

Mezinárodní telekomunikační unie

ITU - R

Radiokomunikační
sektor ITU

Doporučení ITU-R BS.1770-2
(03/2011)

**Algoritmy pro měření hlasitosti zvukových
programů a skutečné špičkové úrovně
zvuku**

série BS
Vysílací služby (zvuk)

ITU
Mezinárodní telekomunikační unie

Předmluva

Úkolem Radiokomunikačního sektoru je zajišťovat racionální, rovnoměrné, efektivní a ekonomické využívání radiofrekvenčního spektra všemi radiokomunikačními službami včetně satelitních služeb a provádět studie bez omezení frekvenčního rozsahu, na jejichž základě se schvalují Rozhodnutí.

Regulační a metodické funkce Radiokomunikačního sektoru se realizují prostřednictvím Světových a Regionálních radiokomunikačních konferencí a Radiokomunikačních shromáždění podporovaných Studijními skupinami.

Zásady ochrany duševního vlastnictví (IPR)

Přístup ITU-R k ochraně duševního vlastnictví je popsán ve Společných patentových zásadách pro ITU-T/ITU-R/ISO/IEC uvedených v Příloze 1 Rozhodnutí ITU-R 1. Formuláře určené pro předkládání patentových prohlášení a licenčních prohlášení držitelů patentů jsou k dispozici na <http://www.itu.int/ITU-R/go/patents/en>, kde také najdete směrnice pro realizaci Společných patentových zásad pro ITU-T/ITU-R/ISO/IEC a patentovou informační databázi ITU-R.

Řada Doporučení ITU-R

(k dispozici také online na <http://www.itu.int/publ/R-REC/en>)

Řada	Název
BO	Satelitní přenos
BR	Záznamy pro výrobu, archivaci a přenos; film pro televizi
BS	Vysílací služby (zvuk)
BT	Vysílací služby (televize)
F	Pevné služby
M	Mobilní, radiolokační, amatérské a související satelitní služby
P	Šíření rádiových vln
RA	Rádiová astronomie
RS	Dálkové snímací systémy
S	Pevné satelitní služby
SA	Vesmírné aplikace a meteorologie
SF	Sdílení frekvencí a koordinace mezi systémy pevných satelitních a pevných služeb
SM	Správa spektra
SNG	Satelitní předávání zpravodajských příspěvků
TF	Vysílání časových signálů a frekvenčních standardů
V	Terminologie a související témata

Pozn.: Toto Doporučení ITU-R bylo schváleno v angličtině postupem popsaným v Rozhodnutí ITU-R 1.

Elektronická publikace
Ženeva, 2011

© ITU 2011

Všechna práva vyhrazena. Žádná část tohoto dokumentu se nesmí uveřejňovat za použití žádných prostředků bez písemného souhlasu ITU.

Doporučení ITU-R BS.1770-2

DOPORUČENÍ ITU-R BS.1770-2¹

Algoritmy pro měření hlasitosti zvukových programů a skutečné špičkové úrovně zvuku

(Otázka ITU-R 2/6)

(2006-2007-2011)

Rozsah

Toto doporučení popisuje algoritmy měření zvuku za účelem stanovení subjektivní hlasitosti pořadu a skutečné špičkové úrovně signálu.

Radiokomunikační shromáždění ITU

S uvážením,

- a) že moderní digitální metody přenosu zvuku nabízejí extrémně široký dynamický rozsah;
- b) že moderní digitální metody produkce a přenosu zvuku nabízejí směs mono, stereo a vícekanálových formátů a že zvukové pořady se vyrábějí ve všech těchto formátech;
- c) že si posluchači přejí, aby subjektivní hlasitost zvukových pořadů byla stejnoměrná u různých zdrojů a typů pořadů;
- d) že je k dispozici mnoho metod měření úrovně zvukového signálu, ale že stávající měřicí metody používané při výrobě programů neposkytují informace o subjektivní hlasitosti;
- e) že pro účely řízení hlasitosti při výměně pořadů, aby se omezilo rušení posluchačů, je důležité mít jednotný doporučený algoritmus pro objektivní hodnocení subjektivní hlasitosti;
- f) že budoucí komplexní algoritmy založené na psychoakustických modelech mohou poskytnout objektivní měřítka hlasitosti pro širokou škálu zvukových pořadů;
- g) že digitální média se mohou náhle přebudit, a je tedy třeba eliminovat i okamžité přebuzení,

a dále s uvážením

- h) že špičkové úrovně signálů mohou narůstat v důsledku běžně používaných procesů, jako je filtrování nebo snižování bitového toku;
- j) že stávající měřicí technologie neodrážejí skutečnou špičkovou úroveň obsaženou v digitálním signálu, protože skutečná špičková hodnota se může objevovat mezi vzorky;
- k) že vzhledem ke stavu zpracování digitálního signálu je praktické zavést algoritmus, který bude věrně odhadovat skutečnou špičkovou úroveň signálu;
- l) že používání algoritmu pro určení skutečných špiček umožní přesné stanovení pásma mezi špičkovou úrovní digitálního zvukového signálu a úrovní omezení,

¹ Radiokomunikační studijní skupina 6 provedla redakční úpravy tohoto Doporučení v květnu 2011

doporučuje,

- 1** aby se v případě, kdy bude potřeba objektivní měřítko hlasitosti zvukového kanálu nebo pořadu za účelem usnadnění přenosu a výměny pořadu, použil algoritmus popsany v Příloze 1;
- 2** aby metody používané při produkci a postprodukci pořadů k stanovení hlasitosti programu byly založeny na algoritmu popsáném v Příloze 1;
- 3** aby v případě, kdy bude potřeba údaj o skutečné špičkové úrovni digitálního zvukového signálu, byla měřicí metoda založena na pokynech uvedených v Příloze 2 nebo na metodě, která bude poskytovat podobné nebo lepší výsledky,

POZN. 1 – Uživatelé by si měli být vědomi toho, že naměřená hlasitost představuje odhad subjektivní hlasitosti a zahrnuje určitý stupeň odlišností v závislosti na posluchačích, zvukovém materiálu a poslechových podmínkách.

dále doporučuje

- 1** aby se dále pracovalo na rozšíření algoritmu popsáného v Příloze 1 tak, aby poskytoval údaje o krátkodobé hlasitosti;
- 2** aby se uvažovalo o případné potřebě aktualizace tohoto Doporučení v případě, že se ukáže, že nové algoritmy určování hlasitosti budou mít výrazně lepší výsledky než algoritmus popsany v Příloze 1.

POZN. 1 – Pro testování shody měřičů podle tohoto Doporučení se může používat testovací materiál ze sady popsané ve Zprávě ITU-R BT.2217.

Příloha 1

Popis algoritmu měření objektivní vícekanálové hlasitosti

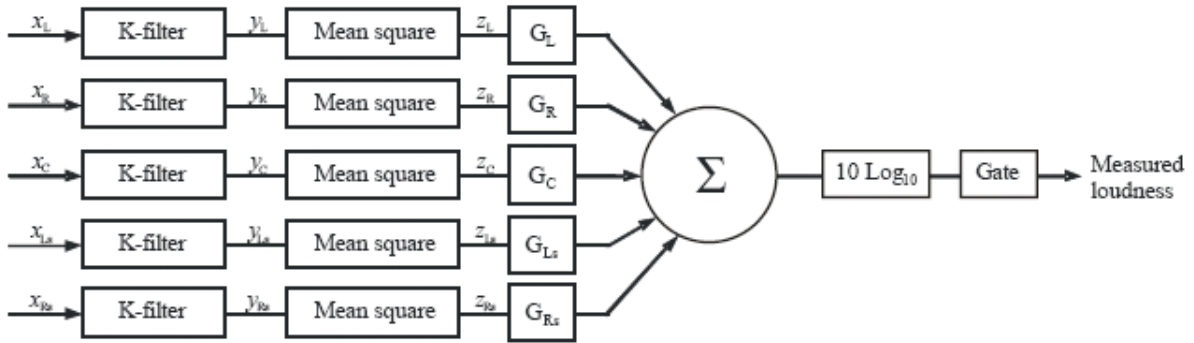
Tato Příloha popisuje algoritmus pro měření vícekanálové hlasitosti.

Tento algoritmus se skládá ze čtyř fází:

- Frekvenční váhování „K“;
- výpočtu střední kvadratické hodnoty pro každý kanál;
- sumarizace s váhování kanálů (prostorové kanály mají větší váhu a kanál LFE je vyloučen);
- hradlování 400 ms bloků (přesah o 75%), kde se používají dvě prahové hodnoty:
 - první na -70 LKFS;
 - druhá na -10 dB od úrovně naměřené po aplikaci první prahové hodnoty.

Obr. 1 obsahuje blokové schéma různých komponentů algoritmu. Na různých místech signálové cesty jsou uvedeny značky, které slouží k popisu algoritmu. Toto blokové schéma obsahuje vstupy pro pět hlavních kanálů (levý, středový, pravý, levý prostorový (surround) a pravý prostorový), což umožňuje měření programů obsahujících jeden až pět kanálů. U pořadu, který má méně než pět kanálů se některé vstupy nepoužijí. Kanál nízkofrekvenčních efektů (LFE) není do měření zahrnut.

