



ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE

Fakulta elektrotechnická

Katedra radioelektroniky

Systém pro klasifikaci vad analogového záznamu

System for Classification of Analog Record Degradation

Václav Moldan

Diplomová práce

Studijní program: Elektronika a komunikace

Studijní obor: Audiovizuální technika a zpracování signálů

Vedoucí práce: Ing. František Rund, Ph.D.

Praha 2019

I. OSOBNÍ A STUDIJNÍ ÚDAJE

Příjmení: **Moldan** Jméno: **Václav** Osobní číslo: **435035**
Fakulta/ústav: **Fakulta elektrotechnická**
Zadávající katedra/ústav: **Katedra radioelektroniky**
Studijní program: **Elektronika a komunikace**
Studijní obor: **Audiovizuální technika a zpracování signálů**

II. ÚDAJE K DIPLOMOVÉ PRÁCI

Název diplomové práce:

Systém pro klasifikaci vad analogového záznamu

Název diplomové práce anglicky:

System for Classification of Analog Record Degradation

Pokyny pro vypracování:

Seznamte se s vadami typickými pro analogový záznam zvuku na gramofonové desce, včetně vad vzniklých při výrobním procesu. Popište jejich typické parametry a vytvořte modely těchto vad. Dále navrhnete metody automatické klasifikace těchto vad a otestujete je na vytvořených modelech vad. Výsledný systém implementujte a ověřte jeho použití na reálných vzorcích vad.

Seznam doporučené literatury:

- [1] Godsill, S, Rayner, P. Digital Audio Restoration. Springer-Verlag, London, 1998.
- [2] Dušek, K. Záznam a reprodukce zvuku. ČVUT, Praha, 1982.

Jméno a pracoviště vedoucí(ho) diplomové práce:

Ing. František Rund, Ph.D., katedra radioelektroniky FEL

Jméno a pracoviště druhé(ho) vedoucí(ho) nebo konzultanta(ky) diplomové práce:

Datum zadání diplomové práce: **13.02.2019** Termín odevzdání diplomové práce: _____

Platnost zadání diplomové práce: **20.09.2020**

Ing. František Rund, Ph.D.
podpis vedoucí(ho) práce

prof. Mgr. Petr Páta, Ph.D.
podpis vedoucí(ho) ústavu/katedry

prof. Ing. Pavel Ripka, CSc.
podpis děkana(ky)

III. PŘEVZETÍ ZADÁNÍ

Diplomant bere na vědomí, že je povinen vypracovat diplomovou práci samostatně, bez cizí pomoci, s výjimkou poskytnutých konzultací. Seznam použité literatury, jiných pramenů a jmen konzultantů je třeba uvést v diplomové práci.

Datum převzetí zadání

Podpis studenta

Prohlašuji, že jsem předloženou práci vypracoval samostatně a že jsem uvedl veškeré použité informační zdroje v souladu s Metodickým pokynem o dodržování etických principů při přípravě vysokoškolských závěrečných prací.

V Praze dne 23. 5. 2019

Václav Moldan

Poděkování

Chtěl bych poděkovat vedoucímu práce Ing. Františku Rundovi, Ph.D. za cenné rady a pomoc při vypracování této práce. Zároveň bych chtěl poděkovat firmě GZ Media, a.s. za poskytnutí nahrávek pro testovací účely.

Tato práce byla podpořena grantem Studentské grantové soutěže ČVUT č. SGS17/190/OHK3/3T/13.

Abstrakt

Tato práce se zabývá tématem modelování a klasifikace vad gramofonových desek. Cílem je vytvořit modely základních typů vad a navrhnout a implementovat v prostředí MATLAB metody automatické klasifikace těchto vad. Modelování i rozpoznávání těchto vad využívá jejich charakteristických zvukových vlastností. Pro klasifikaci vad na základě zjištěných vlastností je použit algoritmus strojového učení k-nejbližších sousedů (k-nearest neighbours, KNN). Funkčnost implementovaného algoritmu je nejprve ověřena na vytvořených modelech vad, kde se podařilo dobře rozlišit všechny typy vad kromě dvojice vad lupanec a přeskok, které mají v samotném místě vady velmi podobný projev. Po natrénování na uměle vytvořených vzorcích vad byl algoritmus ověřen i na dostupných vzorcích vad pořízených z reálných gramofonových desek. Vzhledem k malému množství těchto reálných vzorků však nelze dosažené výsledky považovat za vypovídající.

Klíčová slova: klasifikace, gramofonové desky, vady, strojové učení, KNN, MATLAB

Abstract

The topic of this thesis is modelling of defects in vinyl disc records and their classification. The aim is to create models of the basic types of the defects and propose and implement system for automatic classification of the defects in MATLAB programming environment. The approach for modelling and classification of the defect types is based on their characteristic sound attributes. The classification based on found attributes is performed with k-nearest neighbors (KNN) classification algorithm. The algorithm is tested on the created models of the defects, where we were successfully able to distinguish between all the types of the defects except for pair click and skip, which were often mistaken due to their similar characteristics in the short segment where the defect occurs. After being trained on artificially created models of the defects the algorithm was also tested with available samples of defects recorded from real vinyl discs. Due to the small number of available samples the test results could not be considered conclusive.

Key words: classification, vinyl discs, defects, machine learning, KNN, MATLAB

Obsah

Úvod	13
1. Mechanický záznam zvuku	15
1.1. Výroba gramofonových desek	15
2. Vady v záznamu zvuku na gramofonových deskách	17
2.1. Šum	17
2.2. Houkání	18
2.3. Lupanec	18
2.4. Škrábanec	18
2.5. Škrták.....	18
2.6. Nedolisy.....	19
2.7. Přeskok	20
2.8. Výpadek.....	20
3. Současná řešení problematiky	21
3.1. Rozdílová analýza.....	21
3.2. Vlnková transformace	22
3.3. AR Model	22
3.4. Optická kontrola	22
3.5. Automatická detekce chyb v audio záznamu	23
4. Strojové učení.....	25
4.1. KNN	25
4.2. Další algoritmy	25
4.3. Křížová validace.....	26
5. Vytvoření modelů vad v prostředí MATLAB	29
5.1. Houkání a šum	29
5.2. Škrták.....	30
5.3. Lupanec	31
5.4. Přeskok	33
5.5. Nedolisy.....	33
5.6. Škrábanec	34
5.7. Výpadek.....	34
6. Implementace klasifikačního algoritmu v prostředí MATLAB.....	35

6.1.	Parametrizace – detekování vlastností vad	36
6.1.1.	Výskyt vady na otáčku	36
6.1.2.	Délka a dozvuk vady	36
6.1.3.	Série lupanců za sebou	37
6.1.4.	Výskyt vady v pravém kanále	38
6.1.5.	Šum a houkání v nájezdu	38
6.1.6.	Mezery v signálu.....	39
6.1.7.	Změna ve spektru nahrávky.....	40
6.2.	Test detekce vlastností vad	42
6.2.1.	Detekce vlastností vad bez přítomnosti okolního zvuku	42
6.2.2.	Stanovení prahů a dalších parametrů pro detekční metody	43
6.2.3.	Test algoritmu detekce mezer	44
6.2.4.	Test algoritmu pro detekci šumu a houkání v nájezdu.....	47
6.3.	Detekce vlastností vad přidanych do reálného masteru	50
6.4.	Klasifikace pomocí metody KNN	51
6.4.1.	Testování na modelech vad, výběr parametru k.....	51
6.4.2.	Testování reálných vad	54
Závěr		57
Literatura		59
Přílohy.....		61

Úvod

V této práci jsem se zabýval popisem charakteristických vlastností jednotlivých vad gramofonových desek s cílem vytvořit modely těchto vad a následně tyto modely využít pro vytvoření algoritmu jejich automatické klasifikace. Vzhledem k tomu, že existují a jsou dále rozvíjena relativně pokročilá řešení pro samotné nalezení vad v záznamu zvuku gramofonové desky, nezabývá se tato práce tímto tématem do hloubky a navržený algoritmus vychází z toho, že pozice vad již známe z předchozího kroku, kterým by v praxi bylo právě nalezení pozic vad.

Kontrola vad na gramofonové desce byla v minulosti prováděna převážně pracovníky kontroly bez systému automatické kontroly, ať už poslechem nebo opticky. Tento postup je samozřejmě časově náročný a zároveň i nespolehlivý, protože lidská schopnost soustředit se po delší dobu je také omezená.

Dnes již existují systémy, které pomáhají s automatizací procesu kontroly. Prozatím jsou však dostupné informace pouze o systémech detekce vad, není však doposud popsán algoritmus, který by byl schopen vady dále klasifikovat a určit o jaký typ vady se jedná. Tato informace by byla užitečná k tomu, aby se identifikovala ta část výrobního procesu, při které vada vznikla, případně by napomohla i k určení závažnosti dané vady.

Zkušený pracovník gramofonových závodů je schopen rozlišit vady dle jejich slyšitelných vlastností, jako je například, zda se jedná o krátký samostatný lupanec nebo sérii lupanců, zda se vada vyskytuje na otáčku, jestli je v obou nebo jen v jednom kanále a i podle dalších projevů. Právě tyto poslechem rozlišitelné vlastnosti jsou využity v této práci i pro automatickou klasifikaci vad. Pro následnou klasifikaci, tedy pojmenování vady na základě jejich zjištěných vlastností, je použit klasifikační algoritmus strojového učení k-nejbližších sousedů KNN.

První kapitola této práce je věnována teoretickému popisu principu mechanického záznamu zvuku a výroby gramofonových desek, druhá kapitola navazuje popisem možných vad gramofonových desek. Třetí kapitola popisuje současný stav a existující řešení podobných a souvisejících úloh. Čtvrtá kapitola seznamuje s principy strojového učení, které jsou využity v praktické části, a také těmi, které jsou používané pro podobnou problematiku. Pátá kapitola se zabývá vytvořením modelů vad gramofonových desek na základě jejich zvukových vlastností. Šestá kapitola pak již samotnou implementací a následným testováním klasifikačního algoritmu, který bude schopný vadu na základě jejich zvukových atributů pojmenovat.

1. Mechanický záznam zvuku

Mechanický záznam zvuku je záznam zvuku založený na mechanické deformaci záznamového materiálu během procesu záznamu [1]. Mechanickým záznamem zvuku nejčastěji myslíme záznam zvuku na gramofonové desce. Historicky spadl do kategorie mechanického zvuku ale i záznam na fonografu, tedy rotujícím válečku, ve kterém byl zvuk zaznamenán hloubkovým záznamem (kdy hloubka drážky odpovídá okamžité hodnotě akustického tlaku).

Ačkoliv je záznam zvuku na gramofonové desce stále velmi populární kvůli svému charakteristickému zvuku, který je dán právě mechanickým principem záznamu, plyne z tohoto principu řada omezení a zároveň náchylnost k vadám různých typů.

1.1. Výroba gramofonových desek

Pro návrh algoritmu klasifikace vad na gramofonových deskách je nutné znát základní kroky výroby gramofonových desek. Pro většinu vad je totiž typické, že vznikají v určitých fázích výroby, v některých případech už při přípravě podkladů, nebo naopak až po vyrobení desky při manipulaci s deskou a při jejím přehrávání.

Popis procesu výroby gramofonových desek a použitých technologií je popsán v interním dokumentu firmy GZ Media [2], je zároveň volně dostupný na jejich webových stránkách [3].

Po dodání zvukového obsahu desky, seřazení jednotlivých skladeb a případným přidáním ticha mezi jednotlivé skladby přichází fáze masteringu, kdy je zvukový záznam upraven tak, aby ho bylo možné dále zaznamenat na desku s minimalizací nežádoucích artefaktů a naopak se zachováním zvukových vlastností dodaného podkladu. Změny by měly být velmi citlivé a neměly ovlivňovat charakter hudby. Je v této fázi provedeno také nastavení šířky drážky a hlasitosti desky. Tyto dva parametry spolu úzce souvisí, na jednu stranu je žádoucí mít signál s velkou amplitudou, abychom zvýšili poměr užitečného signálu a nežádoucích artefaktů (šum, vady mechanického záznamu), na druhou stranu však při velké amplitudě hrozí riziko prořezu mezi drážkami a snižuje se možná hrací doba desky [1].

Pro následné řezání se v GZ Media a dalších výrobců používají dvě technologie, tradiční laková a Direct Metal Mastering (DMM).

Princip tradiční lakové technologie výroby spočívá v tom, že se drážka řeže safírovým nožem do lakové fólie, tedy hliníkového plechu s nanesenou vrstvou nitrocelulózového laku. Z lakové fólie není možné, kromě výjimek, kdy je výroba prováděna v malém množství, vyrábět přímo lisovací matrici. Drážka je totiž na lakové fólii velmi citlivá a snadno by se mohla poškodit. Následná výroba lisovacích matric tedy vyžaduje vytvoření meziproductů, a tedy obvykle trojnásobné kopírování, je proto nákladnější a složitější.

Technologie DMM je novější technologií používanou pro řezání. V GZ Media je častější používání této technologie, tradiční laková technologie se používá ve výjimečných případech. Princip DMM je takový, že řezání se provádí speciálně vybroušeným diamantem přímo do čisté měděné vrstvy galvanicky nanesené na plechu z nemagnetické oceli. Výhodou je, že lisovací matrice se pak může vyrábět přímo z tohoto měděného nosiče. Výroba matrice probíhá pomocí niklování měděného nosiče.

Matrice jsou dále kontrolovány, zda neobsahují nějakou vadou. Poté jsou již předány k lisování. Vylisované desky se pak předávají na kontrolu. U plné výroby je kontrola prováděna systematicky tak, aby byla kontrolována průběžně kvalita výroby, není však možné kontrole podrobit každou desku. Kontrolovány jsou jak informační vady (matriční číslo, seznam skladeb apod.), optické vady (skvrny či nečistoty, které sice neruší při poslechu, ale kazí vizuální dojem z desky) tak akustické vady.

Kontrola akustických vad probíhá jak poslechem pracovníky kontroly, tak automaticky. Princip, na kterém je postaven systém automatické detekce, se může lišit pro určité výrobce desek, metody používané konkrétními výrobci nejsou nikde popsány. Principy možných řešení jsou blíže popsány v kapitole 3.

2. Vady v záznamu zvuku na gramofonových deskách

Z výše uvedeného stručného popisu výroby gramofonových desek plyne, že proces je poměrně náročný a je zde mnoho kroků, při kterých mohou vzniknout vady.

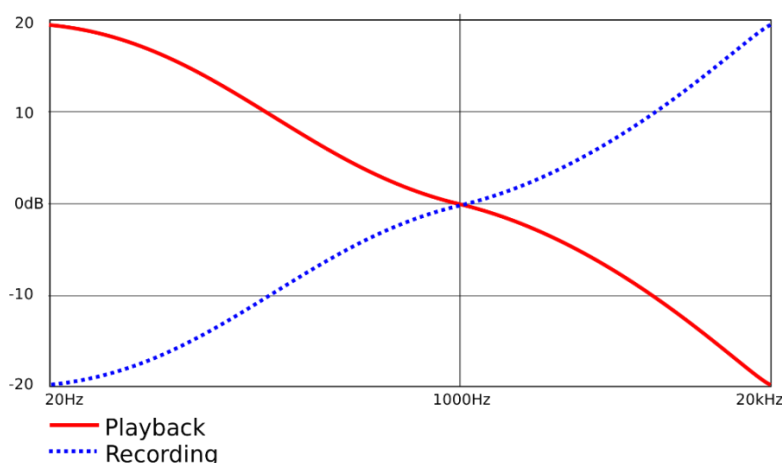
Pro účely této práce pracujeme s několika základními typy vad, které jsou od sebe relativně dobře rozlišitelné. Popis těchto a dalších vad je uveden v interních materiálech firmy GZ Media [2]. Stručný popis je uveden v následující kapitole.

2.1. Šum

Obvyklým nežádoucím artefaktem nejen na gramofonových deskách, ale obecně u záznamu a reprodukce zvuku je šum. U gramofonových desek je nutné rozlišit mezi šumem, který pochází již z podkladu od zákazníka (mohl vzniknout třeba již ve studiu při nahrávání) a šumem, který vznikl v důsledku záznamu zvuku na desku. Určitá úroveň šumu se objeví vždy kvůli tomu, že není technicky možné vyrobit drážky naprosto hladké a nerovnosti se projeví ve zvuku ve formě šumu [1]. Zároveň dojde k přidání šumu samotným přehrávačem, jednak při snímání jehlou a jednak díky elektrickým obvodům v přehrávači. Moderní přehrávače dosahují hodnot SNR = 50 dB [4].

Na gramofonových deskách je výrazný především šum v nájezdu desky. Šum je výraznější v nájezdu proto, že na začátku desky je širší drážka a můžeme tak pozorovat větší úroveň šumu. Výraznější šum pak může být způsoben i dalšími faktory, například tvarováním a broušením matrice nebo lisováním.

