

Děkuji doc. Ing. Janu Holubovi Ph.D. za pomoc při vedení bakalářské práce. Mé poděkování patří též Gabriele Horníčkové, Bc. Marii Valentové za spolupráci při získávání nahrávek pro praktickou část práce a Mgr. Aleně Havlíčkové za korekturu.





## ZADÁNÍ BAKALÁŘSKÉ PRÁCE

Student: **Filip Valenta**

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

Název tématu česky: **Algoritmus pro měření kvality ozvučení prostoru**

Název tématu anglicky: **Sound Distribution Quality Measurement Algorithm**

### Pokyny pro vypracování:

Proveďte shrnutí stávajících metod pro měření kvality ozvučení prostoru (např. algoritmu STI, RASTI a STIPA) a rozbor jejich výpočetní a časové náročnosti.

Vytvořte testovací databázi akustických vzorků, nasnímaných v různých akustických prostředích. Na základě provedeného rozboru navrhnete algoritmus, který s využitím obecného akustického signálu poskytne obdobný výsledek jako STI/RASTI. S využitím vytvořené databáze posuďte jeho přesnost.

### Seznam odborné literatury:

- [1] Zwicker, E., Fastl, H.: Psychoacoustics: Facts and Models, Springer, 1999
- [2] IEC 60849, Sound Systems for Emergency Purposes
- [3] IEC 60268-16, Objective Rating of Speech Intelligibility by Speech Transmission Index.

Vedoucí bakalářské práce: doc. Ing. Jan Holub, Ph.D.

Datum zadání bakalářské práce: 26. listopadu 2014

Platnost zadání do<sup>1</sup>: 31. srpna 2016



Doc. Ing. Jan Holub, Ph.D.  
vedoucí katedry



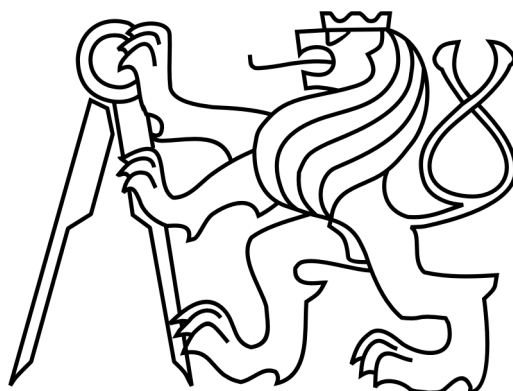
Prof. Ing. Pavel Ripka, CSc.  
děkan

V Praze dne 30. 11. 2014

<sup>1</sup> Platnost zadání je omezena na dobu tří následujících semestrů.



ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE  
FAKULTA ELEKTROTECHNICKÁ  
KATEDRA MĚŘENÍ



## BAKALÁŘSKÁ PRÁCE

Algoritmus pro měření kvality ozvučení prostoru

**Autor:** Filip Valenta

**Vedoucí práce:** doc. Ing. Jan Holub, Ph.D.

Praha, 2016

**Název práce:** Algoritmus pro měření kvality ozvučení prostoru

**Autor:** Filip Valenta

**Katedra (ústav):** Katedra měření

**Vedoucí bakalářské práce:** Doc. Ing. Jan Holub, Ph.D.

**e-mail vedoucího:** holubjan@fel.cvut.cz

**Abstrakt** V předložené práci shrnuji stávající metody pro měření kvality ozvučení prostoru (např. algoritmu STI/RASTI a STIPA) a provedl jsem rozbor jejich výpočetní a časové náročnosti. Ve shrnutí objasňuji, jaké metody jsou výhodné pro jaké aplikace, jak fungují a které metody se již přestávají používat. Vytvořil jsem testovací databázi akustických vzorků, nasnímaných v různých akustických prostředích. Na základě provedeného rozboru navrhl algoritmus, který funguje na základě algoritmu STI, který s využitím obecného akustického signálu poskytne obdobný výsledek jako STI/RASTI. S využitím vytvořené databáze jsem posoudil jeho přesnost a použitelnost v praxi.

**Klíčová slova:** Akustika, STI, Srozumitelnost řeči

---

**Title:** Sound Distribution Quality Measurement

**Author:** Filip Valenta

**Department:** Departement of Measurement

**Supervisor:** Doc. Ing. Jan Holub, Ph.D.

**Supervisor's e-mail address:** holubjan@fel.cvut.cz

**Abstract** In the present work I summarize the existing methods for measuring quality of room (eg. The algorithm STI / RASTI and STIPA) and I performed an analysis of its computational and time consuming. In summary, I illustrate what methods are advantageous for those applications, how they work and which methods are already cease to use. I have created a testing database of acoustic samples taken at different acoustic environments. Based upon my analysis, I designed an algorithm that works on the basis of STI algorithm that uses a general acoustic signal gives a similar result as STI / RASTI. Using the database created, I assess its accuracy and applicability in practice.

**Keywords:** Acoustics, STI, Speech Intelligibility

Prohlašuji, že jsem svou bakalářskou práci napsal samostatně a výhradně s použitím citovaných pramenů.  
Souhlasím se zapůjčováním práce a jejím zveřejňováním.

V Praze dne 23. května 2016

Filip Valenta





# OBSAH

<b>Abstrakt</b>	<b>6</b>
<b>Zadání práce</b>	<b>7</b>
<b>Seznam zkratk a pojmů</b>	<b>11</b>
<b>1 Úvod</b>	<b>12</b>
<b>2 Metody měření kvality ozvučení</b>	<b>13</b>
2.1 Vývoj metod . . . . .	13
2.2 Důležité parametry . . . . .	13
2.3 Srovnání metod . . . . .	14
2.3.1 Subjektivní měření s použitím lidských řečníků a posluchačů . . . . .	14
2.3.2 Prediktivní měření založené na fyzikálních parametrech . . . . .	14
2.3.3 Objektivní měření, které je zajištěno specifickými testovacími signály . . . . .	15
2.4 STI . . . . .	15
2.4.1 Princip . . . . .	15
2.4.2 Výpočet STI . . . . .	18
2.4.3 Limitace STI metody . . . . .	22
2.4.4 Použití řeči při měření STI . . . . .	23
2.4.5 Binaural STI . . . . .	23
2.4.6 Podobnost s AI . . . . .	23
2.4.7 Nevýhodnost a použitelnost . . . . .	23
2.5 RASTI . . . . .	24
2.5.1 Obecně . . . . .	24
2.5.2 Rozdíly STI a RASTI . . . . .	24
2.5.3 Nevýhodnost a použitelnost . . . . .	24
2.6 STIPA . . . . .	25
2.6.1 Obecně . . . . .	25
2.6.2 Princip . . . . .	25
2.6.3 Rozdíly STI a STIPA . . . . .	25
2.6.4 Nevýhodnost a použitelnost . . . . .	26
2.7 STITEL . . . . .	26
2.7.1 Princip . . . . .	26
2.7.2 Rozdíly STI a STITEL . . . . .	26
2.7.3 Nevýhodnost a použitelnost . . . . .	26
2.8 Další používané metody . . . . .	26
2.8.1 AI (Articulation index) . . . . .	27
2.8.2 SAI (Speech Audibility index) . . . . .	27
2.8.3 SII (Speech intelligibility index) . . . . .	27

<b>3</b>	<b>Rozbor výpočetní a časové náročnosti</b>	<b>29</b>
3.1	Náročnost jednotlivých metod . . . . .	29
3.1.1	Počátky metod měření kvality ozvučení a srozumitelnosti . . . . .	29
3.1.2	Současný stav . . . . .	29
3.2	Omezení do budoucna . . . . .	30
<b>4</b>	<b>Výhodnost metod pro rozdílné aplikace</b>	<b>31</b>
4.1	Aplikace měření kvality ozvučení a srozumitelnosti . . . . .	31
4.2	Jaké metody používáme . . . . .	31
<b>5</b>	<b>Čím se měří kvalita ozvučení</b>	<b>32</b>
5.1	Moderní přenosné měřicí přístroje . . . . .	32
5.1.1	2270 Hand Held Analyzer (Brüel & Kjaer) . . . . .	32
5.1.2	Norsonic Nor131 . . . . .	33
5.1.3	Bedrock STIPA meter SM50 . . . . .	34
5.1.4	NTi XL2 analyzer . . . . .	35
<b>6</b>	<b>Testovací databáze akustických vzorků</b>	<b>36</b>
6.1	Tvorba databáze . . . . .	36
6.1.1	Místa, na kterých probíhalo nahrávání . . . . .	37
6.2	Využití databáze mimo rozsah bakalářské práce . . . . .	37
<b>7</b>	<b>Testovací algoritmus na základě STI</b>	<b>38</b>
7.1	Postup vytváření algoritmu . . . . .	38
7.1.1	STI algoritmus . . . . .	38
7.1.2	$STI_{FV}$ algoritmus . . . . .	38
7.1.3	Implementace kódu do Matlabu . . . . .	40
7.2	Použitelnost algoritmu v praxi . . . . .	41
7.3	Porovnání $STI_{FV}$ s STI . . . . .	42
7.3.1	GUI . . . . .	43
7.3.2	Výhody . . . . .	43
7.3.3	Nevýhody . . . . .	43
<b>8</b>	<b>Měření</b>	<b>44</b>
8.1	Přímé a nepřímé měření . . . . .	44
8.2	Kde a jak probíhalo nahrávání . . . . .	45
8.3	Zpracování dat . . . . .	45
8.3.1	Srovnání signálu z krátkou a dlouhou IR . . . . .	49
8.3.2	Zkouška algoritmu . . . . .	50
<b>9</b>	<b>Závěr</b>	<b>51</b>
	<b>Literatura</b>	<b>52</b>
	<b>Přílohy</b>	<b>I</b>

## SEZNAM ZKRATEK A POJMŮ

IEC	Mezinárodní elektrotechnická komise, která vypracovává mezinárodní elektrotechnické normy
ČSN	Označení českých technických norem
STI	Speech Transmission Index
RASTI	Rapid Speech Transmission Index nebo Room Acoustics Transmission Index
STIPA	Speech Transmission Index for Public Address
STITEL	Speech Transmission Index for Telecommunication Systems
AI	Articulation Index
SII	Speech Intelligibility Index
SAI	Speech Audibility Index
RT60	Doba dozvuku - pokles zvukové hladiny o 60 po vypnutí zdroje
IR	Impulzní odezva
SNR	Odstup signál-šum
MTF	Modulační přenosová funkce
CIS	Common Intelligibility Scale - stupnice používaná pro měření STI
RMS	Efektivní hodnota
FIF	Frekvenčně důležitá funkce
CAA	Civil Aviation Authority

## ÚVOD

Ozvučení místností a veřejných prostor je v dnešní době odvětvím elektrotechniky, které je velmi důležité. Lidé chtějí žít v komfortu a nemají rádi přílišný hluk nebo nesrozumitelné hlášení v dopravních prostředcích. Rozhlasy jsou instalovány i na veřejných místech a všude, kde je často velký počet lidí. Při špatné prostorové akustice se tak může stát, že nebudeme rozumět důležitému hlášení, které nás informuje například o odjezdu našeho vlaku či odletu letadla. V horším případě se může jednat o nebezpečnou situaci jako požár, teroristický útok nebo blížící se zemětřesení. Akustická měření se dnes provádějí při kolaudaci každé stavby. Akustika je důležitá při úpravě koncertních sálů a přednáškových místností či veřejných prostor, jako jsou nádraží, letiště či obchodní centra a při instalaci jejich informačních kanálů (reproduktorů). Je důležité, aby informace, které jsou sdělovány, byly také srozumitelně přijaty.

Pokud posloucháme řečníka či vnímáme informace z reproduktorů, je spousta prvků, které rozhodují, jak dobře bude řeč přijata. Záleží na barvě hlasu a hlasitosti projevu, ve většině případů je však nejdůležitější prostředí, ve kterém se řečník (reproduktor) a posluchač nacházejí. Objektivním posouzením kvality se zabývá a řeší jej algoritmus STI a jeho derivace (RASTI, STIPA, STITEL). Tímto algoritmem se budu zabývat ve své práci a to konkrétně úpravou současného algoritmu STIPA, který je v dnešní době nejhojněji rozšířený a používaný. Dříve se pro rozpoznávání řeči a měření jeho kvality používaly také jiné algoritmy jako je AI, SII, SAI nebo %Alcons.

V práci se budu v první části zabývat současnými metodami pro měření kvality ozvučení prostoru a srozumitelnosti řeči. Vycházím z norem, které se tohoto problému týkají, a to z IEC 60268-16, Objective Rating of Speech Intelligibility by Speech Transmission Index a z ní vyplývající evropské a české normy (EN 60268-16 a ČSN EN 60268-16).

V druhé, praktické části práce, budu vysvětlovat vznik mého algoritmu a databáze vzorků nahrávek; jaké úpravy jsem na algoritmu provedl - na rozdíl od algoritmu STI a STIPA - a proč jsem algoritmus takto upravoval. Nakonec rozebírám výsledky měření algoritmu STI a mého algoritmu.

## METODY MĚŘENÍ KVALITY OZVUČENÍ

### 2.1 Vývoj metod

Metody měření srozumitelnosti řeči se postupně vyvíjely již v první polovině 20. století. Za počátek vědeckého bádání nad problémem srozumitelnosti můžeme považovat model srozumitelnosti - Articulation Index - představený v roce 1947 [3]. Jednotlivé metody si prošly svým vývojem.. Některé se stále používají, jiné se ale staly slepou vývojovou větví tohoto oboru. V této kapitole, se kromě STI, RASTI a STIPA metod, podívám také na další čtyři metody měření kvality ozvučení, které nejsou tak známé jako STI. V minulosti se vývoj metod zaměřoval na vylepšování a zjednodušování těchto metod v oblasti zrychlení měření nebo ušetření výpočetního výkonu. V současnosti se vývoj zaměřuje spíše na kvalitu měření a na co nejpřesnější napodobení reálné situace komunikace mezi řečníkem a posluchačem [31].

### 2.2 Důležité parametry

V mnoha situacích, kdy posloucháme něčí projev (řeč), je tento signál určitým způsobem poškozen nebo není úplný. Část informace, která byla řečena, se znehodnotí v přenosovém kanálu. Přenosový kanál představuje místnost, ve které jsme. Přenosový kanál je při měření srozumitelnosti velmi důležitý a ovlivňuje celkový výsledek měření [18]. Srozumitelnost v dané situaci ovlivňuje pouze přenosový kanál. Do přenosového kanálu neřadíme řečníka ani posluchače, kteří mohou svými odlišnostmi a chybami srozumitelnost také ovlivňovat. Při měření předpokládáme „normálního“ řečníka<sup>1</sup> i posluchače<sup>2</sup> [15].