OBR. 1
Zjednodušené blokové schéma algoritmu vícekanalové hlasitosti



K-filter = K filtr

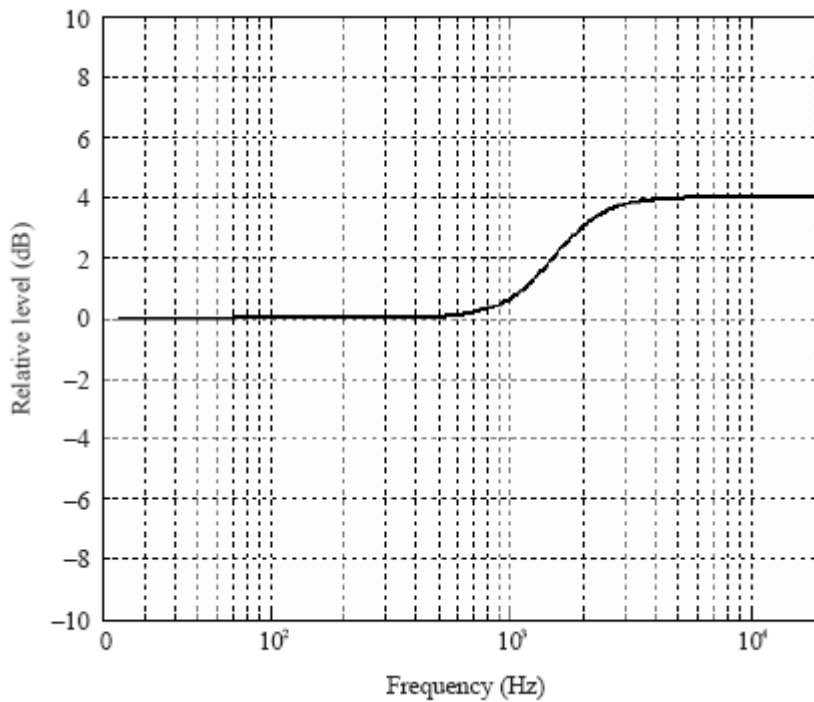
Gate = hradlo

Mean square = střední kvadratická hodnota

Measured loudness = naměřená hlasitost

Prvním krokem algoritmu je dvoufázové pre-filtrace¹ signálu. První fáze pre-filtrace zohledňuje akustické vlivy hlavy, kde hlava je modelovaná jako tuhá koule. Odezva je zobrazena na obr. 2.

OBR. 2
Kmitočtová odezva 1. stupně pre filtru používané k zohlednění akustických vlivů hlavy



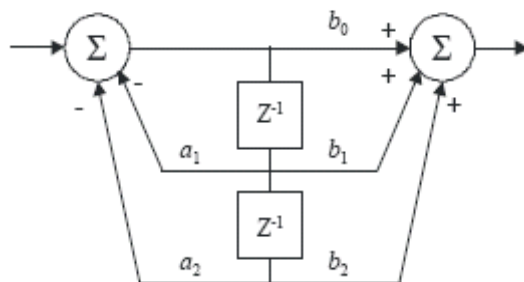
Relative level (dB) = relativní úroveň

Frequency (Hz) - frekvence

1. stupeň filtru je definován filtrem uvedeným na obr. 3 s koeficienty stanovenými v tabulce 1.

¹ Váhovací filtr K se skládá z dvoustupňové filtrace: pásmový (shelving) filtr v prvním stupni a hornopropustný filtr v druhém stupni (váhovací křivka RLB).

OBR. 3
Schéma toku signálu ve filtru 2. řádu



TABULKA 1
Koeficienty filtru pro 1. stupeň pre-filtru pro modelování kulové hlavy

		b_0	1,53512485958697
a_1	-1,69065929318241	b_1	-2,69169618940638
a_2	0,73248077421585	b_2	1,19839281085285

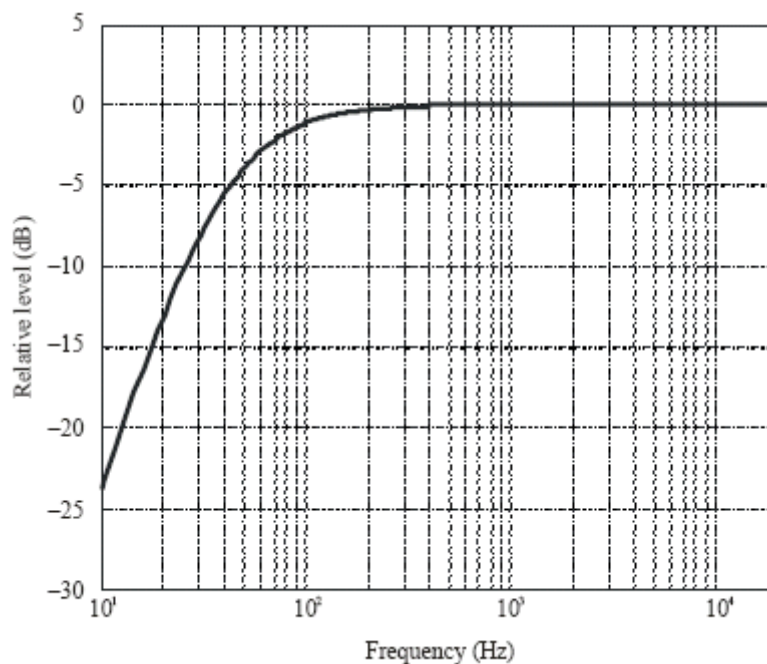
Tyto koeficienty filtru platí pro vzorkovací frekvenci 48 kHz. Realizace pro jiné vzorkovací frekvence budou vyžadovat jiné koeficienty, které je třeba volit tak, aby poskytovaly stejnou frekvenční odezvu, kterou uvedený filtr poskytuje při 48 kHz. Hodnoty těchto koeficientů bude třeba kvantovat kvůli přesnosti dostupného technického vybavení. Testy ukázaly, že platnost algoritmu není citlivá na malé odchylky těchto koeficientů.

Druhý stupeň pre filtru používá váhovací křivku RLB, která se skládá z jednoduchého hornopropustného filtru, jak vidíme na obr. 4.

Váhovací křivka RLB je vytvořena jako filtr 2. řádu, jak je uvedeno na obr. 3, s koeficienty stanovenými v tabulce 2.

Tyto koeficienty filtru platí pro vzorkovací frekvenci 48 kHz. Realizace při jiných vzorkovacích frekvencích budou vyžadovat jiné koeficienty, které je třeba volit tak, aby poskytovaly stejnou frekvenční odezvu, kterou uvedený filtr poskytuje při 48 kHz.

OBR. 4
Váhovací křivka RLB



Relative level (dB) = relativní úroveň
Frequency (Hz) - frekvence

TABULKA 2
Koeficienty filtru pro váhovací křivku RLB

		b_0	1,0
a_1	-1.99004745483398	b_1	-2,0
a_2	0.99007225036621	b_2	1,0

Výkon, střední kvadratická hodnota filtrovaného vstupního signálu v intervalu měření T se měří jako:

$$z_i = \frac{1}{T} \int_0^T y_i^2 dt \quad (1)$$

kde y_i je vstupní signál (filtrovaný 2fázovým pre filtrem, jak je popsán výše a $i \in I$ kde $I = \{L, R, C, Ls, Rs\}$, soubor vstupních kanálů.

Hlasitost v intervalu měření T je definována jako:

$$\text{Hlasitost, } L_K = -0.691 + 10 \log_{10} \sum_i G_i \cdot z_i \quad \text{LKFS} \quad (2)$$

kde G_i jsou váhovací koeficienty jednotlivých signálů.

K výpočtu hradlovaného měření hlasitosti se interval T rozdělí na soubor překrývajících se hradlovaných blokových intervalů. Hradlovaný blok je sada bezprostředně po sobě

následujících zvukových vzorků s trváním $T_g = 400$ ms k nejbližšímu vzorku. Překryv každého hradlovaného bloku bude 75% trvání hradlovaného bloku.

Měřicí interval bude omezen tak, aby končil na konci hradlovaného bloku. Neúplné hradlované bloky na konci měřicího intervalu se nepoužijí.

Výkon, střední kvadratická hodnota j -tého hradlového bloku i -tého vstupního kanálu v intervalu T je:

$$z_{ij} = \frac{1}{T_g} \int_{T_g \cdot j \cdot \text{step}}^{T_g \cdot (j \cdot \text{step} + 1)} y_i^2 dt \quad \text{kde} \quad \text{step} = 1 - \text{overlap} \quad (\text{krok} = 1 - \text{překryv})$$

$$\text{a} \quad j \in \left\{ 0, 1, 2, \dots, \frac{T - T_g}{T_g \cdot \text{step}} \right\} \quad (3)$$

Hlasitost j -tého hradlového bloku je definována jako:

$$l_j = -0.691 + 10 \log_{10} \sum_i G_i \cdot z_{ij} \quad (4)$$

Pro prahovou hodnotu hradlování Γ existuje soubor indexů hradlovaného bloku $J_g = \{j : l_j > \Gamma\}$, kde je hlasitost hradlovaného bloku nad prahovou hodnotou hradlování. Počet prvků v J_g je $|J_g|$.

