

ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ
V PRAZE

FAKULTA ELEKTROTECHNICKÁ

BAKALÁŘSKÁ PRÁCE



VLIV DYNAMICKÉ KOMPRESY SIGNÁLU NA PARAMETR
SROZUMITELNOSTI ŘEČI STI

Autor práce:

Tomáš Štěpánek

Vedoucí práce:

Ing. Marek Brothánek, Ph.D.

leden 2017

Prohlášení autora

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 autora

Poděkování

V první řadě bych rád poděkoval vedoucímu mé závěrečné práce, Ing. Marku Brothánkovi, Ph.D., za odborné vedení, za pomoc a rady při zpracování této práce, jakož i za vstřícnost, trpělivost a veškerý věnovaný čas. Dále děkuji za podporu své rodině a svým přátelům.



ZADÁNÍ BAKALÁŘSKÉ PRÁCE

Student: **Tomáš Štěpánek**

Studijní program: **Kybernetika a robotika**
Obor: **Senzory a přístrojová technika**

Název tématu česky: **Vliv dynamické komprese signálu na parametr srozumitelnosti řeči STI**

Název tématu anglicky: **Influence of Dynamic Signal Compression on Intelligibility Expressed by STI**

Pokyny pro vypracování:

Seznamte se s problematikou:

- metody měření impulsní odezvy používané v akustice (především MLS a TSP),
- určování parametru přenosu srozumitelnosti řeči STI a jeho zjednodušenou variantu STIPA,
- vlastností zařízení, umožňujících dynamickou kompresi úrovně zvuku.

V práci popište možné způsoby měření parametru STI. Při použití různých metod experimentálně ověřte vliv změny dynamických vlastností signálů na výslednou hodnotu STI. Dle specifikace vedoucího práce proveďte výpočty v prostředí MATLAB za použití audio-pluginů, případně u fyzického zařízení.

Seznam odborné 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.
- [4] Stan, G.-B., Embrechts, J.-J., Archangeau, D.: Comparison of Different Impulse Response Measurement Techniques, J. Audio Eng. Soc, 50(4), 249-262, 2002.
- [5] Moriya, N., Kaneda, Y.: Study of harmonic distortion on impulse response measurement with logarithmic time stretched pulse, Acoust. Sci. & Tech. 26(5), 462-464, 2005.

Vedoucí bakalářské práce: **Ing. Marek Brothánek, Ph.D. (K 13102)**

Datum zadání bakalářské práce: **17. ledna 2017**

Platnost zadání do¹: **30. září 2018**

vedoucí katedry



děkan /

V Praze dne 17. 1. 2017

¹ Platnost zadání je omezena na dobu tří následujících semestrů.

Anotace

V této práci je studován vliv dynamické komprese signálu na index přenosu řeči STI (z angl. *speech transmission index*). Jsou popsány teoretické principy jednotlivých metod měření, kterými lze dospět k cílovému parametru STI. Dále je detailně popsáno zařízení pro dynamické snížení rozsahu zvuku, tedy kompresor. Následně jsou všechny uvedené metody implementovány a použity k experimentálnímu ověření. Výsledkem je analýza vlivu kompresoru na parametr STI v rámci jednotlivých metod.

Klíčová slova

STI, kompresor, TSP, MLS, STIPA, Modulační přenosová funkce

Annotation

This thesis studies the influence of dynamic compression of sound signal on the STI (speech transmission index) parameter. Described within are the theoretical principles of the measuring methods which are used to define the STI. There is also described the dynamic range compressor. All the stated methods are implemented and then used for experimental verification. The result is the analysis of the influence of the compressor on the STI parameter considering each of the stated methods.

Keywords

STI, compressor, TSP, MLS, STIPA, Modulation transfer function

Obsah

Zadání	VII
Anotace	IX
1 Úvod	1
2 Teoretická část	3
2.1 Měření parametru STI	3
2.1.1 Přímá metoda	4
2.1.2 Nepřímá metoda	7
2.1.3 STIPA	7
2.2 Měření impulzní odezvy	8
2.2.1 Metoda MLS (maximum length sequence)	8
2.2.2 Metoda TSP (time stretched pulse)	9
2.3 Zařízení pro dynamickou změnu rozsahu zvuku	10
2.3.1 Ovládací parametry kompresoru	10
2.3.2 Výpočet komprese	12
2.3.3 Detekce hladiny	12
2.3.4 Uspořádání systému	14
3 Implementace	17
3.1 Generování testovacích signálů	17
3.1.1 Metoda MLS	17
3.1.2 Metoda TSP	18
3.1.3 Metoda STIPA	18
3.2 Implementace kompresoru	19
3.2.1 Postup výpočtu komprese	19
3.3 Kompresor dostupný v audio programu	20
3.4 Výpočet modulační přenosové funkce a parametru STI	20
3.4.1 Modulační přenosová funkce z impulzní odezvy	20

3.4.2	Modulační přenosová funkce metodou STIPA	21
3.4.3	Výpočet parametru STI	21
4	Experimentální ověření	23
4.1	Měřicí řetězec	23
4.1.1	Ověření funkce implementovaného kompresoru	23
4.2	Měření metodou MLS	25
4.2.1	Vliv změny parametru attack time na STI	26
4.2.2	Vliv změny parametru threshold na STI	27
4.2.3	Porovnání kompresorů	28
4.3	Měření metodou TSP	29
4.3.1	Působení kompresoru na signál TSP	30
4.4	Měření metodou STIPA	31
4.4.1	Měření parametru STI metodou STIPA	31
5	Závěr	35
	Literatura	37
A	Seznam souborů se zdrojovými kódy	39

Kapitola 1

Úvod

Kvalitu přenosu řeči mezi řečníkem a posluchačem lze objektivně určit pomocí parametru STI (z angl. *speech transmission index*, index přenosu řeči). V teoretické části této práce jsou uvedeny metody, které lze ke stanovení tohoto parametru použít. Z pohledu způsobu jeho určení se jedná o metodu přímou a nepřímou.

Přímou metodou měření STI je jeho výpočet z modulační přenosové funkce, která je pro tento parametr základem a k níž lze dospět několika způsoby zmíněnými v dalších částech této práce.

Nepřímá metoda měření parametru STI je provedena pomocí impulzní odezvy (odezva na tzv. Diracův impuls), jež je základním nástrojem k určení vlastností lineárních časově invariantních (LTI) systémů. Teoretická část dále pojednává o způsobech měření impulzní odezvy a speciálně se věnuje metodám MLS (z angl. *maximum length sequence*) a TSP (*time stretched pulse*).

Měření parametru STI může být uskutečněno bez ozvučovací techniky (např. divadla, učebny...), nebo s ní (dopravní prostředky, nákupní centra...). Součástí ozvučovací techniky může být i zařízení pro kompresi dynamického rozsahu zvuku, tedy kompresor, jež je detailně popsán na konci teoretické části.

Praktická část bakalářské práce zahrnuje implementaci uvedených metod výpočtu parametru STI a zejména se věnuje experimentálnímu ověření vlivu dynamické komprese zvuku na tento parametr.

Pro důkladné studium vlivu dynamické komprese na parametr STI byly veškeré algoritmy implementovány v prostředí MATLAB. V praktické části práce je dále uveden celkový popis experimentu, od generování testovacích signálů přes jejich zpracování po následné vyhodnocení v podobě parametru STI. V závěru práce jsou shrnuty výsledky experimentálního ověření vlivu dynamické komprese signálu na parametr STI.

Kapitola 2

Teoretická část

2.1 Měření parametru STI

Metoda měření parametru STI byla poprvé uvedena Steenekenem a Houtgastem v roce 1980 [1]. Podle nich se výpočet parametru STI skládal ze dvou kroků, ke kterým bylo potřeba mít naměřené hodnoty: index modulace m_k a intenzitu \bar{I}_k (tato veličina je v literaturách [1, 2] označena jako intenzita, ale ve skutečnosti se jedná o akustický tlak) pro oktávová pásma, která jsou použita v oblasti řeči:

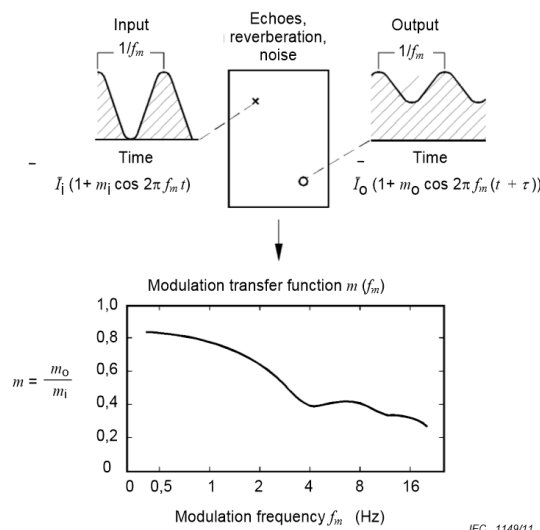
1. hodnoty m_k se použijí pro výpočet hodnot TI (transmission index, index přenosu) reprezentující příspěvky do jednotlivých oktávových pásem (k) ke konečnému parametru STI,
2. hodnoty TI_k jsou následně váženy a jejich sečtením je pak získán výsledný parametr STI.

V této části je čerpáno převážně z literatury [2], ze které pochází všechny následující vztahy. Metoda měření parametru STI je objektivní metoda, která se používá pro zjištění kvality přenosu řeči v přenosovém kanálu. STI je parametr nabývající hodnot od 0 do 1, kde 0 značí nejhorší kvalitu přenosu a 1 naopak nejlepší. Důležitým pojmem pro parametr STI je *modulační přenosová funkce* (MTF, modulation transfer function), která popisuje, jaký má daný přenosový kanál vliv na obálku intenzity signálu řeči (viz obrázek 2.1).

Z této funkce se spočítá tzv. efektivní poměr signál šum $SNR_{\text{eff } k, f_m}$ jako

$$SNR_{\text{eff } k, f_m} = 10 \lg \frac{m_{k, f_m}}{1 - m_{k, f_m}}, \quad (2.1)$$

kde m_{k, f_m} je modulační přenosová funkce pro dané oktávové pásmo (k) a modulační frekvenci (f_m). Tento poměr je omezen na hodnoty mezi ± 15 dB. Dalším krokem



Obrázek 2.1: Modulační přenosová funkce – porovnání vstupu a výstupu (převzato z [2])

k určení STI je vypočítání parametru TI (transmission index, index přenosu), který je dán vztahem

$$TI_{k,f_m} = \frac{SNR_{\text{eff } k, f_m} + 15}{30}. \quad (2.2)$$

Odvozené indexy přenosu (TI) se zprůměrují přes modulační frekvence pro získání MTI_k , indexu modulace přenosu, pro dané frekvenční pásmo k použitím vztahu

$$MTI_k = \frac{1}{n} \sum_{m=1}^n TI_{k,f_m}, \quad (2.3)$$

kde TI_{k,f_m} je ze vztahu (2.2), m je index modulační frekvence a n je počet modulačních frekvencí v oktávovém pásmu.

Parametr STI se použitím indexu modulace přenosu spočítá jako

$$STI = \sum_{k=1}^7 \alpha_k \cdot MTI_k - \sum_{k=1}^6 \beta_k \sqrt{MTI_k \cdot MTI_{k+1}}, \quad (2.4)$$

kde MTI_k je ze vztahu (2.3), α_k je váhovací faktor pro oktávové pásmo k a β_k je faktor redundance mezi oktávovými pásmy k a $(k+1)$. Tyto faktory jsou zobrazeny v tabulce 2.1.