Pro potlačení šumu se při masteringu desky přistupuje k předekvalizaci záznamu [5], kdy jsou vysoké frekvence před výrobním procesem zesíleny (tzv. preemfáze). Opačná ekvalizace je poté použita při přehrávání desky (tzv. deemfáze), výsledek je tedy nahrávka se stejnou frekvenční charakteristikou, jaká byla u původní nahrávky, ale s potlačeným šumem, který při výrobním procesu vznikl na vysokých kmitočtech. Tvar křivky pro preemfázi a deemfázi byl standardizován Asociací amerického nahrávacího průmyslu (Recording Industry Association of America, RIAA) v roce 1954 [6,7].



Obrázek 1: RIAA ekvalizační křivka [6]

2.2. Houkání

Tzv. houkání či foukání se projevuje pulsující frekvencí, může projevem připomínat foukání větru. Obvykle je způsobeno broušením matrice, pokud se objeví na frekvenci 300 Hz, je způsobeno stylem přebrušování norem.

2.3. Lupanec

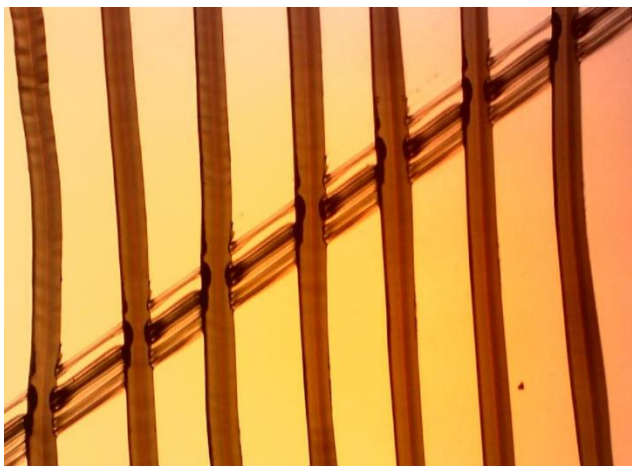
Lupanec může mít více příčin, může být také projevem dalších, níže popsanych vad. Lupanec může mít různou intenzitu.

Speciálním případem je elektrostatický lupanec, který se projevuje velmi krátkým a ostrým lupnutím bez dalšího dozvuku. Při opakovaném přehrání daného místa elektrostatický lupanec neslyšíme vůbec, anebo ho slyšíme na jiném místě (došlo k přesunu elektrického náboje na jiné místo).

2.4. Škrábanec

Škrábanec můžeme opticky pozorovat jako čáru, která je umístěna často přes několik drážek záznamu. Pokud škrábanec nezasahuje do dna drážek, nemusí se akusticky vůbec projevit. Pokud však zasahuje do dna nebo stěn drážek, projeví se akusticky jako lupanec na otáčku.

Pod mikroskopem můžeme také pozorovat, jestli je škrábanec na negativitu nebo pozitivu, je totiž velký rozdíl, pokud je škráblá pouze samotná deska nebo již matrice, ze které byla deska vyrobena, kdy je pak škrábancem poškozená každá deska vyrobená z dané matrice.



Obrázek 2: Škrábanec na desce, poškozeno více drážek desky [2]

2.5. Škrták

Pravý škrták (vzniklý při rozebírání nástrojů) či nepravý škrták (vzniklý škrtnutím desky o matrici) se projevuje souvislým praskáním na desce. Opticky můžeme škrtáky pozorovat jako drobné tečky nebo čárky. V případě pravého škrtáku je vadou zasažen jak pozitiv tak negativ a vadu tedy nalezneme na všech deskách vzniklých z jedné matrice, v případě nepravého škrtáku je zasažena pouze samotná deska.



Obrázek 3: Škrťák na desce [2]

2.6. Nedolisy

Akustický projev nedolisů je podobný jako u škrťáku, praskání se však vyskytuje pouze v pravém kanále. Vada často zasáhne několik drážek a vyskytuje se poté na otáčku. Příčinou je lisování, kde hmota při lisování nedostatečně obteče matrici. Tento problém se obvykle objevuje ke konci stran. Opticky můžeme nedolisy pozorovat jako shluky bílých teček. Nedolisy jsou fyzicky umístěny v závětrí hran matrice.



Obrázek 4: Nedolisy na desce [2]

2.7. Přeskok

Při této vadě dojde k přeskoku jehly do jiné drážky a tím pádem k reprodukci jasně nenačuzujícího úseku skladby. Přeskok může být doprovázen i lupancem.

Důvodů může být více, mohou to být špatně nastavené parametry od operátora, které způsobí prořez drážek do sebe, dále úzká drážka, velká stranová výchylka už na podkladech od zákazníka, poškození drážky nebo výrazná zalísovaná nečistota.

2.8. Výpadek

U výpadku je určitá část signálu nahrazena tichem. Důvodem může být výpadek převodníku při řezání nebo chyba operátora a špatné oříznutí tracku při premasteringu. Chyba však nemusí být přítomna přímo na desce, může se jednat pouze o špatně seřízené vypínání automatického gramofonu, kdy se gramofon vypne a již nepřehraje konec desky.

3. Současná řešení problematiky

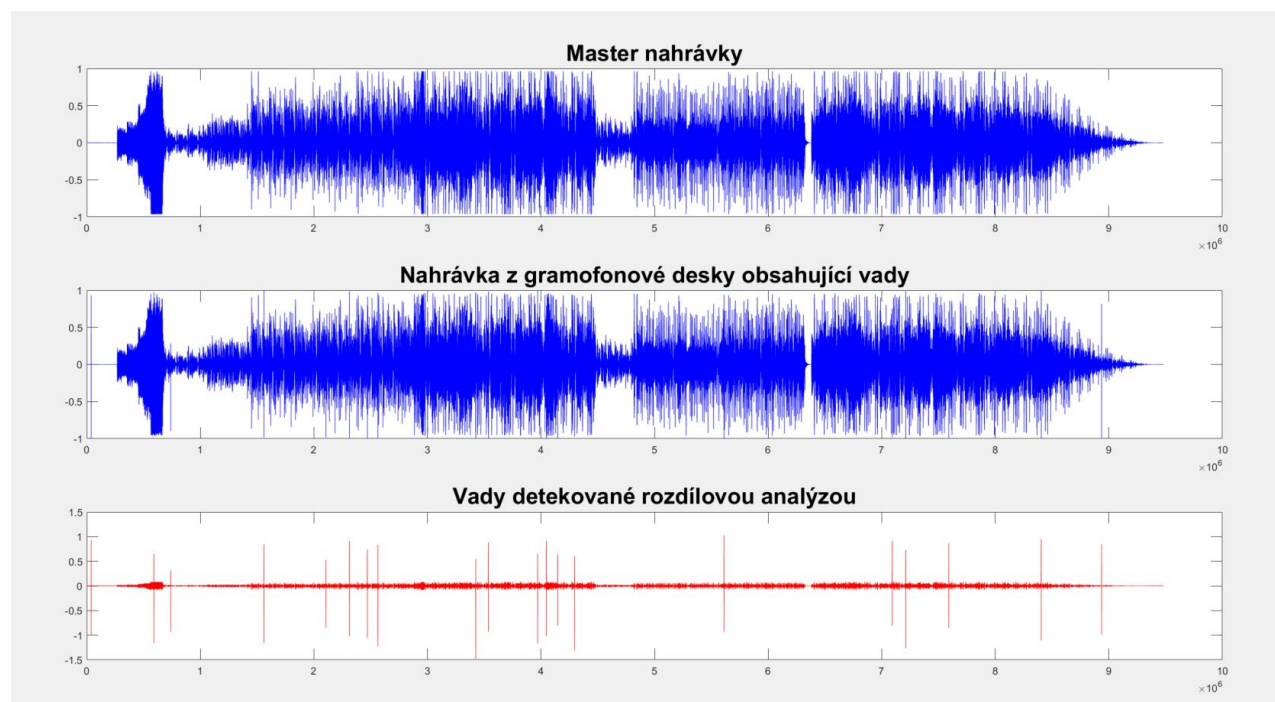
V dnešní době existuje řada řešení pro různé aplikace rozpoznávání a klasifikace zvuku. V oblasti detekce a klasifikace vad existují řešení převážně pro detekci vad, klasifikace je v průmyslu řešena stále převážně poslechem, případně optickou kontrolou operátory kontroly.

V následujících odstavcích jsou stručně popsána vybraná řešení pro automatickou detekci vad.

3.1. Rozdílová analýza

Jedním z možných postupů, který je používán pro kontrolu i u výrobce GZ Media, je metoda rozdílové analýzy. Postup této analýzy je takový, že je nahrávka z gramofonové desky nejprve synchronizována s masterem této nahrávky a následně jsou od sebe nahrávky odečteny. Velmi zásadním krokem je zde samotná synchronizace, protože může vlivem rychlosti otáčení desky docházet i k postupnému rozcházení se nahrávky na desce a jejího masteru v čase, nestačí tedy pouze synchronizovat začátek. Tam, kde je po odečtení výrazný rozdíl mezi masterem a nahrávkou pořízenou z desky, je detekována vada.

Rozšíření této metody o následnou klasifikaci detekovaných vad by jistě přicházelo v úvahu, rozdíl mezi masterem a signálem sejmutým z desky nám dá i relativně přesnou informaci o průběhu samotné vady, ze kterého již bychom měli být schopni odhadnout její další vlastnosti. Návrh podobného rozšíření zatím však není nikde popsán.



Obrázek 5: Rozdílová analýza, ilustrační obrázek

3.2. Vlnková transformace

Algoritmus detekce vad pomocí vlnkové transformace je popsán a implementován v práci [7] a dále ještě modifikován za účelem získání lepších výsledků v práci [8].

Princip detekce vad pomocí vlnkové transformace je takový, že signál je rozkládán pomocí banky filtrů na aproximační a detailní části. Algoritmus detekce využívá koeficienty detailů a porovnává je s dynamickým prahem. Délka filtru by měla být spíše kratší, jelikož délka impulsních vad by měla být kratší než délka vysokofrekvenčních částí signálu. V případě, že bude signál obsahovat vysokofrekvenční složky delší než impulsní vady, může dojít k falešné detekci vady.

V práci [8] byla vlnková transformace porovnána s dalšími modely – metodou AR modelu a metodou matched filtru. Mezi těmito metodami dosáhla vlnková transformace nejlepších výsledků.

3.3. AR Model

Godsill a Rayner [9] ve své práci představují metody automatické detekce a odstranění impulsních vad (tedy lupanců) z gramofonových desek pomocí autoregresního (AR) modelu. Detekční algoritmus založený na AR modelu je také popsán a implementován v diplomové práci [7].

Princip AR modelu je takový, že vzorek signálu získáme váhovaným součtem p předchozích vzorků. Počet těchto prvků p udává řád predikce. Původní signál bez impulsních vad lze lokálně aproximovat AR modelem. Jelikož impulsní vada znamená nepředvídatelnou změnu v signále, zjistíme ji jako rozdíl mezi původním signálem a rekonstruovaným signálem. Tento rozdíl se označuje jako chyba predikce nebo detekční signál. Pro správnou detekci je nutné vhodně stanovit práh. Hodnoty detekčního signálu nad tímto prahem jsou označeny jako impulsní vada.

V práci [7] jsou použity dvě hodnoty prahu zároveň. Tím, že má jeden z prahů vyšší hodnotu, umožní lepší lokalizaci lupnutí, jelikož vzorků nad tímto prahem bude méně.

3.4. Optická kontrola

V diplomové práci [10] byl představen návrh metody detekce vybraných typů vad na základě optické kontroly. Použité metody byly metoda hranových detektorů, matematické binární morfologie a měření objektů. V této práci se podařilo opticky detekovat vady škrťák, směť a křivost desky. Metody v této práci nejsou prozatím připravené k nasazení v průmyslové výrobě desek kvůli nevhodným parametrům použité kamery, nelze ale vyloučit, že v budoucnu dojde k jejich vylepšení a dojde k tomu, že bude kontrola akustických vad prováděna i opticky.

Dnes jsou známy i systémy optického čtení gramofonových desek, toto téma je zpracováno podrobněji v odborném článku [11]. U takových systémů je deska čtena bezkontaktně bez použití jehly, což především zamezí postupnému opotřebení desky. Nejzásadnější technologie bezkontaktního čtení je japonský ELP corporation's laser turntable, který desku snímá s pomocí 5 laserových paprsků.

V budoucnu by pro kontrolu chyb na deskách přicházela v úvahu i podobná metoda, dosud však není takový systém popsán ani znám.

3.5. Automatická detekce chyb v audio záznamu

Bakalářská práce [12] popisuje různé metody rozpoznání vad v audio záznamu obecně, není však zaměřená na gramofonové desky.

Je zde popsán algoritmus pro detekci nežádoucích mezer v audio signálu (blíže popsán v kapitole 6).

Dále je zde popsán algoritmus pro detekci přeskoků signálu na nenavazující úsek. Tento algoritmus využívá autoregresního modelu k detekci nespojitosti v signálu. U této metody se však nedá očekávat dobré uplatnění u gramofonových desek, kdy je přeskok provázen krátkým lupnutím a nemůžeme tedy najít místo, kde je jasně detekovatelná prudká změna fáze signálu.

Další popsán algoritmus nalezení určitého druhu vady je zde detekce náhlého shluk šumu (“noise burst”). Tento algoritmus bohužel však pro naše účely také není příliš použitelný, neboť umožní detekovat silný nárůst šumu, který zásadním způsobem naruší užitečný signál na krátkém úseku, pro naše účely potřebujeme spíše pracovat s relativně stálou hodnotou šumu přítomnou na celé desce.

4. Strojové učení

Pro řadu aplikací v oblasti rozpoznávání zvuku byly úspěšně implementovány metody strojového učení.

Časté je použití klasifikačního algoritmu k-nejbližších sousedů (KNN). Jeho výhodou je jeho jednoduchost a také nižší náročnost na velikost trénovací množiny oproti algoritmu umělé neuronové sítě ANN. Algoritmus byl již úspěšně implementován například pro účel rozpoznání typu obsahu nahrávky (tedy zda se na nahrávce nachází hudba, šum, řeč nebo ticho) [13], pro identifikace konkrétní nahrávky [14], pro identifikaci hudebního žánru [15] nebo pro rozpoznávání řeči [16].

4.1. KNN

Algoritmus k-nejbližších sousedů (KNN) [17] spadá do kategorie algoritmů strojového učení, které využívají učení s učitelem, kdy učení probíhá z trénovacích dat. Při trénování jsou využita trénovací data, která jsou rozdělena do konkrétních tříd, do kterých chceme později klasifikovat neznámé prvky. Každý prvek trénovací množiny zařadíme do určitého místa v N-rozměrném prostoru na základě hodnot parametrů, které si předem zvolíme a které pro daný prvek spočteme. Dále prvek zařadíme do té třídy, do které patří k jeho nejbližších sousedů.

K výpočtu vzdálenosti od ostatních prvků lze použít různé metriky, nejčastější je euklidovská metrika, která je dána vztahem:

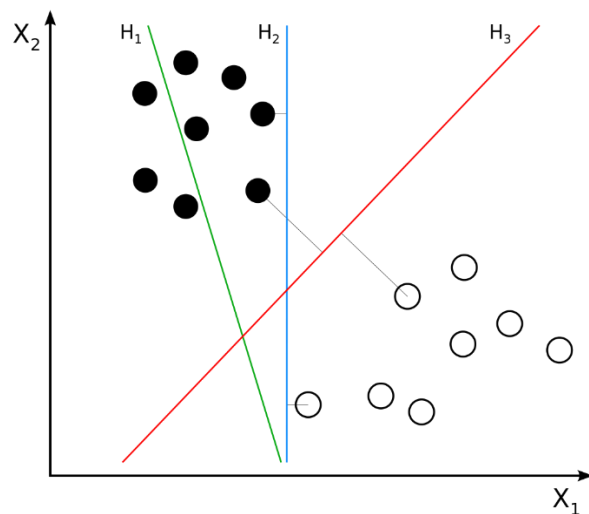
$$d(\vec{x}, \vec{y}) = \sqrt{\sum_i (x_i - y_i)^2}$$

kde \vec{x} a \vec{y} jsou vektory o stejném počtu prvků, a prvky vektorů x_i a y_i jsou hodnoty atributů.

4.2. Další algoritmy

Dalším oblíbeným a běžně používaným algoritmem strojového učení je umělá neuronová síť (artificial neural network) [18]. Princip algoritmu spočívá v propojení velkého množství členů (tzv. umělých neuronů), podobně jako tomu je ve skutečném biologickém mozku. Každý z těchto umělých neuronů má několik vstupů, ze kterých váhovaným součtem získáme výstup. Pokud je struktura vícevrstvá, je výstup neuronu napojen na vstup neuronu další vrstvy. Existují různé varianty implementace neuronové sítě a i způsoby jejího učení, detailnější popis je možné nalézt v článku [18].

