

ALGORITMY SPEKTRÁLNÍHO ODEČÍTÁNÍ PRO ZVÝRAZNĚNÍ ŘEČI V BOJOVÝCH PÁSOVÝCH VOZIDLECH

Ing. Jaroslav Hovorka
 MESIT přístroje spol. s r.o.
 Sokolovská 573, 686 01 Uherské Hradiště
 Email: j.hovorka@msp.mesit.cz

Provozní podmínky pásových bojových vozidel jsou charakterizovány značným hlukem. Tento hluk vstupuje do komunikačních systémů instalovaných na vozidle a výrazně zhoršuje srozumitelnost, zejména však kvalitu komunikace. Algoritmy spektrálního odečítání představují jednu z metod číslicového zpracování řeči, kterou lze pro potlačení těchto rušivých signálů použít. Při posuzování vhodnosti jednotlivých algoritmů pro tuto aplikaci je nutno vycházet z konkrétní znalosti charakteristik hluku ve vozidle.

1. KOMUNIKACE V PÁSOVÝCH BOJOVÝCH VOZIDLECH

Osádky pásových bojových vozidel jsou vystaveny působení hluků, které jsou generovány motorem, samotnou jízdou vozidla a zbraňovými systémy, které jsou na vozidle instalovány. Z tohoto důvodu pro zajištění komunikace mezi jednotlivými členy osádky jsou ve vozidlech instalovány tzv. vnitřní komunikační zařízení, neboli interkomy, které umožňují jednak mezi posádkou vést aktivní komunikaci, jednak prostřednictvím kvalitních náhlavních hovorových souprav ochránit sluch členů osádky.

Hluky dosahující úrovní hladin akustického tlaku až 120 dB vstupují přes komunikační mikrofony do interkomů a palubních radiostanic a podstatným způsobem tak snižují zejména kvalitu komunikace. Nutno zdůraznit, že tyto rušivé signály vstupují do komunikačního kanálu i přes použití velmi kvalitních gradientních mikrofonů.

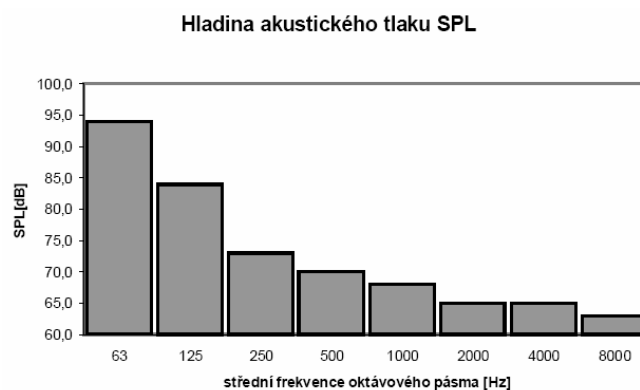
S nástupem digitálních interkomů lze uvažovat o implementaci algoritmů číslicového zpracování signálů pro zvýšení kvality, popř. i srozumitelnosti komunikace.

Tohoto cíle lze dosáhnout implementací algoritmů spektrálního odečítání. Aby bylo dosaženo očekávaného výsledku, je nutno znát základní výhody a nevýhody jednotlivých algoritmů.

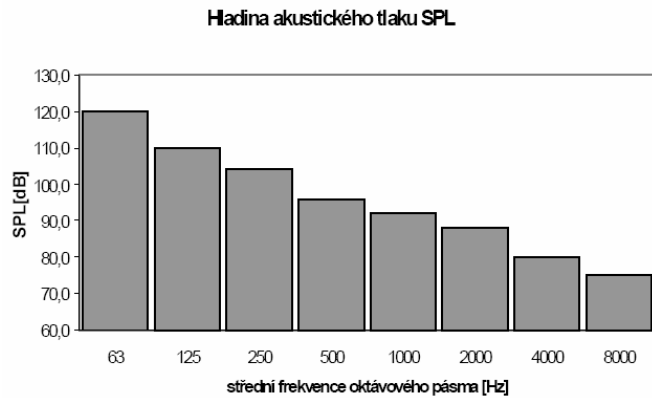
2. HLUK V PÁSOVÝCH VOZIDLECH

Byla provedena řada hlukoměrných měření na bojovém pásovém vozidle. Měření probíhala v mnoha režimech provozu vozidla, tj. u vozidla stojícího v klidu i u vozidla pohybujícího se různými rychlostmi. Vozidlo se pohybovalo ve středně těžkém mírně zvlněném terénu – hlína, kamení, voda. V rámci měření bylo akustickým analyzátozem měřeno hlukové pozadí uvnitř vozidla.

