

Univerzita Karlova v Praze  
Matematicko-fyzikální fakulta

# BAKALÁŘSKÁ PRÁCE



Aleš Zita

## **Zpracování videokymografických záznamů**

Kabinet software a výuky informatiky

Vedoucí bakalářské práce: RNDr. Zitová Barbara, Ph.D.

Studijní program: Informatika

Studijní obor: Obecná informatika

Praha rok 2011

Děkuji RNDr. Barbaře Zitové Ph.D. za její podněty, obětavou podporu a cenné rady, které mi velice pomohly při psaní mé bakalářské práce. Děkuji vřele RNDr. Janu Švecovi, Ph.D. et Ph.D. a Mgr. Davidovi Hauzarovi za jejich práci, na kterou jsem mohl navázat. V neposlední řadě děkuji všem mým učitelům MFF UK za jejich péči a vedení během mého studia.

Prohlašuji, že jsem tuto bakalářskou práci vypracoval samostatně a výhradně s použitím citovaných pramenů, literatury a dalších odborných zdrojů.

Beru na vědomí, že se na moji práci vztahují práva a povinnosti vyplývající ze zákona č. 121/2000 Sb., autorského zákona v platném znění, zejména skutečnost, že Univerzita Karlova v Praze má právo na uzavření licenční smlouvy o užití této práce jako školního díla podle § 60 odst. 1 autorského zákona.

V Praze dne .....

podpis

Název práce: Zpracování videokymografických záznamů

Autor: Aleš Zita

Katedra / Ústav: Kabinet software a výuky informatiky

Vedoucí bakalářské práce: RNDr. Zitová Barbara, Ph.D., Oddělení zpracování obrazové informace, Ústav teorie informace a automatizace, Akademie věd České republiky

**Abstrakt:** V rámci pokračujícího základního výzkumu funkce lidského hlasového orgánu se nyní vědci pokoušejí popsat závislost charakteristik slizničních vln na tvorbě tónu a vlastnostech hlasivek. Jednou z klinicky dostupných zobrazovacích metod je videokymografie – technika založená na snímání lidských hlasivek pomocí speciální řádkové CCD kamery. Jednotlivé řádky poskládány za sebe podle času pak tvoří videokymografický záznam. Videokymografické snímky jsou vhodné pro počítačovou extrakci základních charakteristik hlasivek, včetně charakteristik slizničních vln. V této práci byla vyvinuta a implementována metoda automatické detekce těchto slizničních vln za pomoci technik digitálního zpracování obrazové informace. Metoda je primárně určena pro zpracování videokymografických snímků hlasivek zdravých pacientů jakožto referenční skupiny fyziologické funkce hlasivek. **Klíčová slova:** videokymografie, charakteristiky hlasivek, slizniční vlny, digitální zpracování obrazu, vzájemná korelace

**Title:** Analysis of videokymographic images

**Author:** Aleš Zita

**Department:** Kabinet software a výuky informatiky

**Supervisor:** RNDr. Zitová Barbara, Ph.D., Department of Image Processing, Institute of Information Theory and Automation of the Academy of Science, Czech Republic

**Abstract:** In ongoing basic human vocal chord functionality research, scientists try to describe the dependency of the glottal wave characteristics on tone generation and vocal chord properties. One of the affordable vocal chord visualization techniques for clinics is the Videokymography – technique based on recording the vocal chord vibration using a CCD camera modified to a line-scanner. Recorded rows are then combined to form a spatial-time recording. These records are suitable for computer based extraction of the vocal chord characteristics, including the characteristics of the glottal waves. In this thesis the method for automatic detection of the glottal waves was developed and implemented utilizing the digital image processing techniques. The proposed method is primarily focused on processing the healthy vocal chords videokymographic images as they give us a good reference for the physiology of the human vocal chord organ.

**Keywords:** videokymography, vocal chord characteristics, glottal waves, digital image processing, cross correlation

# Obsah

|  |           |
|--|-----------|
| <b>Obsah</b>                             | <b>1</b>  |
| <b>Předmluva</b>                         | <b>1</b>  |
| <b>1 Úvod do problematiky</b>            | <b>3</b>  |
| 1.1 Lidský hlasový orgán                 | 3         |
| 1.1.1 Charakteristika hlasivek           | 3         |
| 1.1.2 Glotální cyklus                    | 4         |
| 1.1.3 Slizniční vlny                     | 4         |
| 1.2 Snímání a zpracování dat             | 6         |
| 1.2.1 Laryngoskopie                      | 6         |
| 1.2.2 Videokymografie                    | 7         |
| <b>2 Cíl práce</b>                       | <b>9</b>  |
| <b>3 Analýza problému</b>                | <b>10</b> |
| 3.1 Videokymografický záznam             | 10        |
| 3.1.1 Popis snímku                       | 10        |
| 3.1.2 Kvalita                            | 10        |
| 3.1.3 Odlesky                            | 10        |
| 3.2 Metody digitálního zpracování obrazu | 11        |
| 3.2.1 Předzpracování                     | 11        |
| 3.2.2 Segmentace                         | 11        |
| 3.2.3 Vzájemná korelace                  | 12        |
| <b>4 Popis řešení</b>                    | <b>14</b> |
| 4.1 Vstupní data                         | 14        |
| 4.1.1 Videokymografické snímky           | 14        |
| 4.1.2 Předzpracovaná data                | 14        |

|          |  |           |
|----------|--|-----------|
| 4.2      | Metoda detekce slizničních vln                       | 15        |
| 4.2.1    | Předzpracování                                       | 15        |
| 4.2.2    | Určení korelačního jádra                             | 16        |
| 4.2.3    | Iterativní maskovaná normalizovaná vzájemná korelace | 19        |
| 4.2.4    | Výsledek   | 22        |
| 4.3      | Výstupní data  | 23        |
| <b>5</b> | <b>Závěr a otevřené problémy</b>                     | <b>24</b> |
| 5.1      | Závěr  | 24        |
| 5.2      | Otevřené problémy                                    | 24        |
|          | <b>Doslov</b>  | <b>25</b> |
|          | <b>Seznam použité literatury</b>                     | <b>26</b> |
|          | <b>Seznam použitých zkratk</b>                       | <b>28</b> |
|          | <b>Příloha 1</b>                                     | <b>29</b> |
|          | Obsah CD   | 29        |
|          | <b>Příloha 2</b>                                     | <b>30</b> |
|          | Programová dokumentace                               | 30        |
|          | Instalace programu                                   | 30        |
|          | Spuštění programu                                    | 30        |
|          | <b>Příloha 3</b>                                     | <b>32</b> |
|          | Architektura programu                                | 32        |

# Předmluva

Diagnostika lidských hlasivek je ze své podstaty náročný úkol. Největším problémem při vizualizaci funkce lidského hlasového orgánu jsou vysoké frekvence vibrací hlasivek, které znemožňují použití běžných zobrazovacích technik. Typickou základní frekvencí lidských hlasivek jsou kmitočty mezi 80-240Hz. Z tohoto důvodu se jakákoliv technika umožňující vizualizaci jednotlivých kmitů hlasivek potýká s problémem extrémního množství dat. Například pro přehrání 10 sekundového záznamu hlasivek pořízeného systémem s časovým rozlišením 10.000fps je potřeba 2h 46min při rychlosti přehrávání 10fps (1). Je tedy prakticky nemožné posoudit všechny relevantní vlastnosti hlasivek z jediného takového záznamu, a tudíž je snahou použít k určení některých charakteristik metody digitálního zpracování obrazových dat.

Zadání problému zpracování videokymografického záznamu vibrací lidských hlasivek vzešlo od RNDr. Jana G. Švece, Ph.D. et Ph.D., který se touto tematikou dlouhodobě zabývá a zároveň je spoluautorem techniky snímání hlasivek pomocí řádkové kamery do videokymografického záznamu. Videokymografický záznam vzniká opakovaným zaznamenáním jednoho řádku obrazu hlasivkové štěrbin v místě největší amplitudy. Tyto řádky poskládané za sebe vytváří tzv. videokymogram, ve kterém je na jedné ose vynesena řádkový obraz hlasivek a na druhé čas (Obrázek 4).

