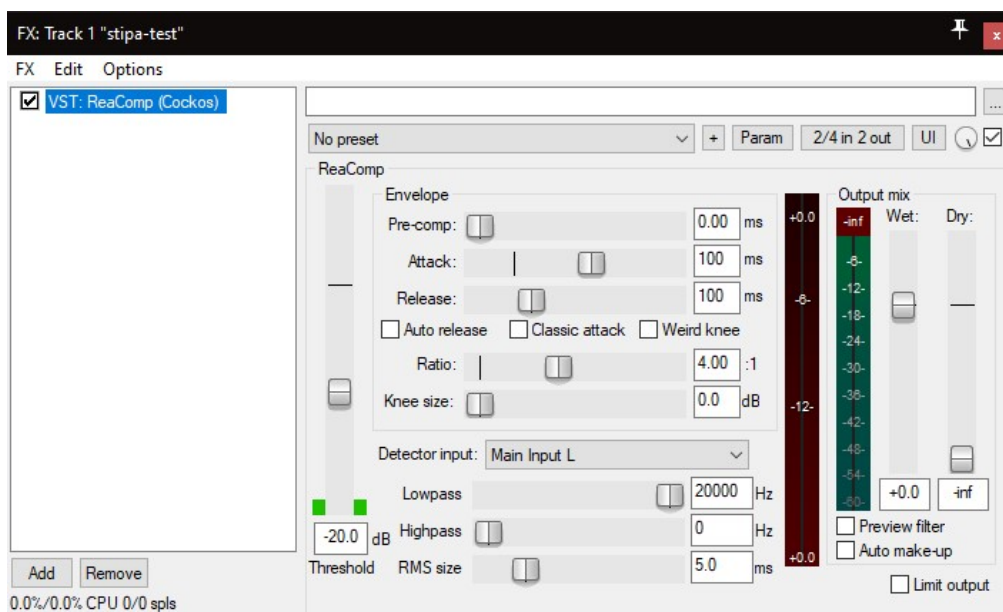
Následuje aplikace attack a release time podle vzorce (5), viz kapitola 3.1.2. Po aplikaci attack a release time následuje převod gainu z decibelové škály do lineární oblasti. Výstupní signál je tedy vytvořen vynásobením vstupního signálu gainem v lineární oblasti.

5.2 Vliv různých parametrů kompresoru na hodnotu STI

Jelikož má kompresor řadu nastavení, je dobré se podívat na vliv každého parametru zvlášť, tedy všechny ostatní parametry musí být konstantní, abychom dostali vypovídající výsledek. Předmětem zkoumání bude attack time, release time, threshold, typ detektoru, délka RMS okna u RMS detektoru a délka zkoumaného audio souboru. Pro všechna měření byla použita vzorkovací frekvence 44100 Hz.

K porovnání výsledků s jiným kompresorem byl zkoumán plug-in kompresor softwaru REAPER, který slouží pro profesionální úpravu zvukových nahrávek. Grafické rozhraní tohoto kompresoru lze vidět na obrázku 8. Je potřeba si všimnout několika detailů. Prvním z nich je nastavení RMS size, které napovídá, že tento kompresor používá RMS detektor. Dále jsou zde nastavení dvou filtrů, Highpass a Lowpass, kterými lze vybrat pouze určité frekvenční pásmo ke kompresi. Pro účely dalších experimentů tato nastavení zůstanou v krajních polohách, tedy kompresor bude zabírat celé frekvenční pásmo.

Problém, který může nastat, je hodnota, ke které je vztažen parametr Threshold. V implementaci v MATLABu je Threshold vztažen k hodnotě signálu 1 pro dynamický rozsah -1 až 1. Experimentem jsem ověřil, že stejný výsledek pro porovnání mezi jednotlivými kompresory vyjde při nastavení thresholdu REAPER kompresoru o 1 dB výše, než MATLAB kompresoru. Například nastavení thresholdu v MATLABu -15 dB odpovídá nastavení -14 dB v REAPERu. S ohledem na tuto skutečnost budou další výsledky prezentovány v hladinách MATLABového kompresoru,



Obrázek 8 - Pluginový kompresor softwaru REAPER

u výsledků REAPER kompresoru je tedy nutné pro reprodukci výsledků použít upravenou hodnotu.

Poslední zajímavou částí je pozorování vlivu aditivního šumu na výsledky měření STIPA při použití kompresoru. Nejčastější typ rušení je růžový šum, proto se měření zaměřilo na tento typ šumu. Nejhorší případ pro měření STIPA je přidaný šum s podobným spektrem, jako má měřicí signál. Pro měření byl tedy použit růžový šum filtrovaný IIR filtrem typu Butterworth prvního řádu se zlomem na frekvenci 500 Hz. Jeho koeficienty přenosové funkce jsou $a = 1$ a $-0,9312$ a $b = 0,0344$ a $0,0344$ pro vzorkovací frekvenci 44100 Hz. Tím vznikne šum s podobnou spektrální charakteristikou, jako má měřicí signál STIPA (podle tabulky 1 v kapitole 4.3.2). RMS hodnota přidaných šumů byla zvolena tak, aby poměrově odpovídala RMS hodnotě původního měřicího signálu. Přidával se tak šum s hodnotami $\frac{1}{4}$ původní RMS (-12 dB), $\frac{1}{2}$ původní RMS (-6 dB), stejné hodnoty RMS (0 dB) a dvojnásobné hodnoty RMS (6 dB). Hodnota RMS testovacího signálu STIPA je v tomto případě 0,1755. K této hodnotě jsou vztaženy hladiny RMS šumových signálů použité v grafech.

5.2.1 Vliv délky testovacího souboru na hodnotu STI

K testování byl použit STIPA signál o délce 60 sekund. Norma uvádí délku jednoho měření 10 – 15 sekund, podle požadované přesnosti. Předmětem tohoto experimentu bylo určit, zda se nějak výrazně liší hodnota STI vypočítaná z dlouhého segmentu signálu a hodnota STI vypočítaná z krátkého signálu. Pro měření byly použity tyto parametry:

- peak detektor,
- threshold -25 dB,
- kompresní poměr 4:1,
- attack time 50 – 250 ms,
- release time 50 – 250 ms.

Výsledkem experimentu byla konstantní hodnota STI pro všechny délky RMS okna. Lehký rozdíl byl v hodnotách STI mezi jednotlivými nastaveními attack time a release time, ty jsou však v souladu s následujícími experimenty. Pro další pokusy byl tedy použit záznam o délce 15 sekund pro ušetření času potřebného k výpočtu STI. Vhodné je zmínit, že při aplikaci v reálném prostředí by mohla kratší doba

zkoumaného signálu způsobit větší náchylnost na možné impulzní rušení vyskytující se v tomto prostoru.