I když tabulka obsahuje faktory pro obě pohlaví, pro měřící účely se používají obecně mužské, protože ženská řeč je považována za více srozumitelnou [2].

2.1.1 Přímá metoda

Přímá metoda měření, známá také jako FULL STI, používá pro výpočet modulační přenosové funkce MTF 14 modulačních frekvencí pro každé ze 7 oktávových pásem.

Oktávové pásmo [Hz]		125	250	500	1000	2000	4000	8000
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	-

Tabulka 2.1: Váhovací faktory MTF pro muže a ženy v jednotlivých oktávových pásmech (převzato z [2])

Při měření je dáno 98 nezávislých testovacích signálů (14×7), z nichž každý má dobu trvání průměrně 10 s. Celý proces metody FULL STI tedy trvá přibližně 15 minut.

Jiný způsob určení modulační přenosové funkce bez použití testovacího signálu (popsáno dále) je z doby dozvuku (2.5) a odstup signál šum (2.6). Tento způsob je jednodušší, nebere v potaz některé aspekty jako například zkreslení. Modulační index vypočítaný z doby dozvuku je pak dán vztahem

$$m(f_m) = \frac{1}{\sqrt{1 + \left(\frac{2\pi f_m T}{13,8}\right)^2}}, \quad (2.5)$$

kde f_m je modulační frekvence (Hz), dána pro jednotlivé oktávové pásmo, a T je doba dozvuku (s). Modulační index určený z poměru signál šum je dán vztahem

$$m = \frac{1}{1 + 10^{(-SNR)/10}}, \quad (2.6)$$

kde SNR je odstup signál šum (dB).

Generování testovacího signálu

Přímá metoda využívá testovací signály, jejichž některé časové a frekvenční vlastnosti odpovídají řečovému signálu. Tyto signály jsou složeny z jednotlivých frekvenčních pásem šumu, jehož intenzita je modulována sinusovou funkcí.

Využívá se růžového šumu, který je rozdělen na sedm frekvenčních pásem, jejichž střední frekvence nabývají hodnot od 125 Hz do 8 kHz a šířka každého pásma je půl oktávy.

Intenzita každého nosného signálu je sinusově modulována jednou nebo více modulačními frekvencemi (v závislosti na použití FULL STI nebo zjednodušených metod) s nejvyšší možnou hloubkou modulace m . Amplitudově modulovaný nosný signál je pak dán vztahem [2]

$$m_{f_m}(t) = \sqrt{0,5 \cdot (1 + m \cdot \cos(2\pi f_m t))},$$

kde f_m je modulační frekvence (Hz), t je čas (s) a m je hloubka modulace (maximálně rovna 1).

Výpočet MTF z testovacího signálu

Výpočet MTF pro každé oktákové pásmo je odvozen korelací obálky intenzity se sinovými a cosinovými signály se specifickou délkou a specifickými modulačními frekvencemi. Modulační přenos každé modulační frekvence je vypočítán odvozením hloubky modulace přijatého signálu (mdr) pro každé oktákové pásmo (k):

$$mdr_{k,f_m} = 2 \cdot \frac{\sqrt{[\sum I_k(t) \cdot \sin(2\pi f_m t)]^2 + [\sum I_k(t) \cdot \cos(2\pi f_m t)]^2}}{\sum I_k(t)}, \quad (2.7)$$

kde f_m je modulační frekvence (Hz), t je čas (s) a $I_k(t)$ je obálka akustického tlaku jako funkce času pro oktákové pásmo k .

Stejný postup je využít i při vypočtení hloubky modulace vyslaného signálu (mdt). Poměr mezi nimi se nazývá modulační přenosový a je označen jako m_{k,f_m} . Společným výsledkem těchto poměrů je tzv. MTF matice. Modulační přenosový poměr je dán vztahem

$$m_{k,f_m} = \frac{mdr_{k,f_m}}{mdt_{k,f_m}}, \quad (2.8)$$

kde mdr_{k,f_m} je hloubka modulace přijatého testovacího signálu pro oktákové pásmo k a modulační frekvenci f_m a mdt_{k,f_m} je hloubka modulace vyslaného testovacího signálu pro oktákové pásmo k a modulační frekvenci f_m . Modulační přenosový poměr nabývající hodnot vyšších než 1 je korigován na hodnotu 1. Hodnoty vyšší než 1,3 jsou velmi nepravděpodobné a mohou být zapříčiněny sinusovým kolísáním nebo impulzivními hluky.

Vypočtené hodnoty jsou pak dále ještě korigovány použitím sluchového maskování vztahem

$$m'_{k,f_m} = m_{k,f_m} \cdot \frac{I_k}{I_k + I_{am,k} + I_{rt,k}},$$

kde m_{k,f_m} je modulační přenosový poměr pro oktákové pásmo k a modulační frekvenci f_m , I_k je akustický tlak v oktákovém pásmu k , $I_{am,k}$ je akustický tlak pro úroveň maskovacího efektu v oktákovém pásmu k a $I_{rt,k}$ je akustický tlak prahu slyšení pro oktákové pásmo k .

Výpočet STI

Po vypočtení korigovaného modulovaného přenosového poměru se dále pokračuje vztahy (2.1) až (2.4) k výpočtu cílového parametru STI.

2.1.2 Nepřímá metoda

Modulační přenosová funkce MTF může být také spočítána z impulzní odezvy přenosového kanálu. Následující rovnice popisuje výpočet modulační přenosové funkce $m_{f,k}$ modulační frekvence f_m v oktávovém pásmu k :

$$m_k(f_m) = \frac{|\int_0^\infty h_k(t)^2 e^{-j2\pi f_m t} dt|}{\int_0^\infty h_k(t)^2 dt} \cdot (1 + 10^{-SNR_k/10})^{-1}, \quad (2.9)$$

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

2.1.3 STIPA

Pro zjednodušení poměrně zdlouhavého postupu 98 měření při FULL STI byla vyvinuta metoda STIPA (*Speech Transmission Index for Public Address Systems*, Index přenosu řeči pro ozvučovací systémy). Ta využívá pouze dvou modulačních frekvencí v sedmi oktávových pásmech, takže počet měření je zredukován na 14.

Modulační frekvence pro metodu STIPA

Každé oktávové pásmo je modulováno dvěma modulačními frekvencemi zároveň ve frekvenčním poměru rovném 5 a hloubka modulace je 0,55 pro sinusové přidání dvou složek s fázovým rozdílem 180° mezi těmito složkami. Modulační frekvence pro metodu STIPA jsou zobrazeny v tabulce 2.2.

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	10,0	6,25	4,00	12,5

Tabulka 2.2: Modulační frekvence pro metodu STIPA (převzato z [2])

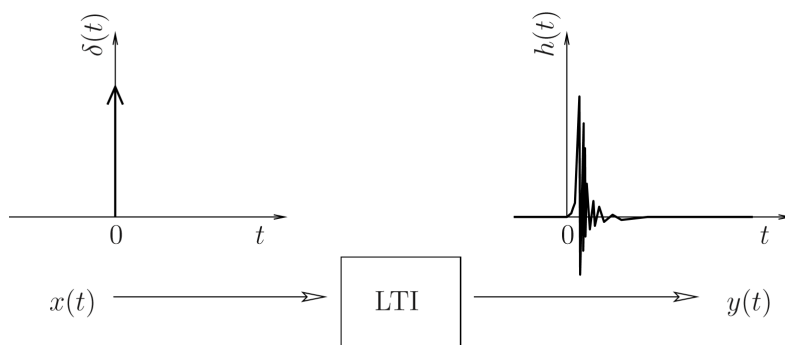
Aplikace a omezení metody STIPA

Tato metoda může být použita na téměř všechny digitální, analogové, elektroakustické a akustické přenosové kanály. Oproti metodě FULL STI se čas měření touto metodou zkrátí přibližně na 15 až 20 s.

Testovacím signálem metody STIPA je šum, takže jednotlivá měření obvykle nedávají stejné výsledky. Tyto výsledky vykazují určitou odchylku od průměru, která je u metody STIPA s časem měření 15 s přibližně 0,03STI pro několik opakovaných měření.

2.2 Měření impulzní odezvy

Impulzní odezva je jedním ze základních popisů LTI (*linear time invariant*, lineární časově neměnný) systémů, protože uceleně popisuje jeho vlastnosti [3]. Jedná se o odezvu na Diracův jednotkový impuls, kterým je soustava buzena (viz obrázek 2.2).



Obrázek 2.2: Odezva LTI systému na Diracův impuls (převzato z [3])

2.2.1 Metoda MLS (maximum length sequence)

Metoda MLS je založena na buzení akustického prostoru binárním periodickým pseudonáhodným signálem, který nabývá hodnot -1 a $+1$. Perioda jednoho takového signálu je dána vztahem

$$L = 2^N - 1, \quad (2.10)$$

kde N je řád MLS.

Vlastnosti MLS signálu

Významnou vlastností MLS signálu je jeho *autokorelační funkce* $R_{p_{xx}}[n]$, jejíž vlastnosti se blíží vlastnostem Diracova impulsu, který je diskrétně popsán jako

$$\delta_p[n] = \begin{cases} 1 & \text{pro } n = 0, \\ 0 & \text{jinak.} \end{cases}$$

Binární signál je považován za MLS, když platí následující vztah:

$$R_{p_{xx}}[n] = \delta_p[n] - \frac{1}{L+1}$$

Z tohoto vztahu je zřejmé, že pro $L \rightarrow \infty$ se bude autokorelační funkce periodického MLS signálu rovnat periodickému Diracovu impulsu.

Impulzní odezva

Vlastnosti MLS signálu z části 2.2.1 se pak využijí při výpočtu výsledné impulzní odezvy. Klíčový je předpis vzájemné korelace:

$$R_{pxy}[n] = x_p[n] \oplus y_p[n] = \frac{1}{L+1} \sum_{k=0}^{L-1} x_p[k]y_p[n+k], \quad (2.11)$$

jehož vlastnosti jsou podobné impulzní odezvě. Tím je dán výsledný vztah [5]

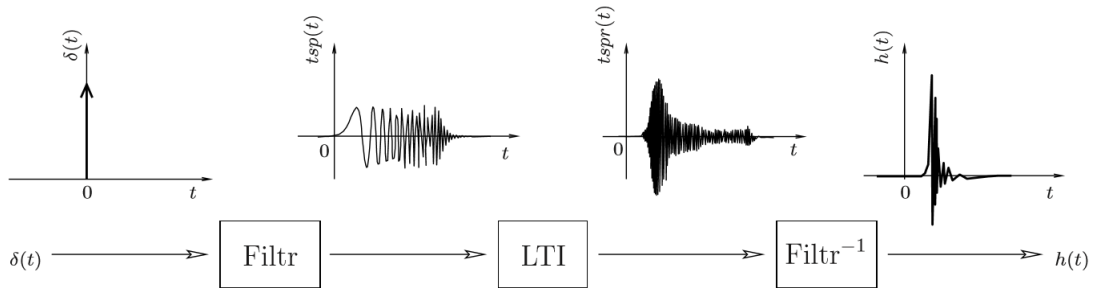
$$h_p[n] \approx R_{pxy}[n],$$

kde $h_p[n]$ je impulzní odezva.

Dále je také důležité zajistit, aby délka MLS signálu byla větší než délka impulzní odezvy, jinak se objeví tzv. časový aliasing.

2.2.2 Metoda TSP (time stretched pulse)

Metoda *Time stretched pulse*, v překladu *puls roztažený v čase*, je založená na filtraci Diracova impulsu $\delta(t)$, jehož energie se postupně rozloží v čase podle kmitočtu. Odezva systému na tento signál je pak filtrována opačným filtrem, pomocí kterého je získána impulzní odezva. Princip metody TSP je zobrazen na obrázku 2.3.



