

## Oponentský posudek na disertační práci

**Doktorand: Ing. Tomáš Ettler**

**Název práce: Glottis Detection and Evaluation in High-Speed Video Recording**

**Školitel: Ing. Pavel Nový, Ph.D.**

**Oponent: doc. PhDr. Ing. Jaroslav Průcha, CSc., Ph.D.**

Posuzovaná disertační práce patří v širším slova smyslu do oboru zpracování a analýzy biomedicínských signálů a dat s cílem přispět informacemi v rámci rozhodovacích procesů k diagnostice hlasového ústrojí člověka. Práce je konkrétně zaměřena na studium fyziologie hlasivek a analýzu jejich dynamického chování na základě obrazového signálu z vysokorychlostní laryngoskopie (Laryngeal High-Speed Videoendoscopy, LHSV).

Práce má bohatý, pro disertační práci příkladně akceptovatelný rozsah 144 stran a je logicky členěna do 9 kapitol, včetně úvodu a závěru. Práce se odvolává na přiměřený počet 56 referencí, které jsou v citacích v textu plně využity. Tématicky lze práci rozdělit do tří částí:

- analýza obrazové informace s výslednou segmentací hlasivkové štěrbiny (glottis),
- výpočet množiny parametrů glottis a jejich vizualizace a následná statistická analýza,
- řešení úlohy klasifikace pomocí vybraných dvojic parametrů, která vede na odhad ohodnocení dynamického chování hlasivek na principu korelačních vazeb.

Disertační práce - kromě účelného řešení aktuální vědecko-výzkumné problematiky - představuje rovněž ucelený a přiměřeně validovaný vývoj nástroje (zdravotnického prostředku) pro efektivní podporu pro lékařské diagnostické rozhodování.

### **a) Zhodnocení významu disertační práce pro obor**

Hlasivky jsou párový orgán umístěný v hrtanu a umožňující vznik hlasu. Jsou tvořeny hlasivkovými vazy (zesílenými fibroelastickými membránami) pokrytými sliznicí a ovládanými vříví řízenými hlasivkovými svaly. Vzduch, který při výdechu prochází štěrbinou mezi hlasivkami, různě napjaté hlasivky rozechvívá, čímž vzniká hlas. V procesu tvorby hlasu se klíčovým způsobem uplatňuje právě štěrbina mezi hlasivkami, tzv. glottis. Hlasivky jsou často napadány řadou chorob, jejichž diferenciální

diagnostika je obtížná. Přitom poměrně časté a velice nebezpečné jsou zhoubné novotvary hrtanu (dg. C32), zasahující hlasivky s prevalencí 4 až 5 tisíc případů v ČR. Jedná se nejčastěji o zhoubné nádory epitelu hrtanu. Dle umístění nádoru v hrtanu se popisují jako rakovina glotická, supraglotická a méně častá rakovina infraglotická. Vznikající novotvar se charakteristicky projevuje právě v oblasti hlasové štěrbiny – glottis. Poněvadž jsou hlasivky poměrně dobře přístupné pohledovému vyšetření, resp. obrazové diagnostice, zdálo by se, že diferenciální diagnostika bude jednoduchá. Ovšem není tomu tak, poněvadž řada dalších, např. zánětlivých postižení hrtanu se může – zejména v počátečních fázích onemocnění - okulárně projevovat obdobně jako iniciální fáze zhoubného novotvaru. Přitom včasná diagnostická zachycení nádorového onemocnění hrtanu je jedinou šancí na účinnou léčbu pacienta s potenciálně přijatelným výsledkem. O všech třech typech zhoubných nádorů hrtanu platí, že nejsou-li zachyceny v zárodku, pak rostou invazivně.

Za této situace je vysoce účelné snímání funkce hlasivek vysokorychlostní endoskopií hrtanu (Laryngeal High-Speed Videoendoscopy, LHSV) s následnou automatizovanou detekcí změn dynamiky glottis.

