

SIMULACE ŘEČOVÉ STRATEGIE ACE A ACE_v V REÁLNÉM ČASE

Martin Vondrášek

Katedra teorie obvodů, Fakulta elektrotechnická ČVUT Praha

Abstrakt

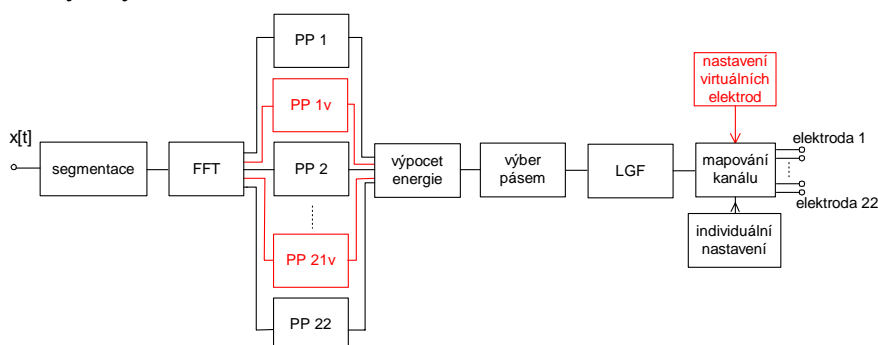
Kochleární implantát [1] je elektronické zařízení nahrazující sluch u pacienta s těžkou sluchovou vadou. Nefunkční vláskové buňky jsou nahrazeny přímou stimulací sluchového nervu proudovými pulsy. Algoritmus převodu řečového signálu na informace o stimulačních pulsech je nazýván řečová strategie. Zpracování řečového signálu probíhá v externí části kochleárního implantátu, v řečovém procesoru. Pro vývoj nových algoritmů a jejich optimalizaci je používána simulace v programovém prostředí Matlab. Komunikace Matlabu s kochleárním implantátem zajišťuje Nucleus Matlab Toolbox. Nucleus Matlab Toolbox také umožňuje implementovat strategie v současnosti používané a jejich modifikace.

1 Nucleus Matlab Toolbox

Nucleus Matlab Toolbox [2, 4] vyvinula australská firma Cochlear l. t. d. původně pro svou vlastní potřebu. V současnosti tento nástroj poskytuje i výzkumným centřům zabývajícím se vývojem kochleárních implantátů. Nucleus Matlab Toolbox je soubor funkcí umožňující vzájemnou komunikaci řečového procesoru, kochleárního implantátu a standardního počítače. Dále obsahuje soubor funkcí realizujících jednotlivé bloky řečových strategií SPEAK, ACE a CIS [3] v současnosti používaných v řečových procesorech pacientů.

2 Řečová strategie ACE a ACE_v

Řečová strategie ACE (Advanced Combination Encoder) je nejnovější a zároveň nejpoužívanější algoritmus zpracování řeči, používaný v kochleárních implantátech Nucleus 24 a Nucleus Freedom [3]. Tato strategie umožňuje pacientům dobrou srozumitelnost řeči. Pro rozpoznání mluvčího a pro poslech hudby je nevyhovující zejména z důvodu malého frekvenčního rozlišení. Kochleární implantáty Nucleus 24 a Nucleus Freedom používají 22 elektrod pro stimulaci. Sluchový vjem pacienta je tak složen pouze z 22 různých tónů. Použití virtuálních elektrod je jedinou možností, jak zvýšit počet různých tónů, který může pacient slyšet v případě stávajícího kochleárního implantátu. Podstatné zvýšení počtu elektrod kochleárního implantátu naráží na technologické bariéry. Výměna kochleárního implantátu za novější znamená také riziko pro pacienta a v neposlední řadě je zde i otázka vysokých nákladů.



Obr. 1: Řečová strategie ACE s virtuálními elektrodami.

Algoritmus řečové strategie ACE je znázorněn na obrázku 1 (černě). Řeč je v bloku segmentace rozdělena na segmenty a na každý segment je aplikována Fourierova transformace. Následuje filtrace ve spektru bankou 22 filtrů s nelineární šířkou propustných pásem. V následném bloku je určena

velikost energie na výstupu z každého filtru. V bloku výběru pásem je vybráno několik pásem s maximální energií. Tyto vybrané energie reprezentují podstatné informace o spektru řečového signálu a jsou použity pro stimulaci. Blok LGF zohledňuje logaritmické vnímání hlasitosti zdravého ucha. Posledním krokem je přepočítání velikostí vybraných energií na konkrétní hodnoty stimulačních proudů s ohledem na individuální potřeby pacienta s kochleárním implantátem.