Obrázek 2.3: Princip metody TSP (převzato z [6])

Vlastnosti logaritmického TSP

Závislost skupinového zpoždění na frekvenci logaritmického TSP je popsána následujícími vztahy [7]

$$TSP_{\log}(k) = \begin{cases} 1 & \text{pro } k = 0, \\ \frac{e^{-j\alpha k \lg(k)}}{\sqrt{k}} & \text{pro } 1 \leq k \leq \frac{L}{2}, \\ TSP_{\log}^*(L-k) & \text{pro } \frac{L}{2} < k < L, \end{cases}$$

$$a = \frac{2M\pi}{\frac{L}{2} \lg\left(\frac{L}{2}\right)},$$

kde L je délka signálu, k diskretní frekvence, M celé číslo udávající polovinu délky úseku signálu nesoucího většinu energie. Hvězdička značí komplexně sdruženou hodnotu. Inverzní signál $iTSP_{\log}$, tedy charakteristika opačného filtru, je definována jako

$$iTSP_{\log}(k) = \frac{1}{TSP_{\log}(k)}.$$

2.3 Zařízení pro dynamickou změnu rozsahu zvuku

Významnou součástí ozvučovací techniky jsou zařízení pro dynamickou změnu rozsahu zvuku. Jako nejpoužívanější zástupci jsou jmenovány čtyři druhy lišící se ve tvaru převodní charakteristiky:

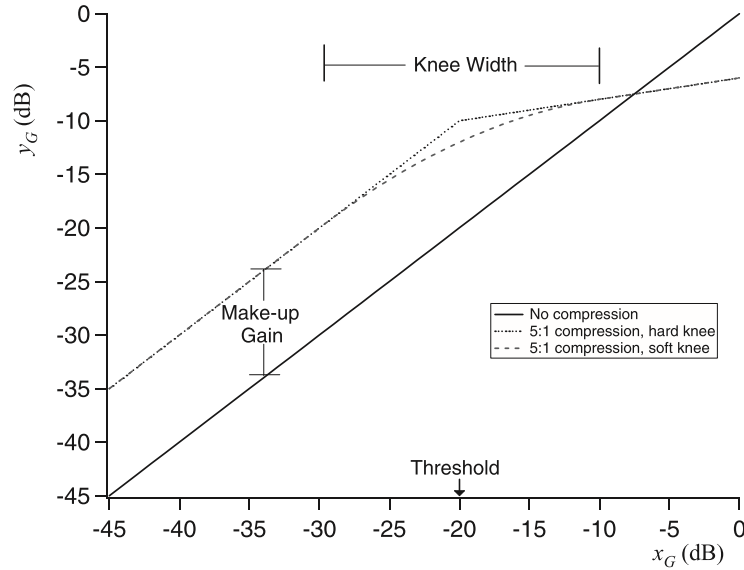
- **Kompresor** je zařízení, které snižuje dynamický rozsah zvuku. Těmto zařízením se bude následující část věnovat detailněji.
- **Limitér** je speciální případ kompresoru, kde se kompresní proměr (viz parametr *ratio* v části 2.3.1) rovná $\infty : 1$.
- **Gate** potlačuje signál, jehož hladina nepřesáhne určitou hodnotu.
- **Expander** funguje opět na podobném principu jako kompresor, ovšem dynamický rozsah zvuku se nesnižuje, ale naopak zvyšuje. Jeho parametr *ratio* bude tedy ve tvaru $1 : n$, kde $n > 1$.

Kompresie dynamického rozsahu (DRC, Dynamic Range Compression) je proces zpracování dynamického rozsahu audio signálu do menšího rozsahu například zmenšením úrovně špiček signálu, zatímco tišší části jsou zachovány. DRC je rozsáhle používána pro nahrávání audiosignálů, vysílání a živá vystoupení [8].

2.3.1 Ovládací parametry kompresoru

Kompresory mají několik parametrů určujících změnu dynamického rozsahu signálu. Obrázek 2.4 popisuje, jak se mění výstupní signál při použití kompresoru. Na ose x je zobrazena vstupní úroveň efektivní hodnoty signálu (dB), na ose y pak výstupní úroveň efektivní hodnoty signálu (dB). Souvislá čára ukazuje převodní charakteristiku bez komprese, tečkovaná a přerušovaná čára s kompresí.

- *Threshold* je parametr, kterým je určeno, pro jakou vstupní úroveň začíná komprese. Signál, který přesahuje threshold je pak redukován.



Obrázek 2.4: Charakteristika statické komprese s parametry *Make-Up Gain* a *hard* nebo *soft knee* (převzato z [8])

- *Ratio* je poměr mezi hladinami vstupního a výstupního signálu, které přesáhly threshold. Udává tedy množství komprese, které se na signál aplikuje.
- Parametry *Attack time* a *Release time* určují, jak rychle kompresor reaguje na signál, který přesáhl threshold (*attack time*) a následně po tom, co se rozsah signálu snížil pod threshold, jak rychle se vrátí do stavu bez komprese (*release time*).
- Kompresí rozsahu se snížil výstupní úroveň signálu. Proto je zaveden *Make-Up Gain*, který udává zesílení komprimovaného signálu umístěné obvykle na výstupu kompresoru.
- Parametr *Knee Width* určuje, zda se hladina signálu vyvíjí lineárně až do thresholdu, kde následuje zlom a dále je hladina dána kompresním poměrem (*ratio*), tento přechod se nazývá *hard knee* (parametr *knee width* je tedy nulový). Přechod ke kompresi může být také hladký, nazýván *soft knee*, kdy je nastavena šířka přechodu ke kompresi.

Hladinu výstupního signálu z kompresoru lze shrnout vztahem [8]

$$y_{\text{dB}}[n] = x_{\text{dB}}[n] + c_{\text{dB}}[n] + M,$$

kde $x_{\text{dB}}[n]$ je hladina vstupního signálu (dB), $c_{\text{dB}}[n]$ je zesílení (dB), M je hodnota *make-up gainu* (dB) a $y_{\text{dB}}[n]$ je hladina výstupního signálu (dB).

2.3.2 Výpočet komprese

Vyčtené parametry v části 2.3.1 se projeví ve výpočtu výsledného výstupního signálu. Kompresní poměr R (ratio) je definován jako [8]

$$R = \frac{x_G - T}{y_G - T} \quad \text{pro } x_G > T,$$

kde T je threshold. Komprimace výstupu je pak popsána vztahem

$$y_G = \begin{cases} x_G & 2(x_G - T) \leq -W, \\ x_G + (1/R - 1)(x_G - T + W/2)^2 / (2W) & 2|(x_G - T)| \leq W, \\ T + (x_G - T)/R & 2(x_G - T) > W, \end{cases} \quad (2.12)$$

kde W (dB) je parametr *knee width*, který určuje, zda přechod ke kompresi bude typu *hard knee* ($W = 0$), nebo *soft knee* ($W > 0$).

2.3.3 Detekce hladiny

Část kompresoru zajišťující detekci hladiny signálu slouží k jejímu hladkému znázornění. Detekce hladiny může probíhat více způsoby. V této části jsou popsány čtyři možnosti detekce.

Změna zesílení probíhá v závislosti na parametrech attack time a release time obvykle uvedených ve filtru, který může být simulován jako digitální jednopólový filtr [8]:

$$s[n] = \alpha s[n-1] + (1 - \alpha)r[n], \quad (2.13)$$

kde α je koeficient filtru, $r[n]$ je vstup a $s[n]$ je výstup. Odezva na jednotkový skok je definována jako

$$s[n] = 1 - \alpha^n \quad \text{pro } x[n] = 1, n \geq 1. \quad (2.14)$$

Časová konstanta τ je definována jako čas, který systému zabere dosáhnout $(1 - 1/e)$ jeho původní hodnoty, tj. $s[\tau F_s] = 1 - 1/e$, takže ze vztahu (2.14) je získán koeficient filtru jako

$$\alpha = e^{-1/(\tau F_s)}, \quad (2.15)$$

kde F_s je vzorkovací frekvence. V analogových detektorech je filtr nahrazen sériovou kombinací rezistoru a kapacitoru s časovou konstantou $\tau = RC$.

Detekce založená na efektivní hodnotě

Detekce hladiny může být založena na měření efektivní hodnoty (RMS) vstupního signálu, což je definováno jako [8]

$$y_L^2[n] = \frac{1}{M} \sum_{m=-M/2}^{M/2-1} x_L^2[n-m].$$

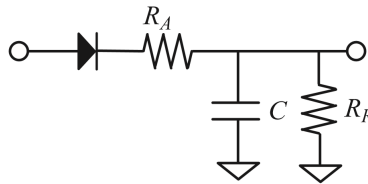
Důvod použití RMS detekce úzce souvisí s vnímáním hlasitosti signálu. Detektor nejprve spočítá efektivní hodnotu vzorku signálu a až následně rozhodne, zda překročil threshold. Tento způsob detekce hladiny je ale nevhodný pro využití v reálném čase kvůli zpoždění o $M/2$ vzorku.

Diferenční rovnice popisující RMS detektor je dána vztahem, který vykazuje obecně ekvivalentní chování jako vztah (2.13), a to [8]

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

Detektor špičky

Analogový detektor špičky je běžně používán v analogových zařízeních pro dynamickou kompresi zvuku [8] a je zobrazen na obrázku 2.5.



Obrázek 2.5: Detektor špičky zobrazený v obvodu (převzato z [8])

Ve zjednodušeném případě je dioda považována za ideální, což výrazně usnadní výpočet a vede k diferenciální rovnici [8]

$$\frac{dV_C}{dt} = \frac{\max(V_{in} - V_C, 0)}{R_A C} - \frac{V_C}{R_R C}.$$

Kapacitor je nabíjen přes resistor R_A podle kladného napětí přes diodu, ale neustále vybíjen resistorem R_R . Ideální analogový detektor špičky je pak simulován [8]:

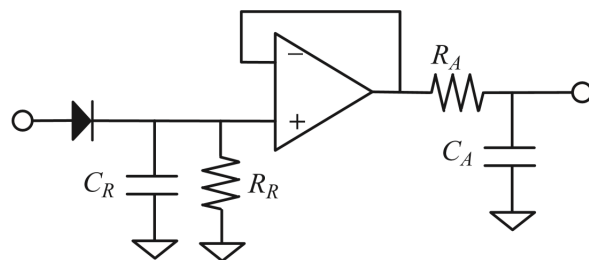
$$y_L[n] = \alpha_R y_L[n-1] + (1 - \alpha_A) \max(x_L[n] - y_L[n-1], 0),$$

kde α_A a α_R jsou tzv. *attack* a *release* koeficienty, které vzešly ze vztahu (2.15), kde *attack time* $\tau_A = R_A C$ a *release time* $\tau_R = R_R C$.

Odezva na jednotkový skok je $\tau_R / (\tau_R + \tau_A)$, což znamená, že správný odhad špičky nastane, pokud bude release time výrazně větší než attack time. Dalším problémem je, že attack time je zmenšován parametrem release time, takže pokud je použit vysoký release time, bude vyšší attack time.

Hladinově upravený detektor špičky

Na obrázku 2.6 je zobrazen alternativní detektor špičky, jehož části obvodu pro attack a release koeficienty jsou odděleny. To eliminuje výkyvy hladiny způsobené použitím jednoduchého detektoru špičky.