Hradlovaná hlasitost v intervalu měření T je potom definována jako:

$$\text{Hradlovaná hlasitost, } L_{KG} = -0.691 + 10 \log_{10} \sum_i G_i \cdot \left(\frac{1}{|J_g|} \cdot \sum_{J_g} z_{ij} \right) \text{LKFS} \quad (5)$$

K provedení hradlovaného měření se používá dvoufázový proces, nejprve s absolutní prahovou hodnotou, potom s relativní prahovou hodnotou. Relativní prahová hodnota Γ_r se počítá změřením hlasitosti za použití absolutní prahové hodnoty $\Gamma_a = -70$ LKFS a odečtením 10 od výsledku, tedy:

$$\Gamma_r = -0.691 + 10 \log_{10} \sum_i G_i \cdot \left(\frac{1}{|J_g|} \cdot \sum_{J_g} z_{ij} \right) - 10 \text{ LKFS}$$

kde

$$J_g = \{j : l_j > \Gamma_a\}$$

$$\Gamma_a = -70 \text{ LKFS} \quad (6)$$

Hradlovaná hlasitost se potom může vypočítat pomocí Γ_r :

$$\text{Hradlovaná hlasitost, } L_{KG} = -0.691 + 10 \log_{10} \sum_i G_i \cdot \left(\frac{1}{|J_g|} \cdot \sum_{J_g} z_{ij} \right) \text{LKFS}$$

kde

$$J_{\xi} = \{j : I_j > \Gamma_r\} \quad (7)$$

Frekvenční váhování v tomto měření, které zajišťuje pre filtr (zřetězení filtru 1. stupně, který kompenzuje akustické vlivy hlavy, a filtru 2. stupně, váhování RLB), se označuje jako váhování K. Za číselným výsledkem hodnoty hlasitosti, která se počítá podle rovnice (2) by mělo následovat označení LKFS. Toto označení znamená:

Hlasitost (L) s váhováním K, vzhledem k jmenovitému plnému rozsahu (FS). Jednotka LKFS je ekvivalentní decibelu v tom, že nárůst úrovně signálu o 1 dB způsobí nárůst hodnoty o 1 LKFS.

Když se na vstup levého, středového nebo pravého kanálu přivede sinusová vlna 0 dB FS 1 kHz, bude naměřená hlasitost rovna -3,01 LKFS.

Váhovací koeficient pro každý kanál je uveden v tabulce 3.

TABULKA 3
Váhování jednotlivých zvukových kanálů

Kanál	Váha, G_i
Levý (GL)	1,0 (0 dB)
Pravý (GR)	1,0 (0 dB)
Středový (GC)	1,0 (0 dB)
Levý prostorový (GLs)	1,41 ($\sim +1,5$ dB)
Pravý prostorový (GRs)	1,41 ($\sim +1,5$ dB)

Je třeba uvést, že zatímco se prokázalo, že tento algoritmus funguje při aplikaci na zvukové pořady, jež mají typický obsah jako vysílání, tento algoritmus není obecně vhodný k určování subjektivní hlasitosti čistých tónů.

Dodatek 1 Přílohy 1

Popis a vývoj algoritmu vícekanálového měření

Tento dodatek popisuje nově vyvinutý algoritmus pro objektivní měření vnímané hlasitosti zvukových signálů. Tento algoritmus se může používat k přesnému měření mono, stereo a vícekanálových signálů. Základním přínosem navrhovaného algoritmu je jeho jednoduchost, jež umožňuje jeho zavedení s velmi nízkými náklady. Tento dodatek také popisuje výsledky formálních subjektivních testů provedených za účelem vytvoření subjektivní databáze, jež se používala k hodnocení funkčnosti algoritmu.

1 Úvod

Existuje mnoho oblastí, kde je potřeba měřit a řídit vnímanou hlasitost zvukových signálů. Příklady takových oblastí zahrnují televizní a rozhlasové vysílání, kde se povaha a obsah zvukového materiálu často mění. V těchto oblastech se zvukový obsah může neustále přepínat mezi hudbou, mluvenou řečí a zvukovými efekty, případně jejich kombinacemi. Takové změny obsahu programového materiálu mohou mít za následek výrazné změny subjektivní hlasitosti. Navíc signály často prodělávají různou formu zpracování dynamiky, což

může mít na subjektivně vnímanou hlasitost signálu podstatný vliv. Samozřejmě, že problematika subjektivní hlasitosti má také velkou důležitost v hudebním průmyslu, kde se zpracování dynamiky běžně používá k maximalizaci vnímané hlasitosti zvukového záznamu. V posledních letech se Radiokomunikační pracovní skupina 6P průběžně snažila najít objektivní prostředky měření vnímané hlasitosti typického programového materiálu používaného ve vysílání. V první fázi svého snažení ITU-R zkoumala výlučně algoritmy objektivní monofonní hlasitosti a ukázalo se, že u monofonních signálů nejlépe funguje vážená střední kvadratická hodnota $Leq(RLB)$ [Soulodre, 2004].

Ovšem uznává se, že pro vysílání je potřeba mít měřič hlasitosti, který by fungoval pro monofonní, stereofonní a vícekanálové signály. Tento dokument navrhuje nový algoritmus měření hlasitosti, který úspěšně pracuje s mono, stereo a vícekanálovými zvukovými signály. Navrhovaný algoritmus je založen na jednoduchém rozšíření algoritmu $Leq(RLB)$. Navíc si nový vícekanálový algoritmus zachovává velmi nízkou výpočetní složitost monofonního algoritmu $Leq(RLB)$.

2 Východiska

V první fázi studie ITU-R byla vyvinuta metoda subjektivního testování za účelem zkoumání vnímání hlasitosti typických monofonních programových materiálů [Soulodre, 2004]. Subjektivní testy byly provedeny v pěti lokalitách po celém světě, aby se vytvořila subjektivní databáze pro hodnocení spolehlivosti potenciálních algoritmů měření hlasitosti. Respondenti přiřazovali hlasitost různých monofonních zvukových sekvencí k referenční sekvenci. Zvukové sekvence byly převzaty z materiálů ze skutečného vysílání (televizního a rozhlasového).

Ve spojení s těmito testy bylo sedmi různými navrhovateli předloženo celkem deset komerčně vyvinutých monofonních měřičů/algoritmů hlasitosti k vyhodnocení do Laboratoře percepce zvuku Výzkumného střediska komunikací v Kanadě.

Navíc Soulodre dodal dva další základní algoritmy hlasitosti, které sloužily jako srovnávací měřítko spolehlivosti [Soulodre, 2004]. Tato dvě objektivní měřítka spočívala v jednoduché funkci frekvenčního váhování následované blokem měření střední kvadratické hodnoty. Jedno z těchto dvou měřítok, $Leq(RLB)$, používá frekvenční váhovací křivku typu horní propust označovanou jako revidovaná nízkofrekvenční křivka B (RLB).

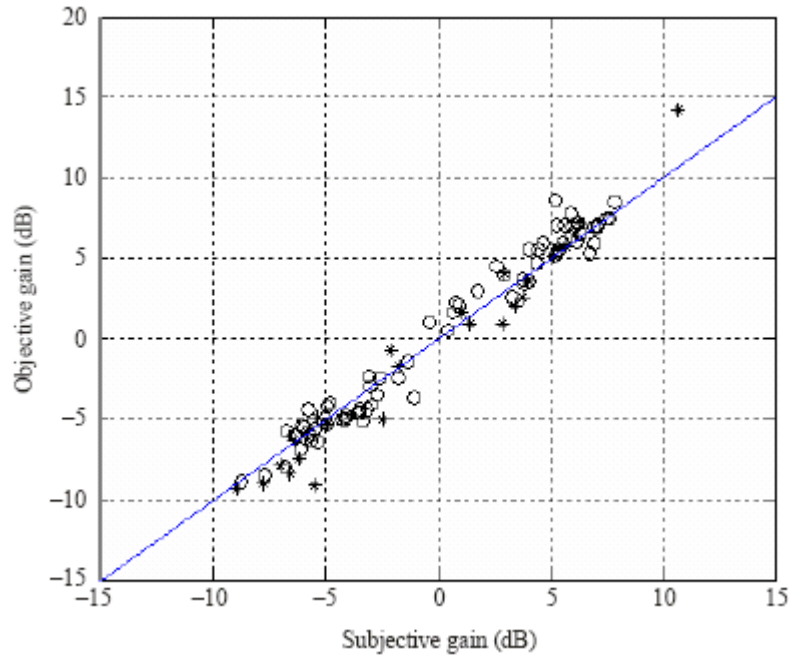
Druhé měřítko, Leq , je jednoduše neváhovaná střední kvadratická hodnota.

Obr. 5 udává výsledky výchozí studie ITU-R měřiče hlasitosti na základě hodnoty $Leq(RLB)$. Horizontální osa udává relativní subjektivní hlasitost odvozenou od subjektivní databáze, zatímco vertikální osa udává hlasitost předpokládanou parametrem $Leq(RLB)$. Každý bod na grafu představuje výsledek jedné ze zvukových testovacích sekvencí v testu. Prázdná kolečka označují zvukové sekvence založené na mluvené řeči, zatímco hvězdičky označují sekvence založené na neřečových signálech. Vidíte, že body jsou velmi těsně nahromaděné kolem úhlopříčky, což signalizuje velmi dobrou spolehlivost měřiče $Leq(RLB)$.

