



University of West Bohemia in Pilsen
Faculty of Applied Sciences
Department of Computer Science and Engineering
Univerzitní 8
306 14 Pilsen
Czech Republic

Detekce a hodnocení videozáznamu pohybu hlasivek z vysokorychlostní kamery

Odborná práce ke státní doktorské zkoušce

Tomáš Ettlér

Technical Report No. DCSE/TR-2017-02
June, 2017

Distribution: public

Abstract

This work summarizes problematic about image segmentation used on laryngoscopic data taken from High Speed Video. There are described most used published method of automatic segmentation and also a new method is introduced. Descriptions of preprocessing and postprocessing methods are also included.

Part of this work describes assessment of input data quality and there are also presented output parameters including ones based on symmetry using the axis, finding of the axis of glottal symmetry is described as well. As an example of parameters of symmetry, location of the center of gravity of glottal area and the border is presented including their progress in time. Several casuistries demonstrate their usefulness for determining the glottal symmetry.

Abstrakt

Tato práce shrnuje problematiku segmentace obrazu, která se používá nad laryngoskopickými daty získané vysokorychlostní kamerou. V práci jsou uvedené nejčastěji používané a publikované metody automatické segmentace obrazu a je také uvedena vlastní vyvinutá metoda založená na shlukové analýze. Součástí je i popis několika metod předzpracování a následného zpracování obrazu.

Dále je popsána problematika kvality snímků, hledání osy symetrie a jsou uvedeny parametry, které lze po detekci hlasivkové štěrbině vypočítat, včetně těch vypočítajících o symetrii hlasivky. Příkladem jsou parametry umístění těžiště plochy a hranice hlasivkové štěrbině a jejich vývoj v čase. Na několika kazuistikách je pak znázorněna užitečnost těchto parametrů pro správné určení symetrie hlasivky.

Poděkování

Velké poděkování za trpělivost, pomoc při hledání informací a za celkové vedení této práce patří mému školiteli Ing. Pavlu Novému, Ph.D.

Dále děkuji Ing. Jiřímu Peštovi, CSc. a MUDr. Monice Vohlídkové za odborné konzultace, poskytnutí expertních znalostí a vytváření datového korpusu.

Obsah

1	Úvod	1
1.1	Anatomie hrtanu a fyziologie hlasu	1
1.2	Vyšetřovací metody	7
1.2.1	Akustické metody	8
1.2.2	Aerodynamické metody	9
1.2.3	Elektrofyziologické metody	10
1.2.4	Optické metody	10
1.3	Metody používané na ORL klinice FN Plzeň	12
2	Vysokorychlostní laryngoskopie	14
2.1	HSV - vysokorychlostní video	14
2.2	Kvalita záznamu a detekce	16
2.2.1	Rozostřený obraz	17
2.2.2	Přesvícení	17
2.2.3	Nedostatečné osvětlení	17
2.2.4	Překrytí	17
2.2.5	Pohyb kamery	18
2.2.6	Šum	18
2.2.7	Štěrbina mimo obraz	18
2.2.8	Přítomnost tekutin	18
2.2.9	Chybějící pohyb	18
2.2.10	Příliš malá hlasivka na snímku	19
3	Metody používané pro detekci hlasivkové štěrbin v obraze	20
3.1	Metody a interakce s člověkem	20
3.1.1	Poloautomatické a ruční	20
3.1.2	Automatické	21
3.2	Metody využívající histogram	21
3.2.1	Thresholding	21
3.3	Metody založené na detekci hran	23
3.3.1	Edge detection	23
3.3.2	Gabor filtering	26
3.3.3	Active contours	27
3.4	Metody využívající topologii obrazu	27
3.4.1	Watershed segmentation	27
3.4.2	Region growing	29
3.5	Předzpracování	29

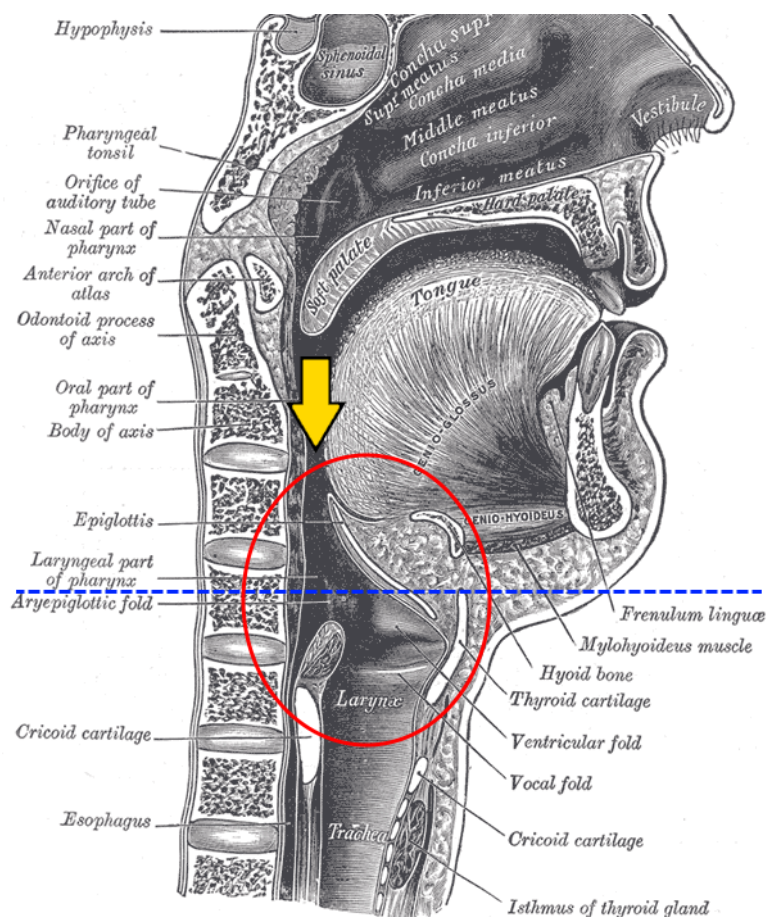
3.5.1	Oblast zájmu	29
3.5.2	Wiener motion estimation	30
3.5.3	Odstranění šumu	30
3.5.4	Vytvoření kymogramu	31
3.6	Postprocessing	32
3.6.1	Region merging	32
3.6.2	Hough transformation	32
3.6.3	Phonovibrogram	32
3.7	Shrnutí	33
4	Vlastní metoda pro detekci hlasivek	34
4.1	Nalezení oblasti zájmu	34
4.2	Segmentace shlukovou analýzou	35
4.2.1	Výsledky testování	39
5	Detekce osy symetrie hlasivkové štěrbiny	41
5.1	Plovoucí osa	44
5.2	Hodnocení výsledků detekce osy symetrie	45
6	Parametry hlasivkové štěrbiny	48
6.1	Základní parametry a symetrie	48
6.2	Parametry doplněné	48
6.3	Těžiště hlasivkové štěrbiny	49
6.3.1	Parametry těžiště	49
6.3.2	Vybrané kazuistiky	52
6.3.3	Shrnutí získaných poznatků	73
6.4	Normování	74
6.5	Parametry přes více snímků	74
6.6	Korelace parametrů s jinými způsoby vyšetření	74
7	Teze dizertační práce	75
7.1	Vylepšení hledání ROI	75
7.2	Registrace obrazu	75
7.3	Validace výsledků	75
7.4	Adaptivní parametry shlukové analýzy	76
7.5	Vyhodnocení dat	76
8	Závěr	79

1 Úvod

Mluvení je pro člověka elementární způsob komunikace. Hlas je formován v ústní dutině, ale tvořen je kmitajícími hlasivkami v hrtanu. Proto jsou hlasivky důležitým orgánem pro komunikaci a při jejich onemocnění nebo dysfunkci je ztížen život člověka a jeho interakce ve společnosti.

1.1 Anatomie hrtanu a fyziologie hlasu

Hrtan (lat. *larynx*) se nachází na začátku dolních cest dýchacích a vykonává řadu funkcí. Kromě dýchání a účasti na polykání je to funkce fonační. Skládá se ze tří nepárových chrupavek (štítná a prstencová chrupavka a příklopka hrtanová), tří párových (hlasivkové a sesamkové chrupavky), svaly a vazy. Řez je na obr. 1.1.



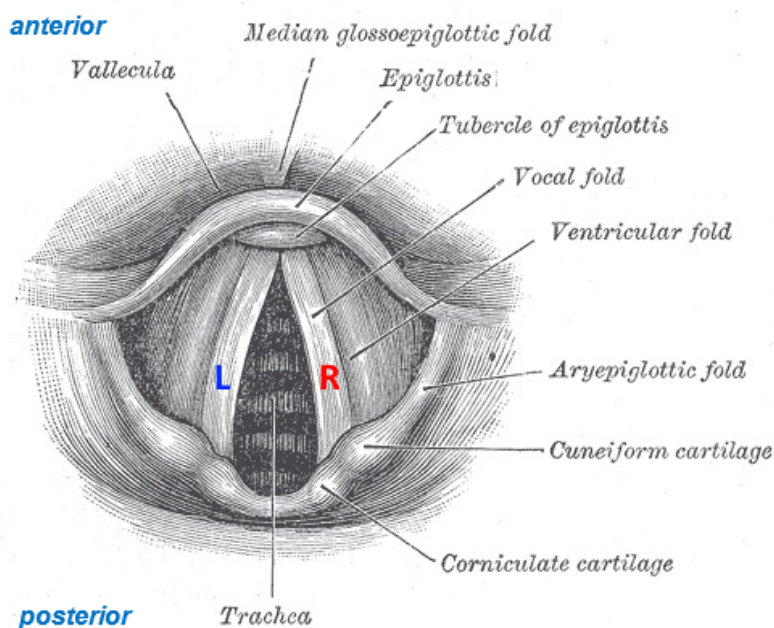
Obrázek 1.1: Anatomie dýchacích cest, sagitální řez ústní dutinou, nosem a hrtanem. Obrázek řezu je doplněn o popis orientace anterior/posterior, transversální rovinu anatomického popisu hlasivek a schématické znázornění směru laryngoskopického pohledu na hlasivky (z [1], fig. 994 - Sagittal section of nose, mouth, pharynx and larynx).

Mezi chrupavkami jsou napjaté dva hlasové vazy, které tvoří hlasivky (viz obr. 1.2). Hlasivky tvoří párový hlasivkový sval, hlasivkové vazy a slizniční hlasivkové řasy (lat. *plicae vocales*), mezi kterými se nachází hlasivková štěrbina (lat. *rima glottidis*) [2], která má u zdravého člověka tvar protáhlého rovnoramenného trojúhelníka. Hlasivky jsou pozorovatelné ústní dutinou za hrtanovou příklopkou. Toho je využíváno laryngoskopickými vyšetřovacími metodami, viz obr. 1.3.

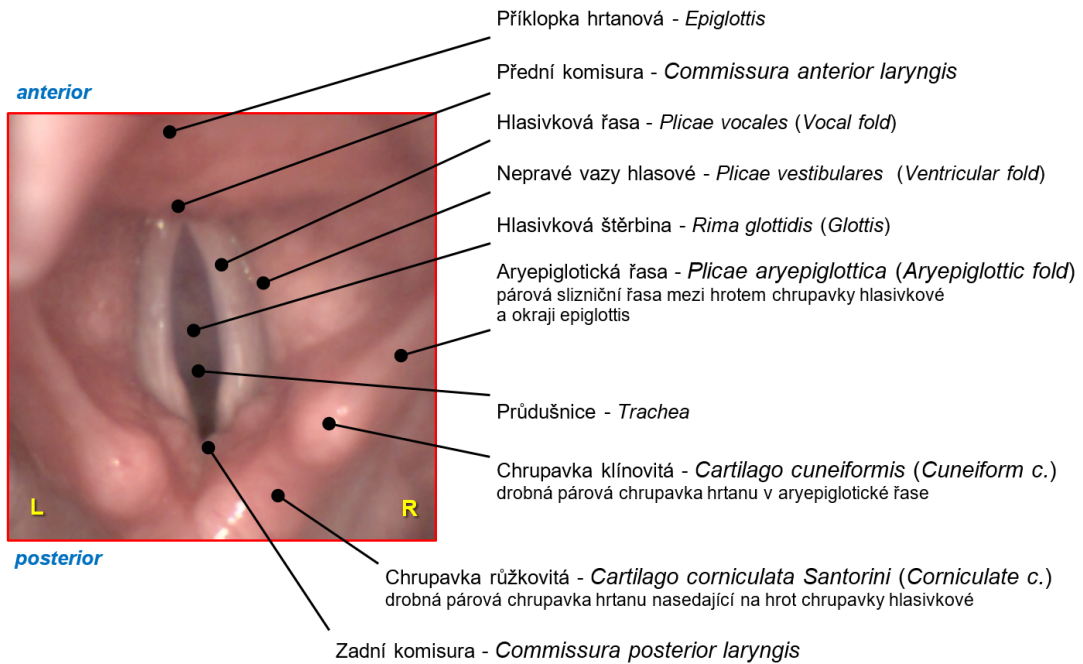
Protože se práce zabývá zpracováním laryngoskopického obrazu hlasivek, je další popis zaměřen právě na anatomické struktury, které ovlivňují a podmiňují výsledky zpracování snímků z vysokorychlostní laryngoskopie. Jedná se hlavně o hlasivkovou štěrbinu (*rima glottidis*).

Na obrázku 1.4 je pak znázorněn frontální řez hrtanem s vyznačením hlasivkových řas. V horní části jsou nepravé hlasivkové řasy (které většinou nejsou příliš výrazné a neovlivňují laryngoskopický pohled na glottis) a pod nimi samotné hlasivkové řasy.

Hlavní funkcí hlasivek je tvorba hlasu, resp. hlasivkového tónu. Principem fonace je vznik přerušovaného proudu vzduchu pomocí hlasivek. Hlas je pak dotvořen rezonančními dutinami (supraglottický prostor, ústní a nosní dutina...) a vzniká jedinečná barva hlasu člověka [2] [3] [4].

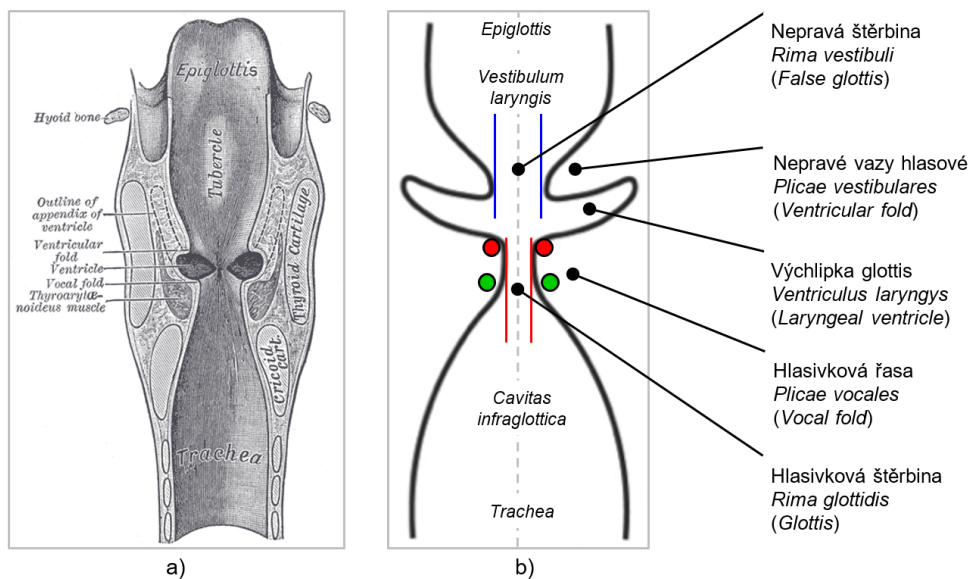


Obrázek 1.2: Laryngoskopický pohled na hlasivky, anatomický popis hlasivek. Popis je veden v transverzální rovině hrtanu, obrázek je doplněn o popis orientace anterior/posterior a L/R (z [1], Fig. 956 – Laryngoscopic view of interior of larynx).



Obrázek 1.3: Laryngoskopický obraz hlasivek.

Snímek z vysokorychlostní laryngoskopie (HSV) je doplněn o popis některých, pro účely této práce významných, anatomických útvarů. (ilustrativní snímek hlasivek je z DB-HSV ORL FN Plzeň)



Obrázek 1.4: Anatomie hrtanu, frontální řez, schéma hrtanu.

a) Anatomický popis hrtanu ve frontálním řezu (z [1], Fig. 954 - Coronal section of larynx and upper part of trachea).

b) Schematický popis hrtanu ve frontálním řezu s označení horní a dolní části hlasivkové řasy pro zobrazení tzv. slizniční vlny.

Červená značka - horní hrana *Plicae vocales*; zelená značka - spodní hrana *Plicae vocales*

Pro způsob tvorby hlasu byly vytvořeny dvě teorie:

Myoelestická teorie (neboli neurodynamická teorie) byla definována v roce 1958 Van Den Bergem. Podle této teorie vzniká kmitavý pohyb souhrou hmoty hlasivek, jejich napětím, elasticitou a poměry, které vznikají při výdechu z plic. Aerodynamicko - aerostatická síla hlasivky od sebe vzdaluje a napětí svalů a vazů vrací hlasivkové řasy k sobě¹.

Podle jiné, neurochronaxické teorie, kterou popsal francouzský fyzik Husson v roce 1953, hlasivky nevykonávají pohyb pasivně, ale pomocí nervových impulzů. Odpůrci však namítají, že žádný nerv není schopen přenést tolik impulsů, kolik je jich potřeba např. při zpěvu, kdy hlasivky kmitají více než 1000x za vteřinu. Dalším argumentem bylo kmitání hlasivky i při jednostranné obrně, kdy by podle této teorie kmitat neměla. Proto je dnes uznávaná myoelastická teorie.

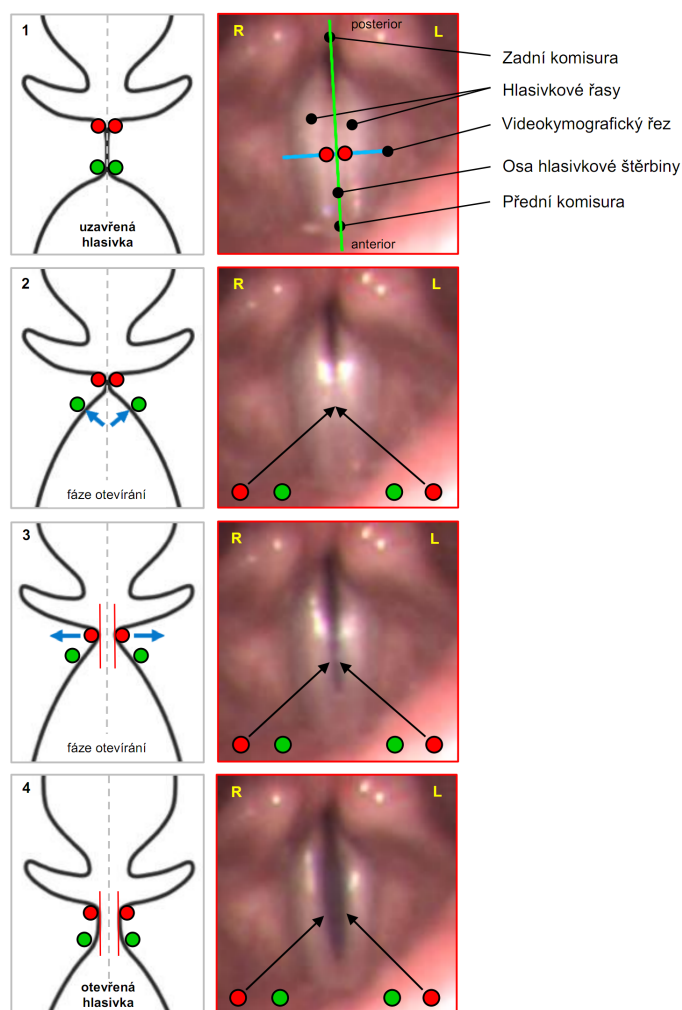
Na následujících obrázcích je znázorněný pohyb hlasivek při fonaci rozdělený do sedmi fází během otevírání hlasivek (abdukce, obr. 1.5) a uzavírání hlasivek (adukce, obr. 1.6), které zobrazují stav hlasivek ve schematickém frontálním řezu a v laryngoskopickém pohledu v jednotlivých fázích pohybu. Obr. 1.7 pak zobrazuje pohyb na kymogramu se srovnáním s průběhem akustického tlaku tvořícím zvuk a signálem z elektroglottografu.

Červené a zelené značky na obrázcích 1.4, 1.5, 1.6 a 1.7 označují horní a dolní část hlasivkové řasy, které ve fázi uzavírání hlasivek vytváří tzv. slizniční vlnu.

Podstatou onemocnění hlasivek je omezení pohybu jedné nebo obou hlasivkových řas z důvodu zmenšení míry elasticity nebo porušení inervace svalů napínající hlasivku [5].

Pro prevenci rozvinutí závažných onemocnění hlasivek resp. hrtanu, je používána řada vyšetřovacích metod.

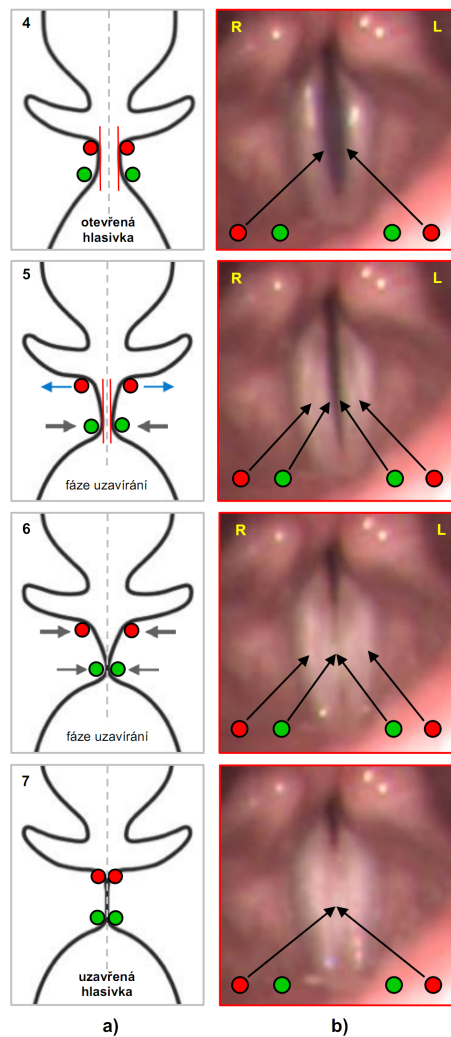
¹Toto fyzikální zdůvodnění myoelastické teorie podal Lieberman v roce 1968



Obrázek 1.5: Popis a rozklad dynamiky hlasivek v průběhu fonační funkce - otevírání (abdukce).

a) schéma vývoje polohy levé a pravé hlasivky během fonace vokálu „a“ v řezu ve frontální rovině;

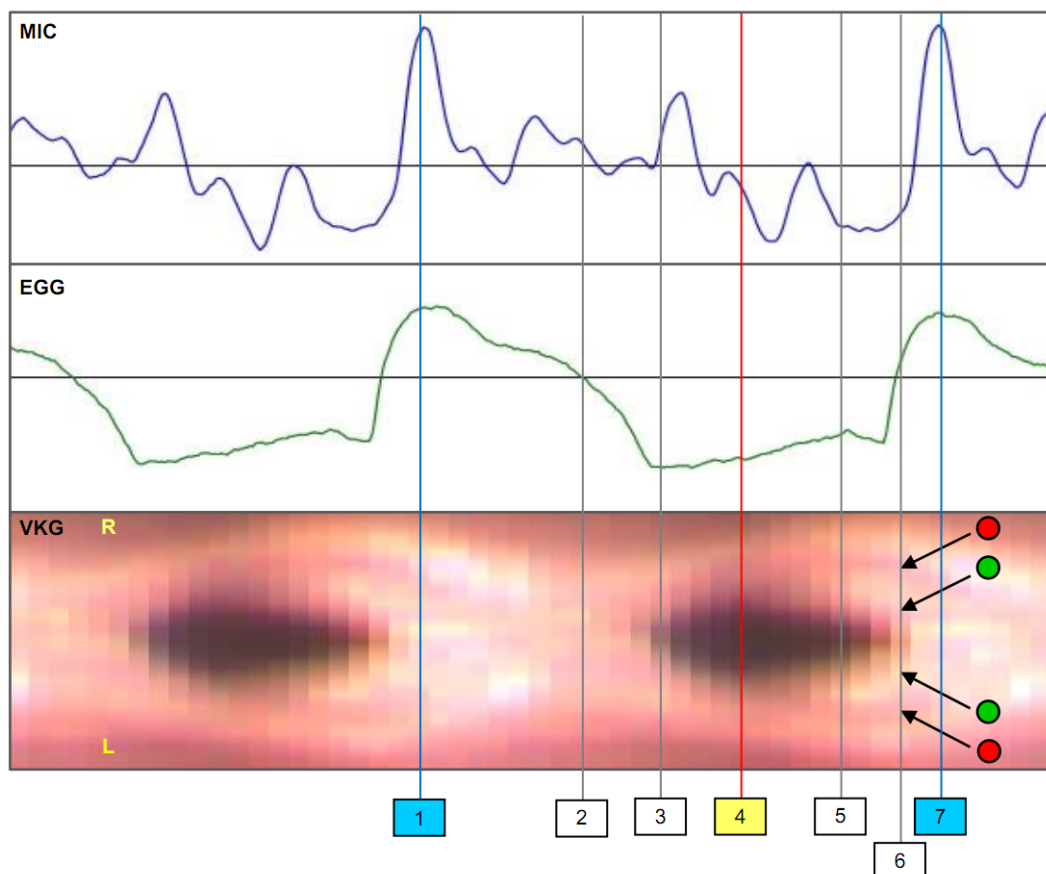
b) reálné snímky vývoje glottis v jednotlivých fázích otevírání hlasivek v transverzální rovině (vysokorychlostní laryngoskopie – HSV).



Obrázek 1.6: *Popis a rozklad dynamiky hlasivek v průběhu fonační funkce - uzavírání (addukce).*

a) schéma vývoje polohy levé a pravé hlasivky během fonace vokálu „a“ v řezu ve frontální rovině;

b) reálné snímky vývoje glottis v jednotlivých fázích uzavírání hlasivek v transverzální rovině (vysokorychlostní laryngoskopie – HSV).



Obrázek 1.7: Rozložení fází otevírání a uzavírání hlasivek ve videokymografickém řezu VKG (znázorněný na obr. 1.5 pozice-snímek 1b);
 MIC - záznam akustického tlaku SPL mikrofonem během fonace vokálu „i“;
 EGG - záznam elektroglottografického signálu;
 VKG - videokymografický záznam v řezu, viz pozice-snímek 1b.

1.2 Vyšetřovací metody

Zpěvák M. Garcia v pol. 19. st. jako první použil metodu nepřímé laryngoskopie [6], kdy zkoumal hlasivky pomocí zrcátka. Mnoho lékařů, a vědců v tomto oboru pokračovalo a vznik obor laryngologie, který se zabývá poruchami hlasivek. Pro celkovou problematiku tvorby hlasu existuje obor foniatrie, zabývající se fyziologií, patofyziologií, diagnostikou i léčbou.

Postupem času vznikaly různé metody vyšetření hlasivek, které lze dělit do čtyř kategorií - metody akustické, aerodynamické, elektrofyziologické a optické.

1.2.1 Akustické metody

První skupinou metod jsou akustické, které využívají záznamu a analýzy zvuku. Jedná se především o tyto metody:

Hlasové pole

P. H. Damsté (1970) zavedl tuto metodu pod názvem fonetografie, ale později byla kvůli nevýstižnosti názvu přejmenována na vyšetření hlasového pole (VRP - Voice Range Profile). Hlasové pole je vyšetřovací metoda, která zkoumá kvantitativní parametry hlasu, a tím získá obraz frekvenčního a dynamického rozsahu projevu.

Metoda spočívá v zaznamenání nejnižšího a nejhlasitějšího hlasu pacienta v celém jeho frekvenčním rozsahu a zanesení hodnot do grafu, kde na vodorovné ose je frekvence a na svislé ose intenzita, která se měří v dB.

Tuto metodu lze použít i na hodnocení zátěžového testu, kdy je porovnáván záznam pořízený před zátěží se záznamem hlasu po zátěži (dlouhé čtení, divadelní představení, zpěv). Únava hlasu pak způsobuje frekvenční nebo dynamické zúžení hlasového pole.

