

Západočeská univerzita v Plzni
Fakulta aplikovaných věd
Katedra informatiky a výpočetní techniky

Bakalářská práce

Zátěžový test pro analýzu únavy hlasu

Plzeň, 2010

Michael Kurdík

Originální zadání

Prohlášení

Prohlašuji, že jsem bakalářskou práci vypracoval(a) samostatně a výhradně s použitím citovaných pramenů.

V Plzni, dne

Michael Kurdík

Poděkování

Tímto chci poděkovat svým rodičům, blízkým a přátelům za vkládanou důvěru, kteří mi byli vždy oporou.

Abstrakt

Práce se zabývá návrhem metod zátěžového testu hlasivek a výpočtem parametrů, které lze považovat jako ukazatel míry únavy hlasivek. Principem zátěžového testu je opakovaná dlouhodobá fonace vokálu „a“, tzv. salva. Ke sledovaným parametrům pak patří délka salvy, útlum fonace během salvy, délka nádechu mezi salvami a energie, která je v salvě obsažena. Samostatným problémem je detekce počátku a konce salvy. Metody jsou implementovány do experimentální programové aplikace včetně databáze pacientů a záznamů fonace, která je určena pro testování na ORL klinice FN Plzeň.

Klíčová slova: foniatrie, chraptivost hlasu, hlasová únava, paréza zvratného nervu, analýza signálu

Abstract

Stress test for analyzing of voice tired.

The Working solves a design of methods of stress test vocal cords and counting parameters, which it can be mark like measure vocal chords tiredness. The stress test is based on longtime repeated phonation vocal “a”, so-called burst. The watched parameters are length of burst, damping of phonation during burst, length of breath between bursts and energies, which is contained in burst. The main problem is detection of begin and end of burst. Methods are implemented in experimental program application including database of patients and phonation records, which is designed for testing in ORL clinic FN Pilsen.

Key words: phoniatory, hoarse voice, voice tired, recurrent nerve paresis, signal analysis

Obsah

1. Úvod	1
2. Otorhinolaryngologie – hlas	2
2.1 Otorhinolaryngologie	2
2.1.1 Foniatrie	2
2.2 Hlas	3
2.2.1 Vývoj hlasu	3
2.2.2 Hlasové ústrojí.....	3
2.2.3 Vznik hlasu	4
2.2.4 Poruchy hlasu a jejich příčiny	5
2.3 Vyšetřovací metody fonačního ústrojí	6
2.3.1 Vyšetření popisná – diagnostická.....	7
2.3.2 Vyšetření funkční – diagnostická	7
3. Zátěžový test	9
3.1 Metoda zátěžového testu	9
3.2 Zpracování a analýza dat	10
3.3 Filtrace signálu.....	11
3.3.1 Filtr FIR	11
3.3.2 Filtrování.....	12
3.3.3 Návrh filtru FIR dolní propust (LPF)	12
3.4 Transformace signálu	15
4. Metody detekce začátku a konce salvy.....	17
4.1 Metoda ruční detekce	18
4.2 Metoda počtu průchodů hranicí	18
4.3 Metoda počtu průchodů hranicí kvadrátu.....	21
4.4 Metoda aritmetického průměru.....	24
4.5 Chyba metod.....	27
5. Parametry salvy	29
5.1 Výpočet délky salvy (pauzy)	29
5.2 Výpočet energie fonací.....	29
5.3 Úbytek intenzity signálu během salvy	32
5.3.1 Detekce lokálního maxima absolutní hodnoty signálu	32
5.3.2 Lineární regrese maxim absolutní hodnoty signálu.....	33

5.3.3 Výpočet koeficientu tlumení	34
6. Implementace	36
6.1 data package	36
6.2 application package	37
6.3 presentation package	39
6.4 Datová úložiště	41
7. Zpracování a vyhodnocení výsledků	42
7.1 Cíle hodnocení výsledků a testování	42
7.2 Získaná data	43
7.2.1 Ruční detekce původní signál $s(k)$	43
7.2.2 Automatická detekce původní signál $s(k)$	43
7.2.3 Ruční detekce filtrovaný signál $s^*(k)$, $M+1=21$	45
7.2.4 Automatická detekce filtrovaný signál $s^*(k)$, $M+1=21$	45
7.3 Hodnocení	46
7.4 Náhledy z navrženého SW	48
8. Závěr	52

1. Úvod

Práce je zaměřena na oblast lékařství, tzv. foniatrii. Cílem práce je navrhnout a realizovat diagnostický systém pro zátěžový test pacientů s onemocněním hlasivek, zejména trpících parézou zvrtného nervu. Zvrtný nerv jako takový je důležitý pro tvorbu hlasu a správné dýchání. Celý test je založen na analýze záznamu hlasového projevu pacienta v podobě opakované dlouhodobé fonace vokálu „a“, tzv. salvy. V průběhu této fonace pacient namáhá s postupujícím časem a opakováním celé hlasové ústrojí, které je v širším smyslu tvořeno plicemi, dýchacími cestami, hlasivkami a rezonančními hlavovými dutinami. U pacientů, kteří trpí onemocněním hlasivek, spec. parézou zvrtného nervu, můžeme na základě již provedených měření vyslovit předpoklad, že se s postupující zátěží budou měnit některé parametry hlasu. Především se zátěž projeví nárůstem obsahu šumu v signálu, zkracováním délky salvy, rostoucím útlumem během jedné salvy a prodlužující se dobou nádechu mezi salvami.

Celý systém zátěžového testu je navržen jako experimentální k testovacím účelům tak, aby bylo možné systematicky testovat předpoklady, týkající se vývoje vybraných parametrů, na větším vzorku pacientů. V případě, že výsledky experimentů splní stanovené předpoklady, bude metoda zátěžového testu používána v klinické praxi na ORL klinice FN Plzeň. Úkolem je tedy do systému implementovat takové metody, pomocí kterých bude možné vyhodnotit stav hlasového orgánu s narůstající zátěží. Jednotlivé metody budou testovány a podle získaných výsledků bude vybrána pro praktické používání konečná kombinace metod. Samostatně řešeným problémem zátěžového testu je detekce počátku a konce salvy. Návrh metody detekce je prováděn s ohledem na kvalitu záznamů fonace u pacientů s onemocněním hlasivek. Signál obsahuje chrapot, a to včetně úseků mezi salvami při nádechu.

Programové vybavení má umožnit vytvoření záznamu opakované dlouhodobé fonace. Systém by měl umožňovat kromě detekce jednotlivých salv vypočítat příslušné parametry. Sem např. patří energie signálu v jedné salvě, délka salvy a délka pauzy mezi salvami a útlum signálu v salvě. Systém pracuje s lokální experimentální databází, s osobní kartou pacienta, zaznamenaným audiosignálem vyšetření a vypočítanými parametry.

2. Otorhinolaryngologie – hlas

2.1 Otorhinolaryngologie

Otorhinolaryngologie, či otorinolaryngologie, je spíše známá pod zkratkou ORL. Jedná se o lékařský obor, který se zabývá diagnostikou a léčbou poruch hlasu, řeči a sluchu. Vedle toho se tento obor zabývá léčbou nemocí hlavy a krku, a proto se tato disciplína celým a správným názvem označuje jako „Otorhinolaryngologie a chirurgie hlavy a krku“, která je mezi laiky obecně známá jako tzv. „ušní, nosní a krční“.

Otologie – věda o chorobách ucha.

Rinologie – věda o chorobách nosu a vedlejších nosních dutin.

Laryngologie – věda o chorobách hrtanu

Problémy člověka spojené s uvedenými oblastmi můžou být vrozené nebo získané v průběhu života¹ bez rozdílu věku či pohlaví. Takovéto potíže potom mohou být jen dočasné nebo dlouhodobějšího rázu, kdy záleží na charakteru a příčinách vzniklých poruch.

Důvodem k vyšetření ovšem nemusí být až vzniklé problémy a je jistě vhodné myslet na prevenci a případným potížím předcházet. Pravidelné prohlídky se především týkají lidí, jako jsou např. tzv. hlasoví profesionálové², kteří hlas daleko častěji oproti „běžnému“ člověku namáhají nebo jen ti, kteří pracují v hlučném prostředí, a pravděpodobnost vzniku poruch je tak daleko vyšší.

2.1.1 Foniatrie

Foniatrie je jedním z nadstavbových oborů ORL, který se zabývá vyšetřováním, léčbou a rehabilitací poruch sdělovacích funkcí člověka jako jsou: lidský hlas, řeč a sluch.

Využívá běžných léčebných metod (medikamentosních, chirurgických), ale zároveň speciálních reedukačních či edukačních metod, kdy se jedná zejména o nácvik správné hlasové tvorby (provádí se především u hlasových profesionálů, a to jak jednotlivě, tak ve skupinách).

¹ např. zánětem, úrazem, neurologickým a jiným onemocněním

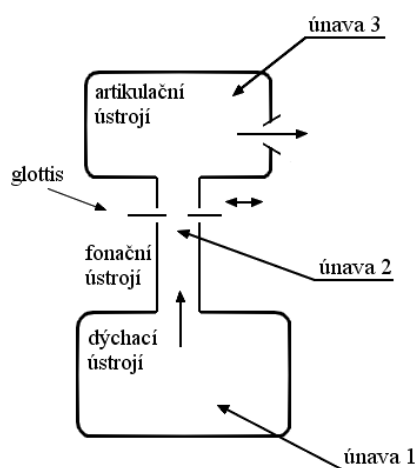
² zpěváci, moderátoři, atd.

2.2 Hlas

2.2.1 Vývoj hlasu

To jakým hlasovým projevem daná osoba disponuje, souvisí s jejím fyziologickým vývojem od okamžiku jejího narození (pokud nyní v úvahu nebereme onemocnění). Mezi jedinci není do doby puberty výraznější intersexuální rozdíl, a proto v dětských letech zvládají vysoké tóny i chlapci, což jim naopak v době puberty činí obvykle výrazné problémy. V době dospívání dochází u chlapců k růstu hrtanu a tím k prodloužení hlasivek, díky čemuž se jejich hlas prohlubuje. Tzv. „mutování“, které je spojeno s dospíváním každého chlapce, způsobují překrvené hlasivky a prosáklá sliznice, a proto zabarvení hlasu přeskakuje z dětské na dospělého výšku hlasu. Během dospívání dívek není nárůst hrtanu tak výrazný a proto nemají výšku hlasu posazenou tak nízko. V dospělosti se výška mužského hlasu pohybuje v rozmezí F-H, tj. 87 až 123 Hz, a u žen o oktávu výše f-c, tj. 174 až 247 Hz, viz [4], u nichž se v období zvané menses může hlasový rozsah zmenšit díky překrveným hlasivkám a prosáklé sliznici. S přibývajícím věkem pak dochází ke změnám na hrtanovém skeletu a k tzv. stařeckým změnám na hlase, které jsou vyvolány hormonálními změnami. Výška mužského hlasu se tak stává vyšší, kdežto u žen se naopak prohlubuje.

2.2.2 Hlasové ústrojí



Obr. 2.1 – Blokové schéma hlasového ústrojí.

fonace podléhá únavě. Jedním z parametrů, který únavě podléhá, je vitální kapacita plic, čímž dochází ke zmíněnému ovlivnění hlasového projevu jedince. Vitální kapacita plic pak představuje objem vzduchu, který může být vydechnut po maximálním vdechu a určuje i energii dýchacího ústrojí.

Hlasové ústrojí se skládá z ústrojí dýchacího, které dodává potřebnou energii v podobě proudu vzduchu o určitém tlaku. To je tvořeno plicemi, průduškami, průdušnicí, mezižeberními svaly a bránicí. Dalšími částmi hlasového ústrojí je ústrojí fonační, které je tvořeno hrtanem s rezonátorem³, a ústrojí artikulační, které se skládá z hltanu a dutiny ústní a nosní. Lidský hlas pak vzniká průchodem tzv. hrtanového tónu hlavovými dutinami.

Obrázek obr. 2.1 schematicky ukazuje podobu hlasového ústrojí. Každá z těchto tří částí ovlivňuje artikulační, v našem případě fonační, schopnosti člověka a s prodlužující se délkou

³ Jsou to tzv. dutiny supraglottis a subglottis (*glottis* – lat. hlasivky)

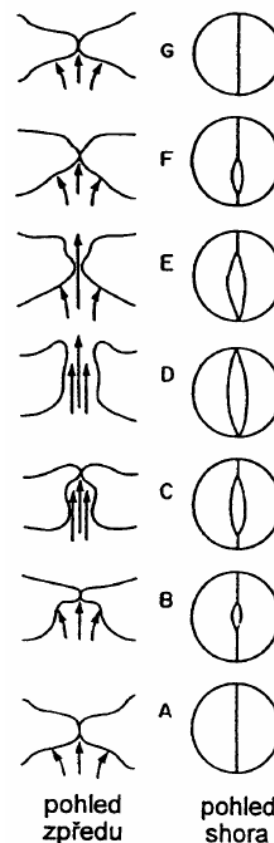
Dalším parametrem, který únavě podléhá, je svalová únava fonačního a artikulačního ústrojí. Protože se však únava při dlouhodobé fonaci projevuje s rozdílným zpožděním, vzhledem k délce fonace, lze v dalším některé projevy únavy zanedbat, viz dále.

2.2.3 Vznik hlasu

Existuje řada teorií popisujících vznik hlasu. Například tzv. „teorie zdroje a filtru“, kterou popisuje ing. Švec v publikaci [9] nebo tzv. „neurodynamická teorie“, která patří k nejrozšířenějším a bude jí tak věnována pozornost i zde.

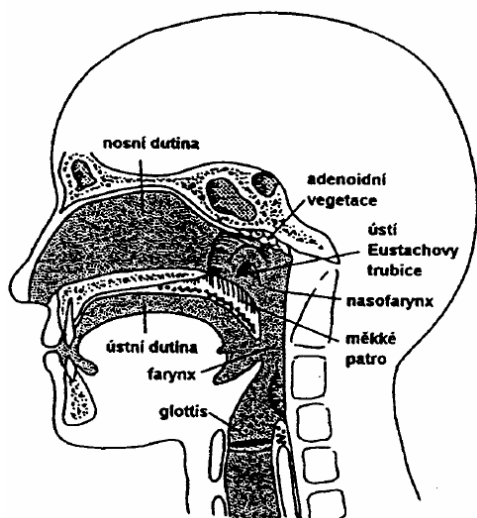
Zvuk hlasu se primárně vytváří v hrtanu. Spočívá ve zvyšování a snižování tlaku vzduchu v oblasti hrtanu, který je následně modifikován v nasadní trubici. [4]

Proud vzduchu z dýchacího ústrojí prochází skrze hlasovou štěrbinu. Periodickým kmitáním hlasivek dojde k jeho resonanci (díky přerušování hlasivkami) a vznikne tak hrtanový tón. Kmitání hlasivek je zajištěno díky jejich elasticitě, napětí a vzduchu procházející štěrbinou. Tlak nad hlasivkami, který se postupně zvyšuje, způsobí roztažení hlasivek a úniku vzduchu, kdy vzniká tzv. „Bernouliho efekt“. Tlak se postupně sníží, což způsobí, že svaly a šlachy vrátí roztažené hlasivky do původní polohy (obr. 2.2).



Obr. 2.2 – Schéma kmitání hlasivek. [7]

Nedílnou součástí pro vznik hlasu jsou zmíněné rezonanční dutiny, tedy dutiny mající tvar nálevky s proměnným tvarem a objemem, viz obr. 2.3. Dutiny supraglottis tvoří hltan, dutiny ústní, nosní a lebeční, subglottis pak průdušnice, průdušky a sklípky plic.



Obr. 2.3 – Schéma rezonančních dutin. [7]

Nosní dutina má tvar neměnný a je naladěna na určitý tón, kdy přidává k některým hláskám jen určité zabarvení. Hltan nemá pro tvorbu hlásek příliš velký význam a podílí se na jejich tvorbě jen minimálně. Hlavní vliv na vznik hlásek má dutina ústní, kdy se její objem a tvar mění pomocí jazyka a spodního patra díky pohyblivé dolní čelisti. Souhlásky vznikají tak, že se do cesty vycházejícímu vzduchu staví překážky, podle čehož se rozdělují na *závěrové, úžinové a polozávěrové*.

Závěrové souhlásky vzniknou nárazem vzduchu mezi rty (p, b), mezi jazyk a patro (t, d) nebo mezi kořen jazyka a měkké patro (k, g). Souhlásky „m“ a „n“ vznikají obdobně jako „k“ nebo „g“, ale vzduch uniká do dutiny nosní. Úžinové souhlásky vznikají průchodem vzduchu zúženým místem. Např. mezi dolním rtem a horními řezáky (f, v), mezi hrotem a výběžkem mezičelistí (s, z) nebo mezi částí hřbetu jazyka a patrem (j, ch, l, r). Souhláska „h“ vzniká přímo v hrtanu průchodem vzduchu štěrbínou mezi hlasivkami. Polozávěrové souhlásky (c, č) vznikají slabším uzávěrem na začátku a vytvořením úžiny na konci jejich vzniku.

Samohlásky jsou pak takovými hláskami, pro něž je při jejich vzniku charakteristickým rysem tón, kdy se jim na rozdíl od souhlásek nestaví do cesty překážka a nevzniká tak šum. Z tohoto důvodu jsou samohlásky (v našem případě „a“ či „e“) vhodné pro fonaci k vyšetření pacienta.

Z výše popsaných řádek je zřejmé, že vznik hlasu je poměrně složitý proces, na němž se podílí řada orgánů, které zároveň ovlivňuje mnoho faktorů. Pro zjednodušení našich případů a výpočtů jsme hlasové ústrojí schematicky rozdělili na tři oblasti (viz obr. 2.1). Vzhledem k délce fonace a únavnosti jednotlivých částí, budeme dále předpokládat, že parametry dýchacího a artikulačního ústrojí jsou konstantní. Mění se jen parametry ústrojí fonačního, které jsou dále předmětem testování pomocí zátěžového testu, kap. 3.

2.2.4 Poruchy hlasu a jejich příčiny

S hlasovými problémy se během života více či méně potýká každý. Tyto obtíže pak mohou být pouze dočasné, způsobené nadměrným namáháním hlasového ústrojí (např. nadměrný křik u dětí) nebo jen nachlazením. V horším případě, pak mohou být nějakého dlouhodobějšího až trvalého charakteru, kdy jsou tyto potíže většinou spojené s poraněním nebo onemocněním. V takovém případě pak, záleží na lokalitě a rozsahu onemocnění (poranění), od čehož se samozřejmě odvíjí doba a způsob léčby.

V každém případě se jedná o hlasové potíže, které jsou způsobeny postižením některých z výše popsaných oblastí hlasového ústrojí a u takového člověka pak nelze přepokládat dobrou artikulační (fonační) schopnost a ve většině případů lze očekávat šum v jeho zvukovém projevu - vokálu (samohlásky „a“ či „e“).

Poruch hlasu, které se mohou u člověka projevit, existuje celá řada. Dělí se podle mechanismu vzniku na organické a funkční. Některé z nich jsou uvedeny výčtem:

dysfonie – patologická změna hlasu (chrapot, přeskokování, změna výšky atd.)

afonie – ztráta zvučnosti hlasu je potom

afázie – projevuje se jako ztráta řeči při poškození mozku

mutismus – ztráta hlasu po fyzickém nebo psychickém traumatu.

S popsanými poruchami jsou pak spojeny diagnózy (určení příčiny poruchy hlasu). Jako příklad některých jsou uvedeny:

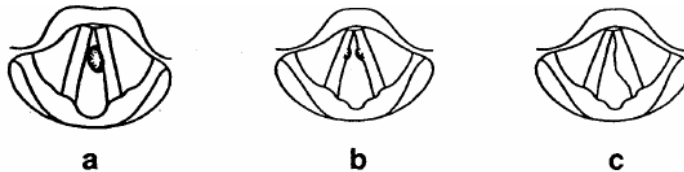
polyp – slizniční výchlípková vyplněná vazivem, je obvykle příčinou diplofonie⁴

uzlík – ohraničené ztlustění sliznice bránící úplnému závěru hlasové štěrbině

edém – dochází k prosáknutí podslizničního vaziva z různých příčin.

hematom – dochází ke krevnímu výlevu do hlasivky a tím se mění její hmota

cysta – zapouzdřený útvar uvnitř hlasivky, je příčinou nerovného okraje hlasivky



Obr. 2.4 – Ukázka, jak vypadá polyp (a), uzlík (b), edém (c). [12]

2.3 Vyšetřovací metody fonačního ústrojí

Základní metodou pro vyšetření hlasové funkce je subjektivní popis kvality hlasu. Tyto popisy ovšem vedou ke značné nejednotnosti v hodnocení kvality hlasu. [4] Kapitola zde rozebírá dosud známé fonační metody, které se omezují na nejednoznačnost v hodnocení hlasové kvality. Každá z uvedených metod je doplněna o stručný popis, kdy získané informace byly čerpány z knihy *Foniatrie a pedaudiologie II*, viz [4], s. 17 až 51.