Reprezentativní výsledky hlukoměrných měření jsou uvedeny na obr. 1 a 2. Měření bylo prováděno v oktávových pásmech se středními kmitočty 63, 125, 250, 500, 1000, 2000, 4000, 8000 Hz bez použití korekční váhové křivky [4].



Obr. 1 Hladiny akustického tlaku SPL na pásovém bojovém vozidle stojícím v klidu



Obr. 2 Hladiny akustického tlaku SPL na pásovém bojovém vozidle pohybujícím se rychlostí 40 km/h ve středně těžkém mírně zvlněném terénu

Z naměřených výsledků jednoznačně vyplývá skutečnost, že v pásových bojových vozidlech jsou velmi vysoké hladiny akustického tlaku zejména v oktávě se středními kmitočty 63, 125 a 250 Hz.

Hodnoty hladin akustického tlaku běžně dosahují v oktávě se středním kmitočtem 63 Hz hodnot až 120 dB. I přes potlačení kvalitními gradientními mikrofony tyto signály následně vstupují do komunikačních systémů, kde je již nelze běžnými filtračními metodami dostatečně potlačit – zejména signály s kmitočty od 300 Hz do 3500 Hz.

Na základě provedených měření lze hluk v pásovém bojovém vozidle charakterizovat takto:

- leží ve spektru hovorového signálu,
- dosahuje velmi vysokých úrovní hladin akustického tlaku,
- hladiny akustického tlaku jsou v jednotlivých oktávách různé, což má za následek nerovnoměrné narušení komunikace,
- maximální hladiny akustického tlaku jsou v oktávách se středními kmitočty 63 a 125 Hz,
- v oktávě se středním kmitočtem 63 Hz jsou hladiny akustického tlaku až 120 dB,
- při akceleraci vozidla, popř. nájezdu vozidla na jiný povrch (hlína, kamení) se mění spektrální charakteristiky hluku.

3. SPEKTRÁLNÍ ODEČÍTÁNÍ

Jednou z efektivních metod, kterou lze použít pro zvýraznění řeči v prostředí bojových pásových vozidel jsou algoritmy spektrálního odečítání.

Princip metody vychází z předpokladu, že hovorový signál je zkrácen hlukem vozidla, který má aditivní charakter. Ze znalosti, nebo odhadu hluku lze určit odhad spektra hovorového signálu, a to jako rozdíl spekter signálu zkráceného hlukem okolí a odhadu spektra aditivního hluku.

Odhad spektra hluku je obecně u většiny algoritmů prováděn v době nepřítomnosti hovorového signálu. Z vypočteného spektra se pomocí IFFT určí časový průběh signálu. Základním předpokladem algoritmu je, že hluk je stacionární, nebo se v čase minimálně mění [3]. Tento předpoklad však u bojových vozidel není vždy splněn, a to zejména při akceleraci vozidla, nebo při nájezdu vozidla na jiný povrch. Toto jsou přítomné jevy, které se u bojových vozidel při jejich operačním nasazení vyskytují často.

Algoritmus spektrálního odečítání je obecně popsán vztahem [1]:

$$\left| \widehat{X}(e^{j\omega}) \right|^p = \left| Y(e^{j\omega}) \right|^p - \left| \widehat{D}(e^{j\omega}) \right|^p \quad (1)$$

kde $\left| \widehat{X}(e^{j\omega}) \right|$ je modulové spektrum výstupního (zvýrazněného) signálu na výstupu algoritmu,

$\left| Y(e^{j\omega}) \right|$ je modulové spektrum vstupního signálu s aditivním hlukem okolí (vozidla),

$\left| \widehat{D}(e^{j\omega}) \right|$ je odhad modulového spektra hluku vozidla.

Mocnina p v definičním vztahu (1) nejčastěji nabývá hodnot 1 nebo 2 podle toho zda je algoritmus realizován v oblasti modulového nebo výkonového spektra.

