

Posudek disertační práce
(Fakulta aplikované informatiky, UTB ve Zlíně)

Ing. Dalibor Slovák

" Frekvenční a amplitudová analýza signálu – Identifikace krátkodobého impulsního poškození v audio datech“.

1. Téma, obsah a struktura práce

Předložená disertační práce jednak mapuje metody oprav chyb v audiosignálech, které vznikly v procesu digitalizace (těchto audiosignálů), jednak předkládá metody, jak tyto chyby identifikovat. Z mnoha různých možných chyb si autor k podrobnějšímu zpracování vybral krátkodobé, autorem nazvané „impulzní“ poškození. Cíl autora směřuje zejména k nahrávkám skladeb, které ve své analogové podobě byly schváleny jako finální pro veřejnou produkci, ale vlivem procesu chybné digitalizace se staly nepoužitelnými.

Předložená disertace má 130 stran s dvaceti sedmi stranami příloh.

Seznam referencí na použité zdroje má 30 položek, seznam publikací, kde je disertant prvním autorem nebo spoluautorem má 17 položek (včetně jednoho uděleného patentu). Grafická úprava práce je na velmi solidní úrovni, až na několik překlepů je jazyková forma kvalitní.

2. Cíl práce a hlavní dosažené výsledky

Cíl disertace je explicitně formulován na straně 38 v pěti bodech. Konstatuji, že body cíle byly splněny. Druhá polovina bodu číslo 5 „Navrhnout další postupy především pro identifikaci a restaurace krátkodobých impulsních poškození“ byla v práci splněna velmi opatrně.

Disertace byla podmíněna silnou motivací, zmíněnou již v předchozím odstavci, *obtížností vyhodnocování korektnosti (chybovosti) digitalizovaných nahrávek* (zejména pro podmínky Českého rozhlasu). Jak sám autor v práci uvádí na několika místech, poškození nahrávek digitalizací se jeví jako velmi komplexní problém a disertace řeší jen jediný typ takového poškození.

Hlavním výsledkem disertace je programový systém (software), který v předložených digitalizovaných nahrávkách detekuje tzv. „impulsní poškození“. Ve skutečnosti to znamená, že označuje v záznamu nahrávky místo, ve kterém se objeví výrazný impuls, odlišný od okolního průběhu záznamu. Toto místo systém detekuje s různou citlivostí pomocí čtyř různých metod: metody využívající Euklidovu metriku, metody s dvojitou konvolucí, metody využívající vlnkové transformace (s různými typy waveletů) a metody zpracovávající časové řady. Autor nezasahuje do teoretického základu těchto metod, ale soustřeďuje se na jejich metodickou aplikaci a implementaci. Doložení funkčnosti tohoto softwaru je v disertaci velmi dobře zpracováno a předložené příklady i přílohy jsou postačující. Předpokládám, že při eventuální prezentaci výsledků práce již budou k dispozici zkušenosti s nasazením vyvinutého softwaru v praktických případech (zejména v podmínkách Českého rozhlasu).

3. Připomínky a poznámky k disertaci

3.1 Konceptuální výhrady

Velkou slabinou disertace jsou konceptuální nepřesnosti, které se týkají popisů, uvedených formulí a výrazů a „definic“.

- Nepřesný je popis vlastní úlohy identifikace. Používání termínu identifikace v disertaci je samozřejmě srozumitelný, ale pokud by chtěl autor zcela vyhovět obsahu identifikace poškození, musel by vyvinout dostatečně vyčerpávající slovník (množinu, seznam) různých poškození, se

kterými by identifikoval nalezené poruchy. Jistěže se autor ohradí, že se zabýval pouze jedním typem poškození (které se projevilo jako výrazný impuls v záznamu signálu). Ale i v tomto případě by se slovník impulsů v různých kontextech hodil (jak dokládá v diskusi ke vlivu dozvuku k obr. 26 a 27. Mimochodem o teoretické a softwarové podobě „dozvuku“ se můžeme domnívat, že ji autor disertace zná.)

- Autor velice volně žongluje s různými zavedenými i nezavedenými termíny (jako v předchozím případě dozvuku). Dovolují si upozornit na: „definici“ kovarianční funkce na straně 22, na „definici“ vnějšího a vnitřního popisu systému na straně 23, na vše!!, co je řečeno v podkapitole 2.10 Využití umělé inteligence na stranách 34 a 35, na zařazení analogie Heisenbergova principu z kvantové mechaniky do metodiky práce s „okny“ na str.42, apod.
- Autor neuvedl seznam použitých symbolů a jejich význam. To vede k různým možným výkladům a k dalším nepřesnostem. Jako příklad uvádím: podkapitolu 2.3.1 o Fourierově transformaci a výraz (1) na straně 21, podkapitolu 2.7 Lineární predikční analýzy – adaptivní prediktor – výrazy (7) a (8) (ve výrazu (8) se vyskytuje „ $d(n)$ “, zřejmě místo „ $x(n)$ “), nejasný vztah výrazů (25) na str. 56 a výrazu (17) na str. 54.