Obrázek 2.6: Oddělený detektor špičky (převzato z [8])

Digitální implementace odděleného detektoru špičky dána vztahy

$$\begin{aligned} y_1[n] &= \max(x_L[n], \alpha_R y_1[n-1]), \\ y_L[n] &= \alpha_A y_L[n-1] + (1 - \alpha_A) y_1[n], \end{aligned}$$

kterými se ale nezabýváme vlivu parametru release time. Sečtením koeficientů α_A a α_R se získá dobrý odhad pro release time.

V digitálním prostředí tyto problémy mohou být odstraněny přidáním větve do diferenční rovnice jednoduchého detektoru špičky.

$$y_L[n] = \begin{cases} \alpha_A y_L[n-1] + (1 - \alpha_A) x_L[n] & \text{pro } x_L[n] > y_L[n-1], \\ \alpha_R y_L[n-1] & \text{pro } x_L[n] \leq y_L[n-1]. \end{cases}$$

Detektor vyhlazené špičky

Všechny analogové návrhy využijí plnou hodnotu parametru release time, když se po špičce vrátí vstupní signál zpět na nulu. Pokud se ale signál ustálí na mezní hodnotě, zastaví se proces vracení zesílení na původní hodnotu. Release time je pak kratší než je očekáváno, což způsobuje skok ve výstupním signálu. Následujícím rozšířením větveného detektoru špičky se zajistí, že bude pokaždé release time plně využit.

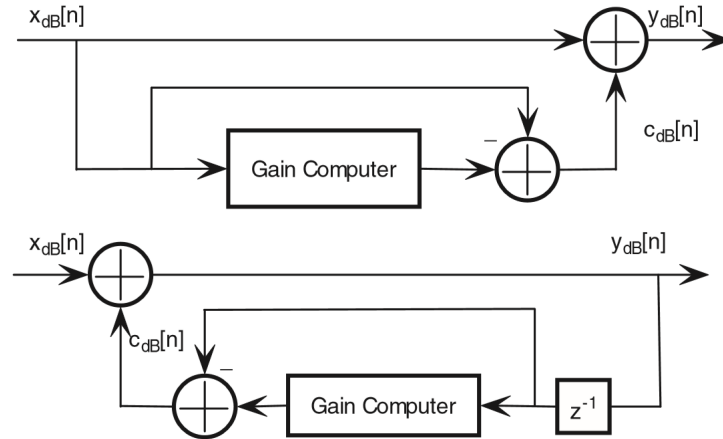
$$y_L[n] = \begin{cases} \alpha_A y_L[n-1] + (1 - \alpha_A) x_L[n] & \text{pro } x_L[n] > y_L[n-1], \\ \alpha_R y_L[n-1] + (1 - \alpha_R) x_L[n] & \text{pro } x_L[n] \leq y_L[n-1]. \end{cases} \quad (2.16)$$

Tento způsob detekce hladiny jednoduše přepíná koeficienty mezi procesy doby náběhu (*attack time*) a doběhu (*release time*).

2.3.4 Uspořádání systému

Postranní výpočetní řetězec může mít zpětnovazební nebo dopřednovazební rozložení. Ve zpětnovazebním rozložení je vstup do postranního řetězce brán po aplikování zesílení. Tento typ byl používán v dřívějších kompresorech a měl tu výhodu,

že postranní řetězec mohl opravit možné nepřesnosti v části výpočtu zesílení. Naopak v dopřednovazebním rozložení je vstup do postranního řetězce před zesílením. Postranní řetězec následně vypočte redukci zesílení vstupního signálu. Většina moderních kompresorů využívá dopřednovazební rozložení postranního řetězce [8]. Obě rozložení jsou zobrazena na obrázku 2.7.



Obrázek 2.7: Dopřednovazební a zpětnovazební uspořádání postranního řetězce (převzato z [8])

Umístění detektoru hladiny

Existuje více možností, kam detektor hladiny v kompresoru umístit. Tři z těchto možností jsou zobrazeny na obrázku 2.8. Jako první způsob umístění detektoru se nabízí v lineární oblasti, před převodem na decibely a před výpočtem zesílení:

$$\begin{aligned} x_L[n] &= |x[n]|, \\ x_G[n] &= 20 \lg y_L[n], \\ c_{dB}[n] &= y_G[n] - x_G[n]. \end{aligned}$$

Při tomto zapojení (obrázek 2.8 (a)) detektor pracuje v plném dynamickém rozsahu vstupního signálu, zatímco zesílení se začne aplikovat, až když úroveň signálu přesáhne threshold. Výsledkem je diskontinuita v doběhu signálu při jeho poklesu pod threshold. Dále je také generováno zpoždění v náběhu zesílení kompresoru, protože detektor potřebuje nějaký čas k nabití na úroveň thresholdu, i když vstupní signál nabíhá okamžitě.

Odečtením thresholdu od vstupního signálu a jeho přičtením po detekci se zamezí předchozím nedostatkům a vzniknou vztahy

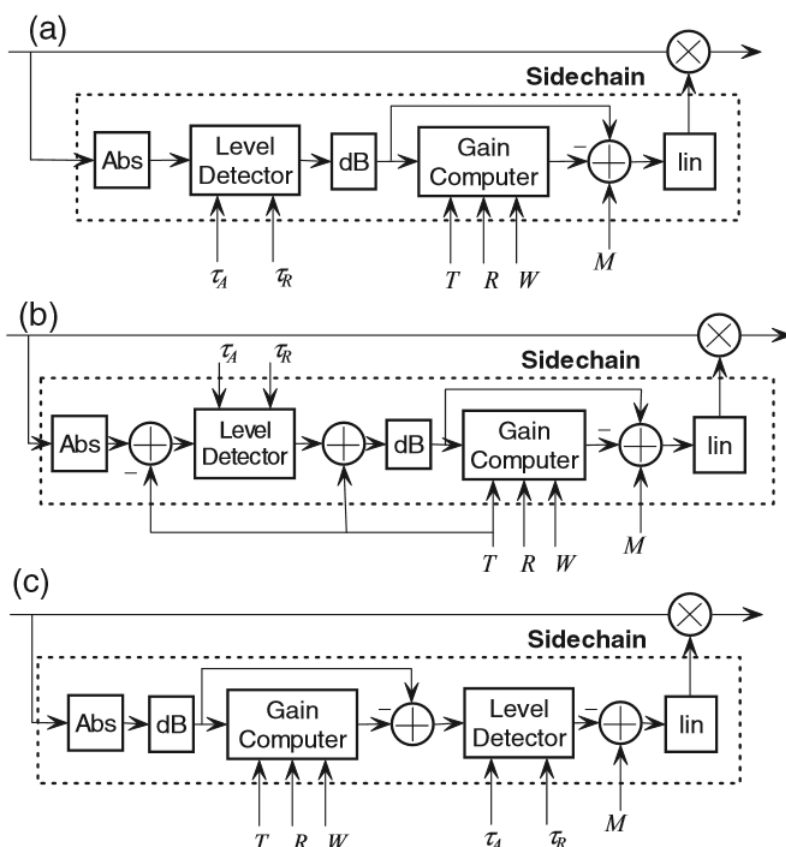
$$\begin{aligned} x_L[n] &= |x[n]| - 10^{T/20}, \\ x_G[n] &= 20 \lg(y_L[n] + 10^{T/20}), \\ c_{dB}[n] &= y_G[n] - x_G[n]. \end{aligned}$$

Toto zapojení je znázorněno na obrázku 2.8 (b). Obálka signálu tedy bude hladce klesat, když spadne pod úroveň thresholdu. Bohužel toto řešení není vhodné pro použití charakteristiky soft knee, protože při ní není dána fixní hodnota thresholdu.

Jako nejvhodnější pro vnímání lidského ucha se jeví umístění detektoru do logaritmické oblasti, kde detektor vyhlazuje přímo řídicí napětí (viz obrázek 2.8 (c)). Podle tohoto uspořádání jsou získány vztahy pro jednotlivé fáze výpočtu jako [8]

$$\begin{aligned} x_G[n] &= 20 \lg |x[n]|, \\ x_L[n] &= x_G[n] - y_G[n], \\ c_{dB}[n] &= -y_L[n]. \end{aligned}$$

Takto detektor přímo reaguje na kontrolní napětí místo vstupního signálu. Pokud je kontrolní napětí nulové při neaktivnosti kompresoru, proces doběhu parametru release time je zajištěn nezávisle na pevně zvoleném parametru threshold. Trajektorie se tedy chová exponenciálně v logaritmické oblasti, což znamená, že release time je nezávislý na množství aplikované komprese [8].



Obrázek 2.8: Blokové znázornění konfigurace kompresoru. (a) detektor vztažený k nule; (b) detektor vztažený k parametru threshold; (c) detektor v logaritmické oblasti (převzato z [8])

Kapitola 3

Implementace

Všechny části měřicího řetězce s výjimkou softwarového kompresoru byly implementovány v prostředí MATLAB. Jejich zdrojové kódy jsou uloženy na CD jako příloha této práce. Seznam všech souborů je uveden v příloze A na straně 39.

3.1 Generování testovacích signálů

Před generováním jednotlivých testovacích signálů je třeba určit jejich jednotné parametry, kterými jsou vzorkovací frekvence F_s (Hz) a délka trvání t (s).

Vzorkovací frekvence je zvolena jako

$$F_s = 44\,100 \text{ Hz},$$

což je běžně využíváno například na CD nosičích.

Délka testovacího signálu by se měla volit podle vlastností konkrétní metody a přenosového systému. Například při volbě signálu MLS je třeba volit jeho periodu delší než je doba dozvuku dané přenosové cesty (prostoru), jinak se objeví nežádoucí aliasing. Přestože bude měření probíhat v části řetězce, kde bude ovlivňovat signál pouze kompresor, délka testovacího signálu bude zvolena jako $t = 15$ s.

3.1.1 Metoda MLS

Pro generování MLS signálu bude využita funkce `mls()`, které je jako vstupní parametr zadán řád MLS. Testovací signál MLS se skládá ze tří po sobě jdoucích stejných posloupností a pro měřicí účely by měla být jeho délka alespoň $t = 15$ s. Řád N MLS signálu bude tedy muset odpovídat nerovnosti

$$\frac{3 \cdot (2^N - 1)}{44100} > 15.$$

Při vyšších hodnotách N může být zanedbáno číslo 1, takže se nerovnost dále dá upravit na vztah

$$N > 17,75.$$

Řád MLS signálu by měl být tedy alespoň $N = 18$.

Vygenerovaný signál je následně vyslán do měřicího řetězce, v našem případě kompresoru. Vstupní testovací a výstupní naměřený signál jsou předány jako vstupy funkci `mls_response()` společně s řetězcem, který udává, ze které části testovacího signálu se finální impulzní odezva vypočítá. Pro výpočet impulzní odezvy je využito vztahu (2.11) na straně 9.

3.1.2 Metoda TSP

Pro generování signálu TSP bude využito funkce `tspgen_lg()` z [6]. Vstupem do této funkce je číslo N (udávající délku signálů jako $L = 2^N$) a vzorkovací frekvence (Hz). Výstupem je pak signál TSP a signál k němu inverzní.

Vlastností logaritmického TSP signálu je jeho rostoucí frekvence v závislosti na čase. Měření kvality přenosu řeči bude probíhat ve slyšitelné oblasti, takže testovací TSP signál začne až od frekvence 27 Hz. Délku celého signálu je tedy třeba volit tak, aby část od 27 Hz byla alespoň 15 s dlouhá. Řád generovaného TSP signálu tedy bude $N = 22$, což zajistí délku testovacího signálu $t = 22$ s.

Testovací signál se vyšle do testovacího řetězce a výstup se zpracuje funkcí `tsp_response()`, které je nutné jako vstup předat také vygenerovaný inverzní signál.