Podle charakteru lze metody rozdělit do dvou skupin:

1. Vyšetření popisná – diagnostická
2. Vyšetření funkční – diagnostická
 - a. Metody založené na technice dýchání nebo proudu vzduchu procházením skrze štěrbinu.
 - b. Metody používající spektrální nebo akustickou analýzu.
 - c. Metody elektrofyziologické.

⁴ porucha projevující se zazníváním dvou hlasů

2.3.1 Vyšetření popisná – diagnostická

Nepřímá laryngoskopie

Základní vyšetřovací metoda. Dříve nebo dnes jen zřídka. K vyšetření se již nepoužívá zrcátka, ale optického vlákna, kdy je možné použít pevnou nebo ohebnou techniku, kde první zmíněná umožňuje přiblížení.

Přímá laryngoskopie

Používá se pouze při fonochirurgických zákrocích v hrtanu, často u velmi malých dětí. V některých případech se dává přednost nepřímé laryngoskopii.

Laryngostroboskopie

Na rozdíl od laryngoskopie, kdy je hrtan osvětlen pevným světlem, u této metody světlo kmitá. Pokud mají hlasivky stejnou frekvenci kmitů, jeví se jako nehybné. Jestliže se jejich frekvence liší, jsou hlasivky osvětleny pokaždé v jiném okamžiku kmitu a lidské oko pohyb vnímá zpomaleně.

Rtg vyšetření

Metoda založená na pořízení snímků hrtanu. Její uplatnění je například pro zobrazení cizích těles v hrtanu (obecně se příliš nevyužívá).

2.3.2 Vyšetření funkční – diagnostická

Ad) a

Pneumografie

Poměrně stará metoda pro vyšetření dýchaní během fonace. Výsledkem vyšetření je tomografie, u níž se především registrují dvě křivky – hrudní a břišní.

Pneumotachografie

Vyšetření je založeno na měření rychlosti proudu a množství vzduchu skrze hlasivky a subglottický tlak (pro některé ukazatele je třeba znát vitální kapacitu plic).

Ad) b

Sonagrafická analýza

Metoda založena na filtraci hlasového signálu akustický filtrem, kdy se filtr mění v nějakém frekvenčním intervalu a propustí vždy jen určitá frekvenční spektra.

Analýza dlouhého spektra (LTAS)

Analýza prakticky libovolně dlouhých vzorků. Na signál se aplikují třetino-oktávové nebo úzko-pásmové filtry a akustická energie jednotlivých pásem se počítá. Výsledkem je záznam dlouhodobého zprůměrovaného spektra.

Třídímenzionální analýza periodicity hlasu (3D-PAN)

Metoda, která rozdělí kvazi-periodický hlasový signál na jednotlivé periody, které se poskládají za sebe do 3D obrazu. Z grafického obrazu se pak určí, jestli se jedná o hlas normální nebo patologický.

Multidímenzionální analýza hlasu

Umožňuje analýzu řady kvalitativních parametrů hrtanového tónu, čímž určí základní kvality hlasu. Znáмым komerčním analyzátořem je program MDVP (Multi-Dimensional Voice Program), který počítá 33 parametrů.

Analýza okamžitého spektra

Umožňuje detailní studium spektra po jednotlivých časových úsecích. V poslední době se používá stále méně.

Hlasové pole

Umožňuje stanovení tzv. mluvního nebo zpěvního pole. V libovolném časovém rozsahu se měří mluvní nebo zpěvní projev. Výsledkem je obraz v jakém frekvenčním a dynamickém rozsahu se projev pohybuje.

Ad) c

Elektroglottografie

Zavedením elektrod nad křídly štítné chrupavky se registrují změny elektrického odporu v důsledku pohybu hlasivek. Na glottogramu je hodnocen čas závěru, fáze otevření a čas otevření hlasivek a hlasové štěrbiny (je vymezeno 6 bodů pro jednotlivé fáze pohybu hlasivek a hlasové štěrbiny).

Elektromyografie

Metoda, která je taktěž založena na zavedení elektrod do oblasti hrtanu. Metoda významně pomohla k objasnění aktivity vnitřních hrtanových svalů během fonace.

Jak již bylo v úvodu zmíněno, cílem práce je navrhnout SW aplikaci, kde použité metody poskytnou parametry pro určení vlivu zátěžového testu na hlasivky vyšetřovaného pacienta. Metodiku, která byla navržena a v rámci práce realizována, lze zahrnout do kategorie funkčních diagnostických vyšetření – tzv. akustické metody.

3. Zátěžový test

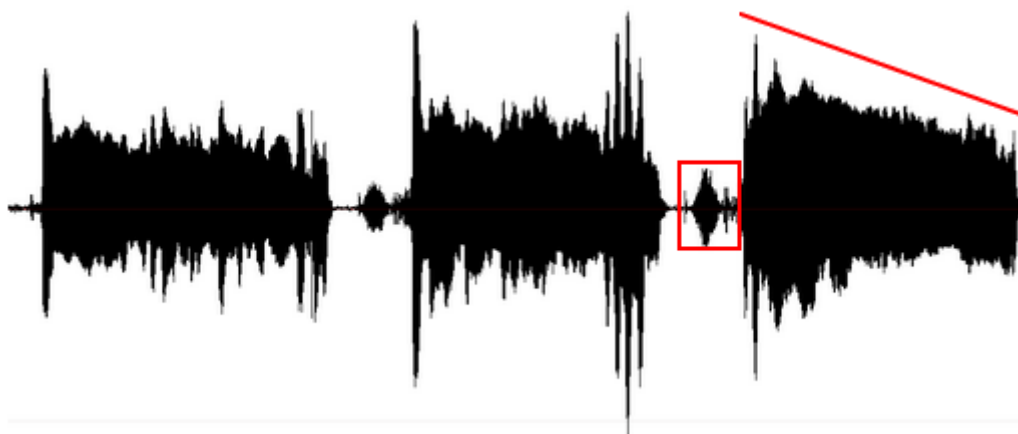
V této kapitole bude popsán vlastní zátěžový test, způsob zpracování signálu dlouhodobé fonace vokálu „a“, nalezení začátku a konce jedné fonace „salvy“ a návrh sledovaných parametrů pro hodnocení změn na hlasivkách vlivem zátěže.

3.1 Metoda zátěžového testu

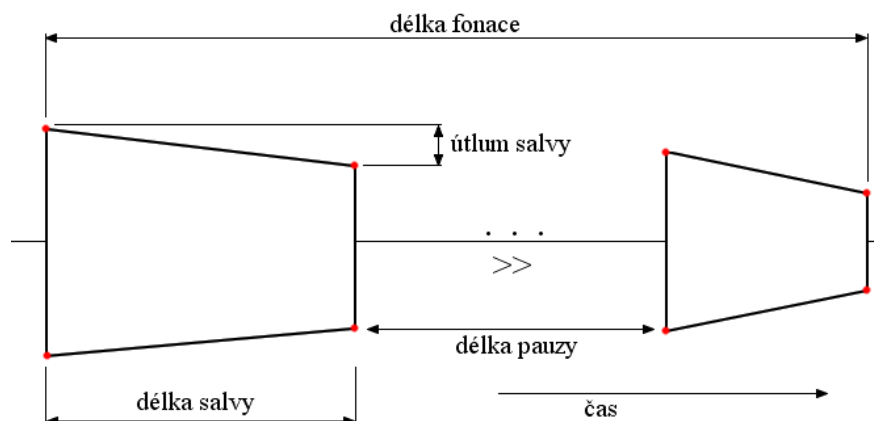
Celý test spočívá ve vytvoření hlasového záznamu $s(k)$ obsahující dlouhodobou fonaci vokálu „a“ v celé jeho délce. Dlouhodobou fonaci chápeme jako sérii několika po sobě jdoucích fonací vokálu „a“, tzv. salv. Salvu jako blok dat, který je v daný okamžik záznamu tvořen jedním nepřerušovaným vokálem „a“.

Záznam je prováděn pomocí elektretového mikrofону. Na výzvu lékaře započne pacient s hlasovým projevem, opakovanou dlouhodobou fonací vokálu „a“. Ukončení fonace se opět děje dle pokynů lékaře, který rozhoduje volbou opakování fonace o míře zátěže. Díky zvolenému počtu opakování souvislé fonace „a“, délce salvy a době nádechu mezi salvami je pak doba záznamu $s(k)$ proměnná a případ od případu se liší. Počet salv v dlouhodobé fonaci se obvykle pohybuje v rozsahu 10 až 20 salv. Od počtu salv se odvíjí i doba trvání jednotlivých vokálů či nádechů, kde značnou roli hraje již zmíněná únava a různé poruchy (viz kap. 2.2.4), kterými pacienti více či méně trpí.

Dle zkušeností z předchozích měření se délka fonace pohybuje mezi cca 30 až 90 sekundami. Jednotlivé salvy, které jsou tvořeny vokálem „a“ na jeden nádech a nutnost pauzy k nádechu pacienta za stálého nahrávání záznamu, stanovuje předpoklad podoby fonace, viz obr. 3.2. V celé délce fonace vznikají mezi jednotlivými salvami mezery, kdy se s postupem v čase délka a intenzita jednotlivých salv zmenšuje a délka pauz se naopak zvětšuje.



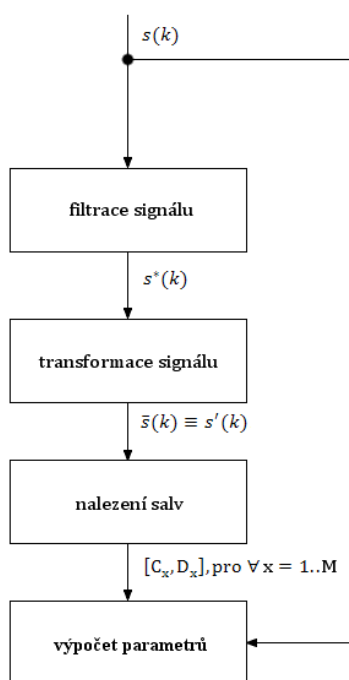
Obr. 3.1 – Ilustrace skutečného záznamu, se kterým se pracuje (diag. paréza zvrtného nervu). Poslední salva je názornou ukázkou útlumu energie. Zároveň ukazuje problém se zpracovávaným signálem (červený čtverec), kdy se projevuje výrazný chrapot při nádechu.



Obr. 3.2 – Předpokládaná podoba dlouhodobé fonace při zátěžovém testu.

3.2 Zpracování a analýza dat

Předpokládejme, že máme k dispozici digitalizovaný hlasový signál $s(k)$. Signál je získán vzorkováním spojitého signálu $s(t)$ vzorkovací frekvencí $f_{vz} = 44100$ Hz. Jedná se o 16-bitový⁵ monofonní audio záznam uložený ve formátu typu wav⁶. Tento signál je dále zpracováván dle schéma na obr. 3.3.



Obr. 3.3 – Schéma ukazující průběh analýzy získaného wav záznamu.

Signál $s(k)$ prochází předzpracováním s cílem, upravit signál tak, aby bylo možné detekovat počátek a konec salvy. Posloupnost jednotlivých kroků předzpracování je volena tak, aby byly odstraněny nežádoucí poruchy zejména v době pauzy. Signál je proto filtrován číslicovým filtrem typu dolní propust, kdy dostáváme dílčí (filtrovaný) signál $s^*(k)$. Dalším krokem v předzpracování signálu je následná transformace-posun signálu o střední hodnotu $E[s^*(k)]$ resp. $E[s(k)]$. Tím je odstraněna stejnosměrná složka. Tato operace může být provedena přímo na výchozí signál $s(k)$ bez nutnosti provádění filtrace. V obou případech dostáváme obecný transformovaný signál $\bar{s}(k)$.

Následující blok implementuje některou z metod pro detekci začátku a konce salvy, od nichž se odvíjí i výpočet parametrů (viz dále). Získané parametry jsou počítány z původního signálu $s(k)$ a slouží k ohodnocení vlivu zátěže na únavu hlasivek.

⁵ Jeden vzorek zvukového formátu wav je velký 16 bitů.

⁶ Formát uložených audio dat v PC.

3.3 Filtrace signálu

3.3.1 Filtr FIR

K filtraci signálu $s(k)$ je využíván nerekurzivní filtr FIR s Hammingovým symetrickým oknem typu dolní propust. FIR filtry jsou číslicové filtry s konečnou impulsní odezvou známé také jako tzv. filtry klouzavých průměrů.

Toto označení je používáno v úlohách modelování signálu a identifikace a souvisí se skutečností, že FIR filtr s libovolnými koeficienty provádí lineární kombinaci (vážený součet) $M+1$ vzorků vstupního signálu, který se s posunem o jeden časový krok opakuje pro každý nový vzorek. Lze tedy říci, že okno, kterým při filtraci vybíráme $M+1$ vzorků signálu, klouže po datech (proto klouzavé součty). [13]

Diferenční rovnice pro filtr řádu M má pak následující tvar

$$y(n) = b_0x(n) + b_1x(n-1) + \dots + b_Mx(n-M), \quad (3.1)$$

kde $x(n)$ je vstupní signál, b_0 impulzní odezva a $y(n)$ je výstupní filtrovaný signál. Rovnice 3.1 zároveň představuje konvoluci mezi posloupností hodnot vstupní proměnné $x(n)$ a posloupností hodnot impulzní odezvy (odezvy na jednotkový impuls) $h(n) = b_0, b_1, \dots$ [13]. Impulzní odezva je vyjádřena vztahem

$$h(n) = b_0\delta(n) + b_1\delta(n-1) + \dots + b_M\delta(n-M), \quad (3.2)$$

kde $\delta(n)$ jsou koeficienty přenosové funkce.

Hlavní výhodou FIR filtrů je jejich stabilita, protože vždy mají M -násobný pól a díky lineární fázi konstantní skupinové zpoždění $M/2$. Struktura filtru FIR (bez lineární fáze) je založena na principu zpoždovací linky s odbočkami pro násobičky. Filtr pak počítá vážený klouzavý průměr z $M+1$ posledních hodnot. Zpoždění, které při filtraci vzniká, nepředstavuje pro naši aplikaci problém, jelikož nejsme omezeni na práci s daty v reálném čase.

Uvažujeme-li, že použijeme filtr FIR s Hammingovým symetrickým oknem, potom H_w chápeme jako klouzavé okénko, které obsahuje vždy lichý počet koeficientů $M+1$ a vyjadřuje velikost filtru pro prvky $[h_0, h_1, \dots, h_M]$. Dolní propust se pak chová tak, že propouští nebo potlačuje složky signálu v určitém frekvenčním pásmu. Tím dochází k potlačení periodických nebo rušivých signálů od signálu užitečného.

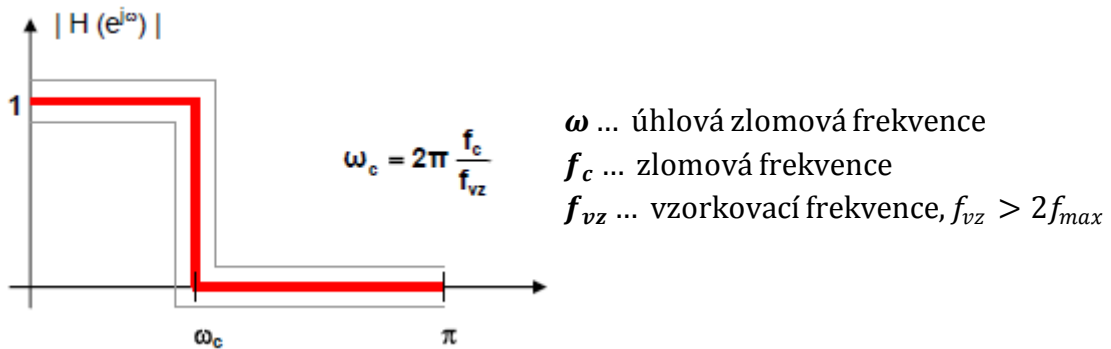
3.3.2 Filtrování

Konkrétní nastavení koeficientů Hammingova okna je v aplikaci počítáno pro zlomovou frekvenci $f_0 = 1,5\text{kHz}$. Velikost Hammingova okna má pak $M + 1 = 21$ prvků (resp. $M + 1 = 31$ a $M + 1 = 51$). Pro volbu hodnoty zlomové frekvence bylo využito zkušeností z práce *Metoda vyšetření hlasového pole a její aplikace*, viz [5], kap. *autokorelační funkce*, s. 4, kdy je jejich funkčnost v praxi ověřená na stejném typu pacientů. Zde se jedná u předzpracování signálu pro výpočet základní hlasivkové frekvence F_0 pomocí autokorelační funkce.

3.3.3 Návrh filtru FIR dolní propust (LPF)

Převzato z literatury [1], s. 50 a s. 59 až 63.

1. požadovaný průběh:



Obr. 3.4 – Požadovaný průběh filtrovaného signálu. [1]

2. požadovaná frekvenční charakteristika:

$$H(e^{j\omega}) = \sum_{n=-\infty}^{+\infty} h(n)e^{-jn\omega} \quad (3.3)$$

3. požadovaná impulsní odezva:

$$h(n) = \frac{1}{2\pi} \int_{-\pi}^{+\pi} H(e^{j\omega}) e^{jn\omega} d\omega \quad (3.4)$$

Filtr navrhujeme pro konečný počet členů n , kde platí: $n = 0, 1, 2, \dots, M$, tj. počet koeficientů filtru je $M + 1$ a impulsní odezva $h(n)$ je omezena pravoúhlým oknem. Jestliže toto okno nabývá hodnoty $w_{RECT}(n) = 1$ resp. $w_{RECT}(n) = 0$, pro $\forall n \in \langle 0; M \rangle$, resp. $\forall n \in \langle -M/2; +M/2 \rangle$, pak můžeme odvodit následující vztah:

$$H(e^{j\omega}) = 1 \quad \text{pro } |\omega| \leq \omega_c$$

$$H(e^{j\omega}) = 0 \quad \text{pro } (|\omega| < \omega) \ \& \ (|\omega| > \omega_c)$$

Pro $h(n)$ potom vypočítáme:

$$h(n) = \frac{\omega_c \sin(n \omega_c)}{2\pi n \omega_c} \quad (3.5)$$

Pro $n \rightarrow 0$ pak platí:

$$h(0) = \frac{\omega_c}{2\pi} \lim_{n \rightarrow 0} \frac{\sin(n \omega_c)}{n \omega_c} = \frac{\omega_c}{\pi}$$

4. výpočet FIR ... LPF:

Pro výpočet použijeme Hammingovo okno H_w . Výsledná impulsní odezva, tj. koeficienty navrženého filtru, pak bude ve tvaru:

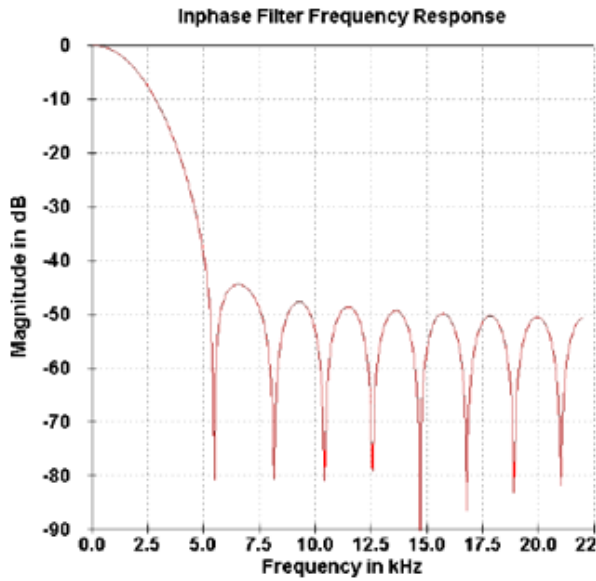
$$h_w(n) = h(n) \cdot H_w(n) \quad (3.6)$$

Obecný tvar pro Hammingovo okno je následující:

$$H_w(n) = w_{RECT}(n) \left[\alpha + \left(1 - \alpha \cos \left(2\pi \frac{n}{M} \right) \right) \right], \quad (3.7)$$

kde $\alpha = \mathbf{0,54}$ pro Hammingovo okno.

5. Frekvenční charakteristika filtru a Výpočet koeficientů Hammingova okna $M + 1 = 21$:



Obr. 3.5 – Frekvenční charakteristika filtru s použitím Hammingova okna H_w .

k	n	h(n)	H_w
0	-10	0,026861	0,080000
1	-9	0,033192	0,102514
2	-8	0,039405	0,167852
3	-7	0,045346	0,269619
4	-6	0,050859	0,397852
5	-5	0,055801	0,540000
6	-4	0,060039	0,682148
7	-3	0,063461	0,810381
8	-2	0,065975	0,912148
9	-1	0,067511	0,977486
10	0	0,068027	1,000000
11	1	0,067511	0,977486
12	2	0,065975	0,912148
13	3	0,063461	0,810381
14	4	0,060039	0,682148
15	5	0,055801	0,540000
16	6	0,050859	0,397852
17	7	0,045346	0,269619
18	8	0,039405	0,167852
19	9	0,033192	0,102514
20	10	0,026861	0,080000

Tab. 3.1 – Výpočet dat v Microsoft Excel.

Realizace návrhu a charakteristika filtrů FIR pro $M + 1 = 31$ a $M + 1 = 51$, viz příloha B.