Měřič $Leq(RLB)$ vykázal nejlepší výkonnost ze všech hodnocených měřičů (ačkoliv v rámci statistické významnosti si některé měřiče založené na psychoakustických parametrech rovněž vedly dobře. Zjistilo se, že parametr Leq byl téměř stejně spolehlivý jako RLB. Tato zjištění naznačují, že u typického monofonního vysílacího materiálu je jednoduchý parametr hlasitosti založený na energii podobně robustní jako složitější parametry, které mohou zahrnovat podrobné percepční modely.

OBR. 5

Monofonní měřič hlasitosti Leq (RLB) versus subjektivní výsledky ($r = 0,982$)



Objective gain (dB) = objektivní zesílení
Subjective gain (dB) = subjektivní zesílení

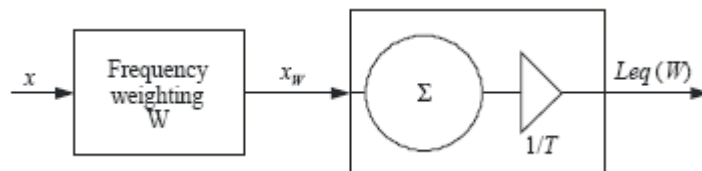
3 Návrh algoritmu Leq (RLB)

Algoritmus hlasitosti Leq (RLB) byl speciálně navržen tak, aby byl velmi jednoduchý. Blokové schéma algoritmu Leq (RLB) je uvedeno na obr. 6. Skládá se z hornopropustného filtru, po němž následuje výpočet časový průměr energie. Výstup filtru prochází blokem zpracování, kde se určí celková energie a vypočítá se průměr v čase.

Účelem filtru je poskytnout určité percepčně relevantní váhování spektrálního obsahu signálu. Jednou výhodou používání této základní struktury pro měření hlasitosti je, že veškeré zpracování lze provádět pomocí jednoduchých časových bloků s velmi nízkými výpočetními požadavky.

OBR. 6

Blokové schéma jednoduchého měření hlasitosti založeného na energii



Frequency weighting = frekvenční váhování

Algoritmus Leq (RLB) uvedený na obr. 6 je jednoduše frekvenčně váhovanou verzí *Ekvivalentní úroveň zvukového signálu* (Leq). Leq se definuje následovně:

$$Leq(W) = 10 \log_{10} \left[\frac{1}{T} \int_0^T \frac{x_W^2}{x_{Ref}^2} dt \right] \quad \text{dB} \quad (3)$$

kde

x_W : signál na výstupu váhového filtru

x_{Ref} : určitá referenční úroveň

T : délka zvukové frekvence.

Symbol W v $Leq(W)$ představuje frekvenční váhování, jež je v tomto případě provedeno revidovanou nízkofrekvenční křivkou B (RLB).

4 Subjektivní testy

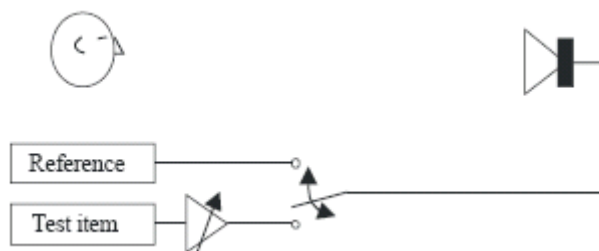
Aby bylo možné vyhodnotit potenciální algoritmy měření vícekanálové hlasitosti, bylo potřeba provést formální subjektivní testy, aby se vytvořila subjektivní databáze. Potom bylo možné vyhodnotit schopnost potenciálních algoritmů hlasitosti předpovídat výsledky subjektivních testů. Databáze poskytla hodnocení subjektivní hlasitosti pro širokou škálu monofonních, stereofonních a vícekanálových programových materiálů. Programové materiály používané při testech byly převzaty ze skutečného televizního a rozhlasového vysílání z celého světa i z nosičů CD a DVD. Sekvence zahrnovaly hudbu, televizní hry a filmy, sportovní přenosy, zpravodajství, zvukové efekty a reklamy. V sekvencích byly zařazeny řečové segmenty v několika jazycích.

4.1 Schéma subjektivních testů

Subjektivní testy spočívaly v úkolu přiřazení hlasitosti. Respondenti poslouchali širokou škálu typických programových materiálů a upravovali úroveň každé testované položky, dokud její vnímaná hlasitost neodpovídala hlasitosti referenčního signálu (viz obr. 7).

Referenční signál byl vždy reprodukován na úrovni 60 dBA, což je úroveň, kterou Benjamin určil jako typickou úroveň poslechu při sledování televize ve skutečných domácnostech [Benjamin, 2004].

OBR. 7
Metodika subjektivních testů



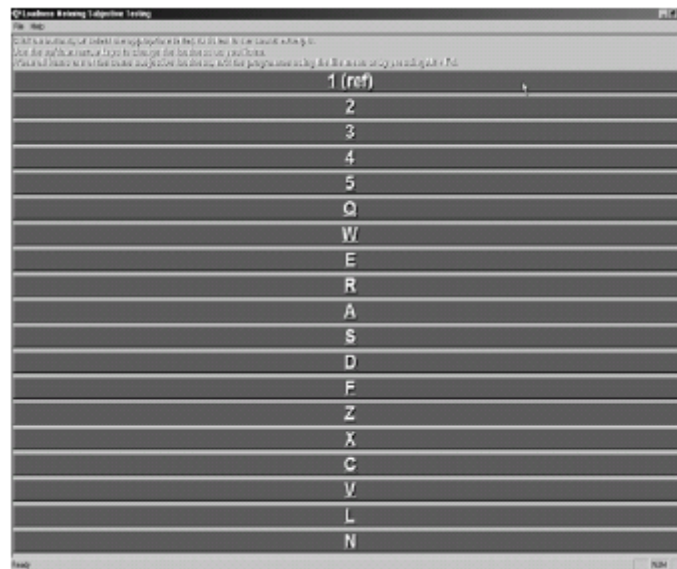
Reference = referenční úroveň

Test item = testovaný materiál

Softwarový vícekanálový systém pro subjektivní testování, vyvinutý a poskytnutý Australskou vysílací korporací, umožňoval posluchači rychle přepínat tam a zpět mezi testovacími materiály a upravovat úroveň (hlasitost) každého materiálu. Ukázka obrazovky testovacího

softwaru je uvedena na obr. 8. Úroveň hlasitosti testovacích materiálů bylo možné upravovat v krocích po 0,25 dB. Při zvolení tlačítka označeného „1“ měl posluchač přístup k referenčnímu signálu. Úroveň referenčního signálu se udržovala pevná.

OBR. 8
Uživatelské rozhraní systému pro subjektivní testování



Pomocí klávesnice počítače zvolil respondent danou testovací položku a upravil její úroveň tak, aby její hlasitost odpovídala referenčnímu signálu. Respondenti mohli okamžitě přepínat mezi libovolnými testovacími položkami pomocí příslušných kláves. Sekvence se během testů přehrávaly nepřetržitě (ve smyčkách). Software zaznamenával nastavení zesílení pro každou položku tak, jak je provedl respondent. Tudíž subjektivní testy poskytly soubor hodnot zesílení (v decibelech) požadovaných k přizpůsobení každé testovací sekvence referenční sekvenci. Tak bylo možné přímo určit relativní hlasitost každé testované položky.

Před provedením formálních testů naslepo prošel každý respondent školením, v němž se seznámil s testovacím softwarem a se svým úkolem v experimentu. Protože mnoho testovacích materiálů obsahovalo směs řeči a jiných zvuků (např. hudby, hluku pozadí, atd.), dostali respondenti specifickou instrukci, aby upravili hlasitost celkového signálu, ne jen řečového komponentu signálů.

Během formálních testů naslepo bylo pořadí, v němž byly testovací položky předkládány každému subjektu, uspořádány náhodně. Žádní dva respondenti tedy neposlouchali testovací materiály ve stejném pořadí. To bylo provedeno proto, aby se eliminovalo případné zkreslení pod vlivem pořadí.

4.2 Subjektivní databáze

Subjektivní databáze používaná k vyhodnocení spolehlivosti navrhovaného algoritmu se vlastně skládala ze tří samostatných datových souborů. Datové soubory byly vytvořeny ze tří nezávislých subjektivních testů provedených v průběhu několika let.

První datový soubor obsahoval výsledky původní studie ITU-R, kde respondenti přiřazovali vnímanou hlasitost 96 monofonních zvukových sekvencí. Pro tento datový soubor byly subjektivní testy provedeny v pěti odlišných lokalitách po celém světě s celkem 97 posluchači. Testovací sekvence a referenční položku zvolila tříčlenná komise složená ze členů Radiokomunikační pracovní skupiny 6P SRG3. Referenční signál představovala řeč v angličtině namluvená ženou. Sekvence se přehrávaly jedním reproduktorem umístěným přímo před posluchačem.

