



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY



**FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH
TECHNOLOGIÍ**

ÚSTAV BIOMEDICÍNSKÉHO INŽENÝRSTVÍ

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION
DEPARTMENT OF BIOMEDICAL ENGINEERING

ZPRACOVÁNÍ ZVUKU PRO KOCHLEÁRNÍ IMPLANTÁT

SOUND PROCESSING FOR COCHLEAR IMPLANT

BAKALÁŘSKÁ PRÁCE

BACHELOR'S THESIS

AUTOR PRÁCE

AUTHOR

TOMÁŠ KRATOCHVÍL

VEDOUCÍ PRÁCE

SUPERVISOR

Ing. VRATISLAV HARABIŠ

BRNO 2014



VYSOKÉ UČENÍ
TECHNICKÉ V BRNĚ

Fakulta elektrotechniky
a komunikačních technologií

Ústav biomedicínského inženýrství

Bakalářská práce

bakalářský studijní obor

Biomedicínská technika a bioinformatika

Student: Tomáš Kratochvíl

ID: 147495

Ročník: 3

Akademický rok: 2013/2014

NÁZEV TÉMATU:

Zpracování zvuku pro kochleární implantát

POKYNY PRO VYPRACOVÁNÍ:

1) Prostudujte základní anatomii a fyziologii lidského sluchu a zaměřte se na jeho různá poškození. 2) Popište principy kochleárního implantátu. 3) Prostudujte možnosti programového prostředí LabVIEW pro zpracování hlasového signálu, zejména pro analýzu řeči. 4) Navrhněte a implementujte program, který bude simulovat funkci kochleárního implantátu, zejména pak filtraci řeči a úpravu signálu pro vnitřní část kochleárního implantátu. 5) Navrhněte způsob pro vyhodnocení funkce implantátu. 6) Proveďte vyhodnocení navrženého řešení a diskuzi dosažených výsledků.

DOPORUČENÁ LITERATURA:

[1] ASHLEY, K. L. Analog Electronics with LabVIEW, New York: Pearson Education, Inc., 2003.

[2] NIE, K., STICKNEY, G., ZENG, F.G. Encoding Frequency Modulation to Improve Cochlear Implant Performance in Noise, Trans. on Biomedical Engineering, vol. 52, no 1, pp 64-73, 2005.

Termín zadání: 10.2.2014

Termín odevzdání: 30.5.2014

Vedoucí práce: Ing. Vratislav Harabiš

Konzultanti bakalářské práce:

prof. Ing. Ivo Provazník, Ph.D.

Předseda oborové rady

UPOZORNĚNÍ:

Autor bakalářské práce nesmí při vytváření bakalářské práce porušit autorská práva třetích osob, zejména nesmí zasahovat nedovoleným způsobem do cizích autorských práv osobnostních a musí si být plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č.40/2009 Sb.

ABSTRAKT

Tato bakalářská práce se zabývá zpracováním zvuku pro kochleární implantát. Hlavním úkolem je prostudování základní anatomie a fyziologie lidského sluchu, popis principu kochleárního implantátu a prostudování možností programového prostředí LabVIEW pro zpracování hlasového signálu, zejména pro analýzu řeči. V druhé části je navrhována a implementována simulace kochleárního implantátu v prostředí LabVIEW.

KLÍČOVÁ SLOVA

Kochleární implantát, sluch, LabVIEW, simulace

ABSTRACT

This bachelor thesis deals with signal processing for cochlear implant. The main task is to study basic anatomy and physiology of human hearing system, describe principles of cochlear implants and to study LabVIEW sound processing capabilities, especially human speech analyse. Second part of this thesis is about design and implementation of simulation of cochlear implant in LabVIEW.

KEYWORDS

Cochlear implant, hearing, LabVIEW, simulation

KRATOCHVÍL, T. *Zpracování zvuku pro kochleární implantát*. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, 2014. 46 s. Vedoucí bakalářské práce Ing. Vratislav Harabiš, Ph.D.

PROHLÁŠENÍ

Prohlašuji, že svou bakalářskou práci na téma Zpracování zvuku pro kochleární implantát jsem vypracoval samostatně pod vedením vedoucího bakalářské práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou všechny citovány v práci a uvedeny v seznamu literatury na konci práce.

Jako autor uvedené bakalářské práce dále prohlašuji, že v souvislosti s vytvořením této bakalářské práce jsem neporušil autorská práva třetích osob, zejména jsem nezasáhl nedovoleným způsobem do cizích autorských práv osobnostních a/nebo majetkových a jsem si plně vědom následků porušení ustanovení § 11 a následujících zákona č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon), ve znění pozdějších předpisů, včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č. 40/2009 Sb.

V Brně dne

.....

(podpis autora)

PODĚKOVÁNÍ

Děkuji Ing. Vratislavu Harabišovi, Ph.D., vedoucímu mé bakalářské práce, za konstruktivní a cenné rady, připomínky a pomoc při řešení problémů vzniklých během tvorby práce. Zároveň děkuji panu Ing. M. Topolovi, technikovi firmy Cochlear, který mi poskytl mnoho cenných informací.

OBSAH

OBSAH.....	
SEZNAM OBRÁZKŮ.....	
ÚVOD.....	1
1 SLUCH.....	2
1.1 Zvuk.....	2
1.1.1 Akustický tlak.....	2
1.1.2 Intenzita zvuku.....	3
1.1.3 Hlasitost.....	3
1.2 Anatomie a fyziologie sluchového systému.....	4
1.2.1 Zevní ucho.....	4
1.2.2 Střední ucho.....	5
1.2.3 Vnitřní ucho.....	5
1.3 Transformace zvukového signálu na elektrický.....	6
1.4 Kódování frekvence a intenzity zvuku.....	7
1.5 Prostorové slyšení.....	8
1.6 Poruchy sluchu.....	8
1.6.1 Převodní poruchy sluchu.....	9
1.6.2 Percepční poruchy sluchu (senzoneurální).....	9
1.6.3 Klasifikace sluchových vad podle stupně postižení.....	10
2 KOCHLEÁRNÍ IMPLANTÁT.....	11
2.1 Mikrofon.....	12
2.2 Řečový procesor.....	12
2.3 Kódovací strategie.....	13
2.3.1 Compressed-Analog (CA).....	13
2.3.2 F0/F1/F2.....	14
2.3.3 MPEAK.....	15
2.3.4 Spectral Peak (SPEAK).....	15
2.3.5 Continuous Interleaved Sampling (CIS).....	16
2.3.6 Advanced Combinational Encoder (ACE).....	16
2.3.7 HiRes120.....	17
2.3.8 Další kódovací strategie.....	19

2.4	Cívka a přijímač-stimulátor.....	19
2.5	Elektrody a stimulece.....	20
3	ANALÝZA ŘEČI V LABVIEW	23
3.1	Vstup	24
3.2	Operace s řečovým signálem.	24
3.2.1	Filtrace	25
3.2.2	Operace s frekvenčním spektrem signálu	25
3.3	Výstup a zobrazení	26
4	SIMULACE.....	27
4.1	Načtení signálu.....	29
4.2	Preemfáze	29
4.3	Filtrace.....	30
4.4	Detekce obálky.....	30
4.5	Přehrání a uložení.....	31
4.6	Výpočet spektrogramu	31
4.7	Popis předního panelu	32
4.8	Obsluha programu	33
4.9	Použité vstupní WAV soubory.....	34
5	ZPĚTNÁ REKONSTRUKCE ŘEČI.....	35
6	VYHODNOCENÍ NAVRŽENÉHO ŘEŠENÍ	36
6.1	Vliv preemfáze	36
6.2	Porovnání zapojení s různým počtem kanálů.....	37
	ZÁVĚR	40
	SEZNAM POUŽITÉ LITERATURY	42
	POUŽITÉ SYMBOLY A ZKRATKY	44
	ELEKTRONICKÁ VERZE PRÁCE.....	45

SEZNAM OBRÁZKŮ

Obr. 1: Graf křivek stejné hlasitosti – izofon (převzato z [2]).....	4
Obr. 2: Anatomicko-fyziologický princip ucha (převzato z [3]).....	6
Obr. 3: Znázornění frekvenční selektivity bazilární membrány (převzato z [10])	7
Obr. 4: Ukázka postupné ztráty zvuku s rostoucím věkem (převzato z [17])	10
Obr. 5: Tok dat kochleárním implantátem (převzato z [5]).....	11
Obr. 6: Schéma kódovací strategie CA (převzato z [21]).....	14
Obr. 7: Schéma kódovací strategie F0/F1/F2 (převzato z [21])	14
Obr. 8: Schéma kódovací strategie MPEAK (převzato z [21])	15
Obr. 9: Schéma kódovací strategie CIS (převzato z [20])	16
Obr. 10: Schéma kódovací strategie ACE (PP – pásmová propust, DO – detekce obálky) (převzato z [22]).....	17
Obr. 11: Porovnání spekter generovaných použitím a nepoužitím virtuálních kanálů. (a) Originální signál, (b) Spektrum generované použitím fixních kanálů, (c) Spektrum generované použitím virtuálních kanálů (převzato z [20]).....	18
Obr. 12: Schéma kódovací strategie HiRes120 (převzato z [20])	19
Obr. 13: Vnitřní část KI s vnitřní cívkou, přijímačem-stimulátorem a elektrodami (součást kochleárního implantátu HiRes 90K firmy Advanced Bionics). (převzato z [23])	20
Obr. 14: Elektrody používané firmou Cochlear (převzato z [26]).....	21
Obr. 15: Základní pracovní komponenty LabVIEW	23
Obr. 16: Ikony bloků Acquire Sound a Sound File Read Simple.....	24
Obr. 17: Ukázka funkčního filtračního bloku Butterworth filter.vi.....	25
Obr. 18: Tok dat zapojením v LabVIEW	27
Obr. 19: Načtení souboru WAV a získání informace o vzorkovací frekvenci	29
Obr. 20: Ukázka zapojení bloku preemfáze.....	29
Obr. 21: Schéma zapojení bloku Butterworth filter v režimu pásmová propust	30
Obr. 22: Schéma zapojení detekce obálky signálu	30
Obr. 23: Schéma zapojení uložení a přehrání zvukových dat.....	31
Obr. 24: Schéma zapojení výpočtu spektrogramu a zobrazení spektrogramu.....	31
Obr. 25: Nastavení bloku Spectral Map.....	32
Obr. 26: Ukázka použitého předního panelu	33
Obr. 27: Ukázka schéma zapojení zpětné rekonstrukce pomocí modulace sinusového signálu pro jeden kanál.	35

Obr. 28: Ukázka vlivu použití bloku preemfáze. Vlevo bez preemefáze, vpravo s preemfází. Vstupní soubor - hudba_zpivana v 8 kanálovém zapojení	37
Obr. 29: Srovnání spektrogramů souboru mluvene_slovo a jeho úprav 8, 16 a 22 kanálovým zapojením	39

ÚVOD

Sluch je jedním z lidských smyslů. Jeho ztráta, nebo poškození má neblahý vliv na hodnotu prožívaného lidského života, protože sluchový vjem nelze nahradit žádným jiným smyslem. S věkem dochází k postupné ztrátě sluchu. U dětí je navíc funkce obou uší významná pro správné utváření orientačního smyslu. Jednou z možných sluchových náhrad je kochleární implantát. Využití kochleárního implantátu vyžaduje chirurgický zákrok, což přináší možná větší rizika než ostatní sluchové náhrady. Základní podmínkou pro implantaci zařízení je fyziologický tvar hlemýždě a funkční sluchový nerv. Lidé s kochleárním implantátem neslyší stejně jako zdraví lidé. Výzvou pro budoucí technologický vývoj kochleárních implantátů bude maximální přiblížení zprostředkovaného sluchového vjemu tomu, jaký mají zdraví lidé. Úlohou této bakalářské práce je vytvořit simulaci zpracování řeči v kochleárním implantátu v prostředí LabVIEW a zároveň se seznámit s problematikou týkající se zadání.

1 SLUCH

Sluch je jedním z lidských smyslů a pro běžný život je velmi podstatný. V této kapitole bude probrána fyzikální podstata zvuku a funkce a poruchy lidského sluchu.

1.1 Zvuk

Zvuk je mechanické vlnění, a zároveň nositel informace, kterou zpracovává sluch. Je šířen pomocí mechanických vln, které způsobují tlakové změny a potřebuje tedy ke svému šíření hmotné prostředí (vzduch, voda, kov atd.). Tyto tlakové změny jsou periodické, neperiodické nebo pulzní. Mechanické vlnění může být příčné, nebo podélné. Zvuk se šíří jako podélné vlnění, v pevných látkách pak také i jako příčné vlnění. Vzduchem se zvuk šíří rychlostí 343 m/s (při teplotě 20 °C a normálním atmosférickém tlaku 1013,25 hPa) a se zvyšující se teplotou daného prostředí rychlost šíření zvuku roste. Jsou-li tlakové změny zvuku periodické, lidské sluchové ústrojí ho vnímá jako tón, jehož výška je přímo úměrná frekvenci daného zvuku. Nejjednodušší tón má sinusový průběh a označuje se jako čistý. Zvuk je dán jeho frekvencí (výškou), zabarvením tónu (určeno harmonickými složkami zvuku), akustickým tlakem a intenzitou. [1,2]

Výška zvuku je určena frekvencí zvukového vlnění. Vyšší frekvence dává větší výšku. Lidský sluch je schopen vnímat zvuk v rozsahu od 16 Hz do 20 kHz, ale s rostoucím věkem a možným onemocněním se tento rozsah snižuje, přičemž je zhoršeno vnímání hlavně vyšších kmitočtů. [1] Více zvuků o stejné výšce může znít jinak. Mohou být odlišeny barvou zvuku, která závisí na počtu vyšších harmonických tónů a jejich amplitudách. [2]

1.1.1 Akustický tlak

Látka kmitající v důsledku průchodu mechanického vlnění vyvolává tlakové změny. Jedná se o akustický tlak. Důležitou vlastností sluchu je sluchový práh, který má několik prahových hodnot. Prah slyšení je nejnižší hodnota akustického tlaku při, které člověk zaznamená sluchový vjem. Normální práh slyšení je potom průměrná hodnota prahu slyšení u lidí s normálním sluchem. Prah bolesti je úroveň akustického tlaku při, které člověk pociťuje bolest. Oblast vymezená prahem slyšení a prahem bolesti se nazývá sluchovým polem.[2]

1.1.2 Intenzita zvuku

Intenzita zvuku je dána energií mechanického vlnění, která projde za určitou dobu určitou plochou položenou kolmo na směr šíření mechanického vlnění. Intenzita zvuku je dána vztahem

$$I = \frac{P}{S} \quad [\text{W} \cdot \text{s}^{-2}], \quad (1)$$

kde I značí intenzitu zvuku, P výkon mechanického vlnění ve wattech a S plochu, kterou mechanické vlnění prochází v metrech čtverečních.[2]

1.1.3 Hlasitost

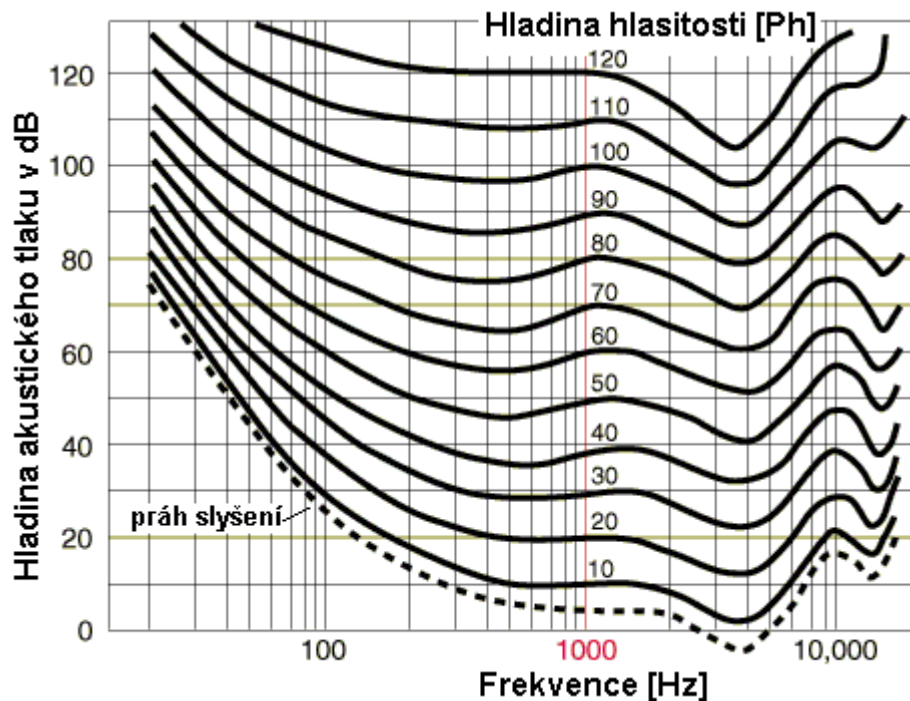
Hlasitost zvuku je subjektivní vjem, který souvisí s intenzitou zvuku. Měřitelnou veličinou je hladina akustického tlaku L_p . Hladina akustického tlaku má velký rozsah, proto je použito logaritmické vyjádření s jednotkou decibel (dB) a je dána vztahem:

$$L_p = 20 \cdot \log \frac{p}{p_0} \quad [\text{dB}], \quad (2)$$

Kde L_p je hladina akustického tlaku, p je daný akustický tlak a p_0 je smluvní vztažná hodnota akustického tlaku, neboli práh slyšení. ($2 \cdot 10^{-5}$ Pa).[2]

Hladina hlasitosti zvuku je subjektivní hlasitost vnímaná sluchem. Její jednotkou je fón (Ph). Jeden fón při frekvenci 1000 Hz odpovídá hladině akustického tlaku 1 dB.[2]

Vyjádřením závislosti vnímání hlasitosti na frekvenci získáme graf křivek stejné hlasitosti – izofon (obrázek číslo 1). Z obrázku číslo 1 je jasné, že člověk nevnímá všechny frekvence při dané hladině akustického tlaku stejně hlasitě.