Další běžně používanou metodou strojového učení je metoda support-vector machines (SVM, metoda podpůrných vektorů), která spočívá v zařazení prvků do prostoru na základě jejich parametrů a v následném stanovení hranice, která nejlépe odděluje prvky dvou různých tříd [19]. Je tedy možné s pomocí této metody rozdělit prvky pouze do dvou tříd. Z principu SVM ale vycházejí další algoritmy, například jádrová analýza hlavních komponent (PCA, principal component analysis), které již umožňují i zařazování do většího počtu tříd.



Obrázek 6.: Hledání hranice mezi prvky dvou tříd pomocí metody SVM [20]

4.3. Křížová validace

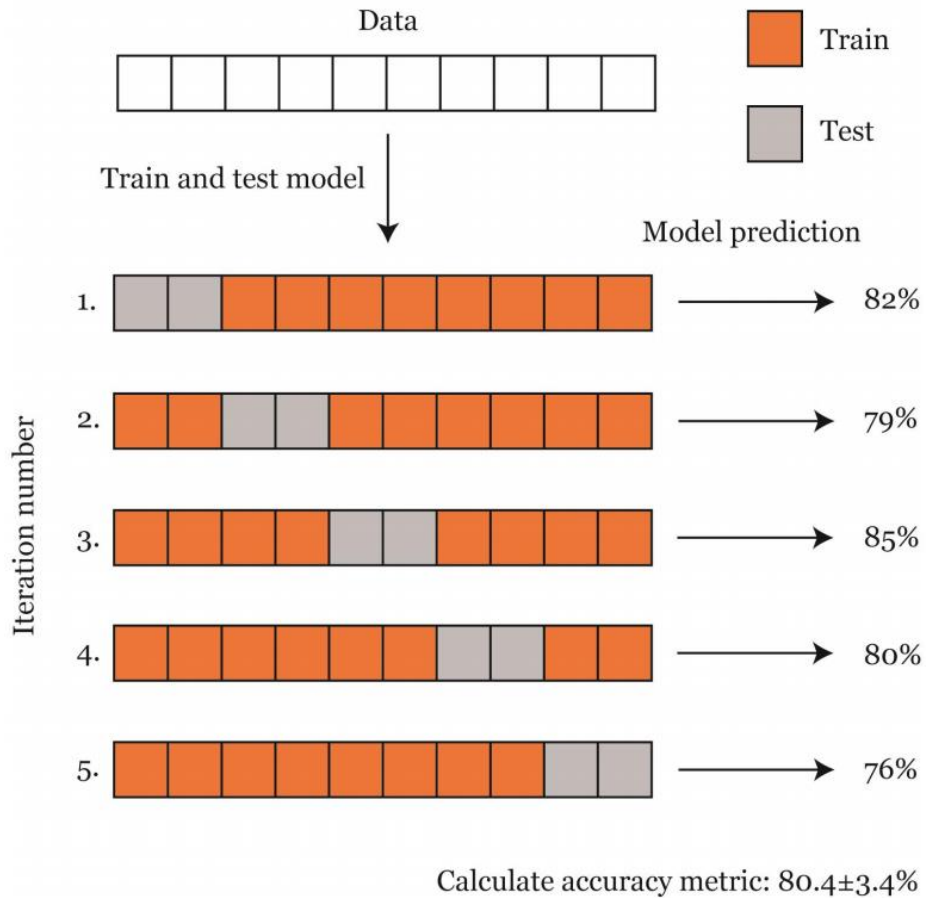
Spolehlivou metodou k ověření funkčnosti statistického modelu je křížová validace [19]. Tato metoda také zajistí, že nedojde k přetrénování algoritmu. Základní přístup této metody je takový, že model je vždy trénován na jiných datech, než na jakých je poté testován. Konkrétních postupů existuje více.

„K-folds cross validation“ funguje tak, že je množina vzorků pro testování rozdělena na k částí. Jedna z těchto částí je vždy vybrána jako testovací a model je natrénován na zbylých datech. Tento postup se opakuje s každou z těchto k částí jako testovací množinou. Výsledná úspěšnost modelu může být získána jako průměr úspěšnosti na jednotlivých dílčích testovacích množinách.

„Leave-one-out cross validation“ je speciálním případem „K-folds cross validation“, kdy k je rovno celkovému počtu prvků, model je tedy natrénován vždy na všech prvcích kromě jednoho, na kterém je testován. Tento postup se obvykle používá pro velmi malé množiny dat. Jeho výhodou je velká přesnost stanovení přesnosti klasifikátoru, nevýhodou mohou být velké odchylky dílčích testů a také velká výpočetní náročnost.

Při „Monte Carlo cross validation“ jsou data rozdělena na testovací a trénovací množinu náhodně. Náhodné rozdělení se obvykle vícekrát opakuje, doporučené je provést nejméně 10 opakování. Výhodou je, že velikost testovací a trénovací množiny není závislá na rozdělení do k skupin, nevýhodou však je to, že některé prvky nemusí být buď do testovací nebo trénovací množiny vůbec vybrány.

K-fold cross-validation



Obrázek 7: Postup křížové validace typu „K-folds“ [19]

5. Vytvoření modelů vad v prostředí MATLAB

Jelikož je velmi obtížné získat dostatečné množství dat ke každému typu vady, jsou v této práci na základě známých vlastností jednotlivých vad tyto vady modelovány pro účely následného trénování a testování klasifikátoru.

K vytvoření modelů je využit popis vad v [2] a zároveň zkoumání vlastností dostupných vzorků vad. Modely jsou zjednodušené tak, aby především zachytily základní zvukový projev vady a při poslechu se velmi blížily projevům skutečných vad. Vytvoření realistických matematických modelů vad by vyžadovalo složitý popis příčin a projevů vad.

Pro modelování vad existuje možnost tzv. aditivního modelu a substitučního modelu [9], kdy aditivní model znamená, že vada je přičtena k původnímu signálu, substituční model znamená, že signál je v místě vady kompletně nahrazen projevem vady. Z podstaty těchto modelů vyplývá, že aditivní model je lepší aproximací méně závažných vad, substituční model lépe vystihuje podstatu závažnější vady, kdy mechanické nebo jiné poškození způsobí úplnou ztrátu původního signálu na daném krátkém segmentu.

Aditivní model můžeme vyjádřit výrazem:

$$y_t = x_t + i_t n_t$$

Kde x_t je původní signál, n_t je samotné rušení (vada) a i_t „přepíná“ mezi hodnotami 1 a 0, a tedy tím, zda je vada přítomna či nikoliv [9].

V naší práci se modelování lupanců, škrτάku, škrábance, šumu a houkání řídí aditivním modelem. Modelování přeskočků a výpadků je vzhledem k podstatě těchto vad modelováno pomocí substitučního modelu.

Pro přidání vad do nahrávky je vytvořen skript *addDefectTypes.m*, který po spuštění celého skriptu, nebo případně pouze sekce pro příslušný typ vady, přidá vady do vybrané nahrávky. Přidání každého typu vady má vlastní funkci, jejím vstupem je kromě dalších parametrů, které se liší pro různé typy vad, počet vad, který chceme do nahrávky přidat. Pozice vad jsou v příslušné funkci vždy vybírány náhodně (pouze u škrábance za sebou následují vady v intervalu daném délkou otáčky). Zároveň jsou pozice vad při vytvoření nahrávky zapsány do souboru tak, aby s nimi bylo případně dále možné pracovat při zpětném testování klasifikace.

Vytvoření modelu každého typu vady je blíže popsáno v následujících odstavcích.

5.1. Houkání a šum

Pro simulaci houkání jsem využil filtrační syntézy, která se v oblasti syntézy zvuků používá často pro syntézu zvuků jako je vítr nebo mořský příboj [21]. Princip této syntézy je takový, že budící signál (v našem případě bílý šum, při využití u elektronických hudebních nástrojů to mohou být však i pulzy) je dále filtrován tak, abychom získali zvuk s požadovanou šířkou pásma a střední frekvencí.

Jelikož houkání se vyznačuje pulsující frekvencí, stanovil jsem řídicí vektor, podle kterého během každých 2 sekund signálu dochází k pulsování s rozsahem 30 Hz.

Vytvoření vektoru frekvence v MATLABu ukazuje následující úryvek kódu. Hodnoty ve vektoru X ukazují umístění bodů v čase (hodnota 0 udává začátek segmentu, hodnota 1 jeho konec), hodnoty vektoru Y ukazují na hodnotu frekvence v daném intervalu ve kterém frekvence pulsuje, 0 by tedy byla hodnota minimální frekvence a 1 hodnota maximální frekvence.

```
X = [0 .15 .2 .3 .4 .5 .6 .7 .85 .95 1];  
Y = [0.8 .3 .1 .3 .1 .3 1 .8 1 .5 0.8 ];
```

```
Fmax=Fmin+30;  
humm=(Fmax-Fmin)*Y+Fmin;
```

```
fr=interp1(X,humm,nT/nT(end));
```

Pro rozlišení mezi houkáním a šumem jsem pro každý případ zvolil jinou šířku pásma a jinou střední frekvenci, pro houkání jsem zvolil frekvenci pulsující v intervalu 250 – 280 Hz a velmi úzké pásmo 10 Hz, kde již budeme moci slyšet místo šumu houkání o dané frekvenci. Jelikož šum na gramofonové desce reálně není bílý šum, byl i ten vytvořen při simulaci stejným principem filtrace, ale abychom zachovali charakter šumu, je pro šum použita střední frekvence 600 Hz a šířka pásma 1000 Hz.

Šum a houkání jsou při simulaci přidány s požadovanou intenzitou do masteru gramofonové desky. Jelikož na masteru není obvykle přítomné dlouhé ticho na začátku (tedy pro naši úlohu podstatný nájezd), je před každý master přidáno 10 sekund ticha, a do výsledné nahrávky je poté přidán šum nebo houkání.

5.2. Škrták

Škrták se sestává ze série lupanců s různou amplitudou. Při zobrazení průběhu každého z jednotlivých kliků můžeme vidět, že se jedná o velmi krátký klik s délkou většinou okolo 3 vzorků.

Při vytváření modelu a přidání uměle vytvořeného škrtáku do nahrávky jsem tedy postupoval tak, že byla každé vadě přiřazena náhodně zvolená délka v okolí 2 sekund s rozptylem 500 vzorků (tedy 2 x vzorkovací frekvence \pm 500 vzorků). Poměrně dlouhá doba je stanovena proto, abychom vytvořili model, kde bude vždy i poslechem zřetelný praskot a který bude jasně rozeznatelný od vady s kratším průběhem. Do tohoto intervalu je umístěno náhodně 25-35 lupanců. Každý z těchto lupanců se skládá ze tří vzorků, jejich hodnota je zvolena náhodně.



Obrázek 8: Průběh reálného škrtáku na gramofonové desce



Obrázek 9: Průběh uměle vytvořeného škrtáku

5.3. Lupanec

Lupanec se posluchači jeví jako krátké ostré lupnutí bez dalšího dozvuku. Při zkoumání průběhu reálného lupance zjistíme, že se jedná o několik jednotek až desítek vzorků, se zřetelným náběhem a odezněním, v jeho průběhu se může objevit několik lokálních minim a maxim.

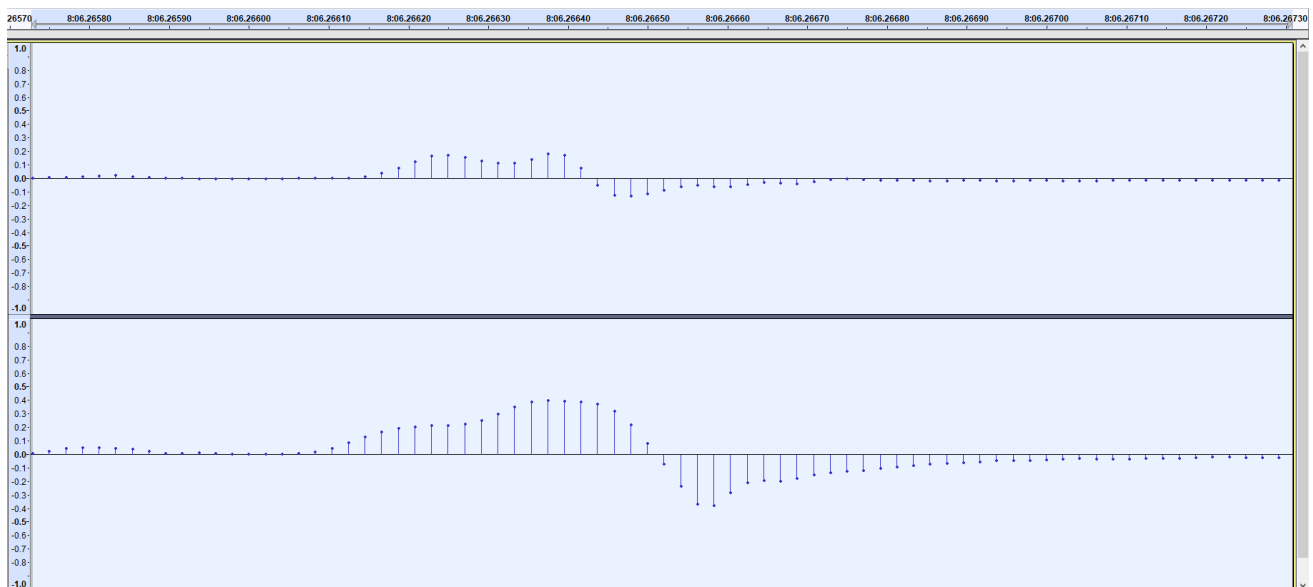
Modelování vady je v MATLABu provedeno následovně:

```
randAmp1 = rand;  
randAmp2 = rand;
```

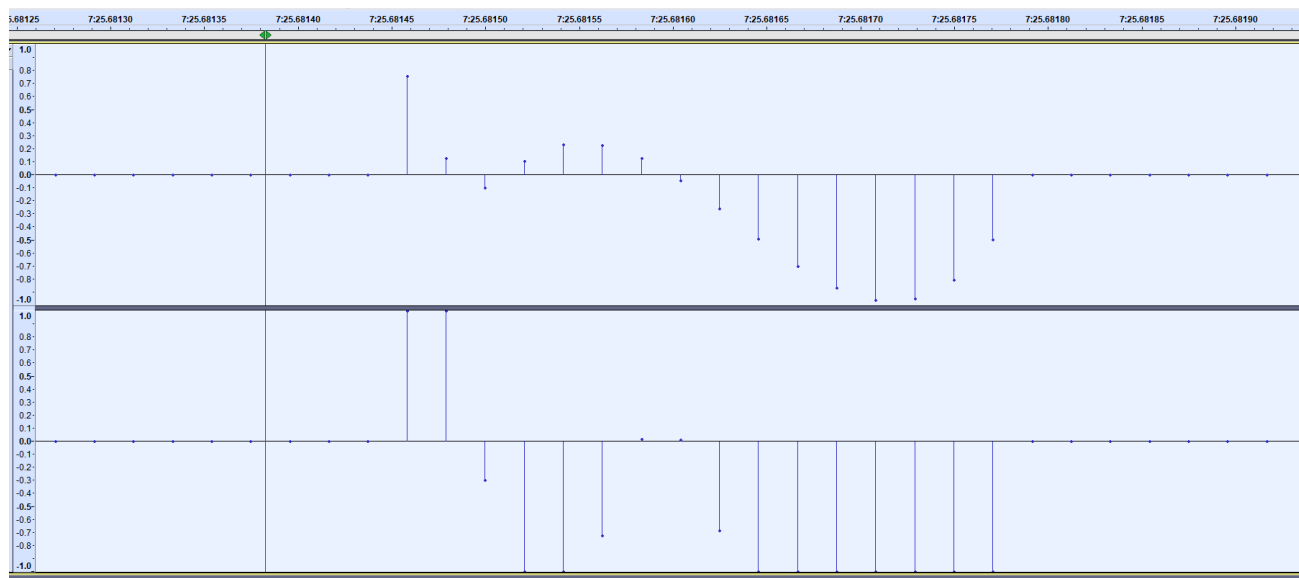
```
X = [0 0.1+rand*0.1 0.2+rand*0.3 0.5+rand*0.1 0.6*rand*0.2 1];  
Y = [0 rand*randAmp1 randAmp1 -rand*randAmp2 randAmp2 0];
```

Obálka signálu je tedy vytvořena stanovením několika bodů a jejich proložením s pomocí lineární interpolace (funkce *interp1* v MATLABu). Hodnoty ve vektoru X ukazují umístění bodů v čase (hodnota 0 udává začátek vady, hodnota 1 její konec), hodnoty vektoru Y ukazují amplitudu těchto bodů. Lupanec tedy začíná na nulové amplitudě, následují dva náhodně zvolené vrcholy s kladnou hodnotou a třetí vrchol se zápornou hodnotou, po kterém dojde k odezdění lupance a návrat na nulovou hodnotu.

Jako délku lupance jsem použil náhodnou délku v intervalu 10 – 50 vzorků.