Mějme dány signály $s^*(k)$ a $s(k)$ s N vzorky. Oba signály procházíme současně po jednotlivých vzorcích. Jestliže $s^*(k)$ chápeme jako aktuální vzorek filtrovaného signálu na pozici k , pak $s(k)$ je aktuálním vzorkem ve vstupním signálu na stejné pozici. Pozice k je omezena hranicemi $k > M/2$ a $k < N - M/2$, což znamená, že tyto krajní vzorky signálu $s(k)$ nejsou filtrovány. Hodnotu filtrovaného vzorku pak získáme podle následujícího vztahu

$$s^*(k) = \sum_{i=-\frac{M}{2}}^{\frac{M}{2}} s(k-i)h(i) \quad (3.9)$$

3.4 Transformace signálu

Transformace signálu $s(k)$ resp. $s^*(k)$ spočívá v jeho posunu o střední hodnotu $E[s(k)]$ resp. $E[s^*(k)]$.

Střední hodnota náhodné veličiny je nejdůležitější charakteristikou polohy. Střední hodnota náhodné veličiny X , označována $E(X)$, je reálné číslo definované pro diskrétní náhodou veličiny dle vztahu 3.9.[7]

Výsledný signál transformace $\bar{s}(k)$ je pak v kartézské soustavě souřadnic umístěn v počátku bez vlivu možných posunů vzniklých při vytváření záznamu.

Mějme daný výchozí resp. filtrovaný hlasový signál $s(k)$ resp. $s^*(k)$ s počtem N vzorků. Necht' se vybraný vzorek $s(k)$ resp. $s^*(k)$ vyskytuje v celé délce svého signálu s pravděpodobností p_x . Pak dostáváme množinu pravděpodobností $\{p_1, p_2, p_3, \dots, p_K\}$ o velikosti K , kde $K \leq N$. Každý prvek této množiny přísluší k vybranému (neopakujícímu se) vzorku signálu. Jelikož k veličině $s(k)$ resp. $s^*(k)$ přistupujeme diskrétně, je střední hodnota definována vztahem

$$E[s(k)] = \sum_{x=1}^K s(x)p_x = s(1)p_1 + s(2)p_2 + \dots + s(K)p_K \quad (3.9)$$

resp.

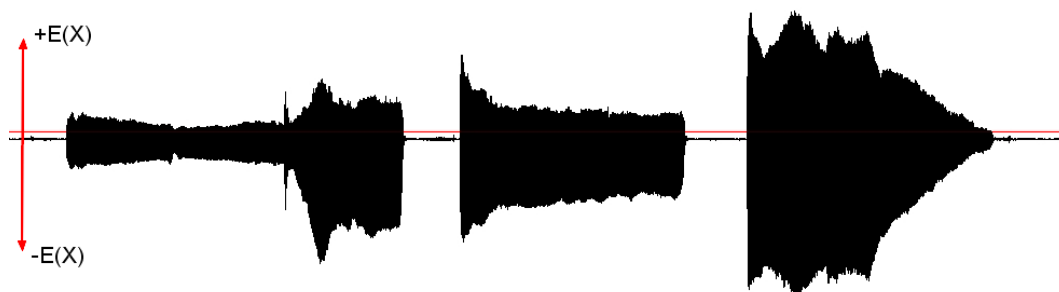
$$E[s^*(k)] = \sum_{x=1}^K s^*(x)p_x = s^*(1)p_1 + s^*(2)p_2 + \dots + s^*(K)p_K. \quad (3.10)$$

Kde $s(x)$ resp. $s^*(x)$ představuje výskyt vzorku v signálu $s(k)$ resp. $s^*(k)$ s pravděpodobností p_x . Nyní máme stanovenou střední hodnotu $E[s(k)]$ resp. $E[s^*(k)]$. Výsledkem posunu je transformovaný signál $\bar{s}(k)$ vyjádřený vztahem

$$\bar{s}(k) = s(k) - E[s(k)] \quad (3.11)$$

resp.

$$\bar{s}(k) = s^*(k) - E[s^*(k)] \quad (3.12)$$

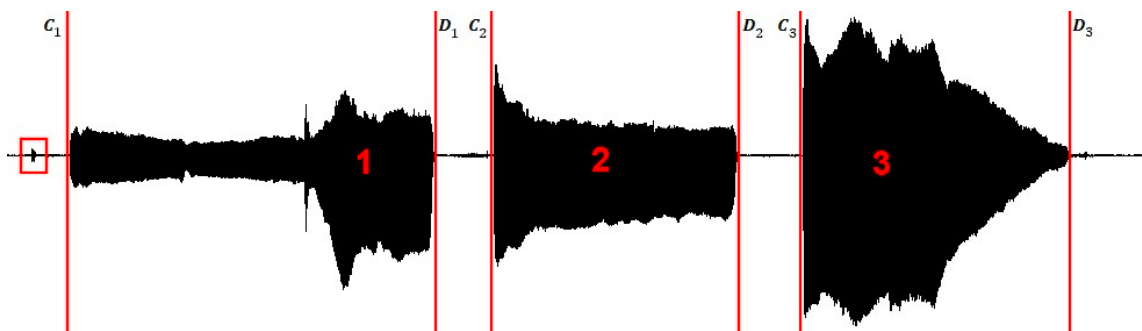


Obr. 3.6 – Ukázka zobrazeného netransformovaného signálu $s(k)$. Transformací se signál posune o hodnotu $E[s(k)]$ resp. $E[s^*(k)]$ po ose y na úroveň vodorovné čáry.

4. Metody detekce začátku a konce salv

Z výše uvedeného je zřejmé, že hlavním předpokladem pro analýzu vlivu zátěže na hlasivky je nalezení salv v signálu $s'(k)$. Kde $s'(k)$ chápeme jako obecný signál, který reprezentuje předzpracovaný transformovaný signál (viz kap. 3.2 a kap. 3.4).

Předpoklad podoby fonace a jednotlivých salv byl stanoven v kap. 3.1. Při hledání hranic salv v signálu $s'(k)$ je nutné brát v úvahu existenci výskytu možných vzruchů, které nejsou brány jako užitečná informace. Nelze je tak jako salvu označit nebo do jiné zahrnout. Tento požadavek pak určuje jak je daná metoda v hledání přesná bez ohledu na výskyt šumů. Jedná se o robustnost metody.



Obr. 4.1 – Ukázka, jak by měly být po analýze signálu $s'(k)$ správně označeny tři salvy zaznamenané v nahrávce $s(k)$ (červený obdélník vyznačuje výskyt vzruchu, který nemá se salvami nic společného).

V obrázku, obr. 4.1, je jednoznačně vidět výskyt salv ve fonaci, které jsou vyznačeny body C_x a D_x . Mezi nimi se nachází „hluché místo“, které představuje v ideálním případě „ticho“ a reprezentuje potřebnou pauzu pro nádech.

Z řady uvažovaných přístupů jsem nakonec přistoupil k diskusi a testování tří metod pro automatickou detekci začátku a konce salv. Jejimi modifikacemi se dostáváme na řadu možností, jak v signálu $s(k)$ resp. $s'(k)$ hranice stanovit. Modifikací chápeme změnu vstupního signálu, kdy se může jednat o signál bez filtrace nebo s filtrací s různými obměnami FIR filtru. Cílem je z těchto možností vybrat nejrobustnější metodu na základě provedené analýzy.

Testované metody pro detekci salv:

- Metoda ruční detekce
- Metoda počtu průchodů hranic
- Metoda počtu průchodů hranic kvadrátu
- Metoda aritmetického průměru

4.1 Metoda ruční detekce

Metoda je založena na ručním určení hranic salvy. V průběhu signálu $s'(k)$ se ručně pomocí myši označí body C_x a D_x , které reprezentují pozici vzorku k v signálu $s(k)$. Současně platí $C_x \ll D_x$, kde x označuje číslo salvy v záznamu signálu $s(k)$.

Tato metoda je implementována, protože počítáme se zpracováním signálů, které jsou velmi zatížené chrapotem pacienta a navrhované metody detekce počátku a konce salvy mohou selhávat. Ruční volba bodů C_x a D_x je také používána při hodnocení „úspěšnosti“ detekce pomocí automatických metod detekce.

Výhodou této metody je, že poskytne přesné označení salv bez případných chyb. Nevýhoda naopak spočívá právě v závislosti na asistenci pozorovatele a správném úsudku, aby určil vybrané body.

4.2 Metoda počtu průchodů hranicí

Metoda je založena na automatickém určení hranic salvy. V průběhu signálu $s'(k)$ se automaticky označí body C_x a D_x , které reprezentují pozici vzorku k v signálu $s(k)$. Určení těchto bodů se zakládá na určení možného začátku salvy, tj. C_x , kdy vzorek $s'(k)$ projde jednou z určených hranic. Tyto hranice se nastavují tak, aby jimi bezpečně prošel vybraný segment $s'(k)$, představující možnou salvu, a zároveň od ní oddělil pauzy pro nádech. Pokud se jedná o hledanou salvu, pak signál $s'(k)$ bude od bodu C_x střídavě procházet určenými hranicemi. V okamžiku, kdy signál opět prochází hranici, se rozhoduje, jestli vzniklý průsečík na pozici k není konec salvy, tedy bod D_x . Jestliže se podaří bod D_x najít, zkoumá se, zdali počet průchodů překročí definovanou hodnotu. Pokud ano, dojde k označení platné salvy. Jeden průchod chápeme jako část půl-periody $s'(k)$ procházející nad resp. pod danou hranicí.

Nechť máme k dispozici obecný signál $s'(k)$ s N vzorky a konstantní funkci f_h' resp. f_d' , která označuje horní resp. dolní hranici. Jestliže existují proměnné ex_h' , ex_d' a p_{ex}' , potom je horní hranice f_h' signálu vyjádřena vztahem

$$f_h' = |ex_h' - ex_d'|p_{ex}'. \quad (4.1)$$

Dolní hranice f_d' nabývá hodnoty opačné oproti funkci f_h' a vyjádříme ji vztahem

$$f_d' = -|ex_h' - ex_d'|p_{ex}' \quad (4.2)$$

nebo také, kdy po dosazení f_h' dostáváme

$$f_d' = -f_h' ,$$

kde ex_h' a ex_d' jsou nalezené extrémy $s'(k)$, pro které platí $ex_h' = \max\{s'(k)\}$ a $ex_d' = \min\{s'(k)\}$, pro $\forall k \in 1, \dots, N$. Proměnná p_{ex}' je definovaná jako procentuální hodnota z $|ex_h' - ex_d'|$.

Hledání začátku salvy

Předpokládejme, že signál $s'(k)$ je harmonického⁷ průběhu. Potom platí, že jeho průběh střídavě prochází kladnou a zápornou polorovinou osy y v kartézském souřadném systému. Procházejme $s'(k)$ po jednotlivých vzorcích nacházející se mezi nastavenými hranicemi. Jestliže $s'(k)$ představuje aktuálně vybraný vzorek na pozici k a platí, že jeho velikost je větší resp. menší nebo rovna konstantní funkci f_h' resp. f_d' .

$$s'(k) \geq f_h' = |ex_h' - ex_d'|p_{ex}' \quad (4.3)$$

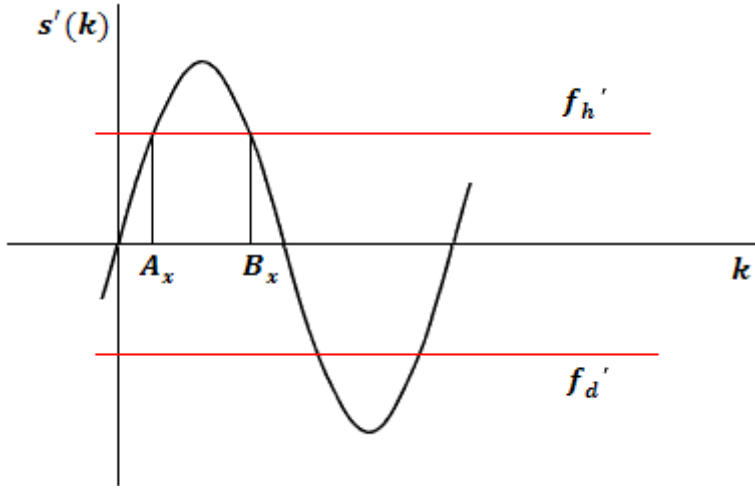
resp.

$$s'(k) \leq f_d' = -|ex_h' - ex_d'|p_{ex}' \quad (4.4)$$

Potom můžeme prohlásit, že se vybraný vzorek nachází v kladné resp. záporné polorovině souřadného systému. Průsečíkem signálu $s'(k)$ a hranice f_h' resp. f_d' na dané pozici dostáváme bod A_x , který nabývá hodnoty $A_x = k$. Dále od bodu A_x předpokládáme, že se nacházíme nad resp. pod definovanou hranicí a každý další vybraný vzorek splňuje podmínku 4.3 resp. 4.4. Hledáme takový vzorek $s'(k)$, který jakožto první tuto podmínku nesplňuje. V tom okamžiku se dostáváme zpět mezi hranice a určíme bod $B_x = k$.

⁷ Periodické funkce jakou jsou sinus nebo kosinus.

Body A_x a B_x pak označují část kladné resp. záporné vlny (půl-periody) $s'(k)$ nad kladnou resp. zápornou hranicí, viz obr. 4.2, a platí $A_x < B_x$. Je-li vzorek A_x prvně nalezený v rámci hledání nové salvy, pak jej prohlásíme za počátek potenciální salvy a určíme bod $C_x = A_x$. Nalezenou půl-periodu potom považujeme za jeden průchod vybranou hranicí.



Obr. 4.2 – Ukázka toho, jak kladná půlvlna signálu $s'(k)$ prochází hranicí f_h' . Body A_x a B_x pak představují pozice vzorků, podle kterých se určuje začátek či konec fonace.

Hledání konce salvy

Definujeme si novou proměnnou $t_s(x)$, která označuje délku pauzy v sekundách, viz obr. 4.3, a proměnnou c_{hd} , která vyjadřuje počet průchodů hranicemi f_h' a f_d' v rámci jedné salvy. Podle následujícího vztahu dostáváme při vzorkovací frekvenci $f_{vz} = 44,1$ kHz, novou proměnnou $t_v(x)$, která vyjadřuje délku pauzy v počtu vzorků.

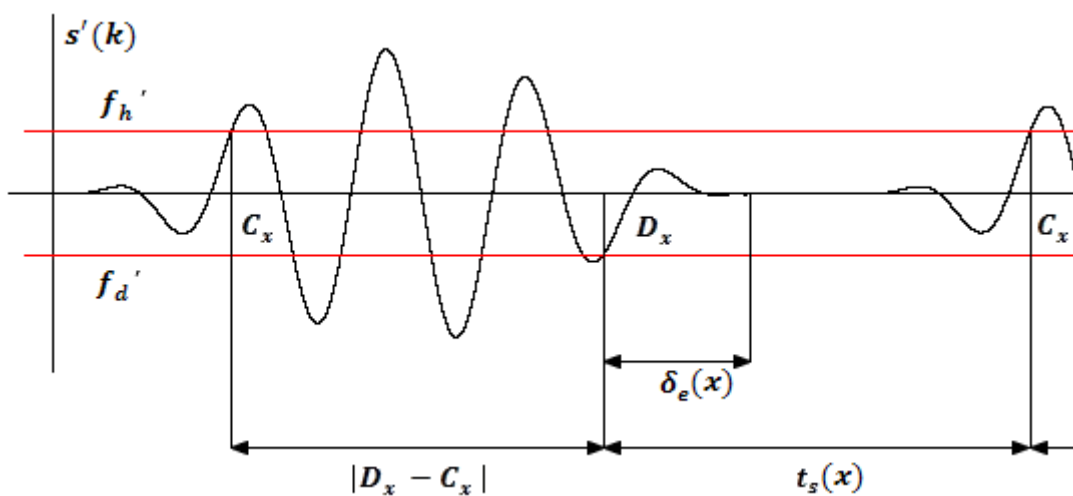
$$t_v(x) = t_s(x) \cdot f_{vz} \quad (4.5)$$

Jestliže při dalším průchodu jednou z hranic f_h' nebo f_d' (nalezení nových bodů A_x a B_x) platí podmínka $|B_{x-1} - A_x| < t_v(x)$, kde B_{x-1} je bod nalezený v předchozím průchodu, pak se nacházíme stále ve stejné salvě a dochází k navýšení čítače c_{hd} o jeden průchod.

Pokud je platnost podmínky opačná, tedy $|B_{x-1} - A_x| \geq t_v(x)$, potom je nalezen začátek nové potenciální salvy $C_{x+1} = A_x$ a pozice B_{x-1} se stává koncem salvy předchozí. Dostáváme tak nový bod D_x , pro který platí $D_x = B_{x-1}$. Jestliže A_x je určen pozicí aktuálního vzorku k v nové salvě, pak od něj dále hledáme konec salvy - bod D_{x+1} .

Potenciální salva se stane skutečně nalezenou, jestliže čítač $c_{hd} \geq \kappa$, kde κ představuje konstantu, která vyjadřuje minimální počet průchodů kladných a záporných půlvln signálu $s'(k)$ vybranými hranicemi v rámci jedné salvy. Tímto krokem zajistíme vyloučení případných vzruchů v signálu a předejdeme tak jejich označení jako správně nalezenou salvu.

Výhoda této metody spočívá především v automatickém hledání salv ve fonaci. Nevýhodou je naopak závislost na nastavených vstupních parametrech, jako jsou zvolená hranice a minimální předpokládaná pauza mezi salvami. I přes opakované nastavení (viz dále – kap. 7) nelze u metody docílit přesného označení salvy a musíme počítat s určitou chybou (viz kap. 4.5).



Obr. 4.3 – Ukázka jak vypadá vyznačení salvy popsanou metodou.

4.3 Metoda počtu průchodů hranicí kvadrátu

Metoda je založena na automatickém určení hranic salvy. V průběhu signálu $s''(k)$, který představuje kvadrát vstupního signálu $s'(k)$, se automaticky označí body C_x a D_x , které reprezentují pozici vzorku k v signálu $s(k)$. Určení těchto bodů se zakládá na určení možného začátku salvy, tj. C_x , kdy vzorek $s''(k)$ projde určenou hranicí. Tato hranice se nastavuje tak, aby jí bezpečně prošel vybraný segment $s''(k)$, představující možnou salvu v $s'(k)$ resp. $s(k)$, a zároveň od ní oddělil pauzy pro nádech. Pokud se jedná o hledanou salvu, pak signál $s''(k)$ bude od bodu C_x procházet určenou hranicí. V okamžiku, kdy signál opět prochází hranicí, se rozhoduje, jestli vzniklý průsečík na pozici k není konec salvy, tedy bod D_x . Jestliže se podaří bod D_x najít, zkoumá se, zdali počet průchodů překročí definovanou hodnotu. Pokud ano, dojde k označení platné salvy. Jeden průchod chápeme jako část půl-periody $s''(k)$ procházející nad danou hranicí.

Necht' máme k dispozici vstupní signál $s'(k)$ s N vzorky a s ním existující proměnné ex_h' , ex_d' a p_{ex}' , kde $ex_h' = \max\{s'(k)\}$ a $ex_d' = \min\{s'(k)\}$ jsou extrémy vstupního $s'(k)$, pro $\forall k \in 1, \dots, N$ a p_{ex}'' je určena jako procentuální hodnota z $|ex_h' - ex_d'|$. Definujme novou funkci $s''(k)$, která je vyjádřena jako druhá mocnina $s'(k)$ a konstantní funkci f_h'' , která označuje její horní hranici

$$s''(k) = [s'(k)]^2, \quad (4.6)$$

a horní hranice f_h'' je vyjádřena vztahem

$$f_h'' = ex_h'' p_{ex}'' , \quad (4.7)$$

kde ex_h'' je nalezené maximum signálu $s''(k)$ určené $ex_h'' = \max\{(ex_h')^2, (ex_d')^2\} = \max\{s''(k)\}$, pro $\forall k \in 1, \dots, N$.

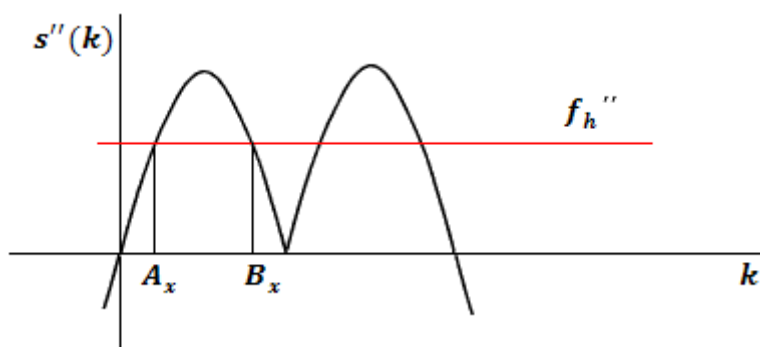
Hledání začátku salvy

Víme, že kvadratická funkce je funkcí kladnou, a tak se průběh funkce $s''(k)$ spolu s konstantní funkcí f_h'' nachází v kladné polorovině osy y kartézského souřadného systému. Procházíme množinu $s''(k)$ po jednotlivých prvcích nacházející se pod stanovenou hranicí. Jestliže pro daný aktuálně vybraný vzorek $s''(k)$ na pozici k platí, že jeho velikost je větší než funkce $f_h''(k)$

$$s''(k) \geq f_h'' = ex_h'' p_{ex}'' , \quad (4.8)$$

potom můžeme prohlásit, že vybraný vzorek představuje průnik s funkcí f_h'' a dostáváme bod A_x , který nabývá hodnoty $A_x = k$. Dále od bodu A_x předpokládáme, že se nacházíme nad definovanou hranicí a velikost každého dalšího vzorku ($k > A_x$) splňuje podmínku 4.8. Hledáme takový vzorek $s''(k)$, který jakožto první tuto podmínku nesplňuje. V tom okamžiku se dostáváme zpět pod hranici a určíme bod $B_x = k$, vzniklý opět průnikem signálu a hranice.