Obr. 1: Graf křivek stejné hlasitosti – izofon (převzato z [2])

1.2 Anatomie a fyziologie sluchového systému

Sluchový systém je tvořen z vnější části (zevní, střední a vnitřní ucho) a centrální části (sluchová dráha a příslušné projekční oblasti mozkové kůry). [1]

1.2.1 Zevní ucho

Zevní ucho je tvořeno boltcem a zevním zvukovodem, který je zakončen bubínkem. V zevním uchu je pomocí boltce zachycen zvukový signál a je dále zvukovodem přiváděn na bubínek. Boltce je důležitý pro určení směru z jakého zvukový signál přichází. Zevní ucho společně s bubínkem funguje jako rezonátor, který zesiluje určité tóny (nejvíce tóny o frekvenci 3 – 4 kHz). [1]

Bubínek je tenká membrána, která odděluje zevní a střední ucho. Skládá se ze zevní, střední a vnitřní vrstvy a je zasazen do spánkové kosti. S věkem se natočení bubínku mění. Novorozenci jej mají natočen téměř horizontálně, dospívající šikmo a dospělí téměř kolmo na podélnou osu zevního zvukovodu. [1]

1.2.2 Střední ucho

Střední ucho je zasazeno do středoušní dutiny spánkové kosti a od zevního ucha je odděleno bubínkem. S vnitřním uchem ho spojují dva otvory překryté pružnými blanami (oválné a kruhové okénko). Ve středním uchu se nachází tři sluchové kůstky: kladívko, kovadlinka, třmínek. Zvukové vlny rozechvívají bubínek a vibrace se přenášejí na sluchové kůstky v pořadí kladívko (rukojetí přirostlé na bubínek) – kovadlinka – třmínek (pružně zasazen do oválného okénka, ve kterém se kývavě pohybuje).

Eustachova trubice propojuje středoušní dutinu s nosohltanem a otvírá se pouze při polykání, žvýkání, nebo zívání. Vyrovnává tlakové rozdíly ve středouší vzhledem k atmosférickému tlaku. K tlakovým rozdílům dochází během náhlé změny výšky. Bubínek se vyklenuje ven, pokud je tlak ve středouší větší než atmosférický tlak. Pokud je atmosférický tlak větší, než tlak ve středouší, dochází k vtahování bubínku do středoušní dutiny. Oba případy jsou provázeny nepříjemným pocitem zalehnutí až bolesti v uchu.

Díky středoušnímu systému nedochází k razantní ztrátě akustické energie, která by jinak vznikla při šíření zvuku ze vzduchu přímo do kapaliny vnitřního ucha. Na rozhraní vzduchu a vody se odrazí 99.9% akustické energie zpět do vzduchu, tomuto jevu je zabráněno zesilujícími mechanismy středního ucha. Plocha bubínku (55mm^2) je sedmnáctkrát větší než plocha oválného okénka ($3,2\text{mm}^2$). Na oválné okénko působí sedmnáctkrát větší akustický tlak. Uspořádání kladívka, kovadlinky a třmínku funguje jako nerovnoramenná páka s kratším ramenem na straně třmínku. Kmity báze třmínku jsou díky tomu, ve srovnání s chvěním rukojetí kladívka, menší amplitudy, ale mají 1,3 krát větší silový účinek. Důsledkem těchto dvou mechanismů je 22 násobné zesílení ($17 \cdot 1,3$) akustického tlaku na straně oválného okénka.

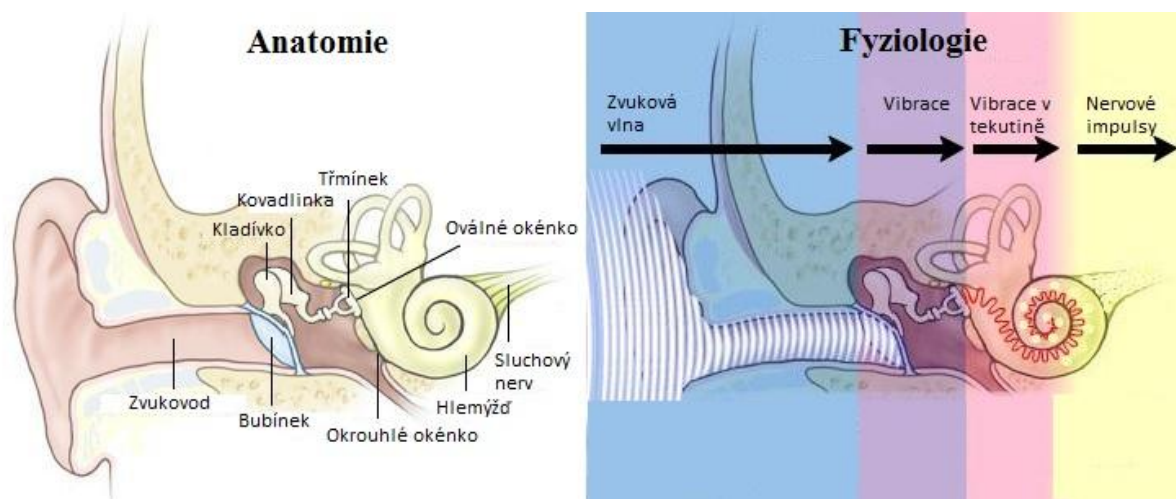
Převod zvuku je možný přes středoušní kůstky (kústkový převod) a pomocí kostního vedení. Během kostního vedení zvuk rozkmitá kosti lebky, kterými se přenáší přímo do kochley (popsán v části „Vnitřní ucho“). Práh slyšení kostního vedení je zhruba o 40 dB výše než u kústkového převodu. Kostní vedení lze ověřit přiložením vibrující ladičky na lebku. [1]

1.2.3 Vnitřní ucho

Vnitřní ucho je uloženo v kostěném labyrintu a obsahuje smyslový epitel sluchového a rovnovážného ústrojí. Sluchová část se skládá z kostěného hlemýžďe (kochlea – kanálek o délce 35mm zatočený do tvaru schránky hlemýžďe) jehož nitrem prochází blanitý

hlemýžď (ductus cochlearis). Ten je vyplněn endolymfou zajišťující přenos zvukového signálu z oválného okénka na membránu Cortiho orgánu, který obsahuje vláskové buňky a slouží jako receptorový orgán sluchu. [1]

Na obrázku číslo 2 je vidět anatomicko-fyziologický princip ucha.



Obr. 2: Anatomicko-fyziologický princip ucha (převzato z [3])

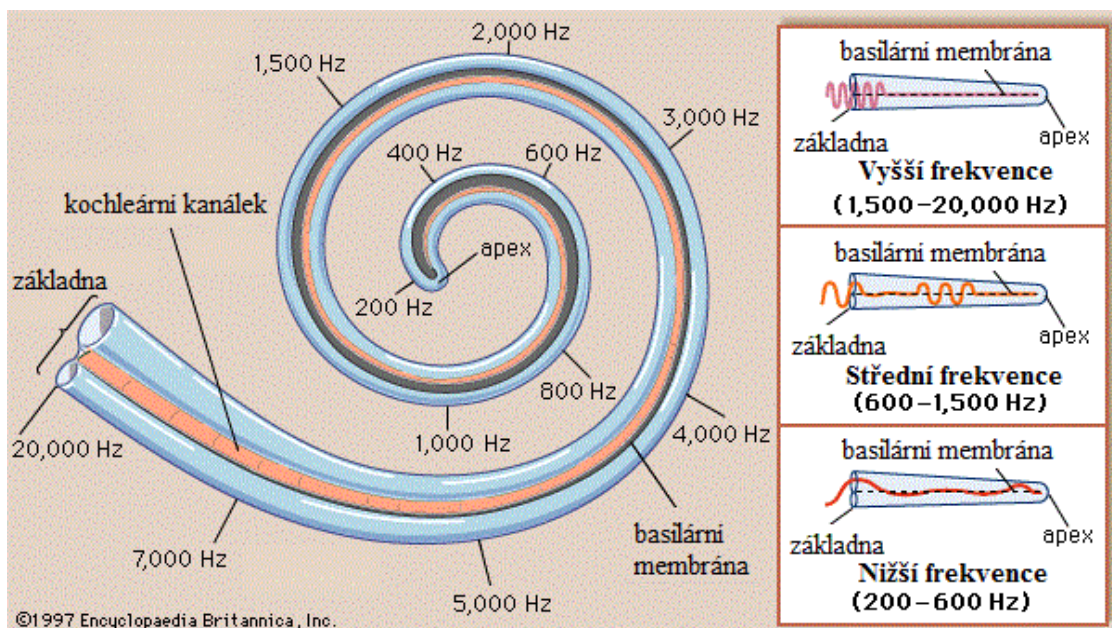
1.3 Transformace zvukového signálu na elektrický

Zvukem způsobené vibrace třmínku vyvolávají tlakové změny perilymfy ve scala vestibuli. Zvuk o slyšitelných frekvencích způsobí tlakovou vlnu šířící se ke kruhovému okénku. Jakmile tlaková vlna dosáhne kruhového okénka, dojde k vyduťtí jeho pružné membrány do středoušní dutiny, čímž se vyrovnává tlaková změna. Interferencí podélných a příčných tlakových vln ve scala tympani, kam obě směřují, dochází k oscilačním pohybům bazilární membrány a Cortiho orgánu. Tento děj vyvolá protipohyb retikulární blány a tektoriální membrány, což způsobí angulární vychylování stereocílií vláskových buněk. Pohyb vlásků způsobuje otevírání a zavírání membránových iontových kanálů receptorové buňky. Díky tomu vzniká generátorový potenciál na vláskové buňce. Stereocilie ohýbané směrem k bazálnímu tělísku vyvolávají depolarizaci, opačný pohyb vyvolává hyperpolarizaci. Vysoká koncentrace K^+ iontů v endolymfě způsobuje vysoký výsledný potenciálový spád mezi intracelulárním prostorem receptorové buňky a endolymfou. To umožňuje snadný vznik generátorových potenciálů. [1]

1.4 Kódování frekvence a intenzity zvuku

Oscilace tlaku nitroušních tekutin, vyvolané vibracemi báze třmínku, vyvolávají postupující vlny, šířící se od oválného okénka k vrcholu hlemýždě. Amplitudy kmiten postupujících vln postupně narůstají a po dosažení maximální amplitudy rychle odezní. Místo výskytu maximální amplitudy vzhledem k oválnému okénku závisí na frekvenci zvuku. Při vysokých frekvencích bude nejvyšší blíže oválnému okénku. Naopak při nižších frekvencích bude maximální amplituda blíže vrcholu kochley (znázorněno na obrázku číslo 3). Z toho plyne, že každé místo bazilární membrány vykazuje maximální amplitudu jen pro určitou frekvenci přichozího vlnění. Těmto místům odpovídá i příslušná skupina receptorových buněk a aferentních nervových vláken. Výška tónu je tedy kódována podle aferentního vlákna sluchového nervu, které vykazuje největší frekvenci akčních potenciálů a odpovídá určitému místu na bazilární membráně.

O kódování intenzity zvuku se soudí, že je prováděna rovnoměrným vzrůstem, nebo poklesem frekvence akčních potenciálů ve vlákních sluchového nervu a počtem aktivovaných axonů tohoto nervu. [1]



Obr. 3: Znázornění frekvenční selektivity bazilární membrány (převzato z [10])

1.5 Prostorové slyšení

Správně fungující sluch je založen na slyšení oběma ušima (binaurální slyšení). Pokud přichází zvuk není ovlivněn rušením odrazy, dokáže člověk velmi přesně rozeznat umístění zdroje zvuku. Za dobrých podmínek v rozsahu 10° v horizontální rovině a hůře ve vertikální rovině.

Pro binaurální slyšení je využíváno časového a intenzitního klíče. Časový klíč je dán rozdílem časových intervalů potřebných pro cestu zvuku do každého ucha. Používá se pro lokalizaci nižších frekvencí zvuku – do 2 kHz. Intenzitní klíč je dán akustickým stíněním hlavy, takže ucho odvrácené od zdroje zvuku, bude přijímat zvuk o nižší intenzitě. Využívá se pro lokalizaci zvuků o vyšších frekvencích. Rozdílná frekvenční detekce je dána změnou vlnové délky při rozdílné frekvenci. Vlnová délka je při frekvencích 2 kHz a vyšších kratší než průměr hlavy a je tedy problematické rozhodnout, jestli na odvrácené ucho dorazila první, nebo některá z dalších zvukových vln. Vyhodnocení intenzitních a časových rozdílů zajišťují binaurální neurony. Konfigurace dendritů těchto neuronů vytváří podklad pro váhový mechanismus vlivu synaptických vstupů z obou uší.