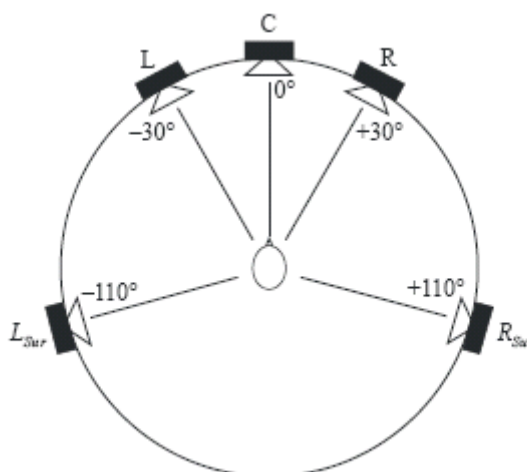
Po původní monofonní studii ITU-R někteří předkladatelé algoritmů spekulovali, že rozsah a typ signálů používané při subjektivních testech nebyly dostatečně široké. Dále spekulovali, že to byl ten důvod, proč si jednoduchý algoritmus $Leq(RLB)$ založený na energii vedl lépe než všechny ostatní algoritmy.

Aby se rozptýlila tato obava, byli předkladatelé požádáni, aby předložili nové zvukové sekvence pro další kolo subjektivních testů. Bylo jim doporučeno, aby předkládali monofonní sekvence, které podle jejich názoru algoritmus $Leq(RLB)$ lépe prověří. Nové sekvence poskytli pouze dva z předkladatelů měřičů.

Za pomoci těchto nových sekvencí byly provedeny nové subjektivní testy v Laboratoři percepce zvuku ve Výzkumném centru komunikací v Kanadě. Celkem 20 respondentů provedlo hodnocení hlasitosti 96 monofonních sekvencí. Při testech se používala stejná subjektivní metodika jako při vytváření prvního datového souboru a také se používal stejný referenční signál. Výsledky těchto testů vytvořily druhý datový soubor subjektivní databáze.

Třetí datový soubor obsahoval hodnocení hlasitosti 144 zvukových sekvencí. Zvukové sekvence se skládaly z 48 monofonních položek, 48 stereofonních položek a 48 vícekanálových položek. Navíc se polovina monofonních položek přehrávala přes středový kanál (mono), zatímco druhá polovina monofonních položek se přehrávala přes levý a pravý reproduktor (duální mono). To se provádělo kvůli dvěma možným způsobům, kterými je možné poslouchat monofonní signál. Pro tento test spočíval referenční signál v řeči v angličtině namluvené ženou se stereofonním prostředím a tichou hudbou na pozadí. Tohoto testu se účastnilo celkem 20 respondentů a používalo se při něm uspořádání reproduktorů uvedené v Doporučení ITU-R BS.775 a vyobrazené na obr. 9.

OBR. 9
Uspořádání reproduktorů použité pro třetí datový soubor



První dva datové soubory byly omezené na monofonní testovací sekvence, takže prostorové modelování nepředstavovalo žádný faktor. U třetího datového souboru, který také obsahoval stereofonní a vícekanálové sekvence, se prostorové modelování stalo důležitým hlediskem,

kteřé bylo potřeba zohlednit. Zdálo se být pravděpodobné, že rozložení kanálů a charakter zvukového prostředí může mít podstatný vliv na subjektivní hlasitost sekvence. Proto byly stereofonní a vícekanálové sekvence vybírány tak, aby zahrnovaly širokou škálu konfigurace kanálů (např. středovou panoramu versus silný pravolevý posuv, zdroje vepředu versus zdroje všude kolem) a různá nastavení dozvuku prostředí (např. suché versus dozvukové prostředí).

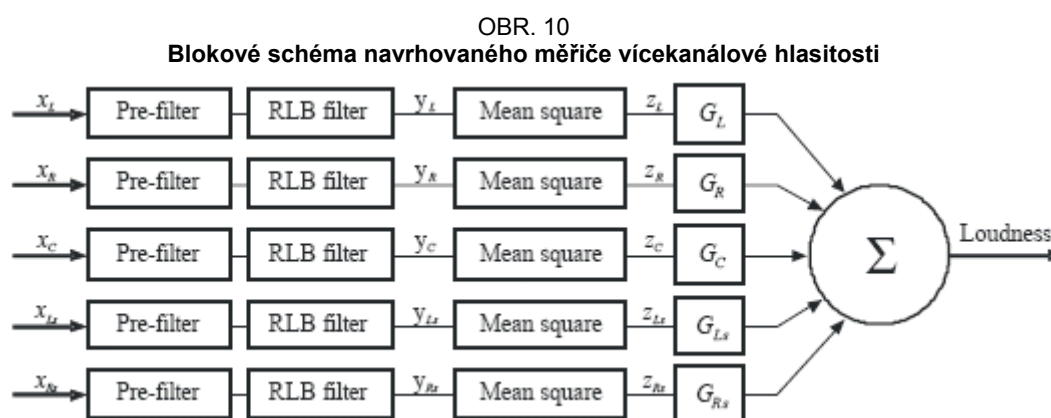
Skutečnost, že respondenti museli zároveň přiřazovat hlasitost monofonních, duálních monofonních, stereofonních a vícekanálových signálů znamenala, že tento test byl z podstaty složitější než s předchozími datovými soubory, které se omezovaly na monofonní signály. Tato složitost byla rozšířena různým rozložením kanálů a různým dozvukovým prostředím. Byla obava, aby v důsledku těchto faktorů nebyli respondenti úkolem zahlceni. Naštěstí předběžné testy napověděly, že tento úkol je zvládnutelný, a skutečně 20 respondentů dokázalo poskytnout konzistentní výsledky.

5 Návrh algoritmu měření vícekanálové hlasitosti

Jak již bylo uvedeno, byl algoritmus Leq (RLB) vytvořen pro práci s monofonními signály a dřívější studie ukázala, že si při tomto úkolu vedl úspěšně. Návrh algoritmu pro měření vícekanálové hlasitosti s sebou přináší několik dalších problémů. Klíčovým požadavkem pro úspěšný vícekanálový algoritmus je, že musí fungovat dobře pro monofonní, duální monofonní a stereofonní signály. To znamená, že na tyto formáty se musí pohlížet jako na speciální případy vícekanálového signálu (i když jsou to velmi běžné případy).

V předkládané studii předpokládáme, že vícekanálové signály odpovídají standardnímu uspořádání kanálů podle Doporučení ITU-R BS.775 5.1. Nijak se nesnažíme zohlednit kanál LFE.

Měřič vícekanálové hlasitosti měří hlasitost každého jednotlivého zvukového kanálu nezávisle pomocí monofonního algoritmu Leq (RLB), jak je uvedeno na obr. 10. Nicméně na každý kanál se aplikuje pre filtrace před vlastním stanovením Leq (RLB).



Pre-filter = pre filtr
 RLB filter = filtr RLB
 Mean square = střední kvadratická hodnota
 Loudness = hlasitost

Účelem pre filtru je zohlednit akustické vlivy, které má hlava na příchozí signály. Zde se hlava modeluje jako tuhá koule. Na každý kanál se aplikuje stejný pre filtr. Výsledné hodnoty

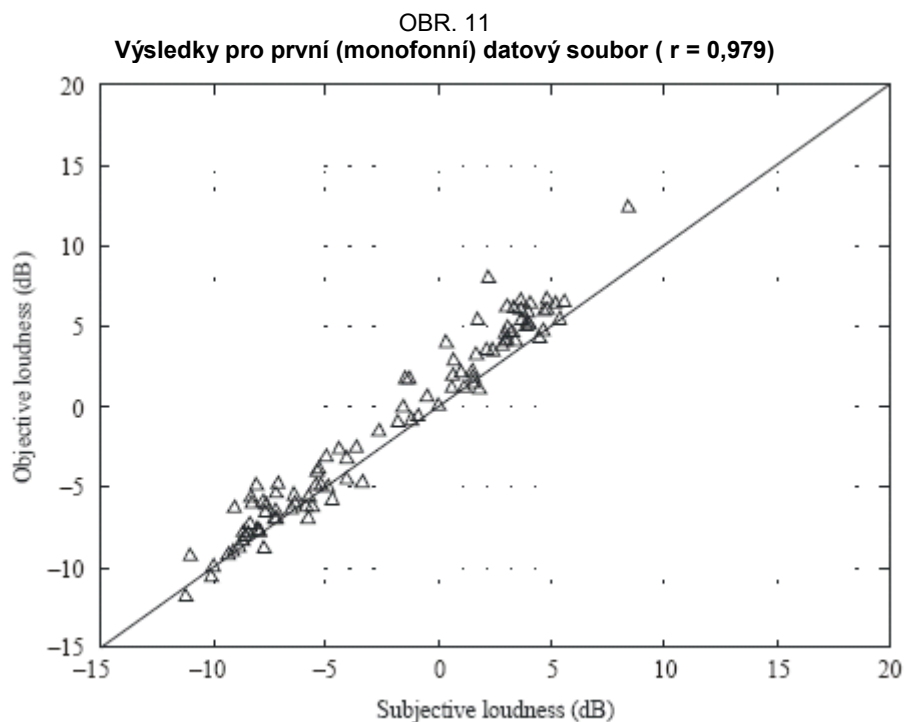
hlasitosti se potom váhují (G_i) podle úhlu, z něhož signál přichází a potom se sčítají (v lineární doméně), aby vytvořily složenou hodnotu hlasitosti. Vážení se používá k zohlednění skutečnosti, že posluchač může zvuky přicházející zezadu vnímat jako hlasitější než zvuky přicházející zepředu.