Tato práce volně navazuje na diplomovou práci s názvem „Zpracování digitálních snímků videokymografických záznamů jako podpůrný nástroj pro diagnostiku hlasivek“ z roku 2010 (2), jejímž autorem je Mgr. David Hauzar. Autor ve své práci představil metodu pro automatickou detekci základních charakteristik hlasivek z videokymografického záznamu. Data z této práce lze s úspěchem použít jako výchozí bod pro další zpracování těchto snímků.

V kapitole *Úvod do problematiky* se nachází shrnutí faktů týkajících se vyšetření lidských hlasivek a metod vedoucích k akvizici obrazových dat. Kapitola *Cíl práce* pak popisuje zadání práce a její cíl. V kapitole *Analýza problému* jsem se zaměřil na popis vstupních dat a rozbor metod vhodných pro jejich zpracování. Na ní navazuje kapitola *Popis řešení*, v jejímž rámci jsou vysvětleny skutečně použité metody a

opodstatnění jejich použití. Poslední kapitola *Závěr a otevřené problémy* shrnuje dosažené výsledky a nabízí náměty pro navazující práce.



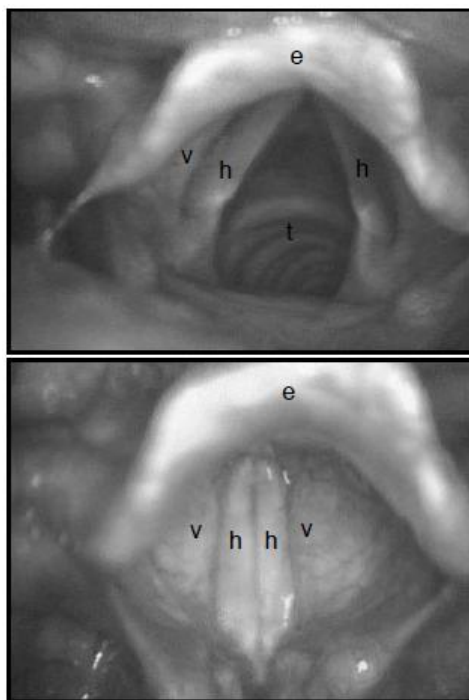
# 1 Úvod do problematiky

## 1.1 Lidský hlasový orgán

### 1.1.1 Charakteristika hlasivek

Zdrojem lidského hlasu jsou kmity hlasivek, které přerušují vzduchový proud, a tak vytvářejí prvotní akustický signál. Při řeči mohou vedle hlasivek působit i jiné akustické zdroje, jako například vzdušné turbulence vznikající zúžením průchozí cesty výdechového vzdušného proudu (např. mezi jazykem a patrem, zuby a rty, mezi rozevřenými hlasivkami při šepotu a podobně). Hlasivky jsou součástí hrtanu. Obrázek 1 ukazuje laryngoskopický pohled shora na hrtan. Hlasivky jsou při dýchání rozevřené a při fonaci přiložené k sobě. Štěrbina, která se při otevření tvoří mezi hlasivkami, se nazývá glottis. Pohyby hlasivek a jejich nastavení lze regulovat pomocí množství svalů, jež jsou upnuty na chrupavčitou kostru hrtanu.

Pomocí svalů je také možno měnit napětí hlasivek a tak ovlivňovat frekvenci kmitů při jejich rozechvění vzdušným proudem. Muži jsou schopni měnit tóny v rozsahu základních frekvencí přibližně 70 – 500 Hz, ženy v rozsahu 150 – 1000 Hz, přičemž zde bývají výrazné interindividuální rozdíly (3).



Obrázek 1. Lidský hlasový orgán: h-hlasivka; v-ventrikulární řasa; e-hrtanová příklopka; t-průdušnice. Obrázek převzat z (3).

### *1.1.2 Glotální cyklus*

Glotalní cyklus je jeden cyklus vibrací lidského hlasového orgánu. Je základním pro tvorbu hlasu. Bývá obvykle popsán následujícími fázemi:

- 1) uzavřená hlasivková štěrbina
- 2) vzrůst tlaku pod hlasivkami
- 3) otevření hlasivek a vypuštění jednoho kvanta vzduchu
- 4) uzavření štěrbiny

Pro naše účely budeme rozeznávat: uzavřené hlasivky (cca 3ms), otevřené hlasivky, otevírací fáze (cca 2ms), stav s největším otevřením štěrbiny v nejširším místě otevřených hlasivek (stručně amplituda) a uzavírací fáze (cca 3ms).

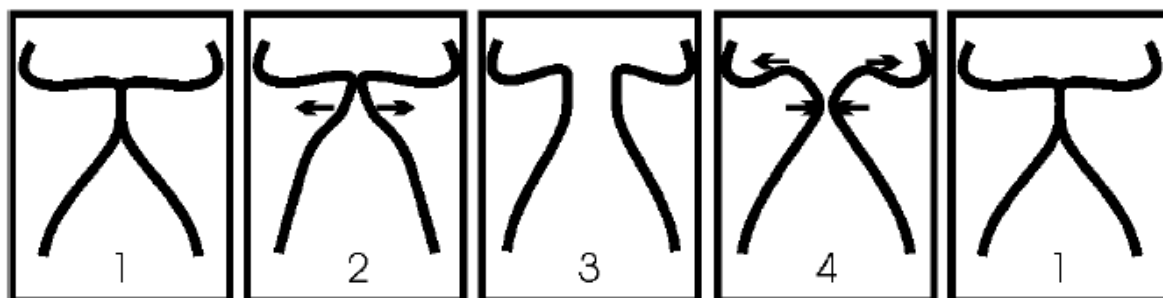
### *1.1.3 Slizniční vlny*

Při hlubokých tónech v tzv. hrudním rejstříku, kdy jsou hlasivky relativně málo napnuté, mají hlasivky nezanedbatelnou vertikální tloušťku a existuje zde fázové zpoždění mezi pohybem dolní a horní části hlasivek, které je pro tyto tóny typické (hlasivky se otevírají odspodu tak, že se nejprve od sebe oddělí spodní část hlasivek a poté horní část, stejně tak se zavírá nejprve spodní a poté horní část). Zároveň se zde objevuje tzv. slizniční vlna, která se šíří po povrchu hlasivek od středu do stran. (Obrázek 2; Obrázek 5)

Slizniční vlny patří mezi sledované charakteristiky. Významné jsou tyto případy: chybějící slizniční vlny, malé slizniční vlny a vzdálené slizniční vlny. Slizniční vlna je malá, pokud je menší než polovina velikosti hlasivek. Čím jsou slizniční vlny výraznější, tím ohebnější je sliznice (3).

**KMITY HLASIVEK**

**Modální/hrudní rejstřík**



Obrázek 2. Grafické znázornění slizniční vlny. Obrázek převzat z (3).

## 1.2 Snímání a zpracování dat

### 1.2.1 Laryngoskopie

Laryngoskopie je metoda nahlédnutí do hrtanu, respektive zobrazení hrtanu a jeho součástí včetně hlasivkové štěrbiny (Obrázek 3).

První známé pokusy o prozkoumání lidského hlasového orgánu jsou starší než 150 let. V roce 1854 se učitel zpěvu Manuel Garcia pokusil o prozkoumání svých hlasivek pomocí dvou zrcátek, zubařského a malého ručního. Své pozorování pak publikoval v roce 1855 (4).

Nezávisle na tom demonstrovali o dva roky později vídeňský neurolog Truck a pražský fyziolog Johann Nepomuk Czermak (mj. vynálezce lékařského reflektoru) použití této techniky v medicíně (4). Od té doby zaznamenaly zobrazovací techniky vhodné pro zkoumání lidského hlasového orgánu obrovský pokrok.



Obrázek 3. Laryngoskopický snímek lidských hlasivek. Obrázek převzat z (8).

V dnešní době jsou diagnostické postupy při určování jednotlivých nemocí v klinické praxi velmi složité. Jsou založené na popisu potíží pacienta, empirických poznatcích a vyhodnocování dat z instrumentálních a histologických vyšetření. Během posledních dvou desetiletí bylo vyvinuto mnoho zobrazovacích technik pro přesné měření kvality hlasu (5; 6).