Multidimenzionální analýza

Multidimenzionální analýza (MDVP - Multi-Dimensional Voice Program) je metoda založena na sledování časových a signálových akustických parametrů. Zkoumá objektivní parametry hrtanového tónu a tím základní parametry hlasu pomocí frekvenční analýzy. Mezi hlavní sledované parametry patří např.:

- Index měkké fonace - SPI (soft phonation index) je poměr energie harmonických frekvencí v rozsahu 70-1600Hz k energii v rozsahu 1600 - 4500Hz.
- Jitter - parametr popisující stupeň poruchy periodicity základní frekvence.
- Shimmer - parametr poruch amplitudové stability [7].

Pro tyto parametry jsou rozlišovány dva typy hodnot:

- absolutní, které jsou vypočítávány jako průměr odchylek period a amplitud následujících cyklů hlasivek,
- relativní, které jsou vztaženy k průměrným hodnotám všech měřených cyklů [8].

Jedním z komerčně dostupných analyzátorů je program MDVP od firmy Pentax (dříve Kay Elemetrics), který počítá až 33 parametrů [9]. Tento program se stal

standardem pro hodnocení akustických parametrů hlasu, protože jeho výsledky lze porovnávat i s naměřenými parametry zdravé populace.

Hodnocení kvality závěru hlasivek

Hodnocení kvality závěru hlasivek (SCORE [10]) je založena na Fourierově analýze jedné periody kmitání hlasivek s cílem detekce přechodového děje jako odezvy na tlakový puls vyvolaný závěrem glottis. Normované amplitudy vypočtené z koeficientů Fourierovy řady pak podle znaleckého hodnocení, které je dáno výsledky vyšetřovacích metod i mnohaletou zkušeností, určují hodnotu SCORE na jednoduché stupnici. Hodnotu Score lze využít k hodnocení funkce hlasivek a případnému stanovení diagnózy hlasové poruchy.

Tato metoda byla vyvinuta na ORL klinice Fakultní nemocnice Plzeň společně s Katedrou informatiky a výpočetní techniky na Fakultě aplikovaných věd ZČU v Plzni.

1.2.2 Aerodynamické metody

Aerodynamické metody jsou založené na sledování dýchání a měření rychlosti a množství vzduchu procházejícího hlasivkovou štěrbinou.

Pneumografie

Pneumografie byla dříve používaná metoda, která se zabývala zkoumáním pohybu hrudníku a břicha při dýchání a fonaci. Tuto metodu poprvé použil Gutzmann ve 20. letech 20. století pomocí tenkostěnné gumové hadičky, která přenášela změnu tlaku na kymograf². Tato metoda byla dále zdokonalována a bylo možné sledovat daleko jemnější pohyby.

Výsledkem vyšetření byly dvě křivky - hrudní a břišní. Podle nich se dále hodnotila frekvence dechu v klidu, poměr vdechu k výdechu, synchronizace mezi křivkami, chování během fonace a další parametry.

Pneumotachografie

Pneumotachografie nahradila v 60. letech 20. století pneumografii. Je založena na měření rychlosti vzduchu a jeho množství procházejícího hlasivkami při fonaci. Součástí této metody bylo i měření subglottického tlaku (tlaku vzduchu pod hlasivkami), ale této hodnoty nebylo možné docílit neinvazivními metodami, proto se od toho upustilo. Jinou možností bylo zavedení katetru a měřícího balónku do oblasti pod hlasivkami, to ale narušovalo přirozenou tvorbu hlasu.

²Kymograf - přístroj zaznamenávající pohyb

Dalším způsobem měření subglottického tlaku je pomocí hadičky umístěné na jazyku. Tuto metodu používá přístroj Phonatory Aerodynamic System od společnosti PENTAX. Měření tlaku v tomto případě probíhá nepřímo při fonaci slabiky „pá“, protože při fonaci konsonantu „p“ je vzduchový sloupec od plic až po rty nepřerušen, a tak je tlak v ústech totožný s tlakem v oblasti pod hlasivkami [2]. Tento způsob se používá i na ORL oddělení FN Plzeň.

Základními parametry jsou fonační objem, maximální fonační doba, průměrná rychlost vzduchu a fonační kvocient, k jehož výpočtu je třeba znát vitální kapacitu plic. Tato metoda však poslouží k posouzení funkce hrtanu a stupni poškození, nikoliv však pro diagnózu onemocnění, jelikož nedokáže nalézt příčinu změny jednotlivých parametrů.

1.2.3 Elektrofyziologické metody

Další skupinou vyšetřovacích metod jsou metody elektrofyziologické, které pro vyšetření hlasivek využívají změny elektrických vlastností při fonaci:

Elektroglottografie

Tuto metodu, která se zabývá sledováním změn elektrického odporu (vodivosti) vysokofrekvenčním měřením při kmitání hlasivek, zavedl v r. 1957 Fabre. Snímací elektrody jsou umístěny nad křídly štítné chrupavky a je vymezeno 6 bodů v glottogramu odpovídajícím jednotlivým fázím pohybu hlasivek a tvaru hlasivkové štěrbiny.

Elektromyografie

Elektromyografie je metoda, která sleduje aktivitu svalů při fonaci. Poprvé ji navrhl Seiffert v r. 1919 a využívá jehlových elektrod zavedených do svalů kolem hlasivek. Tato metoda přispěla k objasnění aktivity svalů během fonace a dnes se využívá hlavně k ověřování přerušování zvrtného nervu při obrně.

Metoda měří aktivitu jednotlivých svalů a velikost amplitud potenciálů.

1.2.4 Optické metody

Poslední skupinou jsou metody optické.

Již v polovině 19. století byla použita laryngoskopie, kdy pomocí zrcátka je možné pozorovat chování hlasivek. Kvůli vysoké frekvenci kmitání bylo pro zachycení pohybu zapotřebí využít další přístroje. Nejdříve se používala laryngostroboskopie, která využívala blikající světlo s frekvencí o trochu větší, než je kmitání hlasivek,

čímž bylo možné zobrazit jejich zdánlivý pohyb. Z důvodu nemožnosti zachytit reálný pohyb hlasivkových řas nebylo možné najít nahodilé jevy v rámci jednoho kmitu.

Videokymografie již dokázala zaznamenat reálný pohyb hlasivkových řas na přímce. Výstupem je kymogram, což je grafické znázornění kmitů, které vznikne poskládáním jednotlivých obrazových řádků vedle sebe [4].

Laryngoskopie

Laryngoskopie je metoda sledování hrtanu a hlasivek pomocí zrcátka (nepřímá laryngoskopie) nebo optických vláken (přímá laryngoskopie). Zrcátko nebo laryngoskop se vloží do úst a odrazem, resp. světlovody, lze pozorovat hlasivky. Sledovat lze ale pouze stav hlasivek, jejich pohyb není kvůli jejich rychlému kmitání okem rozeznatelný. Tato metoda byla poprvé použita už v r. 1854 zpěvákem Garciou.

Laryngostroboskopie

Laryngostroboskopie umožňuje pozorování kmitů hlasivek pomocí stroboskopického jevu³, kdy blikající světlo má nepatrně odlišnou frekvenci od frekvence kmitání hlasivek, které je kvůli své vysoké frekvenci okem nerozlišitelné. Dochází k optickému zpomalení kmitání hlasivek a je tedy možné pozorovat jejich chování. Pro laryngostroboskopické vyšetření je potřeba ustálené kmitání, jinak nedochází k efektu zpomaleného pohybu.

Videokymografie

Videokymografie je metoda založená na matematicko-optickém modelu v příčném řezu hlasivek. Umožňuje je vyšetřit v jakékoliv pozici od přední po zadní komisuru. Přesně graficky zobrazuje frekvenci i amplitudu kmitů hlasivek a lze pozorovat chování při otevírání a uzavírání. Pro záznam je zapotřebí vybavení schopné snímat jeden řádek obrazu vysokou rychlostí (v řádu tisíců řádků za vteřinu) a softwarové vybavení ke zpracování nasnímaných dat a uložení výsledného snímku. Tato metoda byla poprvé popsána a vyzkoušena v [4].

Výsledkem vyšetření je snímek, který je poskládaný z jednotlivých řádků a znázorňuje pohyb hlasivkových řas ve sledovaném místě.

³Stroboskopie je jev založený na setrvačnosti vnímání lidského oka, kdy při určité frekvenci blikajícího světla, která je blízká frekvenci periodicky rychle pohybujícího se objektu, se pohyb objektu zdánlivě zpomalí.

Vysokofrekvenční snímání

Vysokofrekvenční snímání (HSV - High Speed Video) je metoda ke sledování chování hlasivek vysokofrekvenční kamerou, která dokáže zaznamenat více než 1000 snímků za vteřinu. Zaznamenán je na rozdíl od videokymografie celý obraz hlasivek. Tato metoda se objevuje až s vývojem technologie snímání obrazu a nahrazuje předchozí metody.

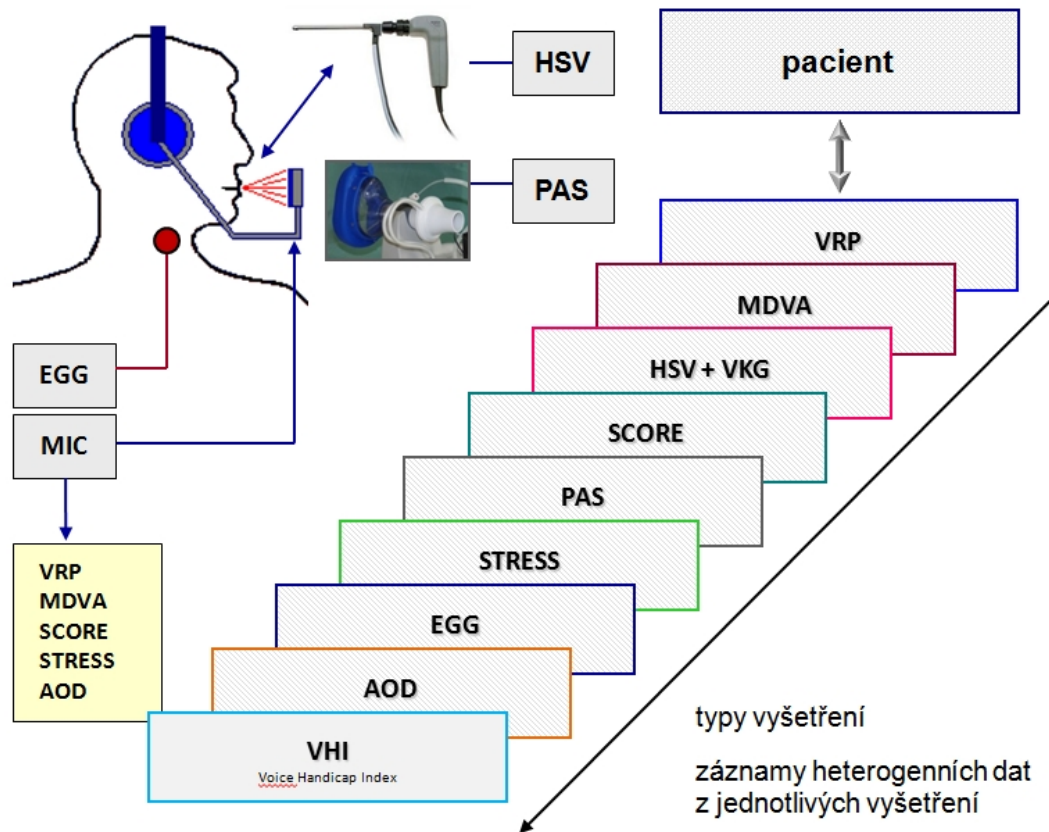
Pro vyšetření a diagnostiku poruchy hlasu se v dnešní době používá nejčastěji kombinace akustických a optických metod, jelikož jsou nejméně náročné na pacienta a dokáží získat velké množství relevantních parametrů. Tato práce se zaměřuje na zpracování a vyhodnocení snímků ze záznamu vysokorychlostní kamery.

1.3 Metody používané na ORL klinice FN Plzeň

Na ORL klinice FN Plzeň se standardně provádí subjektivní hodnocení VHI (Voice Handicap Index) pomocí dotazníku a vyšetření akustickými metodami (hlasové pole, multidimenzionální analýza, hodnocení SCORE), které se doplňují o vyšetření vysokorychlostní kamerou.

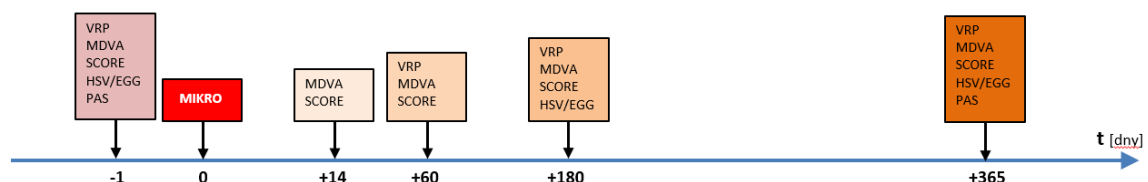
Ve vybraných případech se provádí i další typy vyšetření, jako např. elektroglotografie, zátěžové hlasové testy (STRESS test) [11], aerodynamická vyšetření hlasivek [12] či detekce nestandardního kmitu hlasivek (AOD) [13], [14].

Snahou je vytvořit reprezentativní soubor parametrů nad daty z různých vyšetření, která jsou rozložena v čase v různém stádiu onemocnění či léčby (obr. 1.8). Tato data jsou následně analyzována a jejichž výstupem je ohodnocení hlasivek danými souhrnnými parametry. Pokud léčba choroby vyžaduje mikrochirurgický zákrok, jsou vyšetření rozložena v čase podle schématu na obrázku 1.9. Léčba některých onemocnění však vyžaduje dlouhodobý přístup formou cvičení, pak je sběr dat individuální.



Obrázek 1.8: Přehledové schéma vyšetřovacích metod používaných na ORL klinice FN Plzeň.

- VRP (*Voice Range Profile*) - hlasové pole
- MDVA (*Multi-Dimensional Voice Analysis*) - multidimenzionální analýza hlasu
- HSV (*High Speed Video*) - vysokorychlostní laryngoskopie, vysokorychlostní kamera
- VKG (*Videokymography*) - videokymografie
- SCORE (*Quality of the Glottis Closure*) - analýza kvality závěru glottis
- PAS (*Phonatory Aerodynamic System*) - systém pro měření aerodynamických parametrů
- STRESS (*Stress Test*) - zátěžový test hlasového orgánu
- EGG (*Electroglottography System*) - elektroglottografie
- AOD - analýza nestandardního kmitu ve zvukovém záznamu hlasu
- VHI (*Voice Handicap Index*) - index hlasového postižení



Obrázek 1.9: Příklad uspořádání vyšetření pro mikrochirurgický zákrok.

2 Vysokorychlostní laryngoskopie

Moderní metodou je vysokorychlostní snímání (HSV), které zajišťuje kamera schopná snímat obraz frekvencí až 4000 snímků za vteřinu. Tato práce se zabývá daty získaných právě touto zmiňovanou metodou.

2.1 HSV - vysokorychlostní video

Tato vyšetřovací technika patří mezi optické laryngoskopické metody. Pomocí speciální kamery, která se při laryngoskopickém vyšetření vloží pacientovi do úst, je vytvořen videozáznam hlasivek. Vysokorychlostní kamera dokáže zachytit přesný stav hlasivek několikrát během jedné periody kmitání, celkově až 4000x za vteřinu. Z těchto snímků dokážeme složit přesný pohyb hlasivkových řas.

Mezi největší výrobce speciálních vysokorychlostních kamer patří firmy Pentax¹ a Richard Wolf². Ukázka kamery je na obr. 2.1.



Obrázek 2.1: Ukázka vysokorychlostní kamery (*HRES ENDOCAM 5562, Richard Wolf*)

V záznamu z vysokorychlostní kamery je zobrazen celý obraz hlasivek (pohledem z úst) a reálný pohyb hlasivkových řas (na obr. 2.3). Při vyšetření se většinou poři-

¹Pentax je japonská firma specializující se na optické přístroje. Tato společnost koupila americkou firmu Kay Elemterics Corp. zabývající se výrobou medicínského hardwaru a softwaru, <https://www.pentaxmedical.com/pentax/en/99/1/ENT-Speech/>.

²Německá firma Richard Wolf je výrobcem zdravotnické techniky, <http://www.richard-wolf.com>.

zuje záznam o délce několika jednotek vteřin. Z tohoto množství snímků poté stačí vybrat jen některé, které chceme analyzovat - většinou několik period v době ustálené fonace.

Při snímání je zapotřebí scénu přisvítit, proto mají kamery vlastní zdroj světla. Vyšetřující osoba musí při vyšetření dát pozor, aby hlasivka byla v záznamu správně umístěna, osvětlena a zaostřena (více v sekci kvalita vstupních dat). Na analýzu je nutné použít snímek o rozlišení alespoň 128x128 pixelů. Vyšší kvalita snímků umožňuje přesnější výpočet parametrů

Rozlišení snímků u používaných kamer není zatím příliš veliké. Vlastnosti snímajících čipů se v dnešní době stále zlepšují, zde jsme ale omezení nízkou intenzitou osvětlení, velmi krátkou dobou expozice jednotlivých snímků a velikostí snímače. Nejnovější kamery výrobců ale již umožňují snímání v rozlišení až 4K³

Dalším problémem může být pohyb kamery vůči hlasivkám pacienta. I přes velmi vysokou frekvenci snímání může třes ruky vyšetřující osoby způsobit postupnou změnu pozice hlasivkové štěrbiny v obraze v rámci sekvence.

Při vyšetření pacient fonuje vokál „í“, při kterém je maximálně rozšířen nadhlasivkový prostor. Zde je rozdíl oproti jiným vyšetřením, zvláště založených na akustických metodách, kdy se spíše používá hláska „á“, kde ale není prostor nad hlasivkami tolik přístupný. Rozdíl je vidět na obr. 2.2.



Obrázek 2.2: Rozdíl v postavení hrtanu na snímcích z magnetické rezonance. De-douch, Švec, Horáček, Kršek, Havlík & Vokřál (2003).

Ukázka snímků z vysokorychlostní kamery je na obr. 2.3. Z důvodu pozice kamery jsou snímky otočené o 180°, tudíž pravá hlasivková řasa je vlevo a naopak.

³Kamera ENDOCAM Logic 4K snímá obraz v rozlišení až 3840 x 2160 px. <http://4k.richard-wolf.com/en.html>.



Obrázek 2.3: Ukázka jednotlivých snímků sekvence z vysokorychlostní kamery. Data z DB-HSV ORL pracoviště FN Plzeň.

2.2 Kvalita záznamu a detekce

Kvalita záznamu je asi nejdůležitějším kritériem, které ovlivňuje úspěšnost detekce hlasivkové štěrbinu. Snímky dobré kvality (podle rozdělení níže) jsou segmentovány s velkou přesností a nevykazují velkou chybovost.

Pokud jako vstupní data použijeme pouze kvalitní snímky, téměř všechny metody uvedené dále v této práci vykazují uspokojivé výsledky. Při snížené kvalitě si již většina metod správně neporadí. Proto je potřeba použít nebo vyvinout takovou metodu, která dokáže s velkou spolehlivostí detekovat hlasivkovou štěrbinu i v méně kvalitním záznamu.

Podle kvality vstupních dat lze videosekvence zařadit do 3 kategorií:

1. Snímky dobré kvality, kde většina metod dokáže správně detekovat hlasivkovou štěrbinu.
2. Snímky horší kvality, kde detekce štěrbinu selhává u většiny metod.
3. Snímky špatné kvality, kde již není možné hlasivkovou štěrbinu automaticky detekovat a je potřeba vstup uživatele.

U snímků horší kvality je ale pravděpodobnost chyby větší, ať už ve výběru ROI (region of interest, více v sekci 3.5.1) nebo samotné detekce, hlavně na okrajích (např. na obr. 2.4).

2.2.1 Rozostřený obraz

Rozostřený obraz vede k nepřesnému určení hranice hlasivkové štěrbiny a řasy. Nepřesně určená hranice může vést ke zvětšení nebo zmenšení detekované oblasti nebo k více členitému obvodu, což ovlivňuje velikost plochy, délku obvodu, členitost a další návazné parametry.

Některé kamery umožňují obraz zaostřit, jiné mají ohnisko fixní. Záleží tedy na vyšetřující osobě, aby zaručila ostrost videozáznamu nastavením kamery či použitím vhodné vzdálenosti kamery od hlasivek.

Na některých snímcích může být málo ostrá hrana hlasivkové štěrbiny i v zaostřeném obraze. To může být způsobeno fází zavírání hlasivek, kdy spodní část hlasivky se již uzavírá a na hraně štěrbiny je tato spodní část viditelná jako tmavší než horní strana hlasivkové štěrbiny.

2.2.2 Přesvícení

Přesvícení je jev, kdy určitá oblast obrazu dosahuje vysoké či maximální hodnoty jasu. Rozdíly mezi přesvícenou oblastí a jejího okolí jsou pak často větší, než rozdíl mezi oblastí štěrbiny a hlasivkových řas, detekce pak může chybně určit umístění štěrbiny.

Kamera také může obraz celkově ztmavit a snížit tak kontrast v důležité oblasti.

2.2.3 Nedostatečné osvětlení

Opakem přesvícení je nedostatek světla na snímcích. To může být způsobeno špatným nastavením lampy nebo nevhodným použitím kamery. Například osvětlení z větší vzdálenosti už nedokáže dostatečně osvětlit celou oblast, další příčinou může být špatné natočení lampy nebo zastínění.

Výsledkem jsou pak snímky, kdy je štěrbina špatně detekovatelná kvůli malému kontrastu nebo velkému šumu.

2.2.4 Překrytí

Překrytí je jev, kdy části hrtanu překryjí hlasivkovou štěrbinu a znemožní určit celou plochu. To může vést i k chybně detekované ose. Pokud není problémová oblast na snímku viditelná, nelze ji samozřejmě dále analyzovat.

2.2.5 Pohyb kamery

Pohyb celého obrazu je většinou zapříčiněn pohybem kamery z důvodu třesu ruky vyšetřující osoby. I při vysokých rychlostech snímání je výsledkem postupné posouvání obrazu. Použitá metoda musí s tímto problémem počítat nebo je potřeba použít nějakou formu předzpracování, která pohyb eliminuje (např. registrace obrazu, popsáno v kapitole 7).

2.2.6 Šum

Šum vzniká zvýšením citlivosti snímače, kdy jeho jednotlivé buňky jsou příliš citlivé k vlastnostem okolního prostředí (např. na teplotu či mírné změny napětí) a dochází k rušení snímání obrazu.

Detekce štěrbin na snímcích z větším množstvím šumu je často nepřesná, zvláště pak určení hranice štěrbin. Ta je velmi členitá právě kvůli velkým rozdílům jasů mezi sousedícími obrazovými body zapříčiněným digitálním šumem.

2.2.7 Štěrba mimo obraz

V okamžiku, kdy se štěrba dostane celá mimo obraz, není samozřejmě možné dále pokračovat v detekci. Pokud je mimo obraz pouze část štěrbin, problém je podobný jako u překrytí, kdy nedokážeme určit věrohodně všechny parametry včetně osy. Současně ale na okrajích snímku může být nedostatečné osvětlení, které vede k dalším problémům.

2.2.8 Přítomnost tekutin

V případě, že se v oblasti hlasivek vyskytuje větší objem tělních tekutin (hlenů), zvyšuje se riziko špatné detekce. Již při určování ROI, který se hledá podle pohybu, může dojít k chybné detekci z důvodu pohybu tekutin. Tekutiny také mohou vytvářet odlesky či měnit barvu tkání, což může také ovlivnit detekci hlasivkové štěrbin.

2.2.9 Chybějící pohyb

Pokud neprobíhá fonace (ať už z důvodu mlčení pacienta nebo fyziologické nemožnosti produkovat hlas), může být obtížné hlasivkovou štěrbinu detekovat. Videosekvence v tomto případě degraduje na statický obraz, kde tedy nelze použít metodu lokalizace štěrbin pomocí pohybu v rámci sekvence snímků.

V případě, že se hlasivky pohybují, ale nedohází k otevření, je samozřejmě detekce štěrbin nesmyslná.

2.2.10 Příliš malá hlasivka na snímku

Při umístění kamery příliš daleko (bez použití optického přiblížení) od hlasivek je na výsledném snímku samotná štěrbina příliš malá. Stejný problém vzniká při použití kamery s příliš malým rozlišením snímače.

Zde pak každý obrazový bod má příliš velkou váhu a chybná detekce jednoho bodu již vede k nepřesnostem při výpočtu parametrů.

Výpočet parametrů je také příliš „hrubý“ a malé rozdíly při pohybu hlasivek nemusí být vůbec detekované.



Obrázek 2.4: Tento snímek je nekvalitní z několika důvodů. Je zde vidět rozostření, přesvícení, překrytí, přítomnost tekutin (odlesky) a štěrbina je malá na okraji snímku.

3 Metody používané pro detekci hlasivkové štěrby v obraze

Z detekce hlasivek z videosekvencí HSV se od rozšíření používání těchto systémů stalo poměrně frekventované téma, kterým se zabývají univerzity a výzkumná centra po celém světě. Jedná se o speciální případ aplikace metod zpracování obrazu, které se používají i v ostatních oblastech počítačového vidění.

Hlavní úlohou detekce hlasivek je segmentace hlasivkové štěrby. K tomuto úkolu můžeme použít buď jednotlivé snímky nebo sekvenci jako celek a využít společné prvky a rozdíly mezi snímky. Případně se tyto principy kombinují k dosažení nejlepšího výsledku.

Většinu typů segmentace předchází odpovídající sekvence předzpracování (preprocessing) snímků ve videosekvenci a některé metody využívají i následnou úpravu, tzv. postprocessing.

V této kapitole jsou popsány nejčastěji publikované metody pro detekci hlasivkové štěrby v obraze. Následující popisy obsahují základní principy jednotlivých metod¹.

3.1 Metody a interakce s člověkem

3.1.1 Poloautomatické a ruční

Často jsou v současnosti v praxi používané metody, které pro svůj běh potřebují zásah člověka. Využívá se to například pro označení bližší polohy hlasivkové štěrby, určení osy nebo pro verifikaci výsledků. Takové metody můžeme označit jako poloautomatické. Jejich nevýhodou je, že pro každý snímek nebo sekvenci je tuto interakci nutné provést. Tato činnost ještě často vyžaduje odborné znalosti, aby výstupní data byla správná. Proto jsou tyto metody nevhodné pro hromadné zpracování souborů a jejich výsledky mohou být nekonzistentní.

Těmito metodami se dále nebudeme explicitně zabývat, protože nastupují až v případech, kdy selhávají automatické metody.

¹Mnoho metod je popsáno v anglickém jazyce a některé nemají český ekvivalent nebo by byl zavádějící. Proto jsou v této práci používané anglické názvy metod, u kterých není jednoznačně ustálen český výraz.

3.1.2 Automatické

Automatické metody nevyžadují žádný vstup uživatele. Metody se ze snímku nebo sekvence snaží určit pozici a tvar štěrbinu a získat všechny parametry. Tyto metody se musí vypořádat s různorodostí vstupních snímků.

U těchto metod bychom ale dále měli vyžadovat verifikaci výsledků a detekci chyb (rozpoznat, že výsledek není takový, jaký požadujeme) a označit výsledek jako chybný. Tuto vlastnost jsem zatím u používaných metod nenašel, rád bych se tomuto tématu v budoucnu věnoval (viz kapitola 7).

3.2 Metody využívající histogram

Řada metod vychází z analýzy histogramu snímku a jeho vlastností. Ten lze využít pro detekci objektu, který je možné na snímku rozlišit hodnotou jasu nebo alespoň jedné z barevných složek od pozadí objektu. V ideálním případě získáme bimodální histogram - takový, který obsahuje 2 lokální maxima a mezi nimi lokální minimum, které je nulové nebo výrazně nižší, než zmíněná maxima.