Základní výhodou navrhovaného algoritmu měření vícekanálové hlasitosti je jeho jednoduchost. Algoritmus se skládá výhradně z velmi základních bloků zpracování signálu, které lze snadno implementovat v časové doméně na cenově dostupném zařízení. Další zásadní výhodou tohoto algoritmu je jeho škálovatelnost. Protože je zpracování každého kanálu identické, je velmi jednoduché vytvořit měřič, který zvládne libovolný počet kanálů od 1 do N . Navíc, protože se příspěvky jednotlivých kanálů sčítají jako hodnoty hlasitosti spíše než na úrovni signálu, nezávisí algoritmus na fázi mezi kanály nebo jejich korelaci. Díky tomu je navrhované měření hlasitosti mnohem obecnější a robustní.

6 Hodnocení vícekanálového algoritmu

Navrhovaným algoritmem bylo vyhodnoceno 336 zvukových sekvencí ze tří datových souborů a byly zaznamenány stanovené hodnoty hlasitosti. Pomocí tohoto procesu bylo možné vyhodnotit celkovou spolehlivost algoritmu na základě shody mezi stanoveným hodnocením a skutečnými subjektivními hodnoceními získanými při formálních subjektivních testech.

Obr. 11, 12 a 13 obsahují grafy spolehlivosti navrhovaného měřiče hlasitosti u tří výše uvedených datových souborů. Na každém grafu udává horizontální osa subjektivní hlasitost každé zvukové sekvence v datovém souboru. Vertikální osa udává objektivní hlasitost stanovenou navrhovaným měřičem hlasitosti. Každý bod na grafu představuje výsledek jednotlivé zvukové sekvence. Je třeba poznamenat, že u dokonalého objektivního algoritmu by všechny datové body byly na úhlopříčce se sklonem 1, která by procházela počátkem (jak je vidět na grafech)



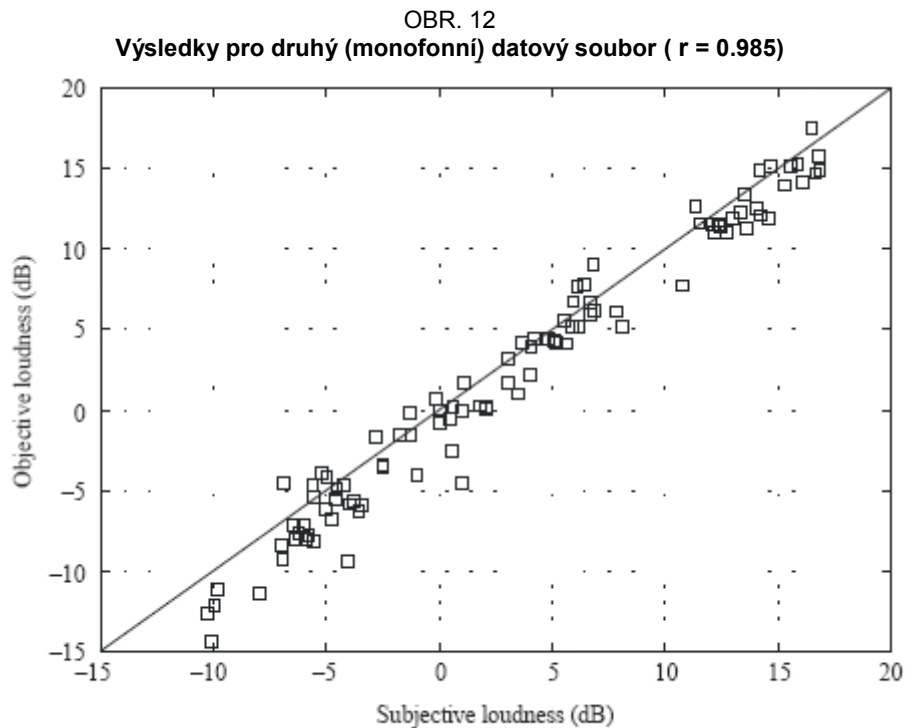
△ 1st dataset

Objective loudness (dB) = objektivní hlasitost

Subjective loudness (dB) = subjektivní hlasitost
1st dataset = 1. datový soubor

Na obr. 11 je vidět, že si navrhovaný algoritmus měření vícekanálové hlasitosti při předvídání výsledků u prvního (monofonního) datového souboru vede velmi dobře. Korelace mezi subjektivním hodnocením hlasitosti a objektivním měřením hlasitosti $r = 0,979$.

Jak je vidět na obr. 12, je korelace mezi hodnocením subjektivní hlasitosti a objektivním měřením hlasitosti u druhého datového souboru také velmi dobrá ($r = 0,985$). Je zajímavé si všimnout, že asi polovina sekvencí v tomto datovém souboru byla hudební.

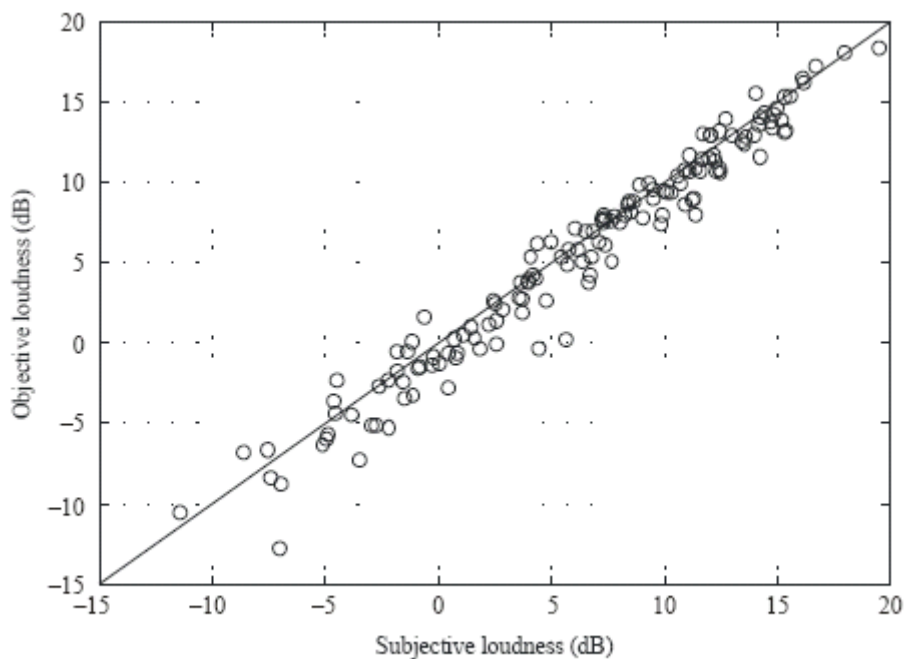


□ 2nd dataset

Objective loudness (dB) = objektivní hlasitost
Subjective loudness (dB) = subjektivní hlasitost
2nd dataset = 2. datový soubor

OBR. 13

Výsledky pro třetí (mono, stereo a vícekanálový) datový soubor ($r = 0,980$)



○ 3rd dataset

Objective loudness (dB) = objektivní hlasitost

Subjective loudness (dB) = subjektivní hlasitost

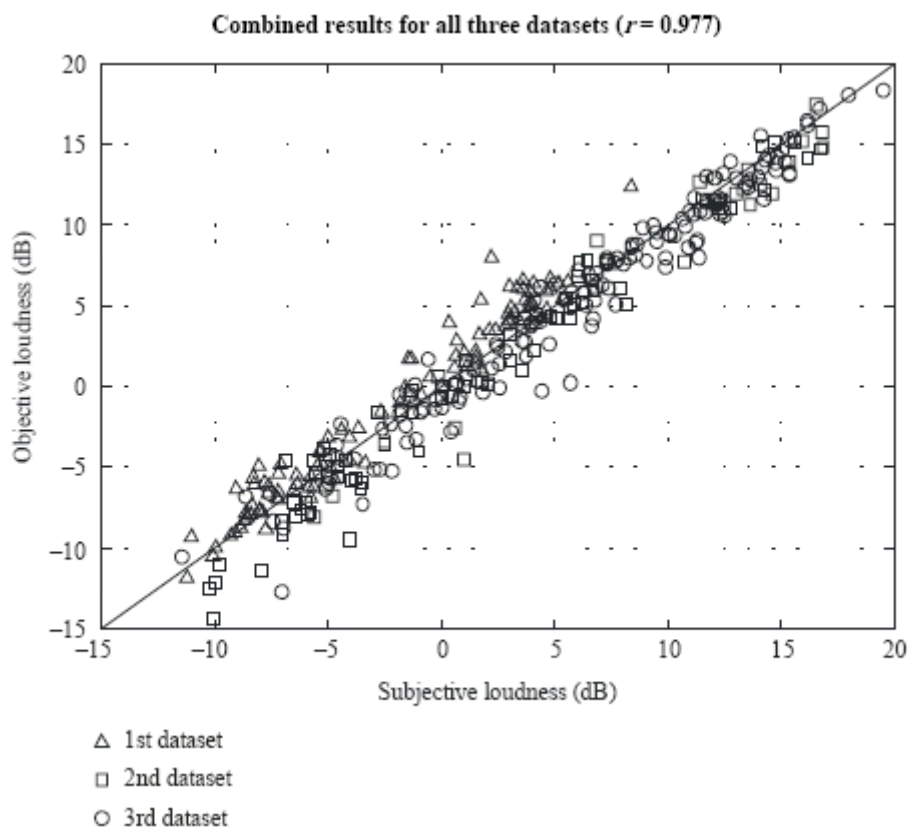
3rd dataset = 3. datový soubor

Obr. 13 uvádí výsledky pro třetí datový soubor, který obsahoval monofonní, duální monofonní, stereofonní a vícekanálové signály. Vícekanálové výsledky obsažené na obr. 13 a 14 platí pro uvedený algoritmus, ale s váhou prostorového kanálu nastavenou na 4 dB (původní návrh) místo 1,5 dB (konečný údaj). Bylo ověřeno, že rozsah od 4,0 dB do 1,5 dB nemá žádný významný vliv na výsledky. Spolehlivost algoritmu je i v tomto případě velmi dobrá, s korelací $r = 0,980$.