3.1.3 Metoda STIPA

Ke generování testovacího signálu pro metodu STIPA bylo využito sedmi půloktávových šumů pro každé oktávové pásmo. Každé z pásem bylo amplitudově modulováno dvěma modulačními frekvencemi podle tabulky 2.2 na straně 7, kde je hloubka modulace 0,55 a fázový rozdíl modulačních signálů 180° . Následně se modulované signály v jednotlivých oktávových pásmech sečetly do výsledného signálu, který byl v metodě STIPA použit. Signál se dal jednoduše přizpůsobit délce $t = 15$ s. Generování testovacího signálu metody STIPA je zajištěno funkcí `stipa_gen()`, které jsou jako vstupy předány vzorkovací frekvence (Hz) a délka signálu (s).

Na rozdíl od nepřímé metody je ze vstupního a výstupního signálu STIPA rovnou počítána modulační přenosová funkce a následně parametr STI. Postup implementace tohoto výpočtu je uveden níže.

3.2 Implementace kompresoru

Jedna z možností zapojení kompresoru do měřicího řetězce je jeho implementace. V tomto případě se jedná o funkci `compressor()`, které jsou předány parametry:

- vstupní signál v rozsahu $\langle -1, 1 \rangle$,
- vzorkovací frekvence (Hz),
- threshold (dB),
- make-up gain (dB),
- ratio (-),
- knee width (dB),
- attack time (s),
- release time (s).

Výstupem je pak komprimovaný vstupní signál.

3.2.1 Postup výpočtu komprese

Po vytvoření kopie vstupního signálu jsou jeho hodnoty v intervalu $\langle -10^{-5}, 10^{-5} \rangle$ nastaveny na hodnotu 10^{-5} , aby se při přechodu do logaritmické oblasti vztahem

$$x_G[n] = 20 \lg |x[n]|$$

zamezilo chybám při převodu částí signálu, kde je jeho amplituda blízká nule.

Následně je proveden výpočet statické charakteristiky podle vztahu (2.12) na straně 12 a z té je získán signál $x_{sc}[n]$. Od výsledku je odečten původní signál v logaritmické oblasti, takže je získáno zesílení, které je třeba aplikovat na vstupní signál:

$$g_c[n] = x_{sc}[n] - x_G[n]$$

Ještě předtím je ale provedena detekce hladiny (viz část 2.3.3 na straně 12). V našem případě je implementován *detektor vyhlazené špičky* podle vztahu (2.16) na straně 14. K signálu g_m upraveném detekcí hladiny je ještě přičten *make-up gain* a poté převeden do lineární oblasti vztahem

$$g_{lin}[n] = 10^{g_m[n]/20}.$$

Jednotlivé složky signálu $g_{lin}[n]$ jsou nakonec vynásobeny se složkami vstupního signálu $x[n]$, a tím je získán výsledný komprimovaný signál.

3.3 Kompresor dostupný v audio programu

Další z možností komprese testovacího signálu je využití kompresoru dostupného v libovolném audio softwaru. V případě této práce se jedná o plug-in kompresor ve volně dostupném nahrávacím softwaru Cubase 5 LE (viz obrázek 3.1). Nacházejí se na něm parametry popsané v 2.3.1 na straně 10, tedy *threshold*, *ratio*, *make-up gain*, *attack time* a *release time*. Chybí volba parametru *knee width*, místo kterého je k dispozici pouze volba *soft knee* nebo *hard knee*. Dále je v pravé části obrázku zobrazena statická charakteristika kompresoru, vstupní a výstupní úroveň a úroveň komprese.



Obrázek 3.1: Ovládací panel Cubase kompresoru (převzato z programu Cubase 5 LE)

3.4 Výpočet modulační přenosové funkce a parametru STI

Modulační přenosová funkce (MTF) je klíčová pro výpočet parametru STI a lze k ní dospět několika způsoby. Vlastnosti MTF jsou popsány v části 2.1 na straně 3.

3.4.1 Modulační přenosová funkce z impulzní odezvy

Metodami TSP nebo MLS (sekce 2.2 na straně 8) je získána impulzní odezva přenosového systému a z ní je vypočítána modulační přenosová funkce vztahem (2.9) na straně 7. V prostředí MATLAB je implementována funkce `mtfIR()`, jejímiž vstupy jsou impulzní odezva systému a vzorkovací frekvence.

Impulzní odezva je nejprve filtrována oktávovým filtrem na sedm oktávových pásem (125 Hz – 8 kHz). Následně je proveden výpočet vztahu (2.9) uvedený v sekci 2.1, ze kterého je pro každé pásmo a modulační frekvenci získána složka MTF. Výsledkem je tedy matice o velikosti 7×14 .

3.4.2 Modulační přenosová funkce metodou STIPA

Výpočet MTF metodou STIPA je proveden v implementované funkci `mtfSTIPA()`, jíž jsou jako vstupy předány vstupní a výstupní (komprimovaný) testovací signály a vzorkovací frekvence.

Nejprve jsou oba signály filtrovány oktávovým filtrem na sedm oktávových pásem. Následuje výpočet hloubky modulace vstupního a komprimovaného signálu podle vztahu (2.7) ze strany 6. Výpočet hloubky modulace bude probíhat pouze na dvou modulačních frekvencích, kterými je dané oktávové pásmo modulováno. Dále jsou nenulové složky vypočítané matice komprimovaného a vstupního signálu mezi sebou vyděleny podle vztahu (2.8) na straně 6, čímž vznikne výsledná modulační přenosová funkce.

3.4.3 Výpočet parametru STI

Závěrečným krokem měření je výpočet konečného parametru STI, který může nabývat hodnot v rozmezí $\langle 0, 1 \rangle$. Vlastnosti parametru STI jsou popsány v části 2.1 na straně 3.

K výpočtu parametru STI nepřímou metodou slouží funkce `STI()`, jejímž vstupem je modulační přenosová funkce a výstupem parametr STI. Postupně jsou provedeny výpočty podle vztahů (2.1) až (2.4) z části 2.1.

U výpočtu parametru STI metodou STIPA je postupováno stejně s tím rozdílem, že se nepočítá s nulovými složkami v matici modulační přenosové funkce. Tento postup je implementován ve funkci `STIPA()`.

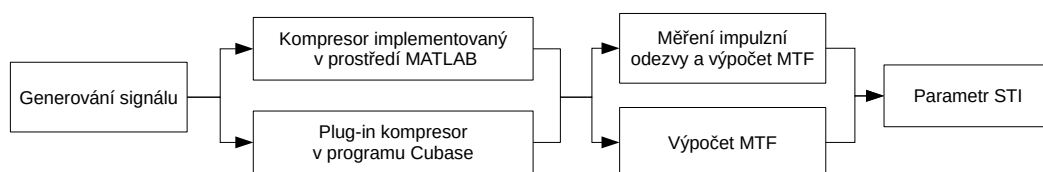
Kapitola 4

Experimentální ověření

Tato kapitola bude věnována vyhodnocení jednotlivých metod výpočtu parametru STI při použití zařízení pro dynamickou kompresi zvuku. Tedy jak se bude měnit hodnota parametru STI v závislosti na změně parametrů zvukového kompresoru.

4.1 Měřicí řetězec

Před vlastním měřením je potřeba vygenerovat testovací signál (MLS, TSP nebo STIPA), který je vyslán do kompresoru. V tomto případě jsou k dispozici dva typy kompresoru – jeden implementovaný v prostředí MATLAB a druhý dostupný v nahrávacím audio softwaru Cubase. Následně se v případě MLS a TSP vypočítá impulzní odezva a poté modulační přenosová funkce (MTF) podle vztahu (2.9) na straně 7. V případě signálu STIPA je modulační přenosová funkce spočítána rovnou podle vztahů (2.7) a (2.8) na straně 6. Po získání MTF je spočítán parametr STI. Pro ilustraci je měřicí řetězec zobrazen na obrázku 4.1.



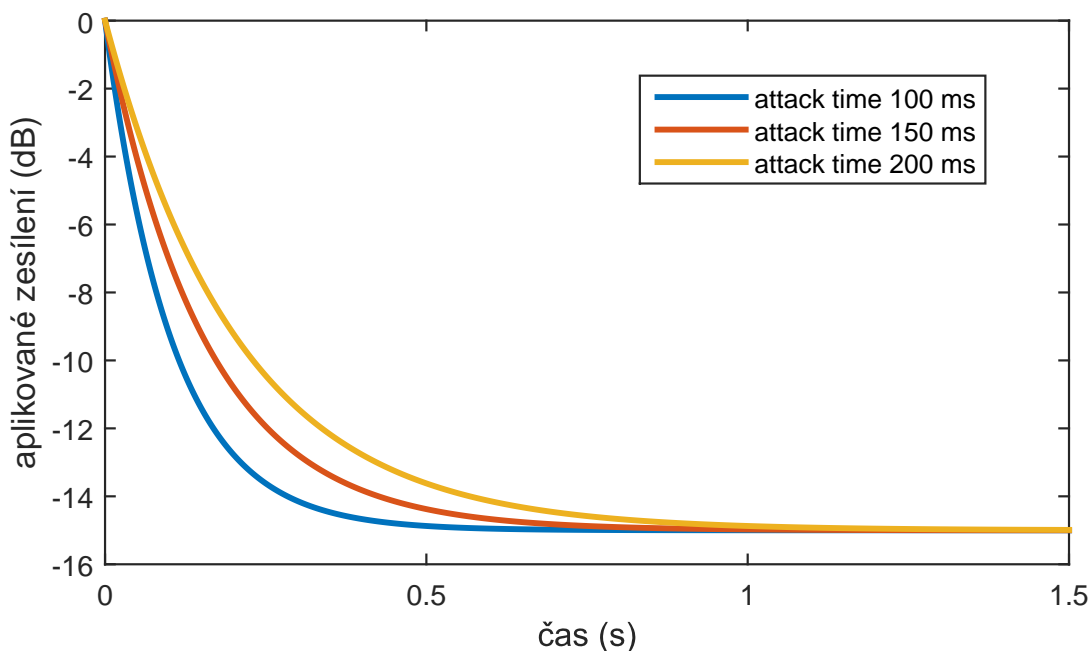
Obrázek 4.1: Měřicí řetězec

4.1.1 Ověření funkce implementovaného kompresoru

Aby mohl být použit implementovaný kompresor, je třeba zajistit jeho správnou funkci. Výhoda oproti softwarovému kompresoru je, že každá výpočetní část může

být podrobně zkoumána a analyzována. Pro ilustraci je na obrázku 4.2 zobrazeno zesílení, které kompresor aplikuje na signál MLS, jehož špičkové hodnoty jsou stále konstantní (viz část 2.2.1 na straně 8). Parametry nastavení kompresoru jsou:

- Threshold: -20 dB
- Ratio: 4
- Make-up gain: 0 dB
- Knee Width: 0 dB
- Attack time: 100, 150 a 200 ms
- Release time: 500 ms