Při měření kvality ozvučení existují tyto parametry, které ovlivňují celkovou srozumitelnost projevu, a tudíž ovlivňují i výsledek měření systému [11]:

- Odstup signál-šum (SNR)
- Maskování<sup>3</sup>
- Hladina akustického tlaku

---

<sup>1</sup>Řečník, který netrpí žádnou vadou řeči a umí správně artikulovat.

<sup>2</sup>Posluchač, který má podle lékařského vyšetření normální sluch.

<sup>3</sup>Hladina intenzity šumu je nad hladinou signálu.

- Doba dozvuku RT60
- Frekvenční odezva
- Šum pozadí
- Odrazy
- Zkreslení
- Vzdálenost řečníka a posluchače

## 2.3 Srovnání metod

Obecně můžeme rozlišit tři metody měření srozumitelnosti [31].

### 2.3.1 Subjektivní měření s použitím lidských řečníků a posluchačů

Pro subjektivní testování se používá několik typů hlasu, které poslouchá několik různých posluchačů. Posluchači hodnotí podle určitých kritérií jeho srozumitelnost a další prvky. Pro testování se často používají různá slova, čísla a věty v kombinaci s různou kvalitou řeči.

Nevýhodou této metody je její časová náročnost, počet lidí, kteří se testování musí účastnit a limitace lidským faktorem, který není bezchybný.

### 2.3.2 Prediktivní měření založené na fyzikálních parametrech

Prediktivní měření se zaměřuje na ztrátu srozumitelnosti v důsledku například omezeného frekvenčního přenosu, maskování šumu, dozvuku, ozvěn nebo nelinearity. Důležitým prvkem je také posluchač a jeho sluchové vlastnosti: schopnost rozlišení, sluchové maskování nebo recepční práh.

Jeden z prvních modelů předpovědi účinku srozumitelnosti na přenosovou cestu byl prezentován Steinbergem (1947) a byl základem pro AI.

### 2.3.3 Objektivní měření, které je zajištěno specifickými testovacími signály

Toto měření se vyvíjí již několik desítek let. V dnešní době jsou vyvinuty přístroje přímo pro objektivní měření srozumitelnosti. Jedna metoda pro objektivní měření srozumitelnosti byla vynalezena pány Steenkenem a Hougastem. Tato metoda je založena na použití konkrétního testovacího signálu. Kvalita přenosu řeči je pak odvozena od analýzy přijatého signálu a je vyjádřena pomocí indexu (STI). Přehled metod a jejich vlastnosti je uvedeno v tab. 2.1.

Typ metody	Omez. pásma	Nelin. zkreslení	Dozvuky, ozvěny	Délka měř.
full STI	ano	ano	ano	15 minut
STI-3	ano	ano	závisí na podm.	4 minuty
STITEL	ano	závisí na podm. <sup>4</sup>	závisí na podm.	15 s
RASTI	ne	ne	ano	15 s
STIPA	ano	závisí na podm.	ano	30 s

Tabulka 2.1: Přehled STI metod, pro co jsou navrženy a jaké jsou jejich vlastnosti

## 2.4 STI

STI byl vynalezen v roce 1971 Holanďany Tammo Houtgastem a Herman Steenekenem, kteří pracovali ve firmě TNO a měli za úkol provést velmi dlouhou sérii akustických měření pro nizozemskou armádu. Tato metoda byla přijata akustickým společenstvím v Americe v roce 1980 a stala se součástí standardu s IEC 60268-16. Standard STI byl několikrát doplňován a dnes má již čtyři revize. Současný, nejaktuálnější, je z roku 2011 [27]. Ve třetí revizi z roku 2003 byl u STI představen koncept maskování, který má velký vliv na srozumitelnost [26].

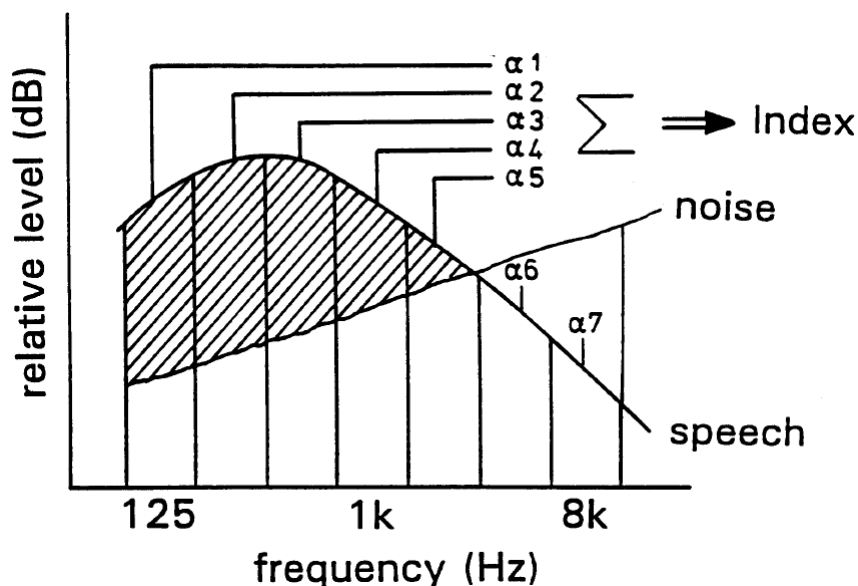
STI je měřítkem kvality přenosu řeči. Měří pouze některé fyzikální vlastnosti přenosového kanálu, které jsou pro měření důležité. STI je založen na zjištění, že nepravidelnosti a výkyvy řečových signálů nesou většinu důležitých informací, které se týkají srozumitelnosti řeči. Tyto nepravidelnosti a odchylky (nazývané jako modulace) kvantifikujeme jako funkci modulačního kmitočtu  $f_m$ , vytvářejícího modulační spektrum [1]. Vyjadřuje, do jaké míry přenosový kanál ovlivňuje srozumitelnost řeči a nahrazuje ucho posluchače. Pomocí stupnice se snaží určit, jak bude na daném místě posluchač rozumět. Některé funkce má společné s AI. Je to vlastně jeho předchůdce, který má několik funkcí navíc.

### 2.4.1 Princip

Měření STI vyžaduje speciální testovací signál, ze kterého získáváme SNR pro každé oktávové pásmo, ve kterém je STI měřen.

Pro výpočet STI je specifické, že testovací signál a jeho vyhodnocení je odolné vůči šumu pozadí, nelineárnímu zkreslení a dalším degradacím, které mohou ovlivňovat výsledek měření. SNR získané z testovacího signálu se měří v každém oktávové pásmu ve frekvenčním rozsahu 125 Hz až 8000 Hz [27]. Tento rozsah je určen rozsahem reálného řečového spektra. Ostatní frekvence nás proto nezajímají a mohou nám také měření rušit a zhoršit. Kromě měření SNR jednotlivě pro každé pásmo má také každé pásmo svůj váhový koeficient. Všechna pásma tudíž

nemají při výpočtu STI stejnou váhu. Při měření musíme také dát pozor na případné maskování šumem. Frekvence, které jsou maskované, pro měření nejsou vhodné. Váhové koeficienty i maskování jsou názorně zobrazeny na obr. 2.1.



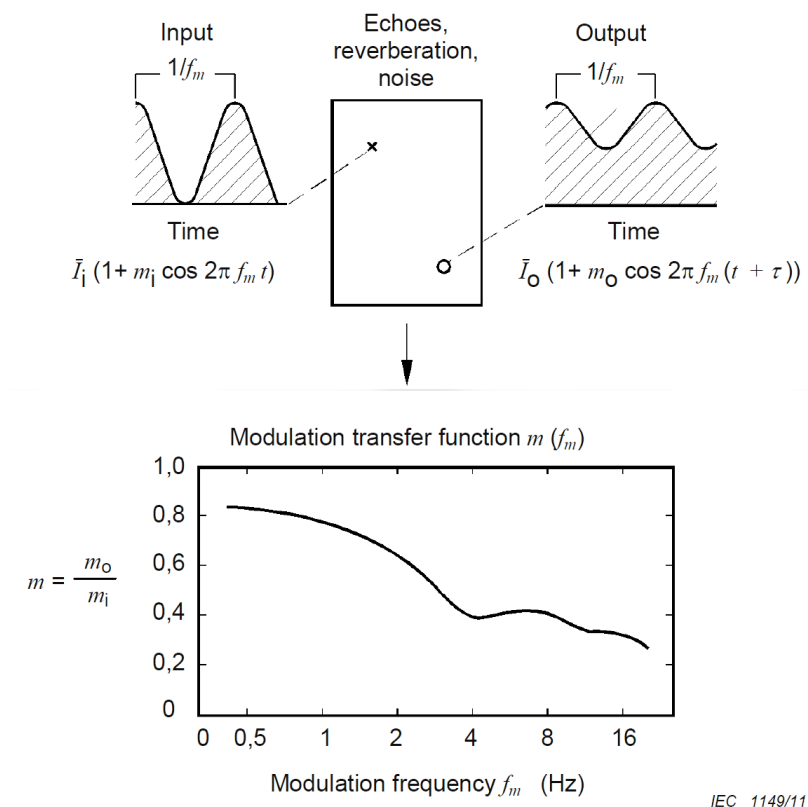
Obrázek 2.1: Frekvenční spektrum řeči maskované šumem a jednotlivá oktávová pásma pro výpočet STI [18]

Při výpočtu se hodnotí relativní rozdíl mezi vstupním a výstupním signálem, který je generován v rámci definovaného systému (modulační přenosová funkce MTF). Při měření porovnáváme modulační index  $m_i$  vyslaného testovacího signálu a modulační index  $m_0$  signálu přijatého [1]. Jak vypadá rozdíl mezi přijatým a vyslaným signálem a jak vypadá vypočítaná MTF je ukázáno na obr. 2.2. Pokud je rozdíl velký, srozumitelnost bude menší a stejně tak STI. STI pro své měření používá umělé signály skládající se z deterministického šumu<sup>5</sup>, který má dlouhou úroveň RMS<sup>6</sup> tak, jako řeč.

<sup>5</sup>Opakem je náhodný šum

<sup>6</sup>Root mean square - efektivní hodnota signálu





Obrázek 2.2: Modulační přenosová funkce [1]

STI testovací signál pro výpočet STI indexu tedy obsahuje sedm oktávových pásem šumu od 125 Hz po 8 kHz. Každé pásmo je modulováno čtrnácti modulačními frekvencemi. Tyto modulační frekvence jsou vybrány v třetinooktávových krocích od 0,63 Hz do 12,5 Hz, což je celkem 98 kombinací a tedy 98 testovacích signálů pro měření modulačních indexů. Cílem STI je určit, kolik modulace zůstalo v každém oktávovém pásmu a v každé modulační frekvenci [7].

Tři důležité prvky při použití algoritmu STI [11]:

- Modulační frekvence mezi 0,63 a 12,5 Hz jsou při měření srozumitelnosti důležité a všechny jsou pro měření stejně důležité.
- Analýza nahrávky prostoru se provádí odděleně na sedmi oktávách v rozsahu 125 Hz až 8 kHz
- Každé oktávové pásmo přispívá do STI jinou vahou podle koeficientů  $\alpha$  a  $\beta$ .

Kolik zůstane v přijímaném signálu modulace, to je vyjádřeno modulační přenosovou funkcí (MTF). MTF může být určena za různých podmínek - v pokoji, při velkém hluku, za použití filtrů ze strany sluchově postiženého posluchače nebo prostřednictvím sluchadla. MTF je založen na myšlence zachování nepravidelností obálky signálu [7].

Abychom stanovili účinnost signálu a množství modulace, srovnáváme testovací signál v každém frekvenčním pásmu. Informace je v lidské řeči uložena v podobě kolísání intenzity signálu.

Ztráta hloubky modulace tedy způsobí ztrátu srozumitelnosti. Testovací signál STI funguje podobně. Má uměle navržené modulační frekvence a měříme, jak přenosový kanál snižuje hloubku modulace.

Základní princip měření STI spočívá v měření umělého signálu místo řeči reálného člověka. Srozumitelnost při měření tohoto signálu zachycuje mikrofon, který představuje ucho. Srozumitelnost je vyjádřena číslem o hodnotě mezi 0 a 1. Pro ohodnocení srozumitelnosti řeči jsou používány nejčastěji dvě stupnice. STI stupnice (tab.: 2.2) a CIS stupnice (tab.: 2.3), kde  $CIS = 1 + \log(STI)$  [11]. Minimální hodnoty, které jsou podle standardu IEC 60628-16 považovány za minimální, jsou 0,45 - 0,50. V praxi každý jednotlivý kanál, který ukazuje hodnoty pod 0,35 STI, není vhodný pro použití v jakýchkoliv podmínkách [15].

Hodnota STI	Ohodnocení srozumitelnosti řeči
$0,75 \leq STI$	Výborný
$0,60 \leq STI < 0,75$	Dobrý
$0,45 \leq STI < 0,60$	Ucházející
$0,30 \leq STI < 0,45$	Špatný
$STI < 0,30$	Mizerný

Tabulka 2.2: STI stupnice[5]

Hodnota STI	Ohodnocení srozumitelnosti řeči
$0,88 \leq STI$	Výborný
$0,78 \leq STI < 0,88$	Dobrý
$0,65 \leq STI < 0,78$	Ucházející
$0,48 \leq STI < 0,65$	Špatný
$STI < 0,48$	Mizerný

Tabulka 2.3: CIS stupnice

## 2.4.2 Výpočet STI

Původní měření STI bylo velmi zdlouhavé. V dnešní době jsou ale vynalezeny metody, které zkrátí čas měření s minimálními dopady na výsledek. Výhoda STI indexu je, že měření může proběhnout přímo, ale také je můžeme zpracovat nepřímo, nahrát si impulzní odezvu místnosti a výsledky zpracovat doma u počítače. Při měření STI je důležité myslet na to, jaký máme zdroj zvuku. Pro zdroj zvuku při měření je důležitá velikost zdroje, jeho směrovost, umístění, hladina akustického tlaku a frekvenční odezva. Pro zdroj zvuku při měření je důležitá jeho velikost, jeho směrovost, umístění, hladina akustického tlaku a frekvenční odezva. V případě použití reproduktoru místo řečníka musíme myslet na směrovost reproduktoru, která by měla být co nejvíce podobná směrovosti reálného řečníka.