Disertační práce má tudíž pro vědecko-výzkumné a vývojové aplikace oboru informatiky a výpočetní techniky zcela klíčový význam. Práci rovněž považuji za kruciólní příspěvek k oboru biomedicínského inženýrství i k oboru samotné klinické medicíny.

#### **b) Vyjádření k postupu řešení problému, použitým metodám a splnění stanoveného cíle**

Disertační práce se zabývá analýzou obrazového signálu získaného z videozáznamů LHSV s cílem detekovat anomálie v dynamickém chování při otevírání a zavírání hlasové štěrbiny - glottis. K tomu slouží detailní analýza zvolených parametrů, které mají geometrický charakter. V oboru zpracování obrazové informace s cílem segmentovat glottis byla již v tomto směru publikována řada metod. I přes tento fakt autor disertační práce našel v této oblasti prostor pro zcela nová a neotřelá řešení detekce oblasti zájmu (Region of Interest, ROI), a to pomocí frekvenční analýzy jasových změn obrazových bodů, které ve snímcích LHSV představují části anatomických struktur.

Autor práce rovněž originálním způsobem řeší vlastní segmentaci glottis, kde disruptivně uplatňuje pro danou úlohu kreativně modifikovanou shlukovou analýzu. Podobně neotřelým a kreativním způsobem selektuje autor práce determinující příznaky v podobě množiny parametrů, které pak dále používá pro hodnocení dynamického chování hlasivek. Jedná se konkrétně o soubor parametrů odvozených od polohy těžiště glottis. Tyto parametry pak mají význam nejen na úrovni vizualizace trajektorie pohybu těžiště, viz např. kasuistiky uváděné v kap. 6.3.3, ale rozhodujícím způsobem se uplatňují rovněž v samotné klasifikační úloze.

Za zcela původní přístup pak považuji analýzu parametrů glottis pomocí korelačních vazeb, viz kap. 7.2, jakož i realizaci klasifikačních algoritmů, viz kap. 7.3, vycházejících z klasifikačních funkcí sestavených pro vybrané parametry glottis, přesněji řečeno pro dvojice parametrů vyznačujících se významnými korelacemi. Takto vytvořený klasifikační systém se učí - trénuje pod dohledem, což autor práce efektivně řeší diskriminací daného pohybu glottis do odpovídající třídy s použitím videozáznamů LHSV, u kterých je ohodnoceno dynamické chování hlasivek expertem ORL.

Vzhledem k tomu, že v současné době byla publikována řada nových původních prací, které se týkají celé problematiky detekce glottis ve videozáznamech LHSV, lze považovat toto téma za vědecko-technicky i medicínsky a též společensky aktuální, zároveň však vyznačující se zcela unikátním postupem. Jedná se o téma interdisciplinární a tím i velmi náročné, leč zároveň vysoce přínosné. Při detailním rozboru předložené disertační práce dospívám k závěru, že doktorand si ve své práci stanovil tři tématické cíle, ke kterým pak relevantně směřuje a postupně je v plné míře dosahuje:

1. Sestavení vhodných sekvencí metod, které vedou k úspěšné segmentaci glottis v každém snímku videosekvence LHSV, spec. pak v případech videosekvencí „horší“ kvality.
2. Analýza kinematiky hlasivek, protože pouhé pozorování videozáznamů LHSV, nebo jiných výsledků vyšetření, nemusí vždy postačovat. V souvislosti s tímto cílem bylo významným činem autora též sestavení souboru parametrů, které mají geometrický základ a využívají osu symetrie hlasivek, resp. glottis, jakož i zkoumání možností nástrojů statistické analýzy pro popis vzájemných vztahů (korelačních vztahů) na základě předpokladu, že většina geometrických parametrů by měla mít silný vztah při standardním chování hlasivek, přičemž porušení tohoto korelačního vztahu může být indikátorem patologického stavu hlasivek, tj. může být významným faktorem z diagnostického hlediska.
3. Vytvoření klasifikačního aparátu pro odhad kvality kinematiky hlasivek, které vychází z obrazové informace extrahované z videozáznamů LHSV, které jsou pro účely klasifikace ohodnoceny expertem ORL, a souborů vybraných parametrů glottis s významnými korelacemi. Cílem tohoto přístupu je při vyšetření hlasivek účinně varovat v případech, kdy jsou zjištěny nesrovnalosti nebo neočekávané korelačních hodnoty.