Obrázek 4.2: Ukázka komprese signálu spolu s aplikovaným zesílením

Byly vybrány tři různé hodnoty parametru attack time, aby byla patrná změna průběhu aplikovaného zesílení. Z obrázku 4.2 je vidět, že při vyšším parametru attack time je doba náběhu zesílení delší. Hodnota, na které se pak zesílení ustálí, je -15 dB, což odpovídá části statické charakteristiky, kde úroveň vstupního signálu překročila threshold. Tu vyjadřuje rovnice

$$y_G = T + (x_G - T)/R.$$

Po dosazení tedy vznikne vztah

$$y_G = -20 + (0 - (-20))/4 = -20 + 5 = -15 \text{ dB}.$$

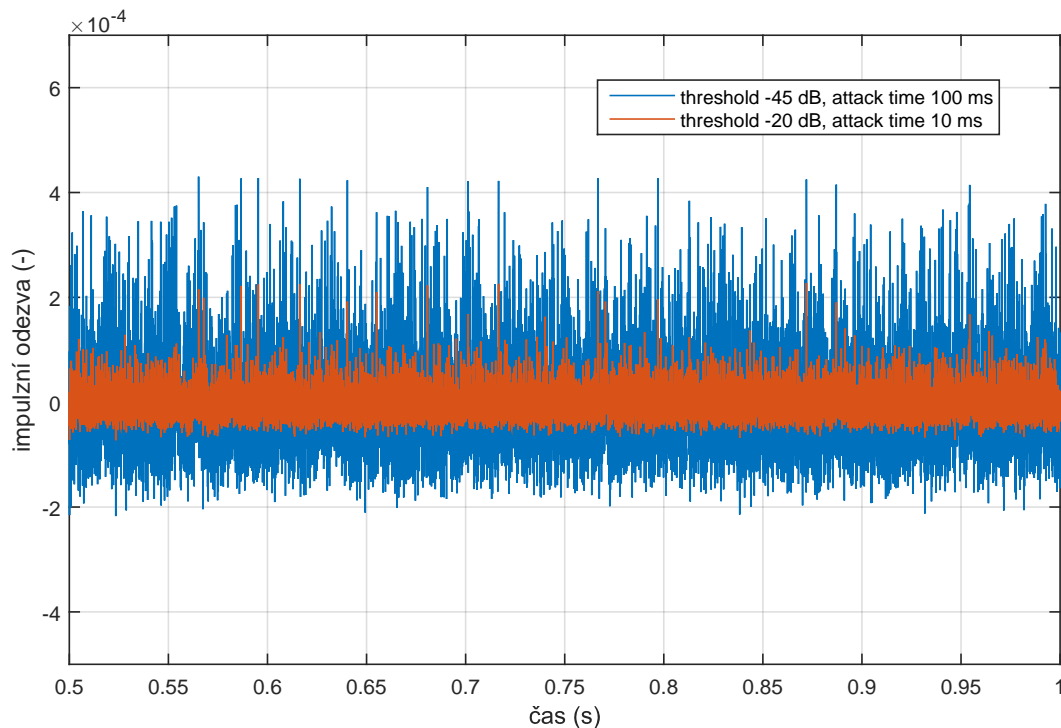
V rámci tohoto výzkumu se při měření vlivu dynamické komprese zvuku na parametr STI měnily dva parametry kompresoru, a to attack time a threshold.

4.2 Měření metodou MLS

K měření pomocí metody MLS se v praxi běžně využívají jeho tři periody, přičemž výsledná impulzní odezva se počítá z té prostřední. Na obrázku 4.2, který zobrazuje aplikované zesílení kompresoru na MLS signál, lze vidět, že změna zesílení proběhne na jeho začátku a dále již zůstává stejné. Graf také zobrazuje průběh zesílení pro tři různá nastavení parametru attack time, je tedy zřejmé, že doba změny zesílení je závislá právě na tomto parametru. Hodnota, na které se následně zesílení ustálí, bude pak dána parametry threshold a ratio.

Co se týče parametru release time, ten v kompresi nebude hrát žádnou roli, protože MLS signál nabývá pouze hodnot 1 a -1 , nedojde tedy k žádnému poklesu hladiny pod threshold.

Naopak u parametrů attack time a threshold se očekává vliv na hodnotu STI při měření impulzní odezvy na první periodě MLS signálu. Na obrázku 4.3 je zobrazen průběh impulzní odezvy pro různá nastavení parametrů attack time a threshold. Je patrné, že při nižších hodnotách parametru threshold a vyšších hodnotách parametru attack time se zvyšuje šum impulzní odezvy, který bude mít zřejmě vliv na výpočet výsledného parametru STI.



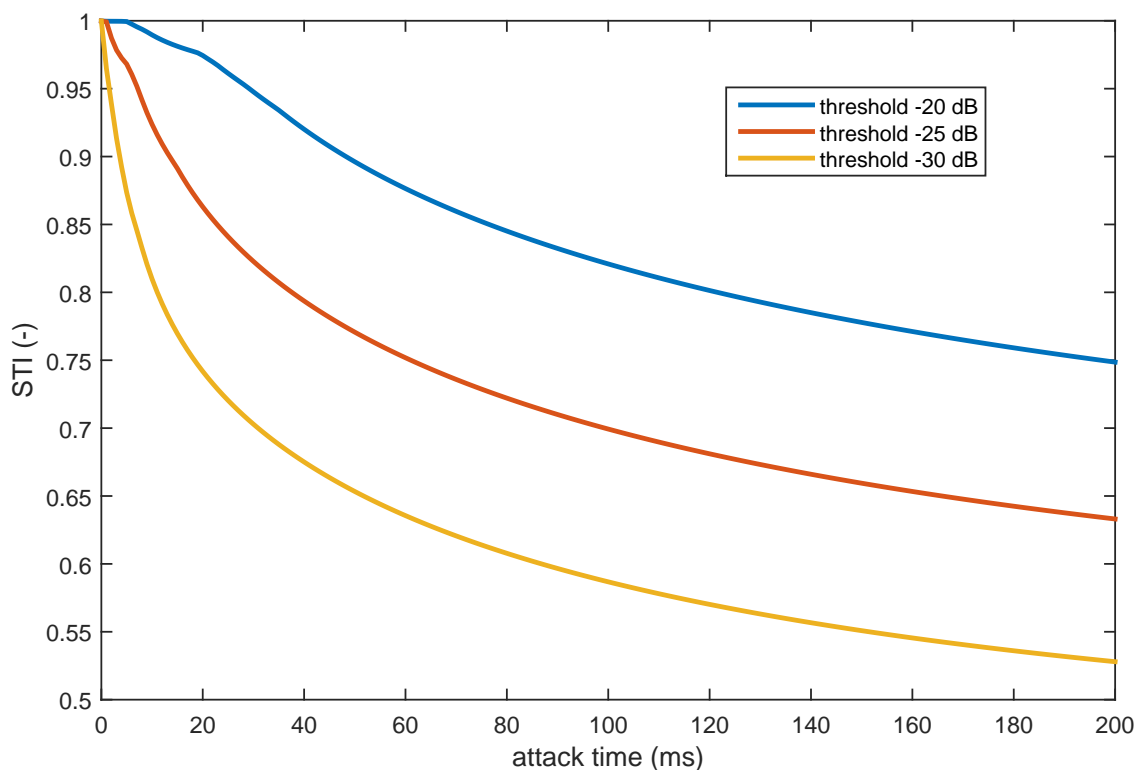
Obrázek 4.3: Průběh impulzní odezvy při různých hodnotách parametrů attack time a threshold.

4.2.1 Vliv změny parametru attack time na STI

Následující měření zkoumalo vliv změny parametru attack time na hodnotu parametru STI. Nejdříve bylo měření prováděno na implementovaném kompresoru s parametry nastavenými takto:

- Threshold: -20 , -25 a -30 dB
- Ratio: 4
- Make-up gain: 0 dB
- Knee Width: 0 dB
- Attack time: 0 až 200 ms s krokem 1 ms
- Release time: 500 ms

Výpočet impulzní odezvy probíhal na první periodě testovacího signálu, aby se projevil vliv změny parametru attack time. Na obrázku 4.4 je zobrazen vývoj parametru STI v závislosti na velikosti parametru attack time při třech různých hodnotách parametru threshold. Je vidět, že s rostoucím parametrem attack time klesá hodnota parametru STI a s rostoucím parametrem threshold klesá STI rychleji.



Obrázek 4.4: Změna parametru STI v závislosti na změně parametru attack time při různých hodnotách parametru threshold.

Při provedení měření na prostřední periodě testovacího signálu se změna parametru attack time neprojevuje. Parametr STI se bude blížit hodnotě 1 (odchylka od

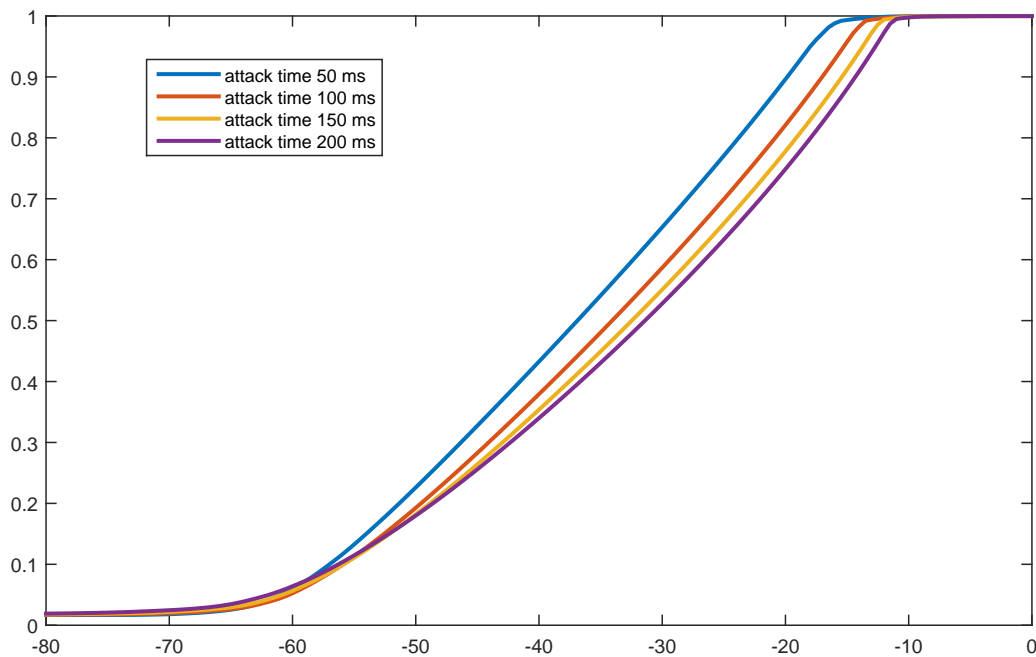
hodnoty 1 je až na 4. desetinném místě). V tomto případě se tedy dá předpokládat, že při měření na prostřední periodě MLS signálu nemá komprese vliv na výsledný parametr STI.

4.2.2 Vliv změny parametru threshold na STI

Následující měření probíhalo opět na první periodě testovacího signálu a v tomto případě byl měněn parametr threshold. Parametry implementovaného kompresoru byly stanoveny takto:

- Threshold: -80 až 0 dB s krokem $0,5$ dB
- Ratio: 4
- Make-up gain: 0 dB
- Knee Width: 0 dB
- Attack time: 50, 100, 150 a 200 ms
- Release time: 500 ms

Na obrázku 4.5 je průběh změny parametru STI v závislosti na nastavení parametru threshold a attack time. Křivka má při různém nastavení parametru attack time stejný tvar, ale při nižších hodnotách je růst parametru STI v závislosti na thresholdu rychlejší.