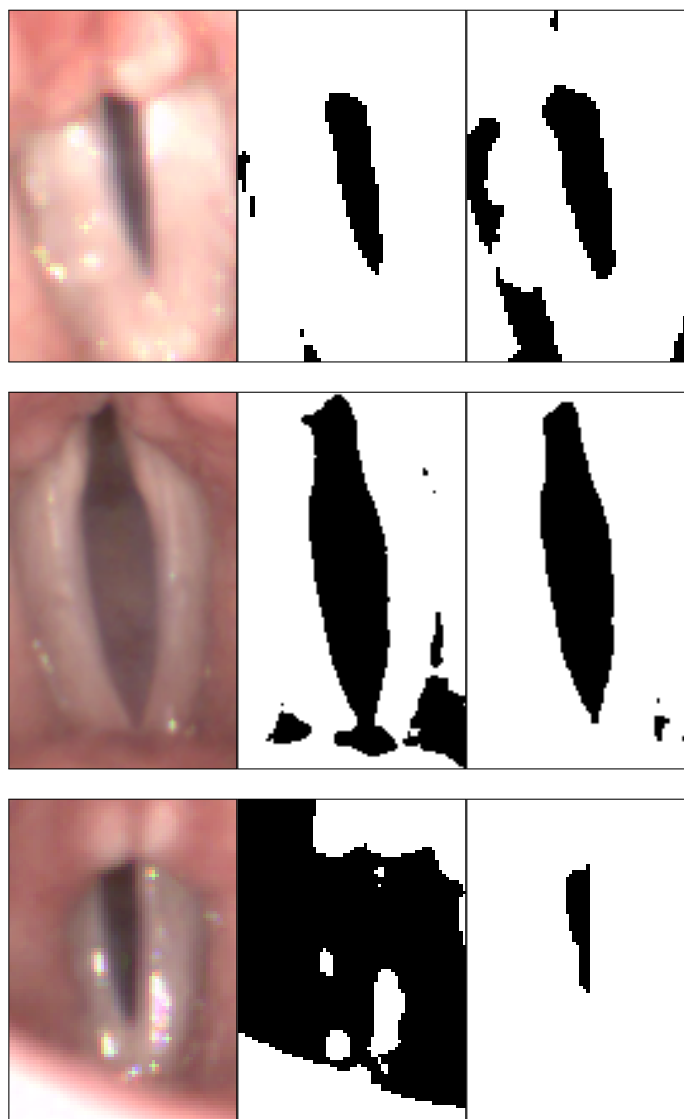
Určení hranice, která odděluje jas bodů náležící objektu od jasu bodů reprezentující pozadí, se zabývá metoda thresholding.

3.2.1 Thresholding

Thresholding (prahování) je metoda používaná pro segmentaci obrazu na dvě nebo více částí. Každý obrazový bod je zařazen do jedné ze dvou (případně více) tříd podle hodnoty jasu nebo hodnoty barevné složky podle vztahu 3.1. x je vstupní hodnota jasu, t je práh a A a B jsou nové výstupní hodnoty, často minimální a maximální možný jas resp. hodnota barevné složky.

$$f(x) = \begin{cases} A & \text{pro } x < t \\ B & \text{pro } x \geq t \end{cases} \quad (3.1)$$

Z charakteru vstupních dat, kdy detekovaný objekt (hlasivková štěrbinu) je tmavá a okolí světlé, lze tato metoda použít. Jednou z možností je výběr prahu ručně tak, aby subjektivně byla štěrbinu detekována „co nejlépe“. Z automatických metod lze použít například metoda OTSU [15] (nebo některé její varianty, např. [16] [17]), metoda minimální chyby ([18]) nebo některé další (ukázka na obr 3.1).



Obrázek 3.1: Ukázka prahování snímků (vstupní obraz vlevo) metodou OTSU (uprostřed) a metodou minimální chyby (vpravo) s různým výsledkem. Na výsledcích je vidět, že ani jedna z uvedených metod nelze bez problémů použít na všechny snímky. Metoda OTSU nedokáže správně reagovat na horší světelné podmínky, metoda minimální chyby obsahuje nesprávně segmentované části obrazu a nepřesné hranice. Další metody přináší podobné nebo horší výsledky.

Pro prahování je nejlepší vybrat snímky s co největším rozptylem jasových hodnot, kterými jsou snímky s maximálně otevřenou štěrbinou (světlé okolí, tmavá štěrbinina). Takový snímek lze najít v sekvenci jako ten s nejnižší sumou jasů bodů. Nalezený práh je pak možné aplikovat na každý snímek v sekvenci.

V ideálním případě tato metoda funguje velmi dobře, avšak u většiny obrázků tato metoda nedokáže správně reagovat na nekonzistentní nasvícení, přesvícení či jiné problémy, kdy jas nebo hodnoty v některé z barevných složek nedostatečně diferencují okolí od štěrbin (histogram snímků není bimodální).

Před prahováním se často přesněji lokalizuje oblast zájmu (ROI, viz dále), aby okolí štěrbin, které může být různě osvětleno, v co nejmenší míře ovlivňovalo výsledek prahování. Metoda prahování je využita např. v článcích [19], [20] a v práci [21].

Jiným způsob pro využití prahování je popsán v [22], kde nedochází k prahování jednotlivých snímků, ale je nejdříve vytvořen kymogram (viz kap. 3.5.4) a ten je pak naprahován. V [22] je ještě využit Gaussian mixture model (proložení histogramu Gaussovými křivkami), pomocí kterých lze vybrat konkrétní práh.

Další způsob je neprahovat celý snímek najednou, ale rozdělit jej na oblasti a ty prahovat samostatně. Prahování obrazu po částech ale v tomto případě nepomůže, protože v některých částech snímku nemusí být možné přesně určit hranici štěrbin.

3.3 Metody založené na detekci hran

Principem těchto metod je nalézt v obraze hranice objektu pomocí nalezení strmého jasového gradientu či náhlé změny hodnoty barevných složek na ploše obrazu.

3.3.1 Edge detection

Operátory

Další metodou pro detekci objektů v obraze je nalezení hran a jejich využití pro segmentaci. Často se využívá masek pro výpočet hodnot nad každým bodem v obraze. Nejčastěji se používají Robert operator (3.2), Prewitt operator (3.3) a Sobel operator (3.4). Tyto masky se postupně použijí konvolucí na každý bod obrazu a z okolních bodů (součet součinů hodnot jasu a hodnoty v masce na stejnohlých pozicích) se vypočte výsledná hodnota nového bodu. Často se používají v kombinaci s maskou otočenou o 90°, aby byly detekovány hrany i v kolmém směru.

$$\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \quad (3.2)$$

$$\begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} \quad (3.3)$$

$$\begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \quad (3.4)$$

Další z používaných operátorů je Laplacian operator (3.5), který není třeba kombinovat s pootočenou verzí. Používají se dvě varianty, kdy první umožňuje detekovat hrany ve čtyřech hlavních směrech, druhá detekuje hrany i v diagonálním směru.

$$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix} \begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix} \quad (3.5)$$

V práci [23] je použit Sobel operátor pro výpočet gradientu jako předzpracování. Práce [24] využívá Sobel operátor pro upřesnění oblasti zájmu (hledání ROI). V článku [25] je ale zmíněno, že tyto operace nevykazují dobré výsledky u snímků nižší kvality.

Canny edge detector

Postup pro detekci popisuje algoritmus Canny edge detector. Ten popisuje základní kroky detekce: Odstranění šumu, určení gradientu, nalezení lokálních maxim a eliminace nevýznamných hran. K odstranění šumu je možné použít Gaussův filtr, pro určení hodnoty a směru gradientu je možné použít např. Sobel operator, lokální maxima lze odstranit ztenčením a prahování pomůže s odstraněním nevýznamných hran - k tomu se používají 2 prahy, body s hodnotou větší než T_1 jsou označeny jako body hrany, body s hodnotou mezi prahy T_1 a T_2 jsou označeny za hrany pouze v případě, že sousedí s body již označené jako hrana.

Mathematical morphology

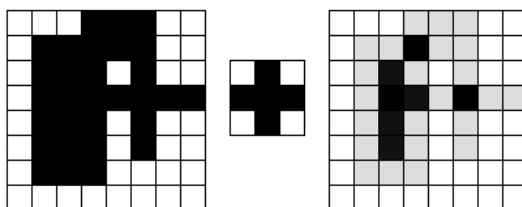
Matematická morfologie je soubor technik pro analýzu a zpracování geometrických struktur. Podrobnější popis je v [26]. Základem je provádět operace pomocí jednoduchých strukturálních elementů (masek) v obraze (např. kruh, čtverec, plus; na obr. 3.2) v obraze. Metody byly určeny pouze pro černobílé obrázky, později proběhla modifikace pro použití na obrázky ve stupních šedi. Následující popis je pro černobílé obrázky.



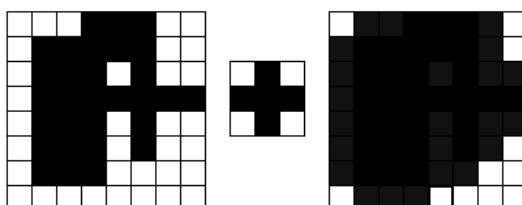
Obrázek 3.2: Příklady strukturálních elementů.

Základní morfologické operace jsou:

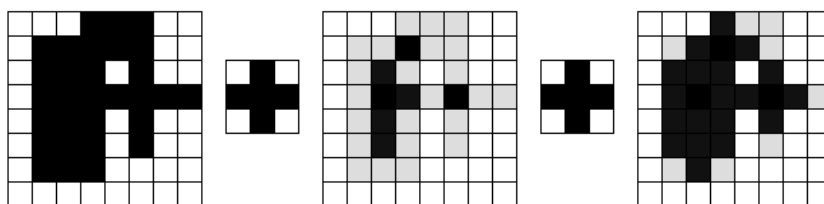
- eroze \ominus - z původního objektu v obraze jsou vybrány pouze ty body, kde zadaná maska sedí na všechny obsažené body (logický operátor AND), objekt v obraze se zmenšuje,



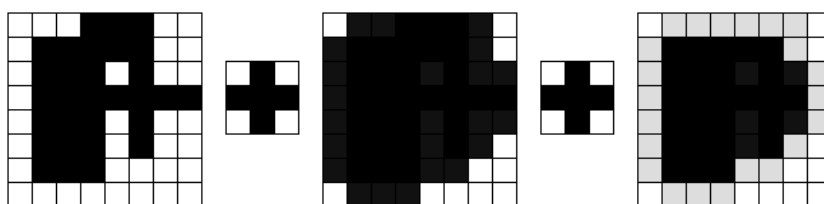
- dilatace \oplus - Do původního objektu v obraze jsou zahrnuty i body z masky (logický operátor OR), objekt v obraze se zvětšuje,



- otevírání \circ - eroze a následná dilatace, vyhlazuje objekt a přerušuje ho v úzkých místech,



- zavírání \bullet - dilatace a následná eroze, vyhlazuje objekt, ale spojuje přerušené objekty.



Pro šedotónové obrazy se používá pro erozi supremum a pro dilataci infimum z hodnot jasu v oblasti masky.

Matematická morfologie je pro detekci hlasivkové štěrbině použita v [25]. V této práci se pracuje se šedotónovými obrazy a operace jsou určeny vztahy 3.6 a 3.7, kde I je snímek hlasivek, B strukturní element, x, y, s, t jsou souřadnice.

$$(I \ominus B)(x, y) = \max(I(x + s, y + t) - B(s, t)) \quad (3.6)$$

$$(I \oplus B)(x, y) = \max(I(x - s, y - t) + B(s, t)) \quad (3.7)$$

Tyto operace jsou použity pro detekci hran použitím vztahů 3.8 a 3.9.

$$M = (I \bullet B) \circ B \quad (3.8)$$

$$Eadg = (M \bullet B) \oplus B - (M \bullet B) \quad (3.9)$$

Výsledek ale nemusí být vždy dokonalý, může být detekováno více hran i mimo hlasivkovou šterbinu. Proto je použit další postprocessing. V uvedeném článku je použit Hough transformation (popsáno dále). V práci je zobrazen lepší výsledek oproti použití operátorů detekce hran, obsahuje také srovnání s metodou Active contours a prahováním.

3.3.2 Gabor filtering

Gaborův filtr (D. Gabor, 1900 - 1979) se často používá pro zpracování obrazu. Je to lineární filtr, jehož impulzní odezva je definována jako Gaussova funkce modulovaná funkcí harmonickou [27].

V 1D prostoru jde například o sinusovou funkci vynásobenou Gaussovou funkcí, ve 2D je to pak sinusoid a gaussovské jádro. Od fourrierovské transformace se Gaborovy filtry liší v tom, že umožňují lokalizaci frekvencí ve snímku.

Gaborův filtr lze vyjádřit vztahem 3.10, kde x a y je umístění, θ frekvence a ϕ orientace. σ je směrodatná odchylka Gaussovského jádra.

$$g(x, y, \theta, \phi) = \exp\left(-\frac{x^2 + y^2}{\sigma^2}\right) \cdot \exp(2\pi\theta i(x \cos \phi + y \sin \phi)) \quad (3.10)$$

Pro aplikaci Gaborových filtrů (například konvolucí) se používá připravená knihovna funkcí, která odpovídá charakteru obrazu a hledaných objektů.

Tato metoda je použita v [23] spolu s metodou Wiener motion estimation (popsaná v sekci 3.5.2), která je použita jako preprocessing pro výběr regionu zájmu. Knihovna Gaborových funkcí je vyjádřena vztahem 3.11.

$$g(f, \theta, x, y) = \exp\left(i(f_x x + f_y y) - \frac{f^2(x^2 + y^2)}{2\sigma^2}\right) \quad (3.11)$$

kde

$$f_x = f \cos \theta, f_y = f \sin \theta, i = \sqrt{-1} \quad (3.12)$$

Úspěšnost detekce uvedená v této práci je kolem 95 %. Autoři uvádějí případy selhání, kdy je šterbina uzavřena nebo je pohyb hlasivek pomalý (malé rozdíly mezi po sobě jdoucími snímky).

3.3.3 Active contours

Princip metody Active Contours je založen na minimalizaci energie křivky, která rozděljuje obraz na jednotlivé oblasti (podle [28]). Nejprve je stanovena počáteční pozice a tvar křivky (nemusí být spojitá) a stanovena energie. Energie má 2 složky, složku vnitřní energie (spojitost a zakřivení křivky) a složku vnější energie, kde je křivka zvýhodněna, pokud v obrazu vede v místě hrany. Postupně se iterativně zmenšuje energie křivky, až dosáhne minima, kdy v ideálním případě křivka kopíruje hranici objektů.

Vnitřní energii lze definovat jako součet energií spojitosti a zakřivení, viz vztah 3.13, kde ω jsou váhové funkce.

$$E_i = \omega_{cont}E_{cont} + \omega_{curv}E_{curv} \quad (3.13)$$

Vnější energie je pak kombinace energie obrazu, hrany a ukončovací funkce. Dále můžou vnější energii ovlivnit další síly. Ve vztahu 3.14 jsou ω opět váhové funkce.

$$E_e = \omega_{line}E_{line} + \omega_{edge}E_{edge} + \omega_{term}E_{term} + F_{con} \quad (3.14)$$

Tato metoda je dobře funkční u obrazů s bimodálním histogramem a není příliš citlivá na šum, dokud nenarušuje bimodalitu histogramu. Záleží ovšem na nastavení parametrů a na počáteční poloze a tvaru křivky. Je více implementací tohoto algoritmu, jelikož je ale princip iterativní, bývá větší časová náročnost na výpočet.

Metoda Active contours je pro segmentaci obrazu hlasivek využita v [29] (kombinace s prahováním) a [30] (3D active contours). Jako počáteční umístění křivky může sloužit preprocessing - nalezení oblasti zájmu a přibližná poloha a tvar šterbiny.

3.4 Metody využívající topologii obrazu

Metody využívající rozložení objektů, barev a tvarů ve snímku. Jako zástupce těchto metod je uvedena metoda watershed segmentation.

3.4.1 Watershed segmentation

Jedna z nejpoužívanějších metod využívající topologii pro segmentaci obrazu se jmenuje Watershed segmentation (segmentace rozvodím). Metoda vychází z topografie a je obdobou hranic rozvodí v krajině.

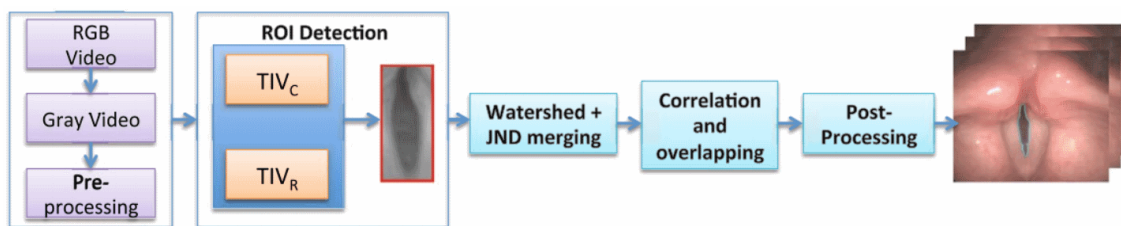
Obraz si můžeme představit jako výškovou mapu, kde hodnota jasu (nebo některá z barevných složek) je výškou v krajině. Metodu watershed si můžeme přestavit jako

simulaci chování vody. Z každého bodu necháme stékat vodu, která si vybírá směr nejstrmějšího klesání, dokud neseče do údolí - lokálního minima. Tento postup se aplikuje na všechny body v obraze a tím je každý bod přiřazen k minimu. Všechny body patřící k jednomu minimu patří do jednoho rozvodí a mezi jednotlivými rozvodími vznikají hranice, čímž získáváme výsledek segmentace [31].

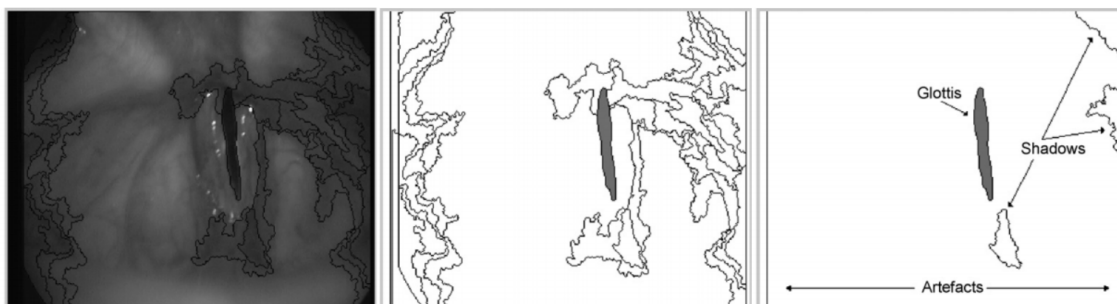
V některých případech může vzniknout velké množství oblastí v místech, kde jich očekáváme pouze několik. Tento stav se nazývá oversegmentation (přesegmentování) a je nejčastěji způsobený šumem ve vstupním obraze, kvůli kterému je v obraze mnoho lokálních minim. Řešením této situace je vstupní obraz předzpracovat a šum odstranit (více v sekci 3.5.3). Pokud ani odstranění šumu nezabrání vzniku velkého počtu oblastí, je možné pro spojování podobných oblastí použít metodu Region merging (více v sekci 3.6.1).

Hojně využívanou technikou je použití watershed metody na gradientní obrázek (hodnoty jasu jednotlivých bodů jsou hodnoty gradientu původního obrazu), který má nejvyšší hodnoty v oblastech hran. Právě tohoto lze využít a metoda watershed pak vyhledá oblasti s podobnými vlastnostmi v původním obraze. V práci [32] jsou ještě hodnoty gradientu omezeny, aby vzniklo méně lokálních minim (if $Y < 2$ then $Y = 0$), což odstranilo 20% nežádoucích oblastí. Ukázka segmentace je na obr. 3.4.

Tato metoda bylo použita pro detekci hlasivkové štěrbiny také v [33]. Schéma procesu segmentace je znázorněné na obr. 3.3.



Obrázek 3.3: Schéma postupu segmentace využívající watershed transformaci a region merging (převzato z článku [33]).



Obrázek 3.4: Ukázka segmentace watershed transformací a následný region merging (převzato z článku [32]).

3.4.2 Region growing

Region growing (růst oblasti) je metoda pro detekci části obrazu rozpínáním oblasti z počátečního bodu podle zadaných kritérií [34]. Nejprve je nutné určit počáteční bod (tzv. seed), ze kterého se postupně procházejí okolní body a testuje se podobnost jasu či barevné složky z počátečním bodem. Pokud jsou si body dostatečně podobné, je tento bod zahrnut do oblasti a pokračuje se dalšími sousedními body.

Problém může být zvolení počátečního bodu, protože na tom závisí výsledek detekce. Např. v práci [35] je počáteční bod vybrán ručně po interakci s uživatelem.

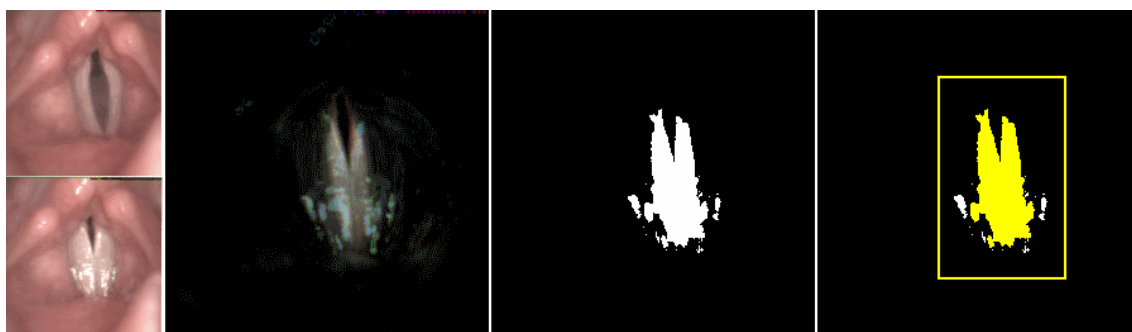
3.5 Předzpracování

Předzpracování (preprocessing) je úprava vstupního obrazu před samotným zpracováním, v tomto případě detekcí hlasivkové štěrbinu.

3.5.1 Oblast zájmu

Ve většině případů není vhodné zpracovávat celý snímek, ale pouze jeho část, která obsahuje objekt zájmu (region of interest, ROI), v tomto případě hlasivkovou štěrbinu. Omezení oblasti je vhodné pro urychlení následujících procesů a i pro zlepšení výsledků, protože omezením oblasti můžou být odstraněny rušivé elementy způsobující falešnou detekci.

Pro nalezení oblasti zájmu na snímcích z HSV lze použít jednoduchou metodu založenou na odečtení snímku s nejvíce otevřenou a s nejvíce uzavřenou hlasivkou. V ideálním případě je jediná nenulová oblast ve výsledném obraze právě oblast hlasivkové štěrbinu. Ve většině případů je ale nutné výsledný obraz naprahnovat, případně nalézt největší souvislou plochu (např. pomocí Connected component labeling [36]). Tento postup je znázorněn na obr. 3.5.

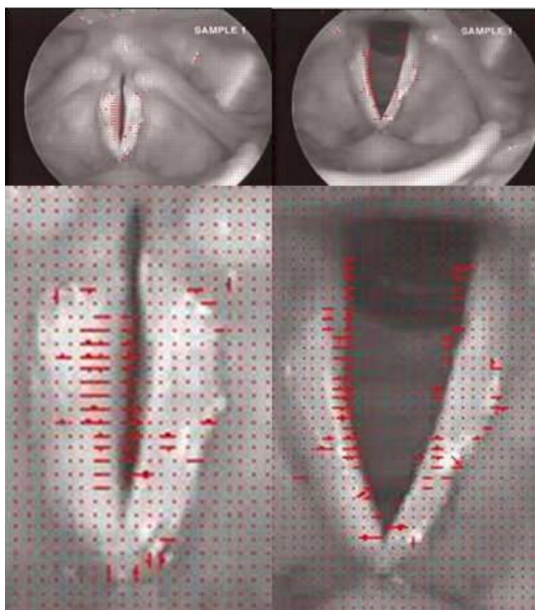


Obrázek 3.5: Nalezení oblasti zájmu. Zleva snímky s otevřenou a zavřenou hlasivkou, odečtený obraz, naprahnovaný obraz (otsu), nalezaná největší souvislá plocha a určený ROI.

3.5.2 Wiener motion estimation

Pro získání regionu zájmu je možné využít motion estimation na po sobě jdoucí snímky v sekvenci. Podle získaných vektorů můžeme z předpokladu, že největší pohyb ve snímku způsobuje kmitající šterbina, vybrat právě tu část snímku s největším pohybem.

Tuto metodu, i když na snímcích ze záznamu využívajícího stroboskopický jev, můžeme vidět např. v [23]. Výsledek této metody je zobrazen na obrázku 3.6.



Obrázek 3.6: Zobrazení vektorů, které jsou výsledkem metody Wiener motion estimation (převzato z článku [23]).

3.5.3 Odstranění šumu

Odstranění šumu, potlačení šumu nebo redukce šumu je proces, který se pokouší odstranit z obrazu (nebo jiných signálů) šum. Pro šum je často typická vysoká frekvence, nízká energie a malá korelace s původním signálem. Pro odstranění šumu je užitečná znalost charakteristik šumu, což může pomoci k výběru konkrétní metody.

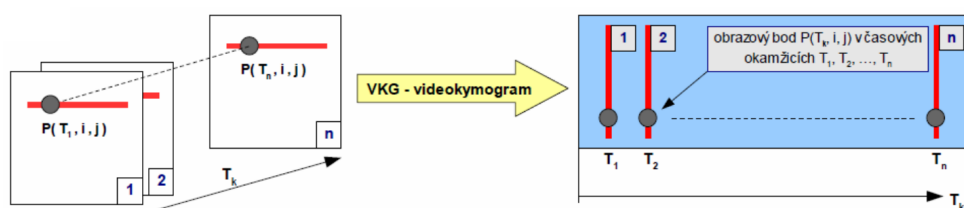
Typy metod pro odstranění šumu v obraze:

- **Lineární filtry** - nejčastěji mají charakter dolní propusti (nepropouští vysoké frekvence typické pro šum), aplikují se například konvolucí.
- **Gaussův filtr** - speciální konvoluce s maskou, která se skládá z elementů určených Gaussovou funkcí. Způsobí rozmazání obrazu, ale dokáže odstranit Gaussův šum.

- **Průměrování** - speciální případ Gaussova filtru (kde směrodatná odchylka se blíží k nekonečnu), kde hodnota pixelu je průměr hodnot okolních bodů.
- **Nelineární filtry** - mezi nelineární filtry patří např. mediánový filtr (nová hodnota pixelu je medián okolních bodů), konzervativní vyhlazení (hodnota pixelu vyšší než maximum okolních bodů je nahrazena tímto maximem, resp. hodnota nižší než minimum je nahrazena tímto minimem), prahování vlnkových koeficientů (použití diskretní vlnkové transformace) či bilaterální filtr.

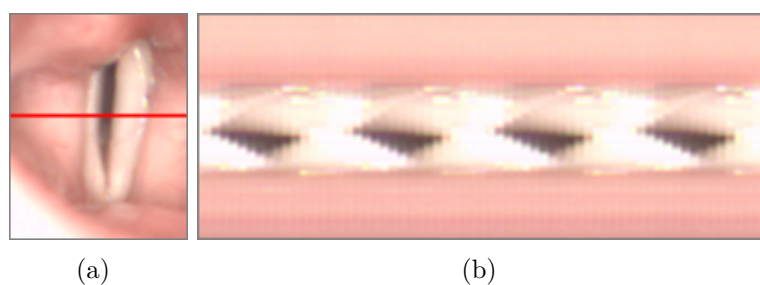
3.5.4 Vytvoření kymogramu

Kymogram je obrazové znázornění pohybu hlasivek na vybraném řezu hlasivkové štěrbině v čase. Jedna osa obvykle vyjadřuje souřadnici obrazového bodu v přímce kolmé na osu hlasivky (většinou se pro zjednodušení metody používá jeden obrazový řádek) a druhá osa je číslo snímku (blíže vysvětleno v [4] a [37]).



Obrázek 3.7: Princip kymografie.

Dříve, než byly k dispozici vysokorychlostní kamery, byly používány speciální snímáčí zařízení pro snímání právě jednoho řezu hlasivek (zjednodušeně řečeno kamera s jedním řádkem). Při použití HSV stačí vybrat jeden z řádků na jednotlivých snímcích. Příslušné řádky ze všech snímků umístěné za sebou tvoří výsledný obraz (obr. 3.8). Tmavší oblast znázorňuje oblast hlasivkové štěrbině.