Body A_x a B_x pak označují část kladné půl-periody $s''(k)$ nad kladnou hranicí f_h'' , pro něž platí $A_x < B_x$, viz obr. 4.4. Je-li vzorek A_x prvně nalezený v rámci hledání nové salvy, pak jej prohlásíme za počátek potenciální salvy a dostáváme nový bod $C_x = A_x$. Nalezenou půl-periodu potom považujeme za jeden průchod vybranou hranicí.



Obr. 4.4 – Ukázka toho, jak kladná půlvlna signálu $s''(k)$ prochází hranicí f_h'' . Body A_x a B_x pak představují pozice vzorků, podle kterých se určuje začátek či konec fonace.

Hledání konce salvy

Definujeme si novou proměnnou $t_s(x)$, která označuje délku pauzy v sekundách, viz Obr 4.5, a proměnnou c_{hd} , která vyjadřuje počet průchodů hranicemi f_h'' v rámci jedné salvy. Podle následujícího vztahu dostáváme při vzorkovací frekvenci $f_{vz} = 44,1$ kHz, novou proměnnou $t_v(x)$, která vyjadřuje délku pauzy v počtu vzorků.

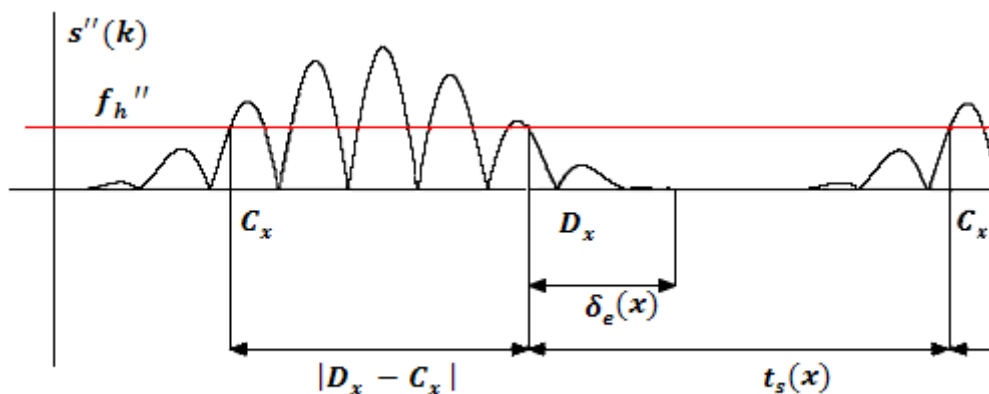
$$t_v(x) = t_s(x) \cdot f_{vz} \quad (4.9)$$

Jestliže při dalším průchodu hranicemi (nalezení nových bodů A_x a B_x) platí podmínka $|B_{x-1} - A_x| < t_v(x)$, kde B_{x-1} je bod nalezený v předchozím průchodu f_h'' , pak se nacházíme stále ve stejné salvě a dochází k navýšení čítače c_{hd} o jeden průchod.

Pokud je platnost podmínky opačná, tedy $|B_{x-1} - A_x| \geq t_v(x)$, potom je nalezen začátek nové potenciální salvy $C_{x+1} = A_x$ a pozice B_{x-1} se stává koncem salvy předchozí. Dostáváme nový bod D_x , pro který platí $D_x = B_{x-1}$. Jestliže A_x je určen pozicí aktuálního vzorku k v nové salvě, pak od něj dále hledáme konec salvy D_{x+1} .

Potenciální salva C_x a D_x se stane skutečně nalezenou, jestliže čítač $c_{hd} \geq \kappa$, kde κ představuje konstantu, která vyjadřuje minimální počet průchodů kladných půlvln signálu $s''(k)$ vybranou hranicí v rámci jedné salvy. Tímto krokem zajistíme vyloučení případných vzruchů v signálu a předejdeme tak jejich označení jakožto nalezená a správná salva.

Výhoda této metody spočívá především v automatickém hledání salv ve fonaci. Nevýhodou je pak nutnost práce se dvěma signály a závislost na nastavených vstupních parametrech – zvolená hranice a minimální předpokládaná pauza mezi salvami. I přes opakované nastavení těchto parametrů (viz dále – kap. 7) nelze u metody docílit přesného označení salvy a musíme počítat s určitou chybou (viz kap. 4.5).



Obr. 4.5 – Ukázka jak vypadá vyznačení salvy popsanou metodou.

4.4 Metoda aritmetického průměru

Metoda je založena na automatickém určení hranic salvy. V průběhu signálu $s'(k)$ se automaticky označí body C_x a D_x , které reprezentují pozici vzorku k v signálu $s(k)$. Základní myšlenka určení těchto bodů vychází z předpokladu stejné velikosti vzorku v době pauzy pro nádech („absolutní ticho“). Tzn., že pokud spočítáme z těchto vzorků aritmetický průměr, pak se hodnota tohoto průměru rovná jakémukoliv vzorku, které průměr zahrnuje. V opačném případě registrujeme vzruch a první vzorek, který je do průměru zahrnut a prochází jednou z vybraných hranic, stanovuje na dané pozici k bod C_x . Od tohoto okamžiku se hledá konec salvy, tj. bod D_x , který splňuje podmínku opačnou – hledáme takové prvky, které se rovnají vypočtenému průměru. Pokud rozdíl hranic salvy, $D_x - C_x$, přesahuje určitou hodnotu (minimální požadovanou délku salvy), dojde k označení platné salvy.

Mějme daný obecný hlasový signál $s'(k)$ s N vzorky a konstantní funkci f_h' resp. f_d' , která označuje horní resp. dolní hranici. Jestliže existují proměnné ex_h' , ex_d' a p_{ex}' , potom je horní hranice f_h' signálu vyjádřena vztahem

$$f_h' = |ex_h' - ex_d'|p_{ex}' \quad (4.10)$$

Dolní hranice f_d' nabývá hodnoty opačné oproti funkci f_h' a je vyjadřuje ji vztah

$$f_d' = -|ex_h' - ex_d'|p_{ex}' \quad (4.11)$$

nebo také, kdy po dosazení f_h' dostáváme

$$f_d' = -f_h',$$

kde ex_h' a ex_d' jsou nalezené extrémy $s'(k)$, pro které platí $ex_h' = \max\{s'(k)\}$ a $ex_d' = \min\{s'(k)\}$, pro $\forall k \in 1, \dots, N$. Proměnná p_{ex}' je definovaná jako procentuální hodnota z $|ex_h' - ex_d'|$.

Dále mějme určenou množinu O velikosti M , kde $M \ll N$. Množina si lze představit jako pomyslné okno posouvající se s pozicí aktuálně vybraného prvku $s'(k)$ a obsahuje vzorky vybraného segmentu – první prvek o_1 je totožný s aktuálním vzorkem $s'(k)$ na pozici k a ostatní prvky množiny pak odpovídají vzorkům signálu od pozice k po $k + M - 1$.

Hledání začátku salvy

Opět platí předpoklad harmonického průběhu $s'(k)$ a tedy fakt, že signál $s'(k)$ střídavě prochází kladnou a zápornou polorovinou osy y v kartézském souřadném systému. Procházejme signál $s'(k)$ po jednotlivých vzorcích nacházející se mezi nastavenými hranicemi spolu s posouvajícím se oknem O , viz obr. 4.6. Při každém dalším vzorku se posouvá okno na pozici aktuálně vybraného a mění se obsah množiny O tak, že

$$o_1 = s'(k), o_2 = s'(k + 1), \dots, o_M = s'(k + M), \quad (4.12)$$

resp.

$$o_1 = o_2; o_2 = o_3; \dots; o_M = s'(k + M). \quad (4.13)$$

Při každém posunu okna se souběžně počítá aritmetický průměr z prvků množiny O . Vzhledem k tomu, že v okamžiku pauzy nedosáhneme absolutního ticha, tak aby měly vzorky v celé její délce konstantní velikost, počítá se průměr z prvních 8bitů⁸ (o_i^{8b})

⁸ Prvních 8bitů, jelikož hodnoty celých 16bitových vzorků jsou celkově dost různorodé.

každého vzorku množiny. Tím dojde k oříznutí hodnot vzorků a dostáváme prvky více-méně stejné velikosti – o malé intenzitě signálu, tak jak během pauzy očekáváme.

$$\bar{r} = \frac{1}{M} \sum_{i=1}^M |o_i^{8b}|, \quad (4.14)$$

kde i je pozicí prvku v množině O . S hodnotou spočteného průměru pak porovnáváme prvních osm bitů hodnoty aktuálního prvku, tedy o_1 . Jestliže se tyto hodnoty rovnají, potom můžeme prohlásit, že tomu tak je i pro ostatní prvky množiny a nacházíme se v oblasti mezi salvami. V opačném případě, kdy $o_1 \neq \bar{r}$ resp. $s'(k) \neq \bar{r}$ zaznamenáváme jakýsi vzruch, který značí možný počátek salvy. Procházíme pomyslné okno od prvního (aktuálního) prvku. Pokud je pro některý vzorek o_i splněna podmínka překročení hranice f_h' resp. f_d'

$$o_i \geq f_h' = |ex_h' - ex_d'|p_{ex}' \quad (4.15)$$

resp.

$$o_i \leq f_d' = -|ex_h' - ex_d'|p_{ex}', \quad (4.16)$$

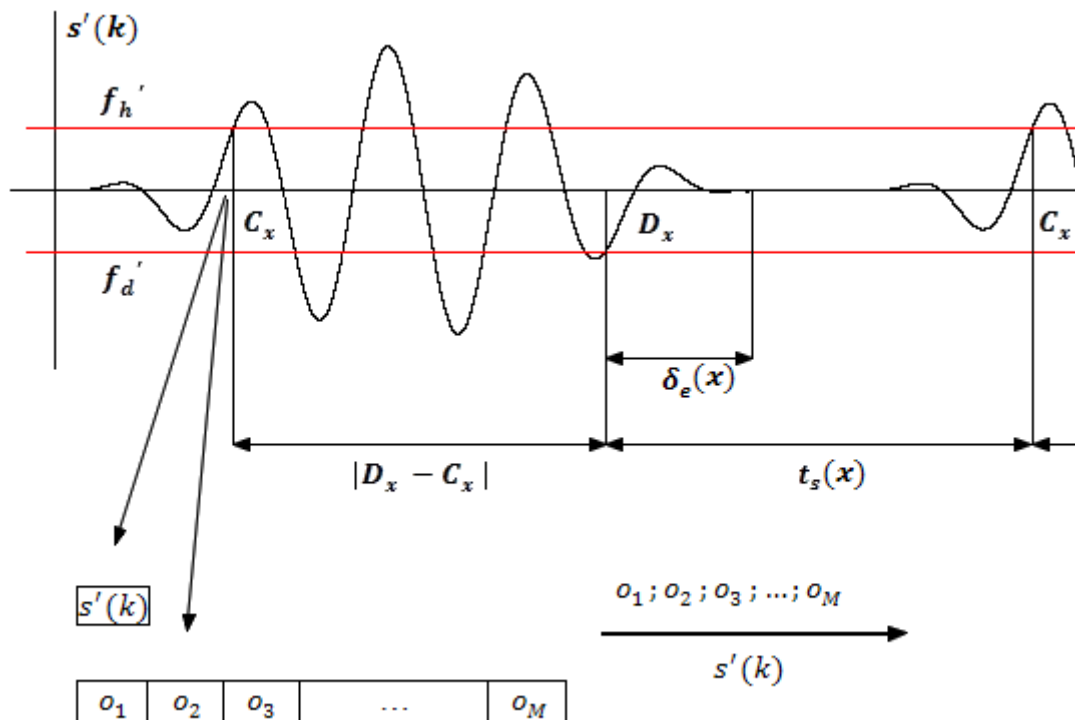
můžeme říci, že se aktuálně vybraný vzorek z množiny o_i nachází v kladné resp. záporné polorovině souřadného systému. Průsečíkem signálu o_i a hranice f_h' resp. f_d' dostáváme bod A_x , který označuje pozici počátku salvy a nabývá hodnoty $A_x = k$.

Hledání konce salvy

Hledání bodu D_x , který označuje pozici konce potenciální salvy, probíhá na stejném principu jako hledání začátku C_x . Dále procházíme signál $s'(k)$ a při každém vzorku počítáme z nově nastavených hodnot v množině O hodnotu průměru dle 4.14. Jestliže jsou všechny prvky v množině O rovny \bar{r} , pak se vybraný vzorek $s'(k) = o_1$ stává koncovým vzorkem potenciální salvy a $D_x = k$.

Potenciální salva se stane skutečně nalezenou, jestliže pro získané body D_x a C_x platí $|D_x - C_x| \geq \kappa$, kde κ představuje konstantu, která vyjadřuje minimální požadovanou délku salvy v počtu vzorků za jednotku času. Tímto krokem zajistíme vyloučení případných vzruchů v signálu a předejdeme tak jejich označení jako správně nalezenou salvu.

Výhoda v popsané metodě spočívá především v automatickém hledání salv ve fonaci. Záporem metody je pak nutnost udržovat duplicitně data pro množinu O a stálé počítání aritmetického průměru, což vede ke značnému prodloužení doby hledání (zejména pokud je množina hodně velká). Dalším nevýhodnou vlastností je závislost na nastavených vstupních parametrech – zvolená hranice a minimální požadovaná délka salvy. I přes opakované nastavení těchto parametrů (viz dále – kap. 7) nelze u metody docílit přesného označení salvy a musíme počítat s určitou chybou (viz kap. 4.5).



Obr. 4.6 – Ukázka jak vypadá vyznačení salvy popsanou metodou.

4.5 Chyba metod

Z obrázků Obr. 4.3, 4.5 a 4.6 je u jednotlivých metod patrné, že po označení bodů C_x resp. D_x v daném signálu, dohází mezi zvolenými hranicemi ještě k několika zákmitům před C_x resp. za D_x . Mějme proměnnou $t_e(x)$, která vyjadřuje délku těchto zákmitů v počtu vzorků za jednotku času. Ačkoliv je zřejmé, že těmito zákmity salva začíná nebo dále pokračuje, tak díky podmínkám 4.3 a 4.4, 4.8 nebo 4.15 a 4.16, do ní nejsou zahrnuty, což je příčinou vzniku chyby $\delta_e(x)$, která je vyjádřena vztahem

$$\delta_e(x) = \frac{t_e(x)}{f_{vz}} [ms]. \quad (4.17)$$

Díky vzorkovací frekvenci $f_{vz} = 44,1$ kHz a definované konstantě κ vyjadřující minimální požadovanou délku salvy v počtu vzorků (*Metoda aritmetického průměru* (3)) nebo čítači c_{hd} vyjadřující tuto délku počtem průchodů hranicemi f_h' a f_d' (*Metoda počtu průchodů hranic* (1) a *Metoda počtu průchodů hranic kvadrátu* (2)), platí předpoklad $|D_x - C_x| \gg t_e(x)$ a docházíme k závěru, že vyjádřená chyba $\delta_e(x)$ je zanedbatelná.

Mějme zavedené proměnné $\delta_{e1}(x)$, $\delta_{e2}(x)$ a $\delta_{e3}(x)$, které vyjadřují chyby jednotlivých metod. Srovnáním metod 1 a 2 můžeme říct, že platí nerovnost chyb $\delta_{e1}(x) \geq \delta_{e2}(x)$. Ostrá nerovnost platí, jestliže hranice f_h'' u metody (2) není definována pro druhý kvadrát (viz 4.3.2) – bude menší, což tuto metodu automaticky zjemňuje vstupním parametrem. Obecně platí rovnost chyb $\delta_{e1}(x) = \delta_{e2}(x)$.

Přesnost metody 3, z níž se odvíjí hodnota chyby $\delta_{e3}(x)$, je závislá na velikosti pomyslného okna (množina O). V případě hledání začátku salvy, lze skutečně dosáhnout přesnějšího označení. Při hledání jejího konce se může projevit tak, že namísto chybějících vzorků salvy, se do ní započítávají vzorky patřící k pauze. I když tato situace může nastat, je vzhledem k vzorkovací frekvenci celkem nepravděpodobná – množina O by musela být značně velká a zároveň krátká fonace.

Shrneme-li získané poznatky o metodách, můžeme říci, že mezi chybami platí následující vztah

$$\delta_{e1}(x) \geq \delta_{e2}(x) \geq \delta_{e3}(x). \quad (4.18)$$

5. Parametry salvy

5.1 Výpočet délky salvy (pauzy)

Vyjádření délky salv a příslušných pauz mezi nimi v jednotkách času [s]. Na základě získaných parametrů je možné posoudit vliv zátěže testu a zároveň potvrdit předpoklad zmenšující se délky salvy s postupujícím časem.

Mějme určené hranice salv, které jsou vyjádřeny pozicí vzorku v signálu $s(k)$, se vzorkovací frekvencí $f_{vz} = 44,1$ kHz. Vezmeme-li body C_x a D_x označující jednu salvu, můžeme určit délku této salvy $t_{salva}(x)$ a délku pauzy $t_{pauza}(x)$ k salvě následující C_{x+1} a D_{x+1} , tj. $t_{pauza}(x) =$ pauza po salvě x .

$$t_{salva}(x) = \frac{D_x - C_x}{f_{vz}} [s] \quad (5.1)$$

$$t_{pauza}(x) = \frac{C_{x+1} - D_x}{f_{vz}} [s] \quad (5.2)$$

5.2 Výpočet energie fonací

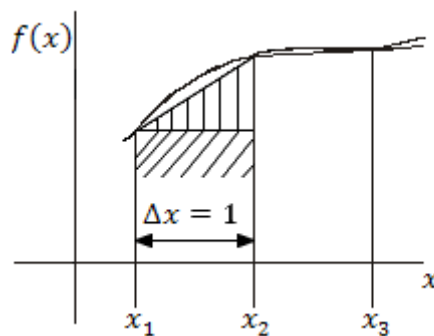
Na základě získané energie jsme schopni blíže určit vliv zátěže na hlasivky. Z praxe víme, že při fonaci pacienti s postupem času „násilím“ natahují délku salvy na úkor její intenzity. Pokud by mělo vyhodnocení probíhat pouze na základě délky jednotlivých salv a pauz mezi nimi, byl by výsledek do jisté míry zkreslený.

Mějme daný obecný signál $s(k)$ s N vzorky a vzorkovací frekvencí $f_{vz} = 44,1$ kHz, a předpokládejme existenci nalezených salv. Pak definujeme funkci $s''(k)$ jako druhou mocninu vstupního signálu $s(k)$

$$s(k)'' = [s(k)]^2. \quad (5.3)$$

Dostáváme tak funkci, jejíž průběh se nachází v kladné polorovině osy y kartézského souřadného systému. K vypočtení energie jedné salvy vyznačenou body C_x a D_x se využívá vztahu známého jako *Lichoběžníkové pravidlo* (LP), jehož tvar je následující

$$\begin{aligned} \Delta L_x &= ((x+1) - x) f(x) + \frac{1}{2}((x+1) - x) (f(x+1) - f(x)) \\ &= (\Delta x) \left(f(x) + \frac{1}{2} f(x+1) - \frac{1}{2} f(x) \right) = \frac{1}{2} (\Delta x) (f(x) + f(x+1)). \end{aligned} \quad (5.4)$$



Obr. 5.1 – Naznačení principu výpočtu elementární plochy.