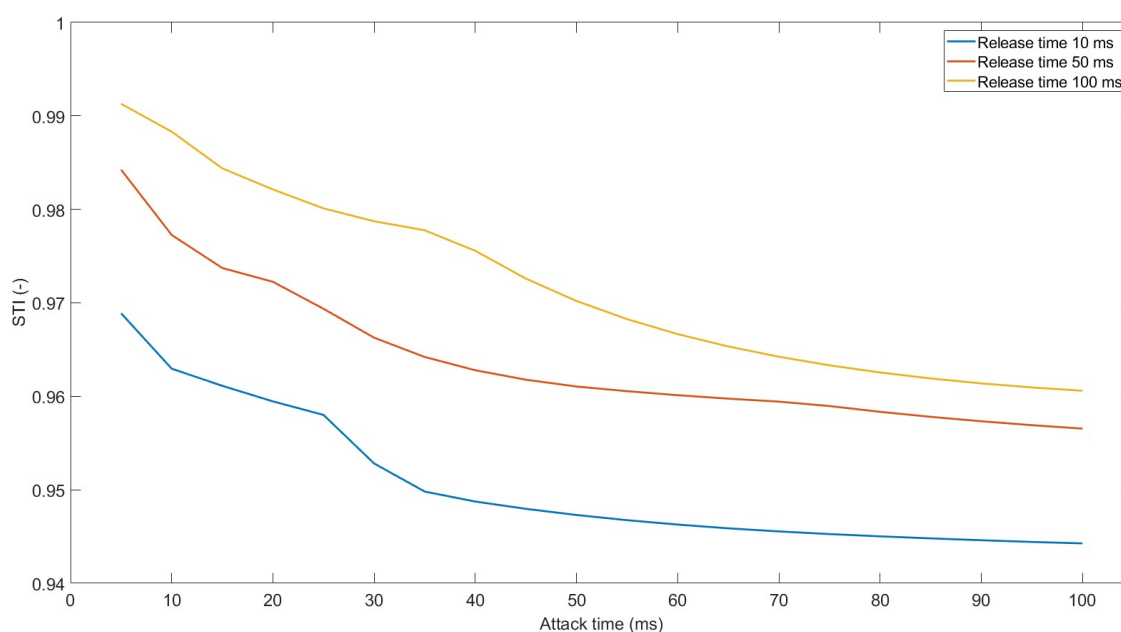
5.2.2 Vliv attack time a release time na STI

Attack time a release time jsou dvě základní časové konstanty kompresoru. Následující experiment měl ověřit jejich vliv na vypočtenou hodnotu STI. Závislost STI na attack time je základní metodou všech následujících měření, u každého z nich se ale mění porovnávaná veličina.

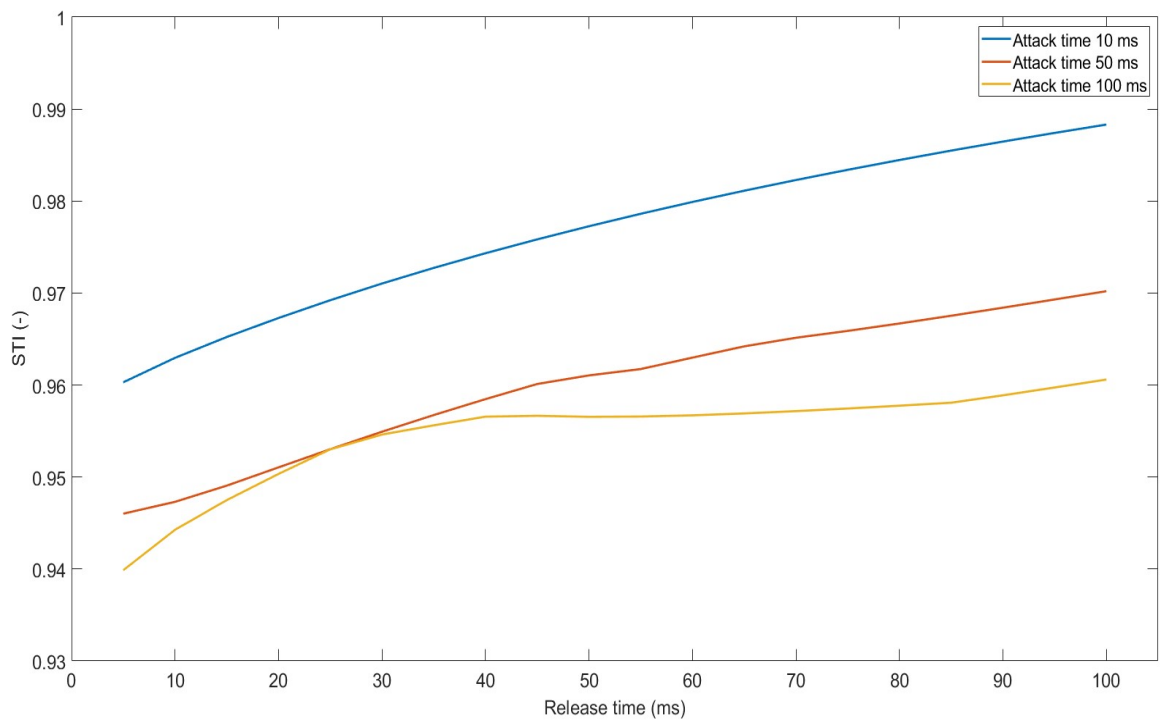
Prvním měřením této série je porovnání vlivu release time na vývoj hodnoty STI s narůstajícím attack time. Parametry kompresoru byly nastaveny takto:

- peak detektor,
- threshold -25 dB,
- kompresní poměr 4:1,
- attack time 5-100 ms,
- release time 10, 50 a 100 ms.

Na obrázku 9 lze vidět výsledek tohoto měření. Z obrázku vyplývá, že rostoucí hodnota parametru attack time má negativní vliv na STI, zatímco se zvyšujícím se parametrem release time hodnota STI stoupá. Toto lze ověřit i na obráceném experimentu, tedy zkoumání závislosti STI na parametru release time pro dané hodnoty attack time (viz obrázek 10).