4. HLAVNÍ OMEZENÍ ALGORITMU

Podstatným omezením všech algoritmů spektrálního odečítání jsou tzv. hudební tóny. Příčinou jejich vzniku je především chybný odhad spektra hluku.

Je-li $\widehat{D}(e^{j\omega}) > Y(e^{j\omega})$, pak výsledné spektrum vypočtené dle vztahu (1) je záporné. Spektrální složky skutečného spektra však nemohou nabývat záporných hodnot. Aby toto bylo zajištěno i v praxi, lze pro tento

případ, tedy že $\widehat{D}(e^{j\omega}) > Y(e^{j\omega})$, výsledné spektrum položit rovno 0.

Platí tedy:

$$\begin{aligned} |\widehat{X}(e^{j\omega})| &= |Y(e^{j\omega})| - |\widehat{D}(e^{j\omega})| \text{ pro } |Y(e^{j\omega})| > |\widehat{D}(e^{j\omega})| \\ |\widehat{X}(e^{j\omega})| &= 0 \quad \text{v ostatních případech} \end{aligned} \quad (2)$$

V důsledku této nelineární operace jsou ve výsledném spektru $|\widehat{X}(e^{j\omega})|$ přítomny izolované vrcholky.

Protože se spektrum hluku neustále mění, mění se rovněž i pozice těchto izolovaných vrcholků. V každém zpracovávaném rámci se tyto vrcholky vyskytují na různých kmitočtech. Po provedení IFFT se následně tyto anomálie ve spektru projevují jako tzv. hudební tóny v časové oblasti. Jejich kmitočet se obecně v každém zpracovávaném rámci mění, což je pro hudební tóny typické.

Vznik hudebních tónů je podstatným omezením algoritmu. Velmi citelné jsou pro posluchače tyto tóny obzvláště v segmentech signálu bez přítomnosti řeči. V mnoha praktických případech jsou tyto tóny podstatně pro posluchače rušivější než samotný hluk.

Za hlavní příčiny generace hudebních tónů lze označit nesprávný odhad spektra hluku, nelineární zpracování záporných spektrálních složek a nesprávnou činnost detektoru řeči - VAD detektorů /Voice Activity Detector/.

Dalším obecným problémem při implementaci algoritmu spektrálního odečítání je skutečnost, že k rekonstrukci signálu je použita fáze původního hlukem zkráceného vstupního signálu.

Praktické laboratorní zkoušky prokazují, že pro $\text{SNR} > 5$ dB použití fáze vstupního signálu pro rekonstrukci signálu nemá za následek podstatné zhoršení kvality výstupního signálu [1]. Pro $\text{SNR} < 0$ dB může však použití fáze vstupního signálu způsobit zhoršení kvality výstupního signálu. Nízké hodnoty SNR jsou dosahovány právě v prostředí bojových pásových vozidel, kde hodnoty hluku dosahují vysokých hodnot. Predikce fáze signálu je však poměrně složitý technický úkol, který výrazně zvyšuje složitost algoritmu. Z tohoto důvodu se při praktické implementaci používá fáze vstupního signálu.

5. METODY SPEKTRÁLNÍHO ODEČÍTÁNÍ

Rodina algoritmů spektrálního odečítání obsahuje celou řadu metod pracujících na principu vztahu (1). Tyto algoritmy se od sebe liší zejména složitostí implementace a množstvím hudebních tónů vznikajících v daném prostředí.

Základní algoritmus tak, jak je popsán ve vztahu (1), je pro použití ve velmi hlučném prostředí zcela nevhodný v důsledku tvorby vysokého množství hudebních tónů. Tato nevhodnost je nadále ještě podpořena skutečností, že při skutečné bojové komunikaci jsou jednotlivé rozkazy prokládány poměrně dlouhými pauzami. Během těchto pauz by byla posádka vozidla značně obtěžována těmito tóny.

V prostředí bojových pásových vozidel lze obecně uvažovat s použitím těchto algoritmů:

- Beroutiho algoritmus,
- nelineární spektrální odečítání,
- pásmové spektrální odečítání,
- algoritmus MMSE,
- rozšířené spektrální odečítání.