Je užitečné prozkoumat spolehlivost algoritmu pro všech 336 zvukových sekvencí, které vytvořily subjektivní databázi. Proto obr. 14 shrnuje výsledky ze všech tří datových souborů, Je vidět že spolehlivost je velmi dobrá v celé subjektivní databázi při celkové korelaci $r = 0,977$.

OBR. 14

Souhrnné výsledky pro všechny tři datové soubory ($r = 0.977$)



Objective loudness (dB) = objektivní hlasitost
 Subjective loudness (dB) = subjektivní hlasitost
 1. dataset = 1. datový soubor
 2. dataset = 2. datový soubor
 3. dataset = 3. datový soubor

Výsledky tohoto hodnocení ukazují, že si algoritmus měření vícekanálové hlasitosti založený na parametru měření hlasitosti $Leq(RLB)$ vede velmi dobře u všech 336 sekvencí subjektivní databáze. Subjektivní databáze poskytla širokou nabídku programových materiálů včetně hudby, televizních her a filmů, sportovních přenosů, zpravodajských pořadů, zvukových efektů a reklam. Do sekvencí byly také zapojeny řečové segmenty v několika jazycích. Výsledky navíc prokazují, že navrhovaný měřič hlasitosti funguje dobře u monofonních, duálních monofonních, stereofonních i vícekanálových signálů.

Literatura

BENJAMIN, E. [říjen 2004] Preferred Listening Levels and Acceptance Windows for Dialog Reproduction in the Domestic Environment, 117th Convention of the Audio Engineering Society, San Francisco, Preprint 6233.

SOULODRE, G.A. [květen 2004] Evaluation of Objective Loudness Meters, 116th Convention of the Audio Engineering Society, Berlin, Preprint 6161.

Příloha 2

Pokyny pro přesné měření „skutečné špičkové“ úrovně

Tato příloha popisuje algoritmus pro odhad skutečné špičkové úrovně v rámci jednokanálového digitálního PCM zvukového signálu. Následující rozbor předpokládá vzorkovací frekvenci 48 kHz. Skutečná špičková úroveň je maximální (pozitivní nebo negativní) hodnota křivky signálu v kontinuální časové doméně: tato hodnota může být vyšší než největší hodnota vzorku v doméně s 48kHz časovanými vzorky. Tento algoritmus poskytuje odhad pro signál tak, jak je, a volitelně, jak by byl v případě, že by nějaké další zařazené zařízení odstranilo složku DC signálu. Přídavné mírné předběžné zdůraznění vysokých frekvencí na trase signálu měření špičkové úrovně může umožnit algoritmu hlásit vyšší špičkovou úroveň pro vysokofrekvenční signály, než je tomu ve skutečnosti. Účelem toho je, že fázové posuny následných fází zpracování signálu (jako jsou Nyquistovy filtry) by mohly způsobit nárůst špiček vysokofrekvenčních signálů a v některých oblastech by tato funkce mohla být užitečná k zajištění další ochrany před ořezáním signálu dále na trase.

1 Souhrn

Fáze zpracování jsou:

- 1 Útlum: útlum o 12,04 dB
- 2 4 × převzorkování
- 3 Zdůraznění: pásmový filtr pro pre-emfázi, nula na 14,1 kHz, pól na 20 kHz (volitelné)
- 4 DC blok (volitelný)
- 5 Absolutní: absolutní hodnota
- 6 Max: detekce nejvyšší hodnoty (volitelná, zařazena, když je zařazen DC blok).

Detekce absolutní hodnoty jak před, tak za DC blokem umožňuje odhadnout špičkovou úroveň signálu v aktuálním bodě měření a rovněž odhadnout špičkovou úroveň, pokud na některé dále zařazeném zařízení dojde k odstranění DC komponentu signálu.

2 Blokové schéma

Vstup	Útlum	4 Převzorkování	Zdůraznění (volitelné)	Absolutní	Výstup 1
				Volitelný DC blok	nebo Výstup 2
				Max. Absolutní	

3 Podrobný popis

První krok spočívá v útlumu signálu o 12,04 dB (2bitový posuv). Účelem tohoto kroku je vytvořit rezervu pro následné zpracování signálu za použití celočíselné aritmetiky. Tento krok není potřeba, pokud se výpočty provádějí s plovoucí desetinnou čárkou.

Filtr 4x převzorkování zvyšuje vzorkovací frekvenci z 48 kHz na 192 kHz. Tato verze signálu s vyšší frekvencí přesněji vystihuje skutečnou křivku, která je signálem reprezentována. Upřednostňují se vyšší vzorkovací frekvence a poměry převzorkování (Viz dodatek 1 této přílohy). Příchozí signály, které mají vyšší vzorkovací frekvence vyžadují v poměru menší převzorkování (např. u příchozího signálu při vzorkovací frekvenci 96 kHz stačí převzorkování 2x).

Volitelná preemfáze zajišťuje, že algoritmus udává vyšší špičkovou úroveň pro složky signálu na nejvyšších frekvencích. To se může provádět na základě úvahy, že je obtížnější měřit a upravovat špičkové hodnoty složek signálu na nejvyšších frekvencích kvůli působení rozptylu (fázového posunu), k němuž dochází v četných Nyquistových filtrech, jež se často vyskytují různě v řetězci vysílacího signálu.

Volitelný filtr blokování stejnosměrné složky se používá pro případy, že signál je silně asymetrický nebo obsahuje nějakou stejnosměrnou odchylku. Vedle měření špičkové hodnoty současného signálu (včetně asymetrie a/nebo stejnosměrné složky) umožňuje zařazení této volitelné části měření signálu tak, jak by vypadal, kdyby nějaké dále zařazené zařízení obsahovalo filtr pro blokování stejnosměrné složky.

Absolutní hodnota vzorků se získá invertováním vzorků s negativní hodnotou; v tomto bodě je signál unipolární, s negativními hodnotami nahrazenými pozitivními hodnotami stejné velikosti. Výstup 1 je proud výstupních hodnot, pokud není použitý volitelný DC blok.

Pokud je volitelný DC blok zařazený, blok „MAX“ vybírá větší z každého vzorku ze dvou signálových cest; v tomto případě se výstup bere z výstupu 2.

Další systémové bloky (nezobrazené ani nepopisované v tomto dokumentu) mohou porovnávat hodnoty výstupního vzorku s jmenovitou 100% špičkovou úrovní signálu (1/4 celého rozsahu, pokud se na vstupu použil útlum o 12 dB), což poskytuje odhad skutečné špičkové úrovně s ohledem na digitální plnou škálu.

Měřiče, které odpovídají těmto pokynům, a které používají převzorkovanou vzorkovací frekvenci minimálně 192 kHz, by měly udávat výsledek v jednotkách dB TP. Toto označení znamená decibely vztažené k 100% plné stupnici, měření skutečné špičkové hodnoty.

Dodatek 1¹ **Přílohy 2**

Úvahy k přesnému měření špiček digitálních zvukových signálů

V čem je problém?

Měřiče špiček v digitálních zvukových systémech často registrují „špičkový vzorek“ spíše než „skutečnou špičku“.

Měřič špičkových vzorků obvykle funguje tak, že porovnává absolutní (rektifikovanou) hodnotu každého příchozího vzorku s aktuální hodnotou měřiče; pokud je hodnota nového vzorku vyšší, nahradí aktuální hodnotu; pokud ne, aktuální hodnota se vynásobí konstantou, která je mírně nižší než jednotka k získání logaritmického úbytku. Takové měřiče se vyskytují

¹ POZN. 1 – Následující informativní příspěvek poskytla Pracovní skupina pro normy AES SC-02-01 prostřednictvím Radiokomunikačního zpravodaje VP 6J o měření hlasitosti

všude, protože se snadno aplikují, ale ne vždy zaznamenávají skutečnou špičkovou hodnotu zvukového signálu.

Takže používání měřiče špičkových vzorků tam, kde je měření špiček programu důležité, může vést k problémům. Bohužel většina digitálních měřičů špiček jsou měřiče špičkových vzorků, ačkoliv to uživateli ne vždy musí být zřejmé.

K tomuto problému dochází, protože skutečné špičkové hodnoty vzorkovaného signálu se obvykle vyskytují spíše mezi vzorky než přesně v okamžiku vzorkování a jako takové je měřič špičkových vzorků nezaznamená.