Formulované cíle práce považuji za plně disertabilní, a to na nadstandardně vysoké úrovni. Z textu práce je patrné, že doktorand projevil značně koncepční přístup spojený s dokonalým přehledem v oboru fyziologie hlasového ústrojí, hlubokými znalostmi v oboru zpracování obrazové informace a statistické analýzy a dokonalou erudicí v informatice a využití výpočetní techniky. Rovněž je zřejmé, že autor účelně prostudoval a analyzoval velké množství odpovídající literatury a komplexně zvládl celou složitou interdisciplinární problematiku.

Při psaní textu doktorand postupuje systematicky, používá přesné a srozumitelné formulace, účelně využívá bohatou rešeršní práci (zejména v části, týkající se způsobů zpracování obrazové informace), kreativně navrhuje neotřelé, potenciálně účinné metody. Zvláště oceňuji, že jím navržené unikátní metody rovněž prakticky implementuje a testuje na značně rozsáhlém souboru videozáznamů LHSV. S údivem oceňuji, že doktorandem uplatněný korpus zahrnuje 692 záznamů. Autor svoji práci navíc doplňuje kazuistikami, na kterých demonstruje nejen úspěchy vyvinutých metod, ale také případy jejich selhání, což je zvláště cenné a obecně méně časté. To vše dokazuje, že autor disertační práce musel provádět velmi mnoho časově vysoce nadstandardně náročných experimentů a podrobně analyzovat příčiny, kdy jednotlivé metody nedosahovaly očekávaných výsledků.

Mohu jednoznačně potvrdit, že stanovené cíle byly splněny.

#### **c) Stanovisko k výsledkům disertační práce a k přínosu doktoranda**

Doktorand ve své práci pečlivě dokumentuje dosažené výsledky pro jednotlivé cíle práce a použité metody. Na základě toho, že autor pracuje s tak rozsáhlým a heterogenním souborem vstupních obrazových dat, výsledky dokládá přehledovými statistikami a relevantně výsledky analyzuje, lze konstatovat, že dosažené výsledky a výstupy práce jsou nezpochybnitelné. Jednotlivé použité postupy jsou původní a společně s již zmíněnou detailní analýzou výsledků jsou přínosem pro obor.

#### **d) Vyjádření k systematicce, přehlednosti, formální úpravě a jazykové úrovni práce**

Ani zde není možno vznést žádné námitky. Jazyková úroveň anglického textu i vložených českých textových částí je skvělá, bez gramatických chyb, s dokonalou srozumitelností. Snad bych pouze doporučil spojky a předložky, zvláště ty o malém počtu písmen a úzce spojené s dalším textem, nenechávat na konci řádek, ale přesouvat je na řádku další, tak jak je běžným zvykem. Formální úprava textu je dokonalá, grafy, tabulky, diagramy a obrázky názorné, srozumitelně popsání, s odvolávkami na správných místech textu. Dalšímu zdokonalení formální stránky práce by prospělo rovněž uvedení samostatného přehledu obrazové a grafické dokumentace. Veškeré zkratky jsou náležitě vysvětleny. Veškerá uváděná literatura je na vhodných místech textu smysluplně citována.

#### **e) Vyjádření k publikacím doktoranda**

V seznamu publikovaných prací doktoranda je uvedeno 7 prací, které se všechny přímo vztahují k tématu disertační práce. Dvě práce jsou kvalifikační, tři konferenční (2x SCOPUS, 1x WoS) a dvě

časopisecké (2x SCOPUS, 1x WoS). Čtyři práce publikoval se svým školitelem, jednu v širším autorském kolektivu. Podle mého názoru je publikační aktivita doktoranda, a to i vzhledem ke kvalitě publikovaných výsledků a jejich citacím, nadstandardní.