5.1. BEROUTIHO ALGORITMUS

Beroutiho algoritmus částečně eliminuje vznik hudebních tónů. Spektrální složky výsledného signálu nepoklesnou pod minimální hodnotu $\beta |\widehat{D}(e^{j\omega})|^2$.

Pro algoritmus platí [1]:

$$\begin{aligned} |\widehat{X}(e^{j\omega})|^2 &= |Y(e^{j\omega})|^2 - \alpha |\widehat{D}(e^{j\omega})|^2 \\ &\text{pro } |Y(e^{j\omega})|^2 > (\alpha + \beta) \cdot |\widehat{D}(e^{j\omega})|^2 \\ &= \beta |\widehat{D}(e^{j\omega})|^2 \text{ v ostatních případech} \end{aligned} \quad (3)$$

kde $\alpha \geq 1$ je subtrakční faktor,

$0 < \beta \ll 1$ je parametr spektrálního pozadí.

Při odečtu odhadu spektra hluku od spektra vstupního signálu s hlukem zůstávají ve výsledném spektru širokopásmové a úzkopásmové vrcholky.

Je-li parametr $\alpha > 1$, dochází k poklesu amplitudy širokopásmových vrcholků ve spektru. Ani toto však není pro dostatečnou eliminaci hudebních tónů zdaleka postačující. Dalšího zlepšení je dosaženo vhodnou volbou parametru β .

Parametr α tedy ovlivňuje zkreslení výsledného signálu. Je-li hodnota parametru α příliš velká, dochází ke značnému zkreslení výstupního signálu. Toto výrazným způsobem negativně ovlivňuje srozumitelnost komunikace.

Parametr β ovlivňuje množství zbytkového hluku, který je ve zpracovaném výstupním signálu přítomen. Rovněž má přímý vliv na množství vznikajících hudebních tónů. Jestliže je hodnota β příliš vysoká, pak nebudou vytvářené hudební tóny patrné. V tomto případě však bude posluchač vnímat zbytkový hluk.

Praktické zkušenosti ukazují, že pro dobré potlačení hluku okolí při minimální tvorbě hudebních tónů musí být pro zpracovávané rámce s vysokým SNR voleno malé α .

Parametr α se určuje pro každý zpracovávaný rámec a je funkcí SNR. Vztah pro výpočet α lze nalézt v [1].

Výrazným nedostatkem Beroutiho algoritmu je předpoklad, že hluk okolí ovlivňuje stejnou měrou všechny kmitočtové složky spektra zpracovávaného signálu. V pásových bojových prostředcích je ve skutečnosti kmitočtové spektrum řeči rušeno hlukem velmi nerovnoměrně, viz obr. 1, 2. Velmi výrazné kmitočty v dolním kmitočtovém pásmu / oktávy se středními kmitočty 63 Hz, 125 Hz a 250 Hz /. Parametr α ve vztahu (3) není kmitočtově závislý a používá se stejný pro odečet všech složek spektra. Z tohoto důvodu tento algoritmus není schopen postihnout nerovnoměrné a zejména v čase se měnící rušení řečového signálu.

Pro použití pro zvýraznění řeči v tomto druhu bojových prostředků není z tohoto důvodu vhodný.

5.2. NELINEÁRNÍ SPEKTRÁLNÍ ODEČÍTÁNÍ

Hlavní výhodou algoritmu ve srovnání s Beroutiho algoritmem je schopnost do algoritmu promítnout spektrální charakter hluku.

Cílem je na kmitočtech s malým SNR odečítat větší hodnoty hluku, na kmitočtech s vysokým SNR pak hodnoty menší.

Tohoto je dosaženo prostřednictvím kmitočtově závislého parametru α .