Obrázek 9 - Závislost parametru STI na délce release time pro dané hodnoty attack time



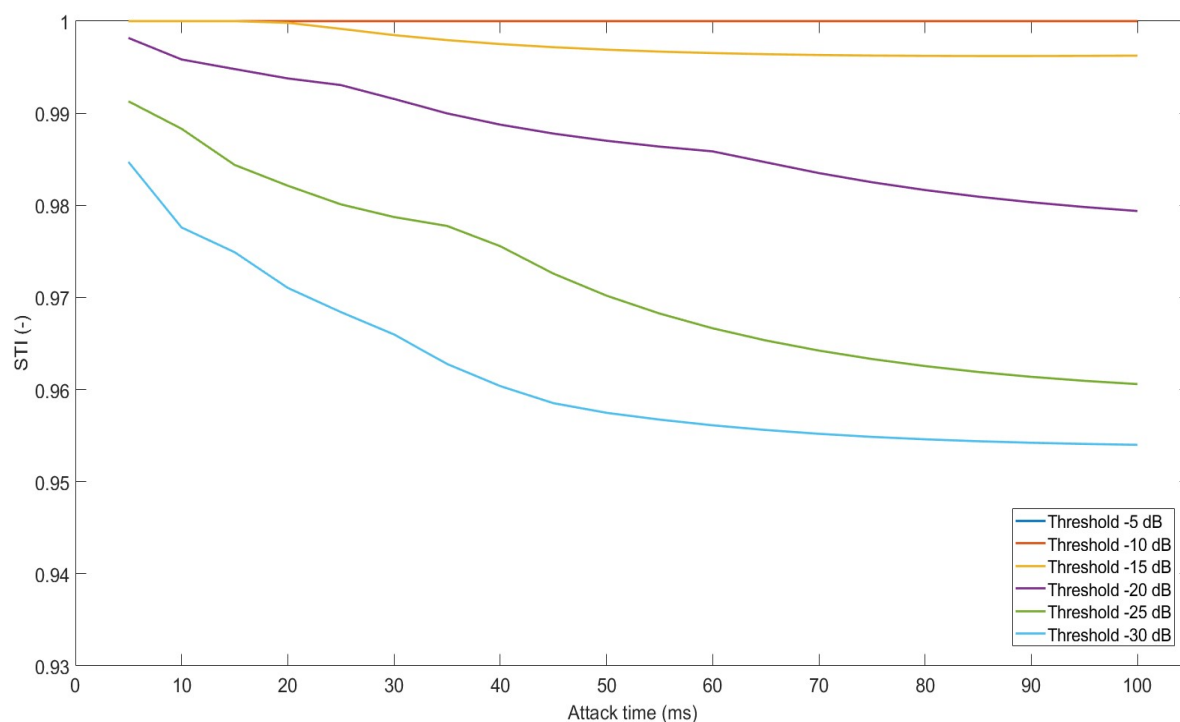
Obrázek 10 - Závislost parametru STI na délce attack time pro dané hodnoty release time

5.2.3 Vliv parametru threshold na hodnotu STI

Tato část se bude věnovat vlivu thresholdu na hodnotu STI. Čím nižší je hodnota threshold, tím se aplikuje více gainu na větší část signálu. Proto předpokládáme, že s nižší hodnotou threshold bude nižší i hodnota STI. Nastavení kompresoru bylo následující:

- RMS detektor,
- threshold -5 až -30 dB,
- kompresní poměr 4:1,
- attack time 5 – 100 ms,
- release time 100 ms.

Z obrázku 11 je patrné, že nastavení threshold -5 dB a -10 dB nemá na parametr STI žádný vliv. Dále se snižující se hodnotou threshold se snižuje i hodnota STI.

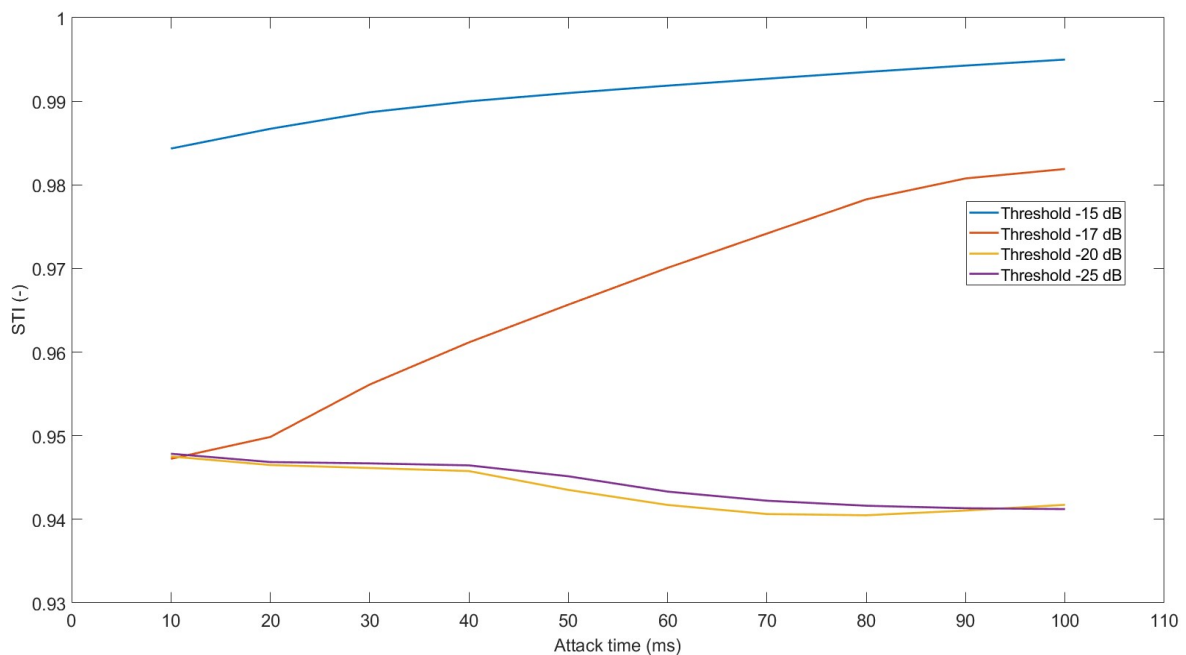


Obrázek 11 - Závislost parametru STI na hodnotě attack time pro dané hodnoty threshold

Na obrázku 12 lze vidět závislost STI na délce attack time pro pluginový kompresor. Z jeho průběhu je jasné, že zpracování signálu se v tomto kompresoru liší od toho implementovaného v MATLABu. Jelikož nemáme přístup k dokumentaci funkce tohoto kompresoru, není možné přesně zjistit, proč se výsledky tohoto měření liší od

měření na kompresoru v MATLABu. Také je patrné, že v reálném využití, tedy při použití neznámého kompresoru, může být dopad na srozumitelnost větší, než při použití přesného modelu kompresoru. MTF při výpočtu STI jak u pluginu, tak MATLABu (tabulka 4) toto pozorování potvrzuje – pluginový kompresor ovlivňuje modulační poměry více, než MATLABový kompresor.