**f) Jednoznačné vyjádření oponenta, zda doporučuje či nedoporučuje disertační práci k obhajobě**

Předložená disertační práce splňuje všechny požadavky, které jsou kladeny na tento typ kvalifikační práce, tj. obsahuje původní výsledky a je samostatnou tvůrčí vědeckou prací s náležitými vědeckými přínosy. Proto práci zcela jednoznačně a bez námitek k obhajobě doporučuji.

**Dotazy do diskuse:**

Na str. 17, obr. 2.8 (*Overview scheme of examination methods used at the ENT clinic of the University Hospital Pilsen*) je uvedena řada vyšetřovacích technik hlasového ústrojí. Byla data, získaná pomocí těchto vyšetření (viz také tab. 3.4 str. 24) přímo využita v rámci řešení disertační práce?

V Praze dne 15.12.2022

doc. PhDr. Ing. Jaroslav Průcha, CSc., Ph.D.

Společné biomedicínské pracoviště FBMI ČVUT a 1. LF UK Praha  
Studničkova 7/2028, 120 00 Praha 2 - Albertov

doc. PhDr. Ing. Jaroslav Průcha, Ph.D.

Digitally signed by doc. PhDr. Ing. Jaroslav Průcha, Ph.D.  
Date: 2022.12.15 21:03:39 +01'00'



## Review report on PhD Thesis

Student: **Tomáš Ettler**  
PhD thesis title: **Glottis Detection and Evaluation in High-Speed Video Recording**  
Faculty: **Faculty of Applied Sciences, University of West Bohemia**  
Supervisor: **Ing. Pavel Nový Ph.D.**  
Reviewer: **doc. Ing. Jana Kolářová, Ph.D., Department of Biomedical Engineering, FEEC, Brno University of Technology**

### Topicality of doctoral thesis:

Doctoral thesis is focused on Glottis Detection and Evaluation in High-Speed Video Recording. Evaluation of glottis function is a very actual and important topic. The submitted text contains few parts about the anatomy of the glottis, methods for image processing used for segmentation of the area of interest and its parametrization. Last part is focused on classification methods, which can help to recognised glottis irregularities. The thesis contains up to 140 pages of detailed text and refers more than 50 literary sources. There are three defined goals in this work, each of them is described in details with results. The importance of the solved topic follows from the presented context.

### Importance for practice or development of the discipline:

Mr. Ettler made a significant contribution to the development of a recognition of glottis irregularities. The work was based on image processing methods and segmentation. The large dataset was used to evaluate proposed methods. The results for detection of region of interest and for segmentation by cluster analysis were described in details. In this field is possible to continue by next steps that are mentioned in the very last chapter.

### Problem solving and dissertation results:

Throughout the thesis, referencing, data collecting, calculations and analysis are conducted rigorously. The conclusions drawn and the results reported are scientifically sound. In my opinion, the results are beneficial for this scientific field.

### Publications and contribution of the author:

Mr. Ettler was working on this topic since his diploma thesis. Until now he has already published one article as first author in a well-ranked journal (Q1, IF 3,88), five conference papers as first author and one conference paper with his contribution.

### Choice of methods:

The methods were chosen properly and the applications are well described and used correctly.

**Formal adjustment of the thesis and language level:**

The thesis is clearly written and at a high level of English language, in particular for a non-native speaker. The goals of doctoral thesis were defined clearly, understandably and are sufficiently demanding and innovative.

**Questions:**

Did you find another database, which you can use for evaluation of your methods? Did you think about processing of database from University Hospital in Pilsen, add notes from medical experts and publish anonymous database with colleagues for research community?

**Conclusion:**

In summary, my conclusion is that the PhD thesis of candidate Tomáš Ettlér presents original results of large importance and I recommend without hesitation that the candidate is awarded the doctoral degree. This thesis is ready to be defended, in front of respective committee.