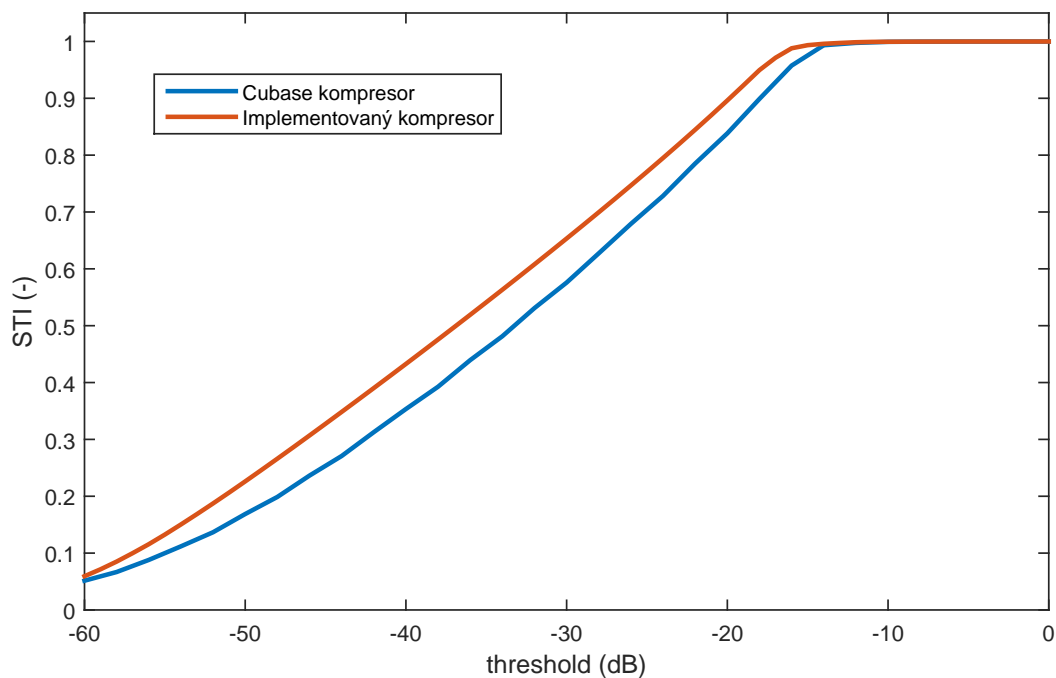
Obrázek 4.5: Změna parametru STI v závislosti na změně parametru threshold a attack time.

4.2.3 Porovnání kompresorů

Pro porovnání implementovaného kompresoru a kompresoru dostupného v programu Cubase byly vybrány parametry:

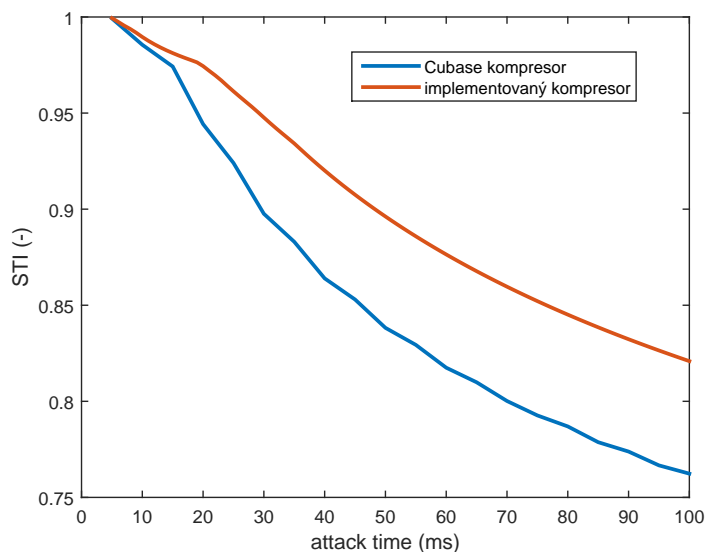
- Threshold: $-60 - 0$ dB s krokem 2 dB u Cubase a 1 dB u implementovaného kompresoru
- Ratio: 4
- Make-up gain: 0 dB
- Knee Width: 0 dB
- Attack time: 50 ms
- Release time: 500 ms

Důvodem menšího rozsahu parametru threshold je skutečnost, že Cubase kompresor má omezenou možnost zadávání rozsahu. Na obrázku 4.6 je porovnání těchto dvou typů kompresoru se stejně nastavenými parametry. Důvodem jejich rozdílného průběhu může být například použití jiného typu detektoru hladiny. Dále také není známo, jak přesně je Cubase kompresor implementován.



Obrázek 4.6: Porovnání funkce implementovaného kompresoru a kompresoru dostupným v programu Cubase při změně parametru threshold.

Obdobně je na obrázku 4.7 zobrazeno srovnání dvou typů kompresorů se stejnými parametry. Tentokrát je ale threshold pevně zvolen jako -20 dB a mění se parametr attack time v rozsahu 5 – 100 ms. Cubase kompresor opět vyšší rozsah neumožňuje, nicméně pro praktické využití kompresoru se delší doba attack time neuzívá.



Obrázek 4.7: Porovnání funkce implementovaného kompresoru a kompresoru dostupném v programu Cubase při změně parametru attack time u metody MLS.

4.3 Měření metodou TSP

Vlastností logaritmického TSP signálu je postupné zvyšování jeho frekvence v čase. V případě této práce začíná testovací signál TSP na 27 Hz, takže komprese signálu proběhne na nižších frekvencích. Výpočet modulační přenosové funkce z impulzní odezvy používá 7 oktávových pásem, z nichž nejnižší je 125 Hz, takže lze předpokládat, že komprese nebude mít na změnu parametru STI téměř žádný vliv.

Při prvním měření byl zjišťován vliv postupného zvyšování parametru attack time na parametr STI při třech různých hodnotách parametru threshold. Parametry implementovaného kompresoru byly tedy nastaveny takto:

- Threshold: -20 , -25 a -30 dB
- Ratio: 4
- Make-up gain: 0 dB
- Knee Width: 0 dB
- Attack time: 0 až 200 ms s krokem 1 ms
- Release time: 500 ms

Všechna tato měření nicméně vykazovala identické výsledky, a to $STI = 1$, kde byla odchylka od hodnoty 1 až na čtvrtém desetinném místě.

Následující měření sledovalo vliv změny parametru threshold při různých hodnotách parametru attack time. Kompresor byl tedy nastaven takto:

- Threshold: -80 až 0 dB s krokem 1 dB
- Ratio: 4

- Make-up gain: 0 dB
- Knee Width: 0 dB
- Attack time: 50 ms, 100 ms, 150 ms a 200 ms
- Release time: 500 ms

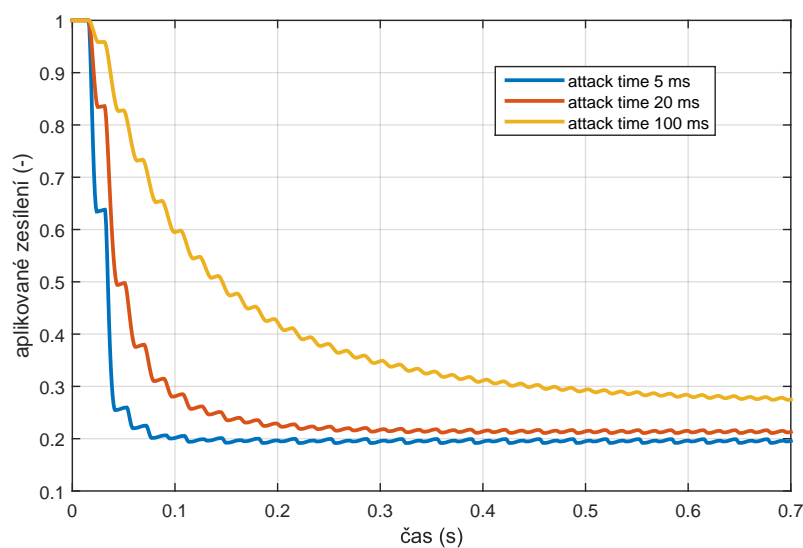
Stejně jako v předchozím případě bylo ale dosaženo stejného výsledku, tedy $STI = 1$ s odchylkou na čtvrtém desetinném místě. Lze tedy předpokládat že při použití této metody nemá kompresor na STI žádný vliv.

4.3.1 Působení kompresoru na signál TSP

V následujícím měření bylo zkoumáno aplikované zesílení kompresoru na signál TSP pro různě volené parametry attack time. Parametry kompresoru byly nastaveny jako

- Threshold: -30 dB
- Ratio: 4
- Make-up gain: 0 dB
- Knee Width: 0 dB
- Attack time: 5, 20 a 100 ms
- Release time: 500 ms

Na obrázku 4.8 je vidět, že průběh signálu TSP není tak hladký jako u signálu MLS (viz obrázek 4.2 na straně 24), který má konstantní absolutní hodnotu. Vlastnosti TSP signálu se promítají i ve vlastnostech aplikovaného zesílení, jehož frekvence také roste v závislosti na čase stejně jako u vstupního signálu. Příčinou tohoto zvlnění je detektor implementovaný v kompresoru, který reaguje okamžitě na špičky v signálu.

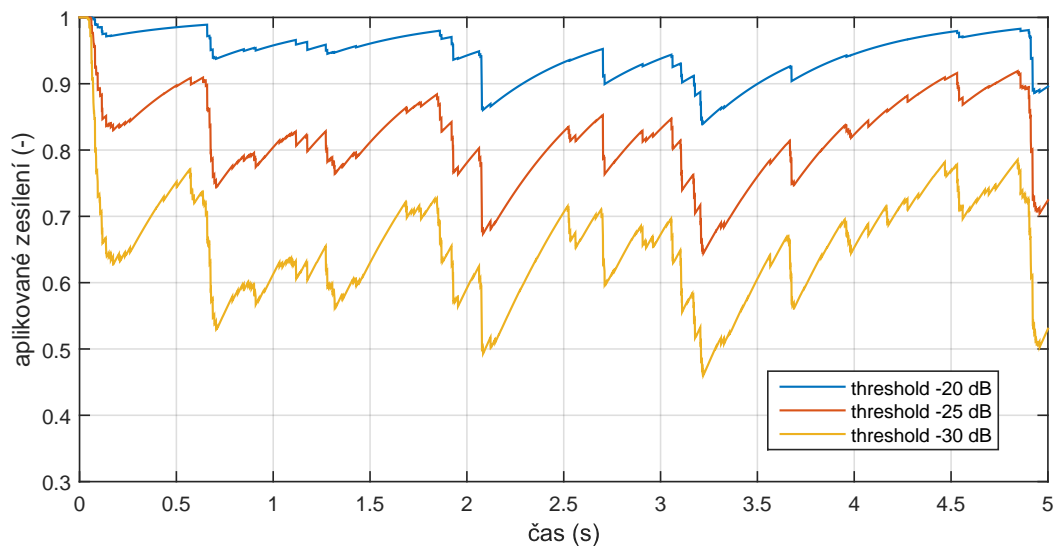


Obrázek 4.8: Aplikované zesílení kompresoru na signál TSP.

4.4 Měření metodou STIPA

Na rozdíl od předchozích dvou metod se při metodě STIPA nevyužívá impulzní odezvy, ale změn modulačních frekvencí v testovacím signálu. Princip měření parametru STI tedy spočívá ve změně modulačních frekvencí při průchodu kompresorem. Díky špičkám v testovacím signálu se bude zesílení kompresoru projevovat více než u předchozích dvou metod. Na obrázku 4.9 jsou zobrazeny tři průběhy aplikovaného zesílení kompresoru na testovací signál STIPA. Parametry kompresoru byly nastaveny takto:

- Threshold: -20 dB, -25 dB a -30 dB
- Ratio: 4
- Make-up gain: 0 dB
- Knee Width: 0 dB
- Attack time: 10 ms
- Release time: 500 ms



Obrázek 4.9: Část aplikovaného zesílení kompresoru na testovací signál metody STIPA.