Pro metodu nelineárního spektrálního odečítání platí [1]:

$$\begin{aligned} |\widehat{X}(e^{j\omega})| &= |\overline{Y}(e^{j\omega})| - \alpha(e^{j\omega})N(e^{j\omega}) \\ \text{pro } |\overline{Y}(e^{j\omega})| &> \alpha(e^{j\omega})N(e^{j\omega}) + \beta|\overline{D}(e^{j\omega})| \\ &= \beta|\overline{Y}(e^{j\omega})| \text{ v ostatních případech} \quad (4) \end{aligned}$$

kde $|\overline{Y}(e^{j\omega})|$ a $|\overline{D}(e^{j\omega})|$ je hladký odhad zkreslené řeči a hluku,

$\alpha(e^{j\omega})$ je frekvenčně závislý parametr,

$N(e^{j\omega})$ je nelineární funkce spektra hluku. Pro hladký odhad zkreslené řeči a hluku platí:

$$\begin{aligned} |\overline{Y}_i(e^{j\omega})| &= \mu_y |\overline{Y}_{i-1}(e^{j\omega})| + (1 - \mu_y) |Y_i(e^{j\omega})| \\ |\overline{D}_i(e^{j\omega})| &= \mu_d |\overline{D}_{i-1}(e^{j\omega})| + (1 - \mu_d) |\widehat{D}_i(e^{j\omega})| \quad (5) \end{aligned}$$

kde $|Y_i(e^{j\omega})|$ je modulové spektrum řeči v i -tém rámci,

$|\widehat{D}_i(e^{j\omega})|$ je odhad modulového spektra hluku v i -tém rámci,

μ_y je konstanta, $0.1 \leq \mu_y \leq 0.5$,

μ_d je konstanta, $0.5 \leq \mu_d \leq 0.9$.

$N(e^{j\omega})$ představuje maximální hodnoty jednotlivých kmitočtových složek modulového spektra odhadu hluku $|\widehat{D}_i(e^{j\omega})|$, které se dle [1] určují z posledních 40 rámců.

Pro výpočet parametru $\alpha(e^{j\omega})$ lze uvažovat různé funkce SNR. Vhodné funkce pro výpočet $\alpha(e^{j\omega})$ lze nalézt v [1], [2], [3].

5.3. PÁSMOVÉ SPEKTRÁLNÍ ODEČÍTÁNÍ

Algoritmus pásmového spektrálního odečítání na rozdíl od algoritmu nelineárního spektrálního odečítání uvažuje konstantní subtrahční činitele pro jednotlivá kmitočtová subpásma, a ne pro jednotlivé diskretní kmitočty.

Vychází ze skutečnosti, že SNR jednotlivých diskretních kmitočtů se mohou mezi jednotlivými zpracovávanými rámci výrazně měnit.

Právě tyto prudké změny jsou rovněž příčinou vzniku hudebních tónů. Na rozdíl od toho se SNR jednotlivých kmitočtových subpásem tak výrazně nemění, a proto lze předpokládat dosažení menšího počtu hudebních tónů ve výsledném signálu.

Princip algoritmu spočívá v rozdělení zpracovávaného signálu do jednotlivých kmitočtových subpásem. V každém z těchto subpásem je následně realizován algoritmus spektrálního odečítání.

Odhad zvýrazněné řeči v i -tém kmitočtovém pásmu je [1]:

$$\left| \widehat{X}_i(e^{j\omega_k}) \right|^2 = \left| \overline{Y}_i(e^{j\omega_k}) \right|^2 - \alpha_i \cdot \delta_i \cdot \left| \widehat{D}_i(e^{j\omega_k}) \right|^2 \quad (6)$$

$$b_i \leq \omega_k \leq e_i$$

kde ω_k jsou diskrétní frekvence,

$\left| \widehat{D}_i(e^{j\omega_k}) \right|^2$ je odhad výkonového spektra hluku,

b_i , e_i jsou počáteční a koncový diskrétní kmitočty i -tého kmitočtového subpásma,

α_i je subtrakční činitel i -tého subpásma,

δ_i přídatný subtrakční faktor pro jednotlivá kmitočtová subpásma.

Záporné spektrální složky, které by mohly vzniknout v důsledku vztahu (6) jsou eliminovány vztahem (7):

$$\begin{aligned} \left| \widehat{X}_i(e^{j\omega_k}) \right|^2 &= \left| \widehat{X}_i(e^{j\omega_k}) \right|^2 \text{ pro } \left| \widehat{X}_i(e^{j\omega_k}) \right|^2 > \beta \cdot \left| \overline{Y}_i(e^{j\omega_k}) \right|^2 \\ &= \beta \cdot \left| \overline{Y}_i(e^{j\omega_k}) \right|^2 \text{ v ostatních případech} \quad (7) \end{aligned}$$

Hodnota parametru spektrálního pozadí β je 0,002 [1].