Obrázek 3.8: Kymogram hlasivek. a) snímek hlasivek a místo videokymografického řezu, b) kymogram.

V kymogramu lze dobře pozorovat pohyb určitého místa hlasivek. Většinou pomocí jednoduchého prahování dokážeme sledovat vývoj šířky hlasivkové štěrbině a pohyb levé a pravé hlasivkové řasy. Nevýhodou je samozřejmě absence celé štěrbině pro komplexnější analýzu chování.

3.6 Postprocessing

Po samotném zpracování obrazu je někdy nutné ještě výsledek upravit, aby vyhovoval požadovaným vlastnostem. Toto následné zpracování se nazývá postprocessing.

3.6.1 Region merging

Pokud obraz po použití metody segmentace (např. watershed segmentation) obsahuje příliš mnoho oblastí, nejlepší metodou proti tomuto „přesegmentování“ je využít metodu Region merging jako postprocessing. Principem je sjednocení oblastí podobných vlastností [38].

V každém kroku se vyhodnotí cena sjednocení oblastí a „nejlevnější“ sjednocení je poté provedeno. Algoritmus končí, pokud bylo dosaženo požadovaného počtu oblastí nebo byl překročen práh ceny.

Varianty této metody se liší především oceňovací funkcí a hodnotou prahu.

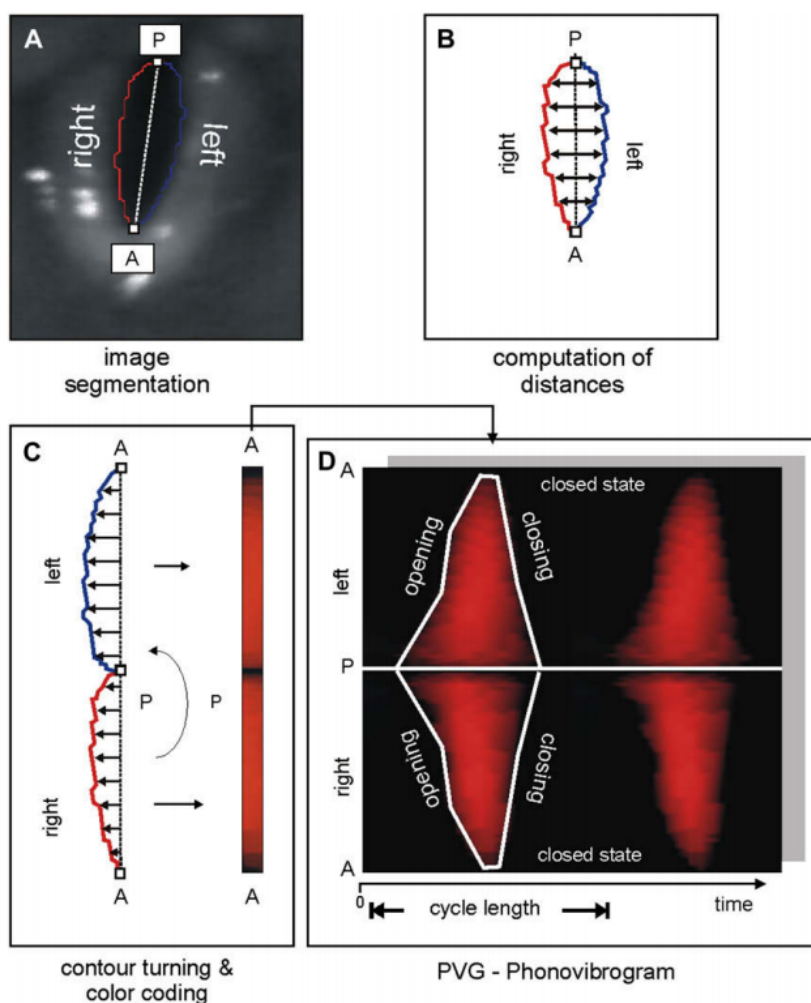
3.6.2 Hough transformation

Hough transformation je metoda extrakce příznaků často používaná v postprocessingu ve zpracování obrazu [36]. Původní Hough transformation dokázala detekovat přímky v obraze, později byla rozšířena pro detekci a lokalizaci libovolných tvarů, často kružnic a elips. Dnes se používá Generalized Hough Transform (zobecněná Houghova transformace) od R. Dudy a P. Harta [39].

Přímky (či jiné objekty) jsou parametrizovány (používá se délka kolmice procházejícím počátkem a její úhel od osy x) a prostor těchto parametrů se nazývá Houghův prostor (akumulační prostor). Postupně se prochází všechny body ve zpracovaném obraze a zaznamenávají se hodnoty parametrů přímek různých úhlů procházejících daným bodem. Protože přímky procházející jedním bodem ve snímku tvoří v Akumulačním prostoru část sinusové křivky, nalezením průsečíku více těchto křivek (např. lokální maximum součtu) nalezneme hledanou přímku procházející danými body v obraze.

3.6.3 Phonovibrogram

Phonovibrogram (PVG) je možností, jak vizualizovat chování hlasivkové štěrbiny [40]. Vstupem této metody je detekovaná štěrbina a osa symetrie. Poté je vypočtena vzdálenost každého bodu hranice od osy na levé a pravé straně štěrbiny, a tyto vzdálenosti jsou pak nanesené jako hodnoty jasu do obrazu. Schéma tvorby phonovibrogramu je na obr. 3.9.



Obrázek 3.9: *Princip tvorby phonovibrogramu (z [40]).*

Výsledný obraz zobrazuje rychlost otevírání a zavírání hlasivek na jednotlivých bodech hranice a lze sledovat amplitudovou asymetrii mezi periodami a mezi hlasivkovými řasami. Další možnosti analýzy jsou popsány např. v [41].

3.7 Shrnutí

Dosažené výsledky jednotlivých metod jsou většinou dobré, ve většině případů je ale použit takový korpus snímků, který je pro metody výhodný. Z dostupných dat z lékařských vyšetření je ale nezanedbatelné množství snímků zhoršené kvality a je tedy žádoucí pro detekci použít více adaptivní metodu, aby dokázala zpracovat i snímky se zhoršenou kvalitou. Proto byla navržena nová metoda (viz kapitola 4), která by měla toto kritérium splňovat.

4 Vlastní metoda pro detekci hlasivek

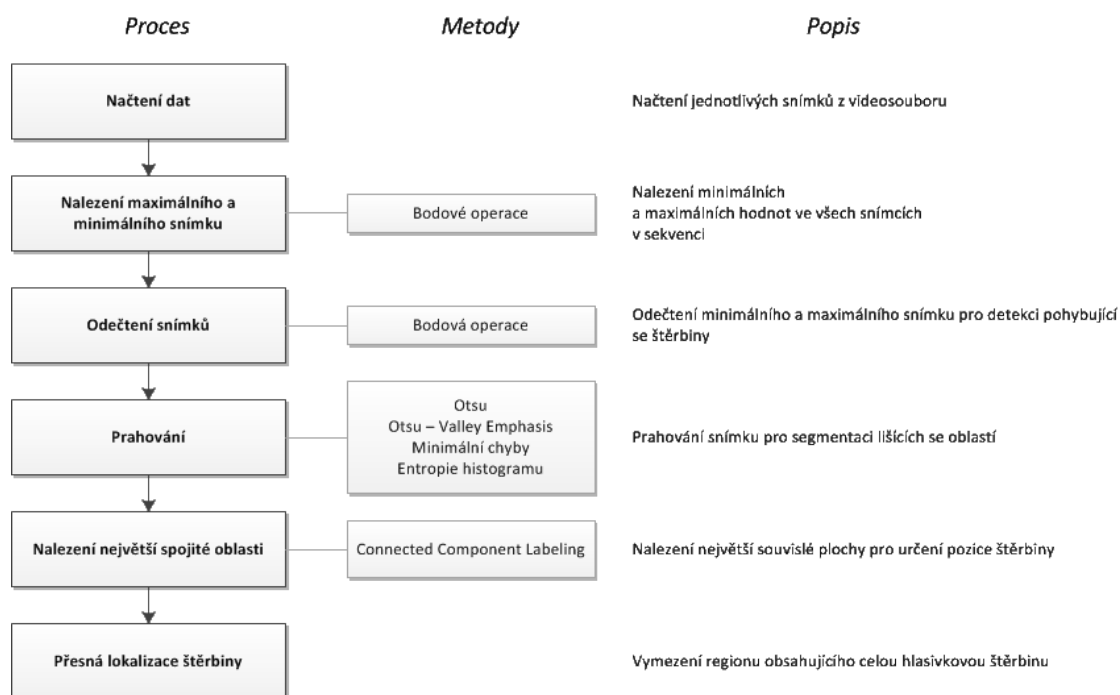
Pro detekci hlasivkové štěrbiny jsem použil 2 metody. V rámci práce [21] byla vyvinuta první z nich, zakládající se na hledání ROI a metodě prahování „Max-Min-Thresholding“.

Druhá metoda také vyžadovala nalezení ROI, ale vlastní detekce byla založena na shlukové analýze, při které byly obrazové body děleny do tříd a jedna z těchto tříd odpovídala hlasivkové štěrbině. Tato metoda vykazuje zajímavé výsledky a úspěšnost je o desítky procent vyšší než u první metody.

4.1 Nalezení oblasti zájmu

Stejně jako u většiny výše uvedených metod je vhodné provést předzpracování snímků, kterým je nalezení oblasti zájmu (ROI - region of interest) - oblasti, která obsahuje štěrbinu, ale co nejméně okolí.

Stávající algoritmus pro hledání ROI je na obr. 4.1:



Obrázek 4.1: Postup při hledání ROI [21])

Výsledky této metody jsou dostatečné pro většinu případů. Na některých snímcích, například s malým pohybem hlasivek či s jinými rušivými elementy (pohyb kamery,

tekutin), však uvedená metoda reaguje nesprávně například tím, že je celá oblast snímku či naprosto odlišná pozice označena jako ROI.

Vylepšení metody nalezení regionu zájmu by mohlo být jednou z oblastí pro další výzkum, protože úspěšnost určení regionu zájmu implikuje úspěšnost dalšího zpracování obrazu.

4.2 Segmentace shlukovou analýzou

Po vyčerpání možností předchozí metody „Max-Min-Thresholding“ byly testovány i jiné metody, jedna z nich byla i metoda shlukové analýzy klasifikace obrazových bodů, která poskytovala již ze začátku slibné výsledky.

Ve většině případů je hlasivková štěrbina tmavší než okolí, a to ještě nejvýrazněji v červené složce barevného modelu RGB. Proto jsme dále jednotlivé obrazové body považovali za objekty s parametry odvozenými z charakteru obrazových dat. Takto chápané obrazové body je pak možné pomocí shlukové analýzy rozdělit do tříd s podobnými vlastnostmi, viz např. [42].

Obecně, shluková analýza reprezentuje postupy seskupování objektů do více či méně homogenních skupin na základě jejich vzájemné podobnosti. Nejpoužívanější metodou shlukové analýzy je metoda K-means, viz [10]. Podle jednotlivých parametrů se objekty klasifikují do k -tříd ($CLASS_j$) s co nejmenšími rozdíly parametrů v rámci třídy nebo s co největším rozdílem parametrů mezi třídami. Jednotlivé klasifikované objekty se umístí do m -rozměrného prostoru, kde m je počet sledovaných parametrů. Souřadnice m -rozměrného prostoru jsou tak indexem k hodnotám parametrů a každý parametr může mít přiřazenu jinou váhu. V tomto prostoru se poté počítají vzdálenosti mezi objekty a hledají se shluky objektů s podobnými vlastnostmi. Vstupem metody jsou zvolená tzv. centra ($Center_j$), kterými mohou být některé vybrané objekty nebo objekty nově vytvořené tak, aby řešení přesně a rychle dospělo k rozdělení objektů do tříd.

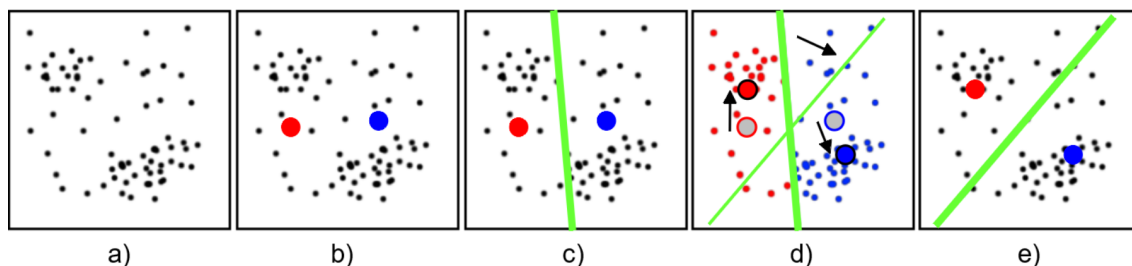
Prvním krokem metody je prvotní rozdělení vstupních dat xx_i , $i = 1$ až n , kde n je počet objektů, do tříd $CLASS_j$ podle zvoleného kritéria. Je použito kritérium nejmenší vzdálenosti objektu vzhledem k jednotlivým prvotním centrům, $Center_j$, viz vztah 4.1, kde $j = 1$ až k , k je počet tříd $CLASS_j$.

$$CLASS_j = \arg \min_{j=1..k} ||xx_j - Center_j|| \quad (4.1)$$

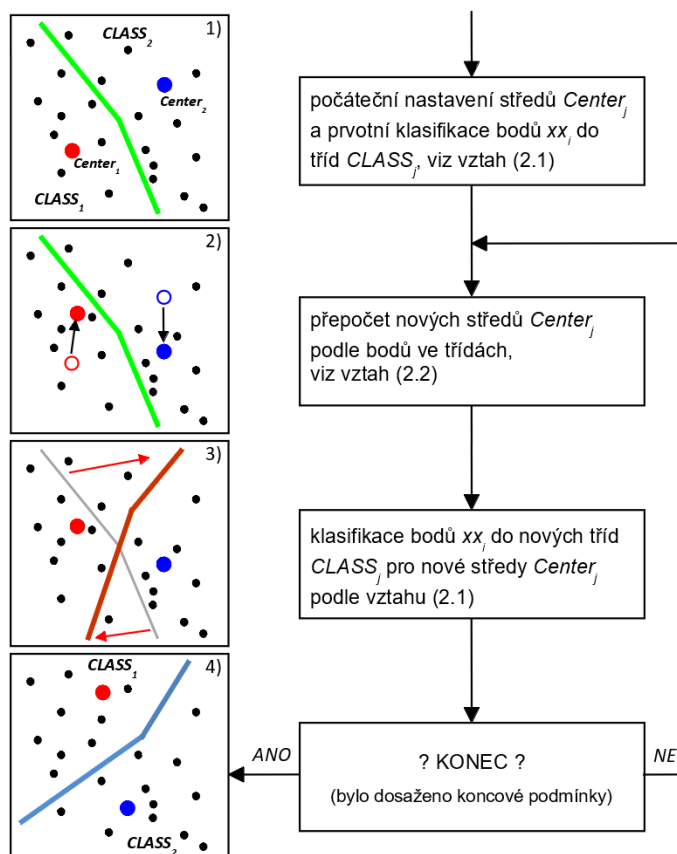
Po rozdělení do tříd se ve druhém kroku stanoví nová centra $Center_j$ jednotlivých tříd, např. jako jejich těžiště, podle objektů xx_i , které náleží do třídy $CLASS_j$, podle vztahu 4.2, kde n_j je počet objektů xx_i ve třídě $CLASS_j$.

$$Center_j = \frac{1}{n_j} \sum_{i \in CLASS_j} xx_i \quad (4.2)$$

Tuto metodu je možné opakovat s novými vstupními daty až do stabilního stavu, kdy se již rozdělení bodů do tříd nemění, případně klasifikaci ukončit po stanoveném počtu cyklů (obr. 4.2). Vizualizace algoritmu je na obr. 4.3.



Obrázek 4.2: Vizualizace principu shlukové analýzy v jednotlivých krocích a) rozmístění objektů xx_n ve snímku region R ; b) zvolená počáteční centra - $Center_j$, $j = 1, 2$; c) prvotní rozdělení xx_j do dvou tříd $CLASS_j$; d) výpočet nových center $Center_j$ a nové přiřazení xx_j do tříd $CLASS_j$; e) uspořádání center $Center_j$ a tříd $CLASS_j$ po prvním cyklu.



Obrázek 4.3: Vizualizace algoritmu shlukové analýzy K-means. Klasifikační třídy: $CLASS_1$, $CLASS_2$ středy klasifikačních tříd: $Center_1$, $Center_2$.

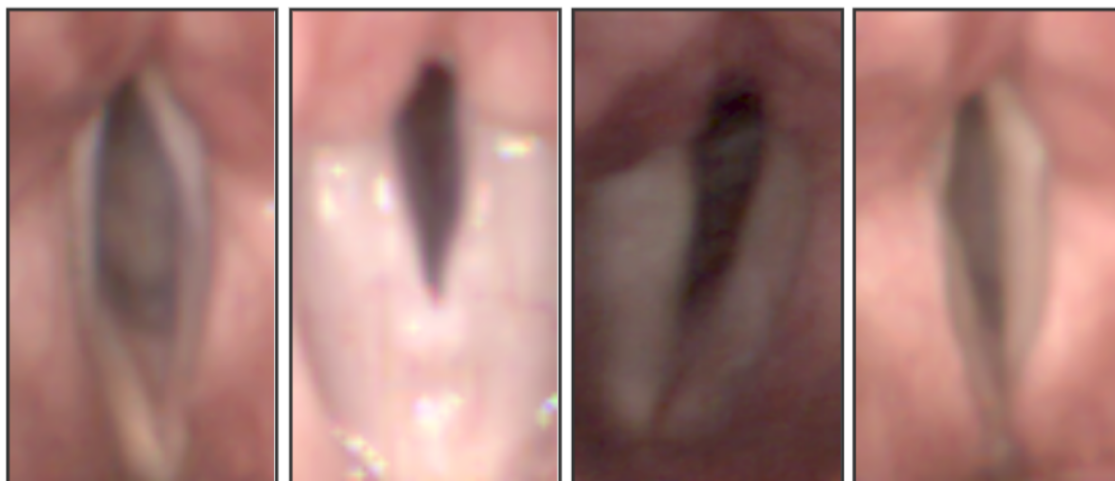
Pokud budeme obrazové body považovat za objekty, nabízí se několik parametrů, které lze zvolit pro klasifikaci včetně nastavení vah jednotlivých parametrů. Optimální nastavení vah je velice složitá úloha typu heuristiky, kdy jsou hodnoty jednotlivých složek určeny experimentálně. Popis a nastavení vah parametrů je uvedeno v tabulce, viz tabulka 4.1.

parametr	obor hodnot	$coef_0$	$coef_1$
R – červená složka	hodnota složky (0 - 255)	1,0	1,0
G – zelená složka	hodnota složky (0 - 255)	0,5	0,1
B – modrá složka	hodnota složky (0 - 255)	0,5	0,1
X – souřadnice x ve snímku	hodnota souřadnice (0 - 255)	1,0	1,0
Y – souřadnice y ve snímku	hodnota souřadnice (0 - 255)	1,0	0,1
R-B – rozdíl R a B složky	rozdíl složek (-255 - 255)	1,0	1,0
C – vzdálenost od středu	vzdálenost v pixelech (0 - 255)	1,0	1,0

Tabulka 4.1: Popis jednotlivých parametrů s počátečním nastavením vah $coef_0$ a výsledným nastavením vah $coef_1$.

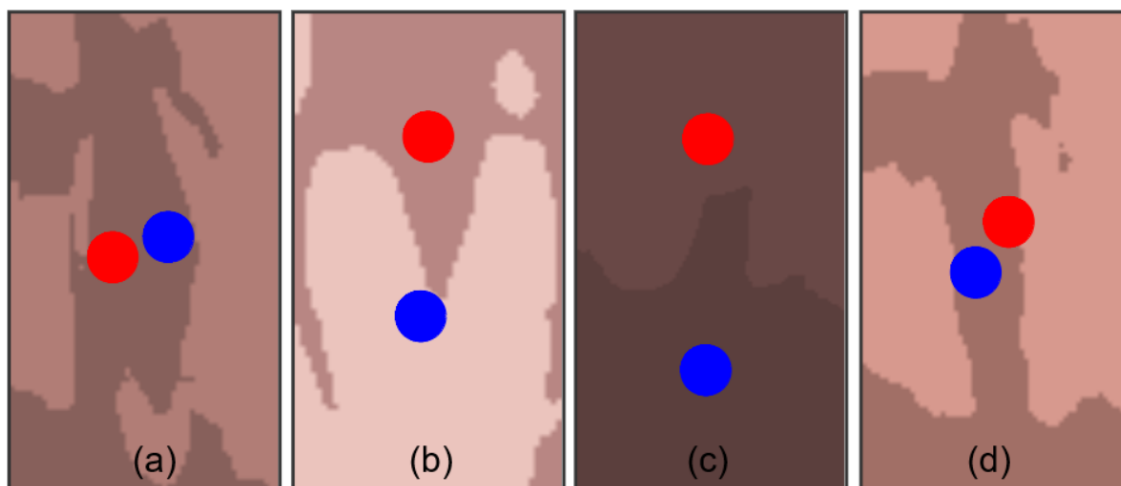
Před aplikací metody shlukové analýzy je nutné určit počet tříd, resp. počet počátečních center. Pro detekci bodů patřících do oblasti hlasivkové štěrby by měly stačit dvě třídy. Ve snímcích s nízkým kontrastem nebo snímcích, kde okolí štěrby obsahuje větší rozptýl jasů, mohou však být dvě třídy nedostatečné.

Pro testování metody byly vybrány čtyři typové snímky s lokalizovanou oblastí hlasivek, obr. 4.4.



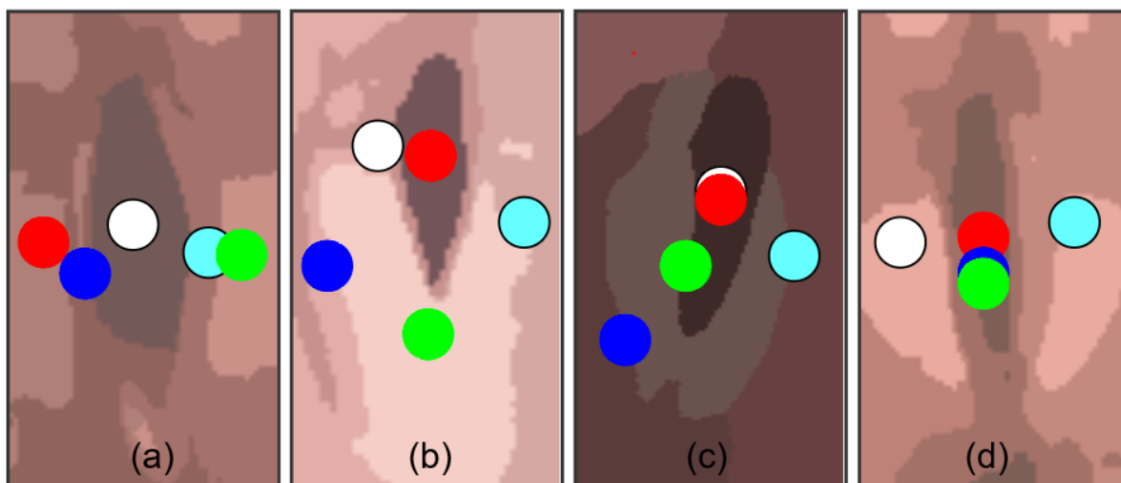
Obrázek 4.4: Snímky s lokalizovanou oblastí hlasivek

Vizualizace výsledků zobrazuje snímek po klasifikaci, obr. 4.5, kdy oblast bodů patřících do jedné třídy je zobrazena stejnou barvou, která je průměrem barvy bodů dané třídy. Označený barevný bod je potom centrem třídy, umístěný na pozici průměrných souřadnic všech bodů.



Obrázek 4.5: Výsledek klasifikace po testu: počet tříd $j = 2$, $coef f_0$

V dalších testech byl postupně zvyšován počet tříd $j = 3, 4, 5$ a byla použita různá nastavení vah parametrů. V závěru se nastavení vah ustálila na $coef f_1$, viz tab. 4.1.



Obrázek 4.6: Výsledek klasifikace po testu: počet tříd $j = 5$, $coef f_1$

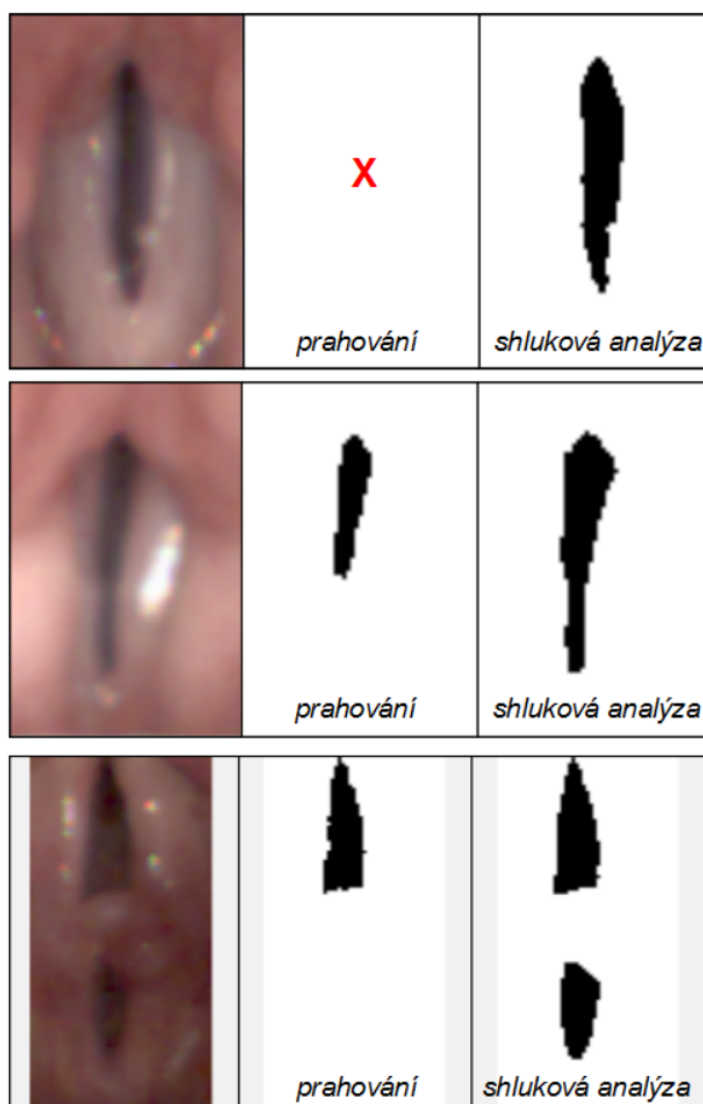
V posledním testu se změny vah parametrů a zvolený počet tříd projeví výrazně a hlasivková štěrbina je dobře detekovatelná nejen ve všech testovaných snímcích, viz obr. 4.6, ale i v dalších typech snímků. Vypočtená centra u nejvíce otevřené hlasivky jsou pak použita pro všechny snímky ve videosekvenci.

Počáteční zvolení center ovlivňuje rychlost a výsledek klasifikace. Pro nejrychlejší klasifikaci je vhodné určit centra, která mají co nejodlišnější parametry. V prostoru by to tedy měly být ty nejvzdálenější. Pokud jsou centra vybrána náhodně, může výsledek být rozdílný při opakovaném použití shlukové analýzy. Metoda totiž nemusí vždy najít globální optimum.

4.2.1 Výsledky testování

Metoda shlukové analýzy výrazně zvyšuje úspěšnost přesné detekce hlasivkové štěrbiny oproti metodě prahování Max-Min-Thresholding. Při použití vzorku vybraných 130 videosekvencí (z celého korpusu viz kap. 7.5), které byly z hlediska detekce štěrbiny komplikované, byla shlukovou analýzou správně detekována štěrbina v 97 případech, prahováním ve 42 případech. Z toho byl výsledek metody prahování pouze ve 4 případech videosekvencí detekce štěrbiny subjektivně lepší než u metody shlukové analýzy. Příklady úplného nebo částečného selhání při nastavení správné hodnoty prahu u metody Max-Min-Thresholding v porovnání se shlukovou analýzou jsou na obr. 4.7.

Testovací soubor obsahoval snímky dobré kvality a snímky horší kvality.



Obrázek 4.7: Příklady úplného a částečného selhání detekce štěrbiny prahovací metodou Max-Min-Thresholding v porovnání s detekcí pomocí shlukové analýzy.