Řečová strategie ACEv (obr. 1) byla upravena použitím virtuálních elektrod (červeně). Počet filtrů v bance filtrů byl zvýšen z 22 na 43. Zároveň byla zmenšena šířka propustných pásem všech použitých filtrů na polovinu. Dále bylo nutno doplnit algoritmus o individuální nastavení parametrů stimulace do virtuálních elektrod. V souvislosti s použitím virtuálních elektrod [5] musel být celý algoritmus přepracován. Zároveň musel být navržen nový algoritmus výběru podstatných informací ve spektru, upravena banka filtrů a změněna segmentace. Dále musely být odstraněny pro pacienta potenciálně nebezpečné stimulační proudové pulsy s malými prodlevami, které značně zvyšovaly hlasitost. Pro zmenšení prosakování ve spektru byla upravena segmentace. Navržený algoritmus byl ověřen na pacientech s kochleárními implantáty a simulacemi na slyšících dobrovolnících [6].

3 Zpětná rekonstrukce řeči z proudových pulsů

Z výstupu řečové strategie lze zpětně rekonstruovat řeč [6]. Takto rekonstruovaná řeč může být použita jako náhled na to, jak by mohl vnímat řeč ideální pacient s kochleárním implantátem. Rekonstruovanou řeč lze použít i pro simulace na slyšících dobrovolnících. Nejčastěji používaná metoda rekonstrukce řeči z proudových pulsů je syntéza pomocí součtu sinusových signálů. Rekonstruovaný signál je složen z několika sinusoid s různými kmitočty podle vzorce

$$s(t) = \sum_{k=1}^N A_k(t) \sin(2\pi * f_k * t),$$

kde $A_k(t)$ je amplituda obálky k -tého pásma. Tato amplituda je nenulová jen v okamžicích, kdy je v příslušném k -tém pásmu vybráno maximum a k je počet pásem analyzující banky filtrů. Kmitočty f_k jsou střední kmitočty propustných pásem analyzující banky filtrů. Počet nenulových amplitud $A_k(t)$ je v každém časovém okamžiku stejný, jako je počet přenášených pásem. Fáze původního signálu není pro rekonstrukci použita.

4 Simulace řečových strategií ACE a ACEv v reálném čase

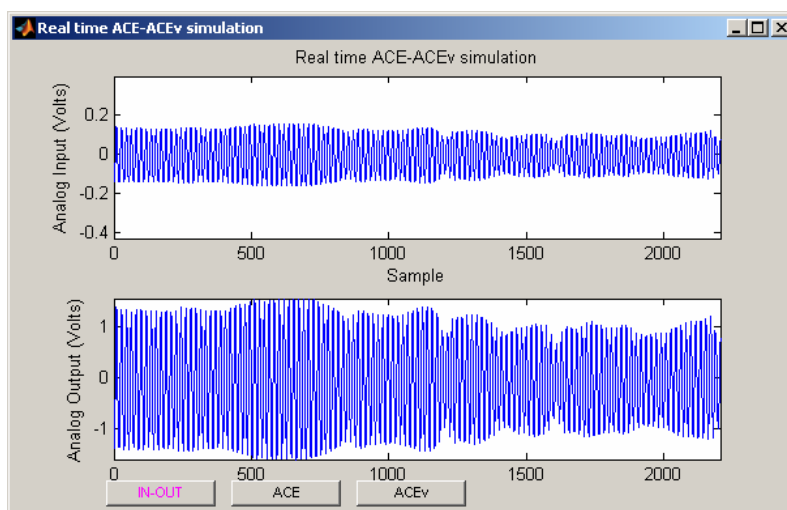
Pro hlubší porovnání řečových strategií ACE a ACEv na slyšících dobrovolnících nebo na pacientech s kochleárními implantáty je nutné, aby testované osoby mohly „poslouchat“ delší čas. Dosud probíhalo testování Off-line. Z předem nahrané databáze slov, krátkých vět nebo ukázek mluveného slova a hudby byla vybrána jedna ukázka zpracována řečovou strategií a poté odstимуlována pacientovi nebo přehrána slyšícímu dobrovolníkovi. Délka ukázky byla limitována dobou zpracování, a tím i čekáním testované osoby, která se blížila délce nahrávky. Ideálním způsobem testování je možnost zpracování v reálném čase. Vstupem řečové strategie by byl jeden nebo více mikrofonů, výstup z televize, rádia nebo MP3 přehrávače. Výstupem pak sluchátka pro případ simulací na slyšících dobrovolnících nebo stimulace do kochleárního implantátu pacienta.