Procházejme signál $s''(k)$ po jednotlivých vzorcích. Jestliže jsme na úrovni počátku salvy, je aktuálně vybraný vzorek na pozici $k \geq C_x$ a $s''(k+1)$ je vzorek bezprostředně následující. Hodnoty těchto vzorků můžeme dosadit do vzorce 5.4, kdy podle definice LP dostáváme element aproximace signálu ΔL_k ohraničený pozicemi vybraných vzorků k a $k+1$.

$$\Delta L_k = \frac{1}{2} (k+1 - k) (s''(k) + s''(k+1)), \quad (5.5)$$

kde $(k+1 - k) = \Delta k$ je perioda vzorkování a vyjadřuje krok mezi jednotlivými vzorky. Konečný vztah 5.5 je

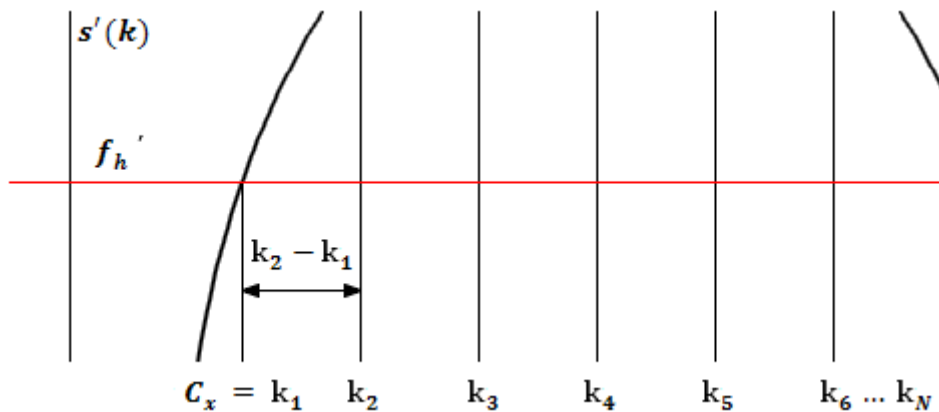
$$\Delta L_k = \frac{1}{2} (s''(k) + s''(k+1)).$$

Procházením signálu po vzorcích, postupně dostáváme dílčí hodnoty ΔL_k . Zavedením množiny $L = \{l_1, l_2, \dots, l_{N-1}\}$, která obsahuje tyto elementy, vypočítáme energii dané salvy jako sumu prvků této množiny, kdy dostáváme následující vztah

$$E_{salvy} = \sum_{k=C_x}^{D_x-C_x} \Delta L_k \quad (5.6)$$

nebo po dosazení můžeme psát

$$E_{salvy} = \frac{1}{2} \sum_{k=C_x}^{D_x-C_x} (s''(k) + s''(k+1)).$$



Obr. 5.2 – Obrázek ukazuje část začátku salvy vzorkovaného signálu $s''(k)$ a způsob aplikace LP.

Parametr energie v salvě bude použit pro analýzu úbytku energie s rostoucí zátěží, tj. počtem opakování dlouhodobé fonace vokálu „a“, tj. jednotlivých salv.

5.3 Úbytek intenzity signálu během salvy

Dle stanoveného předpokladu z kap. 3.1, se s postupem délky fonace vokálu „a“, tj. s délkou salvy, snižuje intenzita fonace. Tedy předpokládáme, že existuje rozdíl mezi intenzitou na začátku salvy a na jejím konci. Tento rozdíl můžeme vyjádřit právě jako úbytek energie během fonace v rámci jedné salvy

K získání informace-měřítko úbytku intenzity signálu, resp. energie během jedné salvy jsou navrženy dva parametry:

- Lineární regrese absolutní hodnoty signálu.
- Výpočet koeficientu tlumení.

5.3.1 Detekce lokálního maxima absolutní hodnoty signálu

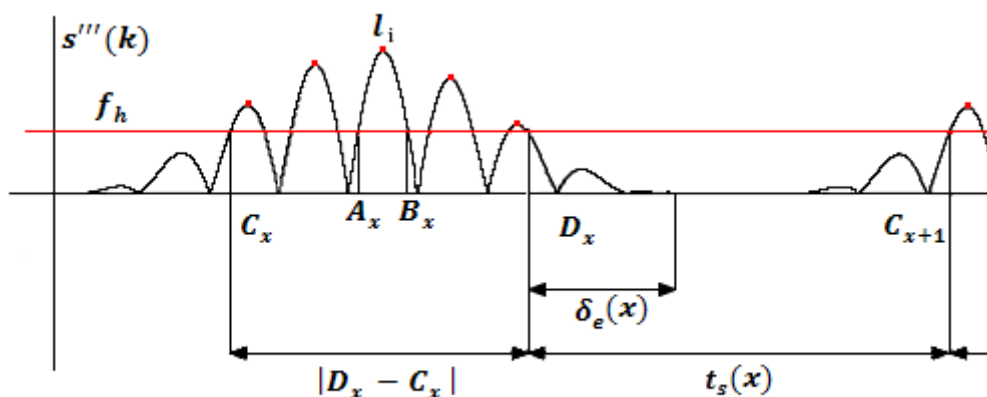
Definujme funkci (signál) $s'''(k)$, která je určena jako absolutní hodnota z obecného signálu $s(k)$ s N vzorky.

$$s'''(k) = |s(k)| \quad (5.7)$$

Potom tento signál procházejme po jednotlivých vzorcích uvnitř nalezených salv $s(k)$. Jestliže body C_x a D_x značí aktuálně procházenou salvu signálu, pak definujeme množinu dvojic

$$L_x = \{(p_1, l_1); \dots; (p_i, l_i); \dots; (p_M, l_M)\}, \quad (5.8)$$

kde l_i představuje lokální extrém, které jsou určeny jako maximum každé půl-periody v dané salvě x a vektor p_i jsou pozice k daným extrémům l_i signálu $s'''(k)$.



Obr. 5.3 – Ukázka jak v salvách signálu $s'''(k)$ nalézt lokální extrémy.

K určení lokálních extrémů, a tedy matice L_x , se využívá existence stanovených hranic f_h resp. f_d , se kterými jsme se setkali v každé z definovaných metod pro automatickou detekci salv (viz kap. 4.), kdy máme zajištěno, že se jedná o maxima salvy. Při získání bodů A_x a B_x , které vyjadřují průchod vybranou hranicí, se od prvního nalezeného bodu A_x porovnávají hodnoty vzorků. Jestliže je aktuální vzorek $s'''(k)$ na pozici $k > A_x$ větší než vzorek předchozí

$$s'''(k) > s'''(k-1), \quad (5.9)$$

potom se hodnota nalezeného extrému l_i a jeho pozice p_i mění. V okamžiku, kdy nalézáme bod B_x končí hledání lokálního maxima v rámci dané půl-periody a přidává se do matice L_x na pozici i , kdy hodnota maxima odpovídá poslední změně podle podmínky 5.9.

5.3.2 Lineární regrese maxim absolutní hodnoty signálu.

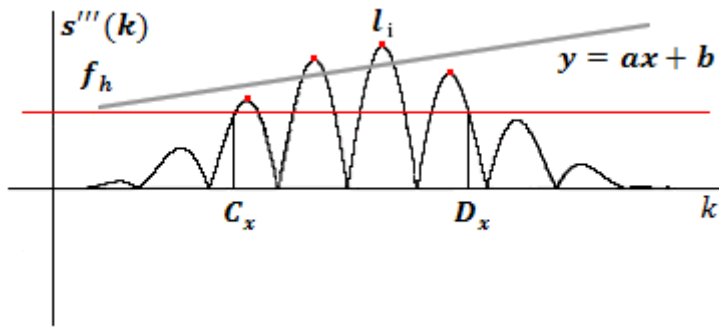
Lineární regrese představuje aproximaci daných hodnot polynomem prvního řádu (přímkou) metodou nejmenších čtverců. Jinak řečeno, jedná se o proložení několika bodů v grafu takovou přímkou, aby součet druhých mocnin odchylek jednotlivých bodů od přímky byl minimální.[12] Metodou nejmenších čtverců dostaneme koeficienty potřebné k vyjádření aproximující přímky $f(x) = ax + b$.

$$S(a, b) = \sum_{k=C_x}^{D_x-C_x} [f(x) - s'''(k)]^2 = \sum_{k=C_x}^{D_x-C} [ax + b - s'''(k)]^2 \quad (5.10)$$

Mějme dānu množinu dvojic L_x k dané salvě signālu $s'''(k)$. Do vzorce 5.3.4 postupně dosazujeme jednotlivé dvojice (p_i, l_i) , kde $p_i \equiv x$ a $l_i \equiv s'''(k)$. Výsledkem aproximace je lineární funkce ve tvaru

$$y = ax + b, \quad (5.11)$$

kde parametr a , představuje míru útlumu energie v dané salvě x (viz obr 5.4).



Obr. 5.4 – Ilustrace proložení přímky lokálními extrémy salvy.

5.3.3 Výpočet koeficientu tlumení.

Jedná se o exponenciální regresní model, resp. proložení maxim exponenciálou. Mějme dánu množinu dvojic L_x , tj. dvojice pozorování lokálních maxim (p_i, l_i) , k dané salvě signálu $s'''(k)$, kde $s'''(k) = |s(k)|$.

Pro útlum signálu $s(t)$ v salvě platí

$$s(t) = Ae^{-\frac{t}{T}} \sin(2\pi f) + \xi_k. \quad (5.12)$$

Tlumičí čára má pak rovnici

$$l(t) = Ae^{-\frac{t}{T}} \quad (5.13)$$

a pro pozorování maxim signálu $s'''(k)$ pak platí

$$l_i = Ae^{-\frac{p_i}{T}} + \eta_{p_i}. \quad (5.14)$$

Logaritmováním tohoto vztahu dostaneme lineární vztah mezi p_i a l_i , resp. p_i a $\ln(l_i)$, tj.

$$\ln(l_i) = \ln(A) - \frac{1}{T} p_i + \zeta_{p_i}. \quad (5.15)$$

Dvojcemi bodů $[x_i, \ln(l_i)]$, proložíme metodou nejmenších čtverců přímkou, tj. bude platit

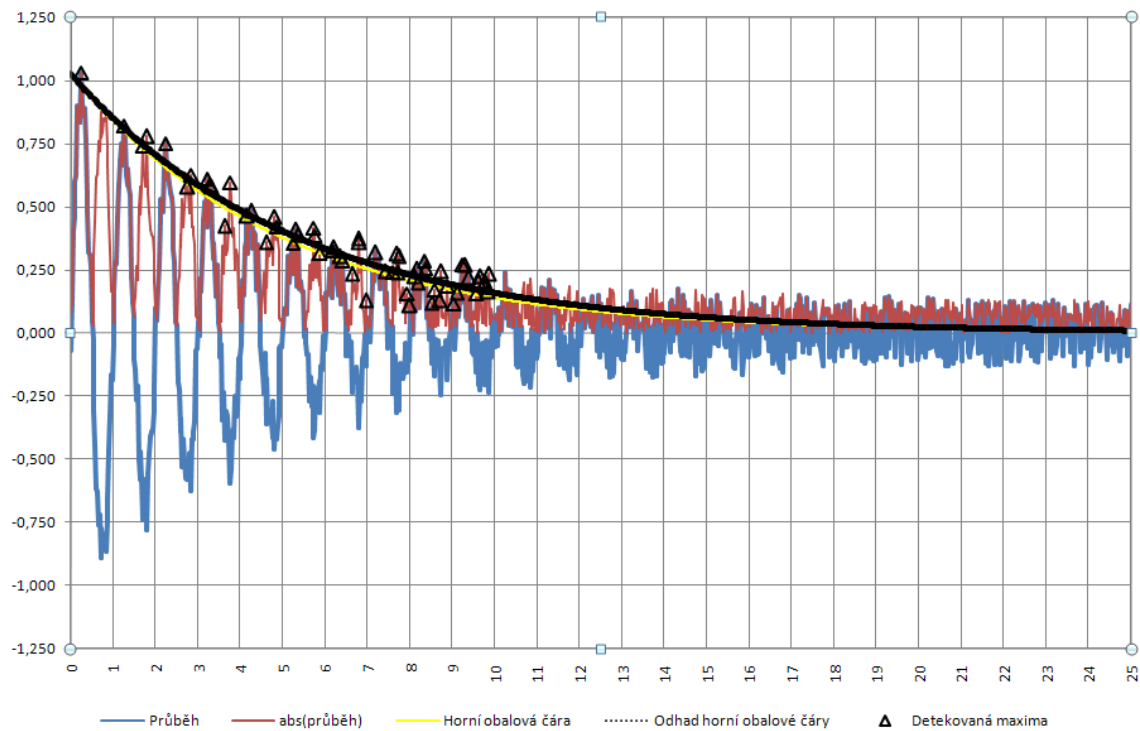
$$\ln(l_i) = C + Dp_i \quad (5.16)$$

Porovnáme-li tento vztah se vztahem 5.3.8, kdy $\zeta_{p_i} = 0$, můžeme psát, že platí

$$A = e^C \quad (5.17)$$

$$\tau = -\frac{1}{D}, \quad (5.18)$$

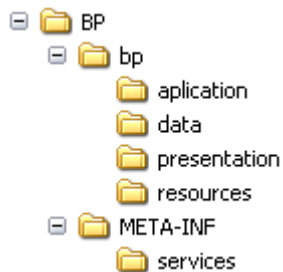
kde τ je tlumení signálu v jedné salvě.



Obr. 5.5 – Ukázka grafu pro koeficient tlumení.

6. Implementace

Program je realizován v jazyku Java, a proto jej tvoří spustitelný *jar* soubor, který má třívrstvou architekturu. Tuto architekturu zajišťují vytvořené balíky, tzv. package – *data*, *aplication* a *presentation*. UML diagram znázorňuje vztahy mezi jednotlivými třídami programu a je k dispozici v příloze A.



Obr. 6.1 – Struktura *jar* archivu.

V aplikaci jsou přidruženy externí knihovny *appframework-1.0.3*, *swing-worker-1.1*, *commons-math-2.0*, *mysql-connector-java-5.1.9-bin*. První dvě knihovny jsou přidány vývojovým prostředím *NetBeans* pro funkci GUI a zbylé dvě výpočet lineární regrese nebo konektivitu s databází.

6.1 data package

Data.java

Třída, která provádí transformaci původního signálu $s(k)$. Transformace se realizuje ve dvou průchodech. Prvním průchodem se vypočte střední hodnota ze vzorků, které se čtou po blocích velkých 22050 resp. 44100 prvků, kdy jeden vzorek tvoří dva byte. K vypočítání střední hodnoty se využívá datové struktury *HashMap*, s níž se provede statistika výskytu jednotlivých vzorků. Ve druhém průchodu se pak od každého vzorku signálu odečítá získaná střední hodnota. Transformované vzorky se pak ukládají jako čísla *double* do soboru *trnsfm.ini*, který má na prvním řádku nastavenou hlavičku ve tvaru:

```
[střední_hodnota] [maximum] [minimum] [počet_vzorků_poslední_ř.] [počet_ř.]
```

Na dalších řádcích za hlavičkou jsou již transformované vzorky uložené po 22050 prvcích, vyjma posledního, kde je počet prvků variabilní, ale zároveň nemůže být větší než u řádků předešlých.

V této třídě jsou pak dvě podtřídy definované jako dvě samostatná vlákna. *ThreadEf* pro provedení statistiky a výpočtu střední hodnoty a *ThreadTrans* obstarávající samotnou transformaci signálu.

Provedení této operace má dva možné stavy:

1. Čtení $s(k)$ z DB – úprava již uloženého signálu
2. Čtení $s(k)$ ze souboru wav na disku – nově nahraný signál

DataF.java

Totožná třída s třídou *Data.java*, kde se signál před transformací ještě podrobí filtraci filtrem FIR (viz kap. 3.3). S aktuálně vybraným vzorkem se posouvá i Hammingovo okno, které je implementováno pomocí pole. Problém, který je třeba při posunu okna řešit, je přechod mezi jednotlivými bloky (22050 prvků), po kterých se data čtou, a tím správně určit hodnotu vzorku na posledních resp. prvních pozicích starého resp. nového bloku. Proto je nutné uchovávat část vzorků z předchozího bloku, jejichž počet odpovídá polovině velikosti Hammingova okna.

DBHandler.java

Třída, která obstarává spojení a vykonává dotazy nad DB. K propojení MySQL databáze a jazyka Java se využívá knihovny *mysql-connector-java-5.1.9-bin*, která je volně dostupná na stránkách *mysql.com*, tak jako databáze samotná.

6.2 application package

ArithmeticAverage.java

Třída implementující metodu *Metoda aritmetického průměru*, jejímž vstupem jsou dva parametry – minimální velikost salvy a hranice průchodu. Tyto parametry nastavují schopnost metody vyhledat salvu(y) v signálu.

Pomocí API třídy „*Scanner*“ se čtou data z pomocného souboru *trnsfm.ini*. V prvním kroku se z hlavičky přečte hodnota maximálního (minimálního) vzorku vstupního signálu $s(k)$ a počet řádek v souboru, z nichž každá obsahuje 22050 vzorků. Výjimku tvoří poslední řádek, kdy se počet jeho vzorků také získává z hlavičky souboru. Po získání základních hodnot se přechází k samotnému čtení vzorků (vzorek po vzorku), přičemž se na každý implementuje metoda hledání (viz kap. 4.4). Samotná metoda je realizována v podtřídě *FindMethod*.

Dále se ve třídě počítá lineární regrese resp. parametr útlumu pomocí matematické knihovny *commons-math-2.0* resp. třídy *DampingRatio.java*. K ukládání pozic vzorků jakožto hranic salv se využívá API třídy *ArrayList*. Výsledek nalezených salv a parametry regrese resp. útlumu se zapisují do souboru *phonation.ini* v podobě vstupní parametry, začátek konec salvy a 1. a 2. parametr rovnic regrese resp. útlumu. Každý údaj na samostatný řádek.

DampingRatio.java

Třída, která slouží k výpočtení koeficientu tlumení (viz kap. 6). Opět využívá matematické knihovny *commons-math-2.0* k výpočtení lineární regrese, z níž se získaný parametr dosazuje do vzorce pro výpočet parametru amplituda a tlumení.

Energy.java

Třída, která počítá energii nalezených salv v $s(k)$ (viz kap. 6). Nejprve se ze souboru *phonation.ini* načtou pomocí API třídy *Scanner* pozice, kde jednotlivé salvy začínají resp. končí. Pak se po vzorcích prochází pomocný soubor *trnsfm.ini*. Ve chvíli, kdy se pozice aktuálního vzorku shoduje s pozicí začátku některé ze salv, začne se z každého následujícího (včetně aktuálního) počítat energie. Po nalezení konce dané salvy výpočet energie končí. Výsledná energie se zapisuje do souboru *en.ini* prostřednictvím API třídy *BufferedWriter*.

ThoughtBounds.java

Třída implementující metodu *Metoda počtu průchodů hranicí*, jejímž vstupem dva parametry – minimální velikost mezery mezi salvami a hranice průchodu. Tyto parametry nastavují schopnost metody vyhledat salvu(y) v signálu.

Pomocí API třídy „*Scanner*“ se čtou data z pomocného souboru *trnsfm.ini*. V prvním kroku se z hlavičky přečte hodnota maximálního (minimálního) vzorku vstupního signálu $\bar{s}(k)$ a počet řádek v souboru, z nichž každá obsahuje 22050 vzorků. Výjimku tvoří poslední řádek, kdy se počet jeho vzorků také získává z hlavičky souboru. Po získání základních hodnot se přechází k samotnému čtení vzorků (vzorek po vzorku), přičemž se na každý implementuje metoda hledání (viz kap. 4.2). Samotná metoda je realizována v podtřídě *FindMethod*.

Dále se ve třídě počítá lineární regrese resp. parametr útlumu pomocí matematické knihovny *commons-math-2.0* resp. třídy *DampingRatio.java*. K ukládání pozic vzorků jakožto hranic salv se využívá API třídy *ArrayList*. Výsledek nalezených salv a parametry regrese resp. útlumu se zapisují do souboru *phonation.ini* v podobě vstupní parametry, začátek konec salvy a 1. a 2. parametr regrese resp. útlumu. Každý údaj na samostatný řádek.

ThoughtBoundsSquare.java

Třída implementující metodu *Metoda počtu průchodů hranicí kvadrátu*, jejímž vstupem jsou dva parametry – minimální velikost mezery mezi salvami a hranice, které nastavují schopnost metody vyhledat salvu(y) v signálu.

Pomocí API třídy „*Scanner*“ se čtou data z pomocného souboru *trnsfm.ini*. V prvním kroku se z hlavičky přečte hodnota maximálního (minimálního) vzorku vstupního signálu $\bar{s}(k)$ a počet řádek v souboru, z nichž každá obsahuje 22050 vzorků. Výjimku tvoří poslední řádek, kdy se počet jeho vzorků také získává z hlavičky souboru. Po získání základních hodnot se přechází k samotnému čtení vzorků (vzorek po vzorku), přičemž se na každý implementuje metoda hledání (viz kap. 4.3). Samotná metoda je realizována v podtřídě *FindMethod*.

Dále se ve třídě počítá lineární regrese resp. parametr útlumu pomocí matematické knihovny *commons-math-2.0* resp. třídy *DampingRatio.java*. K ukládání pozic vzorků jakožto hranic salv se využívá API třídy *ArrayList*. Výsledek nalezených salv a parametry regrese resp. útlumu se zapisují do souboru *phonation.ini* v podobě vstupní parametry, začátek konec salvy a 1. a 2. parametr regrese resp. útlumu. Každý údaj na samostatný řádek.