## Přímá metoda

STI se skládá z 98 testovacích signálů. Každý z těchto signálů obsahuje pouze jednu modulační frekvenci pro jednu konkrétní oktávu a každý je měřen průměrně 10 sekund. Celkové měření STI tedy trvá kolem 15 minut. Tak dlouhé měření v jednom místě limituje z časového hlediska použitelnost STI metody. Z tohoto důvodu se měření full STI používá pouze výjimečně. Výsledkem, který je pro nás důležitý a který simuluje řeč reálného člověka, je redukce modulačního indexu.  $\Omega$  Z hlediska testovacího signálu je MTF pro každé oktávové pásmo derivována korelací intenzity obálky se signály sinus a cosinus, konkrétní délky a modulačních frekvencí. Modulační přenos je pro každou modulační frekvenci spočítán jako první derivace modulační hloubky přijatého signálu [9].

$$m_{dr_{k, f_m}} = 2 \frac{\sqrt{[\sum I_k(t) \cdot \sin(2\pi f_m t)]^2 + [\sum I_k(t) \cdot \cos(2\pi f_m t)]^2}}{\sum I_k(t)} \quad (2.1)$$

kde

$f_m$  je modulační frekvence v Hz

$t$  je čas v s

$I_k(t)$  je intenzita obálky pro konkrétní oktávové pásmo

Výsledný modulační index určíme jako poměr přijatého ku vyslanému testovacímu signálu:

$$m_{k, f_m} = \frac{m_{dr_{k, f_m}}}{m_{dt_{k, f_m}}} \quad (2.2)$$

kde

$m_{dr_{k, f_m}}$  je derivace modulační hloubky přijatého signálu

$m_{dt_{k, f_m}}$  je derivace modulační hloubky vyslaného signálu

Provedení korekce MTF použitím sluchového maskování

$$I_{am,k} = I_{k-1} \cdot amf \quad (2.3)$$

kde

$amf$  je maskovací faktor,  $amf = 0,000316$

$I_{k-1}$  je hladina intenzity pro předchozí frekvenční pásmo

$$m'_{k, f_m} = m_{k, f_m} \frac{I_k}{I_k + I_{am,k} + I_{rt,k}} \quad (2.4)$$

a je dále interpretován jako efektivní hodnota SNR, kterou vyjádříme:

$$SNR_{eff}(k, f_m) = 10 \log \left( \frac{m'_{k, f_m}}{1 - m'_{k, f_m}} \right) \quad (2.5)$$

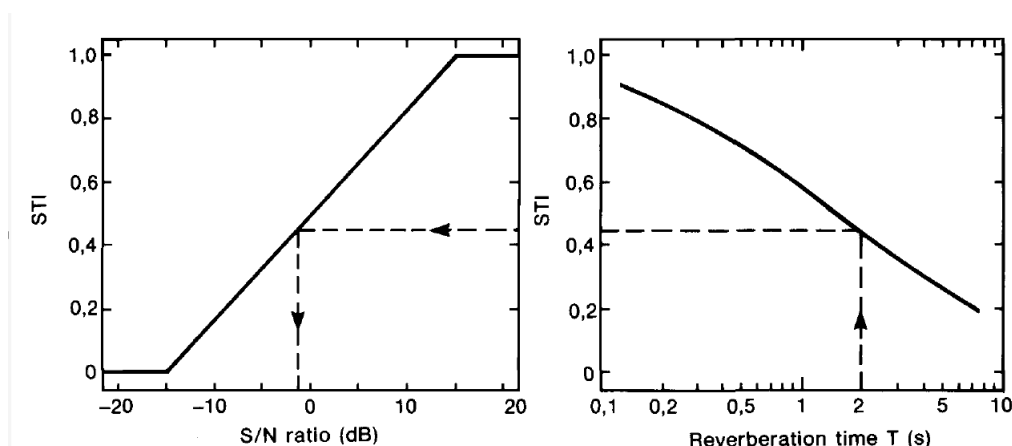
Získané hodnoty SNR musíme znormalizovat tak, že:[30]

$$STI = 1 \text{ pokud } SNR \geq 15 \text{ dB} \quad (2.6)$$

$$STI = 0 \text{ pokud } SNR \leq 15 \text{ dB} \quad (2.7)$$

protože podle STI, je SNR mezi -15 dB a 15 dB lineárně úměrné srozumitelnosti. Normalizaci SNR vidíme na obr. 2.3 vlevo. Z SNR počítáme rovnou STI. Proto můžeme vyjádřit vztah mezi

SNR a STI. Stejně tak můžeme vyjádřit vztah mezi RT a STI, který je na obr. 2.3 vpravo.



Obrázek 2.3: Teoretický vztah mezi SNR a STI a mezi RT a STI [29]

Index, který se už blíží STI pro každé oktákové pásmo a modulační frekvenci, se spočítá ze SNR podle vzorce:

$$TI_{k,f_m} = \frac{SNR+15}{30} \quad (2.8)$$

Tento vzorec průměrujeme přes všechny modulační frekvence a získáme tak pouze sedm hodnot MTI pro každé oktákové pásmo k:

$$MTI_k = \frac{1}{14} \sum_{f_m}^{14} TI_{k,f_m} \quad (2.9)$$

Nakonec, podle revidovaného vzorce  $STI_r$ , se index získá váženým součtem MTI pro všechna oktáková pásma, příslušné korekce redundance a váhové faktory [18]:

$$STI = \alpha_1 MTI_1 - \beta_1 \sqrt{MTI_1 \cdot MTI_2} + \alpha_2 MTI_2 - \beta_2 \sqrt{MTI_2 \cdot MTI_3} + \dots + \alpha_7 MTI_7 \quad (2.10)$$

kde i se v případě full STI rovná 98 kombinacím.

### Váhové faktory a korekce redundance

Faktor  $\alpha_k$  představuje váhový faktor pro každou jednotlivou oktávu a  $\beta_k$  je korekční faktor redundance. Tato korekce redundance souvisí s příspěvkem okolních frekvenčních pásem.

Revidovaný vzorec  $STI_r$  se do standardu IEC 60268-16 dostal až v roce 2002. Tento vzorec rozlišuje faktory zvlášť pro mužský a ženský hlas. Hodnoty váhových a korekčních faktorů redundance pro muže i pro ženu jsou uvedeny v tab. 2.4. Pod hodnotami faktorů je uvedena minimální hodnota hladiny intenzity testovacího signálu [18].

Oktávové pásmo (Hz)		125	250	500	1k	2k	4k	8k
Muži	$\alpha$	0,085	0,127	0,230	0,233	0,309	0,224	0,173
Muži	$\beta$	0,085	0,078	0,065	0,011	0,047	0,095	-
Ženy	$\alpha$	-	0,117	0,223	0,216	0,328	0,250	0,194
Ženy	$\beta$	-	0,099	0,066	0,062	0,025	0,076	-
Min. hodnota intenzity	L	46	27	12	6,5	7,5	8	12

Tabulka 2.4: Váhové a korekční faktory pro mužský a ženský hlas STI,

kde,

$$\sum_{k=1}^7 \alpha_k - \sum_{k=1}^6 \beta_k = 1 \quad (2.11)$$

### Nepřímá metoda

Jak vypočítat STI index nepřímou metodou bylo uvedeno až v poslední edici standardu IEC 60268-16 z roku 2011 [26]. STI může být modifikováno několika způsoby, abychom zkrátili čas potřebný k měření jednoho místa. Pokud je systém, který měříme, považován za lineární, pak existuje mnoho řešení. Budící signál může být modulován všemi modulačními frekvencemi najednou a rozdělení na složky proběhne až po přijetí signálu použitím filtrů nebo Fourierovy analýzy. Znaměřší metoda, kterou budu používat pro výpočet STI v mém algoritmu, a která snižuje čas potřebný pro měření systému je založena na spočítání komplexní modulační přenosové funkce (MTF) z impulzní odezvy testovacího signálu (IR) místnosti (systému), ze které je určena hodnota STI. Pokud můžeme impulzní odezvu považovat jako dobrou odezvu místnosti s exponenciálně rozpadající se obálkou, charakterizovanou dobou dozvuku (RT60), pak modulační přenosovou funkci (MTF) na frekvenci  $f$ , můžeme spočítat přímo z hodnoty RT60 a efektivního odstupu signál-šum (v dB). Zjednodušený vzorec, který nebere v potaz maskování a práh slyšení nám dá následující vztah [20]:

$$m_k(f_m) = \frac{\left| \int_0^{\infty} h_k(t) e^{-j2\pi f_m t} dt \right|}{\int_0^{\infty} h_k(t)^2 dt} \cdot \left[ 1 + 10^{-SNR/10} \right]^{-1} \quad (2.12)$$

Kde

$h_k(t)$  je impulzní odezva oktávového pásma  $k$

$f_m$  je modulační frekvence

$SNR$  je odstup signál-šum v dB

Ze vzorce vidím, že se skládá ze dvou částí. Z části, kde dochází k odrazům, a z části SNR. V případě, že je hladina hluku pozadí zanedbatelná, není potřeba počítat SNR a začleňovat ho do výpočtů. Nejprve ale musím vědět, jak zjistit, že hladina hluku pozadí je již zanedbatelná, a kdy ještě ne. Předpokládá se, že jestliže při výpočtu klesne STI o méně než 5% oproti výpočtu STI bez šumu, pak mohu šum pozadí zanedbat [18]. Dále postupuji obdobně jako při přímé metodě.

- Použiji vzorec 2.5 pro výpočet SNR
- Provedu normalizaci hodnot podle vzorců 2.6 a 2.7
- Spočítám přenosový index pro každou modulaci a každou oktávu (viz. 2.8)

- Spočítám modulační index (viz. 2.9)
- Nakonec spočítám STI z rovnice 2.10

### **Impulzní odezva IR**

Při nepřímém měření je důležité dávat pozor na kvalitu impulzní odezvy. Impulzní odezva je ten prvek, díky kterému bude STI index změřen a pokud zvolíme špatnou impulzní odezvu, bude výsledek špatný a ovlivní nám vyhodnocení testu. Co bych měl při nepřímé metodě u impulzní odezvy kontrolovat? [7]

- Odstup impulzní odezvy od šumu. Odstup by měl být maximální možný a minimálně 20 dB.
- Impulzní odezva musí začínat impulzem systému.
- Délka funkce impulzní odezvy musí být dostatečně dlouhá. Jako pravidlo pro určení délky funkce je, že délka funkce by měla být dvojnásobkem doby dozvuku. Pro kontrolu můžeme sledovat, jestli šum pozadí dominuje nad signálem impulzní odezvy.
- Podle poslední revize normy IEC 60268-16 by doba trvání IR neměla být menší než polovina RT60 a měla by trvat nejméně 1,6 s [2].

### **Testovací signál**

Při použití nepřímé metody pro měření STI používáme testovací signály, ze kterých získáváme IR. Pro tyto signály se nejvíce hodí MLS<sup>7</sup>, bílý šum a sweep. Při použití sweep signálu nám výsledek mohou ovlivňovat přístroje, které způsobují nelineární zkreslení. K chybě může dojít v případě dlouhé doby dozvuku RT60, protože složky zkreslení se mohou dostat do IR. Při použití MLS složky zkreslení naopak mohou vypadat jako šum a nejsou v IR snadno rozpoznatelné.

### **2.4.3 Limitace STI metody**

Tato metoda měření ideálně odráží změny přenosového kanálu, které jsou podstatné pro srozumitelnost. Je ale důležité si uvědomit, že STI je pouze model pro zjednodušení měření srozumitelnosti a nikdy není schopen nahradit lidskou zkušenost. Největší rozdíl mezi testováním srozumitelnosti pomocí STI a reálnou lidskou komunikací je, že při testování srozumitelnosti používáme testovací signál v podobě šumu nebo podobného, umělého signálu.

Výsledek STI ovlivňuje spousta faktorů, které jsou v každých podmínkách jiné. Jaké faktory mohou zkreslovat a ovlivňovat STI? [1]

- Frekvenční posuny
- Crossover zkreslení<sup>8</sup>
- Jitter<sup>9</sup>
- Výpadky přenosu
- Ozvěny

---

<sup>7</sup>Maximum length sequence je typ pseudonáhodné binární sekvence.

<sup>8</sup>Nízko-úrovňové části signálu jsou špatně přeneseny.

<sup>9</sup>Nežádoucí odchylka jedné nebo více charakteristik periodického signálu.

- Nelineární zkreslení
- Šum

#### 2.4.4 Použití řeči při měření STI

V IEC-60268-16 v příloze J jsou uvedeny celkem tři metody, při kterých se používá k měření STI řeč [1].

Jedna metoda, která k měření používá reálnou řeč je založena na odstranění všech tichých míst řečového signálu (místa mezi jednotlivými slovy) před samotným výpočtem. Aby byly překonány spektrální rozdíly mezi signály a úroveň signálu byla dostačující, je metoda založena na váhovém filtrování signálu. Tím se snažíme předejít vlivu tichých míst na výsledek. Při výpočtu  $L_A$  jsou brány v úvahu pouze ty části řečového signálu, které obsahují potřebné informace.

Takto vyfiltrovaný signál je rozdělen na části délky 10 až 20 ms a pro každou část je počítána její energie. Všechna tato energie je zachycena v distribučním histogramu. Použitím tohoto histogramu je vypočítána RMS úroveň signálu řeči. Na základě tohoto postupu jsou tiché části signálu vynechány.

#### 2.4.5 Binaural STI

Při binaurálním<sup>10</sup> poslouchání jsou zvuky prostorově konkrétnější. Binaurální STI zlepšuje srozumitelnost v místech, kde se vyskytuje hodně šumu na pozadí a velké ozvěny. Binaurální zpracování STI je výhodné především v těchto místech a odlišné výsledky jsou v takovýchto místech viditelné [12]. Pro obecné místnosti, jako je domácí pokoj, jsou výsledky STI srovnatelné s monaurálním STI. Binaurální verze STI je založena na interaurální křížové korelaci, kde se signály z jednoho i druhého mikrofону měří odděleně v jednotlivých oktávních pásmech. Interaurální křížová korelace je vypočtena pro každé oktávné pásmo, čímž se získá počet časově zpožděných reprezentací. Tyto reprezentace jsou zpracovány jako v případě jednonárodního STI [19]. Vývoj binaurálního algoritmu STI je teprve v počátcích. Binaurální algoritmus STI není součástí standardu.

#### 2.4.6 Podobnost s AI

STI index v mnohém vychází z Articulation Index a některé funkce má podobné. STI stejně jako AI měří SNR v každém oktávním pásmu v rozsahu frekvenčního spektra 125 Hz až po 8000 Hz. Signál STI je - na rozdíl od AI - amplitudově modulován. Tato modulace způsobuje, že určité nepravidelnosti, které se vyskytují v normální řeči, jsou promítnuty i do STI signálu. STI narozdíl od AI a SII pracuje velmi dobře v prostředí, kde dochází k odrazům a dalším degradacím, ovlivňujícím srozumitelnost, a v komunikačních systémech.

#### 2.4.7 Nevýhoda a použitelnost

Nevýhoda STI je, že pro měření a výpočet používá všech 98 kombinací, a proto měření trvá velmi dlouho. Někdy se proto metoda STI z tohoto důvodu nepoužívá a nahrazuje se jinými, které jsou jednodušší a tudíž i rychlejší. U těchto zjednodušených metod výsledky nebudou tak přesné jako při použití STI. Ne vždy ale potřebujeme výsledky tak přesné, jak nám udává STI.

<sup>10</sup>Přenesení zvuku od zdroje, do uší posluchače v co nejvěrnější podobě. Funguje na principu prostorového zvuku.