Zajímavé je chování REAPER kompresoru pro hodnoty threshold větší, než -20 dB, kde hodnota STI v závislosti na attack time roste místo toho, aby klesala. Důvodem může být přechod hranice mezi částečnou kompresí (hladina signálu se pohybuje těsně okolo zlomu převodní charakteristiky), kde se attack time uplatňuje častěji, a plnou kompresí (hladina signálu se již pohybuje za zlomem charakteristiky), kde se hodnota threshold překročí jednou a parametr attack time je tak použit pouze jednou.



Obrázek 12 - Závislost parametru STI na délce attack time pro dané hodnoty threshold, plugin kompresor

Tabulka 4 - Porovnání vlivu kompresorů na modulační indexy STIPA signálu (RMS detektor, threshold -20 dB, poměr 4:1, attack time 100 ms, release time 100 ms)

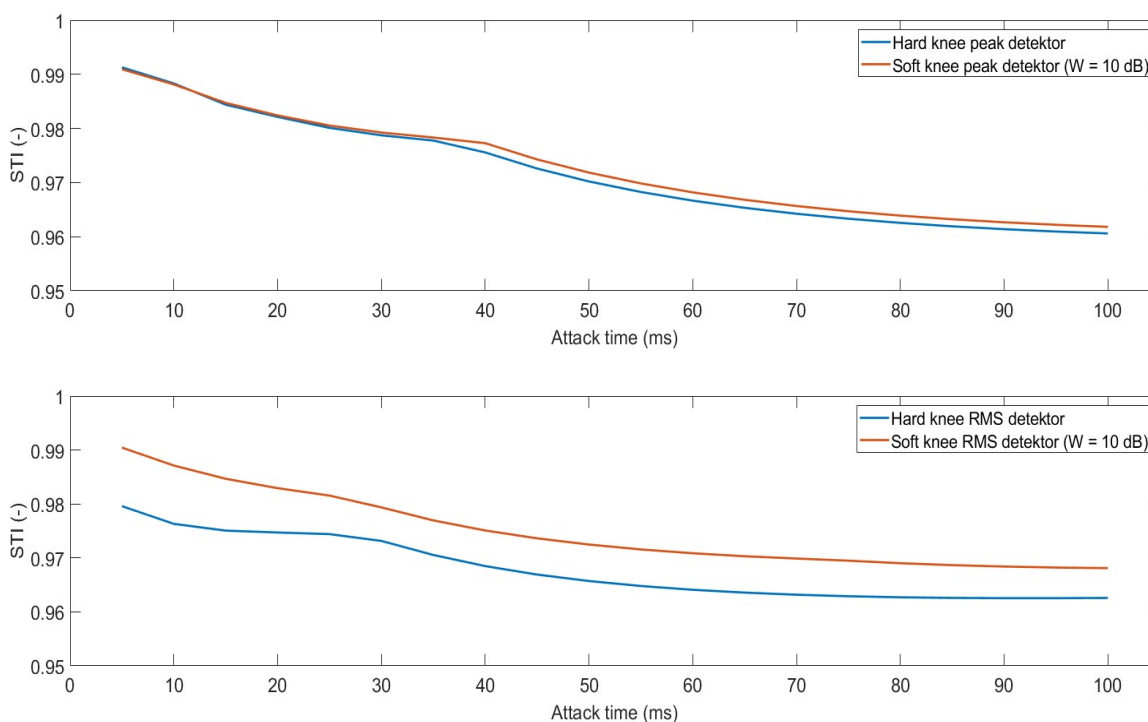
Plugin (STI = 0,94)							
Oktávové pásmo (Hz)	125	250	500	1000	2000	4000	8000
První modulační poměr (-)	0,87	0,73	0,87	1	0,91	0,94	1
Druhý modulační poměr (-)	0,99	1	1	1	1	1	1

MATLAB (STI = 0,97)							
Oktávové pásmo (Hz)	125	250	500	1000	2000	4000	8000
První modulační poměr (-)	0,93	0,84	0,94	1	0,96	0,97	1
Druhý modulační poměr (-)	0,99	1	1	1	1	1	1

5.2.4 Vliv detektoru a zalomení převodní charakteristiky

Dalším zkoumaným parametrem kompresoru byl jeho detektor a převodní charakteristika. Zde jde o porovnání všech čtyřech typů implementací kompresoru, tedy hard knee nebo soft knee kompresor a špičkový nebo RMS detektor. Parametry kompresoru byly následující:

- peak nebo RMS detektor,
- threshold -25 dB,
- kompresní poměr 4:1,
- attack time 5 – 100 ms,
- release time 100 ms.



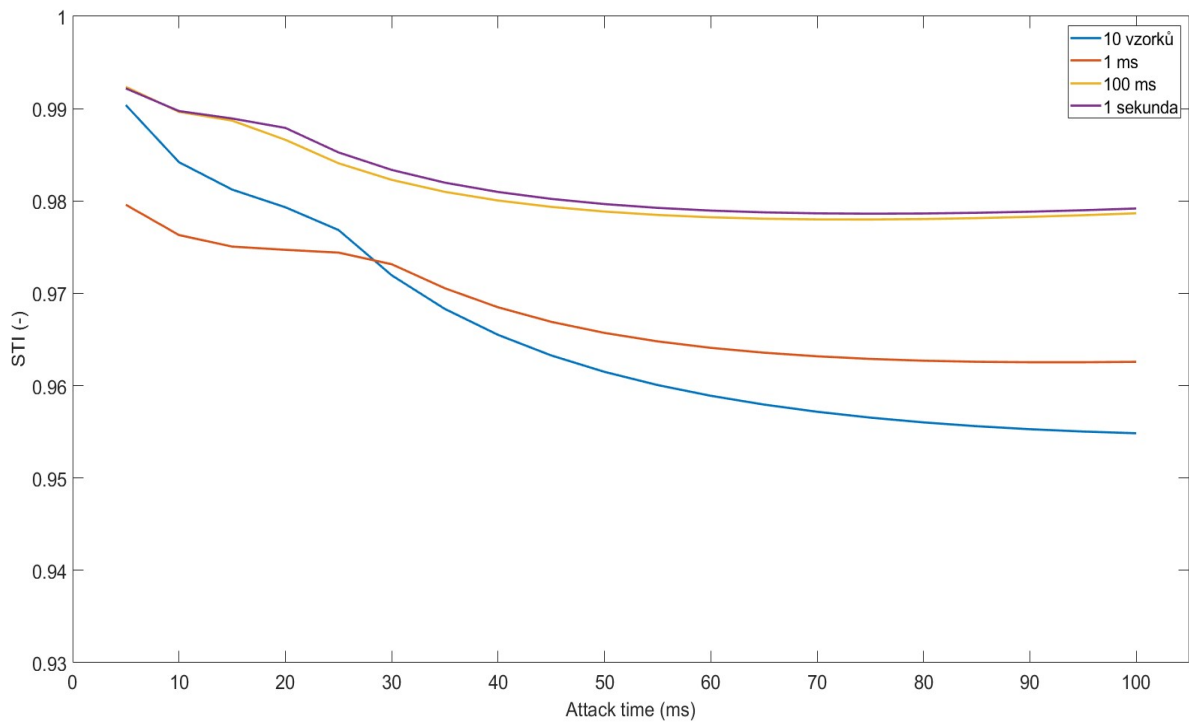
Obrázek 13 - Závislost parametru STI na délce attack time, porovnání hard/soft knee a peak (nahore)/RMS detektoru (dole)