6.3 presentation package

Canvas.java

Třída, která slouží k vykreslení průběhu vybraného signálu. K vykreslení je potřeba existence souborů *trnsfm.ini* s transformovaným signálem a *phonation.ini* s hranicemi salv a parametry útlumu energie salv k danému signálu. Ve třídě jsou pak dvě podtřídy. Třída *SampleBlock*, která se spouští jako několik vláken rozdělující načítání vzorků z *trnsfm.ini* na menší části. A druhá třída *PaintBlock*, která je implementována jako samostatné vlákno a GUI komponenta. Slouží k vykreslení samotného signálu, kdy se podle velikosti zobrazení volí počet (velikost pole) instancí těchto tříd. To znamená, že pokud budeme chtít zobrazit signál v plné velikosti, naskládá se několik těchto GUI komponent o stejné velikosti za sebe, přičemž poslední komponenta je větší nebo rovna velikosti ostatním, ale není větší než dvě komponenty vedle sebe.

Samotné kreslení pak probíhá tak, že se nastaví základní data (jako je maximum, minimum, apod.) z hlavičky pomocného souboru *trnsfm.ini*. Pak se díky třídě *PaintBlock* načtou vzorky do pole – podle velikosti zobrazení, tzn., jak velké bude podvzorkování. Nastaví se pole GUI komponent k vykreslení, načtou se hranice salv a parametry útlumu energie. Nakonec se provede samotné vykreslení signálu s hranicemi, extrémy, atd.

V této třídě je díky možnosti zvětšování signálu implementována metoda pro ruční detekci salv.

CardAndTab.java

Hlavní třída, která obsahuje hlavní okno aplikace (*JFrame*) a spojuje dohromady GUI třídy *Recording.java* a *Replaying.java* prostřednictvím záložkového menu (*JTabbedPane*), a současně obsahuje formulář pro záložku „karty pacientů“ spolu s veškerými interakcemi. Tato třída byla vygenerována (pouze rozvržení komponent formuláře) pomocí SW *NetBeans IDE 6.8* a posléze ručně upravena.

Record.java

Třída, která slouží k pořízení nového záznamu a také jako GUI pro záložku „nahrávání“ v aplikaci. Nejprve se nastaví rozmístění komponent na záložce a akce s nimi spojenými. Vnitřní třída *Recording* obstarává samotné nahrávání záznamu. Nahrávání je

v Javě realizováno prostřednictvím tzv. linky (třída *TargetDataLine*), kterou se data z mikrofonu dostávají do PC. Pro zahájení nahrávání a možnosti data číst, je nutné linku otevřít a spustit. Data jsou čtená po blocích (pomocí pole) a ukládána na danou pozici pomocí API třídy *RandomAccessFile*. První byte nahrávky je ukládán na pozici za hlavičkou wav, která má 44 byte. Hlavička samotná, je po ukončení nahrávání dopočítána a přidána na začátek souboru. Samotný soubor je uložen pod názvem *tmp-rec.wav* na disku ve složce *pacienti* a případně uložen do DB k danému pacientovi. Pro dobrou interakci uživatele s GUI při nahrávání je třída *Recording* implementována jako samostatné vlákno.

Replay.java

Třída, která slouží pro přehrávání záznamu wav a také jako GUI pro záložku „přehrávání“ v aplikaci. Záznam je možné přehrát buď přímo z disku po pořízení nového záznamu, který ještě není uložen v DB, anebo právě z DB, když se otevře starší záznam. Nejprve se nastaví rozmístění komponent na záložce a akce s nimi spojenými. Vnitřní třída *Replaying* obstarává samotné přehrávání záznamu a stejně jako u nahrávání je potřeba linky (třída *SourceDataLine*) k přenosu dat do zvukové karty. Data se čtou po blocích od pozice za hlavičkou ze streamu (*RandomAccessFile* nebo *Blob*) a zapisují se do linky. Pro dobrou interakci uživatele s GUI při nahrávání, je třída *Recording* implementována jako samostatné vlákno. Ve třídě jsou pak dále použity třídy *Canvas.java*, *Data.java* (*DataF.java*), *Energy.java* a *ThoughtBounds.java* (*ThoughtBoundsSquare.java*, *ArithmeticAverage.java*).

Sinus.java

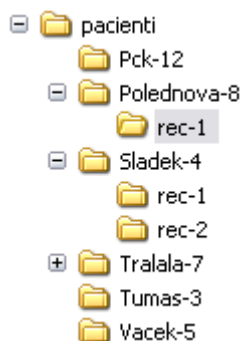
Třída, která se používá k vizualizaci sinusového průběhu při nahrávání nového wav a je použita ve třídě *Record.java*. Průběh funkce sinus se posouvá v čase tak, aby simuloval její pohyb. Před zahájením vizualizace se určí parametry sinusovky jako je počet period a počet bodů, které ji budou tvořit. Do datové struktury *ArrayList* se pak ukládají jednotlivé body sinusovky. Struktura se v cyklu prochází a jednotlivé body se používají jako souřadnice k vykreslování elementárních čar. Třída je pak implementována jako samostatné vlákno, kdy se s určitým zpožděním vykreslování bodů opakuje, a to tak, že se při každé iteraci cyklu ve vlákne připočte určitý fázový posun, který zajišťuje zdánlivý pohyb sinusovky.

MainApplication.java

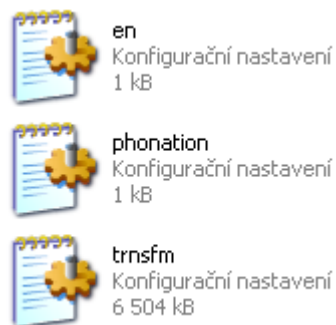
Třída vygenerovaná taktéž pomocí *NetBeans IDE 6.8*, která slouží ke spuštění celé aplikace. Využívá zde některé zdroje a knihovny používající právě *NetBeans* (např. *properties* soubory nebo *org.jdesktop.application*). Nachází se mimo uvedené balíky.

6.4 Datová úložiště

Každý pacient má na disku vyhrazený prostor pro pracovní adresář. Ten se nachází ve složce *pacienti*. Podoba pacientova adresáře je *příjmení-kód_pacienta* a obsahuje složky pro data jednotlivých záznamů. Názvy těchto podsložek jsou pak *rec-1*, *rec-2*, atd., v nichž se nachází soubory *trnsfm.ini* s transformovaným signálem $s(k)$, *phonation.ini*, kde se nachází hranice nalezených salv a k nim příslušné informace o úbytku energie (viz kap. 5) a soubor *energy.ini*, který obsahuje velikost energie každé salvy.

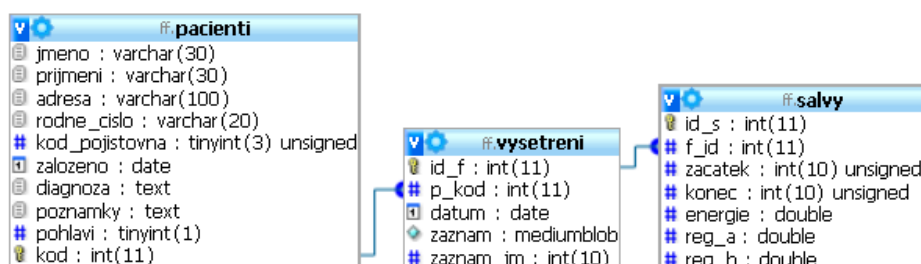


Obr. 6.2 – Struktura pracovního adresáře.



Obr. 6.3 – náhled složky „./pacienti/Polednova-8/rec-1“.

Dalším datovým úložištěm je DB MySQL, která je volně k dispozici na stránkách výrobce. Data se ukládají do databáze s názvem *ff* a obsahuje tři tabulky *pacienti*, *vysetreni* a *salvy*. V první tabulce se ukládají osobní informace o pacientech. Ve druhé to jsou údaje k jednotlivým vyšetřením daných pacientů, včetně nahraného záznamu. A ve třetí tabulce se uchovávají data k salvám – energie, hranice salv a parametry pro rovnice úbytku energie. DB pro uložení dat byla zvolena z důvodů bezpečnosti, cennosti pořízených záznamů a získaných dat.



Obr. 6.4 – Ukázka ERA modelu k DB, vygenerovaný programem phpMyAdmin.

7. Zpracování a vyhodnocení výsledků

7.1 Cíle hodnocení výsledků a testování

Cílem testování je ověření schopnosti metod detekovat počátek a konec salvy, tedy body C_x a D_x , v záznamu opakované dlouhodobé fonace vokálu „a“. Na základě nalezených bodů určit robustnost jednotlivých metod a vybrat jednu z řady možností pro testovací SW aplikaci.

K určení hranic salvy je řada možností, které se během práce implementovaly nebo naznačily. Na základě této myšlenky lze uplatnit další možnosti a docílit tak například lepších výsledků. K testování, kdy budeme testovat výsledky nalezení C_x a D_x pro vybrané verze záznamů fonace (zdraví / nemocní jedinci), se z množiny možností použijí následující kombinace:

- | | |
|--|---|
| 1. původní signál $s(k)$ | ... Metoda počtu průchodů hranic |
| 2. původní signál $s(k)$ | ... Metoda počtu průchodů hranic kvadrátu |
| 3. původní signál $s(k)$ | ... Metoda aritmetického průměru |
| 4. filtrovaný signál $s^*(k)$, $M+1=21$ | ... Metoda počtu průchodů hranic |
| 5. filtrovaný signál $s^*(k)$, $M+1=21$ | ... Metoda počtu průchodů hranic kvadrátu |
| 6. filtrovaný signál $s^*(k)$, $M+1=21$ | ... Metoda aritmetického průměru |

O schopnosti automatické detekce jednotlivých metod rozhodujeme na základě porovnání výsledků ruční a automatické detekce. Porovnáním výsledků jsme schopni určit přesnost metod automatické detekce a podle úspěšnosti nalezení správného počtu salv ve fonaci, popisujeme jejich robustnost. Z toho vyplývá, že asistence experta pro ruční určení hranic má pro celkové testování zásadní význam.

Testování bude prováděno pro záznamy zdravých jedinců (studentů) a záznamy pořízené na ORL klinice FN Plzeň. V době řešení bakalářské práce bylo k dispozici pět záznamů pacientů s diagnózou paréza zvrtného nervu, spec. po měsíci hlasové rehabilitace, kde na prvních dvou zvukových souborech jsou zaznamenány ženy a na posledním muž. Průběh signálu je vlivem různého stádia onemocnění u každého ze záznamů odlišný, a obsahuje vždy 10 salv. K popisu metod z kategorie zdravých jedinců je vybrán jediný zastupující záznam (č. 6), také s počtem 10 salv.

Ačkoliv se může počet nahrávek zdát pro důkladné otestování malý, tak při volbě kombinací nastavení parametrů testování, aplikovaných na uvedených metodách, se dostáváme k poměrně velkému počtu výstupních kombinací. Vyhodnocení metod bude probíhat na základě vstupních a výstupních parametrů. Stěžejními jsou právě nalezené hranice salv, vyjádřené jako pozice vzorků v signálu, od nichž se pak odvíjí vypočtená energie nebo parametry útlumu energie. Pro jednoduchost budou výstupní parametry vypsány pro první (potenciální) salvu, tj. C_x a D_x pro $x = 1$.

7.2 Získaná data

Výstupní parametry (energie, útlum energie) nemají v tomto kontextu z hlediska testování větší význam, a proto nejsou pro ruční detekci počítány. Tyto parametry by pak měly potvrdit či vyvrátit stanovené předpoklady o útlumu fonace, tj. jednotlivých salv, z kap. 3.1.

7.2.1 Ruční detekce původní signál $s(k)$

Záz. č.	Pok. č.	Výstupní parametry		
		Salvy		Doba salvy [s]
		p.	Př. (1. salva)	
1	1	10	49896 - 224798	3.97
2	1	10	30018 - 307013	6.28
3	1	10	40678 - 394375	8.29
4	1	10	55978 - 190598	3.05
5	1	10	43134 - 188369	3.29
6	1	10	83639 - 454072	8.40

7.2.2 Automatická detekce původní signál $s(k)$

Metoda počtu průchodů hranicí:

Záz. č.	Pok. č.	Vstupní parametry		Výstupní parametry					
		hranice [%]	pauza [ms]	Salvy		Energie	Doba salvy [s]	ΔE abs. hod.	ΔE útlum
				p.	Př. (1. salva)				
1	1	0.300	350.0	10	50146 - 224518	6.01E7	3.95	-0.00760	-8.322E5
	2	0.300	200.0	10	50146 - 224518	6.01E7	3.95	-0.00760	-8.322E5
	3	0.200	200.0	10	50146 - 225588	6.01E7	3.98	-0.00846	-6.169E5
2	1	0.300	300.0	10	30276 - 306725	7.372E7	6.27	-0.00258	-1.803E6
	2	0.300	200.0	10	30276 - 306725	7.372E7	6.27	-0.00258	-1.803E6
	3	0.200	200.0	10	30057 - 313328	7.376E7	6.42	-0.00465	-7.782E5
3	1	0.300	400.0	10	29269 - 393825	1.6E8	8.27	-0.00452	-1.172E6
	2	0.300	200.0	10	41620 - 393825	1.60E8	7.99	-0.00461	-1.142E6
	3	0.200	200.0	10	41159 - 394330	1.60E8	8.00	-0.00651	-6.743E5
4	1	0.300	500.0	10	59637 - 189258	5.11E7	2.94	-0.02696	-2.482E5
	2	0.200	400.0	10	56871 - 189447	5.12E7	3.01	-0.02348	-2.483E5
5	1	0.400	300.0	6	27393 - 202964	1.94E7	3.98	-0.01172	-2.684E5
	2	0.650	150.0	10	42399 - 187389	1.94E7	3.29	-0.01031	-3.351E5
6	1	0.300	500	10	83786 - 453926	4.830E7	8.39	0.00979	3.939E5
	2	0.200	500	10	83812 - 454323	4.831E7	8.40	0.00937	3.736E5

Metoda počtu průchodů hranicí kvadrátu:

Záz. č.	Pok. č.	Vstupní parametry		Výstupní parametry					
		hranice [%]	pauza [ms]	Salvy		Energie	Doba salvy [s]	ΔE abs. hod.	ΔE útlum
				p.	Př. (1. salva)				
1	1	0.300	500.0	7	50146 - 404649	1.20E8	8.04	-0.00117	-1.654E6
	2	0.300	400.0	9	50146 - 404649	1.20E8	8.04	-0.00117	-1.654E6
	3	0.300	350.0	10	50146 - 224518	6.01E7	3.95	-0.00760	-8.322E5
	4	0.200	350.0	7	50146 - 405841	1.20E8	8.07	-0.00227	-1.599E6
2	1	0.300	500.0	5	10904 - 306725	7.374E7	6.71	-0.00253	-1.847E6
	2	0.300	300.0	10	30276 - 306725	7.372E7	6.27	-0.00258	-1.803E6
	3	0.200	300.0	10	30057 - 313328	7.376E7	6.42	-0.00465	-7.780E5
3	1	0.300	500.0	8	29269 - 393825	1.60E8	8.27	-0.00452	-1.172E6
	2	0.300	450.0	9	29269 - 393825	1.60E8	8.27	-0.00452	-1.172E6
	3	0.300	400.0	10	29269 - 393825	1.60E8	8.27	-0.00452	-1.172E6
	4	0.200	400.0	8	29159 - 394330	1.60E8	8.28	-0.00599	-7.532E5
4	1	0.300	500.0	10	59637 - 189258	5.12E7	2.94	-0.02696	-2.482E5
	2	0.200	300.0	10	56871 - 189447	5.12E7	3.00	-0.02348	-3.076E5
5	1	0.300	500.0	4	27391 - 205601	1.94E7	4.04	-0.01199	-2.422E5
	2	0.400	300.0	6	27393 - 202964	1.94E7	3.98	-0.01172	-2.684E5
	3	0.650	200.0	10	42399 - 187389	1.94E7	3.29	-0.01031	-3.351E5
6	1	0.300	500	10	83786 - 453926	4.830E7	8.39	0.00979	3.939E5
	2	0.200	500	10	83812 - 454323	4.831E7	8.40	0.00937	3.736E5

Metoda aritmetického průměru:

Záz. č.	Pok. č.	Vstupní parametry		Výstupní parametry					
		hranice [%]	salva [ms]	Salvy		Energie	Doba salvy [s]	ΔE abs. hod.	ΔE útlum
				p.	Př. (1. salva)				
1	1	0.300	500.0	14	13886 - 39051	9.209E4	0.57	-	-
	2	0.300	600.0	12	13886 - 44101	1.087E5	0.69	-	-
	3	0.300	900.0	10	50124 - 227032	6.012E7	4.01	-0.0073	-8.794E5
	4	0.200	900.0	10	50117 - 227032	6.012E7	4.01	-0.0082	-6.336E5
2	1	0.300	900.0	-	-	-	-	-	-
	2	0.300	1200.0	-	-	-	-	-	-
	3	0.200	1200.0	-	-	-	-	-	-
3	1	0.300	500.0	13	41332 - 394787	1.60E8	8.01	-0.0049	1.065E6
	2	0.300	900.0	11	41332 - 394787	1.60E8	8.01	-0.0049	-1.065E6
	3	0.300	1200.0	10	41332 - 394787	1.60E8	8.01	-0.0049	-1.065E6
	4	0.200	1200.0	10	41157 - 394787	1.60E8	8.02	-0.0068	-6.514E5
4	1	0.300	500.0	10	59637 - 191363	5.11E7	2.99	-396.450	-2.494E5
	2	0.300	900.0	10	59637 - 191363	5.11E7	2.99	-396.450	-2.494E5
5	1	0.300	500.0	10	42132 - 188442	1.94E7	3.32	-94.892	-2.249E5
	2	0.650	900.0	10	42399 - 188442	1.94E7	3.31	-79.939	-3.331E5
6	1	0.300	500	10	83786 - 454557	4.831E7	8.41	100.441	3.939E5
	2	0.200	900	10	83755 - 454557	4.831E7	8.41	92.254	3.719E5

7.2.3 Ruční detekce filtrovaný signál $s^*(k)$, $M+1=21$

Záz. č.	Pok. č.	Výstupní parametry		
		Salvy		Doba salvy [s]
		p.	Př. (1. salva)	
1	1	10	49886 - 224233	3.95
2	1	10	30017 - 307367	6.29
3	1	10	40680 - 394916	8.03
4	1	10	56201 - 190589	3.05
5	1	10	43135 - 188230	3.29
6	1	10	83638 - 454062	8.40

7.2.4 Automatická detekce filtrovaný signál $s^*(k)$, $M+1=21$

Metoda počtu průchodů hranic:

Záz. č.	Pok. č.	Vstupní parametry		Výstupní parametry					
		hranice [%]	pauza [ms]	Salvy		Energie	Doba salvy [s]	ΔE abs. hod.	ΔE útlum
				p.	Př. (1. salva)				
1	1	0.300	500.0	8	50147 - 404046	1.200E8	8.02	-6.539	-5.590E6
	2	0.300	200.0	10	50147 - 223976	6.001E7	3.94	-0.008	-8.165E5
	3	0.200	200.0	10	50146 - 225031	6.010E7	3.97	-0.008	-7.321E5
2	1	0.300	500.0	5	10904 - 307100	7.374E7	6.72	-0.0032	-1.346E6
	2	0.300	200.0	10	30264 - 307100	7.373E7	6.28	-0.0032	-1.323E6
	3	0.200	200.0	10	30056 - 315166	7.376E7	6.47	-0.0053	-6.474E5
3	1	0.300	500.0	8	29201 - 393830	1.601E8	8.27	-0.0050	-9.992E5
	2	0.300	200.0	10	41331 - 393830	1.601E8	7.99	-0.0052	-9.590E5
	3	0.200	200.0	10	28815 - 394369	1.601E8	8.29	-0.0061	-6.895E5
4	1	0.300	500.0	10	58992 - 189260	5.115E7	2.95	-0.0255	-2.669E5
	2	0.200	300.0	10	56870 - 189615	5.124E7	3.01	-0.0248	-2.813E5
5	1	0.400	300.0	6	27393 - 202964	1.939E7	3.98	-0.0117	-2.669E5
	2	0.650	200.0	10	42399 - 187388	1.936E7	3.28	-0.0101	-3.448E5
6	1	0.300	500	10	83772 - 453928	4.830E7	8.39	0.0098	3.839E5
	2	0.200	500	10	83812 - 454325	4.831E7	8.40	0.0093	3.740E5

Metoda počtu průchodů hranicí kvadrátu:

Záz. č.	Pok. č.	Vstupní parametry		Výstupní parametry					
		hranice [%]	pauza [ms]	Salvy		Energie	Doba salvy [s]	ΔE abs. hod.	ΔE útlum
				p.	Př. (1. salva)				
1	1	0.300	500.0	8	50147 - 404046	1.20E8	8.02	-6.539	-5.590E6
	2	0.300	200.0	10	50147 - 223976	6.008E7	3.94	-0.00804	-8.165E5
	3	0.200	200.0	10	50146 - 225031	6.010E7	3.97	-0.00797	-7.321E5
2	1	0.300	500.0	5	10904 - 307100	7.374E7	6.72	-0.00315	-1.346E6
	2	0.300	200.0	10	30264 - 307100	7.373E7	6.28	-0.00319	-1.323E6
	3	0.200	200.0	10	30056 - 315166	7.376E7	6.47	-0.00528	-6.474E5
3	1	0.300	500.0	8	29201 - 393830	1.601E8	8.27	-0.00502	-9.992E5
	2	0.300	200.0	10	41331 - 393830	1.601E8	7.99	-0.00517	-9.590E5
	3	0.200	200.0	10	28815 - 394369	1.601E8	8.29	-0.00618	-6.895E5
4	1	0.300	500.0	10	58992 - 189260	5.115E7	2.95	-0.0255	-2.669E5
	2	0.200	300.0	10	56870 - 189615	5.124E7	3.01	-0.0248	-2.813E5
5	1	0.400	300.0	6	27393 - 202964	1.939E7	3.98	-0.0117	-2.669E5
	2	0.690	200.0	10	43324 - 187153	1.935E7	3.26	-0.0095	-3.448E5
6	1	0.300	500	10	83772 - 453928	4.830E7	8.39	0.0098	3.839E5
	2	0.200	500	10	83812 - 454325	4.831E7	8.40	0.0093	3.740E5

Metoda aritmetického průměru:

Záz. č.	Pok. č.	Vstupní parametry		Výstupní parametry					
		hranice [%]	salva [ms]	Salvy		Energie	Doba salvy [s]	ΔE abs. hod.	ΔE útlum
				p.	Př. (1. salva)				
1	1	0.300	500.0	12	13886 - 39051	9.209E4	0.57	-	-
	2	0.300	900.0	10	50126 - 227032	6.012E7	4.01	-0.0076	-8.910E5
	3	0.200	900.0	10	50121 - 227032	6.012E7	4.01	-0.0078	-7.494E5
2	1	0.300	900.0	-	-	-	-	-	-
	2	0.300	1200.0	-	-	-	-	-	-
	3	0.200	1200.0	-	-	-	-	-	-
3	1	0.300	500.0	13	41328 - 394787	1.60E8	8.01	-0.0055	-9.062E5
	2	0.300	1000.0	10	41328 - 394787	1.60E8	8.01	-0.0055	-9.062E5
	3	0.300	1200.0	10	41328 - 394787	1.60E8	8.01	-0.0055	-0.62E5
	4	0.199	1200.0	10	41156 - 394787	1.60E8	8.02	-0.0072	-5.673E5
4	1	0.300	500.0	10	58992 - 191363	5.118E7	3.00	-379.644	-2.675E5
	2	0.300	900.0	10	58992 - 191363	5.118E7	3.00	-379.644	-2.675E5
5	1	0.300	500.0	10	42132 188442	1.94E7	3.32	-94.892	-2.249E5
	2	0.650	900.0	10	42399 - 188442	1.94E7	3.31	-79.939	-3.331E5
6	1	0.300	500	10	83772 - 454557	4.831E7	8.41	0.0098	3.846E5
	2	0.200	900	10	83754 - 454557	4.831E7	8.41	0.0093	-

7.3 Hodnocení

Soustředíme se nyní na metody bez filtrace. Vezmeme-li metody *Metoda počtu průchodů hranicí* a *Metoda počtu průchodů hranicí kvadrátu*, tak je zřejmé, že obě dávají v hledání salv shodné výsledky (při nastavení stejných vstupních hodnot). To potvrzuje předpoklad o rovnosti chyb v kap. 4.5.