Pokud by autor uvažoval o další publikaci disertace pro širší odbornou veřejnost, je nutno provést důslednou korekci popisů, „definic“ a výrazů v celé práci (nejen těch, které jsem zde uvedl).

3.2 Struktura disertace

- Z aktivních 107 stran disertace tvoří 42 stran úvodní pasáže, a popis různých metod a nástrojů. To samo o sobě není tragické. Problémem je hloubka popisu metod a nástrojů, které autor v práci vůbec nevyužije. Typickým příkladem je popis autoregresního modelu v podkapitole 4.7. Tento popis má 4 a půl strany a zachází až do podrobností QR rozkladu matice A a Gram-Schmidtova algoritmu.
- Uživatelské obrázky a menu vyvinutého softwaru (7 stran) by patřily spíše do příloh.

4. Závěr

Autor disertace popsal nelehkou úlohu z oblasti zpracování signálů a vyvinul pro ni funkční software. Cíle stanovené v disertaci splnil. I přes výtky uvedené v odstavci 3. posudku doporučuji předloženou disertaci k obhajobě.

V Praze dne 3.11.2016

Prof. Ing. Jiří Bíla, DrSc.



Doc. RNDr. Jindřich Černohorský, CSc.
Katedra kybernetiky a biomedicínského inženýrství
Fakulta elektrotechniky a informatiky
Vysoká škola báňská – Technická univerzita Ostrava
17. listopadu 15
70833 Ostrava –Poruba

OPONENTNÍ POSUDEK

disertační práce

Téma práce: **Frekvenční a amplitudová analýza signálu – Identifikace krátkodobého impulsního poškození v audio datech**
Doktorand: **Ing. Dalibor Slovák**
Školitel: **prof. Ing. Vladimír Vašek CSc.**
Pracoviště: **Univerzita Tomáše Bati ve Zlíně, Fakulta aplikované informatiky**

Oponentní posudek jsem vypracoval na základě jmenování děkanem Fakulty aplikované informatiky Univerzity Tomáše Bati ve Zlíně panem docentem Mgr. Milanem Adámkem PhD., CSc. do funkce oponenta disertační práce ze dne 30. září 2016.

Předložená disertační práce obsahuje 147 stran textu, z nichž 126 stran je vlastní text, 20 stran pak tvoří přílohy.

Autor v přehledu vědecké a publikační činnosti uvádí 16 publikací a 2 patenty. Tematicky se práce týká zpracování a záznamu signálu a to jak pokud jde o přínos v oblasti software i hardware. Na devíti publikacích a na obou patentech je doktorand uveden jako jediný autor. Z hlediska obhajoby této disertační práce považuji publikace doktoranda za dostačující.

Aktuálnost tématu

Aktuálnost tématu je dána především jeho současnou praktickou potřebou a také faktem, že profesionální komerční nástroje, které předpokládají a jsou koncipovány pro zpracování individuálních záznamů, nevyhovují požadavkům hromadného zpracování záznamů hudebního archivu, které například požadují pracovníci Českého rozhlasu, pro něž je efektivní hledání poruch typu krátkodobé impulsní poškození zásadní věcí.

Na druhé straně se domnívám, že z teoretického hlediska jde o velmi obtížný problém, pokud bychom ho chtěli řešit na sto procent. To znamená zkonstruovat postup resp. algoritmus, který stoprocentně určí všechna pochybení typu „krátkodobé impulsní poškození“, ale současně neoznačí za krátkodobé impulsní poškození něco, co jím opravdu není. To by bylo a pro pracovníky rozhlasu určitě ideální řešení. Ale kdo, nebo co určuje, co je krátkodobé impulsní poškození? Je to „ucho“ člověka, přičemž ale v řadě případů může být toto ucho na pochybách. Pustíme-li totiž neinformovanému posluchači záznam Čajkovského Předehry 1812, bude výstřely předepsané partiturou považovat za náhodné impulsní poškození? Označí tyto výstřely za krátkodobé impulsní poškození autonomní algoritmus? V každém případě je zřejmé, že studovaný problém nabízí řadu praktických i teoretických otázek a mohl by se stát motivem i pro nějaký základní výzkum.

Splnění cílů disertace

Cíle práce, stanovené autorem v kapitole 3 a formulované v pěti bodech, byly splněny.