Z obrázku 13 je vidět, že v případě peak detektoru nemá převodní charakteristika téměř žádný vliv na hodnotu STI. U RMS detektoru je rozdíl sice větší, ale pouze v řádu setin STI, což nevykazuje žádnou zásadní změnu ve srozumitelnosti. Průběhy peak a RMS detektoru jsou si velice podobné na téměř stejné úrovni, proto není nutné dělat všechny experimenty ve všech čtyřech variacích.

5.2.5 Vliv délky výpočetního okna efektivní hodnoty na STI

Co se týče RMS detektoru, je zde další parametr vhodný k prozkoumání, a to délka okna použitého k výpočtu efektivní hodnoty. Jelikož se jedná o offline implementaci, efektivní hodnotu lze vypočítat jak z předešlých, tak budoucích vzorků. Počet těchto vzorků je předmětem tohoto zkoumání. Předpokladem je symetrické okno okolo aktuálního vzorku. Nastavení kompresoru bylo následující:

- RMS detektor,
- threshold -15 dB,
- kompresní poměr 4:1,
- attack time 5 – 100 ms,
- release time 100 ms.



Obrázek 14 - Závislost parametru STI na délce attack time v závislosti na délce RMS okna

Obrázek 14 ukazuje výsledky pokusu s proměnnou délkou RMS okna. Je evidentní, že s narůstající délkou RMS okna stoupá hodnota STI. Samozřejmě s větší délkou okna bude změna RMS hodnoty menší v čase a kompresor tedy bude reagovat pomaleji na změny dynamiky. Proto se také nemění ani hodnota STI. Výpočetní doba se vysoce zvýšila při použití možnosti 100 ms a 1 s. Pro další využití je tedy vhodné použít RMS okno v řádu desítek milisekund, kvůli výpočetní náročnosti.

5.3 Vliv přidaného šumu v kanálu na měření

Pro přiblížení podmínek měření reálnému využití se lze také zabývat otázkou, jak dynamický kompresor ovlivní měření STI, pokud se v kanále vyskytne šumové rušení. V reálných podmínkách mohou vesměs nastat dva případy. Buď se šum vyskytne v místě zdroje signálu (například v kabině strojvedoucího), kdy do kompresoru vstupuje signál smíchaný se šumem pozadí, nebo se rušení objeví v místě poslechu (například ve voze pro cestující ve vlaku). V tomto případě se šum přidá k již komprimovanému signálu. V následující části se budou zkoumat tyto dva případy. Kompresor byl nastaven následujícím způsobem:

- peak detektor,
- threshold -20 dB,
- kompresní poměr 4:1,
- attack time 5 – 100 ms,
- release time 100 ms.

5.3.1 Zpracování v MATLABu

Zřejmým trendem v tomto měření je snižování hodnoty STI s vyšší hladinou přidaného šumu. To lze dokázat dosazením hodnot SNR do vzorečku (9) z nepřímého výpočtu STI. Jediné, co se mění, je přidání šumu, zlomek s integrály se pak vykrátí při dělení modulačních indexů. Nakonec vyjde vzorec:

$$STI = (1 + 10^{\frac{SNR}{10}})^{-1} \quad (14)$$

Jelikož byl použit šum s velice podobnou charakteristikou, jako signál STIPA, SNR bude ve všech frekvenčních pásmech stejné, proto stačí dosadit celkové SNR. SNR je v tomto případě záporná hodnota RMS šumu. V tabulce 5 jsou vypočítané teoretické hodnoty STI

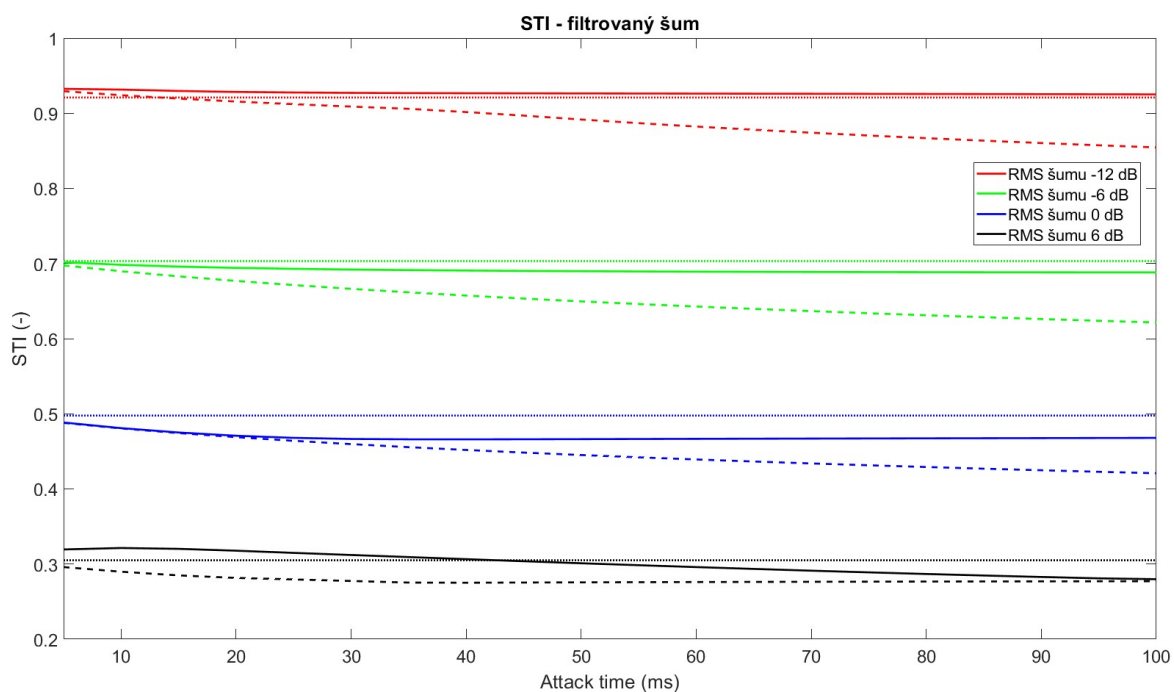
Tabulka 5 - Vypočítané a naměřené hodnoty STI pro signál s růžovým šumem podle vzorce (14)