Porovnáme-li výsledky s hodnotami ruční detekce, tak vidíme, že v případě nalezení začátku nebo konce se výsledky liší řádově v několika stovkách vzorků, což je z pohledu vzorkovací frekvence $f_{vz} = 44100$ Hz zanedbatelné. U metody *Metoda aritmetického průměru* je druhý nastavovaný parametr odlišný, a proto nelze jednoznačně říci, zdali je tato metoda ve srovnání s předešlými přesnější. Na základě porovnání s ruční detekcí jsou pak rozdíly pro některé nahrávky větší a pro některé naopak menší, nic méně stále v řádech stovek vzorků.

Problém metody *Metoda aritmetického průměru* oproti prvním dvěma je jasný. Při špatném nastavení vstupních dat nemusíme některých výsledků docílit, anebo metoda nedokáže najít žádnou ze salv ve fonaci. Z tohoto plyne, že přesnost metod pak nezáleží jen na nastavení vstupních parametrů, ale dost výraznou měrou výsledek ovlivňuje i podoba vstupního signálu. Tato závislost pak pro 3. metodu znamená, že ji nelze označit jako za příliš robustní a spolehlivou. U zbývajících dvou metod je tomu naopak, kdy se hranicím salv dospělo vždy.

U metod *Metoda počtu průchodů hranicí* a *Metoda počtu průchodů hranicí kvadrátu*, které jsou aplikované na filtrovaný signál, se získané výsledky nijak nerozbíhají a opět jsou stejné. V porovnání s předchozí možností (bez filtru) jsou vyznačené salvy nepatrně menší, což je způsobeno ořezem filtru. Lze očekávat, že při vhodném nastavení hranic metod, dostaneme výsledky začátku a konce salv stejné. U metody *Metoda aritmetického průměru*, ve schopnosti hledání, nenabídla filtrace bohužel žádná zlepšení. V případech, kdy metoda salvy vyhledala, pak poskytuje podobně jako u prvních dvou metod (vlivem filtrace) salvy o něco menší.

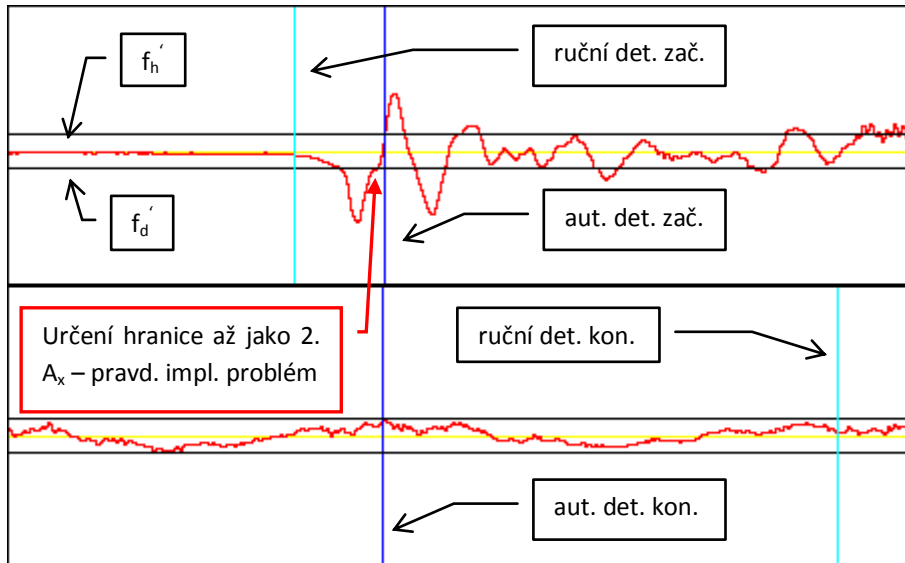
Zhodnotíme-li použitelnost metod, tak pro použití v aplikaci jednoznačně mluví metody *Metoda počtu průchodů hranicí* a *Metoda počtu průchodů hranicí kvadrátu* bez filtrace, které dávají spolehlivé výsledky, jejichž přesnost prakticky závisí pouze na vstupních parametrech. Metody s filtrací, pak poskytnou přesnější výsledky, kdy je odfiltrována šumová složka signálu. *Metoda aritmetického průměru*, pak u dvou záznamů neobstála. Hlavní problém spočívá ve správném nalezení konce salvy a z výsledků je metoda pro nahrávky nemocných zřejmě nepoužitelná. Pro zdravý hlas poskytuje dobré výsledky, podobné výsledkům prvních dvou metod, které jsou dokonce v některých případech lepší – to ovlivňují vstupní parametry. Problém vyhledání hranic fonace nemocných pacientů je zřejmě implementační záležitost nebo by bylo vhodné použít lepší filtraci.

Časové složitosti metod:

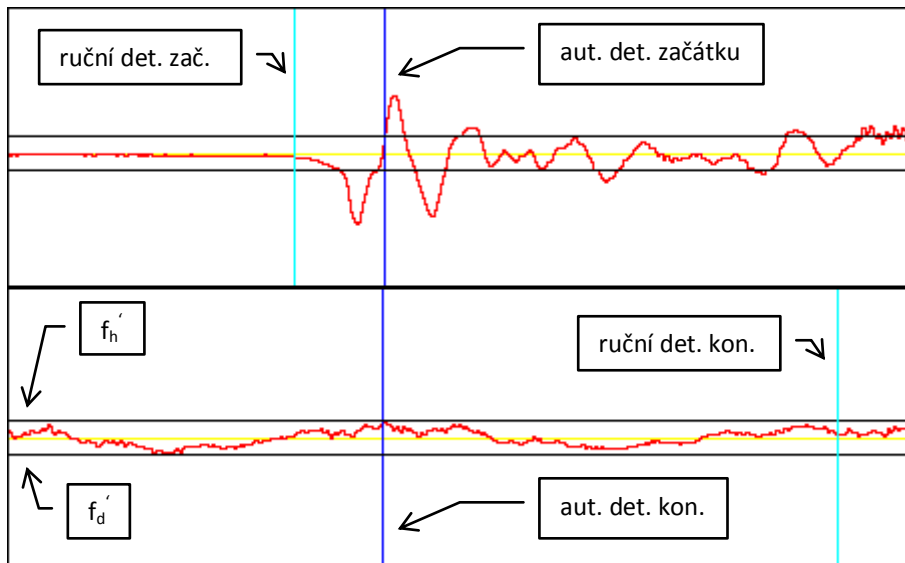
1. Metoda počtu průchodů: $O(n)$, kde n je počet vzorků v signálu
2. Metoda počtu průchodů kvadrátu: $O(n)$, kde n je počet vzorků v signálu
3. Metoda aritmetického průměru: $O(n + km)$, kde n je počet vzorků v signálu, m je velikost okna a k počet průchodů oknem (nalezení začátku, konce)

7.4 Náhledy z navrženého SW

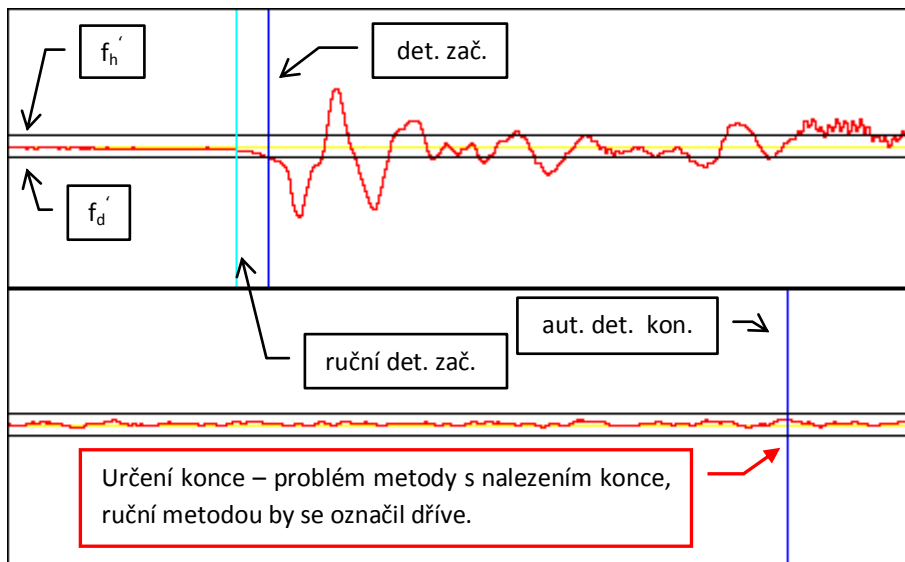
Automatické určení bodů C_x a D_x pro 1. záznam 1. salvy – bez filtrace ($s(k)$) a s filtrací ($s^*(k)$). Náhledy se týkají tučně zvýrazněných výsledků v tabulce.



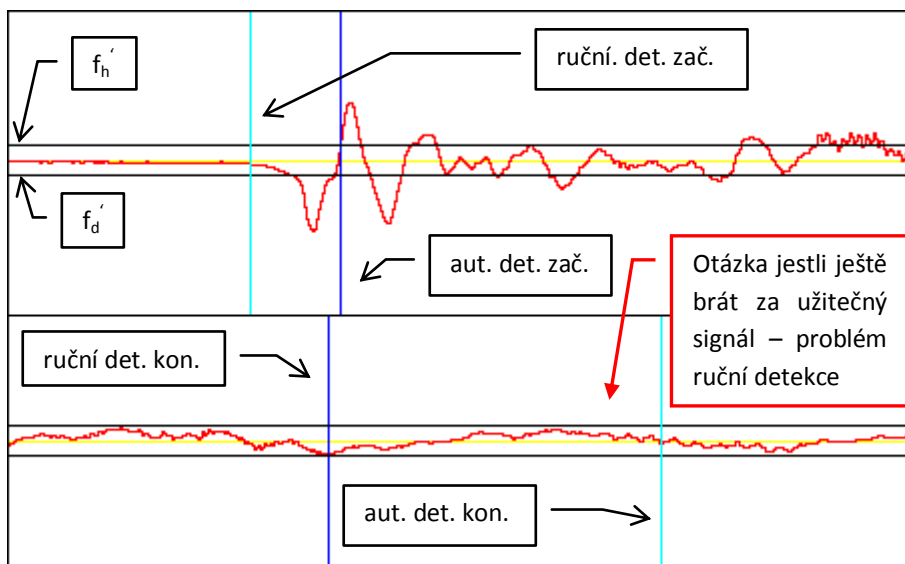
Obr. 7.1 – Metoda počtu průchodů hranicí pro $s(k)$ - nalezení začátku a konce, spolu s ruční detekcí.



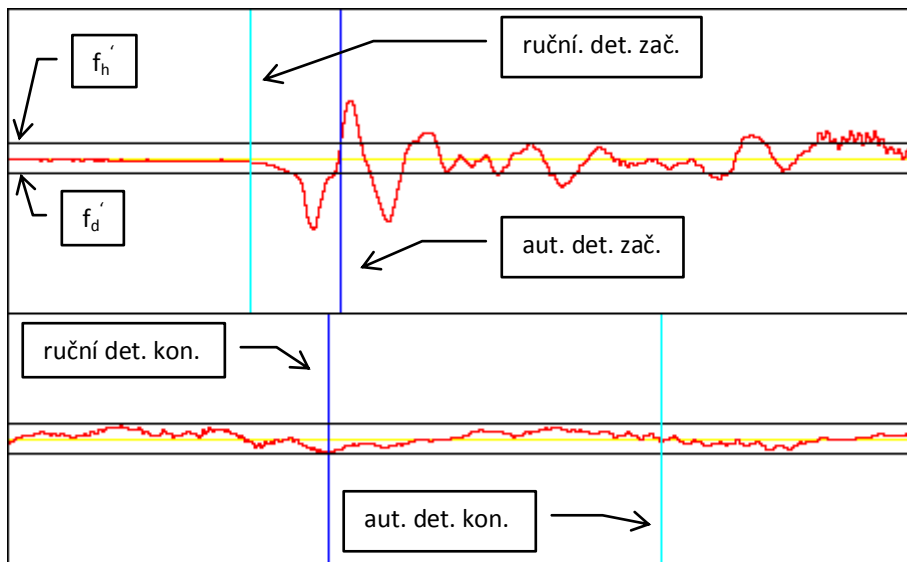
Obr. 7.2 – Metoda počtu průchodů hranicí kvadrátu pro $s(k)$ - nalezení začátku a konce, spolu s ruční detekcí. Obrázek potvrzuje shodnost výsledků hledání s metodou výše (obr. 7.1).



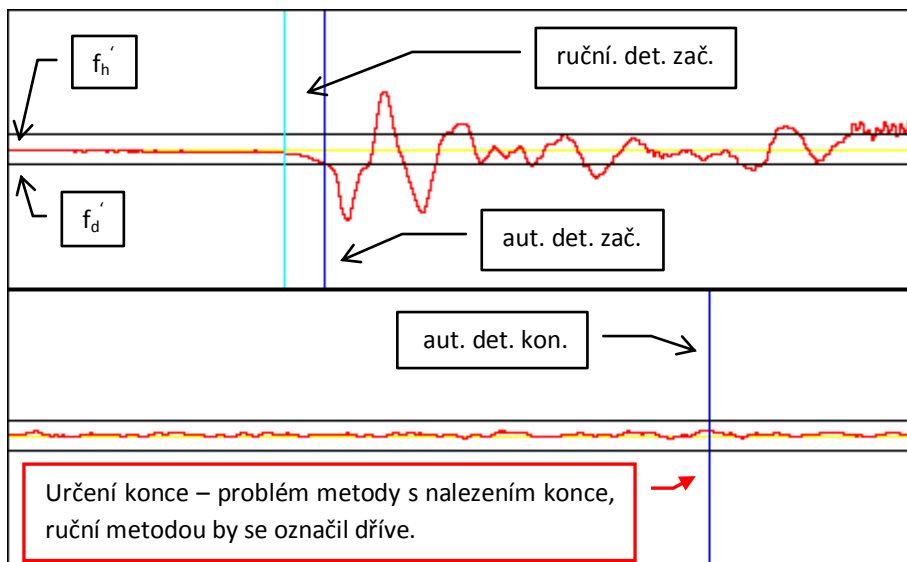
Obr. 7.3 – Metoda aritmetického průměru pro $s(k)$ - nalezení začátku a konce, spolu s ruční detekcí. Nalezení konce s velkou chybou.



Obr. 7.4 – Metoda počtu průchodů hranicí pro $s^*(k)$ - nalezení začátku a konce, spolu s ruční detekcí. Z obrázku není zřejmý výrazný rozdíl ve srovnání nefiltrovaným signálem (na základě dat z tabulky je nepatrný).



Obr. 7.5 – Metoda počtu průchodů hranicí kvadrátu pro $s^*(k)$ - nalezení začátku a konce, spolu s ruční detekcí. Obrázek potvrzuje shodnost výsledků hledání s metodou výše (obr. 7.4). Rozdíly s nefiltrovaným záznamem jsou opět nepatrné.



Obr. 7.6 – Metoda aritmetického průměru pro $s^*(k)$ - nalezení začátku a konce, spolu s ruční detekcí. Nalezení konce s velkou chybou. Filtrace nepřinesla o moc lepší výsledky oproti signálu nefiltrovanému.

Průběhy všech testovaných záznamů viz příloha C.

Možností, jak upravit automatickou detekci metod, je stanovením hranic průchodu zvláště pro začátek a konec salvy. Z analýzy signálu je patrné, že hranice průchodu určená společně pro obě meze salvy nemusí být úplně vyhovující. Začátek salvy je v mnohých případech oproti konci jednoznačnější, a proto pro jeho detekci může být hranice větší. To naopak nevyhovuje detekci konce, kdy je více zatížen šumem a i ruční určení meze může činit problémy.

Další z možností hledání začátku a konce slov je tzv. Určování hraničních bodů promluvy, viz [6], s. 71 až 73. Tato metoda je založena na stanovení hranic H_M , která určuje horní mez intenzity a H_Z určující horní mez počtu průchodů signálu nulou. Analogicky jsou stanoveny i dolní meze D_M a D_Z . *Překročí-li intenzita signálu mez H_M v n mikrosegmentech za sebou (vhodné se ukazuje $n = 3$ až 5), lze usuzovat na výskyt začátku slova. Je-li začátek slova tvořen hláskou s vysokou frekvencí, ale malou intenzitou (např. hlásky „f“, „s“ apod.), projeví se to zvýšenou hodnotou středního počtu průchodu signálu nulou, takže počáteční hraniční bod bude dodatečně odvozen od z překročení H_Z .* Obdobným způsobem se určuje i konec slova, kdy se pracuje s hranicemi D_M a D_Z . Vhodnou modifikací této metody, bychom byli schopni její aplikace na naše případy, které jsou zjednodušeny o fakt, že pracujeme se signálem obsahující vokály (pouze samohlásky). Na rozdíl od její stávající podoby, která počítá s určením hranic celých slov, která jsou sama o sobě výrazně zatížena šumem. Na druhou stranu to pro nás může představovat výhodu, kdy zkoumané nahrávky pacientů obsahují také značný šum.

8. Závěr

Na základě získaných poznatků z kap. 7 jsme jako použitelné metody detekce počátku a konce salvy vyhodnotili následující metody: *Metoda počtu průchodů* a *Metoda počtu průchodů kvadrátu*. Aplikací těchto navržených metod byl na testovaných datech vždy detekován správný počet salv, bez ohledu na podobu vstupního signálu. Lze je proto v této fázi testování považovat za robustní ve smyslu předpokladů a typu zpracovávaných signálů.

Vzhledem k dosažené přesnosti ve srovnání s definovanými hranicemi salvy znalcem, budeme dále testovat v praxi ověřovat právě tyto metody.