Za úspěšně detekované snímky byly hodnoceny takové, kde detekovaná štěrbina odpovídala celé ploše anatomické štěrbině a neobsahovala žádné další větší plochy v okolí. Detekovaná štěrbina u většiny snímků, které byly hodnoceny neúspěchem, odpovídala štěrbině v originálním snímku, obsahovala ale další oblasti z okolí, případně chyběla malá část štěrbině u snímků s menším kontrastem.

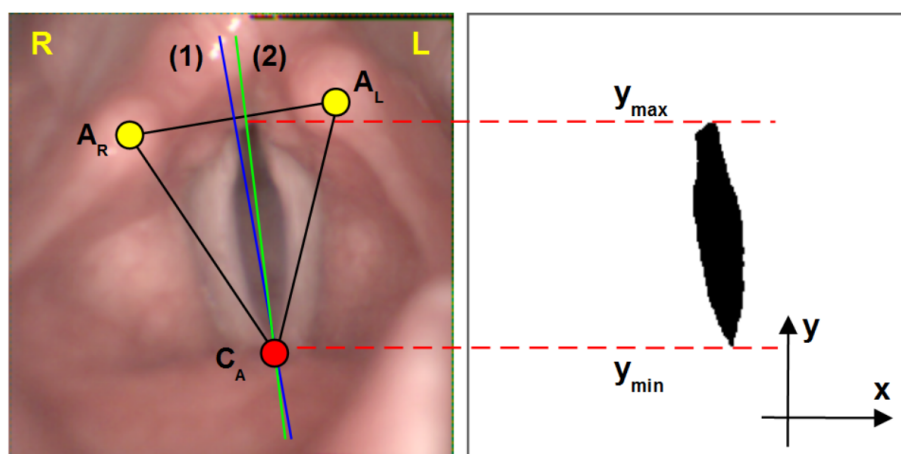
Vybraný počet videosekvencí	130	100 %
Úspěšná detekce	97	74,6 %
Neúspěšná detekce	33	25,4 %

Tabulka 4.2: Úspěšnost detekce.

5 Detekce osy symetrie hlasivkové štěrbiny

V komerčních systémech se většinou automatická detekce osy neprovádí nebo je poloha osy stanovena ručně dvěma body, které odpovídají poloze přední a zadní komisury. Protože dochází během snímání sekvence snímků ke vzájemnému pohybu kamery a hrtanu, může se poloha, velikost i natočení hlasivek v jednotlivých sekvencích snímků lišit. Kromě toho dochází k překrytí části hlasivek některou z chrupavek, k ořezu štěrbiny na okraji oblasti záznamu, zakrytí jinou anatomickou strukturou nebo zamlžení snímků. Metoda detekce osy musí proto tyto odlišnosti v přiměřené míře akceptovat.

Plocha štěrbiny je po její detekci definována černými obrazovými body v černobílém obrázku, viz obr. 5.1.



Obrázek 5.1: Původní snímek hlasivek zachycený vysokorychlostní kamerou. Hlasivková štěrbina je ve fázi maximálního otevření během fonace vokálu „i“: (1): anatomická osa hlasivky; (2): osa symetrie hlasivkové štěrbiny; AL, AR: levý a pravý hrbol c. arytenoidea; CA: - přední komisura (anterior commissura).

Na základě sledování chování hlasivek, resp. hlasivkové štěrbiny vzhledem k poloze přední a zadní komisury a anatomické ose hlasivek, byly stanoveny následující výchozí předpoklady a byla vytvořena robustní metoda pro odhad parametrů hlavní osy hlasivek.

Předpoklady:

- osa kopíruje štěrbinu při uzavřeném stavu;
- při nedovírání hlasivek osa prochází středem štěrbiny v neuzavřenějším stavu;
- umístění hlasivek se v sekvenci snímků nemění, tj. snímky jsou stejnohlé.

Detekce osy ve fázi uzavřené štěrbiny, předpoklad a), je velice obtížná. Její pozici je ale možné určit podle plochy štěrbiny těsně před uzavřením a po jejím otevření. Tato místa jsou evidentně ta, kde se nachází body štěrbiny nejčetněji. Platí tedy, že od okamžiku těsně po otevření až po její uzavření jsou tato místa součástí plochy štěrbiny otevřené hlasivky. Hlasivky se ale nemusí otevírat rovnoměrně po celé jejich délce a jejich pohyb může začínat u jednoho z okrajů. Proto je nutné sledovat chování hlasivek až do jejich úplného otevření, viz obr. 5.2



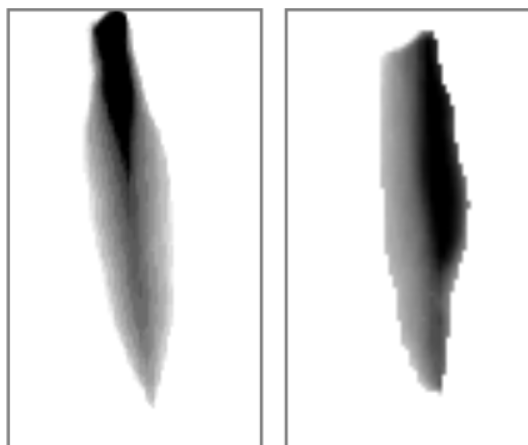
(a) Symetrická hlasivka



(b) Nesymetrická hlasivka

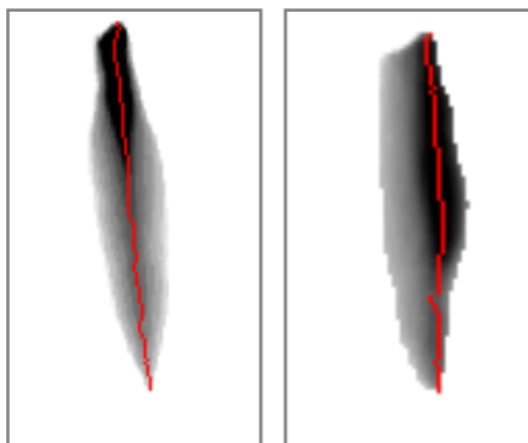
Obrázek 5.2: Změny štěrbiny v sekvenci snímků pro symetrické (a) a nesymetrické (b) chování hlasivek; fáze otevření - uzavření - otevření.

Zde lze využít předpoklad c), že snímky jsou stejnohlé a jejich poloha se v sekvenci nemění. Promítnutím všech snímků sekvence a sečtením stejnohlých obrazových bodů získáme matici, která má význam četností výskytů hlasivkové štěrbiny v jednotlivých bodech snímku. Pro vizualizaci takto získané informace vydělíme hodnotu všech bodů počtem snímků v sekvenci. Tím dojde k efektu bodové filtrace průměrováním podle stejnohlých obrazových bodů a vznikne šedotónový obrázek, ve kterém nejtmaší místa jsou ta, kde je hlasivka nejdéle otevřená. V těchto bodech předpokládáme podle této metody pravděpodobné umístění osy symetrie hlasivkové štěrbiny (obr. 5.3)



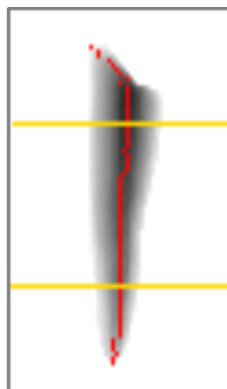
Obrázek 5.3: Šedotónové obrázky získané sečtením jednotlivých snímků v definované sekvenci a následnou bodovou filtrací pro případ symetrického a nesymetrického chování hlasivek.

Díky metodě snímání vysokorychlostní kamerou můžeme dále předpokládat přibližně svislou orientaci osy. Proto v dalším upřesníme její pozici nalezením nejtmašího bodu v každém řádku vypočteného snímku. Při nedovírání hlasivkové štěrbiny dochází k výskytu několika stejně tmavých bodů v jednom řádku. V takových případech podle předpokladu b) označíme prostřední z nejtmaších bodů, viz obr. 5.4.



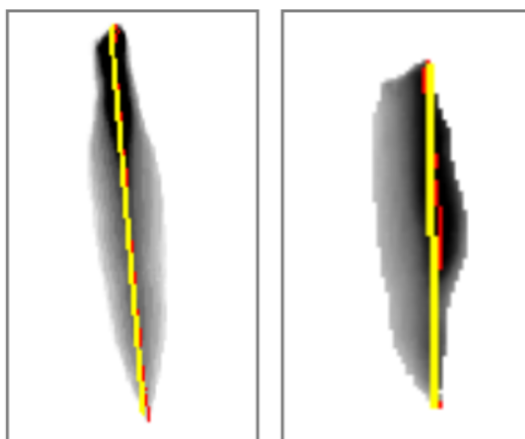
Obrázek 5.4: Vizualizované nejčtější body hlasivkové štěrbiny detekované po řádcích pro symetrické a nesymetrické chování hlasivek.

Na dolním a horním okraji hlasivek může docházet k již dříve zmíněným problémům se zakrytím části štěrbiny či jejímu ořezu. Proto dochází na okrajích štěrbiny k nepřesnostem (obr. 5.5). Z toho důvodu je vhodné k určení osy použít pouze body ve střední části hlasivek.



Obrázek 5.5: Ukázka problému určení osy při zakrytí hlasivky u zadní komisury.

Takto získané body určují podle výše stanovených předpokladů odhad bodů osy hlasivek. Ve většině případů již tvoří přímku nebo ji kopírují v těsné blízkosti. V některých případech poškození hlasivek, např. při rozdělení štěrbiny na více částí, může docházet k nepřesnostem. Proto jsme pro upřesnění bodů přímky, která reprezentuje odhad osy hlasivek, použili metodu lineární regrese. Získanou přímku omezíme pouze na oblast štěrbiny a chápeme ji jako odhad osy symetrie hlasivkové štěrbiny (obr. 5.6).



Obrázek 5.6: Body odhadu hlavní osy hlasivek získané aplikací metody lineární regrese.

5.1 Plovoucí osa

Při pořizování záznamu pohybu hlasivek vysokorychlostní kamerou není neobvyklé, že dochází vlivem vzájemného pohybu hrtanu a kamery ke změně polohy hlasivek v sekvenci snímků. V tomto případě je pak porušen předpoklad c) o stejnolehlosti jednotlivých snímků. Metoda je proto ještě modifikována pro stanovení odhadu tzv. plovoucí osy. Změna polohy hlasivek není vlivem setrvačné hmoty kamery a vysoké rychlosti snímání v jednotlivých po sobě jdoucích snímcích významná. Při rychlosti

snímání 4000 snímků/s a základní hlasivkové frekvenci $F_o = 200$ Hz, připadá na jednu periodu otevření/uzavření hlasivek 20 snímků, tj. doba jedné periody $T_{perioda} = 5$ ms. Pokud budeme uvažovat délku sekvence více než 100 snímků, bude doba sekvence snímků $T_{sekvence} = 25$ ms a může nastat posunutí o několik obrazových bodů. V takových případech není vhodné sčítat všechny snímky v sekvenci, ale jen omezený počet snímků před a po jednom kmitu hlasivek, ve kterém chceme osu detekovat. Počet snímků musí být větší, než je počet odpovídající jednomu kmitu hlasivek, aby se ve výběru vyskytoval stav maximálního uzavření.

5.2 Hodnocení výsledků detekce osy symetrie

Výsledkem aplikace uvedené metody je přímka, která je odhadem osy hlasivkové štěrbiny a zároveň osy předpokládané symetrie. Podle této osy je možné symetrii porovnávat a pozorovat rozdíly vlastností u obou hlasivkových řas.

Metoda byla aplikována na vzorku 233 vybraných videozáznamů hlasivek (z korpusu dat viz kap. 7.5) s různým stupněm poškození, u kterých byla úspěšně detekována oblast hlasivkové štěrbiny. Osa byla úspěšně detekována u všech vzorků, kde bylo také možné správnost osy potvrdit vizuálně. U hlasivek s velkým stupněm poškození, kde již téměř nedocházelo k pohybu nebo byly hlasivkové řasy příliš zdeformované, nelze tuto metodu, a pravděpodobně žádnou automatickou metodu, použít.

Pro ohodnocení kvality (přesnosti) odhadu hlasivkové osy byla použita metoda testování shody dvou přímk AxisConformity (AC), tj. v našem případě odhadované hlasivkové osy a anatomické osy určené expertem ORL. Pro parametr shody dvou přímk AC platí vztahy 5.1 a 5.2,

$$AC = \frac{k}{k + |d_1| + |d_2| + p \cdot (d_1 - d_2)^2} \quad AC \in (0, 1) \quad (5.1)$$

$$d_1 = \frac{x_1 - x_{orl1}}{h} \quad d_2 = \frac{x_2 - x_{orl2}}{h} \quad (5.2)$$

kde k = koeficient tolerance ($k = 0,2$),

p = koeficient penalizace úhlu ($p = 3$),

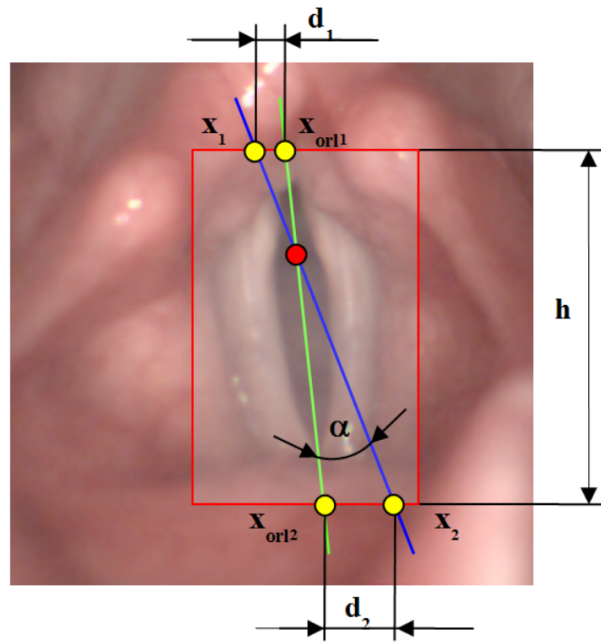
d_i = normovaná vzdálenost bodů,

h = výška výřezu,

x_i = x-souřadnice bodu odhadované osy,

x_{orli} = x-souřadnice bodu anatomické osy podle experta ORL.

Význam AC v závislosti na vzájemném natočení a posunutí obou os znázorňuje obrázek 5.7.

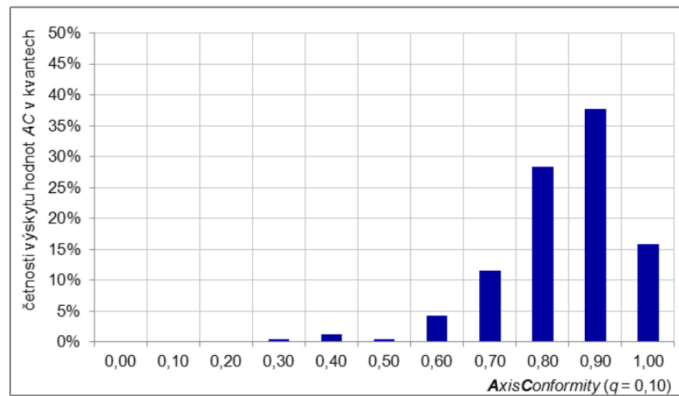


Obrázek 5.7: Schéma metody *AxisConformity* pro ohodnocení shody odhadu anatomické osy (x_1, x_2) a osy stanovené expertem *ORL* (x_{orl1}, x_{orl2}).

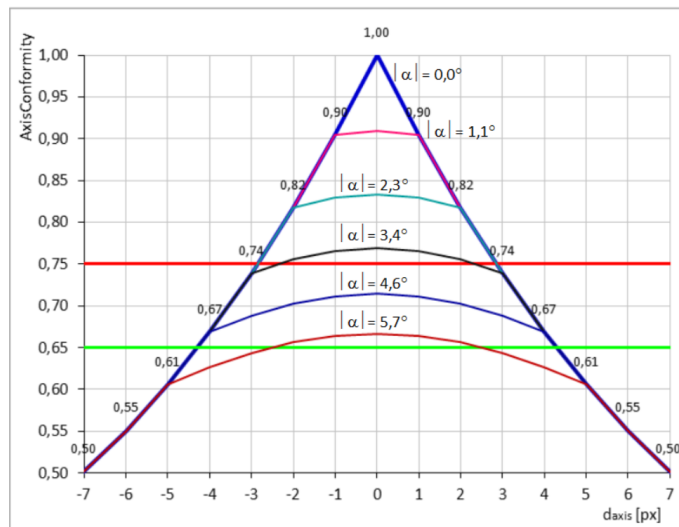
Pro hodnocení úspěšnosti automatické detekce anatomické osy hlasivek byl interval hodnot $AC \in (0, 1)$ rozdělen na 10 disjunktních podintervalů, kde kvantovací koeficient $q = 0,10$. Výsledek přiřazení jednotlivých videozáznamů podle hodnot AC do jednotlivých podintervalů, tj. četností výskytu hodnot AC v testovaném vzorku videozáznamů, je prezentován v tabulce, viz tab. 5.1.

U testovaných 233 videozáznamů se vyskytuje 81,97 % hodnot AC v podintervalech označených $\{0,8; 0,9; 1,0\}$, reálně se jedná o hodnoty $AC \in (0,75; 1,00)$, tab. 5.1. To v případě simulací hodnot AC ($h = 100\text{px}$, $d_{axis} = 0\text{px}$) představuje chybu odhadu v natočení osy $|\alpha| < 3,4^\circ$ a pro simulaci AC ($h = 100\text{px}$, $|\alpha| = 0^\circ$) je chyba odhadu polohy $|d_{axis}| < 3\text{px}$. Zahrneme-li do hodnocení také podinterval $0,7$, tj. reálně $AC \in (0,65; 1,00)$, viz tab. 5.1, obr. 5.8, pak do tohoto intervalu hodnot AC patří 93,56 % testovaných os. Pro simulaci hodnot AC ($h = 100\text{px}$, $d_{axis} = 0\text{px}$) se jedná o chybu odhadu natočení $|\alpha| < 5,7^\circ$ a pro simulaci AC ($h = 100\text{px}$, $|\alpha| = 0^\circ$) je chyba odhadu polohy $|d_{axis}| < 5\text{px}$.

Výsledek simulace hodnot AC pro kombinovanou změnu polohy a natočení odhadu anatomické osy současně je na obr. 5.9.



Obrázek 5.8: Graf četnosti výskytu hodnot AxisConformity.



Obrázek 5.9: Graf detailu vývoje AxisConformity podle úhlu vzájemného natočení os $\alpha \in \langle 45^\circ; +45^\circ \rangle$, posunu os $d_{axis} \in \langle -0,5h; +0,5h \rangle$ a pro parametry výpočtu hodnot AC: $h = 100px$, $k = 0,2$, $p = 3$.

kvantum	četnost	%	AC > 0,65	AC > 0,75
0,00	0	0,00		
0,10	0	0,00		
0,20	0	0,00		
0,30	1	0,43		
0,40	3	1,29		
0,50	1	0,43		
0,60	10	4,29		
0,70	27	11,59	93,56 %	81,97 %
0,80	66	28,33		
0,90	37	15,88		
celkem	233	100		

Tabulka 5.1: Tabulka četností výskytu hodnot AxisConformity v jednotlivých podintervalech AC ($q = 0,10$) pro soubor testovaných videosekvencí.

6 Parametry hlasivkové štěrbiny

V této části práce budou představeny některé parametry hlasivkové štěrbiny, které jsou běžně používány v medicínských systémech vysokorychlostní laryngoskopie, parametry, které jsou publikovány v řadě vědeckých článků, a návrh některých nových parametrů, které jsou testovány ve spolupráci s ORL klinikou FN Plzeň, viz [21], [43].

6.1 Základní parametry a symetrie

Do základních parametrů počítáme dnes již standardní parametry, jejichž měření a výpočet je součástí systémů vysokorychlostní laryngoskopie. Do této skupiny patří parametry velikosti plochy hlasivkové štěrbiny, délky jejího obvodu a výšky a šířky glottis, viz přehled v tabulce 6.1.

parametr	označení
Velikost plochy hlasivkové štěrbiny	A
Poměr max. a min. plochy během kmitu hlasivek	A_r
Délka obvodu hlasivkové štěrbiny	L
Výška hlasivky (délka osy symetrie glottis)	D_{axis}
Šířka hlasivky (délka normály glottis)	D_{norm}

Tabulka 6.1: Přehled základních parametrů pro jednotlivé snímky ve videosekvenci a jejich označení, viz [21].

6.2 Parametry doplněné

Aplikace pro hodnocení snímků z vysokorychlostní laryngoskopie, viz [21] a v [43], je doplněna o množinu testovaných parametrů, tabulka 6.2. Tato množina parametrů je testována na vzájemnou korelaci a korelaci s dalšími parametry některých vyšetření, viz kap. 1.3.

parametr	označení
Velikost levé části plochy hl. štěrbiny (od osy symetrie)	A_{left}
Velikost pravé části plochy hl. štěrbiny (od osy symetrie)	A_{right}
Vychýlení plochy jako rozdíl levé a pravé části plochy	A_{delta}
Délka levé části obvodu hl. štěrbiny (od osy symetrie)	L_{left}
Délka pravé části obvodu hl. štěrbiny (od osy symetrie)	L_{right}
Rozdíl pravé a levé části obvodu	L_{delta}
Členitost obvodu	S
Členitost levé části obvodu	S_{left}
Členitost pravé části obvodu	S_{right}
Podlouhlost	O
Výška elipsy, která aproximuje obvod hl. štěrbiny	H_e
Délka elipsy, která aproximuje obvod hl. štěrbiny	W_e
Podlouhlost elipsy	O_e

Tabulka 6.2: Přehled doplněných parametrů testovaných pro jednotlivé snímky ve videosekvenci a jejich označení, viz [21], [43].

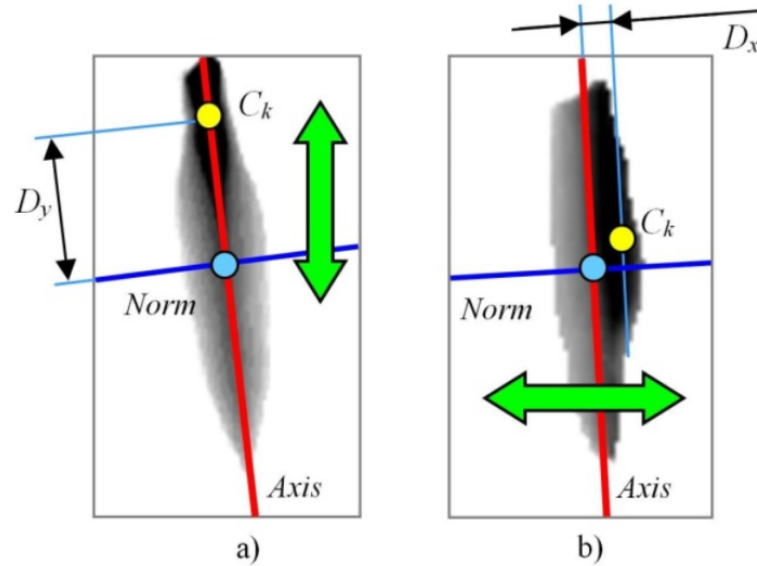
Většina z uvedených parametrů se získává měřením přímo z detekované hlasivkové štěrbiny, resp. hlasivek, ve snímku nebo v sekvenci snímků, některé parametry jsou odvozeny výpočtem z hodnot parametrů naměřených.

6.3 Těžiště hlasivkové štěrbiny

Do souboru tzv. doplněných parametrů byly přidány parametry těžiště hlasivkové štěrbiny, které mají explicitně detekovat symetrické/nesymetrické chování hlasivek. Současně tvoří množinu doplňujících (upřesňujících) parametrů pro rozhodování o symetrii a nesymetrii hlasivek. Prakticky užíváme celou řadu kombinací těchto parametrů, které jsou konstruovány na principu polohy těžiště plochy a polohy těžiště uzavřené obvodové křivky hlasivkové štěrbiny. Poloha těžiště je pak orientována vzhledem k ose symetrie a normále hlasivkové štěrbiny.

6.3.1 Parametry těžiště

Parametry těžiště C_k budeme chápat jako vzdálenost těžiště štěrbiny S_k ve snímku k od hlavní osy (osy symetrie glottis) hlasivek D_x a její normály D_y . Svým charakterem se jedná o parametry jak statické, tak dynamické. Význam sledování dynamiky vývoje těžiště je patrný ze schematického zobrazení na obr. 6.1 (jedná se o schematický popis vývoje polohy těžiště bez ohledu na typ těžiště a způsob výpočtu polohy).



Obrázek 6.1: Schéma a popis pohybu těžiště během jednoho kmitu hlasivek;
a) symetrická hlasivková štěrбина, kdy je dominantní vývoj vzdálenosti D_y , tj. těžiště vzhledem k normále;
b) nesymetrická hlasivková štěrбина, zesiluje vývoj vzdálenosti D_x , tj. těžiště vzhledem k hlavní ose.

V praxi pak používáme dva typy těžiště štěrbiny S_k :

- Těžiště plochy CA_k , kdy pro výpočet polohy těžiště platí vztahy 6.1 a 6.2:

$$CA_k = CA_k(x_c^{(S)}, y_c^{(S)}) \quad (6.1)$$

$$x_c^{(S)} = \frac{1}{A_k} \sum_{x_S \in S_k} \sum_{y_S \in S_k} x_S, y_c^{(S)} = \frac{1}{A_k} \sum_{x_S \in S_k} \sum_{y_S \in S_k} y_S \quad (6.2)$$

x_s a y_s jsou hodnoty jednotlivých souřadnic obrazových bodů, které tvoří segmentovanou oblast štěrbinu S_k detekovanou metodami (viz kap. 3), A_k je velikost plochy štěrbinu S_k .

- Těžiště CH_k hraniční čáry (vnitřní hranice) H_k segmentované štěrbinu S_k ; pro výpočet polohy těžiště platí následující vztahy 6.3 a 6.4:

$$CH_k = CH_k(x_c^{(H)}, y_c^{(H)}) \quad (6.3)$$

$$x_c^{(H)} = \frac{1}{A_k} \sum_{x_H \in H_k} \sum_{y_H \in H_k} x_H, y_c^{(H)} = \frac{1}{A_k} \sum_{x_H \in H_k} \sum_{y_H \in H_k} y_H \quad (6.4)$$

Ve vztahu 6.4 x_H a y_H jsou hodnoty jednotlivých souřadnic obrazových bodů, které tvoří vnitřní hranici H_k oblasti štěrbinu S_k , L_k je délka hranice H_k .

Na základě polohy obou typů těžiště byly definovány další parametry, kterými si klademe za cíl přispět k popisu symetrie hlasivkové štěrbiny a dynamiky hlasivkové štěrbiny, jako jsou vzdálenosti těžiště od osy (D_x) a od normály (D_y). Pro výpočet těchto parametrů platí vztahy 6.5 a 6.6, kde a , b a c jsou koeficienty osy resp. normály, x_c a y_c jsou souřadnice těžiště.

$$D_x = \frac{ax_c + by_c + c}{\sqrt{a^2 + b^2}} \quad (6.5)$$

$$D_y = \frac{a_{norm}x_c + b_{norm}y_c + c_{norm}}{\sqrt{a_{norm}^2 + b_{norm}^2}} \quad (6.6)$$

Celkový přehled parametrů odvozených od polohy těžiště je v tabulce 6.3.

parametr	označení
Pozice těžiště plochy	CA
Vychýlení těžiště plochy od osy	$D_x(S)$
Vychýlení těžiště plochy od normály	$D_y(S)$
Pozice těžiště hranice	CH_k
Vychýlení těžiště hranice od osy	$D_x(H)$
Vychýlení těžiště hranice od normály	$D_y(H)$
Rozdíl polohy těžiště hranice a plochy v ose x	$D_x(diff)$
Rozdíl polohy těžiště hranice a plochy v ose y	$D_y(diff)$

Tabulka 6.3: Přehled parametrů těžiště pro jednotlivé snímky ve videosekvenci a jejich označení, viz [43].