Obrázek 10: Průběh reálného lupance na gramofonové desce



Obrázek 11: Průběh uměle vytvořeného lupance

5.4. Přeskok

Vada typu přeskok se projevuje především tím, že je náhle reprodukován nenavazující úsek. Pro simulaci tohoto artefaktu jsem vytvořil funkci, která na náhodné pozici přeskočí segment s délkou náhodně zvolenou z intervalu 0,5 s – 2 s. Simulujeme tedy přeskok do jiné drážky, jelikož se řádově pohybujeme okolo času jedné otáčky (většinou přibližně 1,3 s). Přeskok na reálné desce může způsobit přeskočení kratšího ale i výrazně delšího intervalu, pro naši simulaci je však lepší mít interval dlouhý natolik, že bude i při poslechu nenavazující úsek naprosto zřejmý. Příliš dlouhý interval by nám ale při přidání většího počtu vad způsoboval zkracování nahrávky a nepřinesl by pro následné trénování a testování klasifikátoru žádnou výhodu.

Jelikož během přeskoků dojde k tomu, že jehla přeskočí jednu nebo více drážek, dojde po několik milisekund k tomu, že není reprodukován vůbec požadovaný hudební signál, ale je slyšet pouze prasknutí. Pro tento účel jsem v místě přeskoků přidal náhodně dlouhý úseku bílého šumu, s dobou trvání 1-500 vzorků a náhodně zvolenou amplitudou.

5.5. Nedolisy

Vada typu nedolisy se posluchači jeví velmi podobně jako škrťák, ale praskot se objevuje pouze v pravém kanále. Nedolisy jsou tedy modelované stejně jako škrťák, ale levý kanál je ponechán a není nijak upraven.



Obrázek 12: Průběh modelu nedolisu

5.6. Škrábanec

Zásadním projevem škrábance je, že se vyskytuje na otáčku. Vytvoření modelu vady tedy probíhá tak, že nejprve určíme kolik projevů vady bude následovat za sebou. Bylo by možné vybrat náhodnou hodnotu, pro zjednodušení úlohy jsem ale stanovil pevnou hodnotu 5 vad na otáčku za sebou. Ze známé rychlosti otáčení vypočteme interval mezi jednotlivými projevy na otáčku jako

$$\text{onTurnDistance} = (60/\text{RPM}) * f_s \text{ [vzorky]}$$

Jsou tedy stanoveny pozice jednotlivých vad, pozice první vady v řadě vad na otáčku je vybrána náhodně a k ní jsou přiřazeny vady na otáčku.

Jednotlivé vady jsou následně modelovány jako samostatné lupance, jak je již popsáno v 5.2.

5.7. Výpadek

Vada typu výpadek se projevuje tím, že po krátký časový úsek audio signál chybí.

Funkce pro přidání mezer tedy na pozici vady vyjme signál a nahradí ho tichem. Jelikož na reálné desce by v mezeře byla vždy přítomna aspoň nějaká hodnota šumu, je do mezery vložen bílý šum s nízkou úrovní, normalizovaný na maximální hodnotu 0,01.

Délka výpadku je vždy náhodná hodnota v předem daném intervalu, který jsem stanovil jako 0,1-1 s.

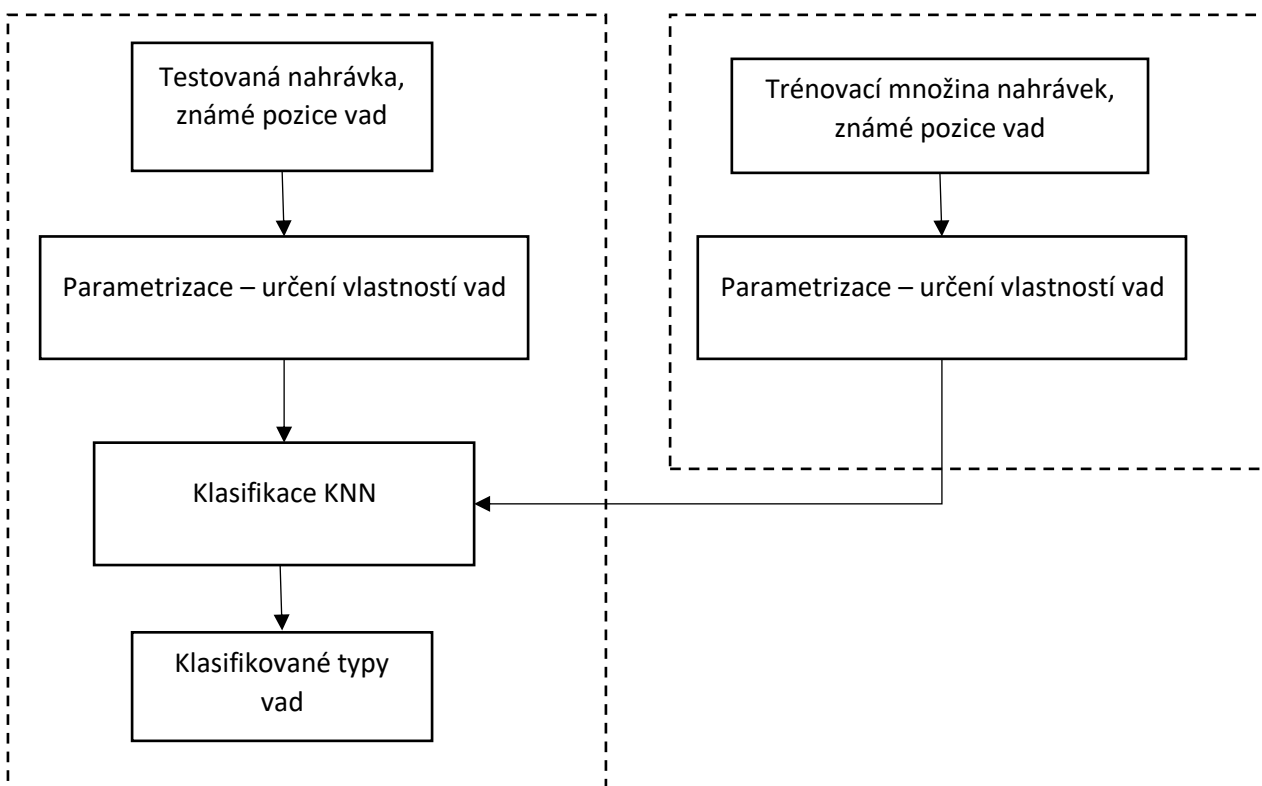
6. Implementace klasifikačního algoritmu v prostředí MATLAB

V této části práce se zabývám návrhem klasifikačního algoritmu, který bude schopný rozlišit jednotlivé typy vad. Na základě konzultace s firmou GZ Media jsem se rozhodl detekovat typy vad, které jsou relativně jednoznačně rozlišitelné, tyto typy byly popsány již v předchozích kapitolách.

Pro zjednodušení úlohy není v algoritmu řešená samotná detekce vad, jelikož pro tuto úlohu již řešení existují (vizte kapitola 3). Za pozice vad na deskách jsou považovány pozice uvedené v protokolech firmy GZ Media, případně známé pozice uměle přidanych vad.

Jako parametry jednotlivých dat jsou detekovány jejich charakteristické vlastnosti, podle kterých by je bylo možné i podle poslechu rozlišit.

Pro samotnou klasifikaci, tedy detekci konkrétních vad na základě detekovaných vlastností vad, je použit algoritmus KNN. Algoritmus je tedy nejdříve natrénován na dostatečném množství vzorků. Vzhledem k malému množství dostupných vzorků reálných vad jsou jako trénovací množina použity pouze uměle vytvořené vady. Algoritmus je následně testován jednak na uměle vytvořených vzorcích a jednak na reálných vadách.



Obrázek 12: Základní blokové schéma navrženého klasifikačního algoritmu

6.1. Parametrizace – detekování vlastností vad

Postup klasifikace v této práci je postaven na popisu zvukových vlastností vad, které jsou pro různé typy charakteristické, a podle kterých by je podle poslechu zvládl rozlišit i posluchač. Jednotlivé typy vad se projevují charakteristickým způsobem a například zkušený pracovník gramofonových závodů je schopen je podle poslechu klasifikovat. Jednotlivé rozlišitelné projevy vad jsou popsány v následujících odstavcích.

6.1.1. Výskyt vady na otáčku

Jedním z popsatelných příznaků je vada vyskytující se na otáčku. Tento typ vady je charakteristický pro škrábanec na desce, může se ale vyskytovat ale i u dalších typů (škrťák, nedolisy).

K nalezení takovéto vady pomocí algoritmu zjistíme interval mezi po sobě jdoucími vadami. Jelikož u každé desky známe počet otáček za vteřinu, jsme odsud schopni určit, jak daleko jsou od sebe vzdáleny dvě po sobě jdoucí vady vyskytující se na otáčku:

$$t = \frac{60}{rpm} [s]$$

Kde t je čas mezi po sobě následujícími vadami a rpm (*revolutions per minute*) udává počet otáček za minutu.

Je očekávatelné, že vzdálenost mezi těmito dvěma vadami nikdy nebude zcela přesně tento stanovený interval. Jednak nemusí být zcela přesně identifikována pozice vady, zároveň bude škrábnutí desky přes několik drážek většinou ne zcela kolmé na desku. Musíme tedy stanovit i práh tolerance, jaké odchylky od očekávané vzdálenosti mezi vadami budeme ještě akceptovat.

6.1.2. Délka a dozvuk vady

Typ vady můžeme také rozlišit podle toho, zda se jedná jen o krátké ostré lupnutí, nebo zda se posluchači jeví jako delší lupnutí, případně zda se subjektivně jeví jako lupnutí, po kterém následuje další dozvuk.

V akustice existují metody detekování například dozvuku místnosti, běžná je například metoda Energy Time Curve (ETC) [22], kdy je měřen pokles energie za čas. Vzhledem k tomu, že dozvuk je především opožděná kopie původního signálu, je k jeho detekci a i případnému odstranění také často používaná autokorelační funkce [23].

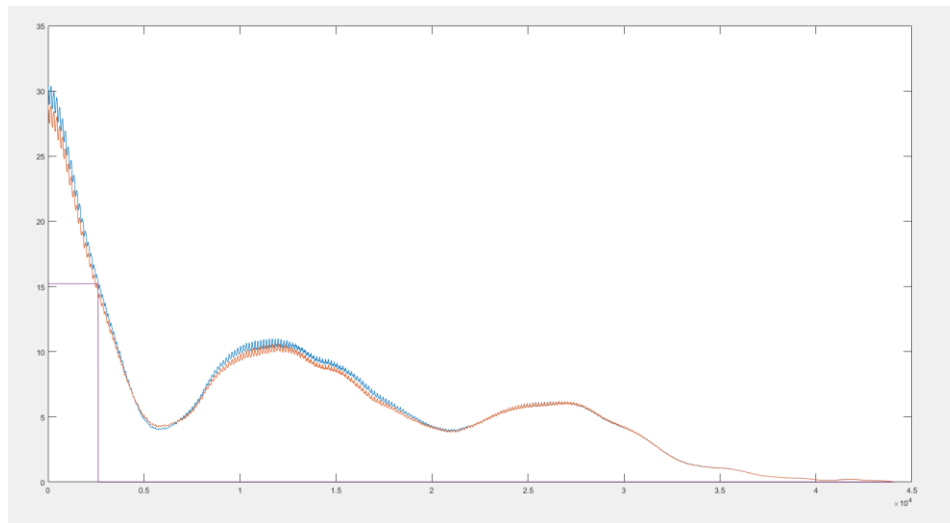
V našem případě je úloha komplikovanější oproti dozvuku místnosti především tím, že máme v okolí vady další hudební signál, a pokles energie tedy nemusí být vždy tolik zřejmý.

Rozhodl jsem se vyzkoušet různé metody, jak rozlišit, jestli se jedná o krátkou a ostrou vadu, nebo vadu obsahující dozvuk, případně vadu s delším trváním.

Detekce dozvuku s použitím autokorelační funkce

Při vyzkoušení metody autokorelační funkce jsem použil takový postup, kde jsem nejdřív stanovil krátký úsek nahrávky v okolí detekované vady, se kterým bude autokorelační funkce pracovat. Následně je spočítána autokorelační funkce v tomto úseku. Autokorelační funkce je dále vyhlazena filtrem využívajícím průměrování. Dále detekujeme místo poklesu autokorelační funkce pod stanovený práh, tedy místo, kdy stanovíme, že došlo k odeznění vady, respektive odeznění dozvuku. Délka autokorelační

funkce nad daným prahem je poté porovnána s délkou, kterou nastavíme jako práh pro délku příslušící vadě obsahující dozvuk.



Obrázek. 13: Průběh autokorelační funkce a nalezení bodu poklesu pod stanový práh

Detekce délky vady s použitím shlukování

Následující postup zkoumá především délku vady a rozložení vzorků pravděpodobně zasažených vadou. Místa zasažena vadou jsou zjištěna pomocí vlnkové transformace. Následně dojde k nalezení shluků takto detekovaných „vadných“ vzorků. Následně jsem se rozhodl sledovat dva ukazatele a vyzkoušet jejich schopnost rozlišit od sebe typ vad. První je délka nejdelšího shluku, kterou zjistíme v ideálním případě trvání vady. Dalším ukazatelem je průměrná délka shluku, která nám dá informaci o tom, zda se zde vyskytují krátce trvající lupance, nebo lupance trvající delší časový úsek.

6.1.3. Série lupanců za sebou

Některé typy vad, především však škrták, jsou charakteristické tím, že se neprojeví jen jedním krátkým lupancem, ale praskotem a více lupanci těsně za sebou.

Postup identifikace takového místa je takový, že je nejprve stanoveno krátké okolí vady, kde by mohly být přítomné další lupance. V tomto okolí je poté použit algoritmus detekce samostatných lupanců pomocí vlnkové transformace. Funkce pro detekci samostatných lupanců pomocí vlnkové transformace je převzata z bakalářské práce [8].

Po nalezení lupanců je určeno, kolik vzorků v daném úseku je detekováno jako obsahující chybu a kolik nikoliv, a je stanoven poměr těchto dvou údajů. Dále je experimentálně určen práh a v případě hodnoty poměru nad tímto prahem je vada označena jako vada s více lupanci za sebou.

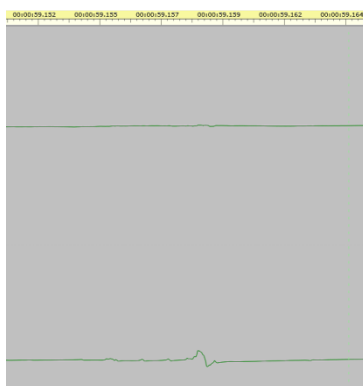
Takto získaný parametr sice dává jen informaci o tom, kolik je „vadných“ vzorků a kolik nikoliv, a nedává další rozlišení o tom, jak jsou uskupeny a kolik je tedy skutečně lupanců, ale v kombinaci s parametry popsány v předchozím bodě již tuto informaci mít budeme.

6.1.4. Výskyt vady v pravém kanále

Vada přítomná v pravém kanále je typická pro nedolisy, může se opět ale vyskytovat u i jiných typů vad. Pro určení, zda se vada nachází v pravém kanále, se nabízí více možných způsobů řešení. Ve své práci jsem navrhl tři parametry pro detekci vady v pravém kanále.

Porovnání amplitudy

Nejjednodušším přístupem pro určení, zda se vada vyskytuje jen v jednom kanálu, je porovnání amplitudy v místě vady v obou kanálech. Abychom zjistili informaci o amplitudě v místě vady, uděláme průměr absolutní hodnoty v okolí detekované vady a zjistíme rozdíl této průměrné hodnoty pro oba kanály. V případě hodnoty větší než je námi předem zvolená hodnota detekčního prahu, můžeme vadu považovat za vadu pouze v jednom kanále.



Obrázek. 14: Příklad průběhu signálu při výskytu vady pouze v pravém kanále

Porovnání počtu lupanců

Vada typu nedolisy se ale sestává ze série lupanců v pravém kanále. Vzhledem ke krátkému trvání jednotlivých lupanců není jisté, že dojde k zásadnímu nárůstu energie v okolí vady tak, abychom ji mohli detekovat postupem popsaným v předchozím bodě. Rozhodl jsem se tedy vyzkoušet i další způsob, kdy vadu detekuji podobně jako více lupanců za sebou pomocí vlnkové transformace. Detekuji, jaký je výskyt lupanců v každém kanále zvlášť, a pokud je v pravém kanále více lupanců než v levém o rozdíl daný stanoveným prahem, je vada detekována jako obsahující více lupanců v pravém kanále.

Porovnání délky vady

Tento postup využívá podobné metody jako výše popsaná metoda detekce dozvuku nebo délky vady pomocí nalezení shluku maximální délky, pouze s tím rozdílem, že je zvlášť získaná maximální délka shluku v pravém i v levém kanále a ty jsou poté porovnány. Pokud je v pravém kanále shluk delší než v levém o stanovený práh, je vada detekována jako obsahující déle trvající vadu v pravém kanále.

6.1.5. Šum a houkání v nájezdu

Pokud se v nájezdu objevuje rušivý zvuk, může být buď šumového charakteru, nebo se může jednat ale i o houkání s danou pulsující frekvencí.