Využití intenzitního a časového klíče je nepoužitelné pro mediálně sagitální rovinu, jelikož rozdíl jejich binaurálního akustického signálu je nulový. Lokalizace v této rovině je monoaurální a využívá směrového účinku ušního boltce, který vytváří rozdílný akustický stín a odrazy pro zdroje s jinou polohou v mediálně sagitální rovině. Tato lokalizace je zhruba pětikrát horší než lokalizace v horizontální rovině.[1]

1.6 Poruchy sluchu

Poruchy sluchu můžeme dělit podle místa vzniku, stupně postižení a doby vzniku. Podrobněji budou popsány první dvě možnosti. Podle místa poškození dělíme poruchy sluchu na centrální a periferní. Centrální porucha sluchu znamená porušení neuronu sluchové dráhy a nemůže být napravena použitím kochleárního implantátu. Periferní porucha sluchu se dělí na převodní a percepční. Její určité případy mohou být kompenzovány použitím kochleárního implantátu. Převodní a percepční vady lze odlišit ladičkovou zkouškou (Weberova, Schwabachova a Rinneho zkouška) a audiometrickými testy. Pro kochleární implantát je podstatná funkčnost sluchového nervu a tvar kochley. [1]

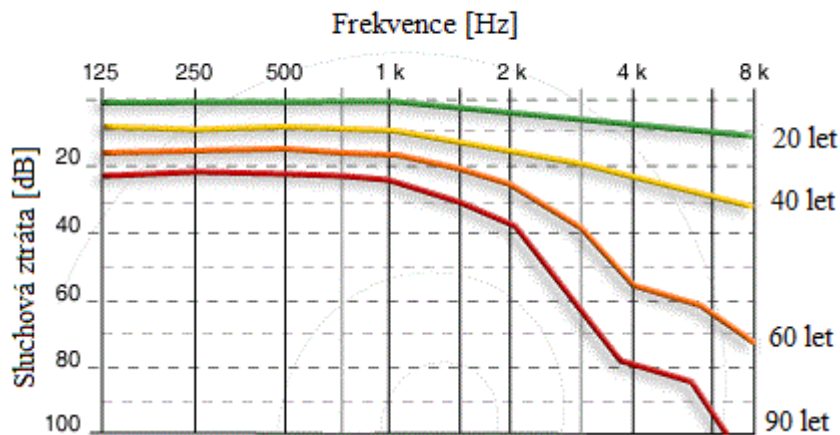
1.6.1 Převodní poruchy sluchu

Převodní poruchy se projevují jako patologická vada v zevním, nebo středním uchu. Dochází k poruše mechanické části sluchové dráhy a narušení přechodu vedení zvuku do vnitřního ucha. Nejčastější příčinou je ucpání zevního zvukovodu mazovou zátkou, perforace bubínku, středoušní zánět, katar Eustachovy trubice, nebo abnormálně tuhé upevnění třmínku v oválném okénku. Převodní vady nikdy nevedou k úplné hluchotě. Zeslabení bývá nejvýše 65 dB. Převodní vady lze kompenzovat pomocí sluchadla, nebo operativně. [1]

1.6.2 Percepční poruchy sluchu (senzoneurální)

Percepční poruchy sluchu můžeme dále dělit na kochleární a retrokochleární. Kochleární poruchy se projevují poškozením kochley. Mezi tyto poruchy řadíme například presbyakuzi a Ménirovu chorobu. Retrokokochleární poruchy se projevují poškozením sluchového nervu a jeho jader. Často následkem zánětu, trauma, tumoru, nebo ototoxickými antibiotiky. Ototoxická antibiotika jsou toxická pro sluchové smyslové buňky. Toxicita vzniká v důsledku pomalejšího vylučování antibiotik z vnitřních tekutin kochley než z krve. Při podání další dávky vzroste koncentrace antibiotik nad kritickou úroveň a hrozí poškození vláskových buněk.

Nejrozšířenější převodní poruchou je presbyakuze neboli stařecká nedoslýchavost. Nejpravděpodobněji je způsobena postupným odumíráním vláskových buněk v Cortiho orgánu. Nastává i u mladší populace po dlouhodobé expozici vysokým akustickým tlakům. Projevuje se zhoršeným vnímáním především vyšších frekvencí (zvoněk, nerozlišitelnost sykavek). Tato porucha je nejčastěji kompenzována sluchadlem. Na obrázku číslo 4 je zobrazena postupná ztráta sluchu s rostoucím věkem. Další možností jsou smíšené poruchy, které vzniknou kombinací převodních a percepčních poruch. [1]



Obr. 4: Ukázka postupné ztráty zvuku s rostoucím věkem (převzato z [17])

1.6.3 Klasifikace sluchových vad podle stupně postižení

Sluchové vady můžeme kategorizovat také podlé míry útlumu sluchového vjemu. Toto rozdělení vystihuje následující tabulka číslo 1.

Tabulka 1: Rozdělení sluchových vad podle míry útlumu sluchového vjemu (převzato z [17])

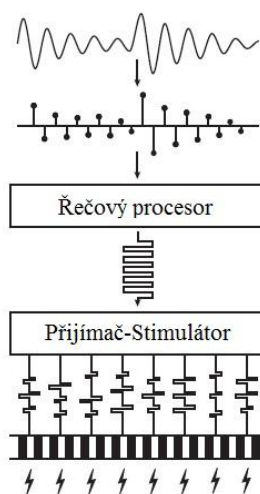
Velikost ztráty sluchu podle WHO	Název kategorie ztráty sluchu	Název kategorie podle Vyhl. MPSV č 284/1995 Sb.
0-25 dB	normální sluch	
26-40 dB	lehká nedoslýchavost	lehká nedoslýchavost (již od 20 dB)
41-55 dB	střední nedoslýchavost	středně těžká nedoslýchavost
56-70 dB	středně těžké poškození sluchu	těžká nedoslýchavost
71-90 dB	těžké poškození sluchu	praktická hluchota
Více než 90 dB, ale body v audiogramu i nad 1 kHz	velmi závažné poškození sluchu	úplná hluchota
V audiogramu nejsou žádné body nad 1 kHz	neslyšící	úplná hluchota

Dalším možným rozdělením sluchových vad je rozdělení podle doby vzniku. Takto můžeme sluchové vady dělit na vrozené a získané. V této práci, ale nebudou blíže popisovány.

2 KOCHLEÁRNÍ IMPLANTÁT

Kochleární implantát (dále jen KI) je implantabilní elektronické zařízení umožňující zvukový vjem lidem s převodní, nebo částečnou percepční sluchovou vadou. Podstatné je, aby byla funkční sluchová dráha. Důležitý je také tvar kochley. KI nahrazuje funkci vláskových buněk a stimuluje pomocí elektrody v kochley sluchový nerv elektrickými impulsy. Provádí se jednostranná i oboustranná implantace. V následující kapitole bude stručně popsána historie kochleárních implantátů, nejvýznamnější výrobci a jejich modely a především základní komponenty a funkční principy KI. Současné KI již snesou vyšetření pacienta pomocí magnetické rezonance a to až do 1,5 T. Je však nutné odebrat magnet, který drží vnější cívku KI na místě. Neexistuje jednotná výrobní norma pro KI, jelikož je nezávisle na sobě vyrábí a vyvíjí několik firem. Mezi hlavní výrobce v současnosti patří: Cochlear Limited, Advanced Bionics, MED-EL, Neurelec a Nurotron. Nejzásadnější roli hrají první tři firmy.[3]

KI se obecně dělí na vnější a vnitřní část. Vnější část obsahuje mikrofon, řečový procesor a vysílač. Vnitřní část se skládá z přijímače-stimulátoru a elektrod. Mikrofon přijímá a převádí zvuk na elektrické impulsy, procesor analyzuje zvuk a určí stimulační úroveň a její vývoj v čase pro každou elektrodu. Výsledný vzor je zasílán pomocí radiových vln vysílačem do přijímače-stimulátoru, který generuje elektrické pulsy pro jednotlivé elektrody. Ve vnitřní části je umístěn generátor bifazických pulsů, který je modulován signálem přijímaným z vnější části KI. Stimulace elektrodami způsobí aktivaci nervových zakončení zvukového nervu a tato aktivita je sluchovou drahou přenášena do centrálního nervového systému a dále zpracovávána s výsledným sluchovým vjemem.[3] Na obrázku číslo 5 je zobrazeno obecné schéma zpracování signálu v KI.



Obr. 5: Tok dat kochleárním implantátem (převzato z [5])

2.1 Mikrofon

Mikrofon je důležitý pro příjem zvukového signálu, který chceme zpracovat. Často používaným mikrofonem je Knowles Acoustics FG-3329, který se skládá z piezoelectrického měniče a tranzistoru JFET. Konstrukce a umístění mikrofonu je důležité pro směrové snímání příchozí zvukové informace. Součástí mikrofonu je i obvod pro automatické vyrovnávání citlivosti. [3]

2.2 Řečový procesor

Základní úlohou řečového procesoru je vyhodnotit příchozí zvukový signál, navzorkovat ho a rozdělit do kanálů pro příslušné elektrody. Tím je zaručeno, že příslušná místa zvukového nervu budou stimulována jen odpovídajícími frekvenčními rozsahy. Dále je nutné signál připravit na odeslání cívkou pomocí radiového signálu. Konkrétní realizace řečového procesoru je u každého výrobce odlišná. Všechny však vychází z určitých základních předpokladů. Dříve se používaly analogové zařízení, dnes už jsou prakticky všechny řečové procesory digitální, využívající technologie CMOS. Řeč je v řečovém procesoru zpracovávána pomocí kódovacích strategií, které budou podrobněji popsány v následující podkapitole. Pro informaci bude uveden stručný přehled současně používaných KI, řečových procesorů a kódovacích strategií, který jsem získal z oficiálních webových stránek výrobců KI. Zajímavostí je řečový procesor Neptune, který je prvním zařízením na světě použitelným během plavání. [3]

Nejnovější nabídkou firmy Cochlear je řečový procesor Nucleus 6 CP910/CP920 (rozdíl je v napájení), Nucleus (CP810) a Freedom, které se používají v KI CI24RE. S využitím kódovacích strategií: ACE, CIS, SPEAK, MP3000.

Firma Advanced bionics využívá řečové procesory Naida, Neptune, Harmony v KI HiRes 90K. S využitím kódovacích strategií HiRes, ClearVoice, CIS, MPS.

Poslední z významnějších firem, MED-EL, používá procesory OPUS 2 a 1 v implantátech CONCERT, SONATA, PULSAR. S využitím kódovacích strategií FS4, FSP a CIS.

Řečový procesor není zprovozněn ihned po operaci, ale zpravidla až po několika měsících. Tento proces vyžaduje audiometrická vyšetření, tak aby byly parametry řečového procesoru nastaveny pro individuální potřeby konkrétního pacienta.[3]

2.3 Kódovací strategie

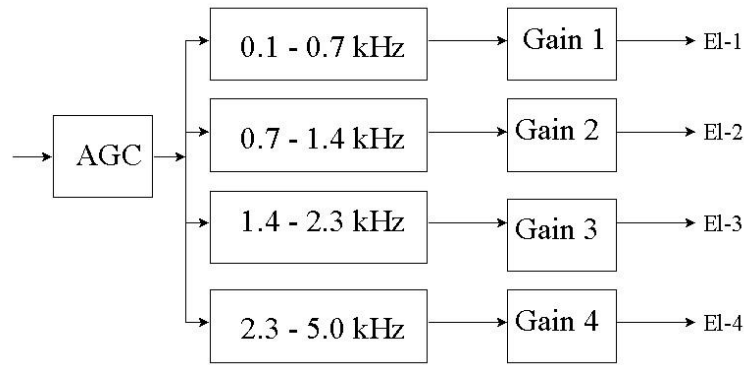
První KI vysílaly analogový signál do každé elektrody současně, což vedlo k součtu vektorů vytvářeného elektrického pole rozdílných úrovní v okolí každé elektrody a následnému vzniku rušivého signálu. Jednalo se o simultánní stimulaci. Tento problém se podařilo vyřešit zavedením kódovacích strategií, které pracují na principu nesimultánní stimulace. To obecně znamená, že v jednom časovém okamžiku je stimulována jen jedna elektroda. Všechny kódovací strategie mají společný výběr frekvenčních pásem z analyzovaného zvuku pomoví banky pásmových propustí s následnou detekcí obálky. Pro stimulaci mohou být vybrány všechny kanály, nebo jen některé. Tato volba záleží na použité kódovací strategii a volbě výrobce. [3] Všechny současné KI využívají při zpracování signálu amplitudovou modulaci. Změnu by mohli přinést nové algoritmy využívající amplitudové i frekvenční modulace. [6] Pro jejich zavedení je ale nutné potvrdit lepší úspěšnost oproti „klasickým“ kódovacím strategiím. Dále je možné kódovací strategie dělit na ty, které využívají pouze fixní kanály a ty, které využívají virtuální kanály. Podstata virtuálních kanálů bude vysvětlena níže.

Před vstupem do banky pásmových propustí prochází signál preemfazickým filtrem. Ten posiluje výkonově slabší složky řeči na vyšších frekvencích oproti výkonově silnějším složkám řeči na nižších frekvencích. V určité míře potlačuje i šum. Obvykle bývá řešen jako filtr typu horní propust s mezní frekvencí 1200 Hz a zeslabením 6 dB na oktávu. [5]

Téměř všichni výrobci využívají kódovací strategie CIS. Každý výrobce má kromě CIS i vlastní kódovací strategie. V následujících několika podkapitolách jich bude několik popsáno.

2.3.1 Compressed-Analog (CA)

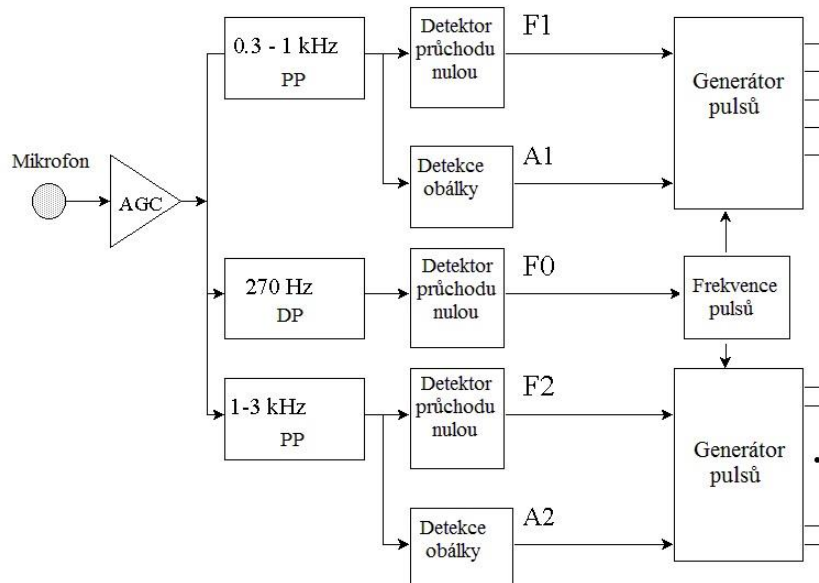
CA strategii lze označit za nejprimitivnější strategii hromadně používanou ve vyráběných a implantovaných KI. Tato strategie byla použita v KI Ineraid vyráběném společností Symbion, Inc, Utah. Nejdříve došlo ke kompresi signálu pomocí obvodu automatického vyrovnávání citlivost a poté k filtraci ve čtyřech kanálech. Se středními frekvencemi 0,5; 1; 2 a 3,4 kHz. Signál poté prošel nastavitelným vyrovnáváním citlivosti a byl poslán perkutánně přímo na čtyři stimulační elektrody v kochley. Signál přicházející na elektrody byl simultánní a analogový. Pro padesát testovaných pacientů s KI Ineraid je uváděna 45% úspěšnost identifikace slov ve větách. Později byl využíván i v KI Clarion firmy Advanced bionics. Na obrázku číslo 6 je schéma kódovací strategie CA.