Zavedením malého množství spektra vstupního signálu obsahujícího řeč s hlukem do výsledného signálu dle vztahu (8) je dosaženo přídatného maskování hudebních tónů:

$$\left| \widetilde{X}_i(e^{j\omega_k}) \right|^2 = \left| \widehat{X}_i(e^{j\omega_k}) \right|^2 + 0.05 \cdot \left| \overline{Y}_i(e^{j\omega_k}) \right|^2 \quad (8)$$

kde $\left| \widetilde{X}_i(e^{j\omega_k}) \right|^2$ je výsledné výkonové spektrum.

Pro každé kmitočtové subpásma je na základě SNR tohoto subpásma určen vhodný subtrakční činitel α_i [5].

SNR je vypočten pro každé kmitočtové subpásma podle vztahu:

$$SNR(dB) = 10 \cdot \log_{10} \left(\frac{\sum_{\omega_k=b_i}^{e_i} \left| \overline{Y}_i(e^{j\omega_k}) \right|^2}{\sum_{\omega_k=b_i}^{e_i} \left| \widehat{D}_i(e^{j\omega_k}) \right|^2} \right) \quad (9)$$

Hodnoty parametru δ_i se stanovují empiricky. Nastavení tohoto parametru lze provést podle [5], viz (10):

$$\begin{aligned} \delta_i &= 1 \text{ pro } f_i < 1 \text{ kHz} \\ &= 2.5 \text{ pro } 1 \text{ kHz} < f_i \leq \frac{F_s}{2} - 2 \text{ kHz} \\ &= 1.5 \text{ pro } f_i > \frac{F_s}{2} - 2 \text{ kHz} \quad (10) \end{aligned}$$

kde f_i je nejvyšší kmitočty i -tého subpásma

F_s je vzorkovací kmitočty.

Smyslem je, aby na nízkých kmitočtech každého subpásma byly používány nižší hodnoty parametru δ_i . Důvodem je, že na nižších kmitočtech je soustředěna většina energie řečového signálu.

Předností algoritmu je závislost subtrakčního parametru na SNR daného kmitočtového subpásma, nikoli na SNR konkrétního diskrétního kmitočtu. Toto vede k menšímu výskytu hudebních tónů ve srovnání s dříve uvedenými algoritmy.

5.4 ALGORITMUS MMSE

Základním nedostatkem dříve uvedených algoritmů je skutečnost, že subtrakční parametry α a β nejsou stanoveny optimálně. Toto následně logicky vede k chybám ve výsledném spektru a po transformaci do časové oblasti k tvorbě hudebních tónů.

Podstatným přínosem algoritmu MMSE (Minimum Mean Square Estimator) ve srovnání s dříve uvedenými algoritmy je určení subtrakčních parametrů tak, aby se spektrum výsledného signálu po realizaci algoritmu co nejvíce shodovalo se skutečným spektrem řeči bez přítomnosti hluku.

Algoritmus spektrálního odečítání lze obecně vyjádřit jako [1]:

$$\left| \widehat{X}(e^{j\omega}) \right|^p = \gamma_p(e^{j\omega}) \left| Y(e^{j\omega}) \right|^p - \alpha_p(e^{j\omega}) \left| \widehat{D}(e^{j\omega}) \right|^p \quad (11)$$

kde $\gamma_p(e^{j\omega})$ a $\alpha_p(e^{j\omega})$ jsou subtrakční parametry,
 p je mocnina,

$\widehat{D}(e^{j\omega})$ je spektrum hluku vypočtené ze segmentů signálu bez přítomnosti řeči.

Parametry $\gamma_p(e^{j\omega})$ a $\alpha_p(e^{j\omega})$ jsou stanoveny na základě minimalizace střední hodnoty chybového spektra $e_p(e^{j\omega})$:

$$e_p(e^{j\omega}) = |X_p(e^{j\omega})|^p - |\widehat{X}(e^{j\omega})|^p \quad (12)$$

kde $|X_p(e^{j\omega})|^p$ je spektrum čisté řeči bez přítomnosti hluku.