Navržená detekce popsaných případů je provedena ve dvou krocích. Nejdříve je určena střední kvadratická hodnota v měřeném úseku nájezdu a ta je porovnána rovněž se střední kvadratickou hodnotou v celá nahrávce. Jejich poměr tedy udává přibližně poměr nežádoucího šumu (či houkání) vůči užitečnému signálu, tedy převrácenou hodnotu běžně používaného parametru signal-to-noise ratio (SNR).

$$\frac{1}{SNR} = \frac{RMS_{signal}}{RMS_{segment}}$$

$\frac{1}{SNR}$ udává poměr šumu (houkání) a užitečného signálu, RMS_{signal} je střední kvadratická hodnota v celém úseku signálu na nahrávce a $RMS_{segment}$ je střední kvadratická hodnota v měřeném segmentu nájezdu.

Pro rozhodnutí, zda se jedná o šum či houkání, je dále použita autokorelační funkce. U autokorelační funkce jsou nalezena lokální maxima, která můžeme očekávat spíše u periodického signálu, tedy u houkání.

Pro nalezení maxim je použita zabudovaná funkce MATLABu *findpeaks.m*. Jako její vstupní argumenty musíme specifikovat argumenty jako minimální vzdálenost mezi vrcholy a tzv. „Minimal peak prominence“, která udává, do jaké míry musí lokální maximum být vyšší než body v jeho okolí.

Abychom stanovili ukazatel míry korelace v tomto úseku, který bude nabývat hodnot v dobře interpretovatelném rozmezí, byl pro tento ukazatel stanoven při implementaci v MATLABu následující vztah:

$$r = \frac{m}{\frac{1}{SNR} * l}$$

Kde r je výsledný koeficient korelovanosti signálu, m je počet nalezených lokálních maxim, $\frac{1}{SNR}$ udává poměr šumu (houkání) a užitečného signálu a l udává délku celého měřeného úseku.

6.1.6. Mezery v signálu

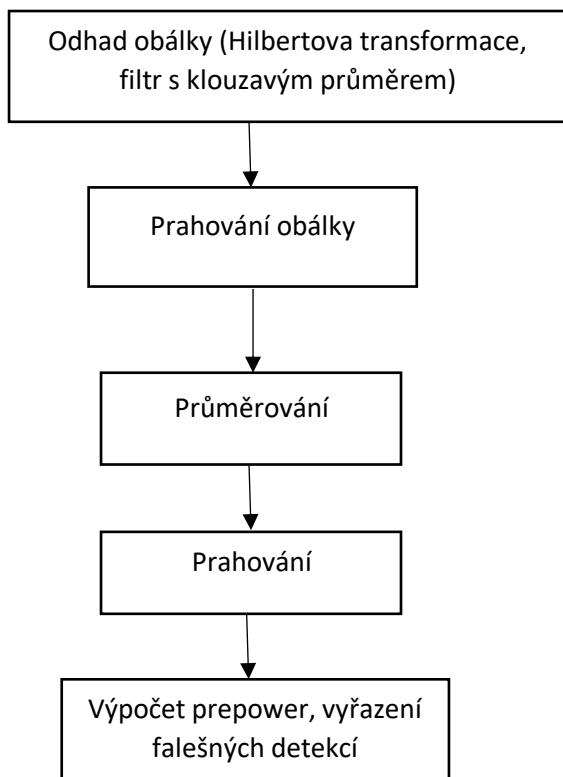
Vada typu výpadek se projevuje tak, že audio signál může po určitou dobu chybět (být nahrazen tichem). Pro detekci mezer v signálu byl použit postup popsaný v práci [12], tento postup zde byl popsán s ohledem na detekci mezer například na CD.

V první fázi je odhadnuta obálka signálu. K tomuto účelu jsem využil Hilbertovu transformaci a následně jsem ještě obálku vyhladil filtrem s klouzavým průměrem o délce okna 4001 vzorků. Obálka je dále prahována, získáme tedy jen vzorky nad daným prahem. Abychom vyloučili lokální výkyvy obálky, je prahovaná obálka ještě jednou průměrována stejným filtrem s klouzavým průměrem a znovu prahována.

Tímto způsobem dosáhneme rozpoznání úseků obsahujících ticho. Dále musíme rozlišit mezi žádoucím tichem (pomlka ve skladbě, mezery mezi jednotlivými písněmi apod.) a skutečnou vadou. V práci [12] je pro tento účel použit výpočet takzvané prepewer. Ta spočívá ve výpočtu energie signálu v krátkém úseku před úsekem ticha.

$$prepower = \text{mean}(\text{abs}(\text{prepowerframe}))$$

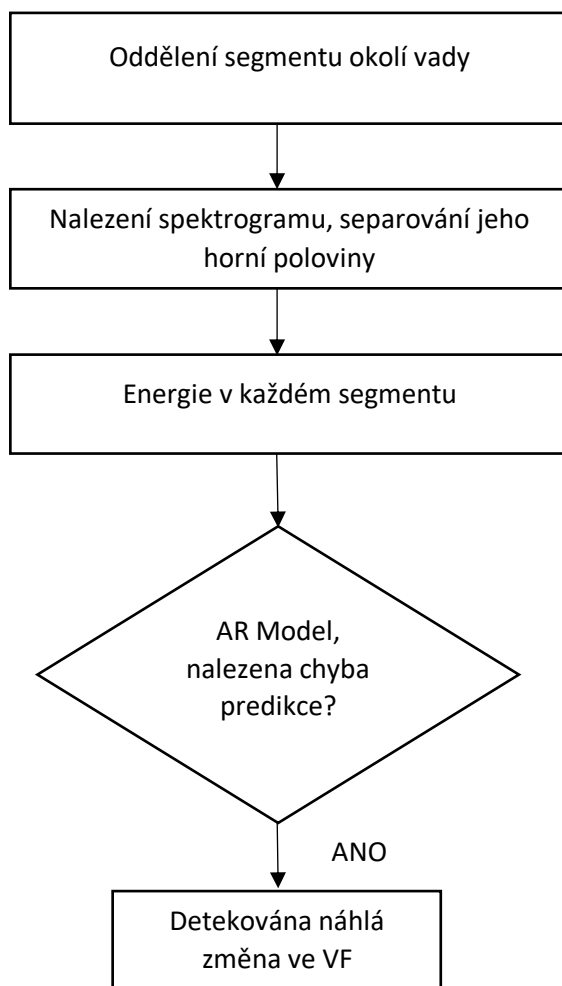
Pokud je tato hodnota menší než daný práh, můžeme mezeru považovat za chybu. Tento postup vychází z toho, že zamýšlené pauzy v hudbě nemají tak rychlý nástup (neboli rychlé odeznění signálu) jako pauzy způsobené chybou.



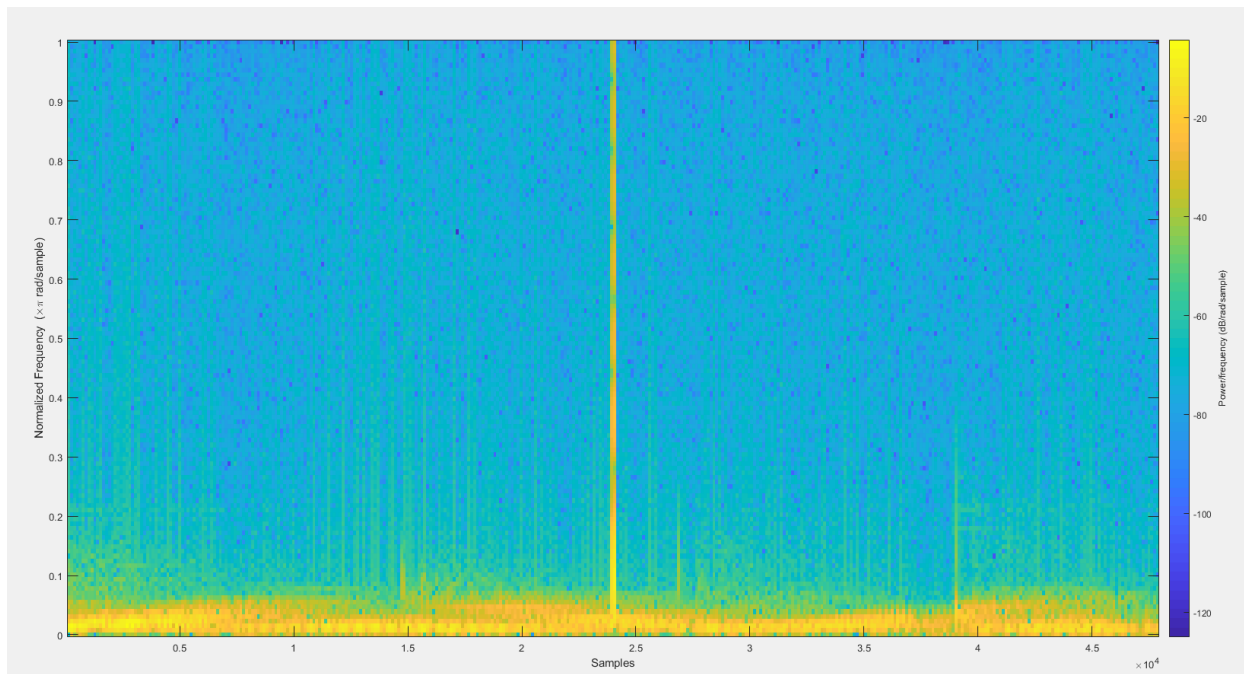
Obrázek 15: Blokové schéma algoritmu pro detekci mezer v signálu

6.1.7. Změna ve spektru nahrávky

Další parametr, který jsem se rozhodl zkoumat a na jeho základě zkusit rozlišit typy vad je informace o změně vysokofrekvenční části spektra nahrávky. Typy vad, které se vyznačují ostrým lupnutím, mívají ploché spektrum [9] a tedy obsahují frekvence i v části spektra, kde se jinak v nahrávce vyskytuje minimální energie. Rozhodl jsem se zkusit nalézt změnu energie ve vysokých frekvencích tak, že jsem nejprve spočítal spektrogram okolí vady, ve spektrogramu jsem vzal jen jeho horní polovinu, a provedl jsem určení celkové energie v každém okně. Tím získáváme informaci o vývoji vysokofrekvenční energie v čase. Abych dále zjistil neočekávanou změnu v tomto průběhu, je k detekování této změny použit AR model.



Obrázek 16: Blokové schéma algoritmu pro detekci náhlé změny ve vysokých frekvencích



Obrázek 17: Spektrum nahrávky obsahující ostrý lupanec

6.2. Test detekce vlastností vad

Rozhodl jsem se testovat klasifikaci vad v několika fázích. V první fázi bylo důležité se ujistit, že základní vlastnosti vad jsme schopni rozpoznat. Proto jsem nejprve testoval rozpoznávání vlastností uměle přidaných vad, v této fázi jsem také nastavil prahy a další parametry jednotlivých funkcí pro detekci vlastností vad. V další fázi jsem již testoval nalezení vlastností vad obklopených okolním zvukem, kdy byly uměle vytvořené vady již přidány do masterů gramofonových desek. Poté jsem již přešel k vyhodnocování těchto vlastností pomocí klasifikátoru KNN a určování konkrétního typu rozpoznávané vady.

6.2.1. Detekce vlastností vad bez přítomnosti okolního zvuku

Pro tento účel byly uměle vytvořené vady přidány do “nahrávky” obsahující pouze ticho, tedy nebyly ovlivněny okolní hudbou či jiným obsahem nahrávky.

Pouze u výpadku a přeskočku (přidání mezer do nahrávky a reprodukci nenavazujícího úseku nahrávky) nemá tento postup smysl, a proto nebyly tyto typy vad v této fázi testovány.

Naopak byly do tohoto testu pro porovnání přidány i „reálné“ vady získané oddělením vady lupanec a vady škrťák z dostupných nahrávek obsahujících tyto vady.

Klasifikovány byly v této fázi pouze jednotlivé atributy nahrávek, jelikož potřebujeme být nejprve schopni rozpoznat tyto vlastnosti, a poté můžeme navázat pojmenováním vady na základě těchto vlastností.

Výsledky testování ukazuje tabulka 1.

Tabulka 1: Výsledky testování vlastností vad bez okolního zvuku

	Škrták - Model	Škrábanec - Model	Nedolisy - Model	Lupanec - Model	Škrták - Reálný	Lupanec - Reálný
Počet přidanych vad	20	20	20	20	20	20
Detekováno více lupanců	95 %	5 %	100 %	0 %	100 %	15 %
Detekován dozvuk (korelace)	35 %	0 %	10 %	0 %	95 %	0 %
Detekován dozvuk (shluk - průměr)	100 %	0 %	100 %	0 %	100 %	0 %
Detekován dozvuk/delší trvání vady (shluk - maximum)	0 %	95 %	0 %	0 %	90 %	100 %
Detekována vada v pravém kanále	0 %	0 %	75 %	0 %	5 %	5 %
Detekováno více lupanců v pravém kanále	0 %	0 %	100 %	0 %	85 %	10 %
Detekováno delší trvání vady v pravém kanále	0 %	0 %	100 %	0 %	5 %	0 %
Detekována vada na otáčku	0 %	100 %	0 %	0 %	0 %	0 %
Detekována změna ve VF	60 %	55 %	0 %	40 %	95 %	85 %

Z výsledků můžeme vidět, že při přidání vad do ticha dopadly výsledky dle očekávání. Model škrtáku byl ve většině případů detekován jako obsahující sérii více vad a obsahující dozvuk. Model škrábance byl detekován jako vyskytující se na otáčku.

U modelu lupance nebyla dle očekávání detekována ani jedna ze zkoumaných vlastností.

Pro reálný škrták a reálný lupanec je situace složitější, neboť jejich vystřížení nebylo díky okolní hudbě na desce možné zcela dokonale a tak nejsou ideálním případem těchto vad. Nicméně i tak byly pro škrták detekovány očekávané vlastnosti, ve většině případů dopadl dle očekávání i reálný lupanec.

U detekce vady v pravém kanále pomocí rozdílu úrovně mezi kanály se sice podařilo dosáhnout rovněž relativně dobrých výsledků, zároveň se ale ukázalo, že je pro dosažení potřeba použít velmi nízkou hodnotu prahu, která na nahrávce obsahující hudbu již jistě nebude použitelná.

U dalších dvou metod detekce vady v pravém kanále se podařilo dosáhnout lepší úspěšnosti, zároveň lze u těchto metod očekávat lepší výsledky i při aplikování na vadu obklopenou okolním zvukem.

6.2.2. Stanovení prahů a dalších parametrů pro detekční metody

Při výše popsaném testu jsem také zkoumal různé hodnoty prahu pro jednotlivé vlastnosti a sledoval, které hodnoty budou schopny vady nejlépe rozlišit. Hodnota prahu je však důležitá především pro test a zhodnocení, zda je daná vlastnost skutečně detekována pro očekávané vady. Pro pozdější vstup do klasifikátoru KNN, ke kterému mají být vlastnosti využity ke klasifikaci konkrétního typu vady, již nebude hodnota prahu tak podstatná, jelikož vstupem budou neprahované hodnoty.

Pro vlastnost více lupanců za sebou jsem zkoušel různou hodnoty prahu poměru mezi vzorky obsahujícími vadu a těmi vadu neobsahujícími. Nejlepší rozlišení jsem dosáhl s hodnotou 0,003. Malá hodnota zde není nijak překvapující, jelikož počet vzorků obsahujících vadu je i v okolí vady mnohonásobně menší než celkový počet vzorků, tudíž hodnota poměru je velmi malá. Délku okolí vady jsem zvolil 2 sekundy, protože to je i délka vady, se kterou jsem pracoval při simulaci.

Pro vlastnost detekovaného dozvuku pomocí korelace jsem nejvhodnějších výsledků dosáhl s hodnotou prahu poklesu 0,35 a při požadované délce úseku autokorelační funkce nad tímto prahem 0,5 sekundy. Při snižování prahu a délky úseku sice můžeme pro tento test dosáhnout většího množství pozitivně detekovaných výsledků pro škrťák a nedolisy (tedy vady kde dozvuk očekáváme), ale roste již falešná detekce u ostatních typů vad a později by v okolí okolního zvuku spíše hrozila falešná detekce.

Pro detekovaný dozvuk pomocí metody shlukování s počítáním průměru shluků jsem nejlepší výsledků dosáhl s délkou okolí vady 2 sekundy, prahem 5 vzorků průměrné délky shluku a minimální vzdáleností 10 vzorků mezi shluky. Při využití metody délky maximálního shluku jsem se snažil aby algoritmus zachytil celou vadu jako jeden shluk, a nastavil jsem tedy stejně jako v předchozím případě délku okolí 2 sekundy, práh maximální délky nalezeného shluku 0,5 sekundy a maximální rozestup mezi vzorky v shluku 0,25 sekundy.