Na základě získaných poznatků o vývoji plochy a obvodu štěrbiny je také možné odvodit parametry symetrie hlasivek vypočítané podle vývoje plochy a detekované vnitřní hranice hlasivkové štěrbiny během fonace. Protože symetrie hlasivek se zdá být důležitým faktorem, jako příklad uvádím grafické zobrazení a popis některých kazuistik. V jednotlivých případech je porovnáván vývoj parametrů plochy a obvodu hlasivkové štěrbiny během videosekvence s vývojem polohy těžiště plochy a hraniční čáry štěrbiny. Kazuistiky jsou vybírány pro případy tzv. symetrických (zdravých) a nesymetrických hlasivek. Výběr případů byl proveden z celkového počtu 400 zpracovaných videozáznamů z vyšetření HSV na ORL klinice FN Plzeň jako typické představitele případů, která se v množině záznamů vyskytují.

6.3.2 Vybrané kazuistiky

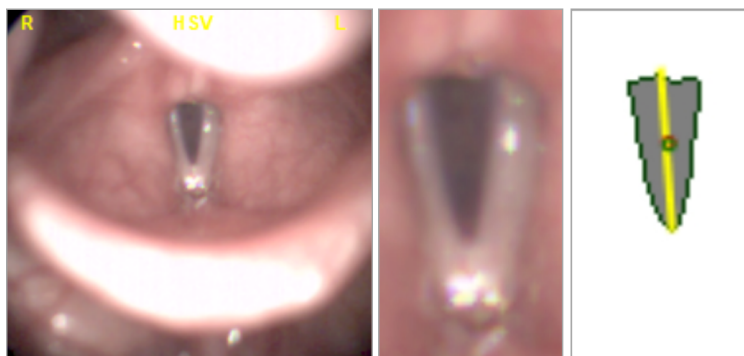
Zdravá symetrická hlasivka

Popis:

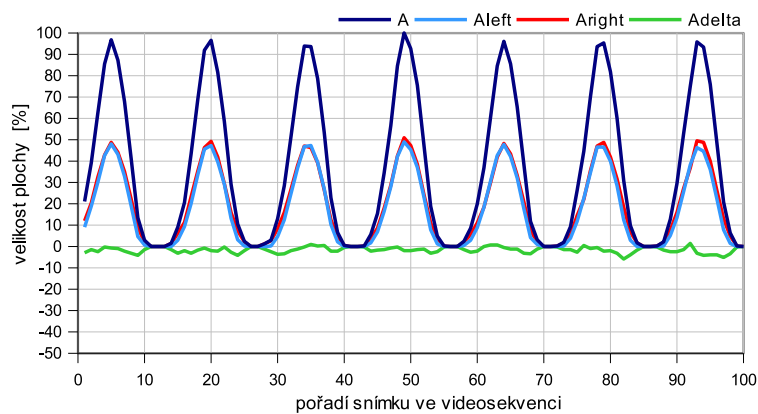
- Zdravá symetrická hlasivka, bez omezení pohybu (obr. 6.2).
- Žena 29 let.
- MIC-HSV, fonace „i“, $SPL_{min} = 92$ dB, $SPL_{max} = 92$ dB, $F_0 = 276$ Hz.
- analýza plochy A_k a délky hranice L_k ukazuje na symetrii, viz parametry A_{delta} a L_{delta} (obr. 6.3 a 6.6).
- Vývoj polohy těžiště plochy a obvodu štěrbiny ve směru normály, tj. $D_x(S)$ a $D_x(H)$ potvrzuje dynamiku symetrické hlasivky, tj. pohyb těžiště ve směru osy symetrie $D_y(S)$ a $D_y(H)$ je doplněn nevýznamným pohybem těžiště ve směru normály $D_x(S)$ a $D_x(H)$ (obr. 6.4 a 6.7).
- Vzájemná poloha těžiště plochy D_S a obvodu D_H štěrbiny během fonace je téměř totožná (obr. 6.5 a 6.8).

Závěr:

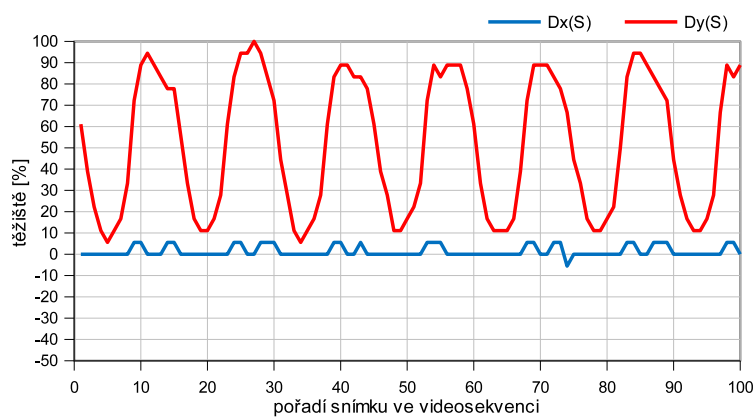
analýza vývoje plochy A_k , délky hranice L_k , polohy těžiště plochy D_s a obvodu H_h štěrbiny během fonace potvrzuje **symetrii hlasivek**.



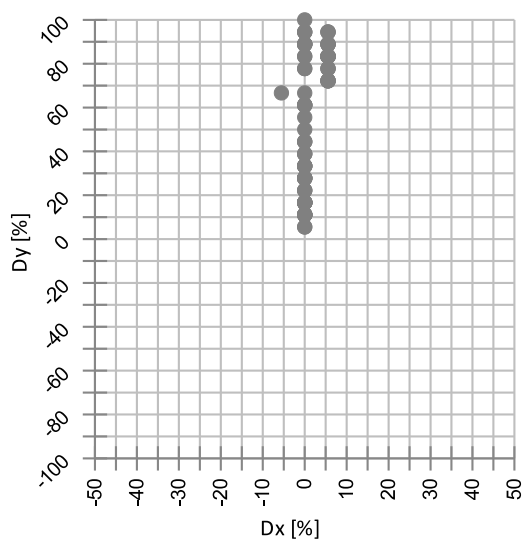
Obrázek 6.2: Příklad zdravé symetrické hlasivky;
 vlevo: snímek z vysokorychlostní laryngoskopie (HSV);
 uprostřed: ROI – tzv. oblast zájmu;
 vpravo: segmentovaná hlasivková štěrбина s osou symetrie, těžištěm a vnitřní hranicí.



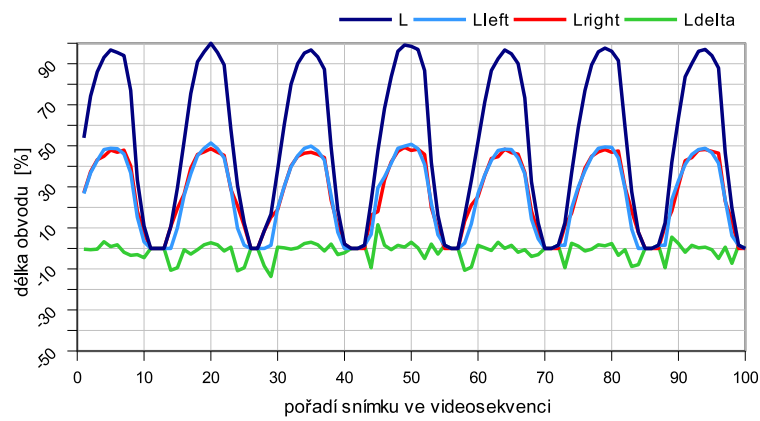
Obrázek 6.3: Vývoj velikosti plochy A hlasivkové štěrbiny.



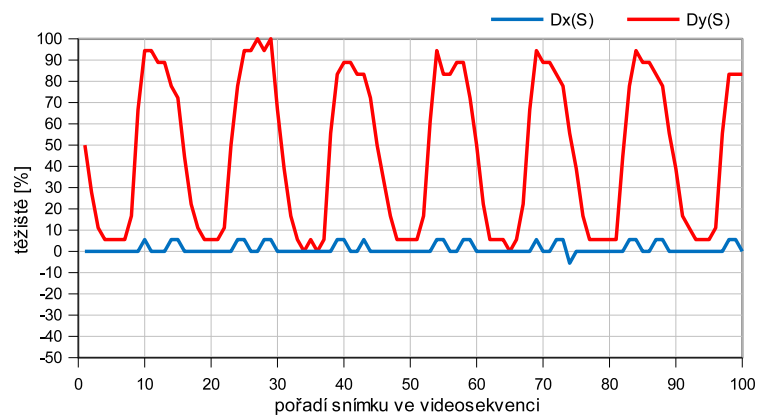
Obrázek 6.4: Poloha těžiště ve směru $D_x(S)$ a $D_y(S)$ podle pořadí snímků.



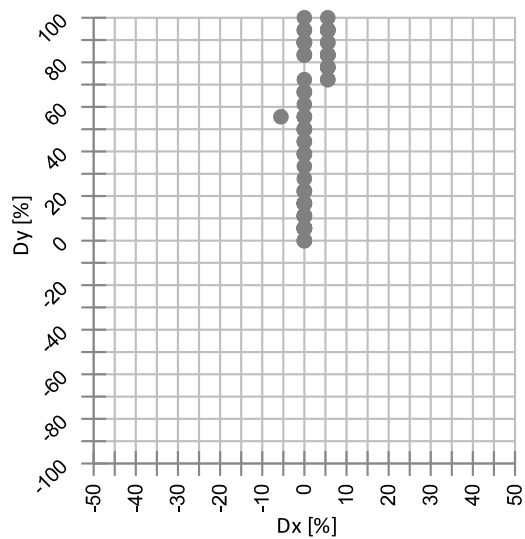
Obrázek 6.5: Vývoj těžiště plochy hlasivkové štěrbiny - normovaná trajektorie pohybu těžiště CA.



Obrázek 6.6: Vývoj délky L vnitřní hranice hlasivkové štěrbiny.



Obrázek 6.7: Poloha těžiště ve směru $D_x(H)$ a $D_y(H)$ podle pořadí snímků.



Obrázek 6.8: Vývoj polohy těžiště vnitřní hranice hlasivkové štěrbiny - normovaná trajektorie pohybu těžiště CH .

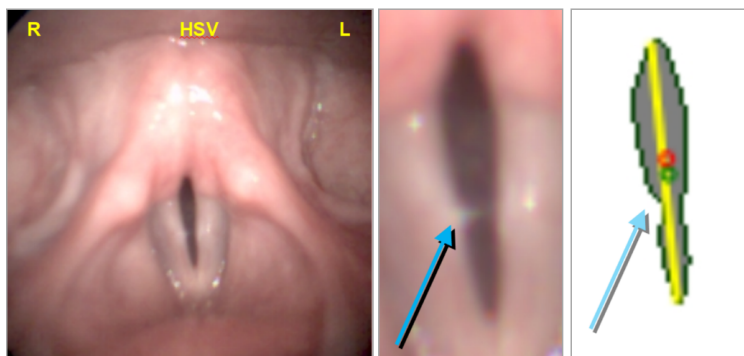
Zdravá symetrická hlasivka s příznaky asymetrie

Popis:

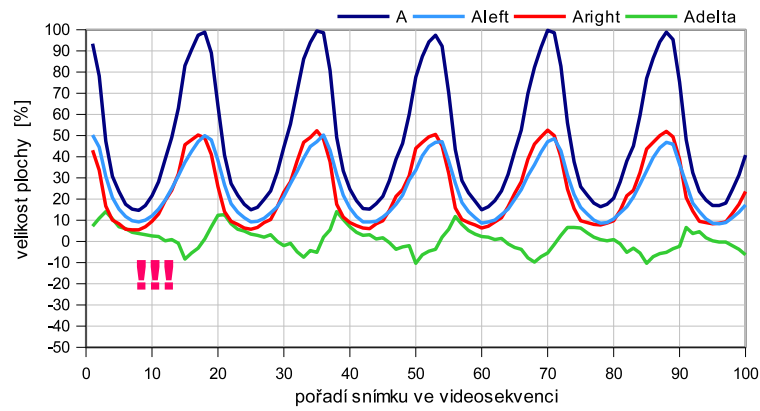
- Zdravá symetrická hlasivka, bez omezení pohybu (obr. 6.9).
- Žena 19 let.
- MIC-HSV, fonace „i“, $SPL_{min} = 77$ dB, $SPL_{max} = 77$ dB, $F_0 = 226$ Hz.
- analýza plochy A_k a délky hranice L_k ukazuje na mírnou nesymetrii, viz parametry A_{delta} a L_{delta} (obr. 6.10 a 6.13).
- Vývoj polohy těžiště plochy a obvodu štěrbiny ve směru normály, tj. $D_x(S)$ a $D_x(H)$, naopak potvrzuje dynamiku symetrické hlasivky, tj. pohyb těžiště ve směru osy symetrie $D_y(S)$ a $D_y(H)$ je doplněn nevýznamným pohybem těžiště ve směru normály $D_x(S)$ a $D_x(H)$ (obr. 6.11 a 6.14).
- Vzájemná poloha těžiště plochy D_S a obvodu D_H štěrbiny během fonace je téměř totožná (obr. 6.12 a 6.15).

Závěr:

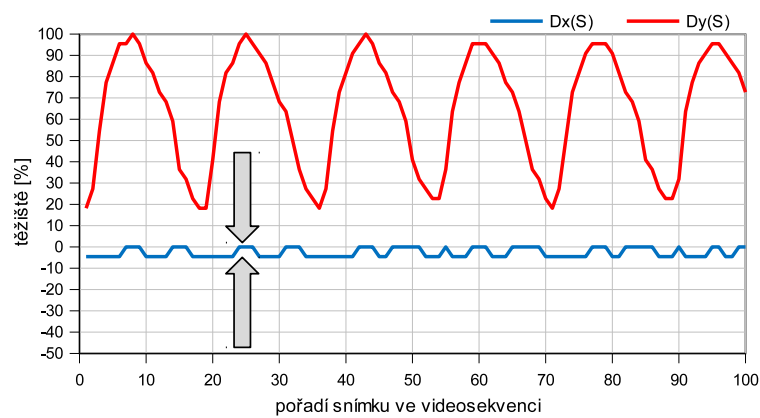
analýza vývoje plochy A_k , délky hranice L_k ukazuje na mírnou nesymetrii, která je způsobená hlenem na pravé hlasivce; vývoj polohy těžiště plochy D_S a obvodu D_H štěrbiny během fonace však ukazuje na **symetrii hlasivek**.



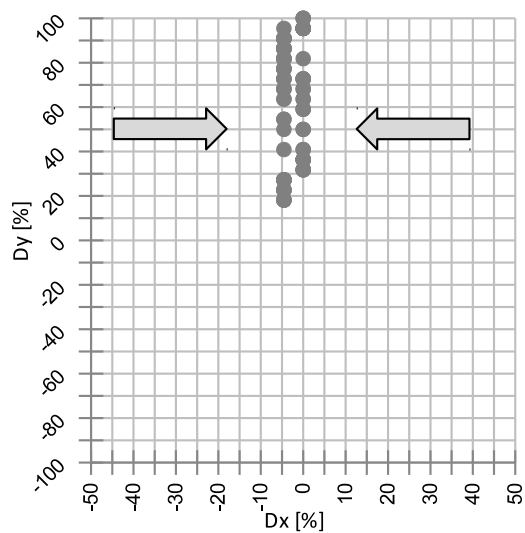
Obrázek 6.9: Příklad zdravé symetrické hlasivky s výskytem hlenu na pravé hlasivce; vlevo: snímek z vysokorychlostní laryngoskopie (HSV); uprostřed: ROI – tzv. oblast zájmu; vpravo: segmentovaná hlasivková štěrbin s osou symetrie, těžištěm a vnitřní hranicí.



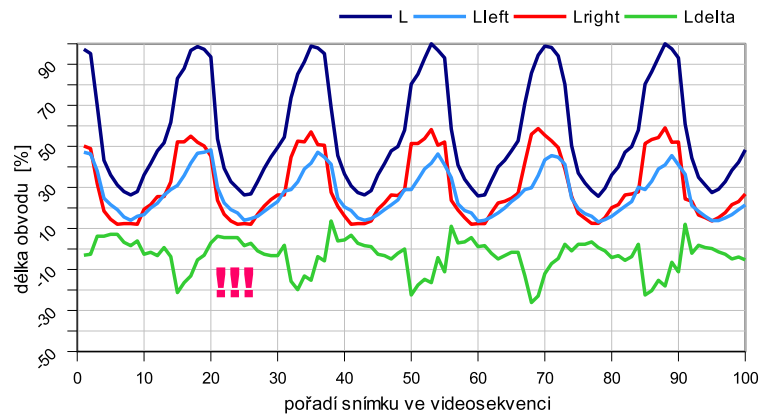
Obrázek 6.10: Vývoj velikosti plochy A hlasivkové štěrby.



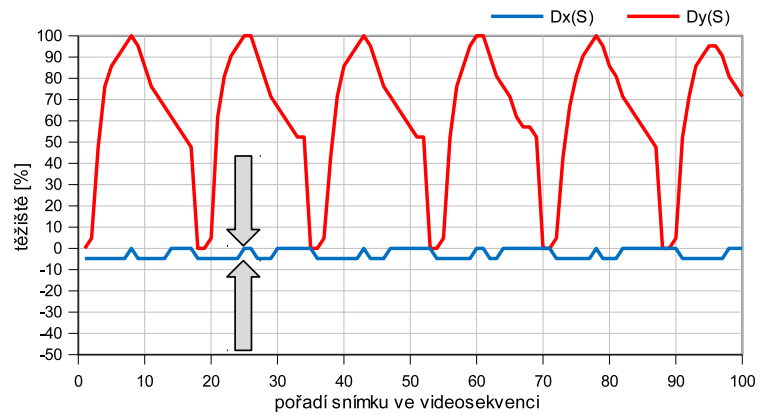
Obrázek 6.11: Poloha těžiště ve směru $D_x(S)$ a $D_y(S)$ podle pořadí snímků.



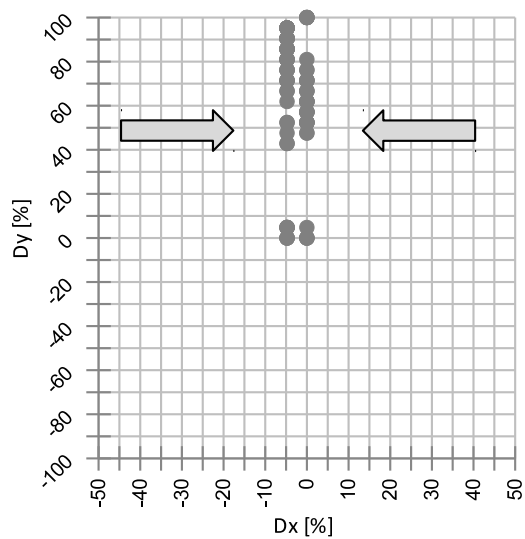
Obrázek 6.12: Vývoj těžiště plochy hlasivkové štěrby - normovaná trajektorie pohybu těžiště CA.



Obrázek 6.13: Vývoj délky L vnitřní hranice hlasivkové štěrby.



Obrázek 6.14: Poloha těžiště ve směru $D_x(H)$ a $D_y(H)$ podle pořadí snímků.



Obrázek 6.15: Vývoj polohy těžiště vnitřní hranice hlasivkové štěrby - normovaná trajektorie pohybu těžiště CH .

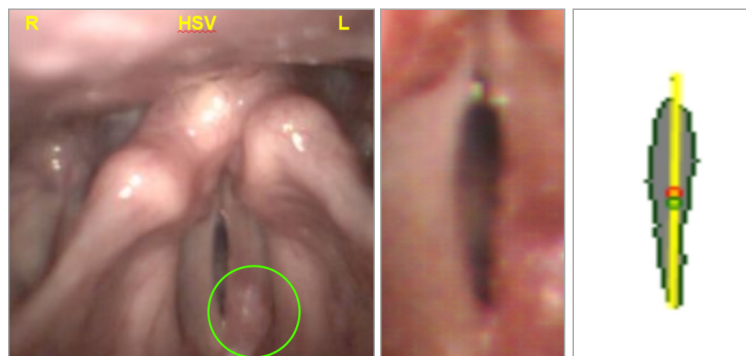
Nesymetrická hlasivka - karcinom

Popis:

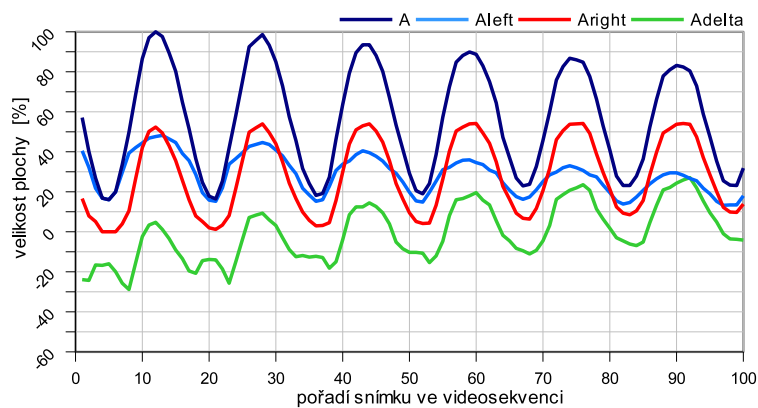
- Nesymetrická hlasivka, omezení pohybu hlasivky vlevo (karcinom) (obr. 6.16).
- Muž 82 let.
- MIC-HSV, fonace „i“, $SPL_{min} = 77$ dB, $SPL_{max} = 81$ dB, $F_0 = 298$ Hz.
- Analýza plochy A_k a délky hranice L_k ukazuje na nesymetrii, viz parametry A_{delta} a L_{delta} (obr. 6.17 a 6.20).
- Vývoj polohy těžiště plochy a obvodu štěrbiny ve směru normály, tj. $D_x(S)$ a $D_x(H)$ potvrzuje dynamiku nesymetrické hlasivky, tj. pohyb těžiště ve směru osy symetrie $D_y(S)$ a $D_y(H)$ je doplněn významným pohybem těžiště ve směru normály $D_x(S)$ a $D_x(H)$ (obr. 6.19 a 6.22).
- Zvětšuje se rozdíl ve vzájemné poloze těžiště plochy D_S a obvodu D_H štěrbiny během fonace (obr. 6.19 a 6.22).

Závěr:

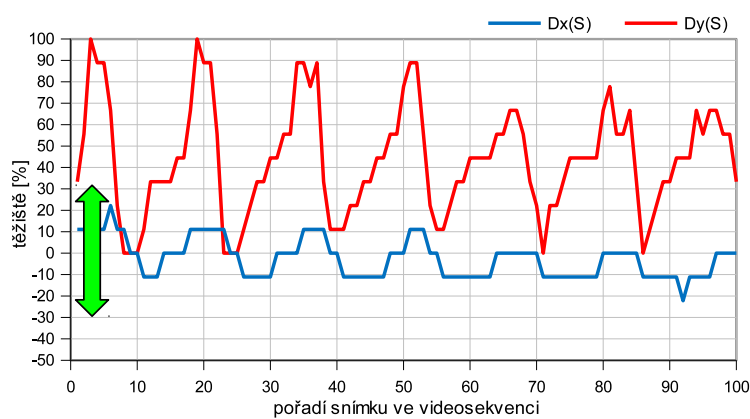
analýza vývoje plochy A_k a délky hranice L_k ukazuje na nesymetrii, která je způsobená hmotou karcinomu na levé hlasivce; vývoj polohy těžiště plochy D_S a obvodu D_H štěrbiny během fonace potvrzuje **nesymetrii hlasivek**.



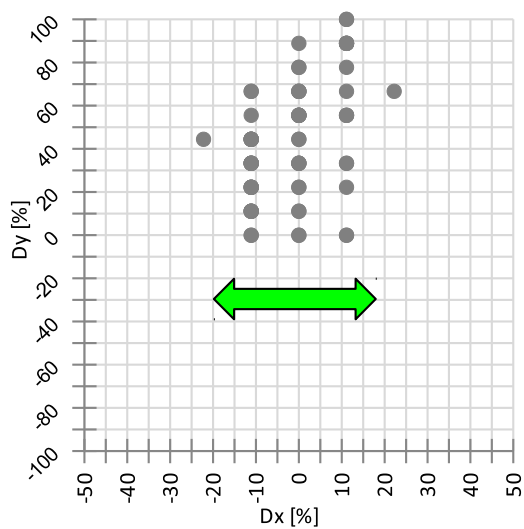
Obrázek 6.16: Příklad nesymetrické hlasivky s karcinomem na levé hlasivce; vlevo: snímek z vysokorychlostní laryngoskopie (HSV); uprostřed: ROI – tzv. oblast zájmu; vpravo: segmentovaná hlasivková štěrbiná s osou symetrie, těžištěm a vnitřní hranicí.



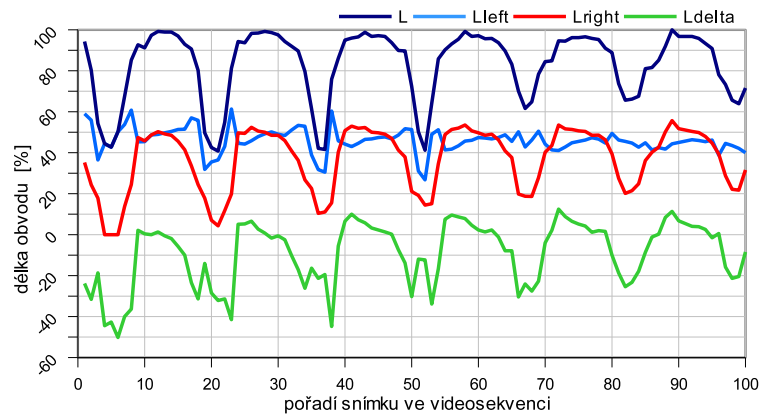
Obrázek 6.17: Vývoj velikosti plochy A hlasivkové štěrby.



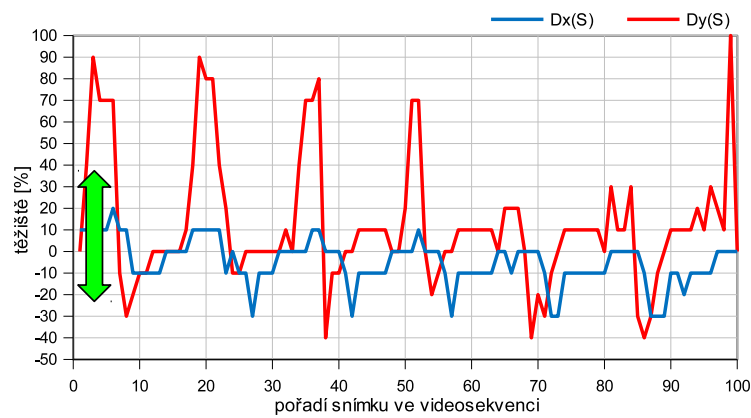
Obrázek 6.18: Poloha těžiště ve směru $D_x(S)$ a $D_y(S)$ podle pořadí snímků.



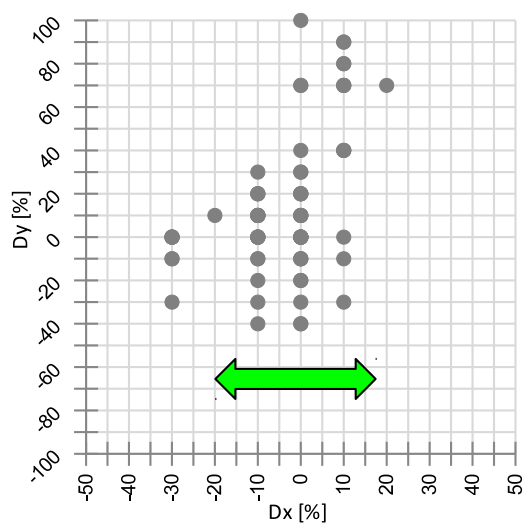
Obrázek 6.19: Vývoj těžiště plochy hlasivkové štěrby - normovaná trajektorie pohybu těžiště CA.



Obrázek 6.20: Vývoj délky L vnitřní hranice hlasivkové štěrby.



Obrázek 6.21: Poloha těžiště ve směru $D_x(H)$ a $D_y(H)$ podle pořadí snímků.



Obrázek 6.22: Vývoj polohy těžiště vnitřní hranice hlasivkové štěrby - normovaná trajektorie pohybu těžiště CH .