Obr. 6: Schéma kódovací strategie CA (převzato z [21])

2.3.2 F0/F1/F2

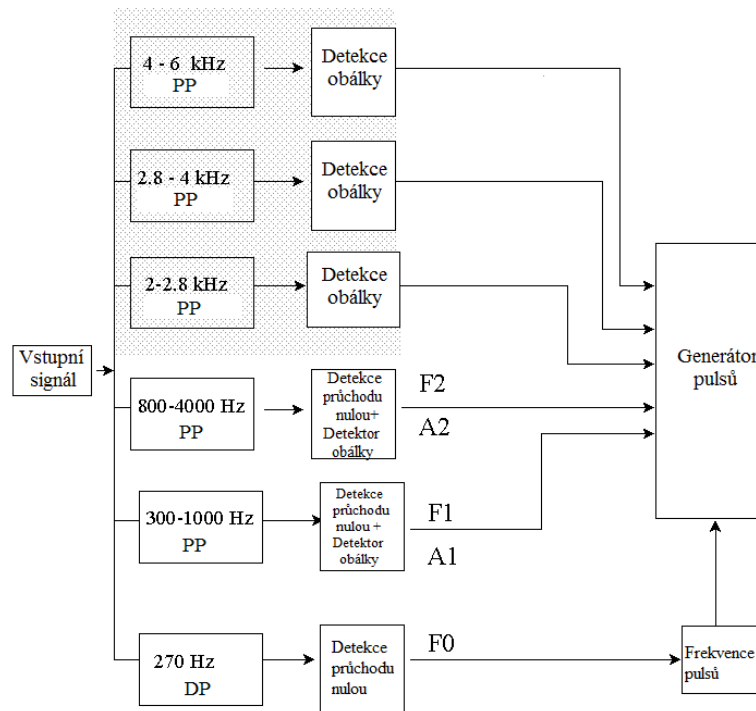
Tato kódovací strategie je vylepšením strategie F0/F2. Spočívá v detekci formantů, což je označení pro oblast lokálního maxima ve spektru složených tónů. Detekujeme formanty F0, F1 a F2. Signál je filtrován bankou pásmových propustí a následně prochází detektorem průchodu nulou a pro formanty F2 a F1 dochází k detekci obálky. Použity jsou dvě sady elektrod, jedna pro F2(báze kochley) a druhá pro F1(apex kochley). Tato strategie dobře funguje pro nízkofrekvenční signály. Pro vyšší frekvence, například frekvenční rozsah souhlásek, bylo nutno vyvinout jiné kódovací strategie. Na obrázku číslo 7 je schéma kódovací strategie F0/F1/F2.



Obr. 7: Schéma kódovací strategie F0/F1/F2 (převzato z [21])

2.3.3 MPEAK

Kódovací strategie MPEAK (někdy také MULTYPEAK) vznikla vývojem kódovací strategie F0/F1/F2. K původní strategii F0/F1/F2 byly přidány tři kanály s pásmovými propustmi. 2 – 2,8 kHz; 2,8 – 4 kHz a 4 – 6 kHz. Šířka pásma pro filtraci formantu F2 byla změněna na 800 – 4000 Hz. Hlavní výhodou strategie MPEAK byla možnost lépe vnímat souhlásky. Na obrázku číslo 8 je schéma kódovací strategie MPEAK.



Obr. 8: Schéma kódovací strategie MPEAK (převzato z[21])

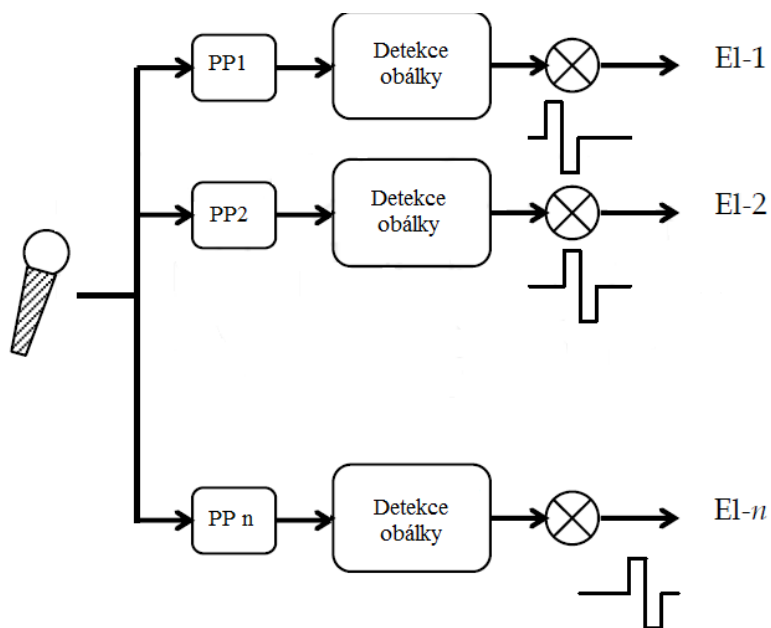
2.3.4 Spectral Peak (SPEAK)

Tato kódovací strategie pracuje s kmitočtovým pásmem (0 – 8 kHz u modelu Nucleus), které je rozděleno na 20 pásem pomocí pásmových propustí. Nejnižší frekvenční pásmo má nejmenší šířku a šířka dalších pásem se nelineárně zvětšuje. V každém pásmu se určí úroveň energie a následně jsou vybrána pásma s největší energií. Počet pásem závisí na kódovací mapě pacienta, intenzitě signálu a tvaru krátkodobého spektra. Může jich být 1 – 10, nejčastěji 6 – 8. Průměrná délka pásma je 4 ms. V jednom aktuálním segmentu zpracovávaného signálu odpovídá počet použitých elektrod počtu vybraných maxim. Hodnota amplitudy stimulačního pulsu je přímo úměrná energii daného pásma.

Vyvolání sluchového vjemu vyžaduje opakování stimulačních pulsů s určitou frekvencí. Stimulační frekvence strategie SPEAK je 250 Hz. [4,5]

2.3.5 Continuous Interleaved Sampling (CIS)

Kódovací strategie CIS byla vyvinuta výzkumníky z RTI (Research Triangle Institute). Jedná se o nesimultánní kódovací strategii. Signál prochází bankou pásmových propustí. Poté je detekována obálka. Následně dochází ke kompresi a přenosu do vnitřní části implantátu. Použití nelineární kompresní funkce zajistí, že výstup bude shodný s dynamickým rozsahem elektricky evokovaného vjemu. Na elektrody jsou přiváděny bifazické pulsy modulované obálkou příchozího signálu s konstantní frekvencí. Podle poskytnutých informací, firma Cochlear používá jako výchozí stimulační frekvenci 1 kHz. Tato hodnota se může změnit v závislosti na pooperačním nastavování u audiologa. KI společnosti Advance bionics využívají 16 kanálů. KI společnosti Cochlear využívají 22 kanálů. Na obrázku číslo 9 je schéma kódovací strategie CIS.

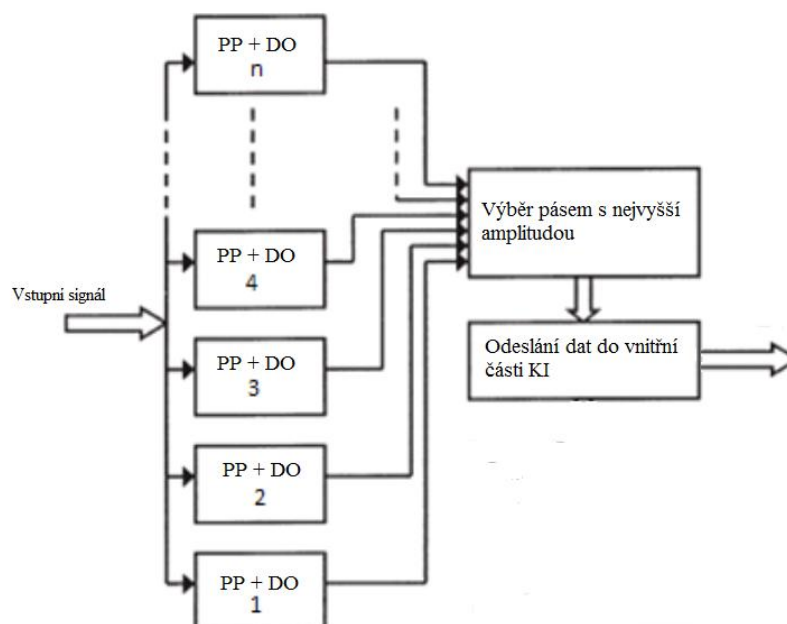


Obr. 9: Schéma kódovací strategie CIS (převzato z [20])

2.3.6 Advanced Combinational Encoder (ACE)

ACE (Advanced Combinational Encoder) je kódovací strategie využívaná firmou Cochlear. Signál je rozdělen do dvaceti dvou kanálů pomocí pásmových zadržů. Poté je detekována obálka. Následně je vybráno 8-10 kanálů s nejvyšší úrovní amplitud (záleží na pooperačním vyšetření audiologa). Signály z vybraných pásem jsou poté zasílány na

příslušné elektrody (nazýváno N z M) princip. Na obrázku číslo 10 je schéma kódovací strategie ACE.

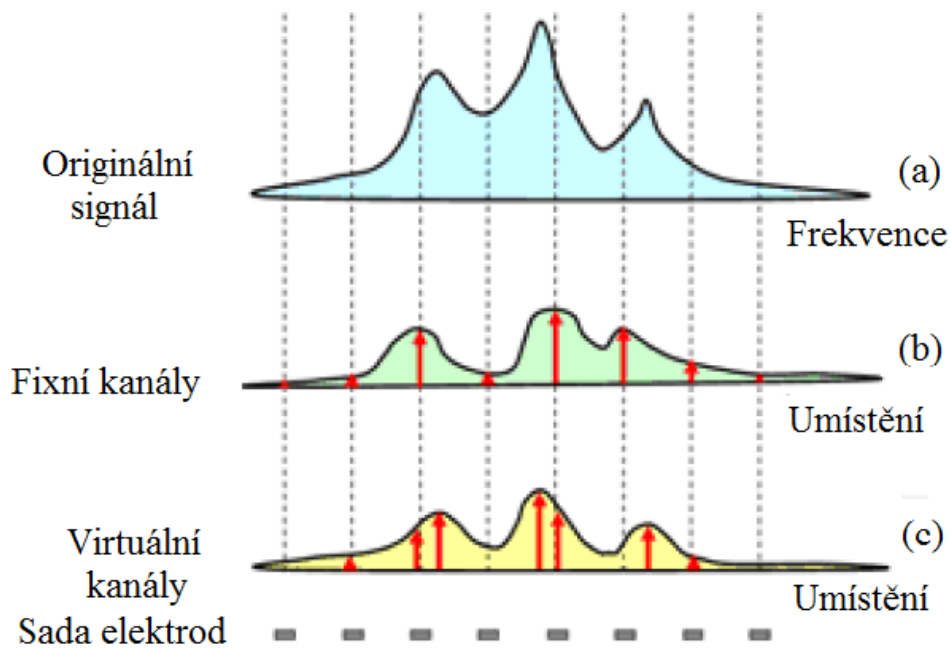


Obr. 10: Schéma kódovací strategie ACE (PP – pásmová propust, DO – detekce obálky) (převzato z [22])

2.3.7 HiRes120

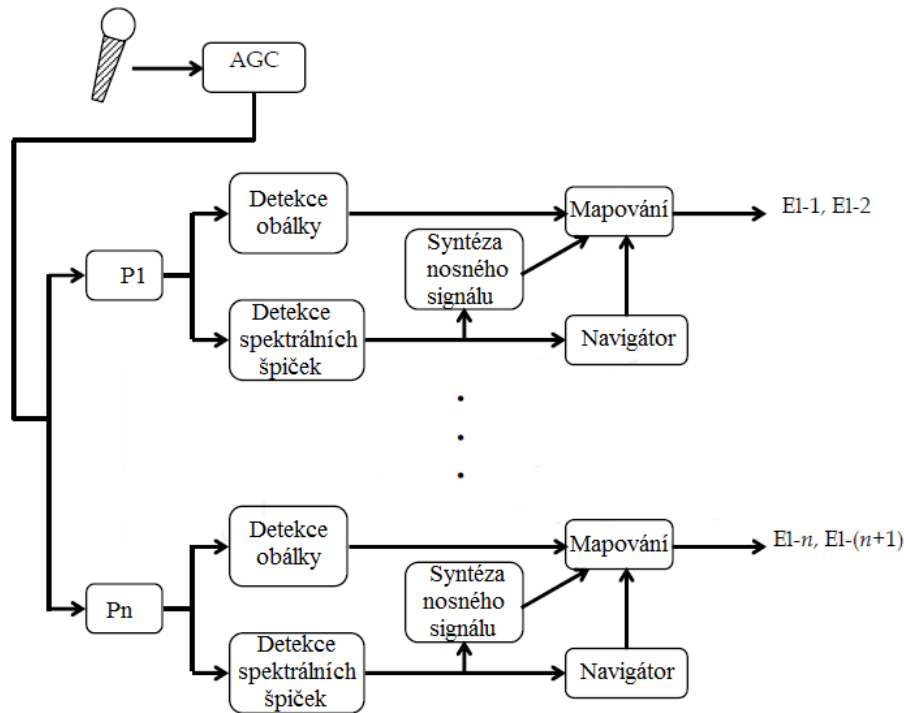
Kódovací strategie HiRes120 je hlavní kódovací strategie používaná firmou Advanced bionics. Je založená na principu virtuálních kanálů a zároveň je první komerčně používanou kódovací strategií, která využívá virtuální kanály. Pro správnou funkci této metody je nutné zajistit simultánní dodávání proudu do více než jedné elektrody. Použitím virtuálních kanálů je navýšeno frekvenční rozlišení/pokrytí stimulace. Podle informací poskytnutých firmou Cochlear tato metoda není výrazně lepší než ACE, nebo CIS.

Jeden virtuální kanál je generován dvěma sousedními elektrodami. Je určen poměrem proudů v obou elektrodách. Pokud dochází ke vhodné stimulaci sousedních elektrod, tak mezi nimi dochází ke sluchovému vjemu o frekvenci, která je dána poměrem stimulačních frekvencí a amplitud proudových pulsů přiváděných na sousední elektrody. Na obrázku číslo 11 je zobrazeno zlepšení použitím virtuálních kanálů.



Obr. 11: Porovnání spekter generovaných použitím a nepoužitím virtuálních kanálů. (a) Originální signál, (b) Spektrum generované použitím fixních kanálů, (c) Spektrum generované použitím virtuálních kanálů (převzato z [20])

Použito je 16 implantovaných elektrod, které tvoří 15 elektrodových párů. Vstupní signál je rozdělen do patnácti pásem a filtrován bankou pásmových propustí. V každém pásmu je detekována obálka. Pomocí Fourierovy transformace je v každém pásmu detekována frekvenční špička. Tyto frekvenční špičky poté pomáhají řídit stimulaci ve virtuálních kanálech. Kódovací strategie HiRes120 používá shodné stimulační frekvence jako strategie CIS. Na obrázku číslo 12 je schéma kódovací strategie HiRes120.