Jak bylo zmíněno v předchozím popisu výsledků, pro detekování rozdílu mezi vadami na základě průměrné absolutní hodnoty amplitudy v daném místě bylo třeba nastavit velmi nízký práh, konkrétně 0,00015.

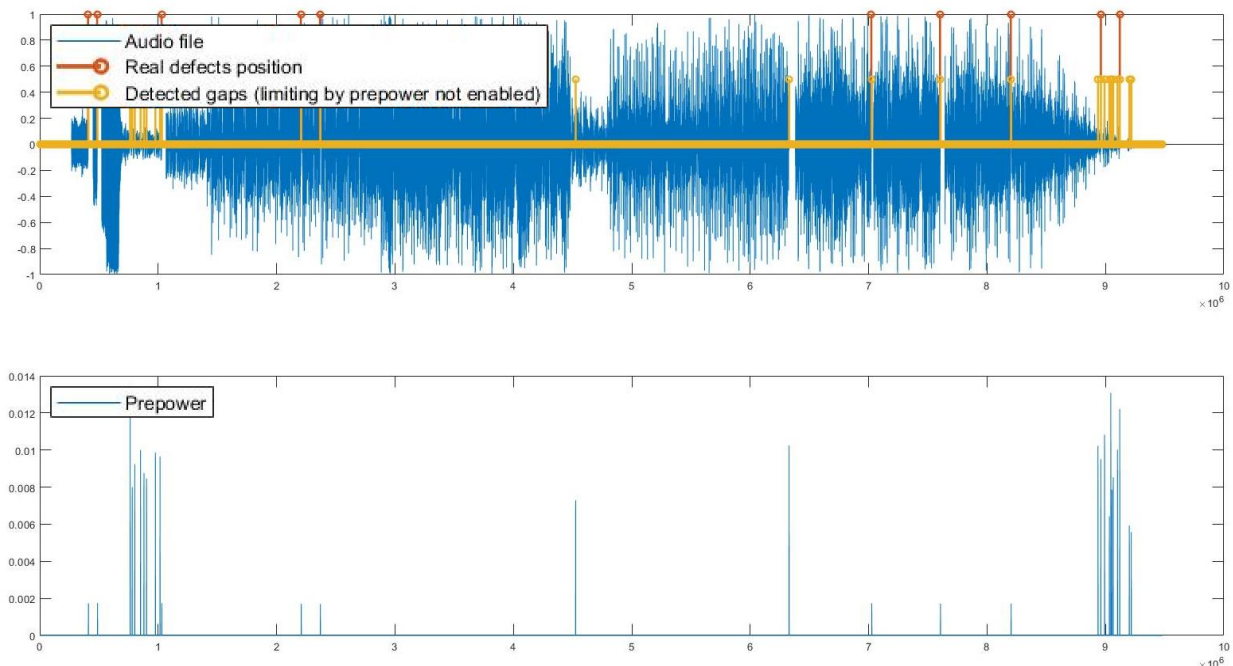
Při detekce lupanců v pravém kanále jsem nejlepší výsledků dosáhl s okolím vady 2 sekundy a prahem rozdílu poměru 0,0002. Malá hodnota je opět dána tím, že již samotný poměr vzorků detekovaných jako vadných vůči ostatním vzorkům je velmi malá hodnota. V podobném řádu se tedy musí pohybovat i práh rozdílu těchto hodnot.

Při detekci vad na otáčku jsem použil toleranci rozdílu od očekávaného intervalu 0,1 sekundy.

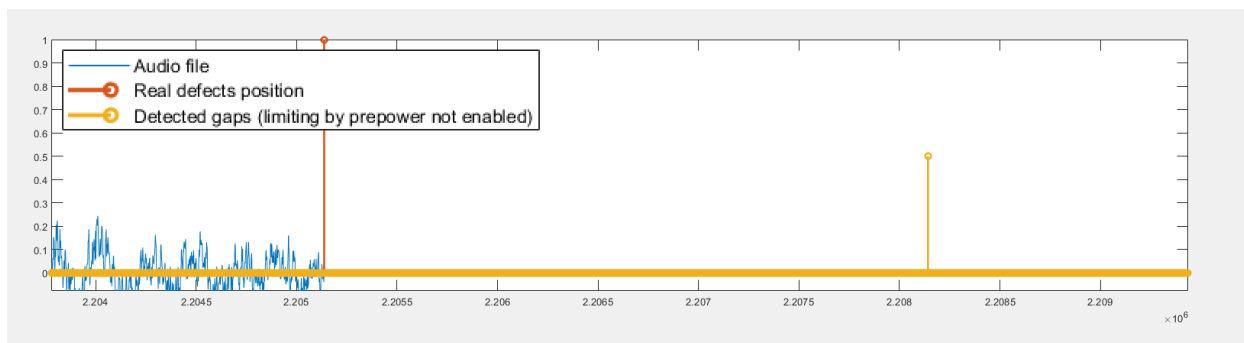
6.2.3. Test algoritmu detekce mezer

Algoritmus pro detekci mezer jsem testoval s nastavením délky průměrovacího filtru na hodnotu 4001 vzorků, s délkou segmentu pro prepover 2000 vzorků a s prahem pro prahování obálky 0,02.

Při prvotním vyzkoušení algoritmu jsem narazil na problém, kdy došlo k mnoha falešně pozitivním detekcím. Z hodnot prepover na druhém grafu je patrné, že by limitace pomocí vhodně nastaveného prahu prepover nebyla možná, neboť je vyšší hodnota v případech, kdy k detekci dojít nemá.



Obrázek 18: Prvotní test detekce mezer, špatná detekce prepower



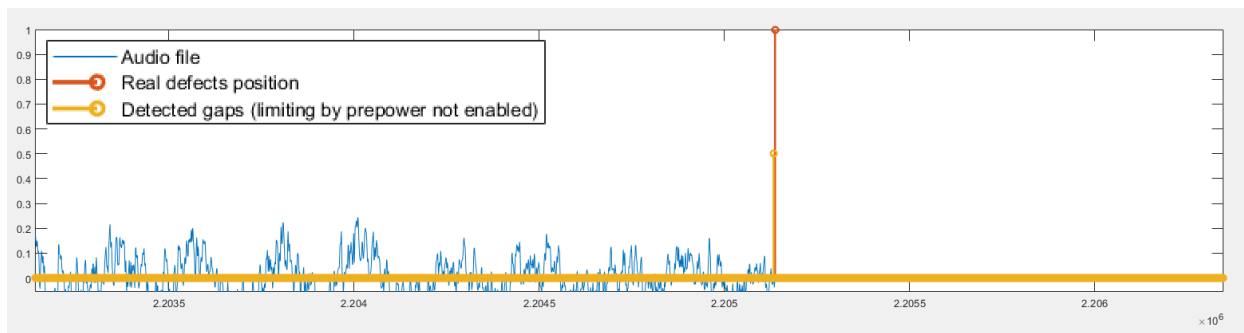
Obrázek 19: Prvotní test detekce mezer, rozdíl mezi detekovanou a skutečnou pozicí vady

Z obrázku 19 je patrné, že k nalezení malé hodnoty prepower došlo kvůli tomu, že hlazení pomocí průměrování způsobilo posun hodnoty detekované vady od skutečného místa vady.

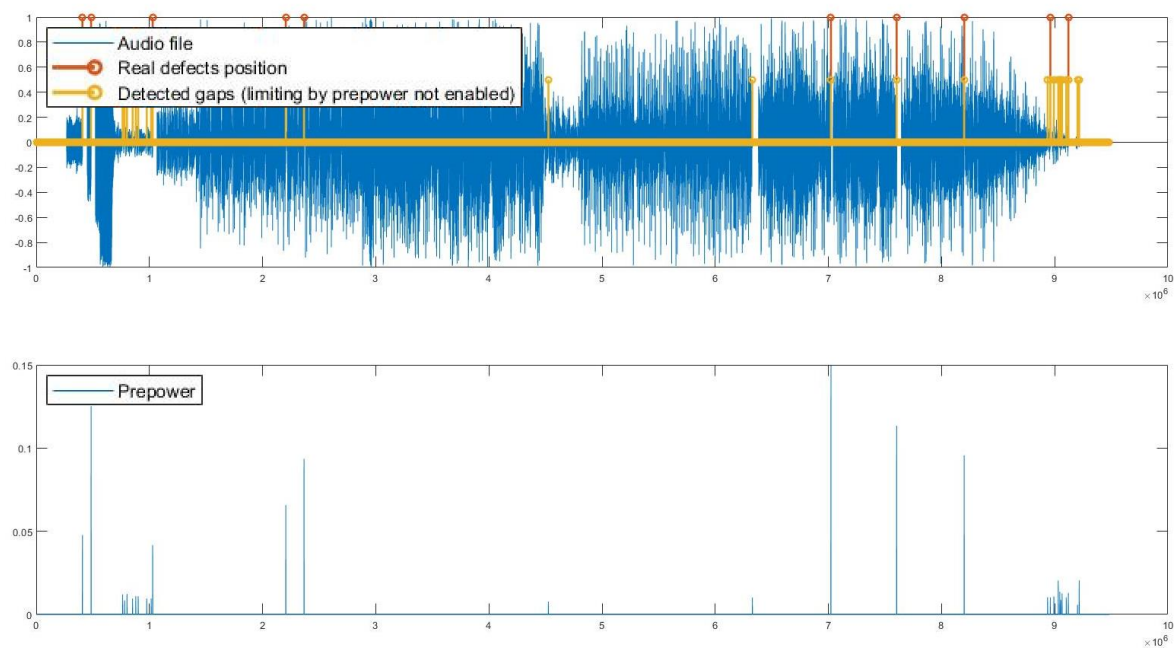
Tento problém můžeme pro účely této práce zjednodušit tak, že jelikož pozice zjištěných vad známe, stačí nám najít existující vadu poblíž detekované mezery.

Možnost, kterou jsem se ale rozhodl implementovat a která zajistí, že by metodu šlo použít i k detekci samostatně (existující algoritmy pro detekci jsou obvykle optimalizované pro detekci impulsních vad a nemusely by nalézt tento typ vady), je metoda, kdy po detekci vady ještě upřesním místo obsahující vadu tak, že místo detekované vady posunu doleva na první vzorek, který překročí daný práh. Jako práh jsem použil stejný práh jako pro prahování obálky. Na obrázku 21 můžeme poté také vidět, že hodnota

prepower už daleko přesněji rozlišuje skutečné vady a falešně pozitivní detekce, pro skutečné vady má ve většině případů vyšší hodnotu. Můžeme vidět, že v některých případech (u nahrávky na obrázku především na konci nahrávky, kdy už se hlasitost nahrávky pohybuje okolo výrazně vyšší úrovně) ani tak nepomáhá rozlišit vady skutečné a falešně detekované vždy. Zkoušel jsem, zda nepřinesou jiné hodnoty délky segmentu pro prepower lepší výsledek, ale výsledky byly pro jiné hodnoty délky velmi podobné.

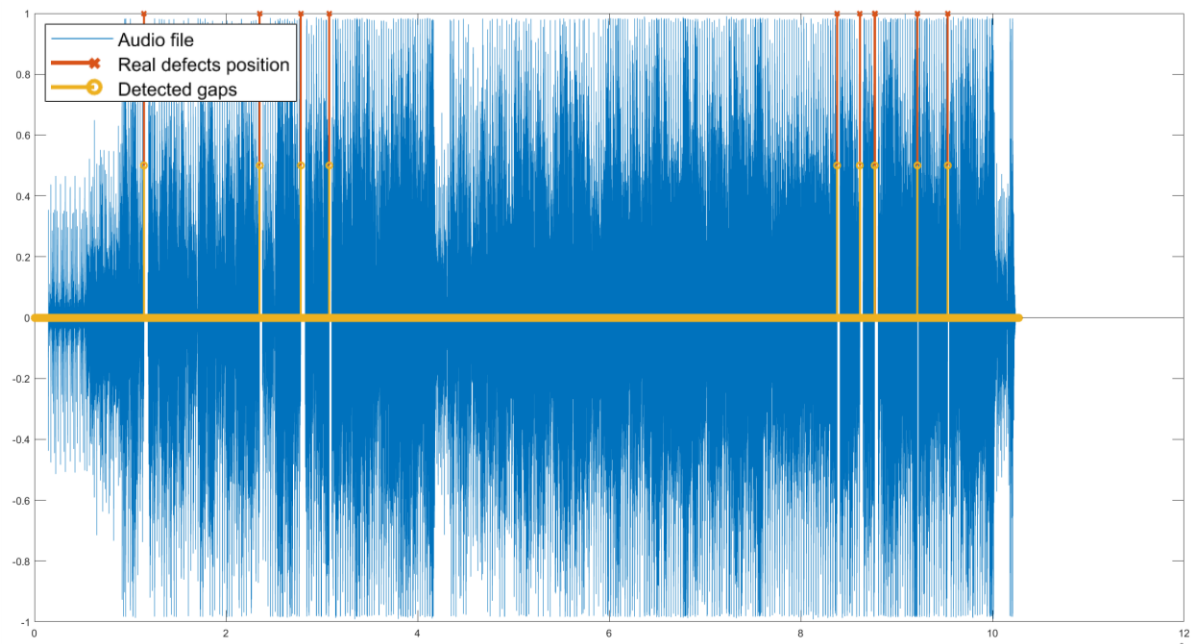


Obrázek 20: Pozice detekované a skutečné vady po modifikaci algoritmu



Obrázek 21: Pozice detekovaných a skutečných vady po modifikaci algoritmu, lze pozorovat již odpovídající hodnoty prepower

Na základě hodnot prepower získaných pro různé nahrávky a přidané mezery jsem stanovil hodnotu prahu na poměrně malou hodnotu, 0,02. Pomocí této metody se povedlo nalézt prakticky všechny přidané mezery, s výjimkou ojedinělých případů jako je mezera přidaná do části nahrávky, kde už je pouze ticho nebo slabý šum na konci nahrávky.



Obrázek 22: Nalezené vady po modifikaci algoritmu a vyřazení falešných detekcí pomocí prepower

Jelikož však stejně pro další testy potřebujeme porovnat pozice detekovaných vad a pozice známých vad, je v této práci algoritmus rozšířen o funkcionalitu, ve které dojde k prohledání známých pozic vad a ke každé dojde k prohledání jejího okolí (které jsem stanovil na 500 vzorků) a nalezení informace, zda se zde vyskytuje nalezená mezera či nikoliv.

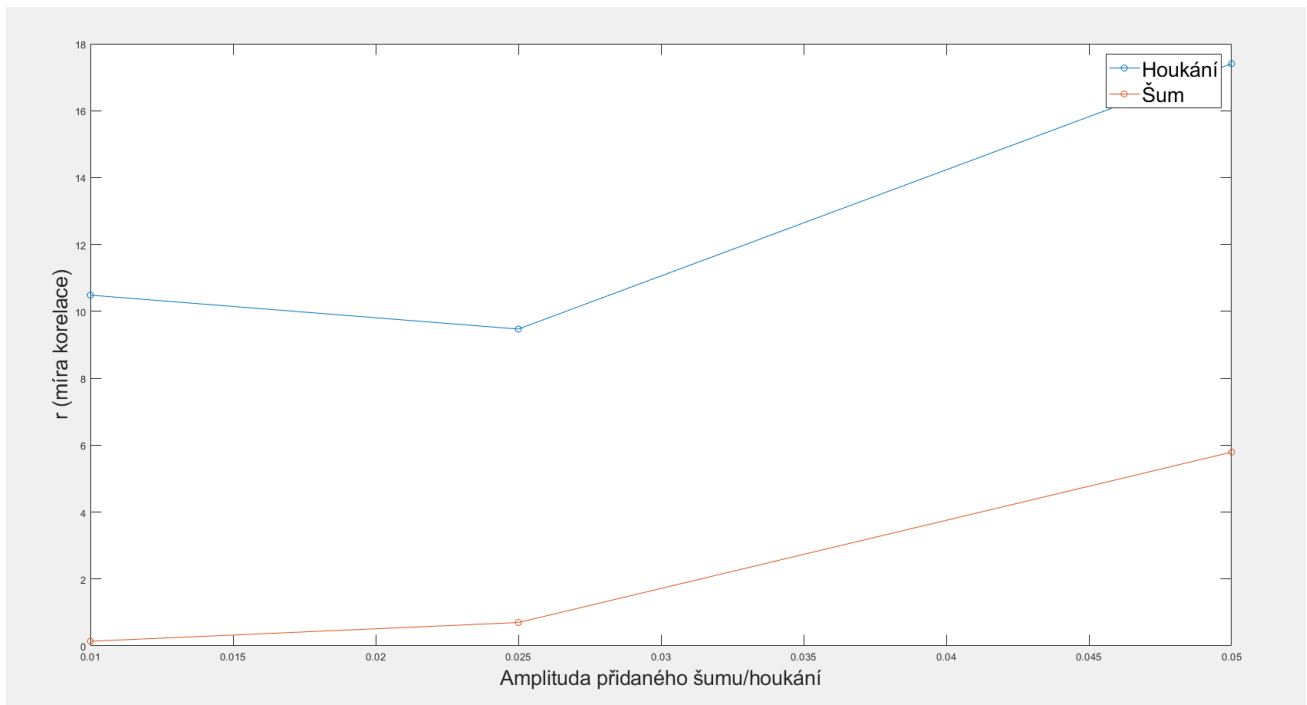
6.2.4. Test algoritmu pro detekci šumu a houkání v nájezdu

Kvůli tomu, že vada šum a houkání v nájezdu je svým projevem velmi odlišná od ostatních, testování této funkcionality probíhalo odděleně od ostatních vad.