Obrovským skokem kupředu bylo například použití počítačové tomografie (CT) a magnetické resonance (MRI) pro vyšetření hlasivek, neboť tyto techniky umožnily

zobrazení oblastí hlasivek, které zůstávají skryty při tradičním endoskopickém vyšetření, nebo umožňují vizualizaci hloubky případného nádoru (7).

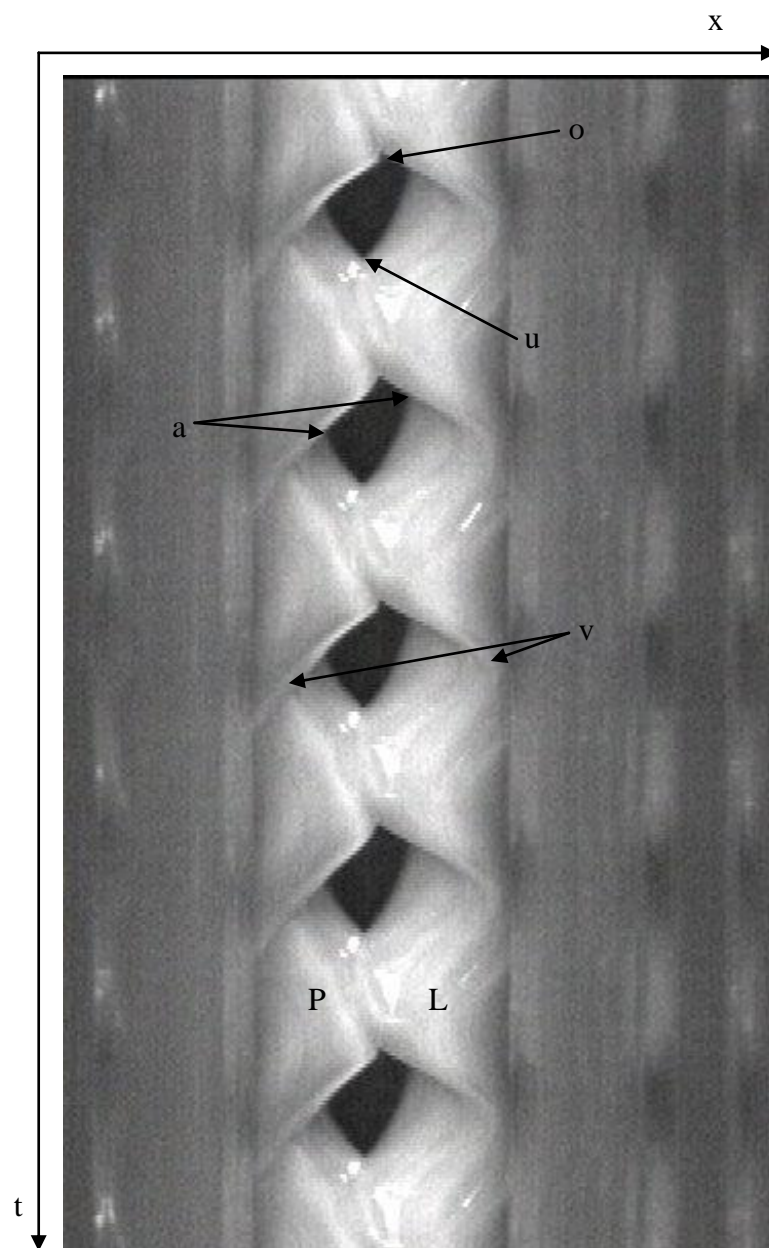
Tyto metody se dnes používají především jako zobrazovací techniky při léčbě zhoubných nádorů a při předoperačním vyšetření k vhodnému plánování chirurgického zákroku (8).

### *1.2.2 Videokymografie*

#### **Charakteristika metody**

Videokymografie (VKG) je původní česko-nizozemská metoda, vyvinuta v roce 1994 v Groningenu (NL) jako levná alternativa k vysokorychlostnímu záznamu. VKG se používá pro optické zaznamenávání mechanických kmitů a byla vyvinuta zejména za účelem zkvalitnění diagnostiky kmitů hlasivek ve foniatrii a laryngologii (9). Základem systému je speciálně upravená CCD videokamera, která je schopna pracovat dvěma různými způsoby - standardně (50 snímků/s - prokládaně) nebo vysokofrekvenčně (7 812.5 řádkových snímků/s). Ve vysokofrekvenčním modu systém registruje snímky pouze z jednoho vybraného horizontálního řádku kamery a skládá je za sebe (Obrázek 4) (10).

Metoda umožňuje rychle a jednoduše zaznamenávat pravidelné i nepravidelné kmity hlasivek, neboť frekvence 7 812.5 snímků/s dostatečně převyšuje frekvenci kmitů hlasivek (cca 70 - 1000 Hz). Ze záznamu lze odečítat důležité vlastnosti kmitů, které napovídají o fyziologickém či patologickém stavu hlasivek a které nebylo možno pomocí dosavadních běžně dostupných laryngologických vyšetřovacích metod (laryngoskopie, laryngostroboskopie) objektivně posuzovat. Kameru je možno použít např. i pro záznam a hodnocení kmitání strun u hudebních nástrojů, či vibrací průmyslových zařízení. Výhodou tohoto systému je příznivá cena, neboť kamera nevyžaduje nestandardní záznamová média a záznam je možno provést např. na běžně dostupný standardní videomagnetofon. Zajímavostí je, že byla tato metoda poprvé nasazena do klinické praxe v r.1996 v Praze (10).



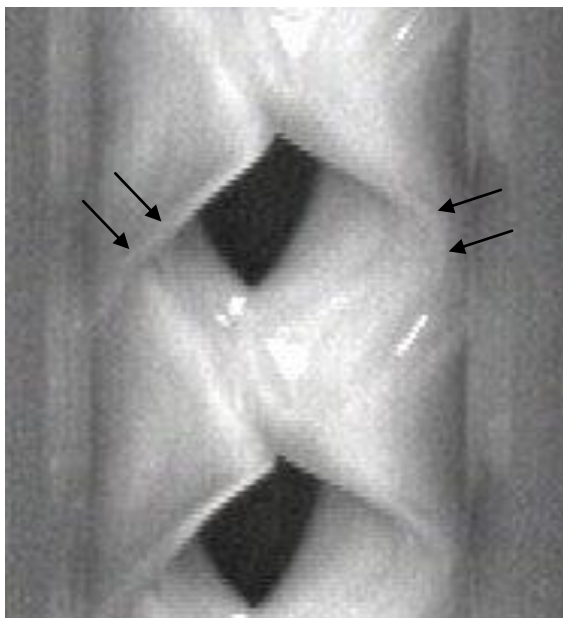
Obrázek 4. Příklad VKG snímku: x – levo-pravá prostorová osa; t – čas; L – levá hlasivka; P – pravá hlasivka; o – bod počátku otevírání štěrbiny; u – bod uzavření štěrbiny; a – amplituda otevření hlasivky; v – slizniční vlna.

## 2 Cíl práce

Jedním z cílů základního výzkumu funkce lidského hlasového orgánu je zjistit, jaké existují souvislosti mezi tvorbou tónu a jeho barvou ve vztahu k charakteru bočně postupujících slizničních vln. Tento výzkum se z důvodů získání spolehlivých referencí provádí zejména na snímcích hlasivek zdravých pacientů.

Mým úkolem v bakalářské práci bude zaměřit se na metody automatické detekce těchto slizničních vln v kymografických záznamech a určení jejich základních charakteristik. Tyto charakteristiky jsou kromě zmíněného také podstatným ukazatelem tuhosti hlasivkové tkáně. Příčinnou tuhé hlasivkové tkáně pak může být například zduření hlasivek, nebo nádor.

Slizniční vlny se na videokymografickém záznamu (Obrázek 5) vyskytují buď s odlesky nebo bez nich (2). Důležitými charakteristikami slizničních vln pro automatické zpracování jsou: dosah slizniční vlny, tedy velikost její amplitudy a úhel, pod kterým vlna postupuje.



Obrázek 5. Slizniční vlna na VKG snímku.

## 3 Analýza problému

### 3.1 Videokymografický záznam

#### 3.1.1 Popis snímku

Na každém kymografickém snímku se nachází záznam několika otevření a uzavření hlasivkové štěrbin. Slizniční vlny se na snímcích projevují někdy více, jindy méně zřetelnou konturou, kterou lze najít v místech zobrazující hlasivkovou tkáň, mezi hlasivkovou štěrbinou a tkání okolní (Obrázek 5).