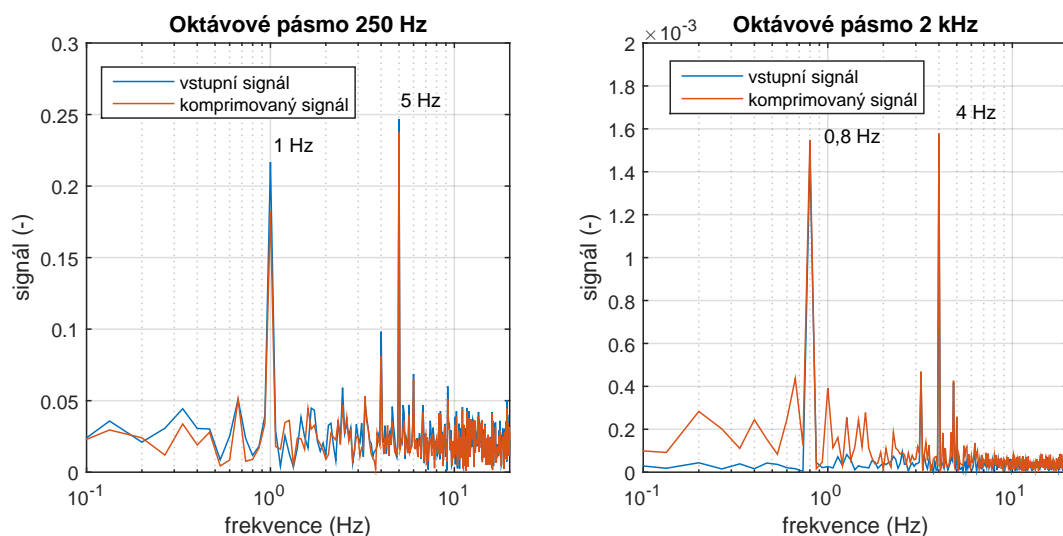
Z těchto křivek je patrné, že kompresor okamžitě reaguje na signál, který překročí hodnotu thresholdu a oproti metodě MLS zde hraje roli i parametr release time. Při nastavení parametru threshold jako -20 dB je dobře patrné, jak se po poklesu signálu pod tuto hodnotu zesílení vrací na původní hodnotu.

4.4.1 Měření parametru STI metodou STIPA

Před měřením samotného parametru STI byla nejprve zkoumána modulační přenosová funkce (MTF), která u metody STIPA nabývá 14 hodnot. Tyto hodnoty by

podle literatury [2] měly být maximálně 1. Pro následující měření byly parametry kompresoru nastaveny takto:

- Threshold: -30 dB
- Ratio: 4
- Make-up gain: 0 dB
- Knee Width: 0 dB
- Attack time: 10 ms
- Release time: 500 ms



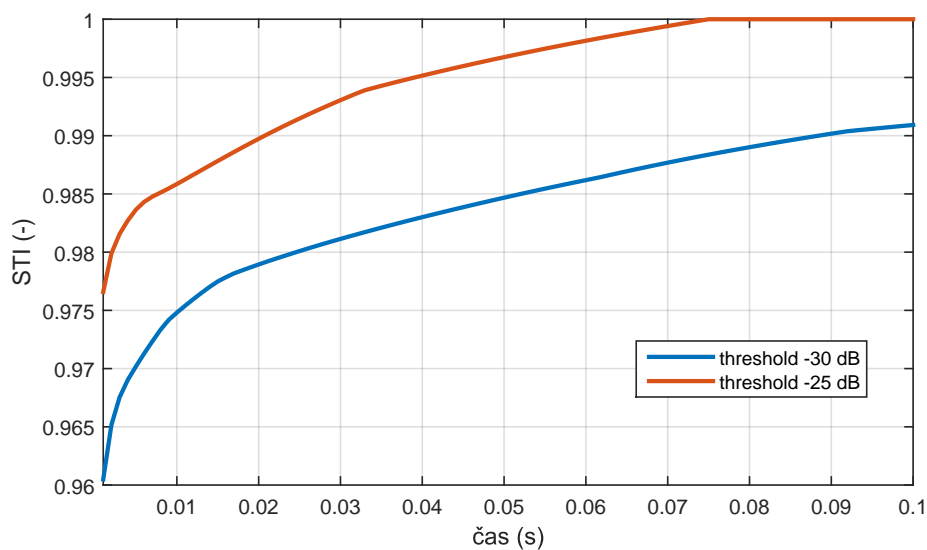
Obrázek 4.10: Hodnoty modulačních frekvencí v oktávových pásmech 250 Hz a 2 kHz.

Na obrázku 4.10 jsou zobrazeny hodnoty jednotlivých modulačních frekvencí pro oktávová pásma 250 Hz a 2 kHz. V oktávovém pásmu 250 Hz jsou to podle tabulky 2.2 na straně 7 frekvence 1 Hz a 5 Hz, pro pásmo 2 kHz jsou to frekvence 0,6 Hz a 4 Hz. Zároveň jsou zobrazeny signály vstupní a komprimovaný. V levém grafu, kde jsou zobrazeny hodnoty modulačních frekvencí v oktávovém pásmu 250 Hz, lze vidět, že hodnoty těchto frekvencí jsou nižší u komprimovaného signálu. Naopak v pravém grafu, tedy oktávové pásmo 2 kHz, jsou hodnoty modulačních frekvencí nižší u vstupního signálu. To znamená, že modulační poměr vypočítaný podle vztahu (2.8) na straně 6 bude vyšší než 1. Pro výpočet parametru STI lze použít pouze hodnoty poměrů maximálně rovno 1, takže hodnoty vyšší jsou omezeny na hodnotu 1. Vyšší hodnoty modulačních frekvencí vpravo na obrázku 4.10 jsou zřejmě zapříčiněny vzniklým šumem, který je v grafu viditelný, na rozdíl od grafu vlevo, kde je vzniklý šum přibližně na stejné úrovni jako je šum vstupního signálu.

Po korekcích z předchozího kroku je zřejmé, že hlavní vliv na hodnotu parametru

STI budou mít modulační přenosové poměry, které nabývají hodnot menších než 1. V následujícím měření byl zkoumán vliv parametru attack time při různých hodnotách parametru threshold na výslednou hodnotu STI. Parametry kompresoru byly nastaveny jako

- Threshold: -25 dB a -30 dB
- Ratio: 4
- Make-up gain: 0 dB
- Knee Width: 0 dB
- Attack time: 10 až 100 ms s krokem 1 ms
- Release time: 500 ms.



Obrázek 4.11: Vývoj parametru STI měřený metodou STIPA v závislosti na parametru attack time při různých hodnotách parametru threshold.

Na obrázku 4.11 je zobrazena změna parametru STI v závislosti na parametru attack time pro dvě různé hodnoty parametru threshold. Průběh parametru STI naznačuje, že při vyšší hodnotě parametru attack time je vyšší i STI. Kvůli hodnotám modulačních poměrů vyšších než 1 je nicméně výsledný parametr STI zkreslený.

Při použití metody STIPA k měření parametru STI v reálném prostředí bude tedy záviset na frekvenční charakteristice jednotlivých částí měřicího řetězce. Z obrázku 4.10 lze předpokládat, že pokud budou dominantní frekvence ve vyšších oktávových pásmech, bude výsledný parametr STI negativně ovlivněn. Pokud naopak budou převládat nižší oktávová pásma, kvůli korekci modulačních poměrů nebude parametr STI výrazně ovlivněn.

Kapitola 5

Závěr

Metodu měření parametru STI lze vedle jiných přenosových kanálů šíření zvuku použít i na zařízení pro dynamickou kompresi zvukového signálu. V této práci bylo měření nejprve provedeno nepřímou metodou využívající impulzní odezvy, a poté zjednodušenou přímou metodou STIPA. Tyto metody byly teoreticky popsány, implementovány v prostředí MATLAB a následně byl experimentálně ověřen vliv činnosti kompresoru na výsledný parametr STI.

Vlastnosti zařízení pro dynamickou kompresi zvuku byly teoreticky nastudovány a dále demonstrovány na kompresoru dostupném v audio programu Cubase a na vlastním implementovaném kompresoru v prostředí MATLAB. Pro tuto implementaci byl vybrán kompresor s detektorem vyhlazené špičky.

Měření vlivu komprese zvuku na parametr STI bylo provedeno při různém nastavení parametrů kompresoru. Jednalo se o parametry *attack time* a *threshold*. Při měření na první periodě MLS signálu byl zaznamenán významnější pokles parametru STI. Tato změna byla očekávána, protože na první periodě probíhá přechodový děj kompresoru. Při měření na druhé periodě MLS signálu byl pak výsledný parametr STI pokaždé roven 1.

Metoda TSP vykazovala pro každé nastavení kompresoru stejnou hodnotu parametru $STI = 1$ s odchylkou na čtvrtém desetinném místě. Aplikace této metody tedy ukázala, že použití kompresoru nemá na srozumitelnost řeči téměř žádný vliv.

Parametr STI měřený metodou STIPA byl při různém nastavení kompresoru odchýlen od hodnoty 1 maximálně na druhém desetinném místě. Bylo nicméně zjištěno, že hodnoty modulační přenosové funkce jsou ve vyšších oktávových pásmech vyšší než 1 kvůli zvýšení šumu v těchto pásmech. Lze tedy předpokládat, že dále bude záležet na frekvenční charakteristice případné další části měřicího řetězce.

Dosažené výsledky této práce lze dále využít k dalšímu výzkumu, například ke zkoumání vlivu kompresoru na parametr STI v reálném prostředí.

Literatura

- [1] STEENEKEN, H. J. M., T. HOUTGAST: A physical method for measuring speech-transmission quality, *J. Acoust. Soc. Am.*, 1980, vol. 67, No. 1, s. 318–326.
- [2] IEC 60268-16:2011: *Sound system equipment – Part 16: Objective rating of speech intelligibility by speech transmission index*, 4. vyd.
- [3] HRDINA, Z., F. VEJRAŽKA: *Signály a soustavy*, 1. vyd. Praha: ČVUT, 1998, 243 s.
- [4] STAN, G., J. EMBRECHTS, D. ARCHAMBEAU: Comparison of Different Impulse Response Measurement Techniques, *J. Acoust. Eng. Soc.*, 2002, vol. 50., No. 4, s. 249–262.
- [5] KADLEC, F.: *Zpracování akustických signálů*, 1. vyd. Praha: ČVUT, 2002, 189 s.
- [6] TUČEK, J.: *Měření impulsové odezvy v akustice*, Diplomová práce, České vysoké učení technické v Praze, Fakulta elektrotechnická, Katedra fyziky.
- [7] MORIYA, N., Y. KANEDA: Study of harmonic distortion on impulse response measurement with logarithmic time stretched pulse, *Acoust. Sci. & Tech.*, 2005, s. 462–464
- [8] GIANNOULIS, D., M. MASSBERG, J. D. REISS: Digital Dynamic Range Compressor Design – A Tutorial and Analysis, *J. Audio Eng. Soc.* 2012, vol.60, No. 6, s. 399–408

Příloha A

Seznam souborů se zdrojovými kódy

Implementace

Funkce

- fconv.m
- mls.m
- tspgen_lg.m
- stipa_gen.m
- compressor.m
- mls_response.m
- tsp_response.m
- mtfIR.m
- mtfSTIPA.m
- STIPA.m
- STI.m

Spouštěcí skripty

- main_mls_matlab_comp.m
- main_mls_cubase_comp.m
- main_tsp_matlab_comp.m
- main_tsp_cubase_comp.m
- main_stipa_matlab_comp.m
- main_stipa_cubase_comp.m