Obr. 12: Schéma kódovací strategie HiRes120 (převzato z [20])

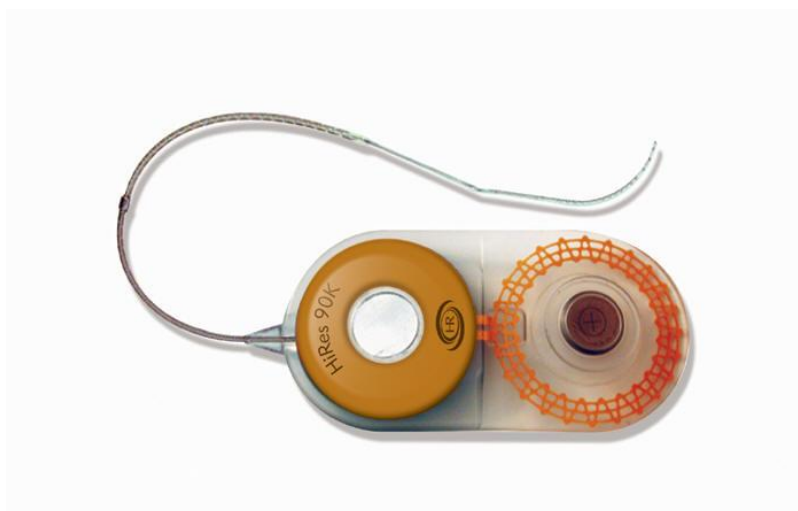
2.3.8 Další kódovací strategie

Mezi nejnovější kódovací strategie patří například kódovací strategie Frequency-Amplitude-Modulation-Encoding Algorithm (FAME), která využívá amplitudovou i frekvenční modulaci a Wavelet zero-crossings stimulation (WZCS). Podle studií překonává FAME i WZCS aktuálně používané kódovací strategie v hlučném i tichém prostředí. WZCS je údajně lepší než FAME. Otázkou však je, jestli tyto výsledky potvrdí i další studie. Podle informací poskytnutých firmou Cochlear v nejbližší budoucnosti nedojde ke komerčnímu využití těchto kódovacích strategií.[6]

2.4 Cívka a přijímač-stimulátor

Informace a napájecí energie mezi vnitřní a vnější částí KI je přenášena pomocí radiových vln. Pro přenos slouží nosná vlna o frekvenci 5 MHz. Vnější cívka je umístěna na povrchu kůže nad uchem. Vnitřní cívka je společně s přijímačem-stimulátorem implantována pod kůži. Ve vnitřní části se nachází permanentní magnet, který drží na místě vnější cívku. Umístění napájecího zdroje mimo vnitřní část KI je nutností, jelikož ve vnitřní části není dostatek místa pro akumulátor potřebné kapacity. Umístění menšího akumulátoru o menší kapacitě by znamenalo nutnost častých chirurgických zákroků a to je značně nevýhodné. [3]

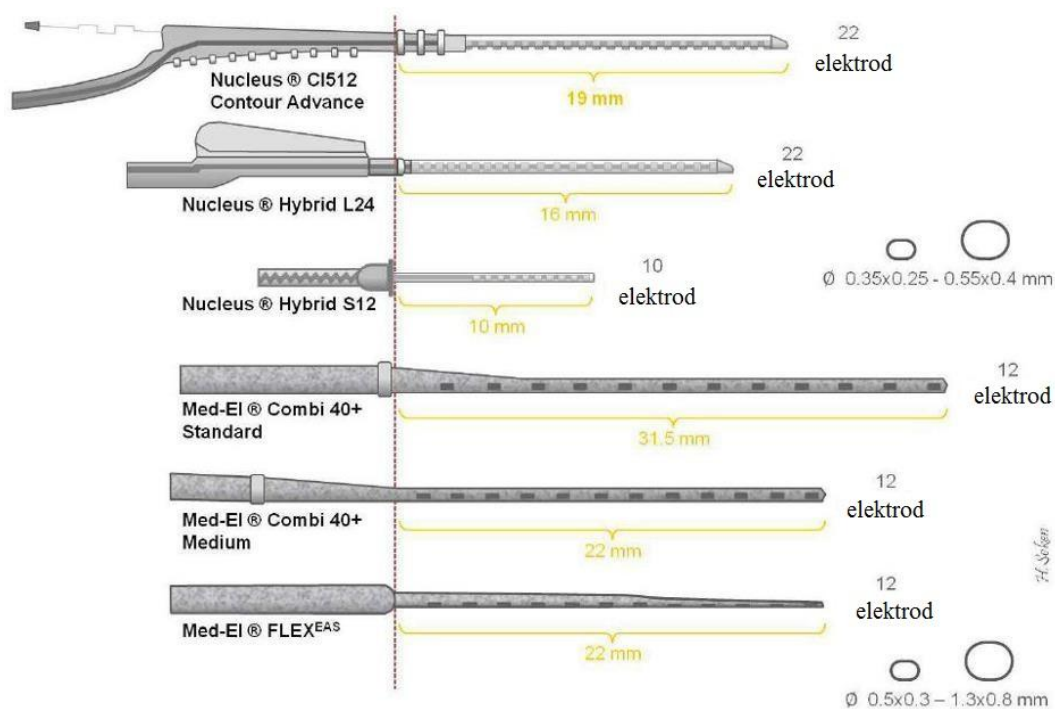
Na obrázku číslo 13 je zobrazena vnitřní část KI s vnitřní cívkou, přijímačem-stimulátorem a elektrodami.



Obr. 13: Vnitřní část KI s vnitřní cívkou, přijímačem-stimulátorem a elektrodami (součást kochleárního implantátu HiRes 90K firmy Advanced Bionics). (převzato z [23])

2.5 Elektrody a stimulace

Elektrody jsou nejčastěji vyráběny z platiny a jsou zasazeny do silikonové formy v podobě malého pásku. Dalšími používanými materiály jsou kombinace platiny s iridiem (Pt-Ir), nitrid titanu, iridium a oxid iridia. Silikonový pásek s elektrodami je volně umístěn v tekutině vyplňující vnitřní prostor kochley. Existují různé techniky implantace elektrod lišící se v síle, kterou je třeba vyvinout pro správné vložení elektrodového pásku do hlemýždě vzhledem k možnému pooperačnímu traumatu. Jednou z nich je například implantační technika AOS. Použité elektrody by měly splňovat určité parametry. Mezi ty hlavní patří efektivní přenos náboje s minimální spotřebou během stimulace, malá impedance s nízkou frekvenční závislostí a biokompatibilní materiál s dlouhou životností. [5, 18] Na obrázku číslo 14 jsou elektrody používané firmou Cochlear. V současnosti je možné implantovat maximálně 22 aktivních elektrod. Počet aktivních elektrod přímo ovlivňuje frekvenční rozlišení vjemu pacienta. Několik elektrod nahrazuje funkci tisíců vláskových buněk.



Obr. 14: Elektrody používané firmou Cochlear (převzato z [26])

Implantace elektrod probíhá formou chirurgického zákroku. Chirurg je schopen umístit silikonový pásek do hloubky maximálně 30 mm. Efektivní délka kochleý je přibližně 40 mm. [5]

Starší KI využívaly jedno kanálovou stimulaci sluchového nervu analogovým signálem. Tento systém byl však neuspokojivý (pacienti dosahovali při testu poslechu vět maximálně 4% úspěšnosti a během testu poslechu samohlásek se úspěšnost pohybovala okolo 0%). Postupně se však vyvinuly nové přístupy a formy stimulace. [3]

Důležitá je forma zapojení aktivních elektrod a země. Dříve se používala metoda common ground. Jedna elektroda byla aktivní – stimulační a ostatní elektrody tvořily zem. Další možností je bipolární mód, kdy je jedna elektroda aktivní – stimulační a vedlejší elektroda je použita jako zem. Dnes nejvíce používanou metodou je monopólní mód. Všechny vnitřní elektrody jsou postupně používány jako aktivní – stimulační elektrody. Jako zemnicí elektroda je použita elektroda umístěná na lebce, nebo zem tvoří kovový povrch implantátu (titan).

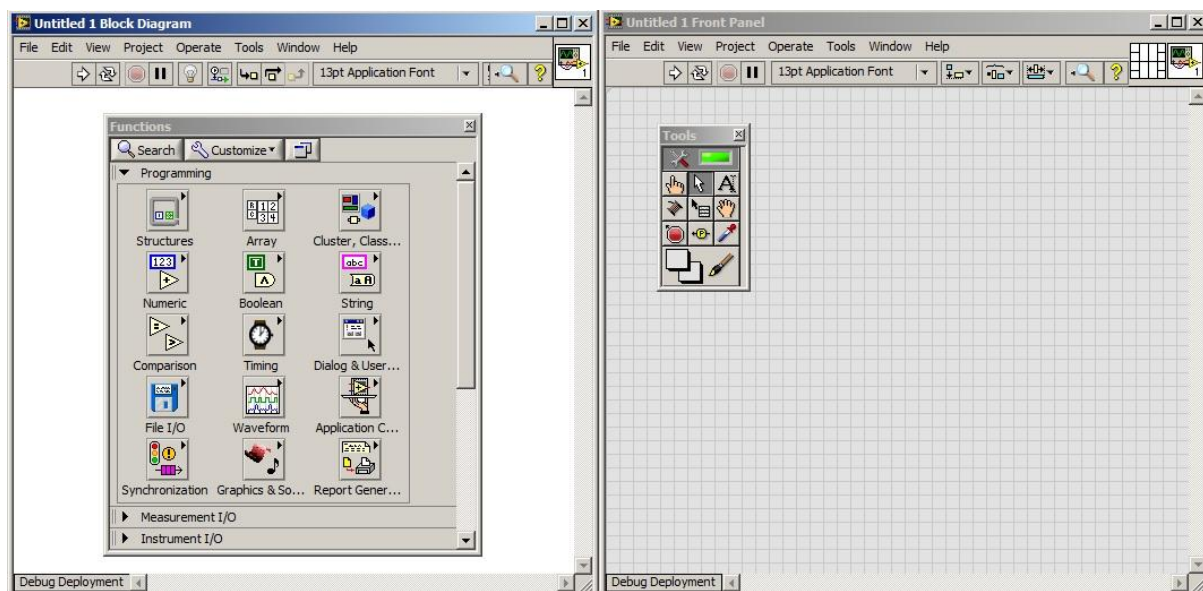
Zásadní je i pořadí stimulace jednotlivých elektrod. Kódovací strategie CA zasílala signál na všechny elektrody současně. Takový způsob stimulace způsobí nárůst šumu a výkyvů hlasitosti vnímaného zvuku. Na základě tohoto zjištění byly zavedeny

nesimultánní způsoby stimulace. Podle informací poskytnutých firmou Cochlear je konkrétní postup stimulace volitelný technikem/audiologem. Pořadí stimulace elektrod může být například od elektrod v bázi kochley k elektrodám v apexu kochley, nebo naopak. Stimulace může probíhat i přeskokově (1-4-2-3).

U moderních KI mají stimulační pulsy podobu bifazických pulsů modulovaných zpracovaným nasnímaným signálem. Tyto bifazické pulsy jsou generovány proudovým zdrojem. Amplituda pulsů se pohybuje v rozmezí 10 – 1750 μA a trvají zhruba 100 μs . Impedance kochley se pohybuje v rozmezí 1 – 10 $\text{k}\Omega$. [19]

3 ANALÝZA ŘEČI V LABVIEW

LabVIEW je grafické vývojové prostředí vyvíjené společností National Instruments (NI). Jeho využití je velmi komplexní a nevyžaduje hlubší znalost programovacího jazyka. Návrh systému probíhá v okně Front panel (Přední panel) a Block diagram (Blokový diagram). Průběžný postup, nebo výsledný program lze uložit ve formátu VI (Virtual Instrument – Virtuální Přístroj). Nejaktuálnější je verze NI LabVIEW 2013, ale připravuje se novější verze LabVIEW 2014. LabVIEW je také vybaveno rozsáhlou kontextovou nápovědou a funkčními VI soubory, které demonstrují řešení dané problematiky. Placená verze obsahuje speciální pracovní balíčky pro práci se zvukem – Sound and Vibration Toolkit a pro práci se signály – Signal Processing. Na obrázku číslo 15 jsou uvedeny základní pracovní komponenty LabVIEW.[8]



Obr. 15: Základní pracovní komponenty LabVIEW

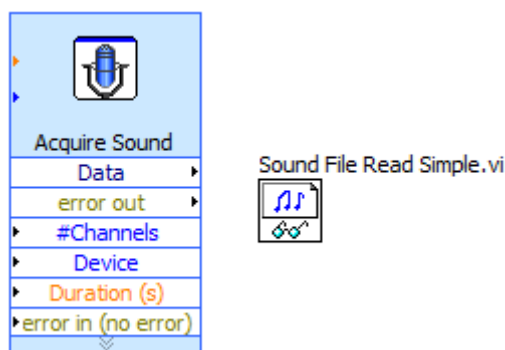
Programové prvky (filtry, smyčky, načítání atd.) jsou umístovány do blokového diagramu. Zobrazovací prvky (grafy atd.) a kontrolní prvky (přepínače, indikátory) jsou umístovány na přední panel. Nabídka Tools umožňuje přepínání aktivního pracovního kursoru. Prvky jsou do blokového diagramu i na přední panel umístovány pomocí palety funkcí. Je možné použít systematický postup nabídkou, nebo název prvku zadat slovně do vyhledávače. Jednotlivé prvky v blokovém diagramu jsou propojovány drátkem (wire). LabVIEW využívá několik druhů datových drátků (scalar, 1D array, 2D array, cluster atd.) pro různé datové typy (numeric, boolean, string, path atd.). Pro funkční zapojení je zapotřebí přivádět na funkční bloky pouze požadovaný datový typ. Datové

typy jsou indikovány rozdílnou barvou. Každý funkční prvek má svou kontextovou nápovědu, která uvádí požadované typy vstupu a nabízené výstupy. Kontextovou nápovědu lze zapnout ikonou žlutého otazníku v horním pravém rohu předního panelu i blokového diagramu.

LabVIEW podporuje tvorbu VI podsouborů (subVI files). Každý takový soubor plní konkrétní operaci (filtrace, zobrazení, načítání) a je možné je zapojit v hlavním VI souboru. V LabVIEW je možné využít i možnosti globálních proměnných.

3.1 Vstup

Před samotnou analýzou dat musíme data nějakým způsobem do analyzačního zařízení dopravit. V případě LabVIEW se nabízí hned několik možností jak nasnímat potřebný zvukový signál. Pro laboratorní účely slouží akviziční karty. V laboratořích FEKT VUT jsou k dispozici akviziční karty NI PCI 6221. Pro nahrání dat do LabVIEW se používá funkční blok DAQ Assistent. LabVIEW. Použít můžeme i funkční blok Acquire Sound.vi, který umožňuje snímání zvuku přímo z jack konektoru zvukové karty. V případě, že chceme použít již nahraný zvukový soubor (ve formátu WAV) můžeme použít funkční blok Sound File Read Simple.vi. Na obrázku číslo 16 je zobrazena ikona bloku Acquire Sound a Sound File Read Simple.



Obr. 16: Ikony bloků Acquire Sound a Sound File Read Simple

3.2 Operace s řečovým signálem.

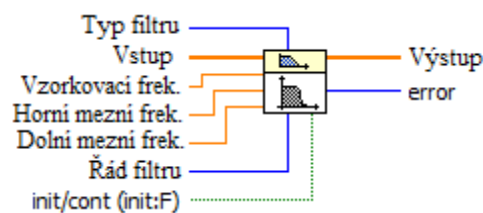
Lidský sluch dokáže vnímat zvuk v rozmezí 20 Hz – 20 kHz. Frekvenční rozsah lidské řeči sahá zhruba k 10 kHz, ale pro přenos srozumitelné řeči stačí pásmo 1 – 3 kHz. Vyš-

ší frekvence pak mají vliv především na barvu hlasu. Proto nás bude při analýze řeči zajímat hlavně frekvenční rozsah zhruba 100 Hz až 7 kHz.

3.2.1 Filtrace

Pro zpracování řeči je důležitá filtrace zvukového signálu. Obecně můžeme digitální filtraci provádět filtry FIR (filtr s konečnou impulsní charakteristikou), nebo IIR (filtr s nekonečnou impulsní charakteristikou). Při nastavování FIR filtru uvádíme délku impulsní charakteristiky. Při nastavování IIR filtru uvádíme číslo řádu. Čím vyšší řád, tím strmější frekvenční charakteristiku bude mít, ale zároveň bude výpočetně náročnější. Výhodou FIR filtrů je jejich lineární fázová charakteristika, nevýhodou je zpoždění signálu. Pokud není lineární fázová charakteristika výchozí podmínkou, je vhodnější použít IIR filtry. Po návrhu IIR filtru je však vždy nutné prověřit jeho stabilitu. Pro analýzu řečového signálu není lineární fázová charakteristika důležitá, proto budou použity IIR filtry. IIR filtry můžeme použít pro zapojení dolních propustí, horních propustí a pásmových zádrží i propustí.

V LabVIEW lze použít jeden filtr samostatně, nebo vytvořit banku filtrů. Použit lze expresní prvek Filter.vi. Po umístění do blokového diagramu se automaticky spustí nabídka nastavení filtru, takže na prvek není nutné připojovat žádné další kontrolní prvky. V nastavení je možné vybrat FIR, nebo IIR filtr. Typ filtru (dolní, horní propust apod.), mezní frekvence filtrace a konkrétní druh filtru (Butterworth, Chebyshev, Bessel). Další možností je využít samostatné bloky příslušných prvků (například Butterworth Filter.vi). Z palety Signal Processing > Filters. Tyto prvky je nutné nastavovat připojením kontrolních prvků v blokovém diagramu. Používány budou filtry typu Butterworth. Na obrázku číslo 17 je ukázka filtračního bloku Butterworth Filter.vi.