Existuje mnoho faktorů ztěžujících automatickou detekci slizničních vln na videokymografickém záznamu. Mezi nejvýraznější patří špatné rozlišení snímků, šum a odlesky.

#### 3.1.2 Kvalita

Videokymografický snímek hlasivek je typicky šedotónový snímek horší kvality, často zatížený šumem. Šum na datech tohoto typu bývá zpravidla aditivního charakteru, způsobeného nedostatečným osvětlením a teplotou CCD snímače.

Zatímco zvýšení rozlišení snímku je těžký úkol, existuje mnoho metod potlačujících nechtěný šum v obrazových datech.

#### 3.1.3 Odlesky

Důvodem zviditelnění bočně postupujících slizničních vln na snímku jsou prakticky výhradně odlesky a stíny vln dané polohou světelného zdroje endoskopu. Tento fakt může způsobit, že na některých záznamech jsou tyto vlny výrazně světlejší než okolní tkáň a na jiných málo zřetelné nebo okem téměř nerozeznatelné. Zpravidla také existuje velký rozdíl v projevu levé vlny oproti pravé na stejném záznamu. Tento fakt je způsoben asymetrií osvětlení levé a pravé hlasivkové tkáně. Dalším



efektem bodového osvětlení jsou velké jasové změny i v rámci jednoho cyklu otevření hlasivkové štěrbin, které komplikují detekci.

## 3.2 Metody digitálního zpracování obrazu

### 3.2.1 Předzpracování

Metody předzpracování snímku jsou zpravidla určeny k odstranění nebo potlačení šumu, nebo pro zvýraznění informace ve snímku obsažené.

#### **Odstranění šumu**

Metody odstranění aditivního šumu jsou založeny na mírném rozmazání snímku, kterým je ze snímku odstraněna vysokofrekvenční informace. Techniky odstranění šumu jsou většinou založeny na konvoluci s příslušným jádrem. Nejčastěji se jako jádro používá gaussovský low-pass filtr (11). Dalšími možnostmi odstranění šumu by mohly být jeden ze způsobů váženého průměrování obrázku nebo mediánový filtr (11). Pro potřeby detekce slizničních vln lze použít gaussovský filtr.

#### **Zvýraznění informace**

Pro zviditelnění informace ve snímku se často používají metody zvýraznění hranové informace buď za pomoci hranových detektorů, nebo opět konvoluce s patřičným jádrem, například LoG (Laplace of Gaussian) (11). K dalším technikám zvýraznění informace v obrazu patří metody modifikace histogramu snímku a to buď klasickou nebo adaptivní ekvalizací histogramu (11). Vzhledem k povaze dat jsem ve své práci volil adaptivní ekvalizaci histogramu.

### 3.2.2 Segmentace

Segmentace je způsob rozdělení obrázku na dvě a více částí se společnými vlastnostmi na základě informací v těchto oblastech obsažených. Metod, které se používají, je více. Uvedu dvě, kterými jsem se za účelem segmentace slizničních vln zabýval. Jedná se o různé varianty prahování, tedy o rozdělení obrázku do segmentů

podle hodnot jasu (11), a dále o metodu s názvem Active Contour Models, známou též pod názvem Snakes (12). Tato je založena na minimalizaci energie asociované s danou konturou, jež je definovaná jako součet interní a externí energie této kontury přiřazené.

Jedním ze směrů mého výzkumu byl pokus o segmentaci slizniční vlny některou z výše zmíněných metod. Bohužel kontury slizničních vln na kymografických záznamech mají v důsledku výše zmíněných problémů s osvětlením a odlesky jednak proměnlivý jas a jednak neostré, často ztrácející se hranice, takže se nakonec nepodařilo ani jednu ze zmíněných metod segmentace uplatnit.

### 3.2.3 Vzájemná korelace

Korelace v oblasti zpracování signálu, do které patří i techniky digitálního zpracování obrazových dat, je technikou určující vzájemnou podobnost dvou signálů. Spojitá 1D varianta korelace:

$$c(t) = \int_{-\infty}^{+\infty} f^*(\tau)g(t + \tau) d\tau$$

Vzorec 1. Spojitá korelace:  $f()$  a  $g()$  jsou funkcemi porovnávaných signálů;  $f^*(\tau)$  - komplexně sdružené číslo k  $f(\tau)$ .

V diskrétní obrazové oblasti se pak jedná o podobnost dvou obrázků. Funguje na principu porovnání dvou signálů pomocí ekvivalentu skalárního součinu.

$$c[m, n] = \sum_{x, y} f[m + x, n + y]t[x, y]$$

Vzorec 2. Diskrétní 2D vzájemná korelace:  $f$  - obrazová funkce;  $t$  - korelační jádro.

Pro použití ve zpracování obrazové informace se z důvodu často lišících se jasových hodnot způsobených například rozdíly v osvětlení používá diskrétní normalizovaná vzájemná korelace, při které je každá ze dvou porovnávaných částí snímku před vzájemným pronásobením nejprve normalizována za pomoci průměrné hodnoty jasu.

$$c[m, n] = \frac{\sum_{x,y} (f[x + m, y + n] - \bar{f}) \cdot (t[x, y] - \bar{t})}{\sqrt{\sum_{x,y} (f[x + m, y + n] - \bar{f})^2} \cdot \sqrt{\sum_{x,y} (t[x, y] - \bar{t})^2}}$$

Vzorec 3: Normalizovaná vzájemná korelace:  $f$  - obrazová funkce;  $\bar{f}$  - střední hodnota funkce ve zkoumané oblasti;  $t$  - korelační jádro;  $\bar{t}$  - střední hodnota korelačního jádra; Převzato z (11).

Budeme-li vycházet z předpokladu, že má slizniční vlna podobný charakter po celé své délce, nabízí se normalizovaná vzájemná korelace jako vhodná metoda pro detekci slizničních vln, jelikož je založena právě na hledání podobností.

## 4 Popis řešení

### 4.1 Vstupní data

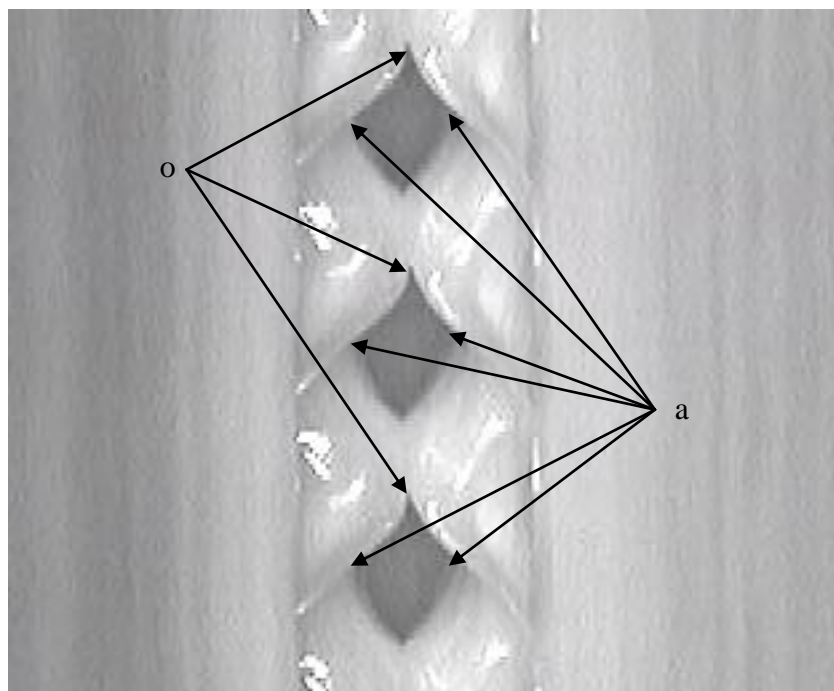
#### 4.1.1 Videokymografické snímky

Vzorové videokymografické snímky pro návrh metody na detekci slizničních vln poskytl RNDr. Jan G. Švec, Ph.D. et Ph.D. Jedná se o modelové snímky zdravých pacientů, na kterých jsou slizniční vlny okem dobře patrné. Bohužel jsou takovéto snímky poměrně vzácné, protože většina pacientů, která se laryngoskopickému vyšetření podrobí, pochopitelně trpí nějakou poruchou. Z tohoto důvodu se podařilo získat za tímto účelem pouze 6 snímků.