RMS šumu	-12 dB	-6 dB	0 dB	6 dB
STI vypočítané	0,94	0,79	0,50	0,20
STI naměřené	0,91	0,72	0,57	0,44

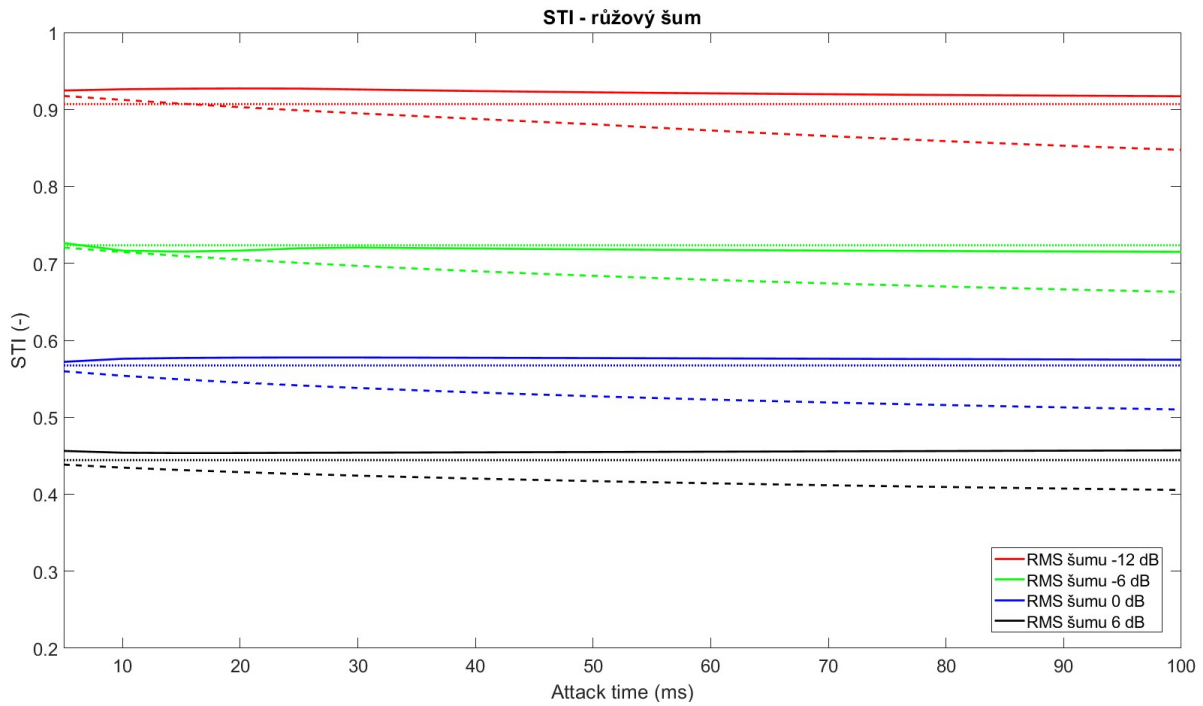
pro RMS šumu od -12 dB do 6 dB. Na obrázcích 15 a 16 lze vidět vliv samotného šumu na hodnotu STI (tečkovaná křivka), vliv kompresoru na zašuměný signál (plná křivka) a vliv šumu na komprimovaný signál (přerušovaná křivka) pro růžový a filtrovaný šum. Hodnoty STI pro filtrovaný šum bez komprese ne zcela odpovídají teoretickému výpočtu. Důvodem může být překrytí signálu šumem, viz text k obrázku 17. Lze vidět, že kompresor ovlivňuje zašuměný signál méně, než signál bez šumu. Přidání šumu ke komprimovanému

vanému signálu pouze sníží hodnotu STI, průběh závislosti STI na attack time zůstane stejný, jako u signálu bez šumu.

Lze si všimnout, že ne ve všech případech použití kompresoru zhoršuje skóre STI. Zatímco u šumu přidaného na konci řetězce se kompresor chová stejným způsobem jako bez šumu, v opačném případě je průběh STI nepředvídatelný a výsledkem může být vyšší STI. Z toho můžeme usoudit, že pokud se šum vyskytuje v poslechové místnosti, je jeho vliv na hodnotu STI vyšší, než kdyby byl ve vysílací místnosti.



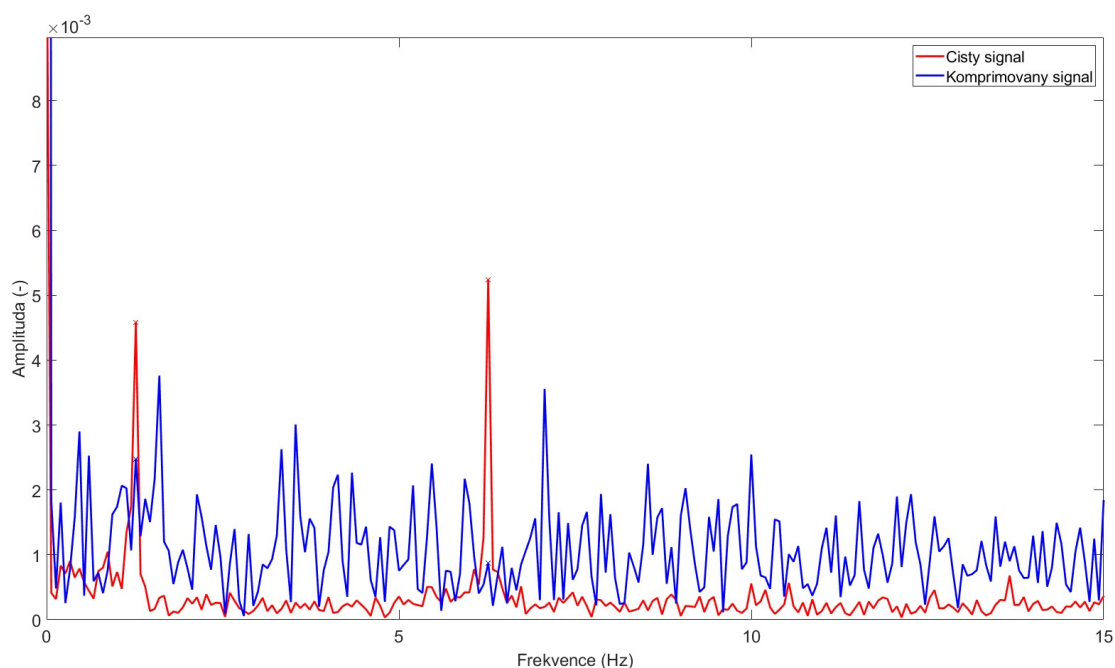
Obrázek 15 - Měření STI s přidaným filtrovaným šumem. Plná křivka představuje kompresi po přidání šumu, přerušovaná křivka kompresi před přidáním šumu. Tečkovaná křivka označuje referenční hodnotu STI pro přidaný šum bez komprese (není závislá na hodnotě attack time).