V Brně, 24.2.2023

Jana Kolářová

Oponentní posudek na dizertační práci:

**Ing. Tomáš Ettler: Detekce a vyhodnocení vysokorychlostního videozáznamu hlasivkové štěrbiny.**

Školitel: Ing. Pavel Nový Ph.D., Katedra informatiky a výpočetní techniky

Oponent: František Vávra.

Práce má formulovány následující cíle:

1. Zpracování vysokorychlostního záznamu a hlavně extrakci „užitečné“ informace ze snímkových sekvencí. Jako dílčí (ale kardinální) problém se jeví vyčlenění a vymezení té části snímků, ve kterých je obsažena tato informace [Region of Interest (ROI)]. Tj. ztráta většiny při zachování podstatného.
2. Rozbor a (parametrickou) reprezentaci pohybu hlasivek v „ne vždy dokonalých“ snímcích nebo i sekvencích snímků.
3. Návrh klasifikačních postupů, které by umožnily vyjádření z dílčího signálu o běžném či neběžném chování hlasivek jako pre-diagnostickou informaci.

Práci lze rozdělit do několika částí:

- Stručné nastínění anatomie hlasivek a mechanismů tvorby hlasu. Některé postupy vyšetřování, včetně metodik užívaných ORL klinikou FN v Plzni. Detailněji je popsáno vyšetření hlasivek prostřednictvím vysokorychlostní kamery, neboť to je zdrojem informací v práci dále analyzovaných. Součástí tohoto bloku je konkrétní popis datového korpusu (datových korpusů), který byl v předložené práci využit. Samozřejmostí je i rozbor kvality záznamů.
- Další částí je extrakce hlasivkové části obrazů (ROI). V ní je obsažen přehled jinde (jinými autory publikovaných) používaných metod. Důležitým a přínosným segmentem je popis vlastních postupů pro takovou extrakci. Ty jsou jedním z hlavních výsledků práce, neboť podmiňují další zpracování bez výrazných redukcí v disponibilních zdrojových záznamech. Předložený autorův postup(y) je zapotřebí zvýraznit, neboť nepřímo prokazuje vysokou profesionalitu, široké oborové znalosti a schopnosti. Prezentované algoritmy extrakce jsou dokumentovány pro regulární a i iregulární snímkové situace – včetně uvedení příčin. V práci jsou obsaženy i popisné (převážně



četnostní) statistiky „úspěšnosti“ pro jednotlivé dílčí korpusy. Ty dokládají, že se jedná o reálné výsledky při aplikacích na reálné záznamy (nikoliv oněch „známých publikačních“ zprofanovaných 99 %, ...). Jako nepřímý důkaz autorových širokých znalostí a schopností je například v této části uvedené využití shlukovacích technik pro zmíněný účel. Taková aplikace v dané oblasti není běžná a je jedním z originálních projevů.

- Tato část je následována zavedením parametrické reprezentace obrazů a pohybů hlasivek. Z daného postupu záznamů přirozeně vyplývají geometrické reprezentace. To, co je obecně chápáno jako „přirozené“, se stává v oblasti záznamových sekvencí náročným technickým problémem. O tom třeba svědčí i to, že většina komerčních aplikací se například „vyhýbá“ automatické detekci „os symetrií“. Pokud jsou využity, vyžadují jejich „ruční doplnění“. Samotná problematika „os symetrií“ je náročná v každém jednotlivém snímku, o jejich sekvencích nemluvě. Autor tento problém vyřešil, včetně „měření kvality“ vlastní automatické detekce. Pro taková „měření kvality“ využil srovnání s expertním „ručním výběrovým“ určením. Automatická detekce osy symetrie (os symetrií u posloupnosti snímků) umožňuje vedle klasických obrazových parametrů (plocha, obvod, poloha těžiště plochy, poloha těžiště obvodu, změna plochy v posloupnostech, ... ) zavést parametry odvozené od polohy osy symetrie (vlevo od, vpravo od, vzdálenost od, ...). Autor za přirozenou reprezentaci považuje záznam hodnot jednotlivých parametrů v čase (daném pořadím ekvidistantních snímků). Vlastní zavedení takových parametrických reprezentací (trajektorií) autor dokládá příklady běžné a patologických hlasivkových situací.
- V předchozí partii zpracované reprezentace jsou bez dalšího vlastně nárůstem objemu informačních projevů. To je „korigováno“ následujícím segmentem, kde je zavedeno „měření podobnosti“ jednotlivých trajektorií. Pro něj je využit klasický výpočetní algoritmus výběrových korelačních koeficientů. Je zde doložen a analyzován jejich geometrický význam, jako míry lineárního vztahu (= podobnost v úměrné obměně a orientaci měřítka a podobnost v posunutí polohy). Pro 14 vybraných „geometrických (a odvozených) parametrů“ je uskutečněno jim odpovídajících 91 „porovnání“. Při posouzení konkrétního záznamu jsou pak označeny páry s významně vysokou (v absolutní hodnotě) hodnotou „korelace“ nebo opačně s významně nízkou. Tím je