#### 4.1.2 Předzpracovaná data

Každá slizniční vlna zdravých hlasivek začíná na záznamu vždy v místě začátku otevření hlasivkové štěrbiny a pokračuje podél hran otevírání obou hlasivek přes místa největšího rozevření neboli amplitudy každé z hlasivek a končí buď na hranici okolní tkáně, nebo se stáčí „dolů“ a vrací se zpět (Obrázek 5).

Za účelem detekce slizničních vln je tedy potřeba znát pozice začátku otevírání hlasivek a jejich amplitud. Jelikož se jedná o informace, které určuje ve své diplomové práci Mgr. Hauzar, rozhodl jsem se jeho práci za tímto účelem použít. Všechny snímky určené pro detekci slizničních vln jsou nejprve předzpracovány pomocí programu Mgr. Hauzara a výsledky jsou uloženy na pevný disk (Obrázek 6).



Obrázek 6. Data předem získaná z (2): o - pozice otevření štěrbin; a - pozice pravých a levých amplitud hlasivek.

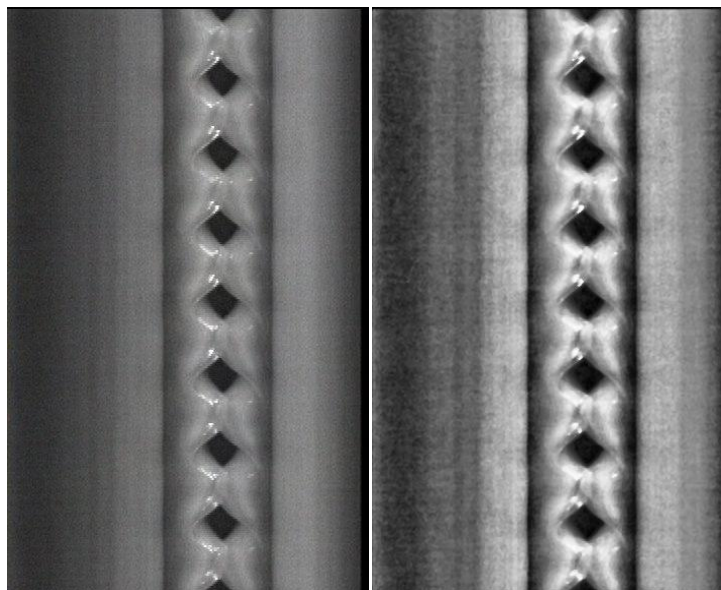
## 4.2 Metoda detekce slizničních vln

Při hledání techniky vhodné k detekci slizničních vln jsem vycházel z předpokladu, že každá slizniční vlna má podobný charakter po celé své délce. Naopak intenzita se v průběhu vln mění poměrně často. V důsledku toho se ukázaly metody segmentace obrazových dat jako nevhodné, takže jsem pro detekci vln volil metody založené na vzájemné korelaci, jelikož je tato určena právě k hledání podobností.

### 4.2.1 Předzpracování

Snímek kymografického záznamu je nejprve zbaven šumu za pomoci gaussovského low-pass filtru. Protože je na většině snímků slizniční vlna málo kontrastní, dalším krokem je adaptivní ekvalizace histogramu, kterou je dosaženo lokálního zvýšení kontrastu a slizniční vlna se stane patrnější. V takto zpracovaném snímku ovšem dojde k opětovnému zvýraznění šumu. Je tedy třeba použít ještě jednu techniku potlačení šumu. Na tomto místě jsem zvolil mediánový filtr pro jeho vlastnost

menšího rozmazání hran v obraze. Tímto postupem je snímek připraven pro vlastní algoritmus hledání slizničních vln (Obrázek 7).



Obrázek 7. Vlevo: Originální snímek; Vpravo: Předzpracovaný snímek

#### 4.2.2 *Určení korelačního jádra*

Jak již bylo zmíněno výše, povaha kontury slizniční vlny zachycené na videokymografickém záznamu napovídá, že by bylo obtížné detekovat tyto vlny pomocí nástrojů segmentace nebo hranových detektorů. Naopak z částečné soběpodobnosti vlny po její délce plyne, že vhodnou technikou pro hledání průběhu vlny je vzájemná korelace. Pro použití vzájemné korelace je třeba nejprve určit velikost korelačního jádra a pozici, ze které mají být použity informace jako hodnoty tohoto jádra.

##### **Velikost korelačního jádra**

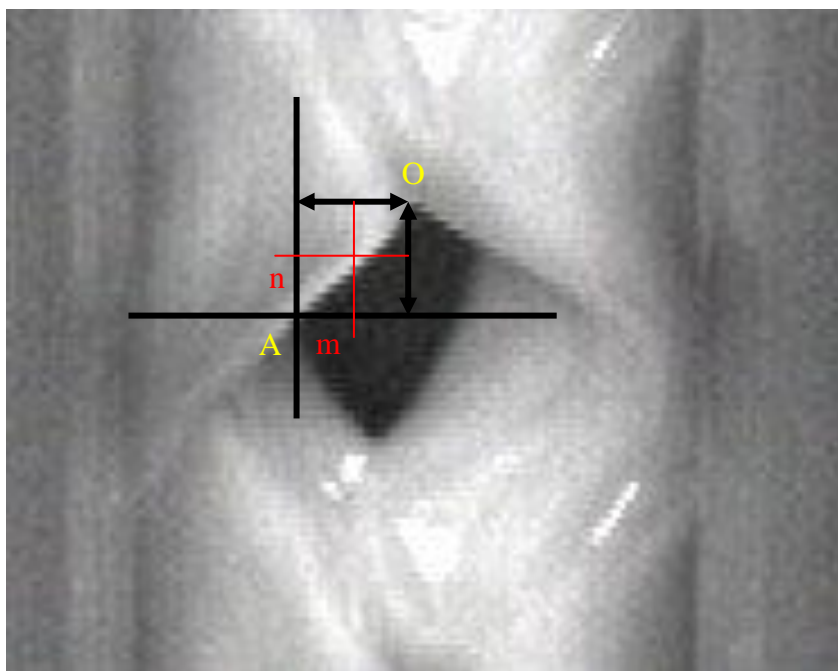
Účelem určení velikosti korelačního jádra je zajistit, aby jádro obsahovalo právě celou informaci o tvaru slizniční vlny. Příliš malé jádro by znamenalo menší množství informace v jádru obsažené. Naopak příliš velké korelační jádro by znamenalo, že obsahuje množství informace o okolí slizniční vlny. Toto okolí ale může být v průběhu proměnlivé.

Rozlišení slizniční vlny na kymografickém záznamu je závislé na rozlišení záznamu a na frekvenci kmitů hlasivek. Při vyšší frekvenci hlasivkových kmitů jsou amplitudy hlasivek menší a doba otevření štěrbiny kratší. Slizniční vlny jsou tedy kratší a méně prostorově výrazné.

Při určení velikosti masky je tedy třeba vycházet z výše zmíněných informací. Nejlepších výsledků bylo dosaženo použitím průměru osové vzdálenosti pozice maxima a počátku otevření hlasivkové štěrbiny (Vzorec 4; Obrázek 8).

$$d = \left\lfloor \frac{|a - o|(x) + |a - o|(y)}{2} \right\rfloor$$

Vzorec 4. d - velikost korelačního jádra; o - pozice začátku otevření; a - pozice amplitudy hlasivky



Obrázek 8. Určení velikosti korelačního jádra: m+n , tak, že: n ... polovina vzdálenosti otevření a amplitudy ve směru osy y.; m ... dtto pro osu x. ; A – pozice amplitudy otevření hlasivky; O – pozice počátku otevírání hlasivky.

Velikost korelačního jádra je určena zvlášť pro každou instanci otevření hlasivkové štěrbiny.