Obrázek 16 - Měření STI s přidaným růžovým šumem. Plná křivka představuje kompresi po přidání šumu, přerušovaná křivka kompresi před přidáním šumu. Tečkovaná křivka označuje referenční hodnotu STI pro přidaný šum bez komprese (není závislá na hodnotě attack time).

Zajímavé jsou výsledky pro vyšší hladiny přidaného šumu. Zde je vhodné se podívat na jednotlivé modulační frekvence zašuměného signálu. Nejvýznamnější pásma pro měření STI jsou podle tabulky 3 v kapitole 4.4 pásma 1000 Hz a 2000 Hz. Zkoumán bude nejhorší případ přidaného šumu, tedy filtrovaný šum s hodnotou RMS 6 dB. Obrázek 17 obsahuje amplitudovou spektrální charakteristiku komprimovaného STIPA signálu s výše uvedeným šumem v pásmu 2000 Hz. Je zřejmé, že měření v tomto případě dává nesmyslné hodnoty, jelikož přidaný šum zcela zakryl svojí energií měřící signál a nejsou tedy vidět ani špičky modulačních frekvencí tak, jako na původním STIPA signálu. Do výpočtu STI pak vstupují hodnoty šumu namísto modulačních indexů.

Tento problém se vyskytuje i u šumu s hodnotou RMS 0 dB, tedy stejnou jako původní signál. I zde zakrývá šum modulační indexy v nejdůležitějších pásmech. U RMS -6 dB tomu tak již není, a proto lze naměřené hodnoty považovat za platné.



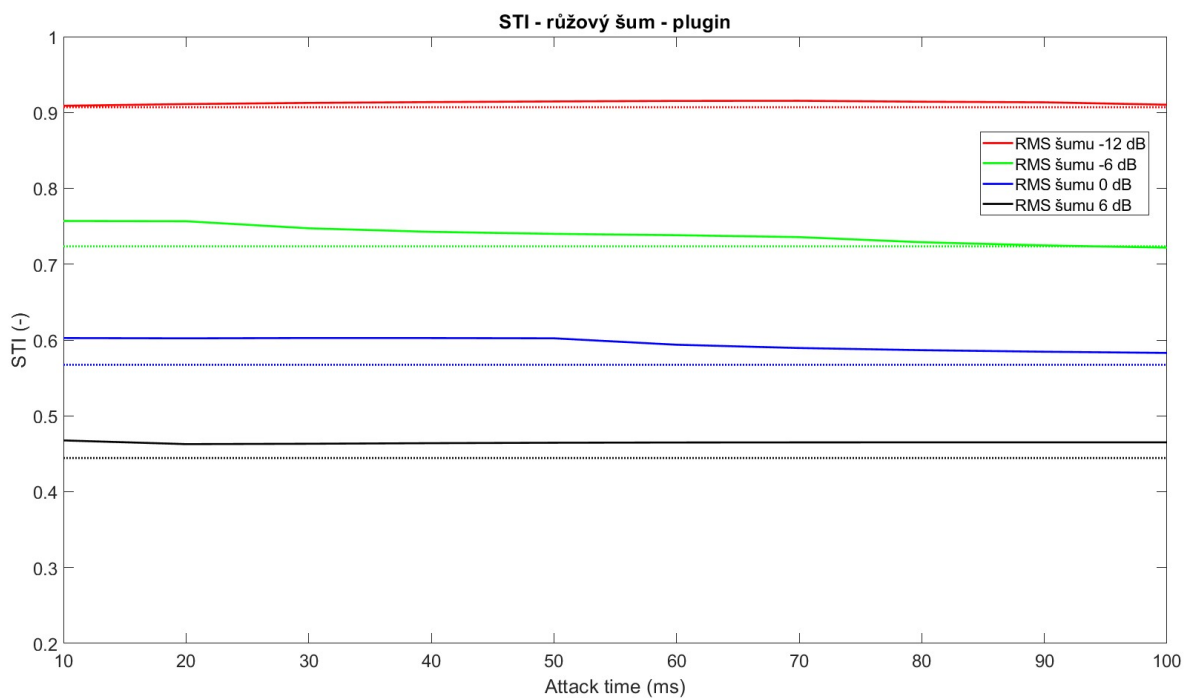
Obrázek 17 - Amplitudové spektrum čistého i zašuměného měřícího signálu v pásmu 2000 Hz

5.3.2 Zpracování v pluginu

Jelikož MATLAB není program na práci se zvukem, rozsah hodnot ve vektoru nemusí být omezen na interval od -1 do 1. To se hodí v tomto případě pro sčítání dvou signálů, kde nejsme omezeni žádnými hranicemi. Toto ovšem neplatí pro software REAPER, kde hodnoty musí být omezeny. Součet signálů byl proveden stále v MATLABu, nicméně funkce *audiowrite* použitá k uložení vektoru do souboru WAV hodnoty větší, než 1 omezí na hodnotu 1 a dojde ke clippingu. Proto je nutné vektor před uložením

normalizovat tak, aby jeho nejvyšší hodnota byla právě 1. Jelikož se normalizuje celý součet signálu a šumu, jejich poměr RMS hodnot zůstane stejný. Problémem je porovnání hodnot threshold v MATLABu a pluginu, jelikož se liší maximální hodnota, ke které je threshold vztahován. Proto je při výpočtech v MATLABu použit stejně normovaný signál, jako v pluginu. Tím se srovnají hodnoty threshold a výsledky jsou porovnatelné.

Po použití pluginového kompresoru se soubory načety zpět do MATLABu a vypočítaly se hodnoty STI pro stejné případy jako u MATLABu. Výsledky jsou velice podobné těm z MATLABu, liší se pouze o setiny STI. Výpočet se šumem přidaným před kompresí je vidět na obrázku 18.



Obrázek 18 - Měření STI s přidaným šumem na začátku řetězce, komprimovaný pluginem. Tečkovaná křivka označuje referenční hodnotu STI pro přidaný šum bez komprese (není závislá na hodnotě attack time).