získána „užitečná“ informace sestávající z několika (označených = labeled) podstatných párů.

- Následuje část, ve které je předchozí aparát využit v klasifikačních postupech. Primární data byla rozdělena do dvou skupin a znalecky ohodnocena. První skupina byla využita jako „trénovací“ a druhá jako testovací. Jednotlivé „korelační míry podobnosti“ byly pro účely klasifikací transformovány tak, aby bylo dosaženo „lepší skupinové oddělitelnosti“. Na tomto místě je nezbytné zdůraznit autorovy široké znalosti, když použil Fisherovu transformaci, která je běžná k zcela jinému účelu. Presentované klasifikační postupy jsou založeny na dvou základech. První, srovnáním polohy mediánu třídy a mediánu jejího doplňku. Protože se autorovi patrně zdá tento přístup příliš redukující (v reprezentacích třídy a jejího doplňku jen mediány) zvolil ještě alternativu. Tu založil na „rozdílových plochách“ mezi empirickými distribučními funkcemi hodnot v třídách a jejich doplňcích jako rozšíření mediánových popisů polohy. Srovnáním ale dospěl k závěru, že obě pojetí jsou prakticky (klasifikačně) rovnocenná. Přístup prostřednictvím „rozdílových ploch mezi EDF“ je třeba označit zajímavým a originálním. Jako u všech předchozích bloků jsou i zde prezentovány výsledky pro „normální a patologické hlasivkové stavy“.

V tomto oponentním posudku je zapotřebí autorovi vytknout jím zvolené pojetí souvislého a téměř jednotného (i když strukturovaného) textu. Taktéž snaha po dokonalosti a (v textu uváděném) důsledném dokladování může být kontraproduktivní. Lze předpokládat, že by forma typu základní text a odkazované přílohy byla přehlednější a snadněji čitelná.

### Závěr a shrnutí:

- Předložená a oponovaná práce má velký význam pro daný obor. Některé postupy a výsledky v ní uvedené jsou využitelné i mimo něj.
- V práci uvedené cíle jsou splněny, rozsahem a kvalitou většinou překročeny. Práce obsahuje i kritický pohled na jejich prvotní formulace. Použité postupy řešení problémů jsou adekvátní cílům práce, často formulaci cílů výrazně překračují.
- Práce je velmi přínosná s originálními výsledky a postupy.

- Práce je napsána systematicky a v textu strukturovaně. Rozsah, uvádění souvislostí a snaha po dokonalém dokladování někdy komplikuje její přehlednost. Je možné, že by strukturace typu základní text a odkazované přílohy byla přístupnější.
- Publikace autora odpovídají oblasti, ve které působí. Ty, které jsou mi známy, významně rozšiřují téma dizertační práce a lze vyslovit tvrzení, že výrazně ovlivní daný obor.
- Práci doporučuji k obhajobě.

V Plzni

28.11.2022



František Vávra.