### **Tvar korelačního jádra**

Normalizovaná vzájemná korelace pracuje primárně se čtvercovými korelačními jádry. Pro dosažení větší přesnosti je třeba, aby byl tento čtverec pokud možno kolmý na průběh slizniční vlny. Tím je dosaženo maximalizace informace v jádru obsažené. Protože matice, ve které je jádro uložené, je čtvercová, je potřeba použít logickou masku pro určení aktivní oblasti jádra. Tato bude použita pro výpočet maskované normalizované vzájemné korelace.

První aproximace úhlu, pod kterým slizniční vlna ve snímku postupuje je spočtena pomocí informací o pozici začátku otevření hlasivkové štěrbiny a pozici amplitudy příslušné hlasivky. Hodnota úhlu je poté použita pro výpočet sklonu masky korelačního jádra. Logickým násobením masky s korelačním jádrem je dosaženo korelačního jádra požadovaného tvaru.

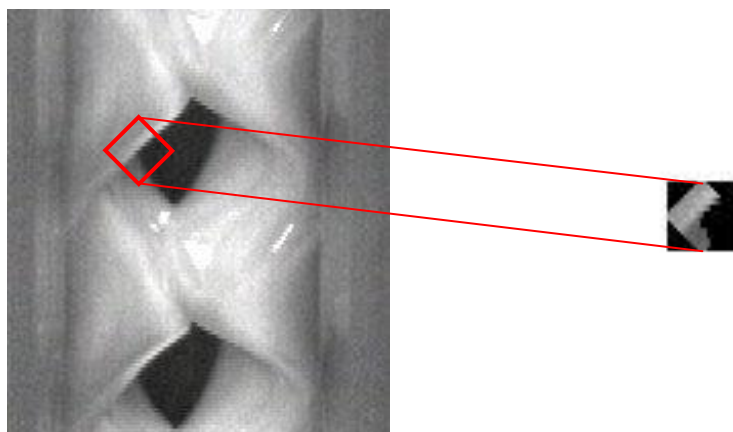
### **Obsah korelačního jádra**

K určení pozice, ve které se nachází informace, jež má být použita jako korelační jádro pro hledání podobnosti, je nutné určit polohu nejreprezentativnější části vlny. Slizniční vlna začíná být dobře patrná zpravidla v první polovině běhu mezi otevřením hlasivkové štěrbiny a amplitudou příslušné hlasivky. Kontura vlny se po celém svém průběhu postupně mění, takže poloha pro data korelačního jádra určená jako pozice v úseku mezi otevřením a amplitudou se ukázala jako málo stabilní.

Dalším nápadem bylo použití pozice amplitudy jako středu korelačního jádra. Za tímto účelem byla upravena maska korelačního jádra pomocí segmentace hlasivkové štěrbiny z korelačního jádra, čímž vzniklo korelační jádro tvaru „L“, které obsahovalo informaci o okolí jak vně vlny, tak uvnitř (Obrázek 9). Toto řešení ale

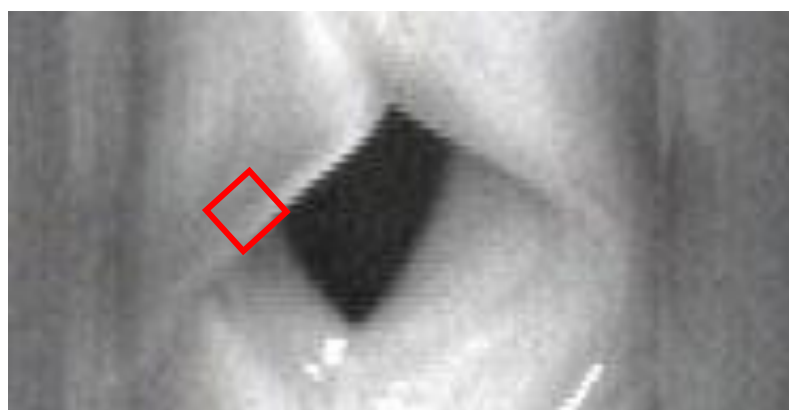


vykazovalo nestabilitu u snímků s měnícím se vnitřním okolím slizniční vlny.



Obrázek 9. Korelační jádro tvaru "L".

Poloha pro informaci použitou v korelačním jádru byla nakonec určena, jako poloha hraničící s pozicí amplitudy otevření vždy na její vnější straně. Na tomto místě je informace o vlně stále reprezentativní a je dobrou aproximací dalšího průběhu vlny (Obrázek 10).



Obrázek 10. Poloha pro určení korelačního jádra.

#### 4.2.3 Iterativní maskovaná normalizovaná vzájemná korelace

##### **Korelace**

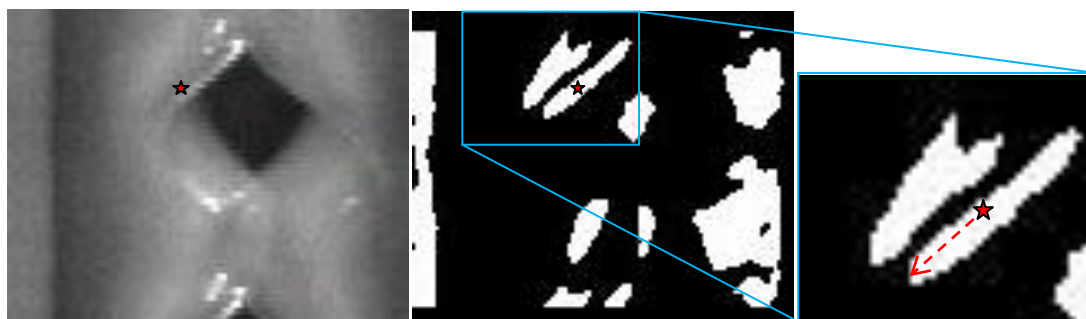
Vlastní implementace vzájemné korelace byla upravena tak, aby zahrnula arbitrární tvar korelačního jádra. Před násobením obou korelačních elementů je každý z nich nejprve pronásoben příslušnou korelační maskou.

Další úprava vzájemné korelace v klasickém pojetí byla vyvolána výše zmíněným faktem, že slizniční vlna se v důsledku bodového osvětlení postupně mění v celém svém průběhu. Ke změnám v kontuře průběhu vlny přispívá i fakt, že se na některých snímcích kontura rozšiřuje nebo plynule stáčí směrem k dolnímu okraji snímku. Zmíněné stáčení vlny zpravidla indikuje blížící se amplitudu, kterou je třeba detekovat.

Tato fakta vedla k návrhu iterativní metody normalizované vzájemné korelace.

### Hledání extrémů

Výsledkem normalizované vzájemné korelace je korelační matice velikosti korelovaného obrázku, jejíž hodnoty se pohybují mezi -1 a 1. Každá hodnota určuje míru podobnosti výřezu tvaru masky příslušné pozice korelovaného obrázku s korelačním jádrem. Kladné hodnoty znamenají pozitivní podobnost, záporné podobnost s negativem. V navrhovaném algoritmu jsou za účelem detekce postupu vlny použity pouze hodnoty přesahující předem stanovený práh. Algoritmus hledání extrémální hodnoty je modifikovaný DDA (Digital Differential Analyzer) algoritmus (13), který postupuje po hodnotách korelační matice ve směru daném aproximací úhlu postupu slizniční vlny tak dlouho, dokud hodnoty korelační matice přesahují stanovený práh (Obrázek 11). V opačném případě se postupuje směrem dolů, kam se předpokládá případné stočení vlny, a úhel pro další postup po korelační matici je náležitě upraven. V případě nemožnosti dalšího postupu po korelační matici je nalezena extrémální hodnota postupu vlny v nejzazším místě, které je podobné startovní pozici.



Obrázek 11. Hledání extrému: Hvězdička vyznačuje počátek hledání, šipka směr hledání; Vlevo - původní snímek; Uprostřed - oprahovaná korelační matice; Vpravo - zvětšená část korelační matice.