Obr. 17: Ukázka funkčního filtračního bloku Butterworth filter.vi

3.2.2 Operace s frekvenčním spektrem signálu

Při práci s řečovým signálem může být důležité zjistit informaci o frekvenčním spektru signálu. To je umožněné díky Fourierově transformaci. Vyhodnocovat můžeme napří-

klad závislost amplitudy na frekvenci, nebo výkonové spektrum. V rámci této práce je výhodné použít porovnání spektrogramů originálního a upraveného signálu.

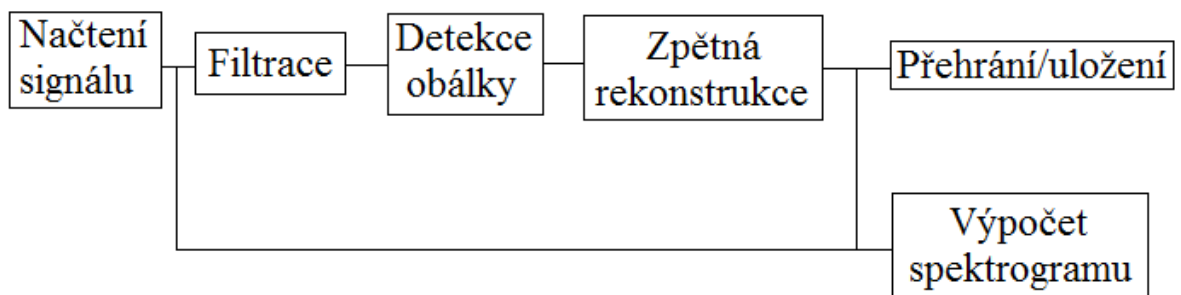
V LabVIEW je možné použít expresní i obyčejné prvky. V případě expresních prvků lze použít blok Spectral Measurements z palety Express > Signal Analysis. Další možností je volba konkrétního prvku. Pro výpočet spektrogramu je možné použít například prvek TFA STFT Spectrogram.vi z palety Signal Processing > Time Frequency Analysis > Spectrogram. Pro výpočet spektrogramu pomocí expresního prvku je možné použít prvek Spectral Map.

3.3 Výstup a zobrazení

Zpracovaná zvuková data je možné v prostředí LabVIEW zobrazovat v grafu (závislost amplitudy na čase), přehrávat, nebo ukládat do WAV souboru. Pro vykreslení do grafu je možné použít prvek waveform graph, který se umístí na přední panel a zároveň se zobrazí i v poli blokového diagramu. Přehrání je možné provést pomocí expresního prvku Play Waveform. Po umístění tohoto prvku do blokového diagramu se zobrazí nabídka výstupního zařízení s informací o jeho možnostech. Je možné vyzkoušet výstupní zařízení přehráním zkušebního tónu. Zvuk můžeme přehrát i pomocí prvku Play Sound File.vi. Uložit zvuková data je možné pomocí Sound File Write Simple.vi do souboru formátu WAV. Tento blok vyžaduje prvek path s informací o umístění výstupního souboru.

4 SIMULACE

Kódovací strategie CIS je využívána prakticky ve všech současných kochleárních implantátech, proto bude použita pro simulaci funkce kochleárního implantátu. V této kapitole bude popsáno zapojení jednotlivých prvků v LabVIEW a jejich funkce. Pro úspěšné porozumění mluvené řeči je dostačující použití osmi elektrod. Firma Cochlear a Advanced Bionics používají 22 a 16 elektrod. Zapojení s 8, 16 a 22 kanály se liší pouze počtem paralelních větví zapojení, ale princip a umístění funkčních prvků v jednotlivých kanálech je stejný. [5, 15] Obrázek celkového schéma by se stal po umístění do textu nečitelný, proto bude popis rozdělen na části. Na obrázku číslo 18 je zobrazeno obecné schéma toku dat simulovaným zapojením. Data jsou nejprve načtena a následně zpracována v daném počtu kanálů. Výsledný signál je možné přehrát, nebo uložit. Pro porovnání původního signálu a výstupního signálu jsou vypočteny spektrogramy. Průběh zpětné rekonstrukce je popsán samostatně v kapitole 5. Na konci této kapitoly budou také popsány jednotlivé prvky umístěné na předním panelu a obsluha programu.



Obr. 18: Tok dat zapojením v LabVIEW

V zapojení není bráno v potaz automatické vyrovnávání citlivosti, které je používáno na vstupu kochleárních implantátů. Obvykle bývá součástí mikrofону. V simulaci není zahrnut ani přenos radiovými vlnami s vyrovnáním úrovně intenzity a generování bifazických pulsů. Vzhledem k použitému druhu rekonstrukce signálu to není zapotřebí.

Hodnoty frekvenčních pásem pro 8 a 16 kanálové zapojení je převzato z [15]. Hodnoty frekvenčních pásem pro 22 kanálové zapojení byly z důvodu nenalezení vhodného zdroje experimentálně stanoveny. V následujících tabulkách číslo 2, 3 a 4 jsou uvedeny frekvenční údaje pro pásma jednotlivých zapojení.

Tabulka 2: Frekvenční údaje pro zapojení s 8 kanály

Kanál	1	2	3	4	5	6	7	8
Střední frekvence pásma [Hz]	384	692	1064	1528	2109	2844	3740	4871
Šířka pásma [Hz]	265	331	431	516	645	805	1006	1257

Tabulka 3: Frekvenční údaje pro zapojení s 16 kanály

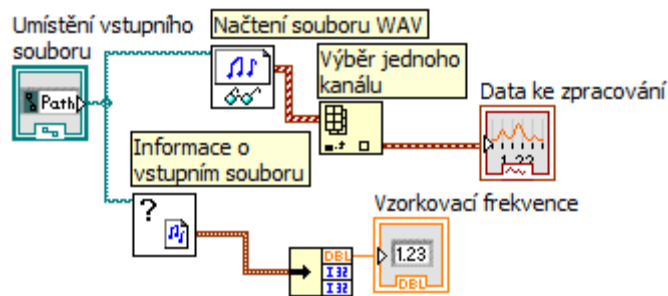
Kanál	1	2	3	4	5	6	7	8
Střední frekvence pásma [Hz]	216	343	486	647	828	1031	1260	1518
Šířka pásma [Hz]	120	135	151	170	192	216	242	273
Kanál	9	10	11	12	13	14	15	16
Střední frekvence pásma [Hz]	307	345	389	437	492	553	622	700
Šířka pásma [Hz]	1000	1000	1000	1000	1000	1000	1000	1000

Tabulka 4: Frekvenční údaje pro zapojení s 22 kanály

Kanál	1	2	3	4	5	6	7	8
Střední frekvence pásma [Hz]	200	280	375	450	550	660	765	890
Šířka pásma [Hz]	110	125	140	160	182	204	228	260
Kanál	9	10	11	12	13	14	15	16
Střední frekvence pásma [Hz]	1025	1165	1310	1460	1615	1775	1940	2110
Šířka pásma [Hz]	296	334	374	416	460	506	554	604
Kanál	17	18	19	20	21	22		
Střední frekvence pásma [Hz]	2285	2585	3085	3785	4485	5185		
Šířka pásma [Hz]	656	736	836	836	836	836		

4.1 Načtení signálu

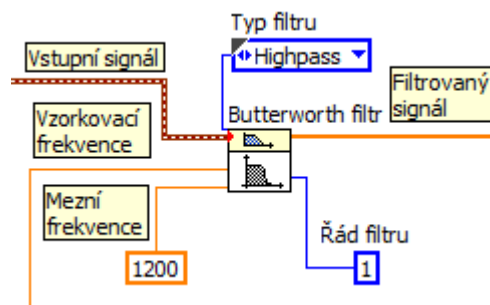
Zvukový signál je načítán z WAV souboru pomocí bloku Sound File Read Simple.vi. Cesta ke vstupnímu souboru je zadána pomocí bloku path. Po načtení je vybrán pouze jeden kanál a data jsou odeslána dále ke zpracování do všech pracovních kanálů zapojení. Blok Sound File Info.vi umožňuje získat informace o načítaném souboru. Pro další zpracování je využita informace o vzorkovací frekvenci. Na obrázku číslo 19 je schéma načtení souboru WAV.



Obr. 19: Načtení souboru WAV a získání informace o vzorkovací frekvenci

4.2 Preemfáze

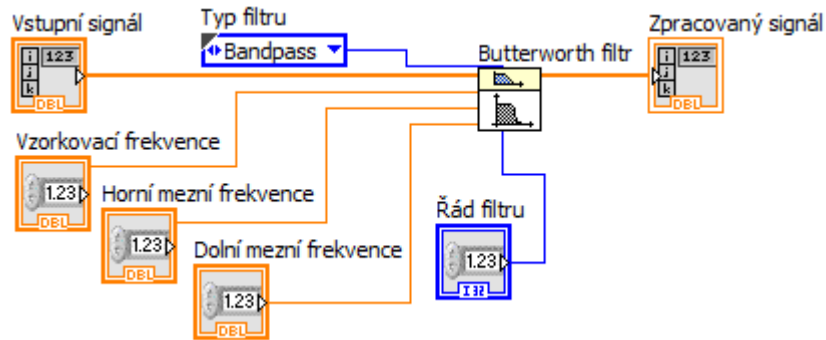
Před vstupem do banky filtrů prochází signál blokem preemfáze. Účelem tohoto bloku je zeslabit vliv řečových složek o větší intenzitě na nižších frekvencích oproti složkám o nižší intenzitě na vyšších frekvencích. Obvykle je použit filtr typu horní propust s mezní frekvencí 1200 Hz a zeslabením 6 dB na oktávu. Těmto potřebám vyhovuje IIR filtr Butterworth prvního řádu v zapojení horní propust, který vykazuje zeslabení 6 dB na oktávu. Na obrázku číslo 20 je ukázka použitého zapojení bloku preemfáze. [5]



Obr. 20: Ukázka zapojení bloku preemfáze

4.3 Filtrace

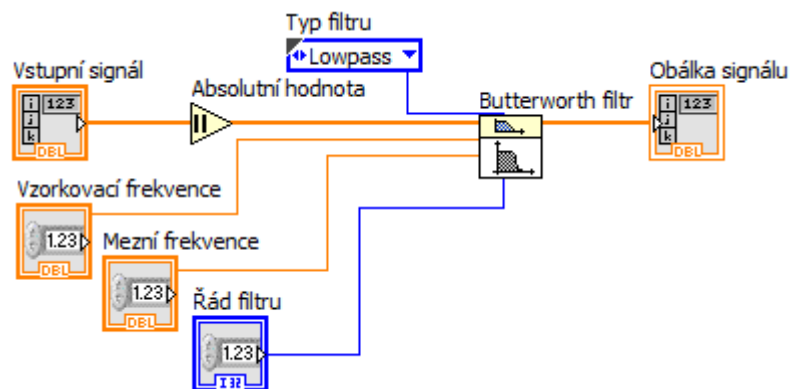
Filtrace probíhá pomocí filtračního bloku Butterworth filter.vi v režimu pásmová propust. Tento blok pro správnou funkci potřebuje vstupní data, nastavení režimu (pásmová propust), horní a dolní mezní frekvenci, vzorkovací frekvenci a řád filtru. Na obrázku číslo 21 je ukázáno schéma zapojení filtru Butterworth v režimu pásmová propust.



Obr. 21: Schéma zapojení bloku Butterworth filter v režimu pásmová propust

4.4 Detekce obálky

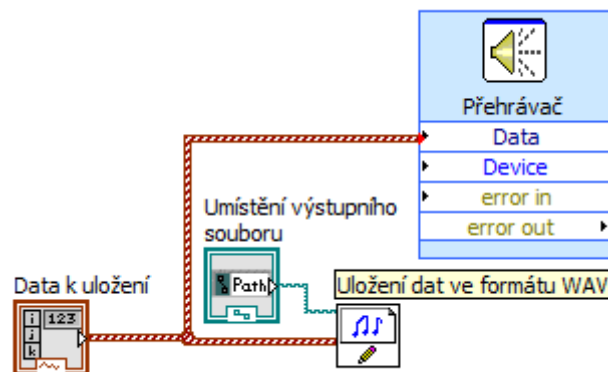
Pro detekci obálky byl použit jednoduchý princip. Nejdříve je signál převeden na absolutní hodnotu blokem Absolute Value a dále je filtrován blokem Butterworth filter v režimu dolní propust s mezní frekvencí 200, nebo 400 Hz (záleží na rozmezí daného kanálu). Tím je zajištěna detekce pomalu se měnícího signálu, který „obaluje“ signál původní. Schéma detekce obálky je uvedeno na obrázku číslo 22.



Obr. 22: Schéma zapojení detekce obálky signálu

4.5 Přehrání a uložení

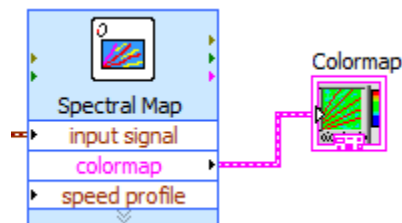
V programu má uživatel možnost přehrát, nebo uložit soubor ve formátu WAV. Přehrání je realizováno funkčním blokem Play Waveform. Uložení dat umožňuje blok Sound File Write Simple.vi. Tomuto bloku pro správnou funkci postačuje pouze vstup dat a zadání umístění a názvu výstupního souboru. Umístění výstupního souboru je nastaveno prvkem path. Na obrázku číslo 23 je znázorněno schéma uložení a přehrání zvukových dat.



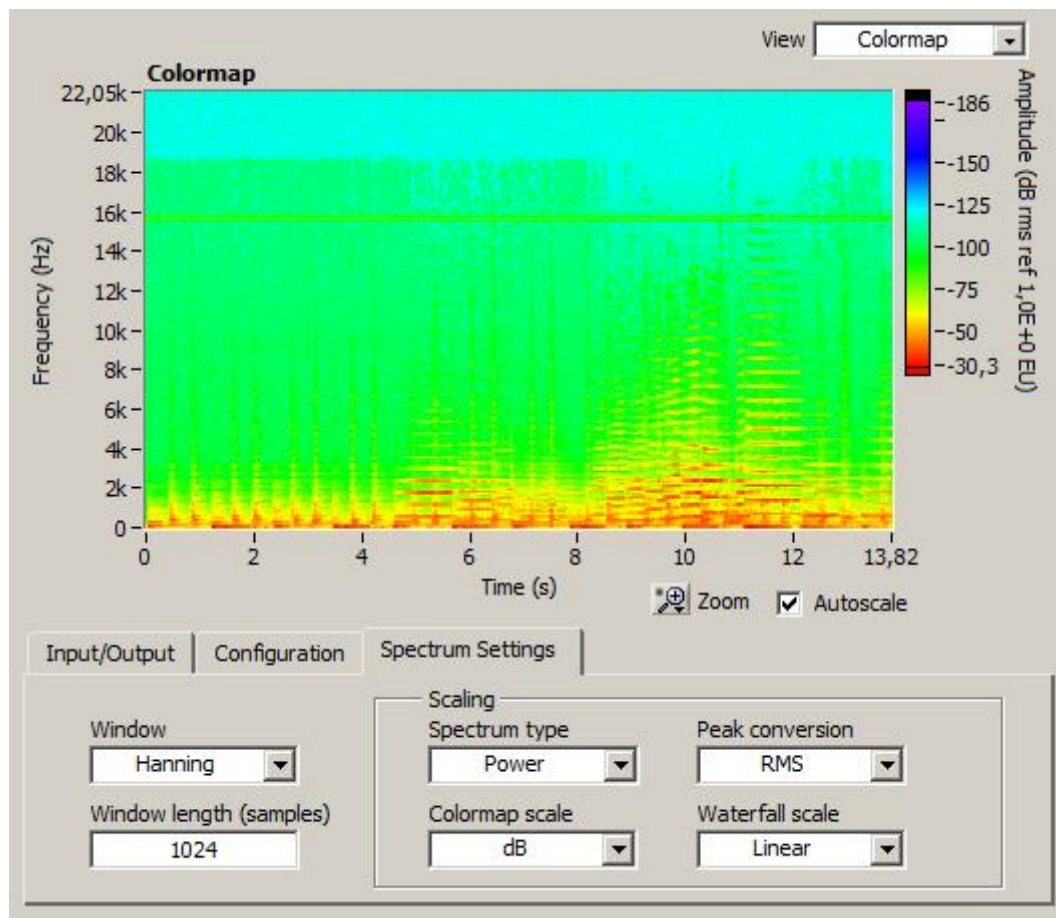
Obr. 23: Schéma zapojení uložení a přehrání zvukových dat

4.6 Výpočet spektrogramu

Spektrogram je získán pomocí bloku Spectral Map a zobrazen blokem Colormap. Po umístění bloku Spectral Map se zobrazí konfigurační nabídka s nastavitelnými parametry. Po potvrzení je blok automaticky nastaven a je zapotřebí připojit pouze vstupní data. Na obrázku číslo 24 a 25 je schéma zapojení bloku Spectral Map s blokem Colormap a nastavení bloku Spectral Map.