Patnáct let zkušeností s vývojem a používáním STI ukázalo potřebu mnoha zlepšení metody pro různé podmínky a aplikace: například při velmi omezeném frekvenčním přenosu nebo při používání reproduktoru. Musíme také rozlišovat pohlaví mluvčího a individuální přístup k vyjádření souhlásek a samohlásek. Tato zlepšení byla popsána v roce 1992 [1].

## 2.5 RASTI

### 2.5.1 Obecně

Metoda RASTI se používá při měření v uzavřených malých místnostech (Room Acoustic STI). Při měření v místnosti musíme myslet na to, že zde dochází k odrazům, dozvukům a ozvěnám. Ty ale v případě malých místností nejsou tak velké a neovlivňují výsledek měření. [6] Při mluvení v uzavřené místnosti s delším dozvukem, je pro posluchače důležitá hlasitost řečníka a její distribuce v čase. Princip měření RASTI vychází z metody STI a zaměřuje se na měření srozumitelnosti přímo mezi řečníkem a posluchačem - bez přítomnosti reproduktorů nebo mikrofonů. Pokud dodržíme zásady pro měření s touto metodou, může být, díky vyloučení reproduktorů a mikrofonů, velmi zjednodušená, aniž by zjednodušení mělo velký vliv na výsledek. Zjednodušení nám přinese výhodu v rychlosti měření, která je v případě RASTI 10 - 15 sekund [11]. RASTI používá (stejně jako STI) pro výpočet srozumitelnosti MTF a může být měřena přímo i nepřímo podle 2.4.2 (přímá metoda) a 2.4.2 (nepřímá metoda).

### 2.5.2 Rozdíly STI a RASTI

RASTI měří pouze devět hodnot místo 98, které měří metoda STI. Měří na dvou oktávových pásmech (500 Hz a 2 kHz), pro čtyři modulační frekvence (pro 500 Hz) a pět modulačních frekvencí (pro 2 kHz) [3]. Pro vyhodnocení se používá stejná STI/CIS stupnice jako pro metodu STI.

Modulační frekvence	$f_1$	$f_2$	$f_3$	$f_4$	$f_5$
Oktávové pásmo: 500 Hz	1 Hz	2 Hz	4 Hz	8 Hz	-
Oktávové pásmo: 2000 kHz	0,7 Hz	1,4 Hz	2,8 Hz	5,6 Hz	11,2 Hz

Tabulka 2.5: Body měření RASTI metody

### 2.5.3 Nevýhodnost a použitelnost

RASTI metoda poskytuje rychlé vyhodnocení srozumitelnosti a objevení míst se špatnou srozumitelností nebo místa s nízkou odolností proti šumu pozadí. Metodou se podle standardů provádí měření i pro některé elektroakustické aplikace (např.: CAA specifikace 15 pro PA<sup>11</sup> systémy letadel) [8]. Nevýhodou RASTI je, že při nelinearitách v systému, jako použití reproduktorů nebo mikrofonů, tato metoda přestává měřit správně. Z těchto důvodů se tato metoda postupně přestala používat a dnes ji zcela ve všech aplikacích nahradila metoda STIPA. V poslední revizi normy EN 60268-16:2011 se již metoda RASTI uvádí jako zastaralá.

<sup>11</sup>Public address.



## 2.6 STIPA

### 2.6.1 Obecně

STIPA vznikla a byla zařazena do standardu IEC 60268-16 až ve třetí revizi v roce 2003 [26] a stala se jeho nejdůležitější součástí, protože je v dnešní době nejpoužívanější a nejrozšířenější. Vznikla za účelem zlepšení přesnosti srozumitelnosti měření oproti RASTI a zároveň zjednodušení a zrychlení původní full-STI metody. Metoda dokáže pracovat s dozvukem místnosti a dalšími elektro-akustickými a akustickými efekty, které ovlivňují kvalitu přenosu řeči a tím pádem srozumitelnost v daném prostředí. Je důležité, aby byl přenosový kanál stabilní a reprezentativní. Měření by mělo probíhat za normálního provozu, kvůli úrovni šumu pozadí a dalším vlastnostem, které mohou hrát při měření roli. Například pokud testujeme obchodní centrum, měli bychom měřit při normálním provozu obchodu a za přítomnosti zákazníků, kteří zde srozumitelnost ovlivňují. Tato metoda se nejvíce používá v uzavřených prostorech, jako jsou kostely, nádraží nebo letištní haly. Odtud se metoda nazývá STIPA jako public address (PA) [18]. STIPA také dobře funguje pro uzavřené místnosti ve kterých v současnosti nahradila RASTI metodu (viz. 2.5).

### 2.6.2 Princip

STIPA používá celý rozsah oktávových pásem (125 Hz až 8 kHz), a modulační frekvence od 0,63 Hz po 12,5 Hz, ale pro každé pásmo pouze dvě frekvence [11]. Celkově tedy probíhá 14 měření a získáme 14 modulačních indexů. Na celé měření potřebujeme pouze jeden testovací signál. Délka měření je podobná jako RASTI, a to 10 - 15 sekund [1]. Protože STIPA používá všech sedm oktávových pásem, může být metoda použita téměř ve všech případech, včetně prostorů s elektro-akustickými efekty a uzavřených místností.

Oktávová pásma (Hz)	125	250	500	1000	2000	4000	8000
1. modulační frekvence (Hz)	1,00	1,00	0,63	2,00	1,25	0,80	2,50
2. modulační frekvence (Hz)	5,00	5,00	3,15	10,00	6,25	4,00	12,50

Tabulka 2.6: Body měření STIPA metody

### 2.6.3 Rozdíly STI a STIPA

STIPA metoda je založena opět na metodě STI (viz.2.4). Na rozdíl od RASTI ale nesnižujeme počet oktávových pásem, které měříme, ale každé pásmo modulujeme pouze dvěma frekvencemi (oktávová pásma a k nim přiřazené modulační frekvence jsou ukázané v tab. 2.6). Pro vyhodnocení se používá stejná STI/ CIS stupnice jako pro metodu STI.

Na rozdíl od full-STI metody, která se (z důvodu dlouhého trvání měření) používá výjimečně, se metoda STIPA používá často.

Pro metodu STIPA se proto nepoužívá nepřímá metoda, kdy srozumitelnost počítáme z IR až později, například pomocí Matlabu. Při měření musíme použít přímo konkrétní STIPA testovací signál, jak je uvedeno v IEC-60268-16, a používáme speciální přístroje, které byly vyvinuty přímo pro měření STIPA (viz. 5.1). Přesnost měření STIPA je 0,02 - 0,03 (na stupnici STI) [15]. Vyšší přesnosti lze dosáhnout průměrováním více měření.

## 2.6.4 Nevýhodnost a použitelnost

Metoda STIPA by se neměla používat v místech, kde dochází k frekvenčním posunům, kde jsou vokodéry<sup>12</sup>, má šum pozadí, který je impulzivní, nebo v místech, které obsahují silně nelineární zkreslující přístroje. Tato metoda se používá téměř pro všechny současné aplikace měření srozumitelnosti. Plně nahradila metodu RASTI a z hlediska rychlejšího a jednoduššího měření se používá i místo STI.

## 2.7 STITEL

### 2.7.1 Princip

Tato metoda funguje opět na principu metody STI stejně jako předešlé tři metody RASTI (2.5) a STIPA (2.6). Rozdíl je v tomto případě větší, a to z hlediska použití. STITEL se používá pro měření srozumitelnosti telekomunikačních kanálů a testování telefonních systémů [31]. K měření používá všech sedm oktávních pásem, ale na každé pásmo používá pouze jednu modulační frekvenci (viz. tab. 2.7).

Oktávné pásmo (Hz)	125	250	500	1000	2000	4000	8000
Modulační frekvence	1,12	11,33	0,73	2,83	6,97	1,78	4,53

Tabulka 2.7: Modulační frekvence pro metodu STITEL

### 2.7.2 Rozdíly STI a STITEL

STITEL používá stejně jako STI k měření oktávních pásem v rozsahu 125 Hz až 8 kHz. Tato pásma nemoduluje čtrnácti modulačními frekvencemi, ale pouze jednou, pro každé pásmo jinou. V tomto případě nejsou použity stejné modulační frekvence jako pro full-STI. Zjednodušení pouze na sedm měření podstatně snižuje měřicí čas, který je v případě STITEL metody mezi 10 a 15 sekundami pro jedno měření.

### 2.7.3 Nevýhodnost a použitelnost

STITEL není vhodný pro práci v podmínkách, kde dochází ke zkreslení v časové oblasti a pro nelineární systémy. Jeho použití se, již podle názvu, zaměřuje především na komunikační systémy, jako jsou rádiové linky, interkomy, telefonní spojení, a další. Protože funguje podobně jako STIPA, platí pro ni stejné podmínky používání (2.6.4).

## 2.8 Další používané metody

Pro měření řeči neexistují jen metody na základě STI, ale jsou zde také další metody, které fungují více či méně odlišně. Metody zde nebudu detailně popisovat, jelikož toto téma není obsahem mé práce, ale zmiňuji je, protože s tématem úzce souvisí.

<sup>12</sup>Zařízení, které analyzuje a syntetizuje řeč a zvuk.

### 2.8.1 AI (Articulation index)

Jeden z prvních modelů popisujících předpovídání efektů přenosové cesty na srozumitelnost řeči byl prezentován pány Frenchem a Steinbergem v roce 1947. Jejich práce položila základ pro metodu s názvem Articulation Index, která byla popsána a prezentována panem Kryterem (1962) a stala se součástí standardu ANSI S3.5-1969.

AI pomocí stupnice rozsahu od nuly do jedné udává poměr SNR, který dorazí k posluchači v daném prostředí. Základní interpretace hodnoty AI je vyjádřena buď jako číslo od nuly do jedné (při nule posluchač neslyší nic, co bylo řečeno a při jedničce uslyší všechno), nebo pomocí procent [3]. Je založen na rozdělení odezvy komunikačního systému do dvaceti frekvenčních pásem, kde každé pásmo částečně přispívá hodnotou celkové srozumitelnosti a celkový příspěvek všech pásem je součtem příspěvků jednotlivých pásem.

V současné době je AI nahrazen SII.

### 2.8.2 SAI (Speech Audibility index)

Je definován jako poměr užitečného řečového signálu (přímá řeč a rychlé dozvuky), který je nad hladinou efektivního šumu (aktuální šum plus pomalé/pozdní dozvuky). SAI je z části podobný STI. V základní teorii SAI můžeme předpokládat, že užitečná řečová informace ve všech frekvenčních pásmech je rovnoměrně distribuovaná přes rozsah 30 dB, od -15 dB, do 15 dB. Tedy, že příspěvek daného frekvenčního pásma pro SAI vzrůstá od 0 po jeho maximální hodnotu, jako efektivní odstup signál-šum v konkrétním pásmu vzrůstá od -15 do 15 dB. Když odstup signál-šum dosáhne 15 dB ve všech významných frekvenčních pásmech, SAI je 1 nebo 100%. [4] Pokud budeme předpokládat, že odstup signál-šum je stejný ve všech frekvenčních pásmech, pak je SAI dán jako [9]:

$$SAI = \frac{SNR + 15}{30} \quad (2.13)$$

Kde

SAI je Speech Audibility Index s mezí 0 a 1

SNR je celkový decibellový rozdíl mezi užitečným řečovým signálem a efektivním šumem.

### 2.8.3 SII (Speech intelligibility index)

Při vzniku STI a SII byly mezi těmito metodami velké rozdíly. Postupem času a díky vývoji se v posledních revizích rozdíly obou metod smývají a mají mnoho společného. Například měření ekvivalentní hladiny řeči a hladiny šumu nebo koncept MTF, který byl v SII uveden v druhé revizi. I když obě dvě metody používají MTF pro udržení efektivního SNR na dobré úrovni, nosná pásma a modulační frekvence jsou jiné.

Ve standardu SII jsou zapsány čtyři postupy měření, které SII povoluje. Každý používá jiný počet a jinou velikost frekvenčního pásma. Zde jsou vypsány (Postupy jsou řazeny sestupně podle přesnosti)[28]:

- Kritické pásmo (21 pásem)
- Třetinooktávové pásmo (18 pásem)
- Rovnoměrně přispívající kritické pásmo (17 pásem)

- Oktávové pásmo (6 pásem)

SII je velmi přesná metoda měření srozumitelnosti v případě dobrých podmínek. Má široké pásmo měření (150 Hz až 8,5 kHz) zejména v postupu kritického pásma (viz. 2.8.3). SII se počítá z akustických měření řeči a šumu. Měří se v rozsahu mezi 0 a 1 a vyjadřuje míru srozumitelnosti řeči. V praxi při SII 0,5 znamená, že řečový signál je slyšitelný, srozumitelný a použitelný v daném prostředí a za daných podmínek. Pro většinu konverzací je SII = 0,5 velmi blízko 100% srozumitelnosti řeči pro normálně slyšící posluchače. SII je stejně jako AI nepřímá metoda měření srozumitelnosti. Přenosovou funkcí (MTF) SII můžeme určit srozumitelnost řeči. SII je v současnosti náhradou AI a je mu v mnoha ohledech podobný. SII je popsán ve standardu z roku 1997 a generalizuje rámec výpočtů ze standardu z roku 1969 (AI). Tento rámec byl navržen, aby uvolnil definování základních vstupních proměnných potřebných pro výpočet. Další rozdíly zahrnují korekci vzestupného šíření maskování a vysoko prezentačních stupňů.

Pro výpočet SII (AI) potřebujeme určité základní informace. Potřebujeme frekvenční specifické informace o řečových úrovních, úrovni šumu, sluchovém prahu a důležitosti řeči. SII je vypočítána stanovením podílu řečové informace, která je slyšitelná přes konkrétní počet frekvenčních pásem. Abychom SII mohli spočítat, porovnáme úroveň špiček řeči sluchového prahu nebo RMS hladinu šumu v konkrétních frekvenčních pásmech. Podíl slyšitelné řeči ve frekvenční oblasti je násobena relativní důležitostí dané frekvenční oblasti. Na konec jsou výsledné hodnoty sečteny přes celkový počet použitých frekvenčních pásem.[17]

$$SII = \sum_{i=1}^n I_i A_i \quad (2.14)$$

Kde

- n je počet jednotlivých frekvenčních pásem použitých ve výpočtu
- $I_i$  představuje důležitost daných frekvenčních pásem (i) pro porozumění řeči. Hodnoty pro  $I_i$  jsou známy jako frekvenčně důležitá funkce (FIF).
- $A_i$  představuje hodnoty s rozsahem od 0 do 1, které indikují poměr řečového signálu, které jsou slyšitelné v daném frekvenčním pásmu.

Současný SII standard je flexibilní v tom, že si můžeme vybrat, jak konkrétní chceme frekvenční měření (více frekvencí specifikuje měření).