Zvolené metody zpracování

S ohledem na to, co jsem napsal v odstavci aktuálnost tématu jsem přesvědčen, že autor postupoval při řešení v podstatě jediným možným způsobem. Metody zpracování korespondují s dílčími cíli. Autor provedl rešerši prostředků použitelných pro řešení problematiky DP. Jde jednak o komerční nástroje, především programy pro zpracování zvukových signálů, kterými jsou SoundForge, Adobe Audition (dříve CoolEdit), Wavelab jako referenční nástroje použitelné pro komparaci s výsledky vytvořených algoritmů. Dále zkoumá autor vlastnosti konkrétních metod zpracování signálů a jejich vhodnost pro zpracování krátkodobého impulsního poškození. Výsledkem zkoumání mnoha metod a přístupů je autorem vytvořená aplikace DispApp, která je jedním z hlavních výsledků práce. Ověřování vhodnosti jednotlivých metod provádí autor na třech vhodných reálných audio záznamech prezentujících tři zdroje hudebního zvuku: zpěv+klavír, kytara a mluvené slovo. Na těchto záznamech diskutuje autor problematiku analýzy a vlastnosti jednotlivých testovaných metod. Používá k tomu vytvořené aplikace DispApp, WaveletApp,

TimeSeriesApp a další metody jako EuclidIdentification, Dvojnou konvoluci, a další kombinace. Pro srovnání pak v závěru testuje i uvedené komerční produkty.

Význam práce pro praxi, přínosy práce

Práce byla inspirována praktickou potřebou. Je zaměřena především na potřeby Českého rozhlasu, jehož pracovníci se setkávají resp. střetávají s problémem identifikace krátkodobého impulsního poškození rozhlasových audiozáznamů. Jde ovšem o problém obecný, konkrétní potřeba neomezuje použití výsledků kdekoliv, kde se tento problém vyskytuje.

Přínos práce spočívá v otestování řady metod vhodných pro identifikaci fenoménu náhodných impulsních poškození a výběru těch z nich, které dávají největší šanci na úspěšnost použití. Konkrétním přínosem autora je pak vytvoření testovacích aplikací, v nichž autorem vyzkoušené metody byly použity. Přínosem práce pak je i to, že autor konkrétně na praktických příkladech demonstruje nemožnost jednoznačného automatického řešení. Práci možná chybí zpracování příkladu audiozáznamu s plnějším zvukovým obsazením.

Formální a jazyková úroveň

Práce má potřebné formální náležitosti pokud jde o její strukturu a kapitoly. Ku prospěchu práce je opatřena vhodným počtem obrázků - grafických záznamů. Avšak její čtení je poměrně obtížné, protože velká část textu je stylizována jako návod pro uživatele a prolíná se s textem vysvětlujícím „neuživatelské“ aspekty jednotlivých metod. Práci by také prospělo, kdyby její jednotlivé části, kapitoly resp. subkapitoly byly uspořádány tak, aby bylo možné lépe sledovat jednotlivá rozhodnutí autora, „proč právě se v konkrétním místě textu rozhoduje, kam se bude v dalším zkoumání ubírat“.

Otázky: K práci nemám, kromě toho co jsem uvedl kritické připomínky, velmi by mně však zajímaly dvě otázky:

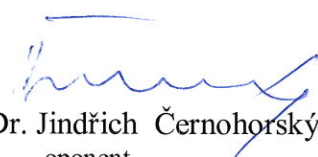
1. V práci byly použity nástrojově poměrně chudé záznamy, kde bych očekával i poměrně průzračný zvukový signál. Zkoušel autor použít svoji metodu například na záznamu se zvukem plného symfonického orchestru? Pokud ne, nabízím kromě Čajkovského 1812, 4. větu z Beethovenovy symfonie č. 7 anebo Haydnovy symfonie č. 94 „S úderem kotlů“.
2. Jaké, resp. v jakém smyslu by mělo přinést zlepšení s využitím neuronových sítí ?

Závěrečné hodnocení

Doktorská disertační práce pana Ing Dalibora Slováka je zpracována na dobré odborné úrovni. Přináší nové poznatky pro řešení problému identifikace krátkodobého impulsního poškození, ukazuje na jeho vysokou erudici v oblasti zpracování signálu i jeho způsobilost k samostatné tvůrčí vědecké práci a dle mého názoru splňuje náročné podmínky stanovené pro doktorské disertační práce v souladu s § 22 zákona č. 172/90 Sb., a proto ji

doporučuji k obhajobě

V Ostravě, dne 31.10.2016


doc. RNDr. Jindřich Černoňorský, CSc.
oponent

Oponentní posudek disertační práce Ing. Dalibora Slováka s názvem „Frekvenční a amplitudová analýza signálu-Identifikace krátkodobého impulsního poškození v audio datech“.

Předložená práce má 147 stran, z toho 128 stran textu. Na 63 stranách popisuje vlastní řešení zadaného problému. Jistému teoretickému úvodu řešené problematiky je věnováno 42 stran. Tento poměr je odpovídající a akceptovatelný.