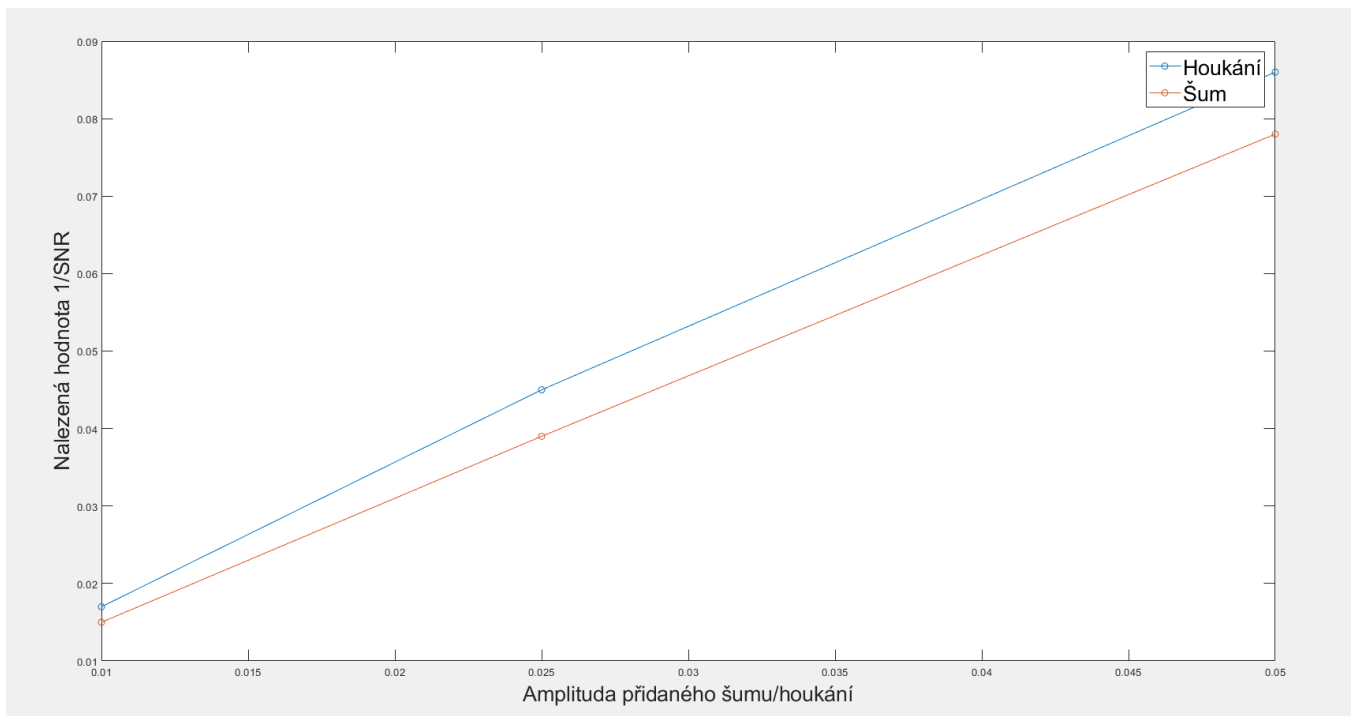
Pro testování jsem využil mastery, do kterých jsem s různou intenzitou přidal šum a houkání. Výsledky testování ukazuje tabulka 2. Obrázky 23 a 24 dále ukazují závislost zjištěné hodnoty $1/\text{SNR}$ a koeficientu korelace na hodnotě přidaného šumu a houkání po zprůměrování hodnot pro všech 5 masterů.

Tabulka 2: Výsledky testování detekce přidaného šumu a houkání

Nahrávka	1/SNR	r (zjištěný koeficient korelace)	Hodnota přidaného šumu	Hodnota přidaného houkání
Master 1	0,018	11,08	0,01	0
Master 2	0,013	12,8	0,01	0
Master 3	0,017	9,78	0,01	0
Master 4	0,02	8,53	0,01	0
Master 5	0,017	10,24	0,01	0
Master 1	0,046	9,87	0,025	0
Master 2	0,033	12,4	0,025	0
Master 3	0,052	8,18	0,025	0
Master 4	0,047	8,27	0,025	0
Master 5	0,045	8,64	0,025	0
Master 1	0,089	24,11	0,05	0
Master 2	0,069	27,89	0,05	0
Master 3	0,09	12,48	0,05	0
Master 4	0,093	10,38	0,05	0
Master 5	0,089	12,18	0,05	0
Master 1	0,016	0,13	0	0,01
Master 2	0,012	0,17	0	0,01
Master 3	0,016	0,13	0	0,01
Master 4	0,016	0,13	0	0,01
Master 5	0,016	0,13	0	0,01
Master 1	0,04	1,14	0	0,025
Master 2	0,028	0,79	0	0,025
Master 3	0,042	0,62	0	0,025
Master 4	0,044	0,5	0	0,025
Master 5	0,041	0,44	0	0,025
Master 1	0,081	7,14	0	0,05
Master 2	0,058	8,28	0	0,05
Master 3	0,08	4,11	0	0,05
Master 4	0,092	4,49	0	0,05
Master 5	0,077	4,94	0	0,05



Obrázek 23: Vývoj koeficientu korelace v závislosti na míře přidaného šumu/houkání (hodnoty zprůměrovány pro všech 5 masterů)



Obrázek 24: Vývoj nalezené hodnoty 1/SNR v závislosti na míře přidaného šumu/houkání (hodnoty zprůměrovány pro všech 5 masterů)

Můžeme tedy pozorovat, že se daří získat odlišné hodnoty koeficientu korelace pro nahrávky obsahující houkání a šum, zároveň můžeme pozorovat prakticky lineární závislost zjištěné hodnoty $1/\text{SNR}$ a hodnoty přidaného šumu.

6.3. Detekce vlastností vad přidaných do reálného masteru

V dalším kroku jsem již vyzkoušel přidat stejným způsobem vady do masterů jednotlivých gramofonových desek. Výsledky ukazuje tabulka 3.

Tabulka 3: Výsledky testování vlastností vad přidaných do masterů gramofonových desek

	Škrták	Škrábanec	Nedolis	Lupanec	Přeskok	Výpadek
Počet přidaných vad	20	20	20	20	20	20
Detekováno více lupanců	92 %	6 %	84 %	11 %	1 %	2 %
Detekován dozvuk (korelace)	74 %	89 %	79 %	73 %	72 %	75 %
Detekován dozvuk (shluk - průměr)	95 %	0 %	92 %	2 %	1 %	2 %
Detekován dozvuk/delší trvání vady (shluk - maximum)	82 %	0 %	79 %	0 %	0 %	0 %
Detekována vada v pravém kanále	13 %	15 %	60 %	15 %	12 %	14 %
Detekováno více lupanců v pravém kanále	4 %	0 %	89 %	1 %	0 %	1 %
Detekováno delší trvání vady v pravém kanále	6 %	0 %	92 %	0 %	0 %	0 %
Detekována vada na otáčku	8 %	99 %	8 %	2 %	11 %	4 %
Detekována změna ve VF	54 %	61 %	55 %	60 %	77 %	84 %
Detekovány mezery	0 %	0 %	0 %	0 %	0 %	82 %

Výsledky se dle očekávání mírně zhoršily, zároveň ale stále detekované vlastnosti ve většině případů odpovídající očekávaným vlastnostem vady. Detekce více lupanců ve škrtáku dosahuje stále velké úspěšnosti, škrábanec je stále detekován jako vada na otáčku, pro nedolisy stále platí, že je detekujeme pomocí velkého rozdílu počtu kliků v pravém a levém kanále. Lupanec vykazuje pouze prudkou změnu ve vysokých frekvencích a detekovaný dozvuk pomocí autokorelace. Ostatní vlastnosti dle očekávání u lupance detekovány nejsou. Bohužel však velmi podobné výsledky jako lupanec vykazuje i přeskok. To je však očekávatelné, jelikož se nepodařilo navrhnout parametr, od kterého bychom očekávali, že bude tyto vady schopný rozlišit.

Také je vidět, že v přítomnosti hudebního signálu na desce již selhává detekce dozvuku pomocí autokorelační funkce. Metoda pomocí shlukování a počítání délky shluků dává ale stále výsledky, které jsou od sebe schopny rozlišit vady, které se projevují krátkým projevem bez dozvuku, a vady, u kterých čekáme delší projev, respektive dozvuk. Mezery se podařilo nalézt ve většině případů. Případy, kdy se mezery najít nepodařilo, jsou dány většinou tím, že byla mezera přidána do tiché části signálu a tudíž buď nebyla detekovatelná, nebo byla pomocí hodnoty prepower detekovaná jako hudebně zamýšlená mezera.

6.4. Klasifikace pomocí metody KNN

V dalším kroku jsem již přistoupil k využití klasifikátoru KNN. Pro přehlednost předchozích testů bylo výhodné pracovat u každého parametru vady s konkrétním prahem a stanovit tak pravdivostní hodnotu, zda vada obsahuje či neobsahuje danou vlastnost. Pro zpracování v klasifikátoru KNN je však výhodnější mít větší rozlišení hodnot parametru a tak jsem po každou vlastnost stanovil škálu, podle které je dále rozpoznávána. Aby byla zajištěna lepší přehlednost, je vždy škála vztažena vůči stanovenému prahu tak, aby při hodnotě prahu byl výsledek 0,5 a tedy můžeme u všech vlastností stanovit, že hodnota větší než 0,5 je považována za detekovanou vlastnost.

Pro implementaci algoritmu byla upravena funkce *KNN.m* z bakalářské práce [13]. Vstupem do funkce jsou trénovací data, tedy matice spočtených vlastností vad z nahrávek z trénovací množiny, spolu s vektorem udávajícím název každé z těchto vad a také data udávající spočtené vlastnosti jednotlivých vad přítomných na desce. Tyto vady jsou postupně porovnávány pomocí algoritmu KNN s vlastnostmi vad z trénovací množiny. Prvky (vady) trénovací množiny jsou poté seřazeny dle vzdálenosti od aktuálně zkoumané vady a je nalezeno k nejbližších sousedů této vady. Vada je poté klasifikována jak ten typ, který se mezi nalezenými sousedy vyskytuje nejčastěji.

Pro klasifikování uměle přidaných vad klasifikátorem KNN jsem vytvořil skript *CLASSIFICATION_artificial.m*. Skript vznikl úpravou skriptu *CLASSIFICATION_main.m* z bakalářské práce [13]. Ve skriptu nejprve uživatel vybere nahrávku ke klasifikaci a dle názvu nahrávky najde příslušný soubor s pozicemi vad. Pro nahrávku jsou dále spočítány vlastnosti jednotlivých dat a ty jsou předány klasifikátoru. Na základě předchozích testů jsem jako vstup do klasifikátoru vybral tyto vlastnosti: výskyt vady na otáčku, výskyt více lupenců, výskyt více lupenců v pravém kanále, průměrná délka shluku, maximální délka shluku a delší maximální shluk v pravém kanále.

6.4.1. Testování na modelech vad, výběr parametru k

Běžným způsobem pro ověření funkčnosti modelu strojového učení je metoda křížové validace. Z této metody jsem vycházel při vyzkoušení různých hodnot parametru k a hledání hodnoty, která bude přinášet nejlepší výsledky. Běžný postup při křížové validaci je postupné přesouvání prvků mezi trénovací a testovací množinou, kdy vždy by mělo platit, že algoritmus je testovaný na jiné množině, než byl natrénován. V našem případě jsme toto však schopni zajistit i se stálou trénovací množinou, jelikož generování vad vždy proběhne náhodně na různá místa nahrávky, a jsme tak schopni i vícekrát použít stejnou nahrávku, ale přidat do ní vady na jiná místa.

Pro ověření úspěšnosti klasifikace algoritmu a také výběr parametru k , který je nutné stanovit v klasifikátoru KNN, jsem použil následující postup. Nejprve jsem typy vad přidal do 5 masterů nahrávek z gramofonové desky. Pro zjednodušení byla vždy použita pouze první skladba na desce. Do těchto

nahrávek bylo vždy přidáno 20 vad daného typu. Na vzniklých nahrávkách byl dále algoritmus natrénován. Získáme tak relativně velkou trénovací množinu, kdy máme od každé vady 100 prvků, pokaždé vygenerovanou náhodně s mírně odlišnými vlastnostmi a zároveň obklopenou jiným okolním zvukem.

Následně jsem provedl vygenerování testovací množiny stejným způsobem, kdy do stejných 5 nahrávek bylo přidáno stejné množství vad, ovšem díky náhodnému generování byly přidány na jiná místa a jejich vlastnosti se mohly opět mírně lišit. Na všech těchto vadách jsem poté otestoval klasifikační algoritmus s různými hodnotami parametru k . Výsledky tohoto testování jsou uvedeny v tabulce 4.

Tabulka 4: Výsledky testování klasifikace vad modelů vad pomocí klasifikátoru KNN

		Simulovaná vada					
k	Detekovaná vada	Lupanec	Přeskok	Mezery	Nedolisy	Škrábanec	Škrták
1	Lupanec	80 %	87 %	42 %	30 %	10 %	20 %
	Přeskok	15 %	9 %	3 %	3 %	12 %	1 %
	Mezery	0 %	0 %	0 %	0 %	0 %	0 %
	Nedolisy	1 %	1 %	33 %	44 %	1 %	21 %
	Škrábanec	2 %	2 %	3 %	5 %	69 %	11 %
	Škrták	2 %	1 %	19 %	18 %	8 %	47 %
3	Lupanec	78 %	81 %	16 %	4 %	5 %	4 %
	Přeskok	14 %	15 %	5 %	1 %	12 %	2 %
	Mezery	0 %	0 %	0 %	0 %	0 %	0 %
	Nedolisy	1 %	1 %	48 %	65 %	0 %	28 %
	Škrábanec	6 %	2 %	5 %	7 %	80 %	13 %
	Škrták	1 %	1 %	26 %	23 %	3 %	53 %
5	Lupanec	80 %	77 %	16 %	4 %	6 %	3 %
	Přeskok	13 %	19 %	4 %	0 %	2 %	1 %
	Mezery	0 %	0 %	0 %	0 %	0 %	0 %
	Nedolisy	1 %	1 %	55 %	67 %	0 %	33 %
	Škrábanec	5 %	3 %	3 %	5 %	89 %	6 %
	Škrták	1 %	0 %	22 %	24 %	3 %	57 %
7	Lupanec	84 %	79 %	15 %	4 %	8 %	4 %
	Přeskok	9 %	17 %	4 %	0 %	2 %	0 %
	Mezery	0 %	0 %	0 %	0 %	0 %	0 %
	Nedolisy	1 %	1 %	59 %	67 %	0 %	41 %
	Škrábanec	5 %	3 %	4 %	4 %	87 %	6 %
	Škrták	1 %	0 %	18 %	25 %	3 %	49 %

Z tabulky je zřejmé, že výsledky odpovídají očekávaným výsledkům dle předchozího experimentu. Dochází k častým záměnám u dvojice vad dvojice lupanec a přeskok. Je také patrné, že se nedaří klasifikovat vadu

typu mezery. Důvodem je to, že klasifikátor bere v potaz u této vady i všechny další vlastnosti, které jsou však pro mezery irelevantní.

Rozhodl jsem se tedy, vzhledem k velké odlišnosti této vady od ostatních, vadu typu mezery vyjmout z klasifikátoru KNN a ještě před vstupem do klasifikátoru rozhodnout, zda je vada považována za detekovanou mezeru pouze na základě získání této samotné vlastnosti bez ohledu na další parametry. Poté jsem znovu provedl otestování klasifikace stejným způsobem. Výsledky ukazuje tabulka 5.