---

## ROZBOR VÝPOČETNÍ A ČASOVÉ NÁROČNOSTI

### 3.1 Náročnost jednotlivých metod

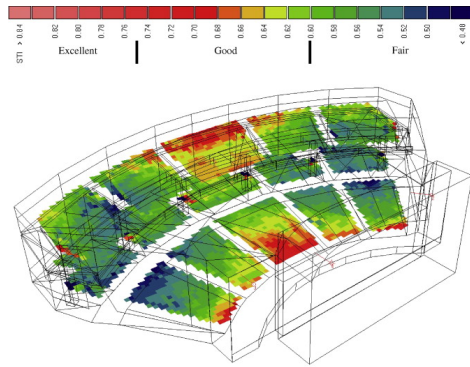
#### 3.1.1 Počátky metod měření kvality ozvučení a srozumitelnosti

V druhé polovině 20. století nebyla výpočetní technika rozvinuta do dnešní podoby a po druhé světové válce byla teprve v počátcích. Proto se tyto metody počítaly ručně (AI, SII,...). V 70. letech, kdy byl vyvinut STI již počítače byly dostupné, ale jejich výpočetní výkon nedosahoval takových možností jako dnes. Proto výpočet těchto metod trval dlouho, nebo se počítače k měření a zpracování výpočtů vůbec nepoužívaly. Dříve se také více vyráběly embedded systémy, které sloužily pouze jednomu konkrétnímu účelu [14].

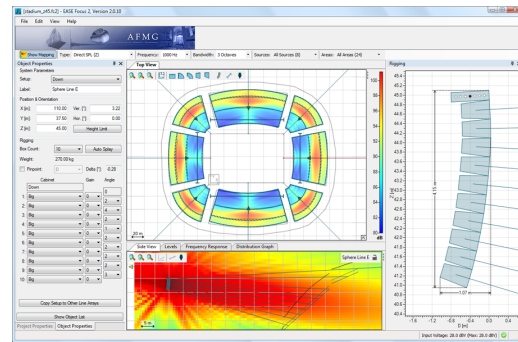
#### 3.1.2 Současný stav

V současnosti máme pro měření kvality ozvučení a srozumitelnosti dostatečný výpočetní výkon a v případě měření srozumitelnosti není problém s výpočtem během několika vteřin. Proto se dnes kromě embedded systémů používají ve velké míře také osobní počítače a laptopy. Ty nejsou kalibrovány pro počítání měření kvality ozvučení, ale jsou levnější a pro informativní měření nám postačí. Pokud ovšem potřebujeme v případě STI, které se dnes již téměř nepoužívá, při přímém měření není časová náročnost problémem nedostatku výpočetního výkonu, obtíže spíše souvisí s potřebou nějakou dobu přehrávat nahrávku, která se teprve posléze vyhodnotí.

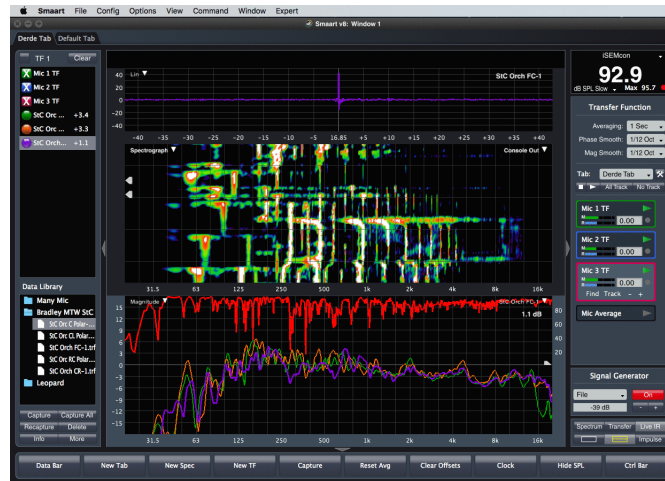
Pro měření kvality ozvučení existují v dnešní době specializované programy (ODEON, EASE, Smaart), jejichž ukázky jsou na obr. 7.2. Tyto programy nepočítají pouze STI index. Jsou velmi komplexní a obsáhlé. Z tohoto důvodu jsou také velmi náročné na výpočetní výkon. Tyto programy jsou schopny vypočítat a pomocí modelu zobrazit celkovou situaci místnosti, srozumitelnost v jejích jednotlivých místech, včetně započítání reproduktorů, mikrofonů a dalších prvků umístěných v místnosti, které ovlivňují ozvučení a srozumitelnost. Jsou schopny spočítat RT60 pro jednotlivé oktávy, odrazy, a další [31].



(a) ODEON software



(b) EASE software



(c) Smart v8 software

Obrázek 3.1: Ukázka programů pro práci s akustikou

### 3.2 Omezení do budoucna

V budoucnosti se budou omezení současných metod pořád zmenšovat. Již v dnešní době nemáme problém spočítat během několika vteřin STI nebo STIPA. Je ale pravděpodobné, že informací o měřeném prostoru budeme chtít - s rostoucím výpočetním výkonem a zmenšováním a vývojem přenosných zařízení - čím dál víc. Proto i výpočetní a časová náročnost bude větší než v současnosti.

## VÝHODNOST METOD PRO ROZDÍLNÉ APLIKACE

### 4.1 Aplikace měření kvality ozvučení a srozumitelnosti

Metody mohou být použity pro měření různých akustických prostředí, míst anebo elektronických systémů, kde chceme měřit srozumitelnost. Mají však i své limity a pro některé aplikace se hodí méně. Proto se uživatel, který potřebuje měřit srozumitelnost, musí rozhodnout, která metoda je pro něj ta nejvhodnější [14, 31].

- Měření veřejných míst a ozvučovacích systémů
- Měření a certifikace alarmů a bezpečnostních systémů
- Měření bezdrátové komunikace, telefonů, interkomů a vysílaček
- Měření přednáškových sálů, učeben a místností, kde se bude mluvit
- Měření míst, kde dochází k přímé komunikaci, jako jsou menší místnosti nebo automobily
- Měření systémů pro neslyšící

### 4.2 Jaké metody používáme

V současnosti se nejvíce používají STI (a STIPA, STITEL) a SII (bývalé AI). Každá z těchto metod funguje na jiném principu, a tudíž má každá své výhody a nevýhody (viz. 2.4 a 2.8.3). Dříve byla nevýhoda SII, že algoritmus neřešil problémy časových a nelineárních deformací [29]. Dnes ale metoda používá hodně prvků z metody STI a rozdíl ve výsledcích mezi těmito metodami je pro praktické použití minimální [25]. Samotná metoda STI se dnes používá minimálně - z důvodu dlouhého měření. STI se používá pouze v případě nepřímého měření pomocí IR. Pro většinu aplikací se používá metoda STIPA, která sice moduluje signál v jednotlivých oktávách pouze dvěma frekvencemi, ale výsledky jsou velmi podobné full-STI a jsou dostačující. Jediná aplikace, kde se metoda STIPA nepoužívá, jsou telekomunikace, vysílačky a bezdrátová komunikace. V těchto aplikacích se používá metoda STITEL. V minulosti se ještě často používala RASTI k posouzení kvality komunikačních systémů (míst, kde docházelo k přímé komunikaci mezi lidmi).

## ČÍM SE MĚŘÍ KVALITA OZVUČENÍ

### 5.1 Moderní přenosné měřicí přístroje

Při vypracování své práce jsem nepracoval s přístrojem, který by mi STI index rovnou spočítal, ale můj výsledek jsem získal pouze z mikrofonu (diktafonu). Proto jsem si vytvořil v Matlabu skript, který mi STI index spočítá z dat, které jsem naměřil.

Když ale probíhá měření při kolaudaci budovy, používají se obvykle profesionální přístroje, které slouží jen pro tyto účely. Přístroje přímo změří hodnotu STI (STIPA) a ukáží ji na displeji. Všechno ostatní zajistí hardware a software přístroje.

#### 5.1.1 2270 Hand Held Analyzer (Brüel & Kjær)

2270 je velmi univerzální, modulární zařízení s mnoha volitelnými dokoupitelnými moduly jako je FFT, nahrávání zvuku anebo pokročilé protokolování. Při koupi zařízení si můžeme nainstalovat mobilní aplikaci pro jednodušší a dálkové ovládání 2270, ukládání nahraných dat na cloud, kde můžeme nahraná data zpracovávat [13].

- Frekvenční rozsah 4,2 Hz - 22,4 kHz (pro mikrofon 4189)
- Dynamický rozsah 16,6 - 140 dB (pro mikrofon 4189)



- Splňuje podmínky IEC 60628-16 rev.4



Obrázek 5.1: Brüel & Kjær 2270

### 5.1.2 Norsonic Nor131

Další ze zástupců přístrojů, které měří srozumitelnost. Dokáže měřit jak STI, tak STIPA metodou a má mnoho dalších užitečných funkcí, např. výpočet RT60, šumu pozadí a další [22].



Obrázek 5.2: Norsonic Nor131

### 5.1.3 Bedrock STIPA meter SM50

Bedrock STIPA meter SM50 byl navržen lidmi, kteří vynalezli metodu STIPA a jsou zodpovědní za vývoj STI během posledních několika let. SM50 má barevný dotykový displej a vzhledem k intuitivnímu ovládání je vhodný i pro amatéry. Zároveň však obsahuje všechny funkce, které bude po přístroji tohoto typu požadovat zkušený odborník [10].

- Frekvenční rozsah 20 Hz - 20 kHz
- Dynamický rozsah 30 - 124 dB
- Splňuje podmínky IEC 60628-16 rev.4



Obrázek 5.3: Bedrock SM50

### 5.1.4 NTi XL2 analyzer

XL2 analyzer je přístroj od firmy NTi audio, která se na akustiku zaměřuje již více než patnáct let. Firma má dokonce pobočku v Praze. Zařízení má taktěž spoustu pokročilých funkcí. Zařízení je přizpůsobitelné tak, aby bylo jednoduše ovladatelné pro běžné uživatele [23].

- Frekvenční rozsah 6,3 Hz - 20 kHz
- Dynamický rozsah (nejnižší) 10 - 110 dB nebo nejvyšší 50 - 150 dB



Obrázek 5.4: NTi XL2

## TESTOVACÍ DATABÁZE AKUSTICKÝCH VZORKŮ

### 6.1 Tvorba databáze

Tvorba probíhala v průběhu jednoho týdne na rozličných vybraných místech. Tvořil jsem ji v programu Microsoft Access. Při zpracování a vyhodnocení dat pro mě databáze bude klíčovým prvkem při porovnávání hodnot mých dvou algoritmů. Pro vytvoření databáze potřebujeme znát objekty, které jsou podstatné pro akustiku a představu o měřeném prostoru. Pro databázi nahrávek prostorů považuji za důležité tyto objekty:

- Číslo nahrávky
- Místo nahrávky
- Datum
- Rozměry místnosti
- Nahrané signály
- Účel prostoru
- Index STI
- Index  $STI_{FV}$

Databáze je s názvem „Testovaci\_databaze\_STI“ uložena v příloze na CD.

### 6.1.1 Místa, na kterých probíhalo nahrávání

Místa, na kterých měření probíhalo, jsem vybíral tak, aby byla co nejvíce rozdílná a zabírala velkou škálu prostorů, ve kterých se běžně pohybujeme, a tudíž potřebujeme, aby zde byla dobrá srozumitelnost. Nahrávání probíhalo za běžných podmínek, takže nahrávky ve většině případů ovlivňuje šum na pozadí.

- Obchod s potravinami
- Chodba bytového komplexu
- Obývací pokoj
- Ložnice
- Kuchyně
- Kavárna
- Kostel
- Veřejná chodba
- Bar s kulečnickovým stolem
- Nádraží
- Přednášková místnost

Vzorky jsem nahrával za běžných podmínek během dne. Databáze obsahuje jedenáct vzorků a má osm polí (6.1).

## 6.2 Využití databáze mimo rozsah bakalářské práce

Mnou vytvořená databáze nahrávek a STI indexu by mohla sloužit studentům, kteří se o toto téma budou zajímat a budou potřebovat uvedené podklady použít pro tvorbu své bakalářské či diplomové práce.

Databáze by také mohla být dostupná na internetu pro širokou veřejnost, kde by si každý mohl najít informace o akustice dané budovy. Tyto informace mohou být užitečné pro konání různých akcí. Například pro konání konference budou organizátoři chtít prostor, kde je dobrá akustika a výborná srozumitelnost posluchače.

Naopak, pro pořádání koncertu je dobré, pokud je zde srozumitelnost horší kvůli dlouhému dozvuku, protože hudba se pak lépe poslouchá a je krásnější.

## TESTOVACÍ ALGORITMUS NA ZÁKLADĚ STI

### 7.1 Postup vytváření algoritmu

Při psaní algoritmu jsem vycházel ze znalostí algoritmu STI. Pokud měřím STI nepřímo (mám nahrávku s výsledkem měření, kterou zpracovávám až později v počítači), potřebuji impulzní odezvu systému. Tu získám ze svých změřených nahrávek signálu MLS pro každou místnost. Předpokládám, že by výsledky indexu STI a mého algoritmu  $STI_{FV}$  pro jednu místnost měly dopadnout velmi podobně. Impulzní odezvu systému použiji jako vstupní parametr funkce, která bude počítat index STI. Dále potřebuji pro výpočet indexu vzorkovací frekvenci a třetino-oktávová pásma, ve kterých budu  $STI_{FV}$  počítat. Metodu neupravuji z důvodu ušetření času nebo výpočetního výkonu, ale z důvodu zkoumání změny indexu při omezení rozsahu oktávových pásem a modulačních frekvencí. Současně se snažím při měření docílit co nejpřesnější napodobení reálnému hovorů.

