

# Posudek na disertační práci Ing. Ladislava Ptáčka

## Automatická identifikace a verifikace ptáků

**Doktorský studijní program, Studijní obor: Kybernetika  
Západočeská univerzita v Plzni, Fakulta aplikovaných věd**

Pan Ing. Ladislav Ptáček předložil disertační práci na téma „Automatická identifikace a verifikace ptáků (Automatic bird identification and verification).“ Zvolené téma disertační práce patří do oboru disertace a je aktuální z hlediska současného stavu vědy.

Obsahem disertační práce je návrh a evaluace nových metod pro automatickou identifikaci a verifikaci ptáků. Využívá živé nahrávky (51 audiogramů, pro 47 x druhů ptáků a 4 x agregované – viz příloha disertační práce) bez nutnosti jejich předzpracování. Identifikace ptáků je možná bez nutnosti kroužkování nebo kontroly DNA. Součástí disertační práce je návrh banky filtrů, která je optimalizovaná na ptačí zpěv - Bird Adapted Filter, BAF (výzkum je zacílen především na budníčka, ale je testován i na jiných ptačích druzích). Nové metody pro rozpoznání a identifikaci ptáků a jiných zvířecích druhů (v práci jsou konkrétně uvedeny rypoši, kteří k dorozumívání používají poměrně složitý systém vokalizace) byly ověřeny experimentálně.

Práce je psána anglicky (úroveň jazyka nechci posuzovat, nejsem rodilá mluvčí), je členěna do 13 kapitol (včetně úvodu, závěru a přílohy) a je doplněna shrnutím (česky a anglicky), seznamem obrázků (84 položek), tabulek (28 položek), dále seznamem zkratk, 93 položkami citované literatury (včetně 7 prací, jejichž je autorem nebo spoluautorem). Práce obsahuje samostatný přehled autorových publikací a dalších aktivit. Rozsah disertační práce je 152 stran. Na základě přehledu publikací uvedených v disertační práci a v tezích mohu konstatovat, že podstatná část disertační práce byla publikována na odpovídající úrovni. Za velmi užitečné považuji také velké množství citací literárních pramenů.

Disertační práce je členěna do dvou celků. První z nich seznamuje čtenáře se současným stavem výzkumu v ornitologii a v oblasti rozpoznávání mluvcího, ve druhé autor popisuje použité technické prostředky.

V úvodu je stručně uvedena motivace k řešení této tematiky včetně stručných zmínek o podobných úlohách a použitých metodách s procentuálním vyjádřením přesnosti identifikace. Je uveden požadavek ornitologů na přesnost identifikace vyvíjeného automatického systému (nejméně 70%). V druhé kapitole si disertant vytýčil 6 cílů disertační práce (2 hlavní cíle a 4 související dílčí cíle).

Ve třetí kapitole je popsán současný stav výzkumu v této oblasti. Zabývá se zaměřením ornitologie, vztahy mezi lidskou řečí a zvuky zvířat. Část textu je věnována ptákům-pěvcům (autor je dále rozděluje na „křiklavé“ a „zpěvní“ ptáky), jejich sluchu, vysvětlení audiogramu u lidí a ptáků, včetně jejich porovnání. Zde si je autor zcela vědom chyby, která při porovnání vzniká z důvodu nerovnoměrnosti počtu posuzovaných subjektů. U zpěvu ptáků rozlišuje druhy vokalizace, zabývá se popisem vokálního traktu. Vychází z metod používaných pro rozpoznávání lidského hlasu, porovnává proto způsob rozpoznání hlasu u lidí a zpěvu ptáků, dává do souvislosti hmotnost zvířat s frekvencí jejich hlasových projevů, zabývá se šířením ptačího zpěvu a problematikou kroužkování.

Čtvrtá kapitola je věnována použitým metodám výzkumu. Byl vytvořen programový systém v Matlabu (není však uvedena verze, která byla používána, i když se od sebe liší), část kódu byla napsaná v jazyce C<sup>++</sup>. Součástí programu byly moduly pro jeho správu, klasifikaci nahrávek, extrakci proměnných, dále přídatné moduly (VAD, BAF a slučování dat), moduly s metodami odhadu, ověření a vyhodnocení experimentů (byly uloženy jako excelovské soubory).