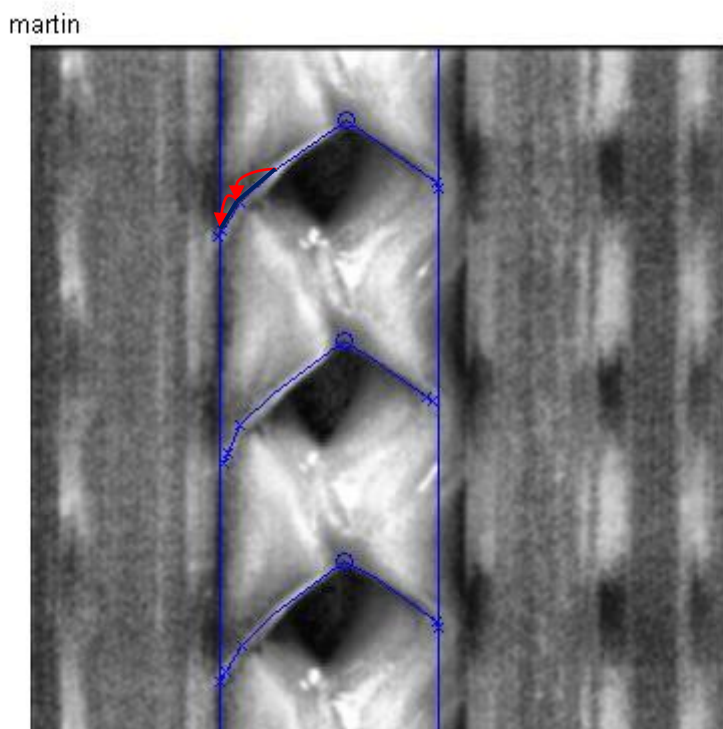
## Iterace

Protože se charakter vlny v jejím průběhu postupně mění, není zpravidla možné jedním průchodem algoritmu nalézt vlnu v celé její délce. Z tohoto důvodu je třeba celý postup iterovat.

V každém kroku iterace je nejprve spočítána metodou maskované vzájemné korelace korelační matice a v ní je poté nalezena extrémní pozice výše zmíněnou metodou. Na základě těchto hodnot je stanovena pozice pro určení hodnot nového korelačního jádra. Pro další hledání extrému v nové korelační matici je také přepočítán úhel, pod kterým vlna v dané iteraci postupuje.

Tento proces pokračuje tak dlouho, dokud se buď vlna nestočí tak že je rovnoběžná s osou y směrem dolů, čímž indikuje nalezení amplitudy, nebo dokud nedojde ke konvergenci řady nalezených extrémů, tj. vlna končí v okolní tkáni a její návrat není z obrázku zřetelný, nebo úplně chybí.

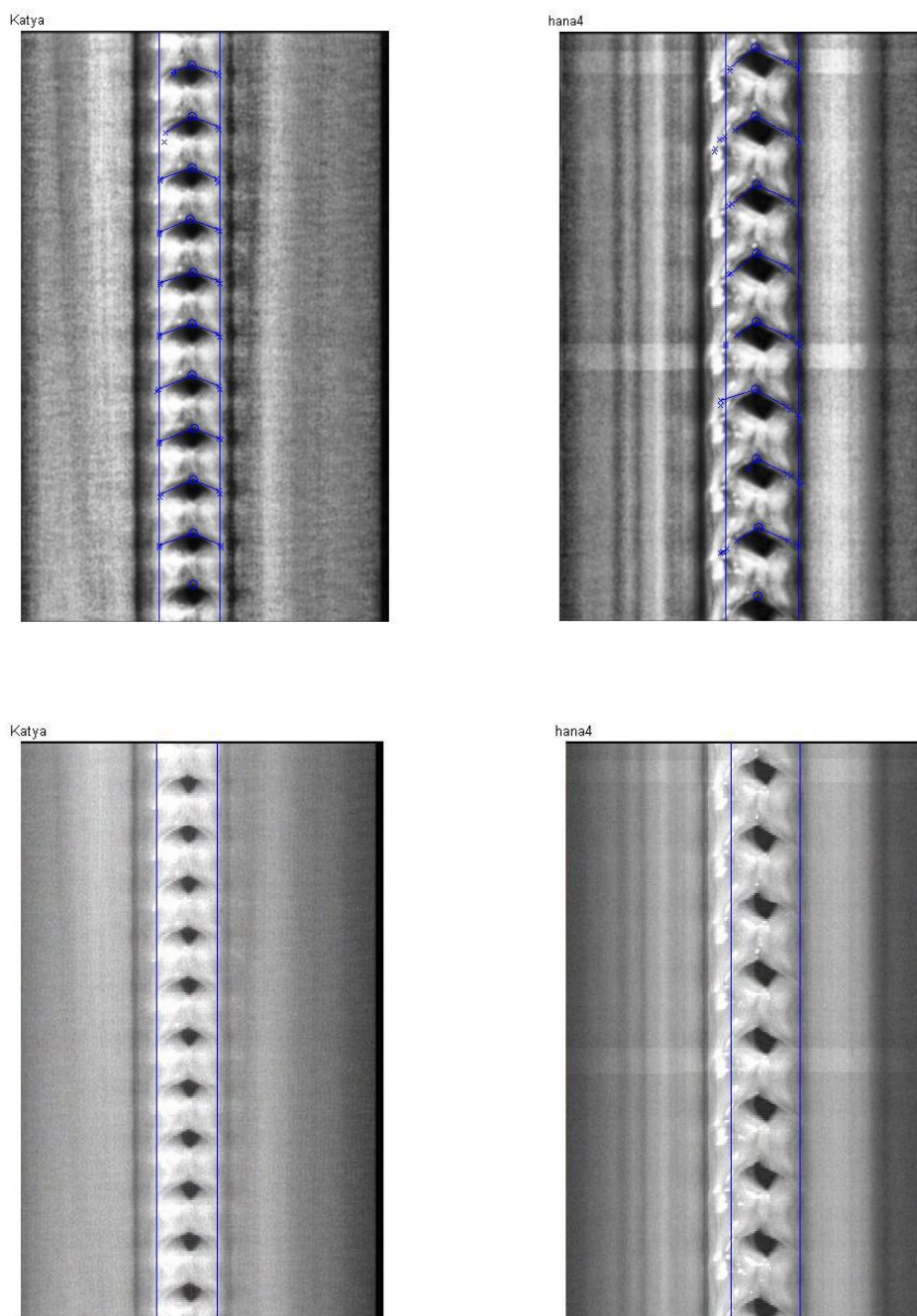
Tímto postupem je přibližně určen tvar a průchod vlny v záznamu a stanoveny hodnoty důležité pro určení charakteristik slizniční vlny (Obrázek 12).



Obrázek 12. Iterace. Tento aktuální snímek ukazuje postup hledání extrémů.

#### 4.2.4 Výsledek

Popsaným algoritmem jsou spočteny úhly a amplitudy pro každou instanci otevření hlasivkové štěrbiny (Obrázek 13). Výsledná hodnota úhlu slizniční vlny daného snímku je určena jako průměr naměřených hodnot. Hodnota levé a pravé amplitudy slizniční vlny je určena jako medián spočtených hodnot z důvodu větší robustnosti algoritmu.



Obrázek 13. Příklady kroků algoritmu a příslušných výstupních obrázků.

### 4.3 Výstupní data

Výstupem navrhovaného algoritmu jsou hodnoty úhlů slizničních vln v radiánech a horizontální hodnoty amplitud v pixelech pro daný snímek. Tyto hodnoty jsou připraveny pro další zpracování. Horizontální souřadnice snímku zachycuje prostorovou dimenzi a lze z její hodnoty pomocí znalosti fyzických rozměrů snímku spočítat fyzickou velikost amplitudy. Vertikální souřadnice snímku odpovídá časové ose, tedy informace o úhlu ukazuje na rychlost šíření slizniční vlny.

## **5 Závěr a otevřené problémy**

### **5.1 Závěr**

Hlavního cíle práce bylo dosaženo. Pomocí popsané metody se mi podařilo automaticky najít pro každý snímek amplitudu slizničních vln a jejich úhel postupu na daných videokymografických snímcích pacientů se zdravými hlasivkami. Metoda úspěšně pracuje i se snímky zhoršené kvality a snímky zatížené odlesky.

### **5.2 Otevřené problémy**

Tato metoda, ostatně jako mnoho metod v oblasti digitálního zpracování obrazových dat, je citlivá na kvalitu vstupních dat a závislá na správném určení parametrů programu. Otevřeným problémem tak zůstává otázka robustnosti algoritmu. V případných rozšířeních tohoto programu se tedy lze zaměřit například na automatické určení příslušných parametrů ze vstupních snímků.