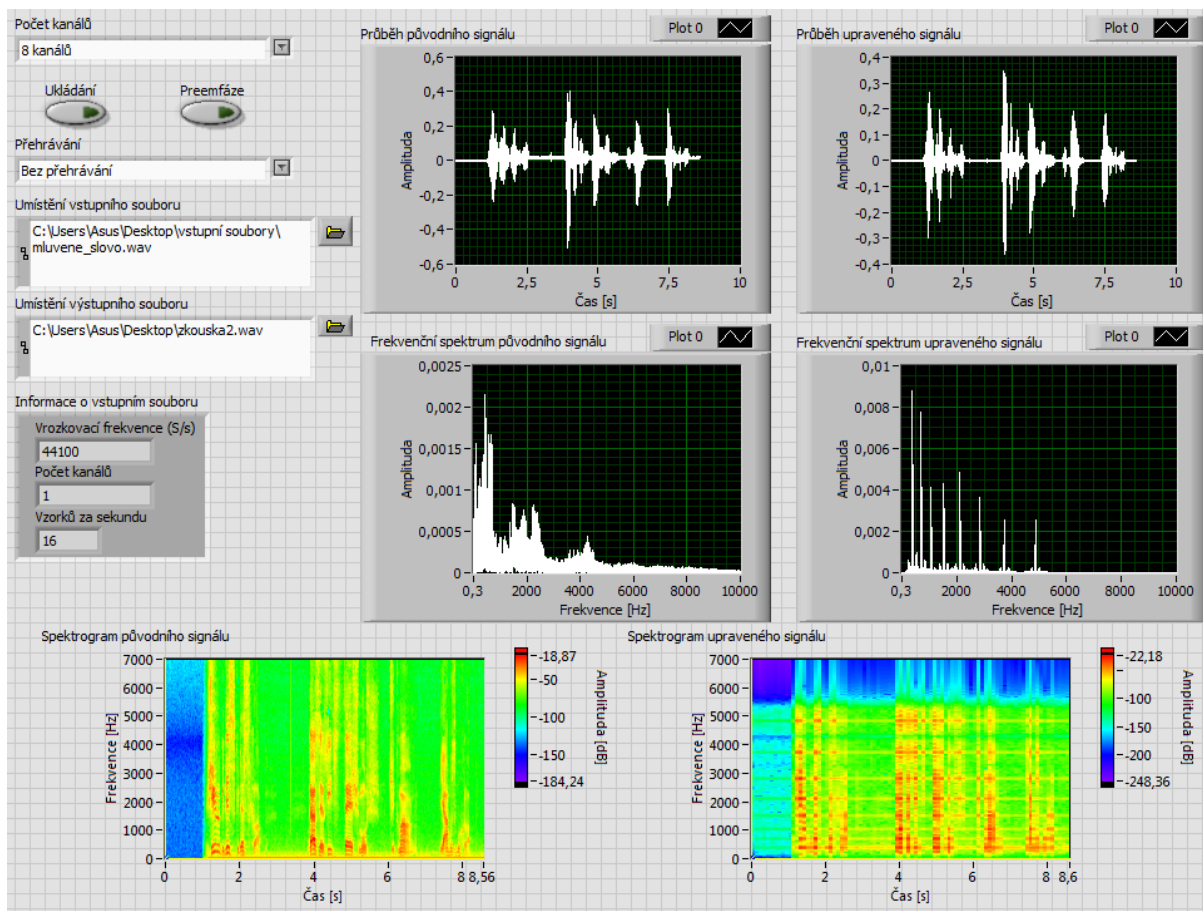
Obr. 24: Schéma zapojení výpočtu spektrogramu a zobrazení spektrogramu



Obr. 25: Nastavení bloku Spectral Map

4.7 Popis předního panelu

Přední panel je forma uživatelského rozhraní, které kontroluje funkční průběh prvků na blokovém diagramu. Je rozdělen na levou (kontrolní) část a pravou (indikační) část. V kontrolní části je možné nastavit počet použitých kanálů, nastavit režim přehrávání a ukládání, vypnout a zapnout filtr preemfáze. Důležitou částí je nastavení správného umístění vstupního a výstupního souboru ve formátu WAV. Pod těmito kontrolními prvky se nachází informační okno o vstupním souboru. Indikační část zobrazuje časový průběh amplitudy původního signálu a upraveného signálu. Pod těmito grafy jsou umístěny grafy frekvenčních spekter původního a upraveného signálu. V nejnižší části jsou vykreslovány spektrogramy původního a upraveného signálu. Na obrázku číslo 26 je snímek předního panelu.



Obr. 26: Ukázka použitého předního panelu

4.8 Obsluha programu

Při otevření programu je jako výchozí funkce z palety nástrojů nastavena ikona Operate Value. Tím je zajištěna možnost ovládání kontrolních prvků na předním panelu. Nejprve je nutné vybrat umístění vstupního souboru pomocí prvku ‚Umístění vstupního souboru‘. Pokud chceme výstupní soubor uložit, zadáme jeho umístění pomocí prvku ‚Umístění výstupního souboru‘ a zmáčkeme tlačítko ‚Ukládání‘. Prvek ‚Počet kanálů‘ umožňuje volbu typu stimulace (8, 16 a 22 kanálů). Prvek ‚Přehrávání‘ umožňuje volbu přehrávání (Bez přehrávání, Přehrávání původního zvuku, Přehrávání upraveného zvuku, Přehrávání původního i upraveného zvuku). Tlačítko ‚Preemfáze‘ zapíná/vypíná filtr preemfáze. Po nastavení požadovaných parametrů je nutné stisknout tlačítko ‚Run‘ (bílá šipka) na horní liště předního panelu. Následně se zobrazí charakteristiky v prostoru indikačních grafů. Tento proces je možné opakovat libovolně dlouho.

4.9 Použité vstupní WAV soubory

Pro simulaci bylo použito několika konkrétních audio souborů. S přihlédnutím na pokrytí možných poslechových situací. Poslech mluveného slova v klidném a hlučném prostředí (zvuk stavby v pozadí). Poslech hudby se zpívaným textem a poslech vážné hudby. V následující tabulce budou uvedeny popisné informace o použitých audio souborech.

Tabulka 5: Tabulka s informacemi o použitých WAV souborech

Název souboru	Délka [s]	Počet kanálů	Vzorkovací frekvence [Hz]	Velikost souboru	Vzorků za sekundu
hudba_vazna	13	2	44100	2,32 MB	16
hudba_zpivana	10	2	44100	1,70 MB	16
mluvene_slovo	8	1	44100	738 kB	16
mluvene_slovo_hluk	8	2	44100	1,44 MB	16
mluvene_slovo_vice_hluku	8	2	44100	1,44 MB	16

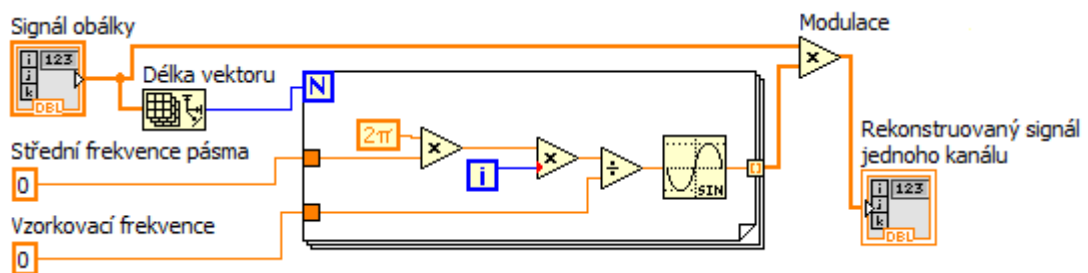
Soubor hudba_vazna.wav je převzat z [25]. Soubor hudba_zpivana.wav je převzat z [11]. Soubor mluvene_slovo.wav je převzat z [24]. Soubory mluvene_slovo_hluk.wav a mluvene_slovo_vice_hluku.wav jsou kombinací souboru mluvene_slovo.wav a hluku ze stavby o různé intenzitě.

5 ZPĚTNÁ REKONSTRUKCE ŘEČI

Zpětnou rekonstrukci řeči je možné provést pomocí součtu sinusových signálů. Detekovaná obálka z každého kanálu se vynásobí sinusovým signálem o frekvenci rovnající se střední frekvenci filtračního pásma. Výsledný rekonstruovaný signál bude složen z několika různých sinusoid. Zpětná rekonstrukce bude provedena pomocí vzorce:

$$s(t) = \sum_{k=1}^N A_k(t) \sin(2\pi \cdot f_k \cdot t), \quad (3)$$

kde $A_k(t)$ je amplituda obálky k -tého pásma a k je počet pásem analyzující banky filtrů. Frekvence f_k jsou středními frekvencemi každého k -tého filtračního pásma. Fáze původního spektra není pro rekonstrukci použita, protože kódovací strategie ACE a CIS přenášejí informaci pouze o modulu spektra a ne o fázi. [12] Na obrázku číslo 27 je ukázáno schéma zapojení zpětné rekonstrukce v LabVIEW.



Obr. 27: Ukázka schéma zapojení zpětné rekonstrukce pomocí modulace sinusového signálu pro jeden kanál.

Informace o amplitudě obálky je přiváděna na součin s generovaným sinusovým signálem daných parametrů. Pomocí bloku Array size je zjištěna délka vstupních dat, která následně nastaví počet iterací for cyklu. Všechny násobené obálky je poté nutno sečíst. Dále je možné získaný signál rovnou přehrát, nebo uložit.

6 VYHODNOCENÍ NAVRŽENÉHO ŘEŠENÍ

V této kapitole proběhne vyhodnocení získaných výsledků. Hlavními výstupními parametry jsou spektrogramy a uložené zvukové stopy.

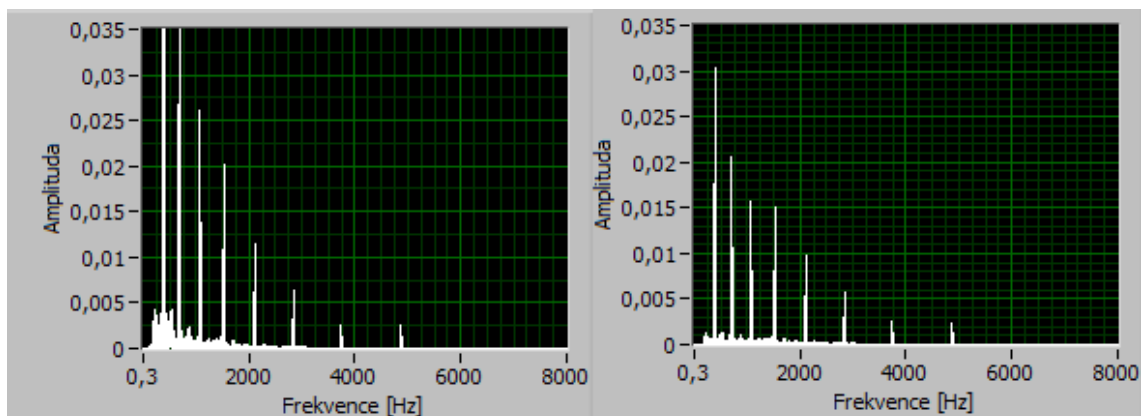
6.1 Vliv preemfáze

Princip preemfáze již byl v textu popsán. V programu je možnost volby mezi použitím a nepoužitím tohoto prvku. Bylo provedeno poslechové porovnání při zapnutém a vypnutém prvku preemfáze. Z měření lze konstatovat, že změna výstupního signálu oproti vstupnímu při použití prvku preemfáze je znatelná. Nejzřetelnější změnu bylo možné pozorovat u hudby se zpívaným textem pro 8 a 22 kanálové zapojení. Rozdíly v ostatních porovnáních jsou velmi slabé, nebo neznatelné. K lepšímu vjemu mluveného slova vedlo použití preemfáze u nahrávek mluveného slova smíchaných s hlukem stavby a to hlavně pro 8 a 16 kanálové zapojení. Tabulka číslo 6 obsahuje shrnutí získaných výsledků poslechovou metodou.

Tabulka 6: Shrnutí porovnání získaných zvukových souborů se zapnutým a vypnutým prvkem preemfáze

Název souboru	Počet kanálů	Rozdíl s použitím preemfáze	Poznámka
hudba_vazna	8	slabě znatelný	
hudba_vazna	16	slabě znatelný	
hudba_vazna	22	neznatelný	
hudba_zpivana	8	velmi zřetelný	
hudba_zpivana	16	slabě znatelný	
hudba_zpivana	22	velmi zřetelný	
mluvene_slovo	8	neznatelný	
mluvene_slovo	16	slabě znatelný	
mluvene_slovo	22	slabě znatelný	
mluvene_slovo_hluk	8	slabě znatelný	
mluvene_slovo_hluk	16	slabě znatelný	
mluvene_slovo_hluk	22	neznatelný	
mluvene_slovo_vice_hluku	8	slabě znatelný	velmi špatně srozumitelné
mluvene_slovo_vice_hluku	16	slabě znatelný	velmi špatně srozumitelné
mluvene_slovo_vice_hluku	22	slabě znatelný	velmi špatně srozumitelné

Rozdíl při použití prvku preemfáze je patrný i z porovnání grafických zobrazení frekvenčních spekter. Signál zpracovaný při použití prvku preemfáze vykazuje ve frekvenčním spektru úbytek míry zastoupení frekvenčních složek (v rozmezí zhruba 0 – 1400 Hz) oproti frekvenčnímu spektru signálu zpracovaného bez prvku preemfáze. Srovnání je ukázáno na obrázku číslo 28.



Obr. 28: Ukázka vlivu použití bloku preemfáze. Vlevo bez preemfáze, vpravo s preemfází. Vstupní soubor - hudba_zpivana v 8 kanálovém zapojení

Lze konstatovat kladný vliv použití bloku preemfáze. Zlepšení při jeho použití je znatelné i navzdory jednoduchosti použitého zapojení. Pro zlepšení kvality výstupního zvuku by bylo vhodné realizovat zapojení s použitím kvalitního mikrofonu a komplexnějšího předzpracování zvuku. Schémata pro předzpracování zvuku se mi bohužel nepodařilo vyhledat. Z velké části se jedná o „know-how“ konkrétních společností vyrábějících kochleární implantáty, nebo k nim příslušné komponenty. Výrobci a obecná pojednání uvádějí, že je použití preemfáze vhodné i pro potlačení šumu. To se částečně potvrdilo v simulaci pro soubor s mluveným slovem zaneseným hlukem stavby. Zlepšení však nebylo výrazné.