Úvodní rešerše problému se v předložené práci nekoná. Z uvedené literatury v počtu 30-ti odkazů se nevyskytují práce v recenzovaných mezinárodních časopisech napsaných po roce 2000. K rešeršní části bych měl výhradu v tom, že podstatnou většinu uvedené literatury tvoří knihy. Obvykle to, co je uvedeno v knihách, je sice docela prověřené, ale staré. Stav problému se dá zjistit pouze studiem aktuálních příspěvků v recenzovaných časopisech a z této literatury není v práci uvedeno vůbec nic.

V teoretické úvodní části se autor až příliš často odkazuje v podstatě pouze na několik učebnic a popisů software a o teorii vlastně vůbec nejde. Teoretická část práce je silně zjednodušená. Matematický popis algoritmů chybí, a pokud se v náznaku objeví, působí jako přepis vytržených tvrzení z uvedených knih.

Vlastní téma práce, které vychází z praktických požadavků pracovníků rozhlasu, považuji za velmi pěkné a aktuální a moc se mi líbí.

Formální úprava práce je diskutabilní, kde obrázky a text si docela vzájemně zavazí. Nerad formální úpravy hodnotím, taktéž nerad se vyslovuji ke způsobu vyjadřování autorů, každý má svůj způsob řeči a sdělování myšlenek, ale zde pouze mohu konstatovat, že předložená práce se mi nečetla plynule. Opakují se tatáž tvrzení s neustálým lamentem nad složitostí zadaného tématu.

S formálními nedokonalostmi bych se smířil celkem bez větších problémů, ale co mne hodně rmoutí je odborný obsah textu.

Uvedu pouze několik poznámek. Kapitola Současný stav začíná částí 2.1 Popis problému. Kdo by však očekával, že zde setká s popisem problému, tak se velmi mýlí. V celé práci se totiž vůbec nevyskytuje popis „krátkého impulsního poškození“, alespoň ve formě záznamu časového průběhu. Pouze hodně slov s minimem informací. Jak došlo k „digitálnímu impulznímu poškození“ se nikdo také nedoví.

Výklad je veden velmi volně, uvedu pouze dva příklady.

Na straně 34 jsou zmíněny Kohonenovy mapy (píšu zmíněny ne popsány) se spoustou tvrzení bez jakýchkoliv odkazů nebo argumentů. Nejsou uvedeny ani vlastní výsledky, ač je

naznačeno, že byly získány. Nejsou doloženy prostě ničím. Tento přístup se objevuje pořád a pořád na každé stránce. Stejným způsobem je psáno skoro vše.

Na straně 40 je uvedena Pythagorova věta v souřadnicích, které se náhle objevují a stejně náhle mizí.

Výklad je veden nesystematicky až chaoticky, ani u jedné metody nejde do hloubky. Každá metoda je označena za vhodnou bez zdůvodnění na začátku kapitoly, aby po několika stranách byla označena za nevyhovující, opět bez jakékoli hlubší analýzy. Předložená práce je v podstatě neoponovatelná. Nevím, co bych oponoval, jelikož výsledků a myšlenek jsem v práci našel jen velmi málo. Autor práce většinou popisuje postup k použití software stylem, kliknete-li na tuto ikonku, stane se toto, uživatelsky v pořádku, ale kam se schovala vědecká metoda kritického myšlení?

Často se autor vyjadřuje k tomu, že se tématem zabývá hodně pracovišť, ale nejsou uvedena, ani není jasné, odkud autor čerpal, když necituje žádnou současnou literaturu.

V Brně na VUT je velmi silná skupina zabývající se zpracováním zvukových signálů, proč jste jejich práce necitoval? Spolupracoval jste s nimi?


Práce se zabývá adaptivní filtrací, ale nenašel jsem jedinou frekvenční charakteristiku použitých filtrů.

V příložených tabulkách jsou uváděny hodnoty na mnoho platných míst, jaké jsou odhadované nejistoty vašeho modelování?

Docela mi v práci chyběl jasný a srozumitelný výstup ve formě neupravené a upravené hudební ukázky, což by velmi pomohlo rozptýlit mé pochyby vyjádřené výše, doufám, že při obhajobě tato ukázka chybět nebude.

Podle mého názoru, a říkám to s lítostí, předložená disertační práce je na hraně obecně uznávaných požadavků k udělení akademického titulu doktora, nicméně téma ač obtížné je velmi zajímavé a tak doufám, že mne autor předložené disertační práce vyvede z mých mylných představ, a proto ji doporučuji k obhajobě.

Ve Zlíně 18.listopadu 2016


doc. RNDr. Vojtěch Křesálek, CSc.
Ústav elektroniky a měření
Fakulta aplikované informatiky
Univerzita Tomáše Bati ve Zlíně