Od páté kapitoly do jedenácté se autor věnuje vlastním experimentům (byly realizovány tři), jejich rozboru a vyhodnocení na základě nahrávek ptačího zpěvu. Nahrávky byly rozděleny náhodně do dvou skupin (ty byly použity pro první dva experimenty, ve třetím experimentu použil autor všechny nahrávky). Každý experiment se skládal z tréninku zpěvu známých ptáků na základě GMM (Gaussian Mixture Model), z UBM (Universal Background Model) a z testování, pro které byly použity nahrávky zpěvu neznámých ptáků. Jednotlivé nahrávky byly pořízeny v různých časových odstupech, obsahují mnoho ostatních nežádoucích zvuků, např. dopravní ruch, zvuky zvířat a zpěv nesledovaných ptáků, různé úrovně zpěvu (v závislosti na vzdálenosti a okolním prostředí). Konečné zhodnocení dosavadního výzkumu a výhled do budoucna je obsahem závěrečné kapitoly.

Po formální stránce je předkládaná práce na dobré úrovni, i když kvalita grafů je rozdílná, některé jsou černobílé, resp. ve stupních šedé, jiné jsou barevné. Nejspíš jako součást článků v časopisech a příspěvků na konferencích. U některých grafů chybí hodnoty na osách, k popisům os jsou použity různé fonty. To ztěžuje možnost porovnání (obr. 7.2, str. 77). Texty v některých obrázcích jsou nečitelné.

Mezi klady disertační práce patří matematické odvození a popis audiogramů, které se dosud používaly pouze v grafické formě, a vytvoření databáze ptačího zpěvu. K práci mám následující připomínky a dotazy:

Připomínky:

- Některé obrázky, které obsahují tabulky jsou špatně čitelné (jsou příliš malé), např. obr. 4.3., 4.4., 4.6., 7.5., 10.2., 10.3., 10.4., 10.5., 10.6. a obrázky v dodatku.
- Některé obrázky jsou barevné, jiné v odstínech šedi.
- Nepovažuji za vhodné rozdělení tabulky do dvou stránek (tab. 10 na str. 66-67, tab. 27 na str. 108-109) nebo rozdělení obrázku a jeho popisu (obr. 10.5, str. 103-104).
- V tab. 13 na str. 68 nesouhlasí udané hodnoty procent s hodnotami v textu.

Dotazy k práci :


- Při experimentech se zaváděním elektrod do těla ptáků jste pracoval samostatně (musel jste tedy používat znalosti i z jiných než technických oborů) nebo Vám pomáhal v této části ornitolog?
- Vysvětlíte, co je „as-is recordings“.
- Podle čeho jste určil potřebný počet nahrávek a jejich rozdělení do trénovacího a testovacího souboru?
- Upřesněte, co znamená velmi nízká, nízká, střední a vysoká úspěšnost (tab. 8, str. 66).

Na závěr konstatuji, že vytčené cíle byly splněny, předložená disertační práce je zajímavá, odpovídá kvalitě, která je požadovaná pro tento typ práce. Vytvoření metod pro rozpoznání a identifikaci ptáků bez nutnosti kontaktu s nimi při kroužkování považuji za výrazný vlastní přínos disertanta pro výzkum a aplikace. Za získanými výsledky se skrývá velmi mnoho užitečné práce. Na základě získaných výsledků bude možné v budoucnu ve výzkumu pokračovat.

Předložená disertační práce odpovídá, podle mého soudu, obecně uznávaným požadavkům na udělení akademického titulu Ph.D. Pan Ing. Ptáček projevil schopnost podílet se na výzkumu, z předložené disertační práce vyplývá jeho schopnost samostatně vědecky pracovat.

**Disertační práci doporučuji k obhajobě.**

V Praze 13. 3. 2017

  
Prof. Ing. Jana Tučková, CSc.  
Katedra teorie obvodů, FEL ČVUT v Praze

## Oponentský posudek disertační práce Ing. Ladislava Ptáčka

Ing. Zbyněk Tychtl, Ph.D., Spechtech, s.r.o.