6.2 Porovnání zapojení s různým počtem kanálů

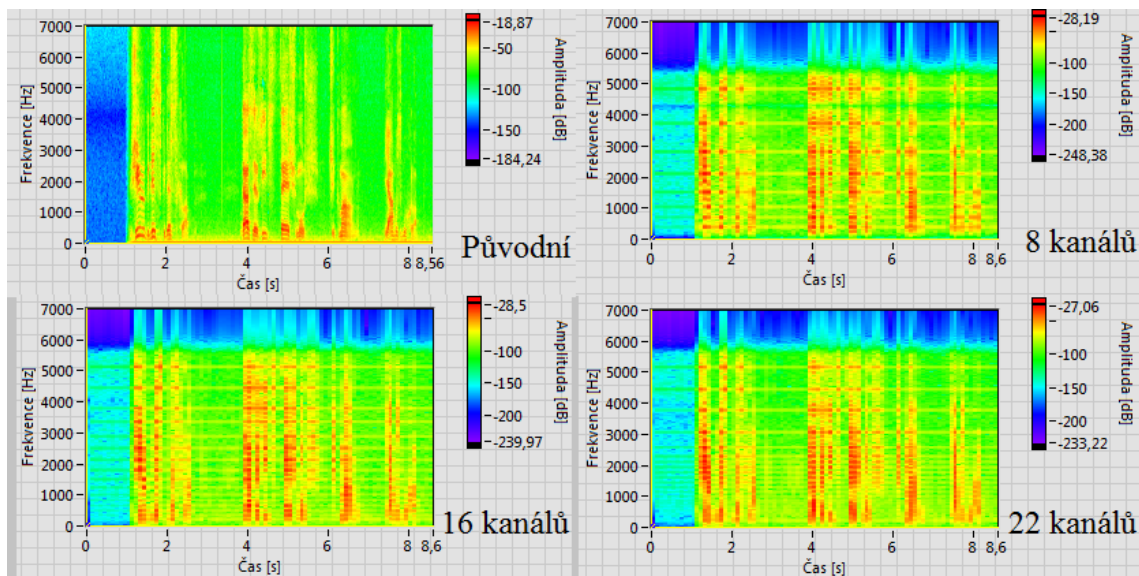
V simulaci byla použita tři možná zapojení lišící se počty kanálů (8, 16, 22). Porovnání se vstupním signálem bylo provedeno pomocí získaných spektrogramů a výstupních zvukových souborů. Poslechové shrnutí je uvedeno v tabulce číslo 7.

Tabulka 7: Porovnání zapojení s různým počtem kanálů

Název souboru	Počet kanálů	Sluchový vjem
hudba_vazna	8	vjem rytmu jinak velmi nepříjemné
hudba_vazna	16	vjem rytmu a částečně i melodie, jinak velmi nepříjemné
hudba_vazna	22	vjem rytmu a částečně i melodie, jinak velmi nepříjemné
hudba_zpivana	8	vjem rytmu a melodie, nesrozumitelný text
hudba_zpivana	16	vjem rytmu a melodie, srozumitelný text
hudba_zpivana	22	vjem rytmu a melodie, srozumitelný text, horší než zapojení s 16 kanály
mluvene_slovo	8	srozumitelná řeč nečitelná barva hlasu
mluvene_slovo	16	srozumitelná řeč, čitelná barva hlasu
mluvene_slovo	22	srozumitelná řeč, čitelná barva hlasu
mluvene_slovo_hluk	8	srozumitelná řeč
mluvene_slovo_hluk	16	srozumitelná řeč, lepší než zapojení s 8 kanály
mluvene_slovo_hluk	22	srozumitelná řeč, horší než zapojení s 16 kanály
mluvene_slovo_vice_hluku	8	velmi špatně srozumitelné až nesrozumitelné
mluvene_slovo_vice_hluku	16	velmi špatně srozumitelné až nesrozumitelné, lepší než zapojení s 8 kanály
mluvene_slovo_vice_hluku	22	velmi špatně srozumitelné až nesrozumitelné, horší než zapojení s 16 kanály

Obecně lze konstatovat, že simulace má nejlepší výsledky při zpracování mluveného slova bez šumu. Při 8 kanálovém zapojení je řeč srozumitelná, ale postrádá barvu hlasu. Zapojení s 16 a 22 kanály vykazují zhruba stejnou srozumitelnost řeči, ale barva hlasu je výraznější. Výsledky 16 a 22 kanálového zapojení jsou srovnatelné. S rostoucí mírou šumu se srozumitelnost řeči zhoršuje. Při poslechu hudby se zpívaným textem je text písně nesrozumitelný s 8 kanálovým zapojením. Při použití 16 a 22 kanálového zapojení je textu rozumět. Problém však představuje vjem melodie a rytmu hudby. Tyto vjemy jsou přirozenému poslechu velmi vzdálené a působí nepříjemný pocit. Nejhorší výsledky vykazuje simulace pro zpracování vážné hudby. Přirozený poslech by měl být plynulý a harmonický. Zvuk zpracovaný simulací, ale zní spíše jako rachocení hrnců. To je v souladu s reálně používanými kochleárními implantáty, které pro zpracování a reprezentaci hudby používají speciální algoritmy a řečové procesory. Navzdory očekávání je zapojení s 22 kanály v některých případech horší než zapojení s 16 kanály. To může být dáno vlastním návrhem frekvenčních rozsahů pro zapojení s 22 kanály. Zároveň je nutné brát v úvahu jednoduchost zapojení. Jednotlivé paralelní větve jsou realizovány pouze pomocí jedné pásmové propusti. Vhodné by bylo implementovat komplexnější banku filtrů.

Další informaci přináší porovnání spektrogramů vstupního souboru a výstupních souborů. Na obrázku číslo 29 je srovnání spektrogramu vstupního signálu ze souboru `mluvene_slovo` se spektrogramy signálů zpracovaných 8, 16 a 22 kanálovým zapojením.



Obr. 29: Srovnání spektrogramů souboru mluvne_slovo a jeho úprav 8, 16 a 22 kanálovým zapojením

Z obrázku číslo 29 je na první pohled patrný vliv průchodu bankou pásmových propustí. Na spektrogramech výstupních souborů se to projevuje jako horizontální linky táhnoucí se spektrogramem, které vykazují maximální frekvenční hodnoty pro středové frekvence jednotlivých pásem. Pro posouzení je důležitá podobnost spektrogramů. Originálnímu spektrogramu se nejvíce podobají spektrogramy odpovídající 16 a 22 kanálovému zapojení. To je logické, jelikož více kanálů přináší větší frekvenční rozlišení.

ZÁVĚR

Předmětem této bakalářské práce je simulace zpracování zvuku pro kochleární implantáty. Samotné simulaci předchází úvodní část, ve které je rozebrán teoretický úvod týkající se anatomie a fyziologie lidského sluchu. Je uvedeno shrnutí sluchových poruch podle několika kategorií a jejich možných kompenzací. Dále jsou popsány základní součásti a charakteristiky kochleárních implantátů s podkapitolou týkající se řečových strategií. V práci jsou popsány základní součásti pro zpracování lidské řeči v prostředí LabVIEW.

Vytvořená simulace načítá vstupní data z audio WAV souborů. Použity byly soubory s vzorkovací frekvencí 44 100 Hz. Podrobné informace o vstupních WAV souborech jsou uvedeny v tabulce číslo 5. Simulace umožňuje poslech v několika režimech a uložení výstupních dat ve formě WAV souboru. Přehrát je možné vstupní i výstupní signál, nebo postupně oba dva. Zpracované vstupní soubory jsou součástí elektronické přílohy této práce.

Realizace zapojení vychází z řečové strategie CIS, která je používána ve většině kochleárních implantátů. Zapojení je doplněno o vypínatelný blok preemfáze. Za tímto blokem následuje větvení do filtračních pásem o příslušných frekvencích rozsazích a počtu kanálů. V práci je použito 8, 16 a 22 kanálové zapojení. Frekvenční rozsahy pro zapojení s 8 a 16 kanály se mi podařilo dohledat v literatuře. Rozsahy pro 22 kanálové zapojení jsem stanovil experimentálně dle vlastní úvahy. Obálka signálu je detekována pomocí usměrnění a filtrace dolní propustí. Zpětná rekonstrukce signálu probíhá podle návodu ze článku uvedeného ve zdrojové literatuře. V každém kanálu dochází k modulaci sinusového signálu s frekvencí střední frekvence daného kanálu obálkou zpracovávaného signálu. Data ze všech kanálů jsou poté sečtena a přehrána, nebo uložena.

Vyhodnocení je provedeno na základě porovnání výstupních audio souborů a na základě porovnání spektrogramů. Porovnána byla i simulace s vypnutým a zapnutým blokem preemfáze. Z hodnocení je zřejmé, že zapnutý blok preemfáze zlepšuje kvalitu a srozumitelnost výstupního signálu a to hlavně pro hudbu se zpěvem a mluvené slovo. Při porovnávání zapojení s rozdílným počtem kanálů byl blok preemfáze aktivní. Nejhorších výsledků dosáhlo 8 kanálové zapojení. To odpovídá praktické realitě. 16 a 22 kanálové zapojení bylo zhruba na stejné úrovni, ale v některých případech bylo 22 kanálové zapojení horší. Tento jev může být způsobem vlastním návrhem frekvenčních rozsahů pro 22 kanálové zapojení. Nejlépe simulace zpracovává mluvené slovo bez šumu a nejhůře vážnou hudbu. Je, ale nutné poznamenat, že i nejlepší výsledek je výrazně vzdálený sluchovému vjemu zdravého člověka. Z porovnání spektrogramů lze dedukovat rozdílnost při použití zapojení s odlišným počtem kanálů. Patrný je vliv průchodu bankou pásmových propustí. Na spektrogramech výstupních souborů se to projevuje jako horizontální linky táhnoucí se spektrogramem, které vykazují maximální frekvenční hodnoty pro středové frekvence jednotlivých pásem. Nejvyšší podobnost s originálním spektrogramem vykazují spektrogramy 16 a 22 kanálového zapojení.

Při posuzování simulace je nutné brát ohled na značnou jednoduchost zapojení. Jednotlivé bloky by bylo možné rozšířit. Například pásmová filtrace by mohla v každém

pásmu obsahovat navrženou banku filtrů. Zajímavá by mohla být i simulace přenosu dat pomocí radiových vln a následné generování bifazických pulsů s možností jejich odměření na oscioskopu. Pozornost by mohla být věnována i důkladnějšímu prostudování obvodů mikrofonu a forma zpracování zvuku na vstupu (směrové mikrofony atd.). Vhodné by mohlo být i vzájemné porovnání jednotlivých řečových strategií. Předností programu je však jeho jednoduchost a přehlednost, která je podstatná pro základní pochopení problematiky zpracování řečového signálu pomocí kochleárního implantátu. Vzhledem k výše uvedeným informacím považuji zadání práce za splněné.

SEZNAM POUŽITÉ LITERATURY

- [1] KRÁLÍČEK, Petr. Úvod do speciální neurofyzologie. Praha: Galén, 2011.
- [2] LEJSKA, Mojmír. Základy praktické audiologie a audiometrie: Učební text. 1. Brno: Institut pro další vzdělávání pracovníků ve zdravotnictví, 1994
- [3] Moctezuma, A., Tu, J. An Overview of Cochlear Implant Systems. USA: University of Illinois, Urbana-Champaign, 2011
- [4] SOMEK, B., FAJT, S., DEMBITZ, A., IVKOVIĆ, M., OSTOJIĆ, J. Coding Strategies for Cochlear Implants. Chorvatsko, 2006
- [5] CLARK, G. Cochlear implants, fundamentals and applications. New York: Springer NY, 2003
- [6] NIE, K., STICKNEY, G., ZENG, F. G., Encoding Frequency Modulation to Improve Cochlear Implant Performance in Noise, Trans. on Biomedical Engineering, vol. 52, no 1, pp 64-73, 2005
- [7] STICKNEY, S., ASSMANN, F., CHANG, J., ZENG, F. G. Effects of cochlear implant processing and fundamental frequency on the intelligibility of competing sentences. USA: Acoustical Society of America, 2007
- [8] Náповěda LabVIEW [online]. [cit. 28-5-2014]. Dostupné z: <http://zone.ni.com/reference/en-XX/help/371361G-01/>
- [9] CAPEL, V. Public Address Systems, USA: Focal Pr, 1992
- [10] Zdroj obrázku číslo 3 [online]. [cit. 28-05-2014]. Dostupné z: <http://media-3.web.britannica.com/eb-media/98/14298-004-99934987.gif>
- [11] Zdroj souboru hudba_zpivana. Tři čuníci – Jaromír Nohavica [CD], Monitor, 1994
- [12] VONDRÁŠEK, M. Zpětná rekonstrukce řeči a virtuální elektroda pro kochleární implantáty. České vysoké učení v Praze, Fakulta elektrotechnická, 2005
- [15] LOIZOU, P., DORMAN, M. On the Number of Channels Needed to Understand Speech. Journal of the Acoustical Society of America, 1999
- [16] Zdroj obrázku číslo 4 [online]. [cit. 27-05-2014]. Dostupné z: <http://www.neuroreille.com/promenade/english/audiometry/audiometry.htm>
- [17] HORÁKOVÁ, R. Sluchové postižení. Úvod do surdopedie. Praha, 2012, ISBN: 978-80-262-0084-0

- [18] LAWAND, S. N., FRENCH, J. P. Long term Cochlear Implant electrode improvement for stimulation and sensing neuronal activity. Electron. Instrum. Lab., Delft Univ. of Technol., Delft, Netherlands, 2012
- [19] NOROGIA, M., TOGNOLA, G., SVELTO, C. Measurement of Electrode Current Pulses From Cochlear Implants. IEEE Transactions on instrumentation and measurement, vol. 54, no. 5, 2005
- [20] CHOI, C. T. M., LEE, Y. A Review of Stimulating Strategies for Cochlear Implants. National Chiao Tung university, National Taichung University of Education, Taiwan, 2012, ISBN: 978-953-51-0582-4
- [21] Zdroj obrázků číslo 6, 7 a 8 [online]. [cit. 28-05-2014]. Dostupné z: <http://ecs.utdallas.edu/loizou/cimplants/tutorial/>
- [22] Zdroj obrázku číslo 10 [online]. [cit. 28-05-2014]. Dostupné z: <http://cdn.intechopen.com/pdfs-wm/41194.pdf/>
- [23] Zdroj obrázku číslo 13 [online]. [cit. 28-05-2014]. Dostupné z: <http://www.advancedbionics.com/>
- [24] Zdroj souboru mluvne_slovo [online]. [cit. 28-05-2014]. Dostupné z: <https://www.vutbr.cz/elearning/mod/resource/view.php?id=259471/>
- [25] Zdroj souboru hudba_vazna. Shostakovich – Šostakovič Dimitrij: Jazz Suites [CD]. 2009
- [26] Zdroj obrázku číslo 14 [online]. [cit. 28-05-2014]. Dostupné z: <http://www.cochlear.com/>

POUŽITÉ SYMBOLY A ZKRATKY

WHO – Světová zdravotnická organizace (World Health Organization)

KI – Kochleární implantát

CA – Compressed-Analog

SPEAK – Spectral Peak

CIS – Continous Interleaved Sampling

ACE – Advanced Combinational Encoder

FAME – Frequency-Amplitude-Modulation-Encoding Algorithm

WZCS – Wavelet Zero-Crossing stimulation

RTI – Research Triangle institute

PP – Pásmová propust

DO – Detekce obálky

ELEKTRONICKÁ VERZE PRÁCE

V této kapitole je popsán obsah disku přiloženého k textu bakalářské práce.

Tabulka 8: Souhrn souborů a složek na přiloženém disku

Název	Poznámka
Tomas_Kratochvil_BP.pdf	Text bakalářské práce ve formátu PDF
WAV vstupní	Složka obsahující vstupní WAV soubory
WAV výstupní	Složka obsahující výstupní WAV soubory
Literatura	Složka obsahující dostupné použité zdroje
Obrázky	Složka obsahující použité obrázky