Tabulka 5: Výsledky testování klasifikace vad modelů vad pomocí klasifikátoru KNN, modifikováno vyhodnocení mezer

k	Detekovaná vada	Simulovaná vada					
		Lupanec	Přeskok	Mezery	Nedolisy	Škrábanec	Škrták
1	Lupanec	80 %	88 %	16 %	6 %	6 %	5 %
	Přeskok	13 %	8 %	1 %	0 %	6 %	0 %
	Mezery	0 %	0 %	82 %	0 %	0 %	0 %
	Nedolisy	2 %	1 %	0 %	60 %	1 %	27 %
	Škrábanec	3 %	2 %	0 %	5 %	77 %	12 %
	Škrták	2 %	1 %	1 %	29 %	10 %	56 %
3	Lupanec	85 %	84 %	16 %	4 %	8 %	4 %
	Přeskok	6 %	12 %	1 %	0 %	8 %	0 %
	Mezery	0 %	0 %	82 %	0 %	0 %	0 %
	Nedolisy	1 %	1 %	0 %	61 %	0 %	22 %
	Škrábanec	7 %	2 %	0 %	7 %	81 %	12 %
	Škrták	1 %	1 %	1 %	28 %	3 %	62 %
5	Lupanec	82 %	82 %	16 %	4 %	4 %	4 %
	Přeskok	11 %	14 %	1 %	0 %	3 %	0 %
	Mezery	0 %	0 %	82 %	0 %	0 %	0 %
	Nedolisy	1 %	1 %	0 %	65 %	0 %	33 %
	Škrábanec	5 %	3 %	0 %	4 %	90 %	6 %
	Škrták	1 %	0 %	1 %	27 %	3 %	57 %
7	Lupanec	87 %	82 %	16 %	4 %	4 %	4 %
	Přeskok	6 %	14 %	1 %	0 %	1 %	0 %
	Mezery	0 %	0 %	82 %	0 %	0 %	0 %
	Nedolisy	1 %	1 %	0 %	63 %	0 %	32 %
	Škrábanec	5 %	3 %	0 %	4 %	92 %	6 %
	Škrták	1 %	0 %	1 %	29 %	3 %	58 %

Z výsledků je patrné, že došlo k velkému zlepšení u detekce mezer a tím pádem i přesnější detekci ostatních vad, i když především dvojice vad lupanec a přeskok je stále zaměňována.

Z tabulky můžeme pozorovat, že různé hodnoty parametru k jsou výhodné pro různé zkoumané vlastnosti vad. Jelikož se ale při větším množství trénovacích vzorků dá očekávat větší spolehlivost při větším množství sousedů, rozhodl jsem se pro další testování použít hodnotu $k = 7$.

6.4.2. Testování reálných vad

Po otestování algoritmu na uměle vytvořených vzorcích vad jsem přistoupil ke kroku ověření algoritmu na vzorcích reálných vad, které jsem měl k dispozici. Postup, který by byl z hlediska využití algoritmu na reálných vzorcích ideální, by byl algoritmus natrénovat na velkém množství reálných vad a poté algoritmus využít pro klasifikaci dalších vzorků. Pro tento postup ale nemáme k dispozici dostatečné množství vzorků. Rozhodl jsem se tedy vyzkoušet testování algoritmu na datech natrénovaných na simulovaných vadách.

Pro testování a optimalizaci algoritmu mi byly firmou GZ Media poskytnuty vzorky nahrávek a zároveň data o tom, kde se na nich nachází vada a v některých případech i jakého je typu.

Celkem jsem měl takto k dispozici data o 7 nahrávkách. Počet vad jednotlivých typů ukazuje následující tabulka. Dle vyjádření GZ Media však nemusí být ve všech případech tato klasifikace přesná, jelikož jejich systém není na klasifikaci dat zaměřen a při analýze poslechem je pro operátory kontroly důležité vadu zachytit, ale vzhledem k velkému množství desek, které mají ke kontrole, by pro ně zabývání se správným pojmenováním vad znamenalo velké zdržení.

Tabulka 6: Typy vad a jejich četnost v testovacích datech

Typ vady	Četnost
Neroznaná vada	122
Škrták lisovna	11
Lupanec	22
Nedolisy	1

Dostupná data pro ověření přítomnosti šumu a houkání v nájezdu bohužel dostupná nemáme, tedy tuto funkcionalitu nebylo možné ověřovat na reálných vzorcích.

Při testování reálných vad je vytvořen skript *CLASSIFICATION_realDefects.m*, který funguje podobně jako skript pro klasifikaci uměle vytvořených vad, ale je v něm řešeno jiným způsobem načtení pozic vad, kdy pro uměle vytvořené vady jsme je měli v souboru používaném pro uložení hodnot proměných z MATLABu (soubor s příponou *.mat*), ale v tomto případě je čteme z *.xls* souboru, ve kterém máme kompletní data o vadách na jednotlivých deskách.

Zároveň se musíme vypořádat s tím, že začátek hraní desky se nikdy nerovná přesně začátku nahrávky a jsou tedy i posunuté i pozice vad. Tento offset jsem stanovil ručně pro každou nahrávku, vždy se pohyboval okolo 10 sekund.

Výsledky spuštění algoritmu klasifikace pro reálné vady ukazuje tabulka 7.

Tabulka 7: Výsledek otestování algoritmu na reálných vadách

		Simulovaná vada					
	Počet vad	Lupanec	Přeskok	Přeskok	Škrábanec	Nedolisy	Škrták
Lupanec	22	31,8 %	9,1 %	0,0 %	40,9 %	9,1 %	9,1 %
Nedolis	1	100,0 %	0,0 %	0,0 %	0,0 %	0,0 %	0,0 %
Nerozpoznaná vada	147	48,30 %	2,04 %	0,00 %	31,97 %	2,04 %	15,65 %
Škrták	12	25,00 %	8,33 %	0,00 %	16,67 %	25,00 %	25,00 %

Z tabulky můžeme vidět, že v testu se nepodařilo dosáhnout příliš dobrých výsledků.

U detekce lupance vidíme, že bylo mnoho lupanců zařazeno jako škrábanec. To bylo patrně způsobeno tím, že se v okolí vyskytovalo mnoho dalších vad a bylo tak možné najít vadu se vzdáleností blížíící se vzdálenosti odpovídající vadě na otáčku.

Vadu typu nedolis máme k dispozici pouze jednu, nedá se zde tedy výsledek považovat za vypovídající.

U detekce škrtáku můžeme pozorovat častou záměnu s nedolisy, což se do jisté míry stávalo i v předchozím testu na uměle přidaných vadách. Záměnou s vadou typu lupanec byla patrně dána charakterem konkrétního škrtáku, kdy se jednalo o méně výrazný škrták, kde praskání netrvalo příliš dlouhý interval.

Závěr

V této práci jsem se zabýval vlastnostmi jednotlivých typů gramofonových desek a možnostmi jejich automatické klasifikace. V teoretické části jsou zmíněny vybrané typy vad a také metody používané pro detekci a klasifikaci audio signálu, které řeší buď podobnou úlohu anebo které se dají pro naší úlohy využít.

V práci jsem se rovněž zabýval vytvořením modelů jednotlivých typů vad, které mohou dále pomoci s trénováním a testováním navržených klasifikačních metod. Typy vad, které byly modelovány jsou vada typu škrťák, lupanec, přeskok, nedolisy, škrábanec a výpadek. Modelování a simulace těchto vad spočívá v jejich přidání na náhodná místa vybrané nahrávky.

Návrh klasifikačního algoritmu vychází z detekování typických vlastností jednotlivých vad. Jsou navrženy metody pro detekci výskytu vady na otáčku, pro detekce délky a dozvuku vady, pro detekci série více lupanců za sebou, vady v pravém kanále, houkání a šumu v nájezdu, mezer v signálu a detekci náhlé změny ve vysokých frekvencích. Vzhledem k tomu, že již existují algoritmy pro detekci vad a jsou dále rozvíjeny, například v paralelně publikované diplomové práci pana Semanského (autora práce [8]), nezabývá se tato práce touto problematikou, ale pouze tím, jak již detekované vady dále klasifikovat.

Detekované vlastnosti jsou dále vstupem klasifikátoru KNN, který po natrénování určí, o jakou vadu se na základě hodnot detekovaných vlastností jedná. Po natrénování se klasifikátoru podařilo s dobrou úspěšností rozlišit uměle přidané vady s výjimkou dvojice lupanec a přeskok. To je způsobeno tím, že v obou případech jde o krátké lupnutí a nepodařilo se najít parametr, který by byl schopen rozlišit návaznost úseku před a po tomto lupnutí.

Po ověření funkcionality při úspěšné klasifikaci uměle přidaných vad byla úspěšnost testována na dostupné množině reálných vad. V tomto testu se již nepodařilo dosáhnout příliš dobrých výsledků. Nicméně kvůli velikosti množiny reálných vzorků, které jsou k dispozici, se tento test nedá považovat za příliš vypovídající. Kromě nízkého počtu vzorků je nutné také uvažovat to, že zařazení vzorků pochází především z kontroly v průmyslu, která byla zaměřena primárně na to vady nalézt a kdy jen u některých byly vady i pojmenovány. A vzhledem k tomu, že to pro operátora kontroly pravděpodobně nebylo klíčové, nemusely být vždy ani pojmenovány zcela správně.

Klasifikační algoritmus prokázal, že je schopen rozpoznat typické vlastnosti vad gramofonových desek a dle nich nahrávky klasifikovat do tříd na základě dat, na kterých byl naučen. Pokud bychom měli k dispozici větší množinu trénovacích dat, bylo by již možné algoritmus natrénovat na této množině a tím pádem by přinesl lepší výsledky na reálných vzorcích.

Probíhající výzkum a předchozí návrh klasifikačního algoritmu byl také prezentován na 23. a 24. mezinárodní studentské konferenci Poster v letech 2018 a 2019 [24, 25].

Literatura

- [1] Dušek, K., Záznam a reprodukce zvuku. 5 vydání. Praha: Vydavatelství ČVUT, 1982
- [2] Interní zdroj firmy GZ Media, a.s.
- [3] Výroba, GZ Media, [cit. 18. 5. 2019], [online], dostupné z:
<http://www.gzvinyl.com/Vyroba/>
- [4] „Gettin´ Serious About Vinyl: Wow, Flutter, and S/N Ratio“, 120studio.com, 2015, [cit. 19. 5. 2019], [online], dostupné z: <http://120studio.com/audio/wowflutter.htm>
- [5] Eargle, J., Handbook of Recording Engineering, 1986
- [6] Wikipedia, „RIAA equalization“, [cit. 11. 5. 2019], [online], dostupné z:
https://en.wikipedia.org/wiki/RIAA_equalization
- [7] Bolek, M., Detekce vad u mechanického záznamu zvuku, FEL ČVUT, 2016
- [8] Semanský, M., Detekce impulzních vad v záznamu zvuku na gramofonových deskách, FEL ČVUT, 2017
- [9] Godsill, S, Rayner, P., Digital Audio Restoration. Springer-Verlag, London,1998.
- [10] Podzemský, O., Vizuální kontrola gramofonových desek, FEL ZČU, 2016
- [11] Hensman, A., Casey, K., Optical Reading and Playing of Sound Signals from Vinyl Records, Dublin Institute of Technology, Dublin 2007
- [12] Mühlbauer, R., Automatic Audio Defect Detection, Technisch-Naturwissenschaftliche Fakultät, Johannes Kepler Universität Linz, 2010
- [13] Koshkina, E., Identifikace obsahu archivních zvukových záznamů, FEL ČVUT, 2015
- [14] Moldan, V., Rozpoznání titulu gramofonové desky podle krátké ukázky, FEL ČVUT, 2017
- [15] Melechovský, J., Klasifikace hudebních žánrů pomocí MFCC, FEL ČVUT, 2017
- [16] Kandaswamy, U., Audio Based Speech Recognition Using KNN Classification Method, Lawrence Technological University, Southfield, Michigan, 2016
- [17] Wikipedia, „K-nearest neighbors algorithm“, [cit. 19. 5. 2019], [online], dostupné z:
https://en.wikipedia.org/wiki/K-nearest_neighbors_algorithm
- [18] Gershenson C., „Artificial Neural Networks for Beginners“, [cit. 19. 5. 2019], [online], dostupné z:
<https://arxiv.org/abs/cs/0308031>
- [19] Rusz., J, Analýza experimentálních dat, Přednáška 10, [cit. 10. 5. 2019], [online], dostupné z:
http://sami.fel.cvut.cz/aed/EDA_Lecture10.pdf
- [20] Wikipedia, „Support-vector machine“, [cit. 18. 5. 2019], [online], dostupné z:
https://en.wikipedia.org/wiki/Support-vector_machine

[21] Čmejla., R, Syntéza audio signálů, Přednáška 5, [cit. 22. 5. 2019], [online], dostupné z:
<http://sami.fel.cvut.cz/syn/SYN05.pdf>

[22] EASERA Tutorial, AFMG Technologies GmbH, Berlin 2012, [cit. 20. 5. 2019], [online], dostupné z:
http://www.afmg-support.de/SoftwareDownloadBase/AFMG/EASERA12/EASERATutorial_EN.pdf

[23] Auto-Correlation and Echo Cancellation Overview, UCLA Samueli School Of Engineering,
[cit. 20. 5. 2019], [online], dostupné z: <http://www.seas.ucla.edu/dsplab/aec/over.html>

[24] MOLDAN, V., Identification and classification of defects in vinyl disc records, Poster Conference, FEL
ČVUT 2018

[24] MOLDAN, V., Experimental System for Classification of Defects in Vinyl Disc Records based on Typical
Defects Attributes: First Proposal, Poster Conference, FEL ČVUT 2019

Přílohy

Na přiloženém CD se nacházejí následující adresáře:

- MATLAB
- xls_results
- modely_vad

Ve složce *xls_results* se nachází soubory formátu *.xls*, ve kterých jsou uloženy výsledky testování popsaných v této práci.

Ve složce *modely_vad* se nachází nahrávky vzniklé přidáním jednotlivých typů vad do nahrávky obsahující ticho. Tam kde je to relevantní, je přidán i *.mat* soubor s pozicemi vad.

Ve složce MATLAB se nachází všechny funkce a skripty využitě ve vytvořeném programu. Zároveň je zde k dispozici stručný návod k práci s programem a soubor *train.mat* obsahující data získaná z trénovací množiny.

- *addDefectTypes.m*: Skript pro přidání vad do masterů z uživatelem vybrané složky
- *addHummNoise.m*: Funkce pro přidání houkání a šumu do masterů z uživatelem vybrané složky
- *AR_segment.m*: Funkce pro detekci chybových vzorků pomocí AR modelu, převzato z [7]
- *CLASSIFICATION_artificial_folder.m*: Skript pro klasifikaci uměle vytvořených vad, pozice vad jsou načteny z formátu *.mat*, využita a pro použití v této práci upravena část kódu ze skriptu *CLASSIFICATION_main.m* z práce [13]
- *CLASSIFICATION_realDefects.m*: Skript pro klasifikaci reálných vad, pozice vad jsou načteny z formátu *.xls*, využita a pro použití v této práci upravena část kódu ze skriptu *CLASSIFICATION_main.m* z práce [13]
- *classifyDefects.m*: funkce, která pro danou nahrávku detekuje jednotlivé vlastnosti vad a vrátí pravdivostní hodnotu jednotlivých vlastností
- *classifyDefectsP.m*: funkce, která pro danou nahrávku detekuje jednotlivé vlastnosti vad a hodnotu míry zastoupení jednotlivých vlastností
- *corrEcho.m*: funkce pro detekování dozvuku pomocí autokorelační funkce
- *cutClicks.m*: funkce pro separování okolí vady
- *defectInDefect.m*: funkce pro detekci více vzorků zasažených vadou
- *defectInDefectR.m*: funkce pro detekci více vzorků zasažených vadou a porovnání, zda je jich více v pravém kanále
- *echoMaxClusters.m*: funkce pro detekci dozvuku pomocí metody shlukování s nalezením délky nejdelšího shluku
- *echoMaxClustersR.m*: funkce pro detekci dozvuku pomocí metody shlukování s nalezením délky nejdelšího shluku a porovnání, zda je větší hodnota v pravém kanále
- *echoMeanClusters.m*: funkce pro detekci dozvuku pomocí metody shlukování s nalezením průměrné délky shluku

- *find_clicks.m*: funkce pro nalezení chybných vzorků pomocí vlnkové transformace, převzato z práce [8]
- *findAttributes.m*: skript, který byl použit pro dílčí testování a ověření nalezení atributů jednotlivých vad
- *findGaps.m*: funkce pro nalezení mezer v signálu
- *inRightCh.m*: funkce pro nalezení vady v pravém kanále pomocí porovnání průměrné amplitudy v okolí vady v obou kanálech
- *KNN.m*: klasifikace pomocí metody strojového učení KNN, převzato z bakalářské práce [13] a upraveno pro využití v této úloze
- *noiseHumm.m*: funkce pro nalezení výskytu šumu a houkání v tichém úseku nahrávky
- *nr2Ltr.m*: převedení čísla sloupce na jeho označení v excelu, pomocná funkce pro funkci zápis výsledků testování do excelu
- *onTurn.m*: funkce pro detekci vady na otáčku
- *runClassificationArtificial.m*: pomocná funkce pro skript *findAttributes.m*, dojde k načtení nahrávky a pozic vad pro tuto nahrávku a následné zjištění jejich vlastností
- *spectralChange.m*: funkce pro nalezení náhlé změny ve vysokých frekvencích v okolí vady
- *testHummNoise.m*: funkce pro otestování detekce šumu a houkání
- *TRAIN.m*: skript pro trénování klasifikátoru KNN, vznikl úpravou skriptu *TRAIN_main.m* z práce [13]