Pro načítání dat do Matlabu a pro přehrávání rekonstruovaných signálů v reálném čase byl použit Data Acquisition Toolbox [5]. Jako interface propojení vstupu s Matlabem byla využita zvuková karta na testovací počítači. Výhodou zvukové karty je, že nemusí být použit žádný další hardware. Testovány byly celkem tři zvukové karty:

1. integrovaná zvuková karta Realtek AC97' Audio
2. zvuková karta Creative Labs Sound Blaster Audigy SE
3. PCMCA zvuková karta Digigram VX Pocket 440

Pro porovnání strategií ACE a ACEv byl u všech testovaných zvukových karet používán jen jeden vstup. V případě zvukové karty VX Pocket 440 je možné použít až 4 nezávislé vstupy a v budoucnu do algoritmu zpracování řeči přidat blok směrového slyšení nebo odstranění šumu. Design testovacího

toolu je uveden na obrázku 1. Umožňuje poslech neupraveného vstupního signálu, signálu vzniklého rekonstrukcí z výstupu strategií ACE a ACEv.



Obr. 1: Design testovacího toolu.

Nevýhod při použití zvukové karty pro vstup a výstup v reálném čase je několik:

Nutnost načítání dat po blocích: data, načítaná nebo odesílaná na zvukovou kartu je nutno zpracovávat po blocích s časovou délkou alespoň 0.1 s z důvodu pomalejší komunikace Matlabu a zvukové karty. Algoritmus zpracování řeči naproti tomu pracuje se segmenty délky 8 ms.

Občasné ztráty dat: v případě načítání dat ze zvukové karty dochází k občasným ztrátám několika vzorků vstupního signálu, jednotlivé načtené úseky pak na sebe nenavazují. Výsledkem těchto ztrát je nepříjemné „praskání“ v průběhu poslechu.

Volba vzorkovacího kmitočtu: vzorkovací kmitočet načítaných dat nelze volit libovolně. Je závislý na konkrétní zvukové kartě.

Poslední nevýhoda nastává jen pro zvukovou kartu Digigram VX Pocket 440. Data Acquisition Toolbox přistupuje k této čtyřvstupové zvukové kartě jako ke dvěma nezávislým stereofonním zvukovým kartám. Načítání dat z kanálu 1-2 a 3-4 probíhá odděleně, a může tak dojít časovému zpoždění.

5 Závěr

Pro porovnání řečových strategií v reálném čase ACE a ACEv byl vytvořen testovací program. Pro načítání dat do Matlabu a pro přehrávání rekonstruovaného signálu byl použit Data Acquisition Toolbox a zvuková karta. Výhodou zvukové karty je jednoduchost a není potřeba žádný další hardware. Použití zvukové karty s sebou nese nedostatky, které ovšem pro první testování v reálném čase nejsou kritické. Zpoždění mezi vstupním a výstupním signálem je minimálně 0.2 s z důvodu načítání a přehrávání dat po blocích délky 0.1 s. Celkové zpoždění včetně zpoždění způsobeného algoritmem zpracování řeči je kolem 0.5 s. V případě, že testovaná osoba má vizuální kontakt s mluvčím, je toto zpoždění již nepříjemné. Testování ale obvykle probíhá bez vizuálního kontaktu, aby výsledky testů nemohly být ovlivněny odezíráním. V takovém případě zpoždění nevádí.

Poděkování

Tato práce vznikla za podpory grantu Grantové agentury České republiky „Předzpracování řeči pro kochleární implantáty“ č. 102/09/P180 a výzkumného záměru „Transdisciplinární výzkum v oblasti biomedicínského inženýrství II“ č. MSM6840770012.

Reference

- [1] CLARK, Graeme. Cochlear implants, fundamentals and applications. New York: Springer NY, 2003. 830 p. ISBN 0-387-95583-6.
- [2] Swanson, Brett. Nucleus[®] Reference manual, Cochlear Ltd. Australia, 2001.
- [3] Nucleus[®] Technical Reference Manual, Cochlear ltd, 1999.
- [4] Nucleus[®] Implant Communicator software, Cochlear Ltd. Australia, 2002.
- [5] Duane C. Hanselman, Bruce Littlefield.: Mastering Matlab, 2006, ISBN 0-13-019468-9.
- [6] Vondrášek, M: Předzpracování řeči pro pacienty s kochleárními implantáty. Distertační práce. CVUT-FEL, 2008.

Ing. Martin Vondrášek, Ph.D.
Katedra teore obvodů, ČVUT-FEL
Technická 2
Praha 6
166 27
vondram3@fel.cvut.cz