Parametry $\gamma_p(e^{j\omega})$ a $\alpha_p(e^{j\omega})$ jsou stanoveny tak, aby bylo dosaženo minimální střední hodnoty chybového spektra – $\min E\{e_p(e^{j\omega})\}^2$.

Potom pro optimální subtrakční parametry platí:

$$\alpha_p(e^{j\omega}) = \frac{\xi^p(e^{j\omega})}{1 + \xi^p(e^{j\omega})}, \quad (13)$$

$$\gamma_p(e^{j\omega}) = \frac{\xi^p(e^{j\omega})}{1 + \xi^p(e^{j\omega})} \{1 - \xi^{-p/2}(e^{j\omega})\}, \quad (14)$$

kde

$$\xi(e^{j\omega}) = \frac{E[|X_p(e^{j\omega})|^2]}{E[|D(e^{j\omega})|^2]}. \quad (15)$$

Po dosazení do definičního vztahu algoritmu spektrálního odečítání (11) platí:

$$\begin{aligned} \widehat{X}(e^{j\omega}) &= \\ &= \left\{ \frac{\xi^p(e^{j\omega})}{1 + \xi^p(e^{j\omega})} [|Y(e^{j\omega})|^p - (1 - \xi^{-p/2}(e^{j\omega})) |\widehat{D}(e^{j\omega})|^p] \right\}^{1/p} \end{aligned} \quad (16)$$

Člen $\xi(e^{j\omega})$ nemůže být nikdy přesně vypočten, protože v reálných situacích není znám čistý signál, tj. signál bez přítomnosti hluku.

Výraz $\xi(e^{j\omega})$ lze z tohoto důvodu pro účely algoritmu dle [1] nahradit výrazem:

$$\xi(e^{j\omega}) \approx (1 - \eta) \max\left(\frac{|Y(e^{j\omega})|^2}{|\widehat{D}(e^{j\omega})|^2} - 1, 0\right) + \eta \frac{|\widehat{X}_{prev}(e^{j\omega})|^2}{|\widehat{D}(e^{j\omega})|^2} \quad (17)$$

kde η je konstanta 0,96,

$|\widehat{X}_{prev}(e^{j\omega})|$ je zvýrazněné spektrum vypočtené v předchozím rámci.

5.5 ROZŠÍŘENÉ SPEKTRÁLNÍ ODEČÍTÁNÍ

Podstatnou výhodou tohoto algoritmu ve srovnání se všemi dříve uvedenými algoritmy je skutečnost, že není nutné při realizaci algoritmu implementovat obvod detektoru řeči (VAD). Odhad spektra hluku je kontinuálně aktualizován, a to i během segmentů s přítomností řeči [6].

Algoritmus používá adaptivní Wienerovskou filtraci pro odhad spektra hluku. Toto spektrum je následně odečteno od vstupního zpracovávaného signálu.

Pro adaptivní Wienerův filtr platí [1],[6]:

$$H_j(e^{j\omega}) = \left(\frac{|\overline{D}_{j-1}(e^{j\omega})|^2}{|\overline{X}_{j-1}(e^{j\omega})|^2 + |\overline{D}_{j-1}(e^{j\omega})|^2} \right)^{1/2}, \quad (18)$$

kde $|\overline{D}_{j-1}(e^{j\omega})|^2$ je hladký odhad výkonového spektra hluku v rámci $j-1$,

$|\overline{X}_{j-1}(e^{j\omega})|^2$ je vyhlazené výkonové spektrum výstupního zvýrazněného signálu v rámci $j-1$.

Hladký odhad hluku je potom [1]:

$$|\overline{D}_j(e^{j\omega})| = c |\overline{D}_{j-1}(e^{j\omega})| + (1 - c) |\widehat{D}_j(e^{j\omega})| \quad (19)$$

kde c je konstanta, $0 < c < 1$,

$|\widehat{D}_j(e^{j\omega})|$ je aktuální odhad spektra hluku určený na základě Wienerovy filtrace spektra vstupního signálu:

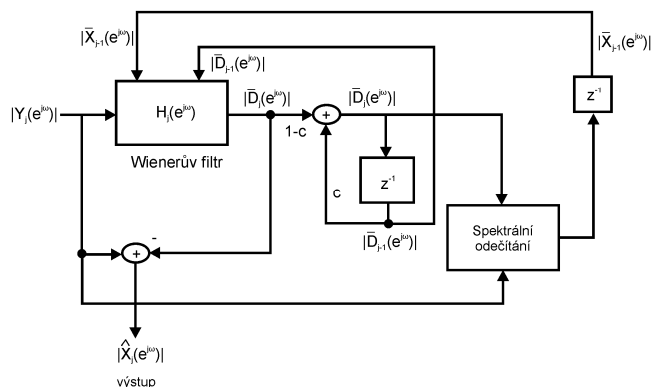
$$|\widehat{D}_j(e^{j\omega})| = H_j(e^{j\omega})|Y_j(e^{j\omega})| \quad (20)$$

Výsledné spektrum je :

$$|\widehat{X}_j(e^{j\omega})| = |Y_j(e^{j\omega})| - |\widehat{D}_j(e^{j\omega})| \quad (21)$$

Předností algoritmu je zejména skutečnost, že umožňuje implementaci i v prostředí s nestacionárním hlukem, což je právě prostředí bojových vozidel, kde hluk se mění v čase – např. při zrychlení vozidla, nebo nájezd na jiný terén.

Parametr c má vliv na funkci a výsledek algoritmu. Optimální chování algoritmu je dle [1] dosahováno při $0,8 < c < 0,99$.



Obr. 3 Rozšířené spektrální odečítání [1].

6. ZÁVĚR

Během minulých let byla vyvinuta řada algoritmů spektrálního odečítání. Zde byly představeny algoritmy, které se jeví jako vhodné pro použití v interkomech vojenských bojových vozidel. Při výběru vhodného algoritmu pro toto prostředí musí být brány v úvahu specifické podmínky v nichž vozidlo operuje. Těmito podmínkami se rozumí především spektrální charakter hluku včetně jeho změn při změně činnosti a velmi vysoké hodnoty hladin akustického tlaku reálně dosahující až 120 dB.

Základní tvar spektrálního odečítání není pro tento účel vhodný z důvodu přítomnosti velkého množství hudebních tónů ve výstupním signálu. Vzhledem k charakteru komunikace, kdy krátké povely jsou prokládány poměrně dlouhými pauzami by hudební

tóny velmi nepříznivě zvyšovaly pracovní zátěž posádky a působily velmi rušivě. Kromě tohoto základního algoritmu byla prezentována řada algoritmů spektrálního odečítání lišících se zejména množstvím hudebních tónů ve výstupním signálu. Tyto algoritmy jsou pro daný účel podstatně vhodnější.

Výsledky mnoha provedených studií prokazují, že algoritmy spektrálního odečítání zlepšují zejména kvalitu řeči. Druhým parametrem je srozumitelnost, ta ale vždy zvýšena touto metodou není. V některých případech může být dokonce v důsledku volby nevhodného algoritmu, nebo nevhodně zvolených parametrů algoritmu srozumitelnost snížena.

7. LITERATURA

- [1] LOIZOU C., Philipos, *Speech enhancement Theory and Practice*, CRC Press, 2007, s. 97-141, ISBN 978 0 8493 5032 0.
- [2] LOCKWOOD, P., BOUDY, J. Experiments with a Non-linear Spectral Subtractor (NSS), Hidden Markov Models and the projections, for robust recognition in cars, *Speech Commun.*, 1992, s. 215-228.
- [3] VASEGHI V., Saeed, *Advanced Digital Signal Processing and Noise Reduction*, John Wiley&Sons, LTD, 2000, s. 333-354, ISBN 0 471 62692 9.
- [4] SMETANA C. a kolektiv, *Hluk a vibrace, Měření a hodnocení*, Sdělovací technika, Praha 1998, ISBN 80 901936 2 5, s. 49-60.
- [5] Kamath, S. , Loizou, P., A multi-band spectral subtraction method for enhancing speech corrupted by colored noise, *Proc. IEEE Int. Conf. Acoust. Speech Signal Process*, 2002.
- [6] SOVKA, P., POLLACK, P., KYBIC, J., Extended spectral subtraction, *Proc. Eur. Conf. Signal Process. Commun.*, s. 963-966., 1996.