V předložené písemné práci „Automatic Bird Identification and Verification“ autor dokumentuje cíle, průběh a výsledky své vědecké práce při návrhu a vytváření automatického systému pro identifikaci jednotlivých ptačích jedinců podle nahrávek jejich zvuků (zpěvu a volání). Autor si kladl za cíl vyvinout takový systém, který by mohl nahradit ornitology doposud užívanou metodu "kroužkování", při které je nutné ptáky odchytil, označit je kroužkem s jedinečným kódem, který umožní při opětovném odchytu jednoznačně identifikovat jedince. Aby měl nově vyvinutý systém potenciál kroužkování opravdu nahradit, musí fungovat především na "otevřené množině" jedinců (tj. vždy lze očekávat libovolný počet nových jedinců, kteří doposud nebyly pozorováni a nejsou tedy k dispozici jejich vzorové nahrávky), měl by pracovat s nahrávkami zvuků ptáků bez nutnosti jejich ručních úprav (jako např. vyřezávání jen malých částí "slabik", "volání", "písní" ap.), a měl by dosahovat poměrně vysoké úspěšnosti (vyšší než 70%) při identifikaci již dříve pozorovaného (nahráného) jedince i detekci "nového" jedince mimo databázi vzorů (tj. doposud nenahráného).

Cíl práce je velmi ambiciózní v porovnání s jinými publikovanými vědeckými pracemi světové vědecké komunity. Jsou známy práce usilující o identifikaci ptačích jedinců, nebo i jiných zvířecích druhů, které dosahují slušných výsledků, ale vždy jen v uzavřené množině jedinců a většinou je nutné systému předkládat pouze předpřipravené "čisté" nahrávky jen malých částí nahrávek porovnávaného jedince.

Pro dosažení hlavních cílů musel autor najít a zkombinovat celou řadu originálních postupů. Od začátku bylo zřejmé, že bude potřeba nalézt a definovat originální hlasové příznaky, které, byť inspirovány příznaky z rozpoznávání lidské řeči, musí zohlednit odlišné akustické vlastnosti vytváření i vnímání zvuků ptáky. Autor využil současné i historické vědecké poznatky o fyziologii produkce a vnímání zvuků ptáky a přispěl vědecké komunitě návrhem vztahu analyticky modelujícího audiogram ptáků. Na jeho základě autor navrhl banku filtrů (Bird Adapted Filters) s rozložením filtrů optimalizovaným pro ptáky. V práci představil modifikace audiogramu pro 47 různých druhů ptáků a několik agregovaných audiogramů pro např. zpěvné a nezpěvné ptactvo a sovy.

Významným počinem bylo přizpůsobení, natrénování a využití technik pro detekci zvuků v nahrávkách a tím jejich automatického "oddělení" od pozadí (ruchů okolního prostředí). K tomu autor využil metod známých jako voice-activity-detection (VAD) a unified-background-model (UBM). Úspěšné přizpůsobení těchto metod především využitím modifikovaného zmíněného VAD detektoru a statistického UBM modelu autoru umožnilo ve svém systému opravdu používat celé, předem ručně neupravené, nahrávky. Lze tedy ve výsledném systému používat nahrávky tak, jak je ornitologové pořídí poměrně běžným postupem.

Autor ve spolupráci s odborníky z Jihočeské Univerzity v Českých Budějovicích definoval podmínky pro pořizování nahrávek a v reálném venkovním prostředí pořídil nahrávky ptačího zpěvu a volání. Pro natrénování a ověření funkce systému se mu podařilo nahrát během měsíce 14 konkrétních jedinců Budníčka menšího. Ty použil k ověření práce systému. Srovnal úspěšnost řady konfigurací a použitých metod a parametrizací. Veškeré výsledky profesionálně statisticky vyhodnotil a v práci prezentoval. Statistické vyhodnocení profesionální úrovně dokládá validitu prezentovaného výsledku úspěšnosti automatické identifikace jeho systému 78,5%. Takovou hodnotu lze hodnotit jako výborný výsledek. Obzvláště na "otevřené množině" jedinců a bez nutnosti ručně předzpracovávat nahrávky.

Práce je napsána v anglickém jazyce na cca 115 stranách plus přílohy. Shrnuje přehled metod používaných pro verifikaci a identifikaci řečníka (z lidské řeči), shrnuje poznatky o fyziologii vytváření zvuků, jejich vnímání ptáky, zmiňuje používané postupy pro zjišťování dat pro tvorbu audiogramů zvířat, která samozřejmě neodpoví na dotazy..., popisuje odvození analytického modelu audiogramu a adaptované banky filtrů, modifikovaný algoritmus výpočtu vektorů příznaků, vytváření databáze nahrávek, diskutuje využití statistických metod modelování akustiky (HMM, GMM, iVectors,...), popisuje návrh a realizaci celého automatického systému, řízení rozsáhlých experimentů pro hledání optimálních nastavení, srovnává řadu metod, precizně statisticky vyhodnocuje experimenty.