6. Závěr

V této práci bylo ověřeno, že lze použít přímé metody výpočtu hodnoty STI pro signál upravený dynamickým kompresorem. Výsledky ukazují, že vliv kompresoru na hodnotu STI je v rámci jedné až dvou desetín STI, tedy jeho vliv není nijak závažný. Zkoumal se vliv délky analyzovaného signálu na hodnotu STI. Výsledkem je překvapivě neměnná hodnota STI u jakékoliv délky signálu od 1 sekundy do 30 sekund. Ukázalo se, že s rostoucí hodnotou attack time má hodnota STI tendenci klesat, zatímco s vyšší hodnotou release time naopak stoupá. Dle předpokladu má nižší hodnota thresholdu větší vliv na STI. Rozdíly mezi hodnotami STI u hard knee a soft knee kompresoru, popřípadě u peak a RMS detektoru jsou také malé. U RMS detektoru se osvědčila délka okna výpočtu efektivní hodnoty v řádu desítek milisekund. Při porovnání výsledků získaných v této práci s předchozí prací [3] lze konstatovat, že přímá metoda měření parametru STI dává stejné výsledky, jako měření metodou nepřímou.

Vliv aditivních šumů na hodnocení STI vyšel lépe, než jaký byl vypočítán pomocí teoretického výpočtu. Vyšší výkon šumu odpovídá menší hodnotě STI, avšak tato hodnota není přesná, z důvodu nízkého odstupů užitečného signálu od šumu v nejdůležitějších oktávových pásmech, s čímž výpočet nepočítá. Kompresor má pak v případě zašuměného signálu menší vliv na hodnotu STI, než při použití signálu bez šumu.

Využití měření STI přímou metodou společně s použitím kompresoru je v praxi možné, avšak výsledky získané měřením neznámého přístroje typu kompresor mohou být odlišné od hodnot naměřených v této práci s pomocí modelového kompresoru. Tato odlišnost je dána nám neznámým návrhem daného kompresoru, tedy nikdo kromě výrobce neví přesně, jak je toto zařízení typu kompresor vnitřně uspořádáno a jaké má nastavení parametrů, které nemusí být pro uživatele přístupné. Výsledky experimentů ale ukazují pouze malé odchylky STI od simulovaného kompresoru, a to jak v případě bez šumu, tak i s ním.

7. Seznam obrázků

Obrázek 1 - Stabilita nelineárního systému. Barevné křivky ukazují splnění podmínek stability.	14
Obrázek 2 - Blokové schéma dynamického kompresoru.....	16
Obrázek 3 - Jednoduchý peak detektor, sestavený z diody, dvou odporů a kondenzátoru (převzato z [1])	17
Obrázek 4 - Převodní charakteristika kompresoru (převzato z [1]). Hodnota threshold definuje převodní charakteristiku, se zlomem soft knee (čárkovaně) a hard knee (čerchovaně).....	20
Obrázek 5 - Zpětnovazební (nahore) a dopřednovazební (dole) uspořádání výpočetního řetězce (převzato z [1])	23
Obrázek 6 - Blokové schéma uspořádání kompresoru. (převzato z [1]) (a) detekce z absolutní hodnoty, (b) detekce s parametrem threshold, (c) detekce v logaritmické oblasti.....	24
Obrázek 7 - Měřicí systém STI (převzato z [2]). Blokové schéma měření metodou STIPA.	30
Obrázek 8 - Pluginový kompresor softwaru REAPER	35
Obrázek 9 - Závislost parametru STI na délce release time pro dané hodnoty attack time	37
Obrázek 10 - Závislost parametru STI na délce attack time pro dané hodnoty release time.....	38
Obrázek 11 - Závislost parametru STI na hodnotě attack time pro dané hodnoty threshold	39
Obrázek 12 - Závislost parametru STI na délce attack time pro dané hodnoty threshold, plugin kompresor	40
Obrázek 13 - Závislost parametru STI na délce attack time, porovnání hard/soft knee a peak (nahore)/RMS detektoru (dole)	42
Obrázek 14 - Závislost parametru STI na délce attack time v závislosti na délce RMS okna	43
Obrázek 15 - Měření STI s přidáním filtrovaným šumem. Plná křivka představuje kompresi po přidání šumu, přerušovaná křivka kompresi před přidáním šumu. Tečkovaná křivka označuje referenční hodnotu STI pro přidání šum bez komprese (není závislá na hodnotě attack time).....	45
Obrázek 16 - Měření STI s přidáním růžovým šumem. Plná křivka představuje kompresi po přidání šumu, přerušovaná křivka kompresi před přidáním šumu. Tečkovaná křivka označuje referenční hodnotu STI pro přidání šum bez komprese (není závislá na hodnotě attack time).....	46
Obrázek 17 - Amplitudové spektrum čistého i zašuměného měřicího signálu v pásmu 2000 Hz.	47
Obrázek 18 - Měření STI s přidáním šumem na začátku řetězce, komprimovaný pluginem. Tečkovaná křivka označuje referenční hodnotu STI pro přidání šum bez komprese (není závislá na hodnotě attack time).....	48

8. Literatura

- [1] GIANNOULIS, Dimitrios, Michael MASSBERG a Joshua D. REISS. Digital Dynamic Range Compressor Design – A Tutorial and Analysis. *J. Audio Eng. Soc.* 2012, vol. 60, No. 6, s. 399-408
- [2] IEC 60268-16:2011. Sound system equipment – Part 16: Objective rating of speech intelligibility by speech transmission index, 4. vyd.
- [3] ŠTĚPÁNEK, Tomáš. Vliv dynamické komprese signálu na parametr srozumitelnosti řeči STI. Praha: *České vysoké učení technické v Praze, Fakulta elektrotechnická*, 2017. 39 s. Vedoucí bakalářské práce Ing. Marek Brothánek, Ph.D.
- [4] ČELIKOVSKÝ, Sergej. Nelineární systémy. Praha: *Nakladatelství ČVUT*, 2006. ISBN 80-01-03435-6.
- [5] RUGH, Wilson J. Nonlinear System Theory: The Volterra/Wiener Approach. Baltimore: *The Johns Hopkins University Press*, 1981. ISBN 0-8018-2549-0.
- [6] HOUTGAST, Tammo., Herman. J. M. STEENEKEN. Evaluation of speech transmission channels by using artificial signals, *Acustica* 25, 1971, 355–367.
- [7] ANDERSON, B. Wayne a Joel T. KALB. English verification of the STI method for estimating speech intelligibility of a communications channel. *The Journal of the Acoustical Society of America*. 1987, **81**(6), 1982-1985. DOI: 10.1121/1.394764. ISSN 0001-4966.
- [8] STEENEKEN, Herman J.M a Tammo HOUTGAST. Validation of the revised STIr method. *Speech Communication*. 2002, **38**(3-4), 413-425. DOI:10.1016/S0167-6393(02)00010-9. ISSN 01676393.
- [9] PLOMP, Reinier. The negative effect of amplitude compression in multichannel hearing aids in the light of the modulation-transfer function. *The Journal of the Acoustical Society of America*. 1988, **83**(6), 2322-2327. DOI: 10.1121/1.396363. ISSN 0001-4966.