Výstupními parametry, které vychází z předpokladu existence nalezených salv, jsou energie, útlum energie, délka salvy a pauzy. Na základě jejich vývoje se pak ověřuje předpoklad z kap 3.1 o vlivu zátěžového testu na hlasivky vyšetřovaného s postupujícím časem. V tabulkách kap. 7 jsou zapsány parametry lineární regrese a resp. útlumu τ . Pokud platí $a > 0$, tak salva roste v čase, pokud $a < 0$ salva klesá. Jestliže $a = 0$ je útlum energie nulový. Parametr τ pak určuje míru útlumu, která je vyjádřena jako exponenciální křivka. Pokud platí $\tau > 0$, tak salva roste v čase, pokud $\tau < 0$ salva klesá. Délka salvy resp. pauzy je pak úměrná počtu vzorků daného segmentu. S ohledem na potvrzený předpoklad z kap 3.1, se délka salvy zmenšuje (nebo intenzita – obecně energie) a délka pauzy mezi salvami roste.

Použitím filtrů FIR, k předzpracování vstupního signálu, se pak hledání začátku a konce salv zpřesňuje, kdy se odstraní rušivá složka signálu. Pro dosažení lepších výsledků lze zvážit použití filtru s větším počtem koeficientů Hammingova okna. S použitím většího Hammingova okna je však třeba počítat s úměrným zvětšením doby potřebné k provedení výpočtů, což nepředstavuje problém vzhledem k práci s daty, které nejsou reálného času.

Z pohledu použití metody prohledání salv ve fonaci se pak přistoupilo v SW k užití metody *Metoda počtu průchodů kvadrátu*. Útlum energie se realizuje pomocí přímky lineární regrese. Důvod je snazší implementace a fakt, že tyto metody byly v aplikaci použity od počátku její realizace.

Zmíněná metoda v kap. 7, tzv. Určování hraničních bodů promluvy, by se mohla uplatnit jako rozšíření stávajících metod, kdy by se po nalezení hranic salv, použila k přesnějšímu určení nalezených začátků a konců v kritických oblastech šum-salva.

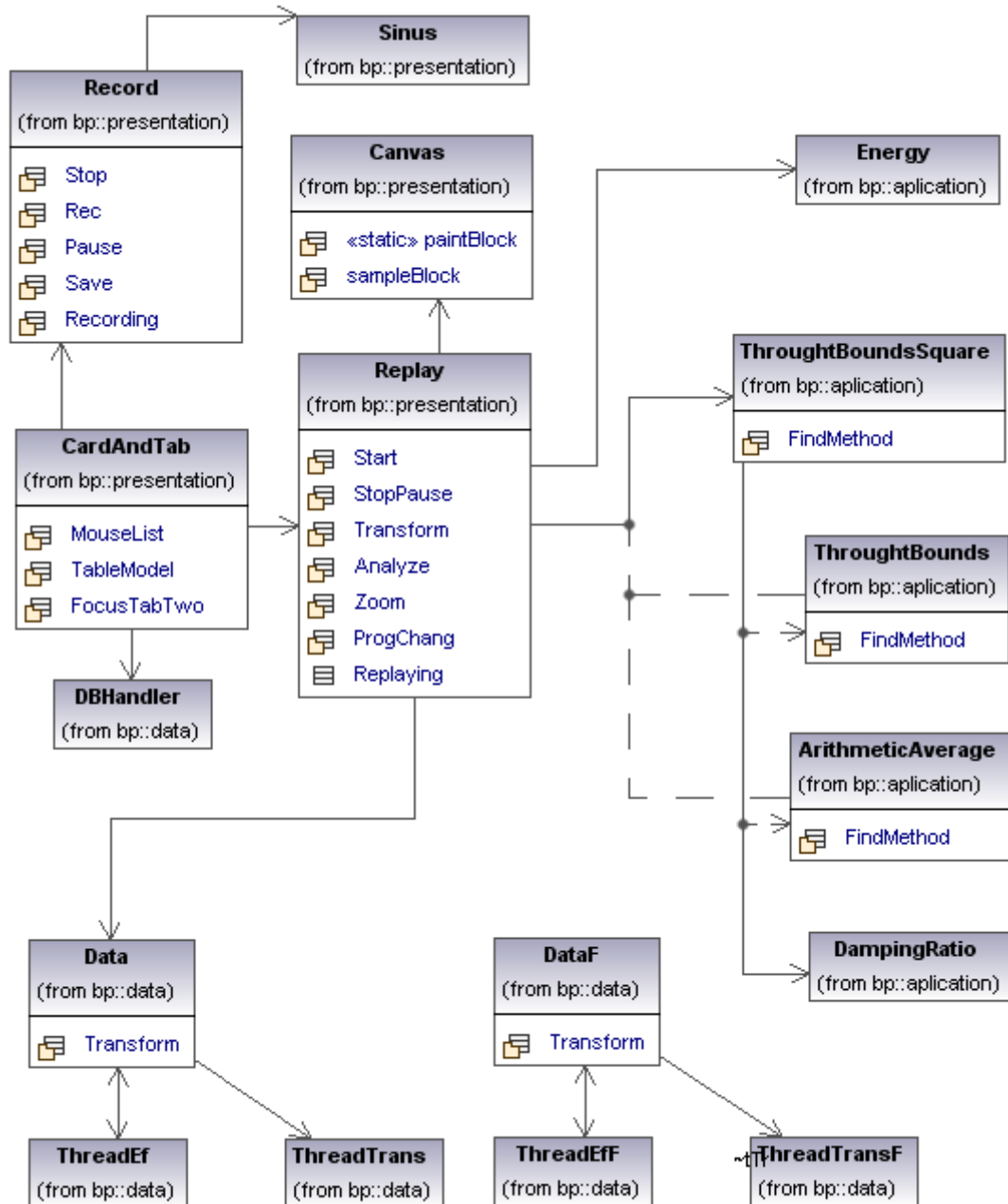
Literatura a zdroje

- [1] Hlaváč, V., Sedláček, F., *Zpracování Signálů a obrazů*, ISBN 80-01-02114-9, ČVUT Praha 2001.
- [2] Hradecká, Z., *Fyziologie lidského hlasového ústrojí*, Brno, 2007. Diplomová práce, Ústav experimentální biologie, oddělení fyziologie a imunologie živočichů. Vedoucí diplomové práce Monika Dušková.
- [3] Míka, S., Brandner, M., *Numerické metody I*, Západočeská univerzita v Plzni, 2000.
- [4] Novák, A., *Foniatric a pedaudiologie II*, Unitisk Praha 1996.
- [5] Nový, P., Vávra, F., Kotlíková, M, *Voice range profile examination method and its applications*, Summer School DATASTAT 03, Proceedings, Folia Fac. Sci. Nat., Univ. Masaryk. Brunensis, Mathematica 15, ISBN 80-210-3564-1, Svratka, 2003.
- [6] Psutka, J., *Komunikace s počítačem mluvenou řečí*, ISBN 80-200-0203-0, Academia Praha 1995.
- [7] Reif, J., Kobeda, Z., *Úvod do pravděpodobnosti a spolehlivosti*, ISBN 80-7043-333-7, Západočeská univerzita v Plzni, 1996.
- [8] Šlapák, I., Janeček, D., Lavička, L., *Základy otorinolaryngologie a foniatric pro studenty speciální pedagogiky*, Elportál, Brno, Masarykova univerzita, 2009, ISSN 1802-128X.
- [9] Švec, J., *Fyziologická akustika zpěvního hlasu: Nový pohled na starý problém*. 60. akustický seminář & 36. akustická konference, ISBN 80-01-02179-3, Kouty, 2000.
- [10] Švec, J., *On Vibration Properties of Human Vocal Folds: Voice Registers, Bifurcations, Resonance Characteristics, Development and Application of Videokymography*, ISBN 90-367-1235-1, Olomouc, 2000
- [11] Švec, J., *Studium mechanicko-akustických vlastností zdroje lidského hlasu*, Olomouc, 1996. Disertační práce na přírodovědecké fakultě Univerzita Palackého v Olomouci na katedře experimentální fyziky, Školitel disertační práce Jaroslav Pospíšil.
- [12] Vávra, F., Plzeň, *Diagnostické metody hodnocení a porovnávání kvality lidského hlasu před a po léčbě*, Plzeň, 1999. Diplomová práce, Fakulta aplikovaných věd ZČU na katedře kybernetiky, Vedoucí diplomové práce Jiří Pešta.

- [13] Když se řekne: Internetový *magazín o automatizaci*. Č.7-8 (červenec–srpen 2005).
URL: <<http://www.automatizace.cz/download.php?d=QXRtX0FydGljbGUscGRmX2FydCw3OTI>> => [cit. 2010-30-5].
- [14] *Matematická knihovna (lineární regrese)*
URL: <<http://commons.apache.org/math/userguide/stat.html>> [2010-20-5].
- [15] *MySQL DB a knihovna pro propojení - Connector/J*
URL: <<http://www.mysql.com/downloads/>> [2010-20-5].
- [16] *WAVE PCM soundfile format*
URL: <<https://ccrma.stanford.edu/courses/422/projects/WaveFormat/>> [2010-18-5].

Příloha A

UML diagram



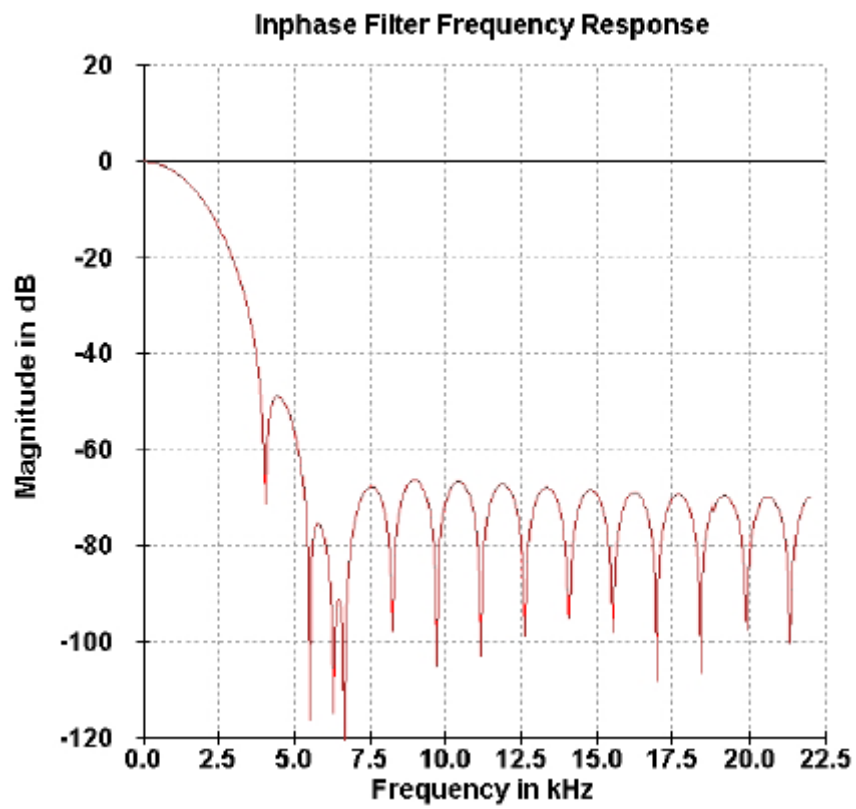
UML diagram – přerušované čáry představují náhradu za stávající třídu, spojenou plnou čarou. Týká se to metod pro hledání salv a výpočtu jejich útlumu energie. Pro rozsáhlost jednotlivých tříd (atributy, proměnné) je zvolena jednoduchá podoba diagramu, který byl vygenerován pomocí programu *Altova UModel*.

Příloha B

1) Příklad výpočtu pro $M + 1 = 31$, $f_{vz} = 44100$ Hz, $f_c = 1500$ Hz

k	n	h(n)	H_w
0	-15	-0,001360	0,080000
1	-14	0,003389	0,090052
2	-13	0,008701	0,119769
3	-12	0,014471	0,167852
4	-11	0,020571	0,232200
5	-10	0,026861	0,310000
6	-9	0,033192	0,397852
7	-8	0,039405	0,491917
8	-7	0,045346	0,588083
9	-6	0,050859	0,682148
10	-5	0,055801	0,770000
11	-4	0,060039	0,847800
12	-3	0,063461	0,912148
13	-2	0,065975	0,960231
14	-1	0,067511	0,989948
15	0	0,068027	1,000000
16	1	0,067511	0,989948
17	2	0,065975	0,960231
18	3	0,063461	0,912148
19	4	0,060039	0,847800
20	5	0,055801	0,770000
21	6	0,050859	0,682148
22	7	0,045346	0,588083
23	8	0,039405	0,491917
24	9	0,033192	0,397852
25	10	0,026861	0,310000
26	11	0,020571	0,232200
27	12	0,014471	0,167852
28	13	0,008701	0,119769
29	14	0,003389	0,090052
30	15	-0,001360	0,080000

Tab. 1 – Výpočet dat v Microsoft Excel.



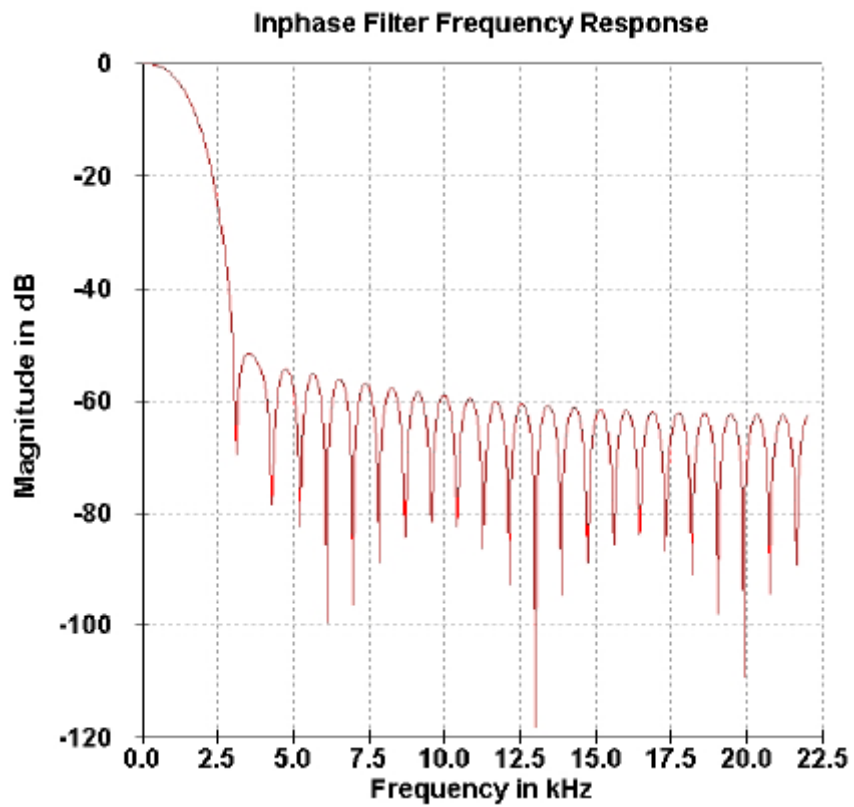
Obr. 1 – Frekvenční charakteristika filtru s použitím Hammingova okna $H_w M + 1 = 31$.

2) Příklad výpočtu pro $M + 1 = 51$, $f_{vz} = 44100$ Hz, $f_c = 1500$ Hz

k	n	h(n)	H_w
0	-25	-0,010285	0,080000
1	-24	-0,012128	0,083627
2	-23	-0,013555	0,094452
3	-22	-0,014468	0,112303
4	-21	-0,014778	0,136899
5	-20	-0,014412	0,167852
6	-19	-0,013318	0,204674
7	-18	-0,011463	0,246785
8	-17	-0,008838	0,293520
9	-16	-0,005456	0,344142
10	-15	-0,001360	0,397852
11	-14	0,003389	0,453805
12	-13	0,008701	0,511116
13	-12	0,014471	0,568884
14	-11	0,020571	0,626195
15	-10	0,026861	0,682148
16	-9	0,033192	0,735858
17	-8	0,039405	0,786480
18	-7	0,045346	0,833215
19	-6	0,050859	0,875326
20	-5	0,055801	0,912148
21	-4	0,060039	0,943101
22	-3	0,063461	0,967697
23	-2	0,065975	0,985548
24	-1	0,067511	0,996373
25	0	0,068027	1,000000
26	1	0,067511	0,996373
27	2	0,065975	0,985548
28	3	0,063461	0,967697
29	4	0,060039	0,943101
30	5	0,055801	0,912148
31	6	0,050859	0,875326
32	7	0,045346	0,833215
33	8	0,039405	0,786480
34	9	0,033192	0,735858
35	10	0,026861	0,682148
36	11	0,020571	0,626195
37	12	0,014471	0,568884
38	13	0,008701	0,511116
39	14	0,003389	0,453805
40	15	-0,001360	0,397852
41	16	-0,005456	0,344142
42	17	-0,008838	0,293520
43	18	-0,011463	0,246785
44	19	-0,013318	0,204674
45	20	-0,014412	0,167852
46	21	-0,014778	0,136899
47	22	-0,014468	0,112303
48	23	-0,013555	0,094452

49	24	-0,012128	0,083627
50	25	-0,010285	0,080000

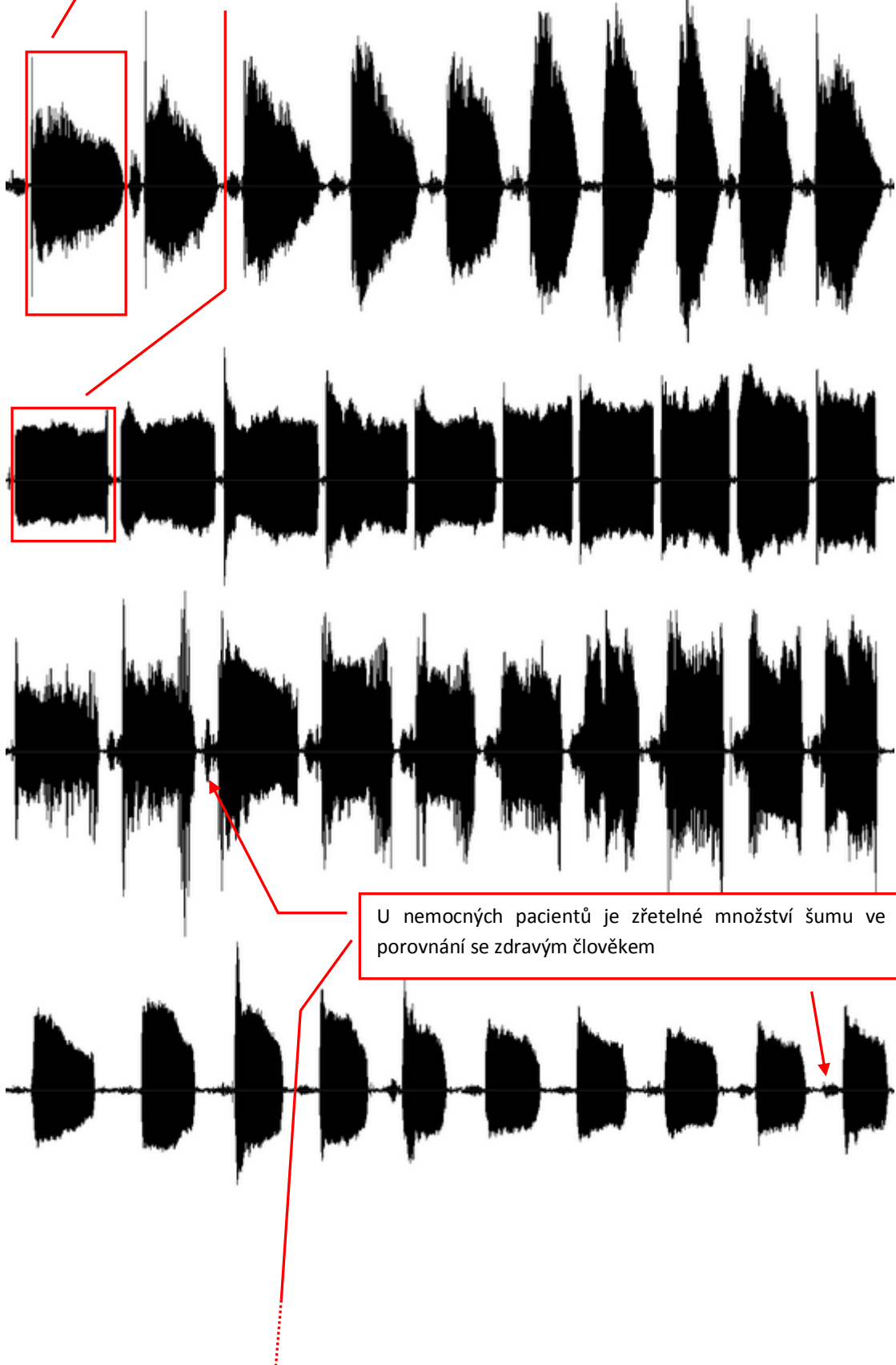
Tab. 2 – Výpočet dat v Microsoft Excel.



Obr. 2 – Frekvenční charakteristika filtru s použitím Hammingova okna $H_w M + 1 = 51$.

Příloha C

Srovnání salv a různě se projevující onemocnění na pacientech.



U nemocných pacientů je zřetelné množství šumu ve fonaci v porovnání se zdravým člověkem



Obr. 1 – Podoba průběhu testovaných souborů (v pořadí 1-6, poslední je zdravý člověk).