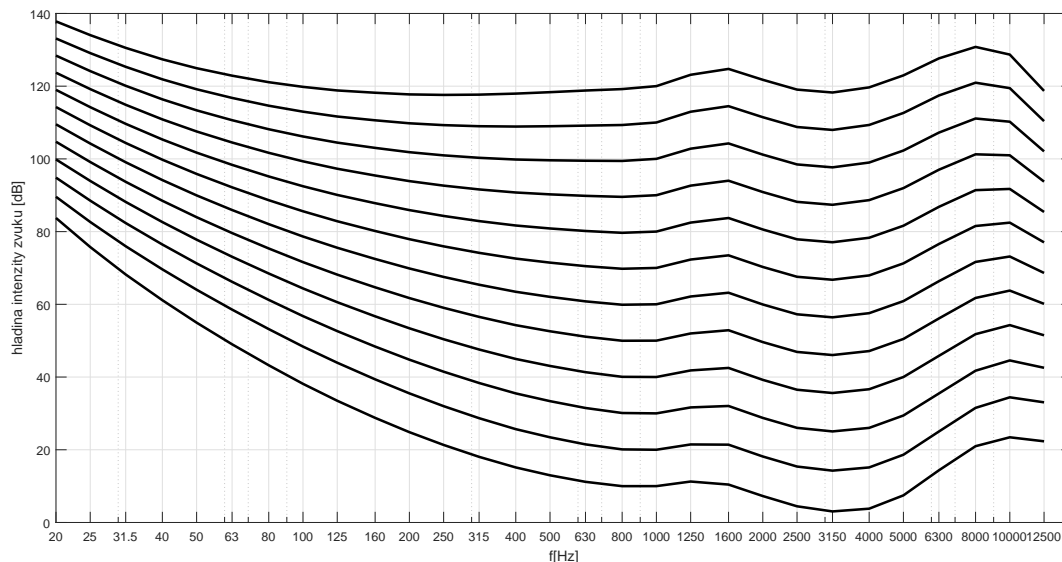
#### 7.1.1 STI algoritmus

V rámci práce jsem použil výsledky měření, jehož vzorky jsou uloženy v databázi (6.1). Naměřené výsledky byly zpracovávány STI algoritmem a upraveným algoritmem na základě STI ( $STI_{FV}$ ). Nejdříve jsem v Matlabu vytvořil skript, který počítá STI z impulzní odezvy systému. Zpracování nahrávek probíhalo až později v počítači. Proto jsem použil nepřímou metodu měření a použil jsem vzorec (2.12), ze kterého jsem počítal index STI. Stejně jsem postupoval i při vytváření mnou upraveného algoritmu. Pro získání impulzní odezvy, jsem musel na měřeném místě pustit z reproduktoru MLS signál, který se nahrával na diktafon. Pro získání impulzní odezvy se může použít i jiný signál než MLS (viz. 2.4.2). Z IR získáme spoustu informací o místnosti, ve které jsem nahrával: dobu dozvuku  $RT60$ , odrazy, ozvěny nebo zkreslení. Impulzní odezvu změřeného místa jsem použil v algoritmu STI a následně algoritmem  $STI_{FV}$  a porovnával jsem jejich výsledky a časovou a výpočetní náročnost.

#### 7.1.2 $STI_{FV}$ algoritmus

Nyní se dostávám k tomu, jak vypadalo vytváření algoritmu  $STI_{FV}$ . Frekvenční rozsah lidského hlasu se pohybuje dle [21] 85 Hz - 1280 Hz a frekvenční rozsah lidského sluchu je 16 Hz - 20 kHz [16]. Z této informace víme, že lidský hlas je jen nepatrná část slyšitelného spektra lidského

hlasu. Podle křivky hladin stejné hlasitosti je lidské ucho nejcitlivější na frekvence okolo 3 - 4 kHz a v případě, že frekvenční rozsah omezíme jen na rozsah lidského hlasu, je ucho nejcitlivější na frekvence 700 - 1000 Hz [24]. Tyto dva rozsahy frekvencí beru jako nejdůležitější pro měření srozumitelnosti v prostoru. Proto ve svém algoritmu budu používat k měření celé frekvenční pásmo, kde mnou vybrané nejdůležitější pásma budou počítat s více modulačními frekvencemi. Algoritmus STI měří hodnoty pro sedm oktávových frekvenčních pásem a každé je modulováno čtrnácti frekvencemi.



Obrázek 7.1: Křivky hladin stejné hlasitosti

Křivky na obrázku 7.1 jsou nazývány hladinami stejné hlasitosti (izofony). Fon je jednotkou hlasitosti a hladiny hlasitosti zkoumaného zvuku ve fonech odpovídající číselně hladinám intenzity zvuku tónů u 1kHz v dB. Křivky stejných hladin hlasitosti udávají tyto hladiny ve fonech<sup>1</sup>. [16]

Při vytváření algoritmu se snažím o takovou jeho úpravu, abych lépe poznal, jak algoritmy měření srozumitelnosti fungují a abych se co nejvíce přiblížil reálné situaci při komunikaci mezi lidmi v konkrétním prostředí. Pro zjištění chování měření při úpravě algoritmu a při snaze o zachování co nejpřesnějších výsledků, jsem pro čtyři – z mého pohledu nejdůležitější - třetino-oktávová pásma se střední frekvencí 800 Hz, 1000 Hz, 3150 Hz a 4 kHz použil všech čtrnáct modulačních frekvencí, kterými je budu modulovat. Rozpis ostatních třetino-oktávových pásem spolu s jednotlivými modulačními frekvencemi jsou rozepsány v tabulce 7.1. Jsou to pásma, ve kterých je lidské ucho nejcitlivější jak na celém frekvenčním rozsahu, tak na zúženém rozsahu (pouze na spektrum lidského hlasu). Ostatní třetino-oktávová pásma z frekvenčního rozsahu jsou modulována pouze dvěma frekvencemi. Při takové úpravě bude algoritmus pravděpodobně (výpočetně i časově) podobně náročný jako algoritmus STI, protože celé frekvenční pásmo dělím

<sup>1</sup>Hladina hlasitosti je definována tak, že hladina hlasitosti 1 fón je při frekvenci 1kHz stejně velká jako jednotka hladiny zvuku 1dB.

1/3 oktávová pásma (Hz)	100	125	160	200	250	315	400	500
1. modulační frekvence (Hz)	1,60	1,00	0,80	1,25	1,00	1,60	0,80	0,63
2. modulační frekvence (Hz)	6,30	5,00	4,00	5,00	5,00	4,00	6,30	3,15

630	1250	1600	2000	2500	5000	6300	8000	10000
1,60	2,00	1,00	1,25	0,63	0,80	1,60	2,50	2,00
8,00	10,00	5,00	6,25	3,15	4,00	8,00	12,50	10,00

Tabulka 7.1: Třetinooktávová pásma a jejich modulace

na třetinooktávy a ne pouze na oktávy. Celkově tedy budu počítat s devadesáti hodnotami (STI počítá s 98 hodnotami).

### 7.1.3 Implementace kódu do Matlabu

STI je počítán z MTF, která se získá z IR testovacího signálu. MTF je definována jako rozsah Fourierovy transformace  $IR^2$ , která je vydělena celkovou energií IR. Algoritmus programuji podle normy [1] a také podle [9]. Nejprve jsem vytvořil kód, který počítá STI a z něj jsem vytvořil program, který počítá podle upravené metody  $STI_{FV}$ . Při vytváření kódu postupuji podle 2.4.2, kde je vysvětlený celý postup výpočtu v případě nepřímé metody a výpočtu STI pomocí IR. Některé kroky jsem v Matlabu musel upravit tak, aby byly výpočetně jednodušší nebo aby se mi lépe do programu zadávaly.

#### Výpočet MTF

Podle vzorce 2.12 se výpočet MTF skládá ze dvou částí. Z části, kde dochází k odrazům, a z části SNR. Nejdříve budu počítat první část, kde je potřeba vyfiltrovat si impulzní odezvy na každou oktávu (třetinooktávu) zvlášť. MTF budu počítat pro impulzní odezvu každého oktávového pásma zvlášť. Poté budu pokračovat následovně.

- Umocnění impulzní odezvy ( $IR^2$ )
- Integrace  $IR^2$  abychom dostali celkovou energii
- Výpočet Fourierovy transformace  $IR^2$
- Normalizace obálky transformovaného spektra
- Vezeme druhé mocniny součtu reálné části na druhou a imaginární části na druhou (absolutní hodnota)
- Se spočítaným MTF pro každou oktávu použijeme 14 modulačních frekvencí.

#### Výpočet SNR

SNR vypočítáme podle vzorce 2.5. Popis postupu při vytváření algoritmu je uveden v odrážkách.

- Převeďme každou z 98 hodnot do SNR v dB
- Omezím rozsah na  $\pm 15$ dB



- Vypočtu střední hodnotu SNR pro každé pásmo
- Počítám s váhou hodnoty každé oktávy

## Výpočet STI

- Převodu SNR na hodnotu TI podle vzorce 2.8, následně na hodnotu MTI a nakonec na námi chtěnou hodnotu STI podle 2.10.

Zde uvádím, jak vypadají mé dvě funkce pro výpočet MTF a TI ( $TI_{FV}$ ). Celý kód pro výpočet STI a  $STI_{FV}$  je v příloze.

```
function [STIfm, STIfw] = STIfv(IR, Fs)
```

Před samotným výpočtem indexu STI ( $STI_{FV}$ ) musím vypočítat MTF (podle vzorce 2.12). Funkce pro výpočet MTF vypadá takto:

```
function MTF = getMTF(ir)
```

Pro výpočet MTF budu potřebovat IR umocněnou na druhou. Poté potřebuji integraci IR, což je vlastně sumace hodnot  $IR^2$ .

## 7.2 Použitelnost algoritmu v praxi

Při používání algoritmu na vzorcích, které jsem naměřil v různých místnostech (6.1), se algoritmus zdá podle předpokladů funkční. Při zkoušení algoritmu na původním signálu MLS je výsledek  $STI=1$ , což byla zkouška pro to, jestli algoritmus funguje správně. Když jsem algoritmus použil na vzorky jednotlivých místností, dostával jsem také rozumné údaje. Bohužel jejich správnost si jiným způsobem ověřit nemohu. Podle výsledků dává algoritmus STI podobné výsledky jako algoritmus  $STI_{FV}$ . Program je vytvořen jako spustitelný soubor v rámci Matlabu. Má své GUI a proto je jednoduše ovladatelný i pro uživatele, kteří s tímto prostředím nemají předchozí zkušenosti. Prostředí je velmi jednoduché a intuitivní. Je zde tlačítko pro načtení IR, výpočet STI a exportu hodnot do textového editoru. Pro uživatele, které zajímá jak algoritmus funguje, je samozřejmě přístupný skript Matlabu.<sup>2</sup> Algoritmus měří revidovaný index  $STI_r$  pro obě pohlaví a dostáváme tak lepší a pravdivější informace o srozumitelnosti. Pokud správně naměříme IR v konkrétní místnosti, pro kterou chceme spočítat STI, a poté pro výpočet algoritmu použijeme korektní IR, algoritmus nám spočítá správný index STI podle definice a standardu IEC 60628-16.

<sup>2</sup>Holý program, který se musí ovládat prostřednictvím příkazového okna Matlabu. Pro ovládání a počítání je potřeba pochopit alespoň základy fungování algoritmu.

## 7.3 Porovnání $STI_{FV}$ s STI

Programování algoritmu  $STI_{FV}$  bylo složitější než naprogramování samotného STI, i když jsem  $STI_{FV}$  vytvářel až z algoritmu STI. V programování  $STI_{FV}$  jsem musel oddělit třetinooktávová pásma do dvou vektorů, se kterými jsem počítal zvlášť, a to z důvodu rozdílného počtu modulací těchto vektorů. Také z důvodu dělení na třetinooktávová pásma byla funkce v některých případech složitější. Zde je ukázka porovnání části výpočtu pro  $STI_{FV}$  a pro STI:

```

% Výpočet pro 4 pásma (F0)
for i=1:Nfc0
    if nargin < 3
        f.F0 = F0(i);
        H = design(f,'butter');
        ir_filtered = filter(H, IR);
    else
        ir_filtered = filter(ThirdOctFilters(i), IR);
    end
    MTF_octband0(:,i) = getMTF(ir_filtered); % MTF pro jednotlivá pásma
end
% Výpočet pro zbytek pásem (F1)
for i=1:Nfc1
    if nargin < 3
        f.F0 = F1(i);
        H = design(f,'butter');
        ir_filtered = filter(H, IR);
    else
        ir_filtered = filter(ThirdOctFilters(i), IR);
    end
    MTF_octband1(:,i) = getMTF(ir_filtered); % MTF pro jednotlivá pásma
end

```

```

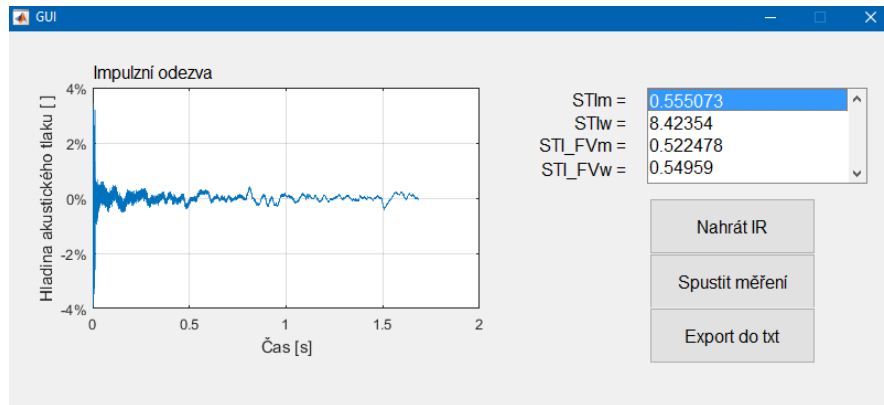
for i=1:Nfc
    if nargin < 3
        f.F0 = F0(i);
        H = design(f,'butter');
        ir_filtered = filter(H, ImpulseResponse);
    else
        ir_filtered = filter(OctaveFilters(i), ImpulseResponse);
    end
    MTF_octband(:,i) = getMTF(ir_filtered);
end

```

Protože je frekvenční rozsah rozdělen na třetinooktávová pásma místo oktávových, musel jsem si upravit váhové a korelační koeficienty  $\alpha$  a  $\beta$  pro každé třetinooktávové pásmo. V podstatě jsem stejné koeficienty použil vždy pro tři třetinooktávová pásma. Úprava musela proběhnout také ve výběru při použití pouze dvou modulačních frekvencí. Pro nosné frekvence z oktáv jsem použil stejné frekvence jako při metodě STIPA (2.6) a doplnil jsem je modulačními frekvencemi podle citu.

### 7.3.1 GUI

Výsledek mého programu jsem implementoval pro lepší obsluhu jako GUI. Toto uživatelské rozhraní je vytvořeno pomocí Matlab Guide. GUI obsahuje pouze jednoduché ovládací prvky. Je zde graf zobrazující průběh IR, tlačítka pro načtení IR, výpočet  $STI$  a pro export výsledků  $STI$  do textového souboru. Nakonec je zde textové pole pro zobrazení vypočtených hodnot  $STI$ . Jak GUI vypadá se můžeme podívat na obr. 7.2.



Obrázek 7.2: Ukázka programu po výpočtu

### 7.3.2 Výhody

Při pracování s algoritmem a jeho porozumění zjistíme, že se nesnažím o jeho zjednodušení, ale přiblížení se realitě v podobě úpravy důležitosti jednotlivých třetinooktávních frekvencí. Rozdělení frekvenčního rozsahu na třetinooktávních pásma nám dává více možností algoritmus upravovat a vylepšovat. Tímto směrem se ubírá metoda SII, která má několik variant výpočtu pro různý rozsah frekvenčních pásem 2.8.3.