To má za následek několik známých anomálií měřičů špičkových vzorků:

- *Nekonzistentní údaje o špičkách:* Často si lze všimnout, že opakované nahrávání analogového záznamu do digitálního systému s měřičem špičkových vzorků vede ke zcela odlišným hodnotám špiček programu při každém přehrání. Podobně když se digitální záznam opakovaně přihrává přes převaděč vzorkovací frekvence před měřením, zaznamenané špičky jsou obdobně odlišné při každém přehrání. To vyplývá z toho, že okamžiky vzorkování mohou spadat na různé části skutečného signálu při každém přehrání.
- *Nečekaná přebuzení:* Protože vzorkované signály mohou obsahovat přebuzení, i když nemají žádné vzorky nad plným digitálním rozsahem ani v jeho blízkosti, je signalizace přebuzení měřičem špičkových vzorků nespolehlivá. Přebuzení může způsobovat ořezávání v následných procesech, např. v určitých D/A převodnicích nebo během převodu vzorkovací frekvence, i když je předtím měřič špičkových vzorků nezaznamenal (a dokonce byly v tom místě neslyšné).
- *Nízké hodnoty a rázy měřených tónů:* Čisté tóny (jako např. nastavovací tóny) blízko celočíselných násobků vzorkovací frekvence mohou vykazovat nižší hodnoty nebo vykazovat neustále proměnné hodnoty, i když je amplituda tónu konstantní.

Jak závažný může ten problém být?

Obecně platí, že čím vyšší je frekvence měřeného signálu špičkových vzorků, tím větší je riziko potenciální chyby.

To lze snadno prokázat u kontinuálních čistých tónů, např. o 3 dB nižší údaj u nešťastně fázovaného tónu na čtvrtině vzorkovací frekvence. Snížení hodnoty u tónu na polovině vzorkovací frekvence může být téměř neomezené; nicméně většina digitálních zvukových systémů nemá na této frekvenci výraznou frekvenci (protože ji do velké míry vyloučí anti-aliasingové filtry v bodě D/A převodu a protože „skutečným“ zvukům obvykle nedominují trvalé vysoké frekvence).

U kontinuálních tónů, které nejsou blízké nízkým celočíselným násobkům vzorkovací frekvence, nedochází u měřičů špičkových vzorků k zobrazování příliš nízkých hodnot, protože frekvence záznějů (rozdíl mezi $n \cdot f_{tone}$ a f_s) je vysoká ve srovnání s reciprokou hodnotou rychlosti doběhu měřiče. Jinými slovy je okamžik vzorkování dostatečně blízko skutečné špičky tónu dostatečně často na to, aby měřič neudával příliš nízké hodnoty.

Nicméně u jednotlivých přechodových jevů se příliš nízké hodnoty tímto mechanismem neskryjí, takže čím vyšší je frekvenční obsah přechodového jevu, tím větší je potenciální snížení naměřené hodnoty. U „skutečných tónů“ je normální, že přechodové jevy se vyskytují s výrazně vysokým frekvenčním obsahem a chyba naměřené hodnoty může v takových případech běžně činit několik dB.

Protože skutečné zvuky mají obecně spektrum, které směrem k vyšším frekvencím klesá a protože se to nezmění zvýšením vzorkovací frekvence, bývá vykazování nižších hodnot u měřičů špičkových vzorků méně závažné při vyšších původních vzorkovacích frekvencích.

Jaké je řešení?

Aby se měřila skutečná špičková hodnota vzorkovaného signálu, je potřeba „převzorkovat“ (nebo „dovzorkovat“) signál v zásadě tak, že dotvoříte původní signál mezi stávajícími vzorky a tak zvýšíte vzorkovací frekvenci signálu. Tento návrh zní pochybně: jak můžeme obnovit informace, které jsou pravděpodobně již ztracené? Ve skutečnosti teorie vzorkování ukazuje, že to možné je, protože víme, že vzorkovaný signál neobsahuje žádné frekvence nad polovinou původní vzorkovací frekvence.

Jaký poměr převzorkování je potřeba? Abychom to zjistili, musíme si odpovědět na pár otázek:

- Jaká je maximální přijatelná chyba měření?
- Jaký je poměr nejvyšší měřené frekvence ke vzorkovací frekvenci (maximální „normalizované“ frekvenci)?

Pokud známe tato kritéria, je možné vypočítat poměr převzorkování, který potřebujeme (i ještě bez zvážení podrobností realizace převzorkování) pomocí jednoduché metody „čtverečkování papíru“. Jednoduše můžeme uvážit, jaká chyba spočívající v příliš nízké hodnotě vyplyne z dvojice vzorků při poměru převzorkování, jež se objeví symetricky na každé straně vrcholu sinusoidy při naší maximální normalizované frekvenci. Toto je „nejhorší případ“ naměření nižší hodnoty.

Takže pro: poměr převzorkování, n
maximální normalizovanou frekvenci, f_{norm}
vzorkovací frekvenci, f_s

vidíme, že:

vzorkovací perioda při poměru převzorkování je $1/n \cdot f_s$
perioda maximální normalizované frekvence je $1/f_{norm} \cdot f_s$

takže:

maximální chyba naměřené hodnoty (dB) je $20 \cdot \log(\cos(2 \cdot \pi \cdot f_{norm} \cdot f_s / n \cdot f_s \cdot 2))$
(2 ve jmenovateli, protože můžeme špičku minout maximálně o polovinu intervalu převzorkování)

nebo:

maximální chyba naměřené hodnoty (v dB) = $20 \cdot \log(\cos(\pi \cdot f_{norm} / n))$

Podle této rovnice byla vytvořena následující tabulka, která pravděpodobně pokrývá oblast zájmu:

Poměr převzorkování	Max. chyba(dB) $f_{norm} = 0,45$	Max. chyba (dB) $f_{norm} = 0,5$
4	0,554	0,688
8	0,136	0,169
10	0,087	0,108
12	0,060	0,075
14	0,044	0,055
16	0,034	0,042
32	0,008	0,010

Jak implementovat měřič skutečné špičkové úrovně?

Převzorkování se provádí vkládáním vzorků s nulovou hodnotou mezi původní vzorky, aby se vytvořil datový proud při požadovaném poměru převzorkování a potom zapojením „interpolačního“ dolnoproústného filtru, aby se vyloučily frekvence nad požadovanou maximální hodnotou f_{norm} . Když teď aplikujeme algoritmus špičkových vzorků na převzorkovaný signál, budeme mít měřič skutečné špičkové úrovně s maximální požadovanou chybou měření.

Je zajímavé zvážít použití takového zařízení pro převzorkování. Obvykle se používá takový filtr s dolní propustí jako symetrický FIR. Tam, kde se takové filtry používají k filtrování vysoce kvalitního zvukového záznamu, např. u (staromódních) převzorkovacích A/D převodníků nebo převodníků vzorkovací frekvence, je potřeba vypočítat vysoký řád filtru, aby se udrželo velmi nízké zvlnění propustného pásma a aby se dosáhlo extrémního utlumení nepropustného pásma a úzké přechodové pásma. Také musí být zachována velká délka slova, aby se uchoval dynamický rozsah a minimalizovalo zkreslení.

Nicméně protože nebudeme výstup našeho zařízení pro převzorkování poslouchat, ale použijeme ho jen k zobrazení výsledku nebo vytvoření sloupcového grafu, pravděpodobně nemáme stejně striktní požadavky. Dokud zvlnění propustného pásma ve spojení s izolovanými komponenty z nepropustného pásma nezhorší přesnost měření přes naši cílovou hodnotu, jsme spokojeni. Tím se podstatně snižuje požadovaný řád, i když i přesto můžeme chtít dosáhnout úzkého přechodového pásma v závislosti na naší cílové maximální normalizované frekvenci. Podobně může stačit, když délka slova bude zaručovat naši cílovou přesnost až ke spodku sloupcového grafu, pokud nebudeme požadovat přesný číselný výstup při nízkých amplitudách.

Takže může stačit, když pohodlně nainstalujeme vhodné zařízení pro převzorkování (eventuelně pro mnoho kanálů) do běžného nízkonákladového DSP nebo FPGA nebo dokonce ještě do skromnějšího procesoru. Na druhou stranu se již používají měřiče s převzorkováním používající vysoce přesné čipy pro převzorkování určené pro použití v D/A převodnicích. I když se jedná možná o plýtvání křemíkem a energií, tato zařízení nejsou drahá a jsou snadno k dispozici.

Nejjednodušším způsobem jak určit požadovaný řád a koeficienty jednotlivých větví pro parametry konkrétního měřiče je použít rekurzivní program pro navrhování FIR filtrů, jako je Remez nebo Meteor.

Také se může objevit požadavek vyloučit u měřiče špiček vliv případné vstupní stejnosměrné složky, protože měřiče zvuku tradičně měly blokování stejnosměrné složky. Na druhou stranu, pokud se zajímáme o hodnotu skutečných špiček signálu pro účely eliminace přebuzení, musí být obsah stejnosměrné složky zachován a změřen. V případě potřeby lze vyloučení stejnosměrné složky dosáhnout s nízkým výpočetním výkonem zařazením filtru IIR nízkého řádu s horní propustí na vstup měřiče.

Někdy se požaduje měření amplitudy špičkového signálu po zařazení určitého typu váhovacího filtru za účelem zdůraznění účinků určitých částí frekvenčního pásma. Vlastní realizace závisí na povaze konkrétního váhovacího filtru.