**Přínos pro obor:** Vše nasvědčuje tomu, že autorem navržené metody (příprava audiogramu, BAF filtrů, i konkrétní konfigurace a kombinace metod) se stanou základem vývoje podobných

automatických systémů pro identifikaci druhů i jedinců různých živočichů z jejich hlasů. Ty nahradí nebo výborně doplní jiné doposud používané metody jako kroužkování ptáků, a stanou se inspirací pro další vědce. Autor mimo jiné dokázal, že aktuální velmi pokročilé metody vyvinuté především pro zpracování lidské řeči lze po jistých modifikacích úspěšně a velmi užitečně využít i v jiných oblastech.

**Postup řešení:** Autor se velmi dobře orientuje v aktuálním stavu výzkumu ve světě, svoje metody zakládá na fungujících aktuálních metodách, které je schopen posoudit, hodnotit i testovat. Metody vylepšuje vhodnými úpravami, navrhuje řešení problémů, kombinuje přístupy. Výsledky experimentů, které provádí systematicky, je schopen precizně statisticky zpracovávat a tím dokládat jejich věrohodnost. Prezentovaný postup řešení problému je vhodný.

**Formální stránka práce:** Rozsah práce je dostatečný, vlastní struktura práce je vhodně navržena, vzájemný poměr rozsahu jednotlivých kapitol je vyvážený mezi popisem současného stavu a vlastním vědeckým příspěvkem autora. Velmi pozitivně lze hodnotit úpravu práce jen s minimem typografických chyb a překlepů. Vyzdvihl bych také jazykovou stylizaci práce. Autor se nepouští do zbytečných "povídavých" opisů, které do technické vědecké práce nepatří. Drží se technického, věcně přesného a hlavně úplného popisu, který nenechává čtenáři prostor pro nezodpovězené varianty.

**Publikační aktivita:** Autor ve své práci uvádí seznam 7 publikovaných prací. Vyzdvihnu jednu ve významném časopise Bioacoustics journal v roce 2015. Všechny práce byly publikovány v letech 2015 a 2016, což vzhledem k době, kterou autor na problému pracoval, může vypadat poněkud nevyváženě. Nicméně je pochopitelné, že mohl publikovat výsledky v podstatě až v pokročilé fázi práce, kdy už bylo zřejmé, že několikaleté snažení opravdu vede k cílům, které si autor vytýčil. Proto naopak, publikační činnost dokazuje aktuálnost a přínos autorovy práce. Publikační činnost je dostačující, a její obsah bude jistě inspirací pro řadu dalších vědeckých prací.

**Závěr:** Na základě prostudování předložené disertační práce a vzhledem k jejímu výše uvedenému zhodnocení je možno konstatovat, že disertační práce pana ing. Ladislava Ptáčka splňuje podmínky samostatné tvůrčí vědecké práce a obsahuje původní a autorem publikované nebo k publikaci přijaté výsledky. **Disertační práci doporučuji k obhajobě.**

V Plzni dne 20. března 2017

  
Ing. Zbyněk Tychtl, Ph.D.  
Speechtech, s.r.o.



## Oponentský posudek na dizertační práci Ing. Ladislava Ptáčka

Dizertační práce Ing. Ladislava Ptáčka s názvem „Automatická identifikace a verifikace ptáků“ se z převážné části zabývá možnostmi automatické identifikace jedinců ptáků na základě akustické struktury jejich zpěvu. Automatické rozpoznávání jedinců nebo druhů živočichů na základě jejich hlasových projevů představuje v současnosti velmi zajímavé a diskutované téma. Zvyšující se kvalita a dostupnost přístrojů pro nahrávání hlasových projevů umožňuje získávání velkého množství nahrávek, které mohou být využity k monitorování populací. Výhodou této metody je minimální dopad na studovaná zvířata, naopak její stinnou stránkou je časově náročné zpracování nahrávek. Najít způsob, jak počítačové zpracování nahrávek zjednodušit, aniž by se snížila spolehlivost výsledků, je tedy vysoce žádoucí a neobejde se bez propojení zoologie s technickými obory. Z tohoto hlediska tedy hodnotím práci jako velmi přínosnou a zároveň oceňuji, že probíhala ve spolupráci se zoology z Jihočeské univerzity v Českých Budějovicích. Své metody a postupy autor testoval na nahrávkách budníčka malého (*Phylloscopus collybita*) a rypoše (*Fukomys darlingi*). Při řešení cílů práce musel autor také nalézt způsob, jak vyjádřit audiogramy ptáků v automatizovaných systémech nebo navrhnout a vytvořit databázi ptačích zpěvů.