Nesymetrická hlasivka - chordektomie

Popis:

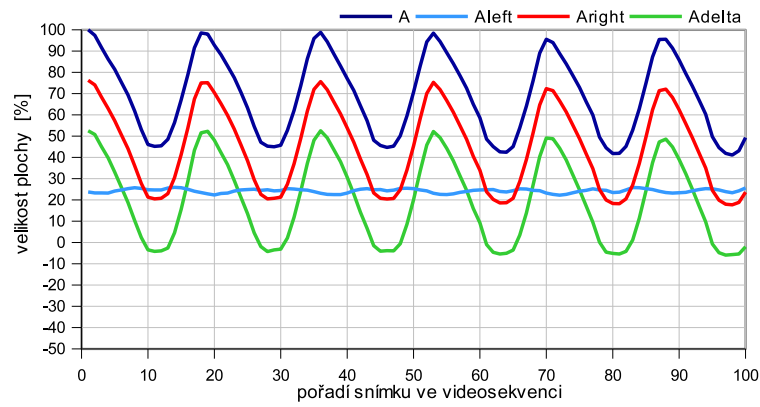
- Nesymetrická hlasivka, stav po chirurgickém vyjmutí hlasivky vlevo (chordektomie) (obr. 6.23).
- Žena 57 let.
- MIC-HSV, fonace „i“, $SPL_{min} = 74$ dB, $SPL_{max} = 84$ dB, $F_0 = 230$ Hz.
- Analýza plochy A_k a délky hranice L_k ukazuje na nesymetrii, viz parametry A_{delta} a L_{delta} , levá část hlasivky se nepohybuje (obr. 6.24 a 6.27).
- Vývoj polohy těžiště plochy a obvodu štěrbiny ve směru normály, tj. $D_x(S)$ a $D_x(H)$ potvrzuje dynamiku nesymetrické hlasivky, pohyb těžiště ve směru osy symetrie $D_y(S)$ a $D_y(H)$ je doplněn jednostranným pohybem těžiště ve směru normály $D_x(S)$ a $D_x(H)$ (obr. 6.26 a 6.29).
- Protože se levá část hlasivky nepohybuje, je rozdíl ve vzájemné poloze těžiště plochy D_S a obvodu D_H ve směru normály štěrbiny během fonace minimální (obr. 6.26 a 6.29).

Závěr:

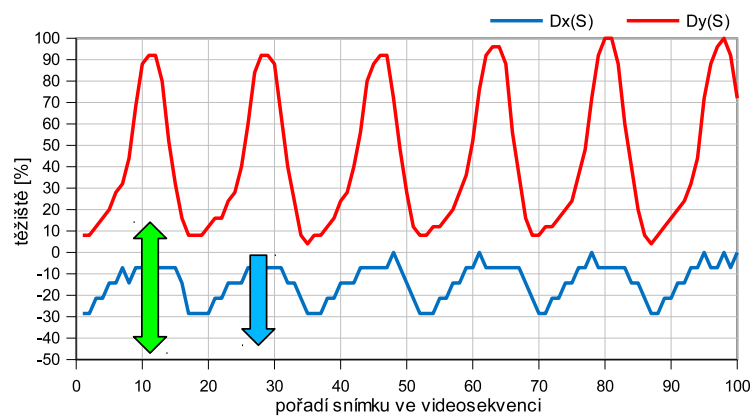
analýza vývoje plochy A_k a délky hranice L_k ukazuje na jednostrannou nesymetrii, která je způsobená chordektomií vlevo; vývoj polohy těžiště plochy D_S a obvodu D_H štěrbiny během fonace tuto **nesymetrii hlasivek** potvrzuje.



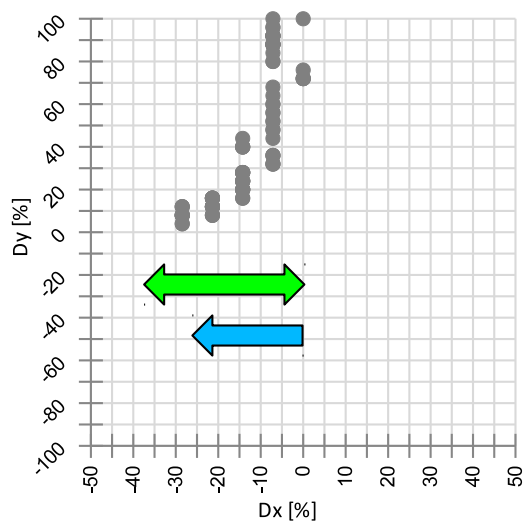
Obrázek 6.23: Příklad nesymetrické hlasivky po chordektomii vlevo; vlevo: snímek z vysokorychlostní laryngoskopie (HSV); uprostřed: ROI – tzv. oblast zájmu; vpravo: segmentovaná hlasivková štěrbina s osou symetrie, těžištěm a vnitřní hranicí.



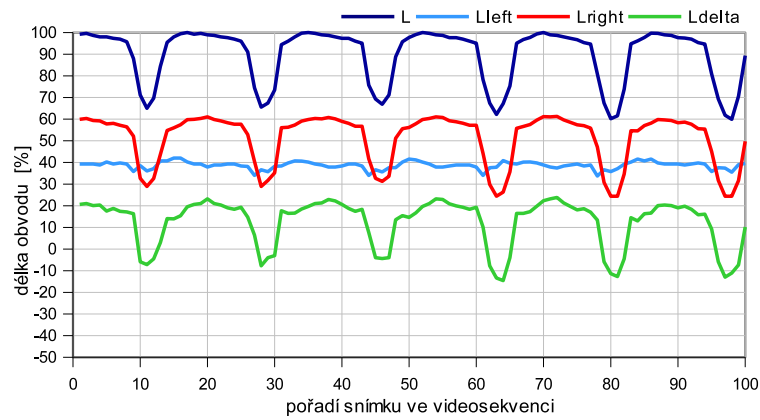
Obrázek 6.24: Vývoj velikosti plochy A hlasivkové štěrby.



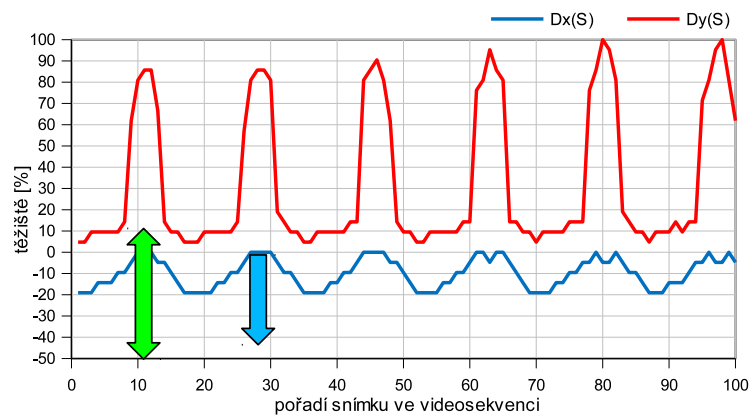
Obrázek 6.25: Poloha těžiště ve směru $D_x(S)$ a $D_y(S)$ podle pořadí snímků.



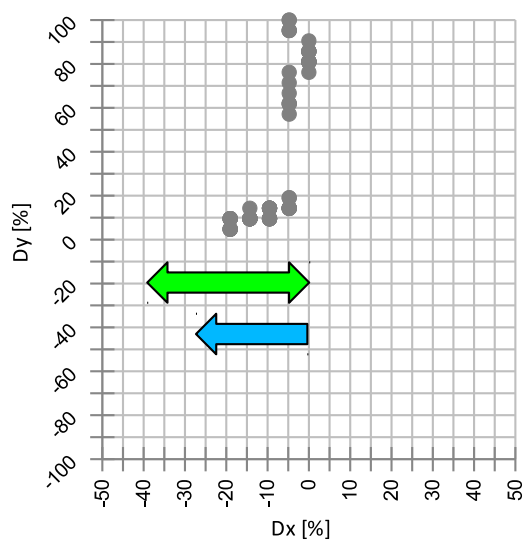
Obrázek 6.26: Vývoj těžiště plochy hlasivkové štěrby - normovaná trajektorie pohybu těžiště CA.



Obrázek 6.27: Vývoj délky L vnitřní hranice hlasivkové štěrby.



Obrázek 6.28: Poloha těžiště ve směru $D_x(H)$ a $D_y(H)$ podle pořadí snímků.



Obrázek 6.29: Vývoj polohy těžiště vnitřní hranice hlasivkové štěrby - normovaná trajektorie pohybu těžiště CH .

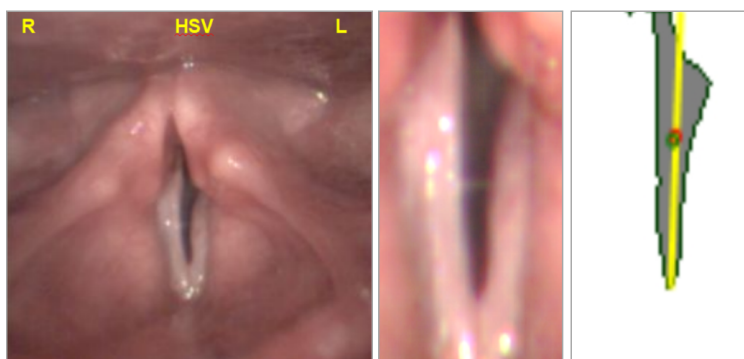
Nesymetrická hlasivka - paréza zvratného nervu

Popis:

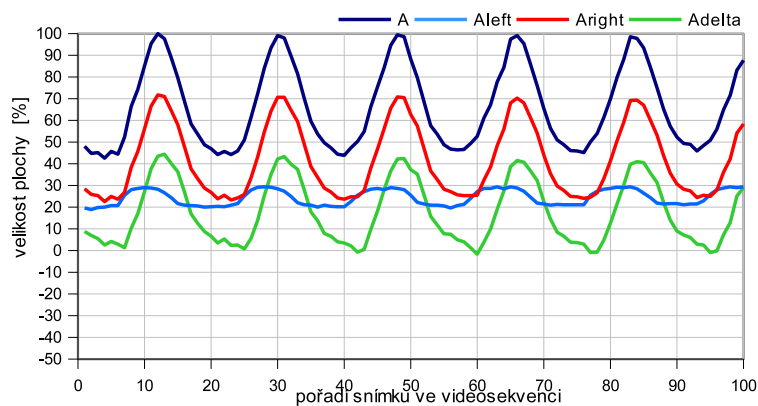
- Nesymetrická hlasivka, paréza zvratného nervu vlevo (obr. 6.30).
- Žena 29 let.
- MIC-HSV, fonace „i“, $SPL_{min} = 53$ dB, $SPL_{max} = 78$ dB, $F_0 = 240$ Hz.
- Analýza plochy A_k ukazuje na nesymetrii, viz parametr A_{delta} (obr. 6.31).
- Analýza délky hranice L_k ukazuje na nesymetrii, viz parametr L_{delta} (obr.6.34).
- vývoj polohy těžiště plochy a obvodu štěrbiny ve směru normály, tj. $D_x(S)$ a $D_x(H)$ potvrzuje nesymetrii v dynamice hlasivky, tj. těžiště ve směru normály $D_x(S)$ a $D_x(H)$ je systematicky vychýleno v jednom směru (obr. 6.32 a 6.35).
- Rozdíl ve vzájemné poloze těžiště plochy D_S a obvodu D_H štěrbiny během fonace je minimální (obr. 6.33 a 6.36).

Závěr:

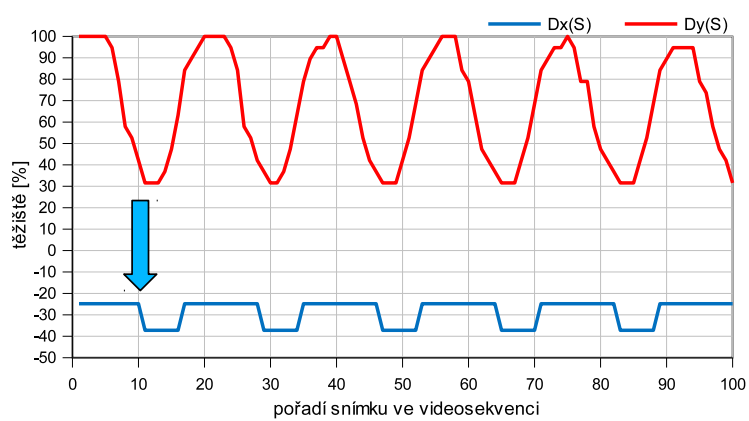
analýza vývoje plochy A_k a délky hranice L_k ukazuje na nesymetrii, která je způsobená parézou vlevo; vývoj polohy těžiště plochy D_S a obvodu D_H štěrbiny během fonace tuto **nesymetrii hlasivek** potvrzuje.



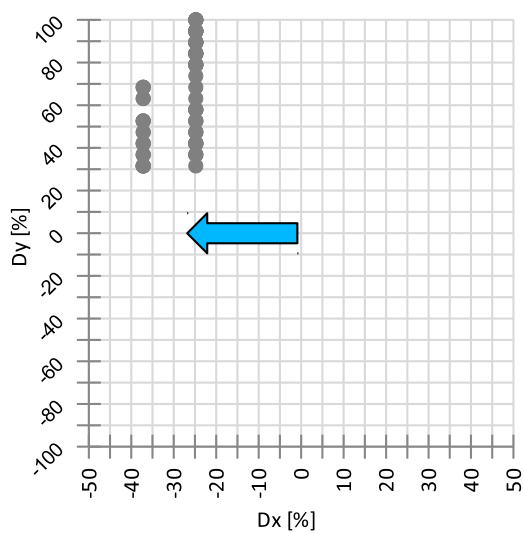
Obrázek 6.30: Příklad nesymetrické hlasivky s parézou zvratného nervu vlevo; vlevo: snímek z vysokorychlostní laryngoskopie (HSV); uprostřed: ROI – tzv. oblast zájmu; vpravo: segmentovaná hlasivková štěrbina s osou symetrie, těžištěm a vnitřní hranicí.



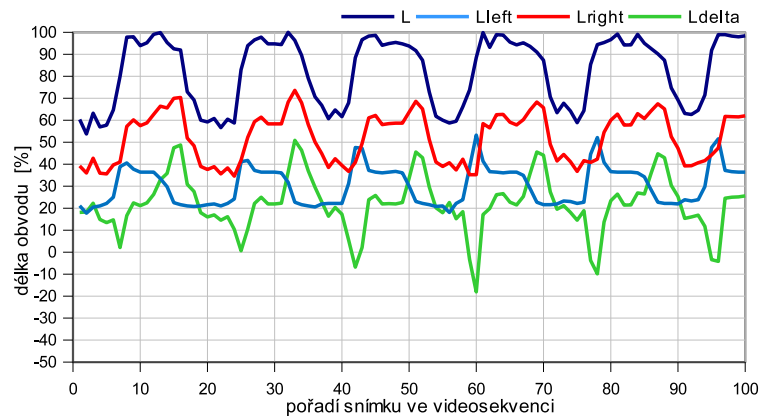
Obrázek 6.31: Vývoj velikosti plochy A hlasivkové štěrby.



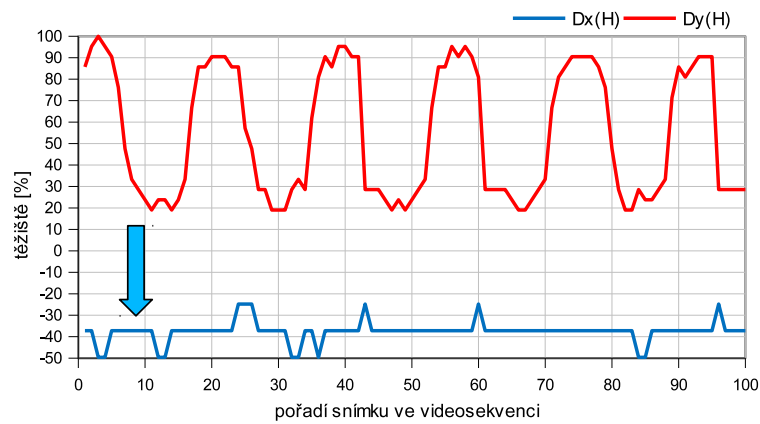
Obrázek 6.32: Poloha těžiště ve směru $D_x(S)$ a $D_y(S)$ podle pořadí snímků.



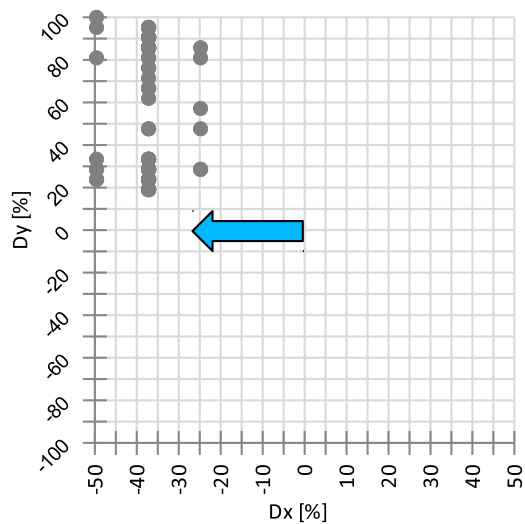
Obrázek 6.33: Vývoj těžiště plochy hlasivkové štěrby - normovaná trajektorie pohybu těžiště CA.



Obrázek 6.34: Vývoj délky L vnitřní hranice hlasivkové štěrby.



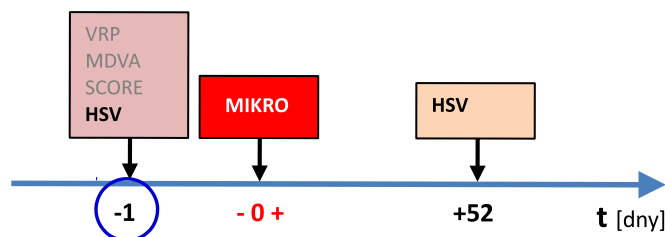
Obrázek 6.35: Poloha těžiště ve směru $D_x(H)$ a $D_y(H)$ podle pořadí snímků.



Obrázek 6.36: Vývoj polohy těžiště vnitřní hranice hlasivkové štěrby - normovaná trajektorie pohybu těžiště CH .

Nesymetrická hlasivka - uzlík

(stav před zákrokem)



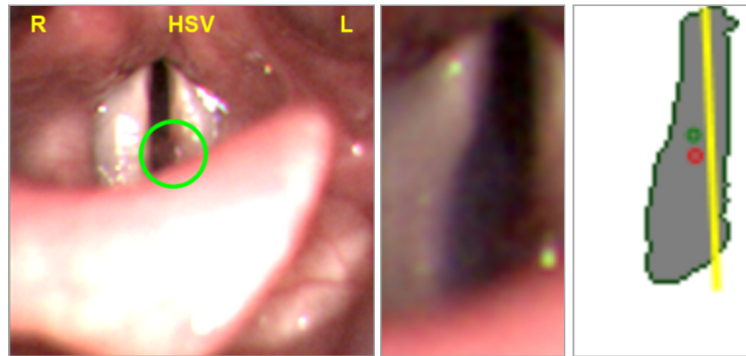
Obrázek 6.37: Časová osa vyšetření před a po mikrochirurgickém zákroku - stav před zákrokem.

Popis:

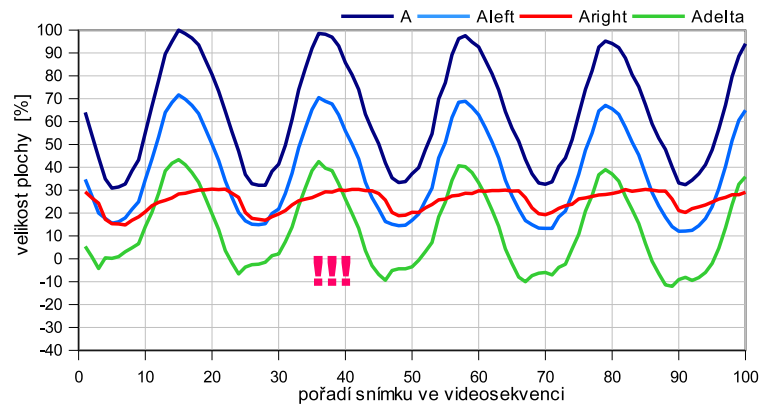
- Nesymetrická hlasivka, uzlík vpravo, stav před zákrokem (obr. 6.38).
- Žena 48 let.
- MIC-HSV, fonace „i“, $SPL_{min} = 78$ dB, $SPL_{max} = 87$ dB, $F_0 = 192$ Hz.
- Analýza plochy A_k a délky hranice L_k ukazuje na výrazné omezení pohybu pravé hlasivky a nesymetrii vpravo, viz parametry A_{delta} , A_{right} , L_{delta} a L_{right} (obr. 6.39 a 6.42).
- Poloha těžiště plochy a obvodu štěrbin ve směru normály, tj. $D_x(S)$ a $D_x(H)$ potvrzuje závěry z analýzy plochy a délky hranice štěrbin, tj. pohyb těžiště ve směru osy symetrie $D_y(S)$ a $D_y(H)$ je doplněn významným pohybem těžiště ve směru normály $D_x(S)$ a $D_x(H)$ (obr. 6.40 a 6.43).
- Došlo k vychýlení polohy těžiště D_S a D_H vpravo od osy hlasivek (obr. 6.41 a 6.44).

Závěr:

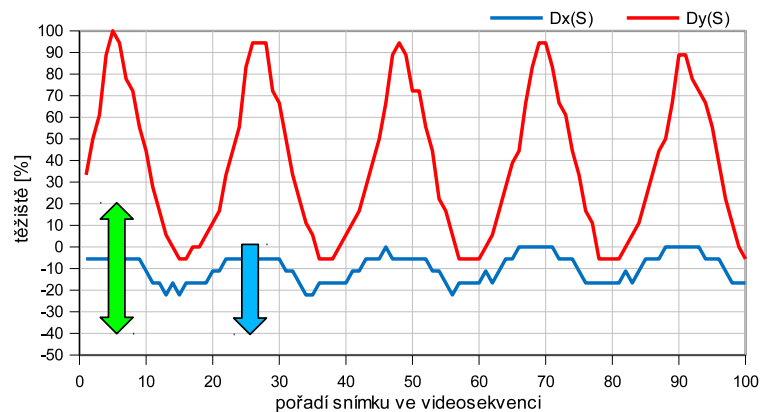
analýza vývoje plochy A_k a délky hranice L_k ukazuje na nesymetrii, která je způsobená uzlíkem vlevo; vývoj polohy těžiště plochy D_S a obvodu D_H štěrbin během fonace tuto **nesymetrii hlasivek** potvrzuje.



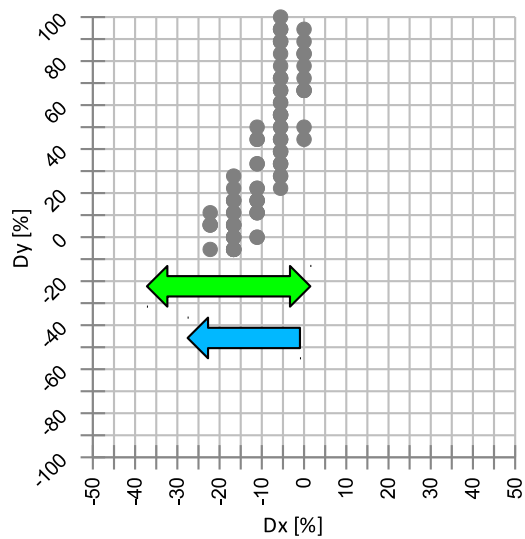
Obrázek 6.38: Příklad nesymetrické hlasivky s uzlíkem vlevo;
 vlevo: snímek z vysokorychlostní laryngoskopie (HSV);
 uprostřed: ROI – tzv. oblast zájmu;
 vpravo: segmentovaná hlasivková štěrbinina s osou symetrie, těžištěm a vnitřní hranicí.



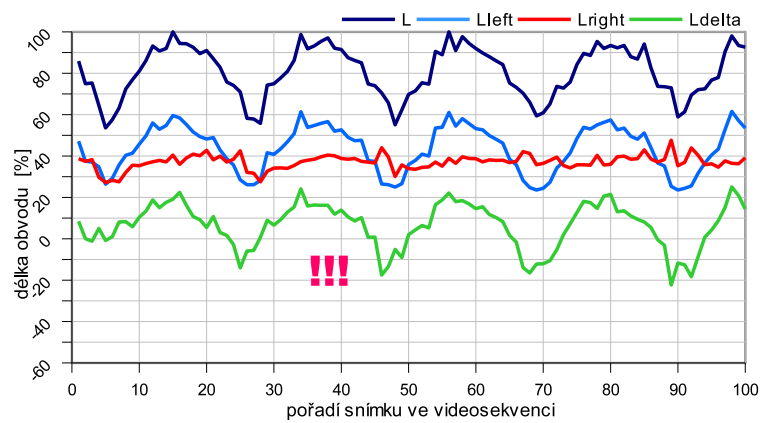
Obrázek 6.39: Vývoj velikosti plochy A hlasivkové štěrbiny.



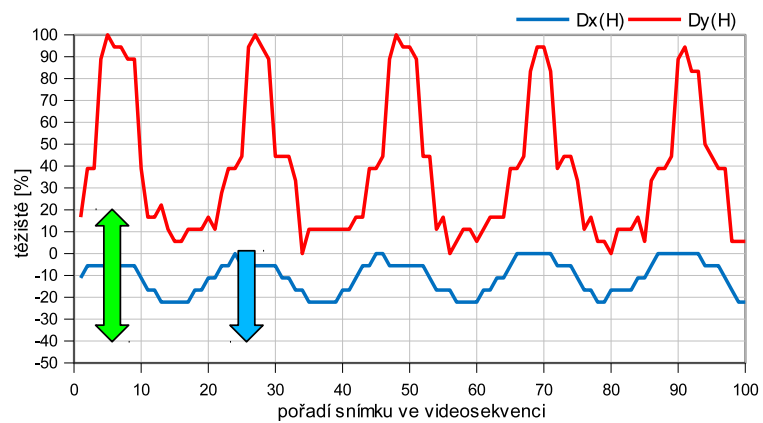
Obrázek 6.40: Poloha těžiště ve směru $D_x(S)$ a $D_y(S)$ podle pořadí snímků.



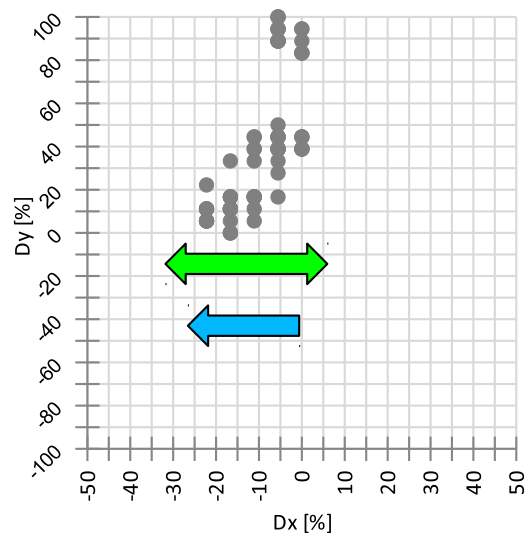
Obrázek 6.41: Vývoj těžiště plochy hlasivkové štěrby - normovaná trajektorie pohybu těžiště CA.



Obrázek 6.42: Vývoj délky L vnitřní hranice hlasivkové štěrby.



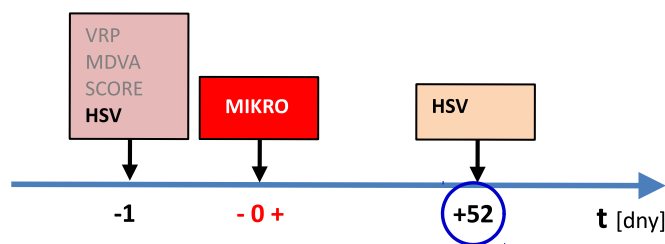
Obrázek 6.43: Poloha těžiště ve směru $D_x(H)$ a $D_y(H)$ podle pořadí snímků.



Obrázek 6.44: Vývoj polohy těžiště vnitřní hranice hlasivkové štěrbině - normovaná trajektorie pohybu těžiště CH.

Nesymetrická hlasivka - uzlík

(stav po zákroku)



Obrázek 6.45: Časová osa vyšetření před a po mikrochirurgickém zákroku - stav 52 dnů po odstranění uzlíku vlevo.