### 7.3.3 Nevýhody

Nevýhodou algoritmu  $STI_{FV}$  je jeho rozdělení na třetinooktávních pásma, což je v jednom případě, jak popisují výše výhoda, ale také to prodlužuje dobu výpočtu a zvětšuje počet měřících bodů. Rozdílem proti algoritmu  $STI$  jsou jeho odlišné výsledky. Zatím jsem nezjistil, jestli je algoritmus  $STI_{FV}$  pro některá místa měření lepší než algoritmus  $STI$ , nebo jestli jsou místa, pro která se algoritmus nehodí. Takové informace mohou získat až po dlouhodobějším měření na více místech a po porovnávání algoritmu s jinými.

Měření srozumitelnosti STI, RASTI a STIPA se obvykle provádí na specializovaných přístrojích a na hardwaru vyrobeném přímo pro toto konkrétní použití. Hardwarem pro měření kvality ozvučení a srozumitelnosti se zabývám v 5.1. Při měření pro účely této práce nebyly použity žádné specializované přístroje pro měření STI.

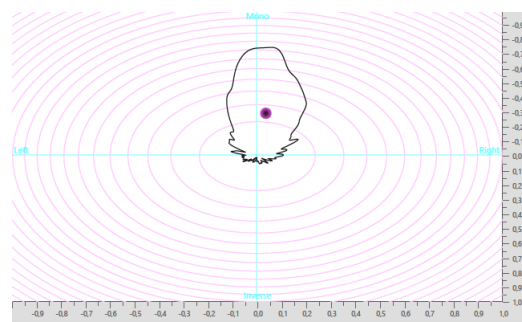
Zároveň je ale pro výsledky měření velmi důležité, čím měření provádíme. Správné nebo špatné přístroje mohou ovlivnit výslednou hodnotu STI. Proto musíme dbát na velikost reproduktoru, jeho směrovost a další vlastnosti, které jsou uvedeny v 2.4.2.

Pro měření byly použity:

- Jednoduchý reproduktor s průměrem satelitu 10 cm, poměrně plochou charakteristikou na frekvenčním pásmu od 250 Hz po cca 11 kHz a směrovost podobnou lidské řeči při komunikaci dvou osob. Parametry použitého reproduktoru jsou zobrazeny na obr. 8.1.
- Diktafon Zoom H4 Handy Recorder.



(a) Frekvenční charakteristika



(b) Směrovost

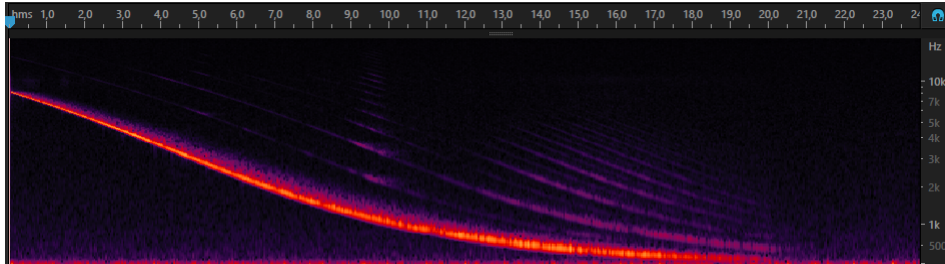
Obrázek 8.1: Parametry reproduktoru

## 8.1 Přímé a nepřímé měření

Při použití specializovaných přístrojů se měření a vyhodnocení většinou provádí přímo v místě měření. Já, jelikož používám obvyčné přístroje, udělám nahrávky na místě a zpracování proběhne až později v počítači (viz. 2.4.2). Jeho výhodou je, že přesnost vzorkovací frekvence

mezi signálem zdroje a přijímače je méně chybová (méně než  $0,5 \cdot 10^{-6}$ ) než z přímého měření ( $20 \cdot 10^{-6}$ ) [1].

Měření probíhalo na několika různých místech a za různých podmínek. Nejdříve jsem reproduktorem, který představoval ústa, vysílal 20 sekund MLS signál a 20 sekund sweep signál ve výšce 1,5 m nad povrchem podlahy; vzdálenost reproduktoru a mikrofonu byla dva metry. Při zpracování naměřených dat jsem používal pouze testovací signál MLS z důvodu lepších výsledků 2.4.2. Při použití sweep signálu je ale na spektrogramu (viz. obr. 8.2) krásně vidět jednotlivé odrazy.



Obrázek 8.2: Spektrogram testovacího signálu sweep v kostele

## 8.2 Kde a jak probíhalo nahrávání

Měl jsem vytipovaná místa, o kterých jsem věděl, že jejich akustika je ideální pro potřeby mé práce. Prostor se špatnou akustikou (dlouhou dobou dozvuku a velkými ozvěnami), prostor, který je určen k přednášení, klasická domácí místnost, malý prostor, velká hala. Nahrávky jsem pořizoval celkem na jedenácti místech za běžného provozu. Ukázkou z nahrávání ukazují na obr. 8.3. Poté, co jsem měl místa vybraná a nahrál jsem potřebné impulzní odezvy, přišlo na řadu zpracovávání dat v počítači.

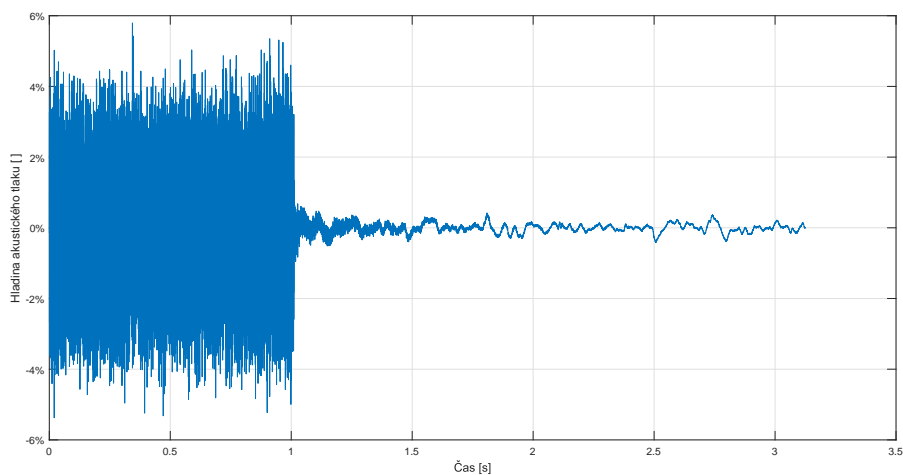


Obrázek 8.3: Nahrávání srozumitelnosti v kostele

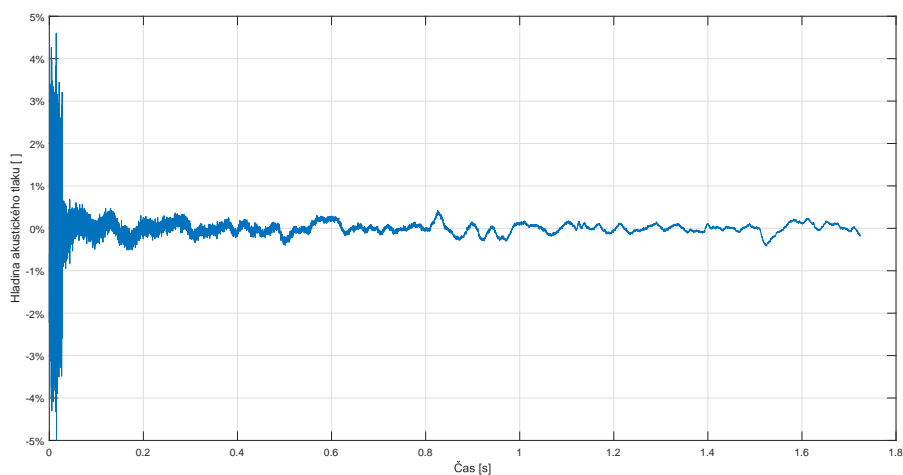
## 8.3 Zpracování dat

Při zpracování nahrávek je důležité mít co nejpřesněji zpracovanou impulzní odezvu systému. Program v Matlabu funguje tak, že zpracuje všechny čísla z IR, ze které počítá STI. Pokud je

před impulzní odezvou dlouhý průběh signálu MLS nebo je po vypnutí signálu a jeho odeznění velký vektor šumu pozadí, velice to ovlivňuje výpočet STI. Proto musím přesně vědět, jak bude průběh impulzní odezvy dlouhý a také kde bude začínat a kde končit. To mohu dělat metodou pokus-omyl a zkoušet, kdy dosáhnu nejlepších výsledků anebo sledovat, kdy odezní signál, který jsem vysílal. Na obrázku 8.4<sup>1</sup> vidíme, že STI algoritmus zpracovává jednu vteřinu průběh MLS signálu. Tato vteřina nám ale velmi ovlivňuje konečný výsledek hodnoty STI. Na obrázku 8.5 vidíme správně zpracovanou IR. Výsledek STI pro špatně zpracovanou STI je  $STI = 0,1570$  a pro správně zpracovanou IR  $STI = 0,5881$ . Já jsem všechny naměřené testovací signály ořezal na IR sledováním jeho průběhu.



Obrázek 8.4: Špatně zpracovaná impulzní odezva

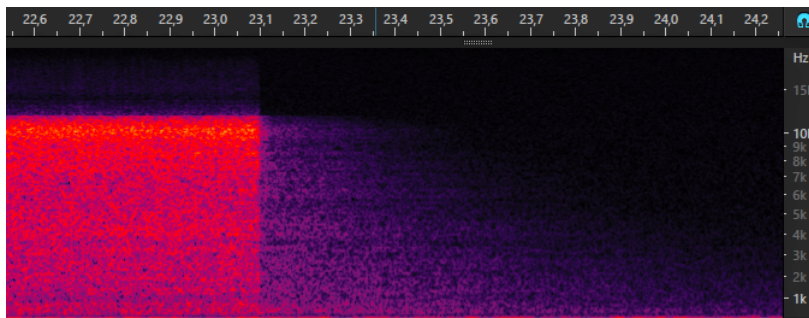


Obrázek 8.5: Správně zpracovaná impulzní odezva

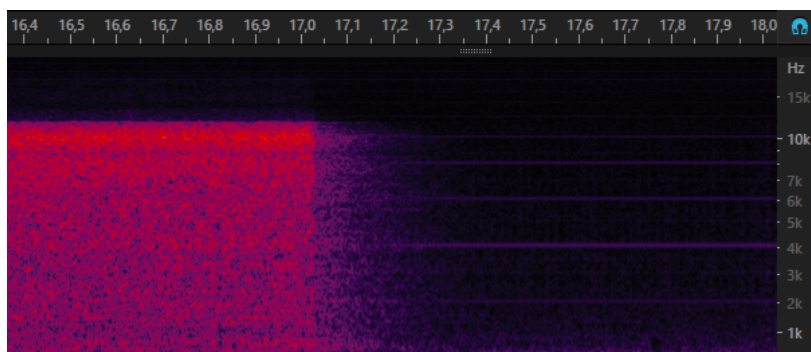
Při IR má každá frekvence jinou dobu dozvuku. Nejpomalejší jsou nízké frekvence, ale může

<sup>1</sup>Hladina akustického tlaku je v grafu uvedena v jednotkách procent, kde 0% je  $-\infty$  ... a 100% (-100%) je 0 dB

se stát, že v konkrétní místnosti bude rezonovat z nějakého důvodu frekvence třeba někde uprostřed spektra. V případě naměřených nahrávek ze spektrogramu IR pro kostel na obrázku 8.6 vidíme, že v čase  $t = 23,1$  s přestal být vysílán signál a ještě v po jedné sekundě můžeme pozorovat ozvěny a odrazy, které nejdéle zůstávají na nízkých frekvencích. Impulzní odezvu MLS signálu na obrázku 8.6 můžeme porovnat s IR v obývacím pokoji (8.7). Proto je potřeba aby IR trvala minimálně 1,6 s (viz. 2.4.2). Při místnostech s kratší dobou dozvuku můžeme IR zkrátit, aby nám šum pozadí neovlivňoval výsledek STI.

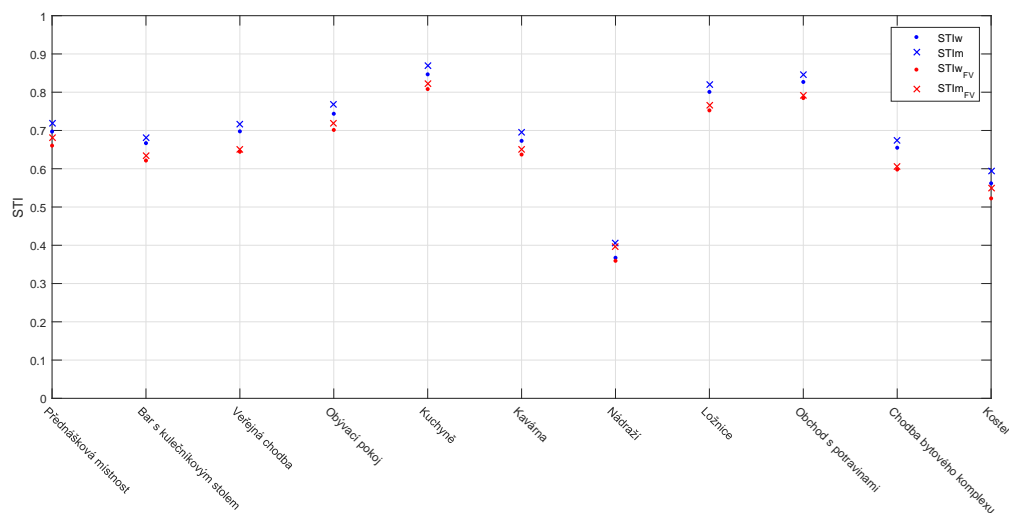
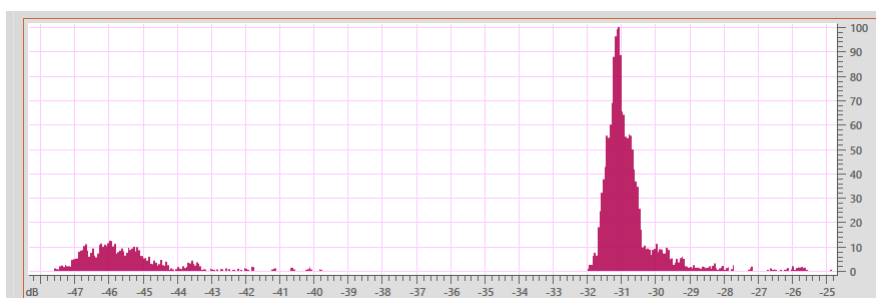


Obrázek 8.6: Kostel: prostor s dlouhou dobou dozvuku



Obrázek 8.7: Obývací pokoj: prostor s krátkou dobou dozvuku

Výpočet algoritmu STI a  $STI_{FV}$  jsem provedl pro všechny prostory, kde jsem měřil IR. Výsledky výpočtů jsou znázorněny na obr. 8.8 a vypsány v tab. 8.1. Při měření na nádraží pozorujeme v grafu menší hodnoty akustického výkonu. To je způsobeno špatnou akustikou tohoto místa, i subjektivně hodnotím srozumitelnost jako špatnou. Výpočet mohl ovlivnit šum pozadí, který byl při normálním provozu velmi vysoký a SNR bylo asi 15 dB (8.9). Všimněte si také, že v případě hodnot pro nádraží, které vypočítal algoritmus pro měření STI a pro měření  $STI_{FV}$ , se vypočítané hodnoty téměř shodují. Ostatní vypočítané hodnoty se od sebe navzájem trochu liší, ale tato odlišnost je způsobena odlišností obou algoritmů; s rozdíly v měření a jejich výsledcích jsme počítali.