Jako zoolog neznalý technické stránky dizertační práce se těžko vyjádřím k aplikovaným metodám a postupům, nicméně cíle práce jsou v zásadě uspokojivě splněny. Mám jen pocit, že kdyby práce byla psána ucelenějším a srozumitelnějším způsobem, byly by možná metodické části a výsledky o něco více srozumitelné i čtenářům mimo počítačové obory. Usuzuji tak z kapitol, které mi byly bližší, například kapitol uvádějících do problematiky nebo pojednávajících o ptačím zpěvu či nahrávání hlasů. Některé odstavce nebo i kapitoly dizertační práce jsou psané poněkud zmateným způsobem, takže jsem se v nich i jako člověk relativně znalý bioakustiky poměrně těžko orientovala. Například již v úvodní kapitole (str. 2) autor poněkud zmateně vyjmenovává práce zabývající se automatickým rozpoznáváním hlasů zvířat, přičemž u některých z nich není z textu jasné, s jakými druhy se vlastně pracovalo. Mezi pracemi o pěvcích se tak objevuje práce zabývající se hmyzožravci (Zsebök et al. 2015) a autor jejich hlasové projevy trochu nešťastně označuje jako zpěvy. Srozumitelnosti a čtivosti textu by také pomohlo, kdyby se autor držel názvů kapitol. Například kapitola „Speaker recognition methods in ornithology“ (str. 24) zahrnuje i odstavec o psím štěkání, což podle mne zbytečně zatěžuje a odvádí pozornost od tématu. V některých částech práce jsou informace podávány na přeskáčku, například bych jako první očekávala informaci o tom, kde a v jakých podmínkách jsou rypoši chováni a teprve poté informaci o tom, jakým způsobem byly nahrávány jejich hlasové projevy („Mashona mole-rat identification“, str. 106). Také bych informaci o podmínkách chovu rypošů nehledala v podkapitole „Data“. V práci se také informace někdy zbytečně opakují, například v kapitole „Data recording“ (str. 19) autor opakovaně zmiňuje některé problémy při nahrávání ptáků v terénu.



Formální a jazyková úroveň práce je víceméně v pořádku, je jen škoda, že celkový dojem kazí opakované překlepy pravděpodobně způsobené automatickou opravou pravopisu použitého programu (např. křiklaví místo křikaví na str. 6 nebo Fukomys darling místo Fukomys darlingi na str. 106) a také nekonzistentnost v užívání kurzívy u vědeckých názvů zvířat.

Jak jsem již zmínila, předložené dizertační práci bych hlavně vytkla malou čtivost a srozumitelnost, které jsem si všimla i u kapitol mně nejbližších. Větší srozumitelnost dizertační práce i pro čtenáře mimo počítačové obory by možná nebyla na škodu, neboť výsledky práce budou jistě zajímat zejména ornitology a zoology. Na druhou stranu autor výsledky své práce již úspěšně publikoval v časopise Bioacoustics a já se domnívám, že není vážný důvod, aby nebyla doporučena k obhajobě. Proto ji k obhajobě doporučuji.

K práci mám několik otázek:

1. Autor se snažil identifikovat 13 jedinců budníčka malého a dosáhl úspěšnosti klasifikace 78.6%. Je takový počet jedinců pro nějaké smysluplné ornitologické studie dostačující, nebo ornitologové pracují v terénu s větším počtem kroužkovaných ptáků?
2. Na straně 110 v souhrnné kapitole autor uvádí, že výsledky u ryposů prokázaly, že samice preferovaly hlasy podřízených samců. Nic takového se ale v celé kapitole o ryposích neřeší, je zaměřená na testování individuality v hlasech samic. Probíhala tedy souběžně na ryposích nějaká další práce?

RNDr. Irena Schneiderová, Ph.D.

V Praze 6.3.2017