

REDUKCE HLUKU PRO MOBILNÍ TELEFON SE DVĚMA MIKROFONY

Autor: Bc. David Botka, Vedoucí: Doc. Ing. Zbyněk Koldovský, Ph.D.
Fakulta Mechatroniky, informatiky a mezioborových studií, Ústav informačních technologií a elektroniky

Abstract

- A novel method of noise reduction for dual-microphone mobile phones is proposed.
- The method is based on a set (bank) of target-cancellation filters derived in a noise-free situation for different possible positions of the phone with respect to the speaker mouth.
- The set of the cancellation filters is used to accurately estimate the noise of the environment, which is then subtracted from the recorded signal via standard Wiener filter.
- Experiments with recorded data show a good performance and low complexity of the system.

Úvod

- Motivace
 - Obrovské množství mobilních telefonů po celém světě
 - Odstraňování hluku – stále aktuální problematika
 - Dva vstupní kanály – možnost dosažení kvalitnějších výsledků, než při jednorázovém zpracování

Formulace

- Dvoukanálové nahrávání s fixní pozicí zdroje:

$$\begin{aligned}x_L(n) &= \{h_L * s\}(n) + y_L(n) \\x_R(n) &= \{h_R * s\}(n) + y_R(n)\end{aligned}\quad (1)$$

- n ... časový index ($n = 1 \dots N$)
- $*$... konvoluce
- $x_L(n), x_R(n)$... signály z levého a pravého mikrofonu
- $s(n)$... target signál
- $y_L(n), y_R(n)$... signály obsahující hluk
- $h_L(n), h_R(n)$... impulsní odezvy mikrofonů

Target Cancellation filtr (CF)

- Ideální systém pro úplné potlačení target signálu s – dva filtry g_L a g_R

$$g_L * h_L * s = g_R * h_R * s \quad (2)$$

- Výstup ideálního CF

$$z = g_L * x_L - g_R * x_R = \dots = g_L * y_L - g_R * y_R \quad (3)$$

Výpočet CF1

- Běžná konstrukce $\rightarrow g_R(n) = \delta(n - D)$
- Výpočet g_L

$$\hat{g}_L = \arg \min_{g_L} \|g_L^T X_L - g_R^T X_R\|_2^2 \quad (4)$$

- X_i ... Toeplitzovská matice signálu x_i , $i \in \{L, R\}$
 - Toeplitzovská matice
 - První řádek ... $[x_i(1), \dots, x_i(N), 0, \dots, 0]$
 - První sloupec ... $[x_i(1), \dots, 0]$
 - Rozměry ... $L \times (N + L - 1)$
 - L ... délka filtrů g_L, g_R

Výpočet CF2

- Nezarušený target signál a referenční hluk
- Výpočet g_L, g_R

$$\hat{g}_L, \hat{g}_R = \arg \min \|g_L^T X_L - g_R^T X_R\|_2^2 + \varepsilon \|g_L^T Y_L - g_R^T Y_R - y\|_2^2 \quad (5)$$

- ε ... Kladné číslo
- Y_i ... Toeplitzovská matice signálu y_i , $i \in \{L, R\}$ (referenční hluk)
- y ... Vektor hluku, který chceme pozorovat na výstupu CF (signál z)
- Příklad vektoru y – zpožděný signál y_L o D vzorků

System redukce hluku

Banka CF pro mobilní telefony

- Běžné použití mobilních telefonů – mluvčí je v bezprostřední blízkosti mikrofonů
- Příprava banky CF pro nejpravděpodobnější pozice telefonu vůči mluvčímu
- Pořízení nahrávky mluvčího v tichém prostředí pro každou pozici (pozorování – CF závisí především na konstrukci telefonu, mluvčím a pozici. Závislost na ostatních okolních objektech je zanedbatelná.)
- Transformace vypočtených CF z časové oblasti do frekvenční oblasti a uložení do paměti telefonu pro následné použití