Obrázek 8.8: Hodnoty STI a STI<sub>FV</sub>

Obrázek 8.9: SNR Nádraží

V tabulce 8.1 vidíme jednotlivé hodnoty naměřené pomocí mého algoritmu pro výpočet STI a STI<sub>FV</sub> a jejich výsledek na STI stupnici. Ve většině soukromých i veřejných prostranstvích, kde je srozumitelnost důležitá, je výsledek STI dobrý nebo výborný. Pouze na nádraží a v kostele nebyly výsledky pro srozumitelnost dobré. Horší výsledky na těchto dvou místech se daly předpokládat, a to vzhledem k velkému prostoru a vzdáleným stěnám, díky kterým je zde dlouhá doba dozvuku a odražené zvuky doznívají ještě při další informaci, která je tím méně srozumitelná. Pro veřejný prostor (jako např. nádraží) je ale srozumitelnost velmi důležitá. Lidé potřebují vědět odkud a v kolik jim jede vlak a tato informace je pomocí rozhlasu mnohdy velmi špatně srozumitelná. Do takových míst by se měly instalovat akustické podhledy a panely. V dnešní době by se v takových místech mohlo experimentovat s aktivním potlačením hluku.

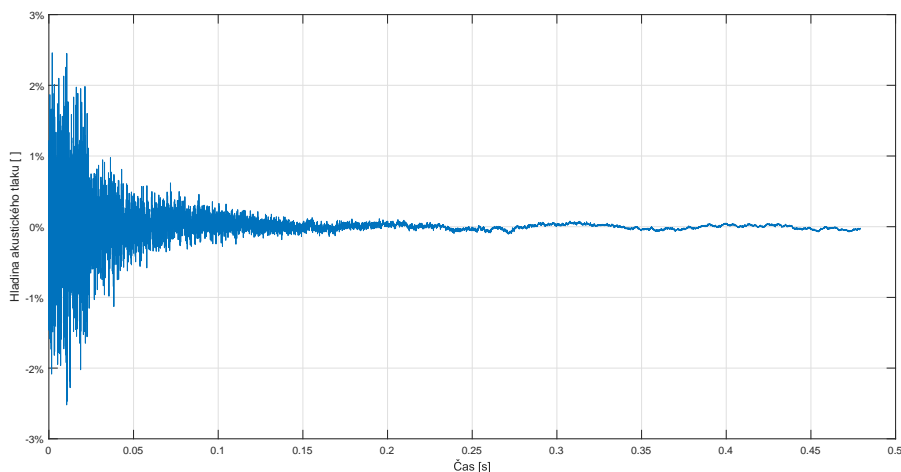


	STIm	STIw	STIfm	STIfw	STI
Přednášková místnost	0.6972	0.7194	0.6604	0.6816	Dobry
Bar s kulečníkovým stolem	0.6669	0.6807	0.6214	0.6338	Dobry
Veřejná chodba	0.6977	0.7168	0.6444	0.6511	Dobry
Obývací pokoj	0.7437	0.7683	0.7014	0.7187	Dobry
Kuchyně	0.8467	0.8704	0.8082	0.8222	Výborný
Kavárna	0.6729	0.6954	0.6371	0.6497	Dobry
Nádraží	0.3674	0.4055	0.3594	0.3955	Špatný
Ložnice	0.8008	0.8208	0.7525	0.7653	Výborný
Obchod s potravinami	0.8268	0.8459	0.7847	0.7923	Výborný
Chodba bytového komplexu	0.6549	0.6741	0.5979	0.6050	Dobry
Kostel	0.5619	0.5934	0.5225	0.5496	Ucházející

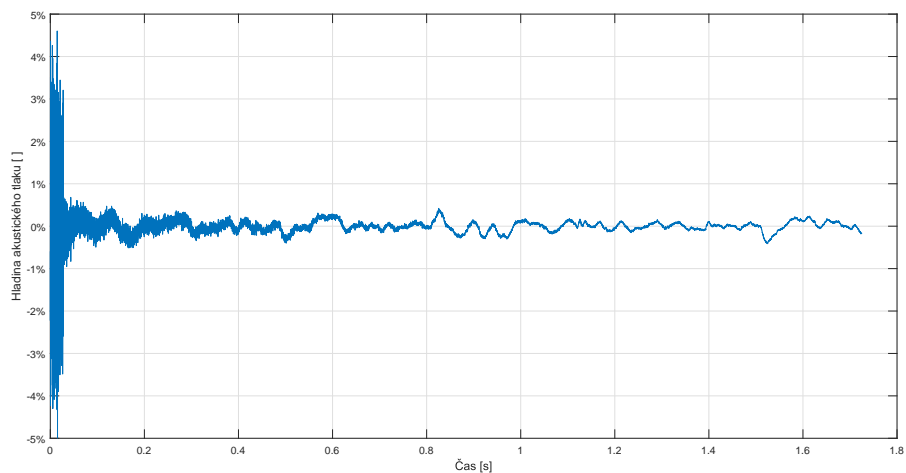
Tabulka 8.1: Tabulka naměřených hodnot STI a  $STI_{FV}$ 

### 8.3.1 Srovnání signálu z krátkou a dlouhou IR

Měření v různých prostorech nám přináší některé problémy, které musíme při následném zpracování signálu v počítači řešit. IR není ve všech místnostech stejná a naopak každá místnost má specifickou IR. Díky tomu můžeme IR používat pro měření kvality srozumitelnosti. V tomto případě nás nejvíce zajímá délka IR. Při používání tohoto algoritmu musíme brát ohled na délku IR. Pokud v IR necháme dlouho jen šum pozadí, ovlivní nám to výsledek, který nebude přesný. IR proto musíme přizpůsobovat konkrétnímu signálu. V případě prostoru s dlouhou dobou dozvuku (viz. obr. 8.11) je vidět, že signál odeznívá minimálně 1,2 s. V případě krátké doby dozvuku (obr. 8.10) nám stačí IR dlouhá 0,4 s.



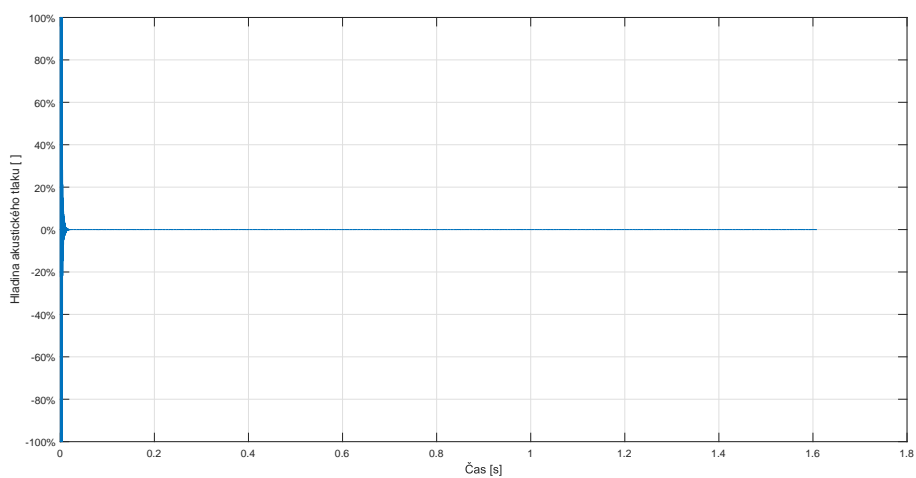
Obrázek 8.10: Signál v místnosti s krátkou dobou dozvuku



Obrázek 8.11: Signál v místnosti s dlouhou dobou dozvuku

### 8.3.2 Zkouška algoritmu

Protože jsem algoritmus STI a  $STI_{FV}$  psal sám, provedl jsem nejdříve zkoušku správné funkčnosti algoritmu. Funkčnost jsem ověřoval tak, že jsem na oba algoritmy použil původní, vygenerovaný signál MLS, který by mi měl dát výsledek STI a  $STI_{FV}$  roven jedné. Zkouška vyšla pro oba algoritmy podle předpokladů rovna jedné.



Obrázek 8.12: Impulzní odezva původního MLS signálu

## ZÁVĚR

Výsledkem mé práce je algoritmus, který funguje na principu metody STI, ale snažil jsem se ho upravit vzhledem k tomu, jak funguje lidský hlas a sluch. Rozdělení na třetinooktávová pásma a kladení důrazu jen na některé frekvence nebyla špatná úvaha, protože výsledky algoritmu jsou velmi dobré (podobné samotnému algoritmu STI). Při vytváření algoritmu jsem bral v potaz všechny současné standardy a myslel na nejnovější trendy, které se v oblasti akustiky vyskytují. Mým původním cílem bylo vytvořit zjednodušený algoritmus, ale tímto směrem se měření kvality ozvučení a srozumitelnosti nevydává. Dnes je cílem co nejvěrnější napodobení lidské komunikace s cílem získat srozumitelnost na místě měření. Přitom chceme, aby měření bylo komfortní, rychlé a levné. Můj algoritmus jistě není dokonalý, ale cílem ráce nebylo vymyslet dokonalý algoritmus. Chtěl jsem pochopit základy problematiky akustiky a měření kvality ozvučení a seznámit s touto problematikou případné čtenáře této práce.

Na mé bakalářské práci bych rád pokračoval a zaměřil se na současný vývoj metod měření jako je třeba binaural STI a další moderní metody, které jsou teprve ve vývoji.

Z hlediska současného trendu mobilních zařízení se budou programy a přístroje pravděpodobně rozvíjet směrem k virtuální realitě a mobilním zařízením. Brýle virtuální reality by nám mohly testovací signály vysílat reproduktorem a zároveň je přijímat mikrofony, které budou umístěny na místě lidských uší. Při chůzi po místnosti nám bude virtuální realita ukazovat spoustu informací o akustice v místnosti na konkrétním místě. Mohli bychom si zasimulovat umístění řečníka a podívat se, jaká je při konkrétní vzdálenosti srozumitelnost. Výhledy do budoucna jsou zajímavé a uvidíme, jakým směrem se budou technologie ubírat.



## LITERATURA

- [1] Sound system equipment bs en 60268-16:2011, 2011.
- [2] Rational Acoustics. *Smaart 7 impulse response measurement and analysis guide*. *Smaart*, 2015.
- [3] Jerry L. Punch PhD Aryn M. Amlani, MS and PhD Teresa Y. C. Ching. Methods and applications of the audibility index in hearing aid selection and fitting. *Something*, 99:250–257, 2002 Sep.
- [4] Ph.D. Arthur Boothroyd. Room acoustics and speech perception. *Room Acoustics*, 2002.
- [5] NTi audio. Introducing speech intelligibility. 2008.
- [6] NTi audio. Myths & facts about speech intelligibility. 2009.
- [7] NTi audio. Measuring speech intelligibility using dirac type 7841, 2013.
- [8] United Kingdom Civil Aviation Authority. Specification no. 15, public address systems. 2012.
- [9] Farrel Becker. Guide to computing the speech transmission index. *Synergetic Audio Concepts*, March 2010.
- [10] Bedrock Audio. *Bedrock SM50 STIPA meter*.
- [11] Ole-Herman Bjør. Measure speech intelligibility with a sound level meter. *Sound and Vibration*, pages 14 – 17, October 2004.
- [12] Jens Blauert, editor. *The Technology of Binaural Listening*. Springer Science & Business Media, 2013.
- [13] Brüel & Kjær. *Hand-held Analyzer Types 2250 and 2270*.
- [14] Michael Vorländer David Havelock, Sonoko Kuwano. *Handbook of Signal Processing in Acoustics*. Springer Science & Business Media, 2008.
- [15] Jan A. Verhave Dr. Sander J. van Wijngaarden. Speech intelligibility measurements in practice; obtaining accurate and reliable data using stipa tools. November 2014.
- [16] Jan Hálek. *Biofyzika pro bakaláře*. Univerzita Palackého v Olomouci, 1999.

- [17] Benjamin W.Y Hornsby. The speech intelligibility index: What is it and what's it good for? *The Hearing Journal*, 57:p. 10 – 17, October 2004.
- [18] Tammo Houtgast and Herman Steeneken. Basics of the sti measuring method.
- [19] Lloyd A. Jeffress. Binaural phase difference and pitch variation. *The American Journal of Psychology*, 61(4):468–486, 1948.
- [20] Akil Lau and Deon Rowe. Comparative analysis of on-site stpa measurements with ease predicted sti results for a sound system in a railway station concourse. *Building Sciences, Acoustics 2015 Hunter Valley*:3, November 2015.
- [21] Alan Lofft. Frequency ranges of male female and children's voices. *Axiom Home Theaters*, July 2009.
- [22] Norsonic. *Norsonic Nor 13x*.
- [23] NTi audio. *XL2 Handheld Audio and Acoustic Analyzer*.
- [24] Juhani Parmanen. Some reasons to revise the international standard iso 226:2003: Acoustics—normal equal-loudness-level contours. *Open Journal of Acoustics*, December 2012.
- [25] Jian Kang Peisheng Zhu, Fangshuo Mo. Experimental comparison between sti and sii metrics for the objective rating of speech intelligibility. 2014.
- [26] Jan Verhave Sander van Wijngaarden and Herman Steeneken. The speech transmission index after four decades of development. *Embedded Acoustics*, 40:134 – 138, August 2012.
- [27] JA Verhave SJ van Wijngaarden. Recent advances in sti measuring techniques. 28, 2006.
- [28] Meyer Sound. Machine measures of speech intelligibility. 2010.
- [29] Herman J.M. Steeneken. The measurement of speech intelligibility. 1985.
- [30] H.J.M. Steeneken and T. Houtgast. Rasti: A tool for evaluating auditoria. 1985.
- [31] Herman Steeneken Tammo Houtgast. *Past, present and future of the Speech Transmission Index*. TNO Human Factors, 2002.

Tento seznam literatury byl vytvořen pomocí Bib<sub>T</sub>E<sub>X</sub>u s použitím stylu plain. Citace jsou automaticky seřazeny podle abecedy.

# **Přílohy**

---

Programy pro výpočet algoritmu STI a  $STI_{FV}$  jsou přiloženy na CD. Je zde přiložena také testovací databáze a dokument práce ve formátu pdf.