Popis:

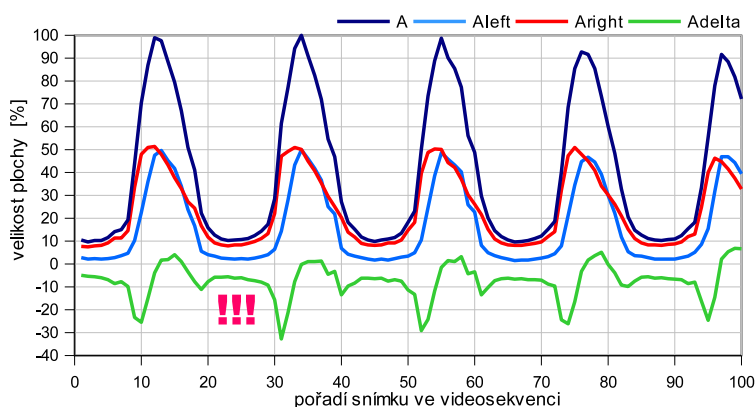
- Nesymetrická hlasivka, uzlík vlevo, stav 52 dnů po zákroku (obr. 6.46).
- Žena 48 let, $SPL_{min} = 82$ dB, $SPL_{max} = 84$ dB, $F_0 = 191$ Hz.
- Analýza plochy A_k a délky hranice L_k dále vykazuje nesymetrii, viz hodnoty parametrů A_{delta} a L_{delta} , i když se již pravá hlasivka pohybuje, viz A_{right} a L_{right} (obr. 6.47 a 6.50).
- Vývoj polohy těžiště plochy a obvodu štěrbině ve směru normály, tj. $D_x(S)$ a $D_x(H)$ potvrzuje zlepšení dynamiky hlasivek, pohyb těžiště ve směru osy symetrie hlasivky $D_y(S)$ a $D_y(H)$ je výrazný, pohyb těžiště ve směru normály $D_x(S)$ a $D_x(H)$ je naopak minimální (obr. 6.48 a 6.51).
- obě těžiště se pohybují po ose symetrie hlasivek s minimální vzdáleností (obr. 6.49 a 6.52).

Závěr:

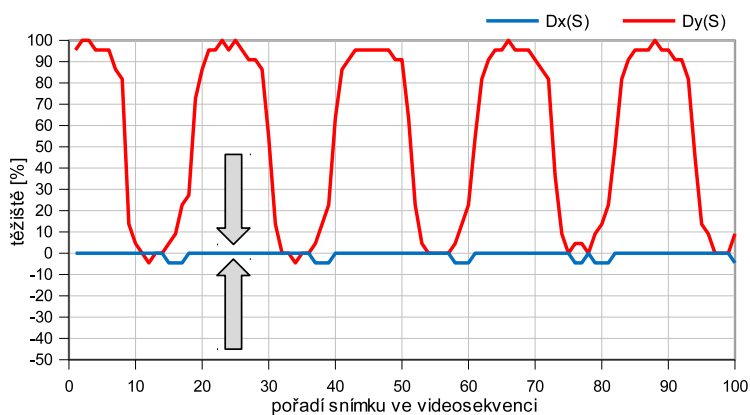
analýza vývoje plochy A_k a délky hranice L_k ukazuje na mírnou nesymetrii, která je způsobená pooperačním stavem; vývoj polohy těžiště plochy D_S a obvodu D_H štěrbiny během fonace však již tuto **nesymetrii hlasivek nepotvrzuje**.



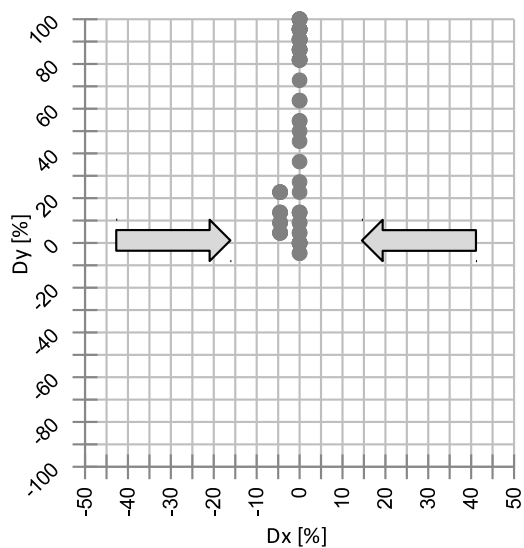
Obrázek 6.46: Příklad hlasivky po mikrochirurgickém zákroku (uzlík vlevo); vlevo: snímek z vysokorychlostní laryngoskopie (HSV); uprostřed: ROI – tzv. oblast zájmu; vpravo: segmentovaná hlasivková štěrбина s osou symetrie, těžištěm a vnitřní hranicí.



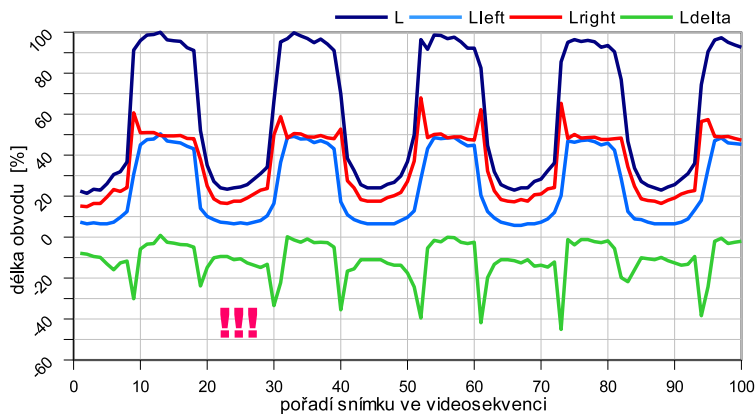
Obrázek 6.47: Vývoj velikosti plochy A hlasivkové štěrbiny.



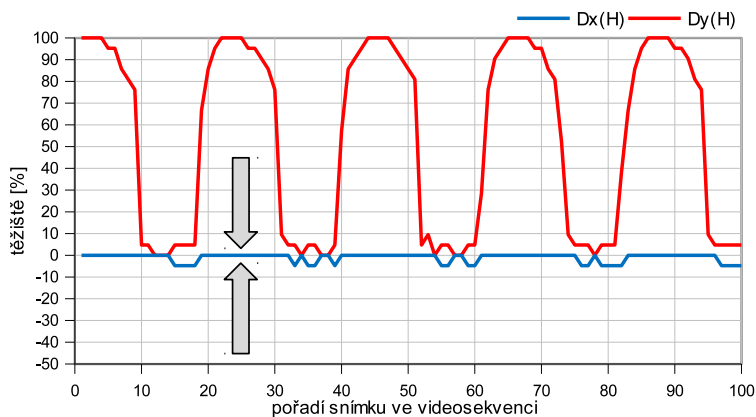
Obrázek 6.48: Poloha těžiště ve směru $D_x(S)$ a $D_y(S)$ podle pořadí snímků.



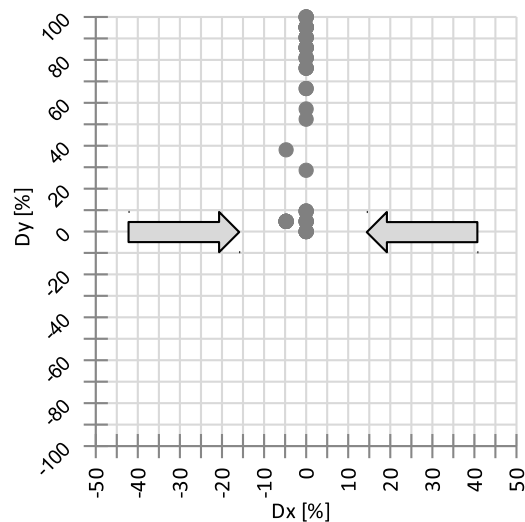
Obrázek 6.49: Vývoj těžiště plochy hlasivkové štěrby - normovaná trajektorie pohybu těžiště CA.



Obrázek 6.50: Vývoj délky L vnitřní hranice hlasivkové štěrby.



Obrázek 6.51: Poloha těžiště ve směru $D_x(H)$ a $D_y(H)$ podle pořadí snímků.



Obrázek 6.52: Vývoj polohy těžiště vnitřní hranice hlasivkové štěrbině - normovaná trajektorie pohybu těžiště CH.

6.3.3 Shrnutí získaných poznatků

Význam nově zavedených parametrů těžiště D_S plochy a D_H obvodové čáry hlasivkové štěrbině lze podle dosavadních výsledků shrnout a porovnat s parametry plochy A_k a obvodu L_k . Výsledky srovnání jednotlivých typových případů jsou uvedeny v tabulce 6.4. Obě těžiště vykazovala téměř shodný pohyb, pro symetrii je zajímavý pohyb zvláště ve směru x .

hlasivka	plocha, obvod	pohyb těžiště	popis
symetrická	symetrie	$D_x \rightarrow \min$	shoda parametrů - potvrzená symetrie
symetrická	nesymetrie	$D_x \rightarrow \min$	vývoj těžiště potvrzuje symetrii, i když vývoj A_k a L_k na to neukazuje
nesymetrická	nesymetrie	$D_x \rightarrow \max$	zvýrazněný pohyb těžiště ve směru normály – potvrzení nesymetrie
nesymetrická	nesymetrie	$D_x \rightarrow$ vychýlení	jednostranné vychýlení od osy symetrie – potvrzení nesymetrie

Tabulka 6.4: Přehled typových případů zjištěných po analýze kazuistik.

Spolehlivost a přesnost parametrů plochy obvodu a těžiště jsou závislé na kvalitě videozáznamu HSV, tj. na velikosti detailu hlasivek a úhlu, pod kterým jsou hlasivky snímány.

6.4 Normování

Většinu parametrů nelze běžně porovnávat mezi jednotlivými soubory dat, protože nejsou ve výchozím stavu nijak normovány. Normování je ale možné dopočítat ex post na základě znalosti dalších parametrů (např. podle velikostí plochy, délky osy symetrie apod.). Relativní parametry již nepotřebují být normovány, jako např. členitost či podlouhlost.

Přesnost parametrů závisí na kvalitě záznamu. Šum na snímcích přináší nepřesnosti, malé rozměry hlasivkové štěrbiny na snímku zase malý počet pixelů, které reprezentují oblast glottis, a tím klesá přesnost. Proto je nutné dbát na kvalitu vstupních dat.

6.5 Parametry přes více snímků

Kromě parametrů hlasivkové štěrbiny používané pro jednotlivé snímky je možné extrahovat parametry pro části videosekvence, např. jedna perioda kmitu hlasivek, definovaný počet snímků videosekvence či celá sekvence. Může jít o minima a maxima výše uvedených parametrů, průměry, rozptyly apod., ale i parametry přímo závislé na čase, jako jsou rychlost otevírání a zavírání hlasivky, délka jedné periody nebo obecně vývoj parametrů v čase.

6.6 Korelace parametrů s jinými způsoby vyšetření

V rámci studia jsem se zabýval korelací mezi jednotlivými parametry, pro zjištění, jak jsou na sobě parametry závislé. Experimentálně byly testovány nejen korelační vazby mezi parametry stanovenými z videozáznamu HSV navzájem, ale také s parametry dalších vyšetření (MDVA, VRP, SCORE, viz kap. 1.3).

Některé parametry spolu korelují ve většině případů, ale ojediněle se liší. Tyto odchylky mohou být dalším zdrojem informací, že dochází k nestandardnímu průběhu kmitání hlasivky.

Zajímavostí je, že žádný z parametrů získaných pomocí analýzy HSV nekoreluje (práh $>75\%$) s některým z parametrů získaných jinými typy vyšetření. V této souvislosti lze vyslovit hypotézu, že tato analýza přináší nová data, které nelze získat jiným způsobem. Příčina je pravděpodobně v odlišné metodice získávání dat a ne úplně srovnatelnými způsoby měření (např. akustické vs. optické metody). V tomto směru analýz budu dále pokračovat, viz kap. 7.5.

7 Teze dizertační práce

Současným problémem je spolehlivost automatické detekce štěrbin a ROI u snímků s horší kvalitou záznamu. Změna metody segmentace sice velmi zvýšila úspěšnost, nyní se ale začíná projevovat nepřesná lokalizace ROI jako nejslabší článek řetězce. Proto předpokládám další směřování práce podle následujících bodů.

7.1 Vylepšení hledání ROI

Prvním úkolem je najít vhodný preprocessing, např. Registrace snímků, aby byla přesnější lokalizace prostoru se štěrbinou – jediná oblast snímků, kde by mělo docházet k pohybu, je právě hlasivková štěrbinina. Z další části by měla vyplynout metoda, která ROI co přesněji určí, aby další automatické zpracování snímku proběhlo s minimální chybou. Např. použitím pouze části obrazu s vyšší frekvencí změny jasu ve stejnohlých bodech.

7.2 Registrace obrazu

Jak již bylo zmíněno, u některých záznamů dochází k pohybu kamery vůči hlasivce pacienta (z důvodu třesu rukou či pohybu pacienta). Již byl popsán způsob eliminace tohoto rušivého vlivu pomocí plovoucí osy. Z důvodu nízkého rozlišení snímků ale může docházet k příliš výrazným změnám mezi jednotlivými snímky, kdy i jeden obrazový bod je příliš velký rozdíl.

Použitá metoda ale nepočítá s tím, že by hlasivková štěrbinina opustila ROI, který obsahuje právě z těchto důvodů těsné okolí hlasivky.

Jednou z možností, jak problémům předejít a pravděpodobně i zjednodušit a zpřesnit další výpočty, případně umožnit použít statickou masku pro maximální oblast hlasivkové štěrbin, je registrace snímků ve videosekvenci.

Registrací je myšleno sesazení jednotlivých snímků tak, aby statické oblasti hrtanu v okolí hlasivek na snímcích v rámci jedné videosekvence měly stále stejnou polohu, jediný pohyb by pak vykazovaly pouze hlasivkové řasy (pokud nebudeme uvažovat tekutiny).

7.3 Validace výsledků

Během studia literatury jsem se nesetkal s metodou, která by řešila problém validace výstupních dat, tzn. kontrolu, jestli výsledek segmentace opravdu odpovídá

požadovanému výstupu. U svých metod jsem používal jednoduchou validaci, kdy detekovaná oblast štěrbinu nesmí zasahovat do rohu snímku, tato metoda je ale nedostatečná.

7.4 Adaptivní parametry shlukové analýzy

Dalším, paralelním, krokem by měla být modifikace metody shlukové analýzy s adaptivními parametry. První pokusy již byly provedeny a v tomto směru by mohl být další potenciál ke zlepšení výsledků segmentace.

Parametry s nízkou vahou z předchozího testování byly úplně vyřazeny jako nepodstatné a zbytečně snižovaly rychlost výpočtu. Naopak byl přidán nový parametr - rozdíl hodnoty červené složky bodu vlevo a vpravo od daného bodu.

Nastavení vztahů pro jednotlivé parametry probíhalo opět experimentálně. Jako nejvhodnější byly vybrány vztahy 7.1. Uvedené vztahy pomáhají zmenšovat nebo zvětšovat parametry podle rozsahu hodnot daného parametru.

$$\begin{aligned}
 R &= 2 \cdot \min\left(5, \frac{80}{R_{max} - R_{min}}\right) \\
 R - B &= R \\
 X &= \frac{40}{X_{max}} \\
 C &= \frac{R}{2} + X \\
 LR_{diff} &= \frac{R_{max} - R_{min}}{120}
 \end{aligned} \tag{7.1}$$

Navíc byl maximální jas omezen na hodnotu 150, aby se eliminovaly přесvícené části obrazů a snížily se rozdíly mezi různými videosekvencemi. Tato změna snížila kontrast mezi oblastí štěrbinu a okolí, nepřinesla však zhoršení detekce, naopak zvýšila stabilitu detekce u videosekvencí s nižší kvalitou záznamu.

Testování a vyhodnocování nejvhodnějších vztahů stále probíhá na dostupném vzorku dat. Zatím bylo dosaženo mírné zlepšení detekce u méně kvalitních záznamů při zachování stávající nebo lepší detekce případů správně detekovaných předchozí metodou.

7.5 Vyhodnocení dat

Dalším tématem je zpracování naměřených a vypočítaných dat, jejich statistická analýza a vyhodnocení. Jednou oblastí využití statistických metod je testování statistické významnosti a vzájemných vazeb parametrů podle vysokorychlostní laryn-

goskopie HSV. Cílem tedy bude definovat takové parametry, které mají význam pro odlišení videozáznamů podle kvality závěru glottis. Další oblastí statistické analýzy je vzájemné testování parametrů podle vysokorychlostní laryngoskopie HSV s parametry získanými pomocí jiných, nezávislých vyšetřovacích metod, zejména parametry hlasového pole (VRP), multidimenzionální analýzy (MDVA) a parametry kvality závěru glottis (SCORE).

Metody vyšetření VRP, MDVA a SCORE jsou standardně používané vyšetřovací metody, tj. k řadě videozáznamů v datovém korpusu DB-HSV, viz tab. 7.1 a 7.2, jsou dostupné také výsledky těchto vyšetření, viz tab. 7.3.

Korpus dat se stále průběžně doplňuje, upřesňují se diagnózy a přidávají další data z HSV, VRP, MDVA a SCORE (ke každému videozáznamu HSV existuje audiozáznam fonace vokálu „í“ během vyšetření vysokorychlostní laryngoskopií a tedy také výpočet (ohodnocení) upravenou metodou SCORE(HSV) pro fonaci vokálu „í“ oproti standardnímu „á“.)

datový korpus DB-HSV	videozáznamy			vyšetřované osoby		
	celkem	muži	ženy	celkem	muži	ženy
dg	549	190	359	203	75	128
cysta (cystis vocalis)	7	0	7	1	0	1
polyp (polypus vocalis)	31	13	18	14	8	6
vyjmutí hlasivek (chordektomie)	5	0	5	1	0	1
papilom (papilloma)	13	13	0	3	3	0
uzlík (noduli cant.)	16	6	10	8	4	4
karcinom (carcinoma)	8	8	0	6	6	0
granulom (granuloma)	2	2	0	1	1	0
Reikeho edém (reinke oedema)	18	1	17	10	1	9
paréza zvratného nervu (n. laryngeus rekurens paresis)	104	32	72	33	9	24
stav po odstranění krčních mandlí (tonsilektomie)	64	10	54	21	1	20
hemangiom (hemangioma)	3	0	3	1	0	1
stav po operaci štítné žlázy (glandula thyreoidea)	23	2	21	12	1	11
zdravá hlasivka	133	34	99	50	17	33
diagnóza neurčena	122	69	53	42	24	18

Tabulka 7.1: Přehled o skladbě používaného datového korpusu záznamů videosekvencí vysokorychlostní laryngoskopie HSV, DB-HSV ORL klinika FN Plzeň.

	věkové rozložení vyšetřovaných osob				
	min	max	medián	průměr	sm.odch.
celkem	13.00	82.00	49.00	47.93	16.46
muži	21.00	82.00	52.00	51.08	13.57
ženy	13.00	80.00	47.50	46.08	17.73

Tabulka 7.2: Věkové rozložení vyšetřovaných osob zařazených do datového korpusu.

videozáznamy HSV + vyšetření VRP, MDVA, SCORE	počet
HSV	549
HSV AND VRP	244
HSV AND MDVA	216
HSV AND SCORE	206
HSV AND VRP AND MDVA AND SCORE	153

Tabulka 7.3: Počty dostupnosti dat z vyšetření VRP, MDVA a SCORE k záznamům videosekvencí vysokorychlostní laryngoskopie HSV, DB-HSV ORL klinika FN Plzeň.

8 Závěr

Tato práce se zabývá zmapováním automatických metod analýzy obrazu hlasivek. Nejpoužívanější metody včetně předzpracování a následného zpracování byly popsány a byly zhodnoceny jejich výsledky, výhody a nevýhody. Dále byla uvedena vlastní metoda pro rozpoznávání hlasivkové štěrbině založená na shlukové analýze obrazových bodů.

V rámci této práce byla také uvedena problematika kvality vstupních obrazových dat a její vliv na detekci hlasivkové štěrbině.

Kromě detekce obrazu byla rozebrána i metoda pro hledání osy symetrie hlasivkové štěrbině. Pro ohodnocení symetrie byly uvedeny parametry pozice těžiště plochy a hranice hlasivkové štěrbině a jejich vývoj v čase. Na několika kazuistikách byla ukázána užitečnost těchto parametrů.

Vstupní data pro detekci hlasivkové štěrbině dodalo ORL oddělení FN Plzeň.

Na závěr jsou uvedena témata pro další práci v této oblasti.

Literatura

- [1] Gray, H.: Anatomy, published in 1918. Elsevier, 2008.
- [2] Čihák R.: Anatomie 2. 1. vyd. ISBN 80-060-88. Avicenum Praha, 1988.
- [3] Vokřál J.: Akustické parametry chraplavosti, Doktorská disertační práce. Fakulta elektrotechnická, České vysoké učení technické v Praze, 1998.
- [4] Švec, J.: Studium mechanicko-akustických vlastností zdroje lidského hlasu. dizertační práce, Přírodovědecká fakulta, Univerzita Palackého v Olomouci, Olomouc, 1996.
- [5] Novák A.: Foniatrie a pedaudiologie II, Poruchy hlasu - základy fyziologie hlasu, diagnostika, léčba, reedukace a rehabilitace. UNITISK, Praha, 1996.
- [6] Laryngoskopie. uLékaře.cz.
URL:<<http://www.ulekare.cz/clanek/laryngoskopie-1031>>, 13.4.2011.
- [7] Titze I.R.: Principles of Voice Production, Prentice Hall (currently published by NCVS.org), ISBN 978-0137178933. 1994.
- [8] Kučera M., Frič M, Halíč M.: Praktický kurz hlasové rehabilitace a reedukace. ISBN 978-8025465929. Opočno, 2010.
- [9] Multi-Dimensional Voice Program (MDVP), Operations manual, Kay Elemetrics Corp. 1995.
- [10] Pešta J., Slípka J., Nový P., Vávra F.: Hodnocení kvality závěru glottis. Otorinolaryngologie a foniatrie, 4 / 2010.
- [11] Kurdík, M.: Zátěžové hlasové testy. Diplomová práce, ZČU v Plzni, FAV, Katedra informatiky a výpočetní techniky, Plzeň, 2013.
- [12] Puchrová, L.: Aerodynamická vyšetření ve foniatrii. Bakalářská práce, ZČU v Plzni, FAV, Katedra informatiky a výpočetní techniky, Plzeň, 2016.
- [13] Kroupa, L: Detekce nestandardního kmitu hlasivek. Diplomová práce, ZČU v Plzni, FAV, Katedra informatiky a výpočetní techniky, Plzeň, 2015.

- [14] Kroupa, L., Vávra, F., Nový, P.: Statistic of Quasi-periodic Signal with Random Period - First Application on Vocal Cords Oscillation. 16th Conference on Applied Mathematics Aplimat 2017, Proceedings, ISBN 978-80-227-4650-2, Institute of Mathematics and Physics, Faculty of Mechanical Engineering, Slovak University of Technology in Bratislava, Bratislava, 2017.
- [15] Otsu N.: A threshold selection method from gray-level histograms, *IEEE Trans. Sys., Man., Cyber.* 9 (1): 62–66, 1979.
- [16] Morse B.: Thresholding, Brigham Young University, 2000.
- [17] Ng H. F.: Automatic thresholding for defect detection, *Pattern Recognition Letters*, (27): 1644-1649, 2006.
- [18] Kittler J., Illingworth, J.: Minimum error thresholding, SERC Rutherford Appleton Laboratory, Chilton, Didcot, Oxon OX11 0QX, U.K, 1986.
- [19] Blanco M., Chen X., Yan Y.: A Restricted, Adaptive Threshold Segmentation Approach for Processing High-Speed Image Sequences of the Glottis, *Scientific Research*. 2013.
- [20] Demeyer J., Dubuisson T., Gosselin B., Remacle M.: Glottis Segmentation with a High-Speed Glottography: a Fully Automatic Method. 3rd Advanced Voice Function Assessment International Workshop. Madrid, 2009.
- [21] Ettlér, T.: Analýza vysokorychlostního záznamu kmitání hlasivek. Diplomová práce, ZČU v Plzni, FAV, Katedra informatiky a výpočetní techniky, Plzeň, 2012.
- [22] Koç T., Çiloglu T.: Automatic Segmentation of High Speed Video Images of Vocal Folds, *Journal of Applied Mathematics*, Volume 2014, Article ID 818415. 2014.
- [23] Méndez A., Ismaili Alaoui E.M., García B., Ibn-Elhaj E., Ruiz I.: Glottal Space Segmentation from Motion Estimation and Gabor Filtering, 31st Annual International Conference of the IEEE EMBS. Minneapolis, Minnesota, USA, September 2-6, 2009
- [24] Gilles Degottex: Glottal source and vocal-tract separation. Signal and Image processing. Université Pierre et Marie Curie - Paris VI, 2010. English.
- [25] Aghlmandi D., Faez K.: Automatic Segmentation of Glottal Space from Video Images Based on Mathematical Morphology and the Hough Transform, *International Journal of Electrical and Computer Engineering (IJECE)*. ISSN: 2088-8708, 2012.
- [26] Serra, J.: Image Analysis and Mathematical Morphology , ISBN 0-12-637240-3 (1982)
- [27] Palm C., Keysers D., Lehmann T., Spitzer K.: Gabor Filtering of Complex Hue/Saturation Images for Color Texture Classification. Proc JCIS 2000, Atlantic City, USA, pp. 45-49, 2000.

- [28] Kass M., Witkin A., Terzopoulos D.: Snakes: Active contour models. In First International Conference on Computer Vision, pages 259–268, London, June 1987.
- [29] Allin S., Galeotti J., Stetten G., Dailey S. H.: Enhanced Snake Based Segmentation of Vocal Folds, 2004.
- [30] Schenk F., Ursler M., Aigner C., Roesner I., Aichinger P., Bischof H.: Automatic glottis segmentation from laryngeal high-speed videos using 3D active contours. 2014.
- [31] Bleau A., Leon L.: Watershed-based segmentation and region merging. *Computer Vision and Image Understanding* 2000;77(3):317–70.
- [32] Osma-Ruíz V., Godino-Llorente J.I., Sáenz-Lechón N., Fraile R.: Segmentation of the glottal space from laryngeal images using the watershed transform. *Computerized Medical Imaging and Graphics* 32(3), 193 – 201. 2008.
- [33] Miranda G.A., Llorente J.I.G., Velázquez L.M., García J.A.G.: An automatic method to detect and track the glottal gap from high speed videoendoscopic images. *BioMed Eng OnLine* (2015) 14:100 DOI 10.1186/s12938-015-0096-3. 2015.
- [34] Wood, S.A., Hoford, J.D., Hoffman, E.A., Zerhouni, E., Mitzner, W.: A method for measurement of cross sectional area, segment length, and branching angle of airway tree structures in situ. *Comput. Med. Imag. Graph.* 19 (1), 145–152, 1995.
- [35] Lohscheller J, Toy H, Rosanowski F, Eysholdt U, Döllinger M. Clinically evaluated procedure for the reconstruction of vocal fold vibrations from endoscopic digital high-speed videos. *Med Image Anal*, 11:400-413, 2007.
- [36] Shapiro, L., Stockman, G.: *Computer Vision*. Prentice Hall. pp. 69–73. 2002.
- [37] Kastner J.: Videokymografie a digitální kymografie, kymografické nálezy u pacientů po thyreoidektomii. *Endoskopie* 18 (2). Praha, 2009.
- [38] Peng B., Zhang L.: Automatic Image Segmentation by Dynamic Region Merging. *IEEE Trans. Image Processing*, vol. 12, no. 12, pp. 3592-3605, 2011.
- [39] Duda R. O., Hart P. E.: Use of the Hough Transformation to Detect Lines and Curves in Pictures. *Comm. ACM*, Vol. 15, pp. 11–15, January, 1972.
- [40] Döllinger M, Lohscheller J, McWhorter A, Kunduk M.: Variability of normal vocal fold dynamics for different vocal loading in one healthy subject investigated by phonovibrograms. *J Voice* 2009; 23(March 2):175–81, 2009.
- [41] Voigt D, Döllinger M, Braunschweig T., Yang A., Eysholdt U., Lohscheller J.: Classification of functional voice disorders based on phonovibrograms. *Artificial Intelligence in Medicine* 49 51–59, 2010.

-
- [42] Tatiraju, S., Mehta, A.: Image Segmentation using k-means clustering, EM and Normalized Cuts. Department of EECS (2008): 1-7. 2008.
- [43] Pešta, J., Slípka, J., Vohlídková. M., Ettler. T., Nový, P., Vávra, F.: Kinematika hlasivek - nové parametry hodnocení. Otorinolaryngologie a foniatrie, 65, č. 2, 2016.