Postup redukce hluku

- Krátkodobá Fourierova transformace
- Každý blok zfiltrován všemi filtry banky
- Jako odhad hluku vybrán výstup filtru s nejvyšším potlačením
- Odečtení odhadnutého hluku od původního signálu pomocí Wienerova filtru

Experimenty

Model telefonu a umělá hlava

- Simulace mluvčích přehráváním nahrávek z databáze TIMIT
- Různé signály s hlukem (řeč, ruch ulice a ruch restaurace)
- Banky filtrů (pro CF1 a CF2): dvě umístění mikrofonů, délka filtrů 1000, délka nahrávek 4 s, referenční hluk pro CF2 – mužský hlas, $\varepsilon=0.5$
- Testování: přehrávání z umělé hlavy, pohyb modelu telefonu
- Vyhodnocení:
 - SNR (Signal-to-Noise Ratio)
 - SDR (Signal-to-Distortion Ratio)
 - Perceptuální kritéria softwaru PEASS
 - TPS (Target-related Perceptual Score)
 - OPS (Overall Perceptual Score)

Vyhodnocení

- Rozmístění mikrofonů: vpředu a vzadu
- Výsledky uvedeny v dB
- Vstupní SNR: -10 dB

	CF1	CF2	PLD	CF1	CF2	PLD	CF1	CF2	PLD
SNR	-4.2	-2.8	-0.2	-2.5	-6.9	6.2	-5.0	-7.3	-7.2
SDR	7.7	8.1	3.2	8.5	8.1	4.8	6.5	7.2	5.4
TPS	53.9	66.4	4.1	52.0	56.7	5.8	47.0	58.4	12.1
OPS	23.2	22.9	23.1	19.6	20.5	25.1	22.1	21.5	19.5

- Vstupní SNR: +10 dB

	CF1	CF2	PLD	CF1	CF2	PLD	CF1	CF2	PLD
SNR	18.0	18.7	19.2	19.6	15.3	21.5	17.6	14.5	14.1
SDR	15.6	11.8	6.4	16.7	12.0	6.7	14.1	11.4	6.6
TPS	70.1	72.3	16.6	58.7	67.6	34.2	74.3	79.0	37.8
OPS	14.3	17.9	16.0	6.9	9.3	11.5	12.4	16.0	21.8

- Zlepšení SNR až o 8 dB při dostatečném SDR
- Lepší výsledky v porovnání s PLD
- Horší kvalita při zpracování nahrávek skutečných osob \rightarrow vhodnější naměřit banku CF na skutečné osobě

Závěr

- Experimentální ověření metody
- Možná různá umístění mikrofonů
- Nízká výpočetní náročnost
- Demonstrační aplikace
 - Implementace v jazyce JAVA
 - Online i offline zpracování
 - Při online zpracování možnost online přehrávání původního signálu, signálu s potlačeným hlasem (výstup banky CF) a signálu s potlačeným hlukem
 - Snadné přizpůsobení banky CF konkrétnímu uživateli
- Možná pokračování práce
 - Pokročilejší metoda pro odečítání odhadnutého hluku
 - Adaptivní modifikace banky CF

Reference

- M. Jeub, C. Herglotz, C. M. Nelke, C. Beaugeant and P. Vary, "Noise Reduction for Dual-Microphone Mobile Phones Exploiting Power Level Differences," *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pp. 1693–1696, Kyoto, Japan, Mar. 2012.
- N. Yousefian and P. C. Loizou, "A Dual-Microphone Speech Enhancement Algorithm Based on the Coherence Function," *IEEE Trans. on Audio, Speech and Language Processing*, vol. 20, no. 2, Feb. 2012.
- S. Gannot, D. Burshtein, and E. Weinstein, "Signal enhancement using beamforming and nonstationarity with applications to speech," *IEEE Trans. on Signal Processing*, vol. 49, no. 8, pp. 1614–1626, Aug. 2001.