## **Doslov**

Je mou velikou nadějí, že tato práce povede k lepšímu porozumění funkce lidského hlasového orgánu zejména také proto, že i tento orgán podléhá tvorbě zhoubných rakovinných tumorů a včasná diagnostika je základem úspěšné léčby této lidstvo sužující choroby.

## Seznam použité literatury

1. **Deliysky, DD, et al.** Clinical Implementation of Laryngeal High-Speed Videoendoscopy: Challenges and Evolution. *Folia Phoniatica et Logopaedica*. 2007.
2. **Hauzar, D.** Zpracování digitálních snímků videokymografických záznamů jako podpůrný nástroj pro diagnostiku hlasivek. *Diplomová práce*. Praha : MFF UK, 2010.
3. *Fyziologická akustika zpěvního hlasu: Nový pohled na starý problém.* **Švec, Jan G.** Kouty : s.n., 2000. 60. akustický seminář & 36. akustická konference.
4. **Alberti, PW.** The history of laryngology: a centennial celebration. *Otolaryngol Head Neck Surg*. 1996, Vol. 114, pp. 345-354.
5. **Mafee, MF, Valvassori, GE and Becker, M.** *Imaging of the neck and head*. 2. Stuttgart : Thieme, 2005.
6. **Uloza, V, Saferis, V and Uloziene, I.** Perceptual and acoustic assessment of voice pathology and the efficacy of endolaryngeal phonomicrosurgery. *J Voice*. 2005, Vol. 19, pp. 138–145.
7. **Gallivan, KH and Gallivan, GJ.** Bilateral mixed laryngoceles: simultaneous stroboscovideolaryngoscopy and external video examination. *J Voice*. 2002, Vol. 16, pp. 258–266.
8. **Verikas, A, et al.** Advances in laryngeal imaging. *Eur Arch Otorhinolaryngol*. 2009, Vol. 266, pp. 1509-1520.
9. **Švec, Jan G and Schutte, HK.** Videokymography : high-speed line scanning of vocal fold vibration. *J Voice*. Vol. 10, pp. 201-205.
10. *Videokymography in 2000: the Present State and Perspectives of the High-Speed Line-Imaging Technique.* **Švec, JG, Šram, F and Schutte, HK.** Jena : s.n., 2000. Advances in Quantitative Laryngoscopy, Voice and Speech Research. pp. 57-62. ISBN:3-00-005636-X.
11. **Gonzalez, RC and Woods, RE.** *Digital image processing*. Upper Saddle River : Prentice-Hall, 2008. ISBN:978-0-13-168728-8.
12. **Kass, M, Witkin, A and Terzopoulos, D.** Snakes: Active Contour Models. *International Journal of Computer Vision*. 1988, pp. 321-331.
13. **Kunihiro Okada, et.al.** *Digital Differential Analyzer*. 4106100 USA, 8 8, 1978. Electrical Computers: Arithmetic Processing And Calculating.



14. **The MathWorks, Inc.** *Matlab: The Language of Technical Computing*.  
<http://www.matlab.com> : The MathWorks, Inc., 2009.

## Seznam použitých zkratek

- CCD - Charged Coupled Device – Je elektronická součástka pro snímání obrazové informace.
- CT - Computed Tomography – Výpočetní tomografie je radiologická vyšetřovací metoda, která za pomoci rentgenového záření umožňuje zobrazení vnitřních orgánů.
- DDA - Digital Differential Analyzer – Metoda postupu po pravidelné mřížce pomocí počítání diferencí.
- LoG - Laplacian of Gaussian – Filtr používaný v digitálním zpracování obrazových dat, který vznikne kombinací laplaceovského filtru (derivace) s gaussovským filtrem (rozmazání).
- MRI - Magnetic Resonance Imaging – Magnetická resonance je zobrazovací metoda, která pomocí silného magnetického pole a elektromagnetického vlnění s vysokou frekvencí umožňuje zobrazení vnitřních orgánů těla.
- VKG - Videokymogram neboli videokymografický záznam. Vzniká snímáním lidských hlasivek v místě největší amplitudy speciální řádkovou kamerou.

# **Příloha 1**

## **Obsah CD**

Na přiloženém CD se nachází následující soubory:

- Doc/Bakalarska prace.pdf ... elektronická verze bakalářské práce
- Program/... ... distribuce programu
- Program/Images ... vstupní snímky zdravých hlasivek
- Program/Data ... předpočítané vstupní hodnoty (viz str.14)

## Příloha 2

### Programová dokumentace

#### *Instalace programu*

Distribuce programu obsahuje zdrojové kódy, sadu testovacích snímků a předem spočítané příznaky těchto snímků potřebné pro běh programu, které byly získány s pomocí diplomové práce Mgr. Davida Hauzara.

#### **Systémové požadavky**

Program pro svůj běh potřebuje prostředí Matlab (14) verze r2008a a novější spolu s přídatným balíčkem Image processing toolbox, ve kterém byl také testován.

#### **Instalace**

Program se instaluje zkopírováním zdrojového kódu programu do libovolného adresáře.

#### *Spuštění programu*

Pro spuštění programu je třeba zpustit prostředí Matlab a nastavit pracovní adresář jako adresář, do kterého byla distribuce programu zkopírována. Program lze spustit z příkazového řádku prostředí Matlab jedním z následujících příkazů.

Příkazem `GWProcess()` se spouští zpracování jednoho záznamu, příkazem `RunAll()` sada všech v distribuci přiložených snímků ve vizualizačním modu (viz níže).

Syntaxe zadávání příkazů je následující:

```
$ vysledek = GWProcess([název snímku])  
$ vysledek = GWProcess([název snímku], [th, masksize])  
$ vysledek = GWProcess([název snímku], [th, masksize],  
                        [mode])  
$ RunAll()  
$ RunAll([th, masksize])
```

Kde *vysledek* je vektor o čtyřech prvcích [*LAmp* *RAmp* *LAng* *RAng*] takových, že :

- *LAmp*, resp. *RAmp* udává velikost levé, respektive pravé amplitudy v pixelech od středu otevření štěrbiny;
- *LAng*, resp. *RAng* průměrnou hodnotu získaných úhlů levých, respektive pravých slizničních vln v radiánech v levotočivém systému s počátkem směřujícím na snímku dolů.

Jako *název snímku* v parametru funkce se udává jméno JPG souboru obsahujícího příslušný kymografický snímek bez přípony.

*th* (nepovinné) – nastavení prahové hodnoty korelační matice určené pro hledání postupu slizniční vlny. Povolené hodnoty jsou desetinná čísla z intervalu (0,1)

*masksize* (povinný údaj při použití *th*) – horní mez pro poloviční velikost korelačního jádra. Smysluplné jsou celočíselné hodnoty z intervalu (3,10)

*mode* (nepovinné) - způsob výstupu dat programu. Povolené hodnoty jsou:

- 'display' – zobrazí grafický výstup získaných amplitud
- 'debug' – použití pouze pro účely ladění programu.

Příklad použití:

```
$[a b c d] = GWProcess('Martin');
```

## Příloha 3

### Architektura programu

Program se skládá z několika souborů obsahujících jednu nebo více tematicky podobných funkcí.

- `GWProcess.m` - Hlavní tělo výpočtu se nachází v souboru `GWProcess.m` spolu s dvěma podprogramy – `ComputAngle()` pro výpočet úhlu z páru vektorů a `CorrelationWaveTracking()` pro iterativní normalizovanou korelaci. Tento soubor dále obsahuje hlavní cyklus procházející jednotlivá otevření na vstupním snímku.
- `CreateMask.m` - Podprogram generující korelační masku podle zadaného úhlu.
- `CrossCorrelation.m` - Podprogram výpočtu maskované normalizované vzájemné korelace.
- `FindExtrema.m` – Podprogram vyhledání extrémního bodu pomocí upraveného DDA algoritmu.
- `FindLeftExtrema.m` / `FindRightExtrema.m` – Rutiny hledání prvních instancí výskytů extrémů z korelačních matic.
- `RunAll.m` – Pomocný program automatického zpracování všech dostupných dat.

Datový tok programu - schéma:

