

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

Fakulta elektrotechniky
a komunikačních technologií

DIPLOMOVÁ PRÁCE

Brno, 2019

Bc. Štěpán Miklánek



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH TECHNOLOGIÍ

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION

ÚSTAV TELEKOMUNIKACÍ

DEPARTMENT OF TELECOMMUNICATIONS

URČENÍ MÍSTA PŮVODU HUDEBNÍCH INTERPRETACÍ ČESKÉ KOMORNÍ A ORCHESTRÁLNÍ HUDBY ZA POMOCI TECHNIK MUSIC INFORMATION RETRIEVAL

MUSIC INFORMATION RETRIEVAL TECHNIQUES FOR DETERMINING THE PLACE OF ORIGIN OF THE
CZECH CHAMBER AND ORCHESTRAL MUSIC INTERPRETATIONS

DIPLOMOVÁ PRÁCE

MASTER'S THESIS

AUTOR PRÁCE

AUTHOR

Bc. Štěpán Miklánek

VEDOUCÍ PRÁCE

SUPERVISOR

Ing. Tomáš Kiska

BRNO 2019

Diplomová práce

magisterský navazující studijní obor **Audio inženýrství**

Ústav telekomunikací

Student: Bc. Štěpán Miklánek

ID: 174465

Ročník: 2

Akademický rok: 2018/19

NÁZEV TÉMATU:

Určení místa původu hudebních interpretací české komorní a orchestrální hudby za pomoci technik Music Information Retrieval

POKYNY PRO VYPRACOVÁNÍ:

V rámci této práce budou shrnuty dosavadní poznatky z oblasti zvané Music Information Retrieval (MIR). Cílem práce bude vytvořit databázi nahrávek hudebních interpretací české komorní a orchestrální hudby s co možno nejpočetnějším zastoupením. Dále bude navržen systém, který bude využívat metody MIR. Dále pak budou analyzovány jednotlivé hudební interpretace z hlediska barvy zvuku, rytmiky a dynamiky. Rovněž budou vybrány parametry, pomocí kterých bude možné rozpoznat místo původu u interpretací české komorní a orchestrální hudby. K rozpoznávání místa původu budou využity klasifikační algoritmy, pomocí kterých bude vyhodnocena diskriminační síla jednotlivých parametrů.

DOPORUČENÁ LITERATURA:

[1] SCHEDL, M.; GOMÉZ, E; URBANO, J. Music Information Retrieval: Recent Developments and Applications. Foundations and Trends in Information Retrieval, 2014: s. 127-261, DOI: 978-1-60198-807-2.

[2] MÜLLER, M. Fundamentals of Music Processing: Audio, Analysis, Algorithms, Applications [online]. Springer International Publishing Switzerland, 2015, 483 s. ISBN 978-3-319-21945-5.

Termín zadání: 1.2.2019

Termín odevzdání: 16.5.2019

Vedoucí práce: Ing. Tomáš Kiska

Konzultant:

prof. Ing. Jiří Mišurec, CSc.
předseda oborové rady

UPOZORNĚNÍ:

Autor diplomové práce nesmí při vytváření diplomové práce porušit autorská práva třetích osob, zejména nesmí zasahovat nedovoleným způsobem do cizích autorských práv osobnostních a musí si být plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č.40/2009 Sb.

ABSTRAKT

Tato práce se zabývá statistickou analýzou nahrávek české komorní a orchestrální hudby. Významná část této práce je věnována problematice extrakce parametrů z nahrávek. V rámci extrakce parametrů jsou využity techniky zpracování hudebního signálu používané v oblasti Music Information Retrieval. Dále jsou v práci představeny kroky při zpracovávání databází nahrávek a výběru stejných notových úseků z jednotlivých interpretací skladeb pomocí dynamického borcení časové osy. Pro účely výběru stejného notového úseků z interpretací byl naprogramován nástroj v prostředí MATLAB. V neposlední řadě je představen vyhodnocovací systém, pomocí kterého je prováděna klasifikace nahrávek podle místa původu. Samotná klasifikace je provedena v podobě binárního rozdělení na dvě skupiny Česko a svět. První část vyhodnocení spočívá v individuální analýze jednotlivých statistických parametrů, přičemž je hodnocena jejich diskriminační síla. Druhá část vyhodnocení spočívá v selekci skupin parametrů, které by měly zvýšit přesnost klasifikace oproti individuální analýze parametrů.

KLÍČOVÁ SLOVA

česká, klasická, komorní, orchestrální, hudba, extrakce, rozpoznávání, podobnost, binární, klasifikace

ABSTRACT

This diploma thesis is focused on the statistical analysis of chamber and orchestral classical music recordings composed by Czech authors. One of the chapters is dedicated to the description of a feature extraction process that precedes the statistical analysis. Techniques of Music Information Retrieval are used during several stages of this thesis. Databases used for analysis are described and pre-processing steps are proposed. A tool for synchronization of the recordings was implemented in MATLAB. Finally the system used for classification of recordings based on their geographical origin is proposed. The recordings are sorted by a binary classifier into two categories of Czech and world recordings. The first part of the statistical analysis is focused on individual analysis of features. The features are evaluated based on their discrimination strength. The second part of the statistical analysis is focused on feature selection, which can improve the overall accuracy of the binary classifier compared to the individual analysis of the features.

KEYWORDS

Czech, classical, chamber, orchestral, music, extraction, recognition, similarity, binary, classification

MIKLÁNEK, Štěpán. *Určení místa původu hudebních interpretací české komorní a orchestrální hudby za pomoci technik Music Information Retrieval*. Brno, Rok, 72 s. Diplomová práce. Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav telekomunikací. Vedoucí práce: Ing. Tomáš Kiska

PROHLÁŠENÍ

Prohlašuji, že svou diplomovou práci na téma „Určení místa původu hudebních interpretací české komorní a orchestrální hudby za pomoci technik Music Information Retrieval“ jsem vypracoval samostatně pod vedením vedoucího diplomové práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou všechny citovány v práci a uvedeny v seznamu literatury na konci práce.

Jako autor uvedené diplomové práce dále prohlašuji, že v souvislosti s vytvořením této diplomové práce jsem neporušil autorská práva třetích osob, zejména jsem nezasáhl nedovoleným způsobem do cizích autorských práv osobnostních a/nebo majetkových a jsem si plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon), ve znění pozdějších předpisů, včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č. 40/2009 Sb.

Brno

.....

podpis autora

PODĚKOVÁNÍ

Rád bych poděkoval vedoucímu semestrálního projektu panu Ing. Tomášovi Kiskovi za odborné vedení, konzultace, trpělivost a podnětné návrhy k práci.

Brno

.....

podpis autora



Faculty of Electrical Engineering
and Communication
Brno University of Technology
Purkynova 118, CZ-61200 Brno
Czech Republic
<http://www.six.feec.vutbr.cz>

PODĚKOVÁNÍ

Výzkum popsany v této diplomové práci byl realizován v laboratořích podpořených z projektu SIX; registrační číslo CZ.1.05/2.1.00/03.0072, operační program Výzkum a vývoj pro inovace.

Brno

.....

podpis autora



EVROPSKÁ UNIE
EVROPSKÝ FOND PRO REGIONÁLNÍ ROZVOJ
INVESTICE DO VAŠÍ BUDOUCNOSTI



Obsah

Úvod	13
1 Music Information Retrieval	14
2 Základní vlastnosti hudebního signálu	15
2.1 Vlastnosti hudebního signálu z hlediska času	15
2.2 Vlastnosti hudebního signálu z hlediska frekvence	16
2.3 Barva zvuku	17
2.4 Dynamika	17
2.5 Tempo	18
3 Zpracování hudebního signálu	19
3.1 Předzpracování hudebního signálu	19
3.1.1 Formát hudebního signálu	19
3.1.2 Vzorkovací kmitočet hudebního signálu	19
3.1.3 Bitová hloubka hudebního signálu	20
3.1.4 Konverze do jednokanálové reprezentace	20
3.2 Parametrizace hudebního signálu	20
3.2.1 MIR Parametry	20
3.2.2 Statistické parametry	21
4 Parametry obsažené v MIR Toolboxu	22
4.1 Parametry popisující dynamiku hudebního signálu	22
4.1.1 Efektivní hodnota signálu	22
4.1.2 Počet průchodů nulovou úrovní	22
4.1.3 Nízká energie	22
4.2 Parametry popisující rytmickou strukturu hudebního signálu	23
4.2.1 Fluktuace	23
4.2.2 Počátek zvukové události	23
4.2.3 Hustota zvukových událostí	23
4.2.4 Tempo	23
4.2.5 Metrický centroid	24
4.2.6 Metrická síla	24
4.2.7 Čitelnost pulzu	24
4.3 Parametry popisující barvu zvuku	24
4.3.1 Doba náběhu	24
4.3.2 Sklon náběhu	24
4.3.3 Skok náběhu	25

4.3.4	Sklon poklesu	25
4.3.5	Pokles spektrální energie	25
4.3.6	Drsnost zvuku	25
4.3.7	Melovské keprální koeficienty	25
4.3.8	Doba trvání	26
4.3.9	Parametry popisující tvar spektrálního rozložení	26
5	Vybrané statistické parametry	27
5.1	Střední hodnota	27
5.2	Rozptyl	27
5.3	Směrodatná odchylka	28
5.4	Medián	28
5.5	Percentil	28
5.6	Mezipercentilové rozpětí	28
5.7	Kvartil	28
5.8	Mezikvartilové rozpětí	28
6	Synchronizace nahrávek	29
6.1	Chromagram	29
6.2	Dynamické borcení časové osy	30
6.3	Nástroj pro synchronizaci nahrávek	32
6.4	Výběr stejného notového úseku	33
7	Vyhodnocovací systém	35
8	Evaluace systému	36
8.1	Klasifikační stromy	36
8.2	Křížová validace	36
8.3	Vyhodnocení algoritmů	36
8.3.1	Matice záměn	37
8.3.2	Přesnost	37
8.3.3	Senzitivita a specificita	38
8.3.4	Precision	38
8.3.5	F1 Score	38
8.3.6	Matthewsův korelační koeficient	38
9	Databáze nahrávek	40
10	Individuální analýza parametrů	42
10.1	1. Janáčkův smyčcový kvartet	42
10.2	2. Janáčkův smyčcový kvartet	44

10.3 Dvořáková 9. symfonie, rozdělení dle dirigenta	45
10.4 Dvořáková 9. symfonie, rozdělení dle orchestru	46
10.5 Zhodnocení individuální analýzy parametrů	48
11 Selekce parametrů pomocí SFFS	49
11.1 1. Janáčkův smyčcový kvartet	49
11.2 2. Janáčkův smyčcový kvartet	51
11.3 Dvořáková 9. symfonie, rozdělení dle dirigenta	52
11.4 Dvořáková 9. symfonie, rozdělení dle orchestru	53
11.5 Zhodnocení selekce parametrů pomocí SFFS	54
12 Závěr	55
Literatura	58
Seznam symbolů, veličin a zkratk	61
Seznam příloh	62
A Přiložené soubory	63
B Seznam vypočítaných parametrů	64

Seznam obrázků

2.1	Zjednodušená časová obálka.	15
2.2	Dělení hudebních signálů.	16
3.1	Vizualizace procesu zpracování hudebního signálu	19
6.1	Proces výběru stejného notového úseku z interpretací	29
6.2	Vizualizace chromagramu	30
6.3	Vizualizace synchronizace dvou chromagramů	31
6.4	Vizualizace omezené oblasti δ	32
7.1	Schéma vyhodnocovacího systému	35
8.1	Grafické znázornění matice záměn	37

Seznam tabulek

2.1	Přehled výrazů pro popis dynamiky a přednesu	18
5.1	Vybrané statistické parametry	27
6.1	Tónové třídy	29
6.2	Porovnání DTW algoritmů	33
9.1	Informace o jednotlivých databázích nahrávek	40
10.1	1. Janáčkův smyčcový kvartet, individuální analýza parametrů, notový úsek A	42
10.2	1. Janáčkův smyčcový kvartet, individuální analýza parametrů, notový úsek B	43
10.3	1. Janáčkův smyčcový kvartet, individuální analýza parametrů, celé nahrávky	43
10.4	2. Janáčkův smyčcový kvartet, individuální analýza parametrů, notový úsek A	44
10.5	2. Janáčkův smyčcový kvartet, individuální analýza parametrů, notový úsek B	44
10.6	2. Janáčkův smyčcový kvartet, individuální analýza parametrů, celé nahrávky	45
10.7	Dvořákova 9. symfonie, individuální analýza parametrů, notový úsek A, klasifikace podle dirigentů	45
10.8	Dvořákova 9. symfonie, individuální analýza parametrů, notový úsek B, klasifikace podle dirigentů	46
10.9	Dvořákova 9. symfonie, individuální analýza parametrů, celé nahrávky, klasifikace podle dirigentů	46
10.10	Dvořákova 9. symfonie, individuální analýza parametrů, notový úsek A, klasifikace podle orchestrů	47
10.11	Dvořákova 9. symfonie, individuální analýza parametrů, notový úsek B, klasifikace podle orchestrů	47
10.12	Dvořákova 9. symfonie, individuální analýza parametrů, celé nahrávky, klasifikace podle orchestrů	48
11.1	1. Janáčkův smyčcový kvartet, selekce SFFS, notový úsek A	49
11.2	1. Janáčkův smyčcový kvartet, selekce SFFS, notový úsek B	50
11.3	1. Janáčkův smyčcový kvartet, selekce SFFS, celé nahrávky	50
11.4	2. Janáčkův smyčcový kvartet, selekce SFFS, notový úsek A	51
11.5	2. Janáčkův smyčcový kvartet, selekce SFFS, notový úsek B	51
11.6	2. Janáčkův smyčcový kvartet, selekce SFFS, celé nahrávky	51
11.7	9. Dvořákova symfonie, selekce SFFS, klasifikace podle dirigentů, notový úsek A	52

11.8	9. Dvořákova symfonie, selekce SFFS, klasifikace podle dirigentů, notový úsek B	52
11.9	9. Dvořákova symfonie, selekce SFFS, klasifikace podle dirigentů, celé nahrávky	53
11.109.	Dvořákova symfonie, selekce SFFS, klasifikace podle orchestrů, notový úsek A	53
11.119.	Dvořákova symfonie, selekce SFFS, klasifikace podle orchestrů, notový úsek B	53
11.129.	Dvořákova symfonie, selekce SFFS, klasifikace podle orchestrů, celé nahrávky	54

Úvod

Hlavním tématem této práce je statistická analýza databází digitalizovaných nahrávek z oblasti české komorní a orchestrální klasické hudby. V rámci této práce bude představen a v jednotlivých krocích podrobně popsán proces, který je nutné během analýzy provést.

V této práci budou využity techniky z rychle rozvíjející se oblasti Music Information Retrieval. Název této oblasti se dá volně přeložit jako získávání informací z hudby. Cílem je snaha získat informace, které nejsou patrné při prostém poslechu nahrávek. Informace jsou z nahrávek získávány v podobě strojové extrakce parametrů, přičemž tyto parametry mají za úkol objektivně popisovat hudební signál na základě jeho vlastností vztahujících se k hudebnímu kontextu. Existuje více způsobů, které je možno využít k naložení s takto získanými informacemi.

Parametrizace hudebního signálu umožňuje například vytváření systémů, pomocí kterých je prováděna automatizovaná klasifikace a zařazování jednotlivých nahrávek do vhodně zvolených kategorií. Nutností je vhodně zvolit tyto kategorie tak, aby výsledkem byla klasifikace o určité vypovídající hodnotě. V praxi se může kupříkladu jednat o rozdělování databáze nahrávek do tříd, které odpovídají daným hudebním žánrům. Klasifikace se také může rovnat procesu rozpoznávání jednotlivých elementů, kterými je tvořen výsledný hudební signál. Takovýto výstup si lze kupříkladu představit jako informaci o tom, jaké hudební nástroje se v dané nahrávce vyskytují.

Jak již bylo zmíněno, analýze budou v případě této práce podrobeny nahrávky české komorní a orchestrální klasické hudby. Konkrétně se jedná o 1. a 2. smyčcový kvartet Leoše Janáčka a 9. Symfonii Antonína Dvořáka. Ke každé ze skladeb bude přidružena databáze co největšího počtu interpretací. Velikost databází bude ovlivněna počtem získaných digitálních nahrávek dané skladby. Smyslem této práce je využití parametrizace hudebních nahrávek a následné využití binární klasifikace. Binární klasifikátor by měl být schopen správně roztřídit databáze nahrávek do kategorií Česko a svět.

Jedním z hlavních výstupů této práce bude provedení individuální analýzy parametrů, které popisují nahrávky na základně dynamiky, rytmiky a barvy zvuku. Bude hodnocena schopnost jednotlivých parametrů správně zařadit nahrávky do výše zmíněných kategorií. Druhým výstupem je provedení strojové selekce skupin parametrů, které by mohly potenciálně zlepšit vlastnosti klasifikátoru v porovnání s předcházející individuální analýzou parametrů.

1 Music Information Retrieval

V úvodu již bylo zmíněno, že je výzkum v oblasti Music Information Retrieval především soustředěn na extrakci vhodně zvolených parametrů z hudby. Tyto parametry jsou extrahovány zpravidla z audio signálů, symbolických reprezentací (notový zápis) nebo z externích zdrojů jako jsou webové stránky. Pomocí extrahovaných parametrů je možné vytvářet systémy, které umožňují vyhledávání a doporučování hudby na základě vlastního obsahu (dle hudebních žánrů a stylů) nebo uživatelská rozhraní, která slouží k procházení obsáhlých hudebních databází.

Zároveň je také nutné zmínit, že Music Information Retrieval – MIR je poměrně nová vědní oblast, která se postupně formovala až ke konci 90. let minulého století. Od této doby zaznamenala tato oblast konstantní růst z hlediska počtu výzkumů a jejich následném využití v praxi. Hlavními důvody úspěchu této oblasti je vývoj technik používaných ke kompresi audio signálů a postupné zvyšování výpočetního výkonu osobních počítačů, které jsou v dnešní době schopny extrahovat parametry v dostatečně krátkém čase. Dále také stojí za zmínku rozšíření dostupnosti mobilních přehrávačů hudby a v neposlední řadě příchod streamovacích služeb jako jsou například Spotify nebo Apple Music.

Zatímco v raných výzkumech se pracovalo především se symbolickými reprezentacemi hudby (digitální reprezentace hudebních partitur v MIDI), se zvýšením výpočetního výkonu bylo umožněno aplikování technik číslicového zpracování signálu přímo na signály hudební. Zpracovávání hudebního signálu umožnilo popis vlastností a kvalit jako jsou rytmus, barva zvuku („témbr“), melodie nebo harmonie.

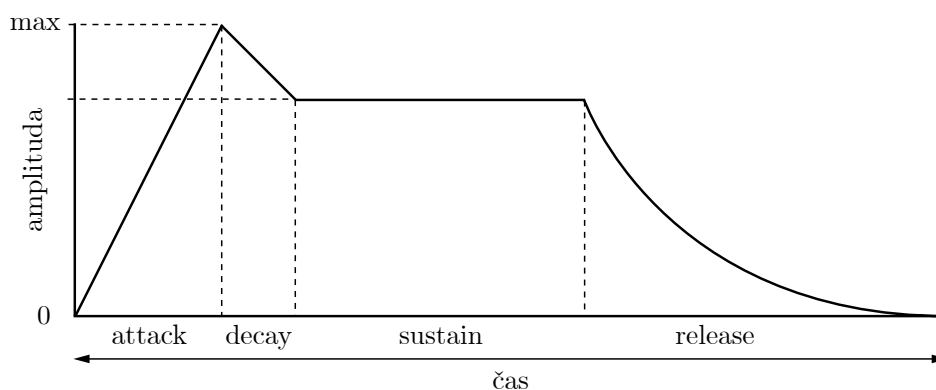
Dodatečně, mnoho vlastností hudby je vztaženo nejenom k hudebnímu obsahu (např. hudební žánry), ale také ke kontextuálním a kulturním aspektům, které mohou být modelovány pomocí informací generovaných uživatelem na základě interakce s danou aplikací nebo webovou stránkou.

2 Základní vlastnosti hudebního signálu

Pro začátek je nutné vymezit základní vlastnosti hudebního signálu. Proto budou v této kapitole vyjmenovány standardizované pojmy, které se pro popis hudebního signálu běžně používají. Hudební signál je také možné reprezentovat a vizualizovat mnoha různými způsoby. Některé z běžných reprezentace budou v této části práce detailněji rozebrány. Součástí této kapitoly budou i běžně používané pojmy z hudební teorie používané pro popis hudebního signálu.

2.1 Vlastnosti hudebního signálu z hlediska času

Časový průběh představuje jednu ze základních abstrakcí hudebního signálu. Časový průběh hudebního signálu je definován jako změna amplitudy (intenzity, akustického tlaku apod.) tónu či zvuku v závislosti na čase. Ze změny tvaru časového průběhu lze usuzovat jeho tónový či hlukový charakter. Pravidelné změny v časovém průběhu indikují spíše tónový charakter, naopak změny nahodilé přísluší šumovému charakteru signálu. Na základě členitosti časového průběhu hudebního signálu je možno posuzovat taktéž jeho harmonickou strukturu. Členité časové průběhy impulzního charakteru naznačují, že signál obsahuje vysoké množství vyšších frekvenčních složek. Naopak signály s jednoduchým „hladkým“ průběhem obsahují vyšších frekvenčních složek méně. Zobrazení časového průběhu je provedeno v podobě tzv. časové (amplitudové) obálky. Zjednodušená časová obálka je znázorněna na obrázku 2.1.



Obr. 2.1: Zjednodušená časová obálka.

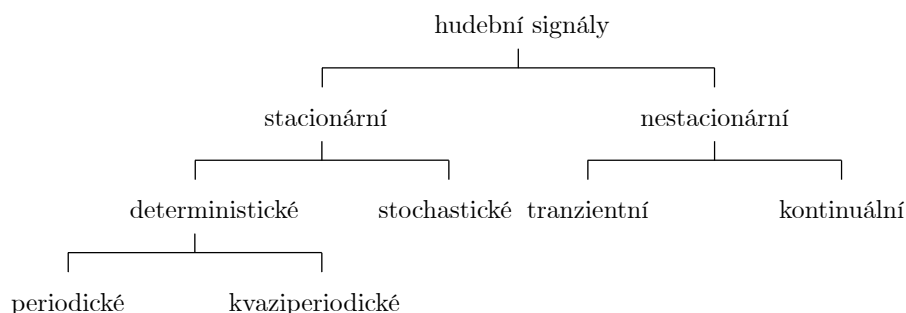
Amplitudovou obálku lze rozdělit na 4 základní dílčí části *A* - *attack*, *D* - *decay*, *S* - *sustain* a *R* - *release*. Označení těchto částí je převzato z terminologie použité ve zvukové syntéze. Části obálky *attack* a *decay* spadají do tzv. nakmitávací fáze signálu, *sustain* reprezentuje zakmitaný stav a fáze *release* je v češtině definována jako

tzv. dokmitávací pochody. Zakmitaný stav signálu je také označován jako kvazistacionární část signálu, nakmitávací a dokmitávací stav je označován též jako tranzient tónu. Kvazistacionární část signálu u většiny tónů mnohonásobně převažuje ostatní fáze signálu z hlediska doby trvání jednotlivých fází, je nositelem informace o výšce, hlasitosti i barvě tónu. V této části mohou být patrné i periodické změny amplitudy signálu neboli *tremolo*. Tyto periodické změny signálu jsou také označovány jako fluktuace ustálené dynamiky a jsou více či méně patrné v závislosti na daném hudebním nástroji [26].

Přirozený hudební signál je signál komplexní, skládá se s z jednotlivých harmonických složek. Výška tónu je určena tzv. fundamentem, první harmonickou složkou vyjádřenou v jednotkách Hertz. Subjektivní vjem výšky tónu je ovšem dále ovlivněn i intenzitou a délkou trvání tónu. Periodické změny okolo fundamentální frekvence jsou označovány jako *vibrato*.

2.2 Vlastnosti hudebního signálu z hlediska frekvence

Vzájemné uspořádání frekvenčních složek hudebního signálu je označováno jako frekvenční struktura. Běžně je tato struktura reprezentována pomocí frekvenčního spektra signálu. Hudební signály můžeme dělit podle několika přístupů k jejich charakteristice viz obrázek 2.2.



Obr. 2.2: Dělení hudebních signálů.

První přístup dělí hudební signály na *stacionární*, tedy takové, které na dostatečně dlouhém časovém intervalu nemění svoje vlastnosti, přičemž opakem jsou signály *nestacionární*. Druhý přístup dělí hudební signály na tzv. *deterministické* a *stochastické*. Deterministické signály je možné v každém bodě v čase definovat pomocí matematického (analytického) předpisu, stochastické (nahodilé) signály nelze popsat matematickou funkcí, ale pouze pomocí pravděpodobnosti a matematické statistiky. Deterministické signály je možno dále rozdělit na signály periodické, ty jsou reprezentovány diskrétním spektrem, které je tvořeno pouze základní frekvencí

a násobky základní frekvence. Kromě periodických signálů jsou definovány také kvaziperiodické signály, kde je možný výskyt frekvenčních složek i v iracionálním poměru. Signály, které vykazují šumový či hlukový charakter a vyznačují se spojitým spektrem jsou označovány jako *stochastické*. Nestacionární signály jsou rozděleny do kategorií *tranzientní* a *kontinuální*. Tranzientní signály se vyznačují impulzním charakterem, jejich spektrum je také spojitě. Kontinuální signály jsou analyticky nejvíce složité a vyznačují se časově proměnnými vlastnostmi, typickým příkladem může být lidská řeč [26].

2.3 Barva zvuku

Barva zvuku, prvotně označována jako „témbr“ (timbre), je soubor vlastností hudebního signálu, který vychází zejména ze subjektivního popisu kvality zvuku. Při slovním popisu barvy zvuku jsou často používána adjektiva převzatá ze smyslových oblastí, jako jsou například zrak nebo hmat. Rozklad periodického i neperiodického signálu je prováděn pomocí Fourierovy řady, přičemž je pomocí tohoto matematického procesu získána informace o frekvenčním obsahu signálu. Frekvenční obsah hudebního signálu má přímý vliv na subjektivní vjem barvy zvuku posluchačem. Barva zvuku je určena zejména počtem a intenzitou jednotlivých harmonických složek. Zdůrazněné frekvenční oblasti spektra hudebního signálu jsou označovány jako *formanty*. Slovní popis barvy zvuku je většinou určen pomocí pojmových párů, například *tupá - ostrá*, *plná - dutá*, *temná - světlá* a další. Mezi důležité teorie barvy zvuku patří Helmholtzova relativní a absolutní teorie barvy zvuku, poznatky o barvě zvuku jsou shrnuty v Helmholtzově práci. Za zmínku také stojí Schumannovy zákony barvy zvuku [21].

2.4 Dynamika

Dynamikou hudby je myšleno označení zvukové intenzity přednesu. Je používán předpis pomocí dynamických znamének. Tři základní hodnoty jsou vyjádřeny zkratkami italského názvosloví: slabě (*p* - piano), středně silně (*mf* - mezzoforte) a silně (*f* - forte). Klasická notace používá celou řadu italských výrazů pro přednes a dynamiku, přehled často používaných výrazů je znázorněn v tabulce 2.1.

V oblasti zpracování hudebního signálu se pro popis používají parametry, které dynamiku hudby popisují v čase. Základním popisem dynamiky je amplitudová obálka hudebního signálu. Další popisy jsou uvedeny v kapitole 4.1 [17].

Tab. 2.1: Přehled výrazů pro popis dynamiky a přednesu

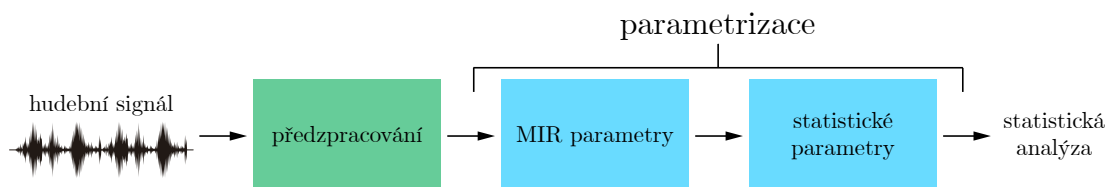
<i>ppp</i>	piano pianissimo	co nejslaběji
<i>pp</i>	pianissimo	velmi slabě
<i>p</i>	piano	slabě
<i>mf</i>	mezzoforte	středně silně
<i>f</i>	forte	silně
<i>ff</i>	fortissimo	velmi silně
<i>fff</i>	forte fortissimo	ci nejsilněji
<i>sf/sfz</i>	sforzando/sforzato	se silným důrazem
<i>cresc.</i>($<$)	crescendo	zesilovat
<i>decresc.</i>($>$)	decrescendo	zeslabovat
<i>dim.</i>($>$)	diminuendo	zeslabovat

2.5 Tempo

Označení tempa udává, jak rychle se má daná skladba hrát. Začátek skladby bývá často označen italským přednesovým označením (např. *allegro*, *lento*), nebo anglickým (např. *fast*, *medium*). Přesnější vyjádření tempa je číselné *metronomické* označení, kde číslo vyjadřuje počet úderů za minutu viz kapitola 4.2.4 [17].

3 Zpracování hudebního signálu

Obsahem této kapitoly je popis postupu při zpracování hudebního signálu za účelem následné statistické analýzy. Existuje řada celá řada podmínek, které musí hudební signál splňovat, aby mohla být provedena parametrizace.



Obr. 3.1: Vizualizace procesu zpracování hudebního signálu

Na obrázku 3.1 je možné vidět všechny kroky, které jsou prováděny v rámci provedení parametrizace hudebního signálu. V následujících podkapitolách budou jednotlivé bloky výše zmíněného řetězce popsány detailněji.

3.1 Předzpracování hudebního signálu

Tato podkapitola je věnována procesům, které předcházejí samotné parametrizaci hudebního signálu. Jsou zde popsány nutné vlastnosti, které musí hudební signál splňovat, aby došlo k jeho konzistentní parametrizaci.

3.1.1 Formát hudebního signálu

Tato podmínka určuje vstupní formát hudebního signálu. Aby měla výsledná statistická analýza dobrou vypovídající hodnotu, musí být zajištěn kvalitní vstupní hudební signál. Je tedy nutné použít bezztrátové formáty hudebního signálu jako vstupní data. Vstupními formáty jsou z převážné většiny nekomprimované soubory s koncovkou `.wav` a bezztrátově komprimované soubory s koncovkou `.flac`. Všechny nahrávky budou sjednoceny na jednotný formát `.wav`.

3.1.2 Vzorkovací kmitočet hudebního signálu

Nejprve je nutné sjednotit všechny nahrávky na jednotný vzorkovací kmitočet. V případě této práce bude použit jeden ze standardních vzorkovacích kmitočtů, konkrétně $f_{vz} = 44,1$ KHz. Tento vzorkovací kmitočet je standardem, který je používán zejména na fyzických nosičích typu Audio CD. Nahrávky s vyšším vzorkovacím kmitočtem budou převzorkovány na výše zmíněnou hodnotu f_{vz} .

Za účelem zachování kvality informace o spektru signálu, je zároveň tento vzorkovací kmitočet stanoven jako požadované minimum pro další zpracování signálu.

3.1.3 Bitová hloubka hudebního signálu

Bitovou hloubkou hudebního signálu je dán počet kvantovacích úrovní, kterých může signál dosahovat. Všechny nahrávky budou sjednoceny na hodnotu bitové hloubky 16 bitů. Tato hodnota společně s výše zmíněnou hodnotou f_{vz} taktéž odpovídá standardu používanému nosiči typu Audio CD a je jí také determinován celkový dynamický rozsah nahrávky.

3.1.4 Konverze do jednokanálové reprezentace

Poslední podmínkou před vypočtením parametrů je převod všech nahrávek databáze do jednokanálové reprezentace signálu. Valná většina dostupných nahrávek je reprezentována dvěma kanály, tedy jedná se o stereofonní nahrávky. Všechny nahrávky jsou v konečném důsledku převedeny na monofonní, tudíž výsledný hudební signál je reprezentován pouze jedním vektorem, kde po sobě jdoucí prvky vektoru reprezentují jednotlivé vzorky hudebního signálu.

3.2 Parametrizace hudebního signálu

Celkový průběh parametrizace hudebního signálu se dá rozdělit na dva na sebe navazující podprocesy, které budou níže podrobněji nastíněny.

3.2.1 MIR Parametry

Prvním z podprocesů je vypočtení MIR parametrů, které se dají považovat za objektivní popisy vlastností daného hudebního signálu. Na finské univerzitě v Jyväskylä je v současné době vyvíjena sada funkcí pro prostředí MATLAB nazvána jako MIR Toolbox [10]. Zmíněné funkce jsou schopny počítat parametry, kterými jsou hudební signály reprezentovány z hlediska dynamiky, tonality, rytmu nebo spektrálních vlastností. Tyto parametry jsou navrženy tak, aby byly schopné sledovat vlastnosti hudebního signálu v čase v závislosti na konkrétním hudebním obsahu nahrávky. Tyto parametry jsou z většiny reprezentovány pomocí vektorů, které odpovídají časové posloupnosti původního signálu. Parametry také mohou být reprezentovány skalárně. V tomto případě je skalární hodnota parametru vztažena k jednomu segmentu původního hudebního signálu. Délka tohoto segmentu je volena uživatelem a pohybuje se v rozmezí desítek, až stovek milisekund. Pokud konkrétní parametr není reprezentován vektorem, je možné vstupní hudební signál segmentovat a ke každému prvku

posloupnosti segmentů přiřadit skalární hodnotu získanou výpočtem parametru na daném segmentu vstupního hudebního signálu. Výsledkem tohoto procesu je vektor hodnot parametru, který ve výsledku opět popisuje hudební signál v celé délce jeho trvání.

Všechny parametry spadající do této kategorie budou dále podrobněji popsány v kapitole 4.

3.2.2 Statistické parametry

Jak již bylo zmíněno výše, výsledkem výpočtu MIR parametrů jsou vektorové reprezentace. Z těchto získaných reprezentací je dále nutné vypočítat statistické parametry. Statistických parametrů existuje celá řada, přičemž pro potřeby této práce budou vybrány pouze některé z nich. Výčet a vlastnosti vybraných statistických parametrů budou popsány v kapitole 5.

4 Parametry obsažené v MIR Toolboxu

V této kapitole budou podrobně popsány všechny vybrané parametry z MIR Toolboxu pro prostředí MATLAB. Všechny tyto parametry budou použity pro následnou klasifikaci nahrávek.

4.1 Parametry popisující dynamiku hudebního signálu

4.1.1 Efektivní hodnota signálu

Efektivní hodnota signálu, typicky označována jako hodnota RMS, může být použita pro odhad hlasitosti nebo jako indikátor zvukových událostí, používaných při segmentaci hudebního signálu. Efektivní hodnota je definována vztahem:

$$\text{RMS}_t = \sqrt{\frac{1}{K} \cdot \sum_{k=t \cdot K}^{(t+1) \cdot K - 1} s(k)^2}, \quad (4.1)$$

kde K je počet vzorků v každém segmentu a $s(k)^2$ je druhá mocnina amplitudy k -tého vzorku [8].

4.1.2 Počet průchodů nulovou úrovní

Tento parametr v původním znění *Zero-Crossing Rate* vyjadřuje počet změn znaménka hodnoty amplitudy signálu na daném časovém intervalu. Definice tohoto parametru je vyjádřena vztahem:

$$\text{ZCR}_t = \frac{1}{2} \cdot \sum_{k=t \cdot K}^{(t+1) \cdot K - 1} |\text{sgn}(s(k)) - \text{sgn}(s(k+1))|. \quad (4.2)$$

Přestože je tento parametr používán zejména v oblasti rozpoznávání řeči, v hudební nahrávce jím mohou být detekovány perkusní zvuky nebo hluk. V oblasti klasické hudby je tento parametr obvykle nižší a více stabilní [5].

4.1.3 Nízká energie

Křivka energie může být použita ke zhodnocení lokální distribuce energie, za účelem zjištění, zda energie zůstává v průběhu signálu konstantní nebo naopak jsou-li některé segmenty hudebního signálu kontrastnější než ostatní. Jedna z možností odhadu této závislosti je výpočet míry nízké energie (*low energy rate*), tedy procento segmentů, které vykazují hodnotu energie menší než průměrnou [27].

4.2 Parametry popisující rytmickou strukturu hudebního signálu

4.2.1 Fluktuace

Jedná se o odhad rytmické periodicity hudebního signálu s využitím modelu lidského slyšení. Tento proces odhadu rytmické struktury je proveden výpočtem spektrogramu, který je transformován pomocí sluchového modelu a poté je proveden odhad spektra v každém frekvenčním pásmu [19].

V prvním kroku výpočtu tohoto parametru je vypočítán výkonový spektrogram, poté je vypočítán Terhardtův model vnějšího ucha. Dále je provedena vícepásmová redistribuce energie za použití *bark* nebo *mel* pásem. Jsou odhadnuty vlivy maskování a nakonec jsou reprezentovány amplitudy v decibelech. Dalším krokem je provedení rychlé Fourierovy transformace (FFT) na každém pásmu. Výsledkem je matice rytmických periodicit pro každé pásmo *bark* či *mel*.

4.2.2 Počátek zvukové události

Ze závislostní křivky tohoto parametru je možné vyčíst po sobě jdoucí lokální nárůsty energie, odpovídající charakteru hudebního signálu. Je prováděna operace zvaná *peak picking*, která by se dala přirovnat k výběru umístění lokálních maxim z výše zmíněné křivky. Tato operace je prováděna za účelem zobrazení odhadu pozic událostí, jako jsou jednotlivé noty, akordy, či jiné zvukové události. Tento parametr je také využíván k zjištění změny tempa nahrávky v čase [10].

4.2.3 Hustota zvukových událostí

Tento parametr se přímo vztahuje k výše zmíněnému parametru počátek zvukové události. Jeho hodnotou je vyjadřován počet detekovaných zvukových událostí za sekundu [10].

4.2.4 Tempo

Parametr udávající tempo skladby je jedním ze základních ukazatelů o dané interpretaci skladby. Hodnota tohoto parametru je udávána v jednotkách BPM (*beats per minute* - počet úderů za minutu). Odhad této závislosti je běžně prováděn detekcí periodicit z výše popsané křivky počátků zvukových událostí. Zejména u klasické hudby jde o velmi proměnlivou hodnotu. Na základě lokálních a globálních změn tempa lze rozlišovat jednotlivé interpretace skladeb.

4.2.5 Metrický centroid

Jedná se o popis metrické aktivity, který je založen na výpočtu centroidu vybraných metrických úrovní. Z výsledné křivky tohoto parametru je možné vyčíst vývoj metrické aktivity v čase. Hodnota je udávána v jednotkách BPM, aby bylo možné tento parametr přímo srovnávat s hodnotou tempa. Vysoké hodnoty BPM metrického centroidu indikují členění na velký počet krátkých rytmických hodnot. Nižší hodnoty naopak korespondují s rytmickým členěním metra na celé noty nebo takty. Jestliže je některá z rytmických úrovní dominuje, hodnota metrického centroidu se přirozeně blíží k hodnotě tempa skladby [9].

4.2.6 Metrická síla

Metrická síla neboli *metrical strength* indikuje, zda se v nahrávce vyskytují čitelné pulzy, pravidelná a silná metrická struktura nebo jsou naopak pulzy z nahrávky těžko čitelné a jedná se o komplexní strukturu rytmických pulzů [9].

4.2.7 Čitelnost pulzu

Čitelnost pulzu je odhad rytmické čitelnosti a je jím indikována síla úderů vypočítaná z křivky tempa. Tento parametr má za úkol popsat míru rozlišitelnosti rytmické a metrické pulzace dané hudební kompozice nebo její části [11].

4.3 Parametry popisující barvu zvuku

4.3.1 Doba náběhu

Jedná se o jeden z parametrů, který je možno odvodit z křivky počátků zvukových událostí. V MIR Toolboxu je tento parametr označován jako *attack time* a pro jeho výpočet je provedena derivace energetické obálky podle času. Na křivce derivace energetické obálky můžeme pozorovat lokální nárůst těsně před začátkem doby náběhu. Počátek doby náběhu nastává tehdy, když křivka derivace energetické obálky dosáhne 20 % svého maxima. Když hodnota křivky klesne pod 20 % svého maxima, jedná se o konec doby náběhu [18].

4.3.2 Sklon náběhu

Sklon náběhu neboli *attack slope* je parametr, který definuje průměrný sklon energetické obálky v průběhu doby náběhu. Nejdříve jsou spočítány lokální příspěvky energie a potom je spočítán vážený průměr těchto příspěvků. Váhy jsou zvoleny tak, aby byla zdůrazněna hodnota sklonu uprostřed doby náběhu.

4.3.3 Skok náběhu

Skok náběhu, anglicky *attack leap* je parametr, který popisuje dobu náběhu na základě rozdílu amplitud počátku a konce doby náběhu. Hodnoty tohoto parametru jsou vyjádřeny ve stejném rozlišení jako byl původní signál.

4.3.4 Sklon poklesu

Parametr sklon poklesu (*decay slope*) udává míru poklesu energie signálu. Tímto parametrem je možné rozlišit rychlé perkusní zvuky od zvuků, které mají delší ustálený stav (*sustain*). Výpočet tohoto parametru je definován vztahem:

$$S(t) = A \cdot e(-\alpha(t - t_{max})) \cdot t_{nmax}, \quad (4.3)$$

kde α je odhadnuto lineární regresí logaritmu obálky energie [20].

4.3.5 Pokles spektrální energie

Původní název tohoto parametr je uváděn jako *spectral rolloff*. Jedná se o jeden ze způsob, jak odhadnout míru zastoupení vysokých frekvencí v signálu. To je provedeno nalezením frekvence, pod kterou je obsažen daný zlomek celkové energie audio signálu. Procentuálně je množství této energie vyjádřeno jako 85 % nebo 95 % [10].

4.3.6 Drsnost zvuku

Tento parametr, původním názvem *roughness*, reprezentuje míru vnímané disonance (nelibozvuku) dané zvukové nahrávky. Jev disonance vzniká, pokud se jednotlivé harmonické složky hudebního zvuku nacházejí blízko sebe ve frekvenční oblasti. Výsledkem tohoto parametru je odhad drsnosti zvuku na základě poměru každého páru harmonických složek, které jsou obsaženy ve frekvenčním spektru daného zvuku. Čím více složek s dostatečně velkou amplitudou se nachází v těsné blízkosti, tím je zároveň větší celková drsnost zvuku [10].

4.3.7 Melovské keprální koeficienty

Jedná se o parametry, které berou v úvahu nelineární a maskovací vlastnosti lidského slyšení a zároveň popisují spektrální rozložení energie hudebního signálu. Tyto parametry jsou již dlouhou dobu používány při zpracování lidské řeči, avšak své využití najdou i při zpracování hudebního signálu. Melovské keprální koeficienty byly již v minulosti použity pro klasifikaci hudebních žánrů nebo při rozpoznávání emocí v hudbě [6],[16]. Výpočet těchto parametrů je prováděn v několika krocích, které budou popsány níže.

1. Rozdělení signálu do rámců (segmentů) o délce 20–40 ms a vynásobení oknem, často používáno Hammingovo okno
2. Provedení rychlé Fourierovy transformace pro získání amplitudového spektra rámce signálu
3. Rovnoměrné rozmístění trojúhelníkových filtrů v kmitočtovém pásmu (MIR Toolbox pro MATLAB používá 40 trojúhelníkových filtrů)
4. Výpočet výkonového spektra rámce hudebního signálu
5. Vynásobení spektra jednotlivými trojúhelníkovými filtry
6. Zpětná diskretní kosinová transformace

4.3.8 Doba trvání

Doba trvání (*duration time*) udává dobu za kterou obálka energie přesáhne jistou prahovou úroveň, obvykle 40 %. Tímto parametrem lze určit např. dobu trvání noty v sekundách nebo rozlišit zvuky perkusní od zvuků s delším ustáleným stavem [10].

4.3.9 Parametry popisující tvar spektrálního rozložení

Šikmost spektra

Tímto parametrem je udávána míra asymetričnosti spektra okolo jeho průměrné hodnoty, anglický název je *spectral skewness*. Jestliže je hodnota rovna nule, jedná se o symetrické rozložení. V případě, že je hodnota menší než nula, je zastoupeno více energie v nízkých frekvencích. Při hodnotě větší než nula je více energie zastoupeno ve vysokých frekvencích.

Špičatost spektra

Jedná se o parametr, jímž je udávána míra špičatosti frekvenčního spektra kolem jeho průměrné hodnoty, anglický název je *spectral kurtosis*. Jestliže je tato hodnota rovna 3, jedná se o normální (Gaussovo) rozložení. Hodnota menší než 3 znamená plošší rozložení, hodnota větší než tři pak („špičatější“) rozložení.

Pločnost spektra

Tento parametr (anglicky *spectral flatness*) popisuje spektrum signálu na základě toho, jak moc se blíží charakter signálu šumu nebo naopak má signál spíše tónový charakter. Výpočet parametru je proveden jako podíl geometrického a aritmetického průměru signálu.

$$\frac{\sqrt[N]{\prod_{n=0}^{N-1} x(n)}}{\left(\frac{\sum_{n=0}^{N-1} x(n)}{N}\right)} \quad (4.4)$$

5 Vybrané statistické parametry

V této kapitole je představen výběr statistických parametrů, které budou počítány z vektorových reprezentací parametrů získaných v prostředí MATLAB pomocí MIR Toolboxu. Jejich výčet je možné vidět v níže uvedené tabulce 5.1.

Tab. 5.1: Vybrané statistické parametry

Označení parametru	Popis
mean	střední hodnota
var	rozptyl
std	směrodatná odchylka
median	medián
1p	1. percentil
99p	99. percentil
ir	interpercentilový rozsah
Q1	1. kvartil
Q3	3. kvartil
IQR	interkvartilový rozsah

5.1 Střední hodnota

Bývá nazývána jako očekávaná hodnota (odtud označení E - *Expected*, anglicky očekávaný) nebo populační průměr. Je to statistický parametr rozdělení náhodné veličiny, který je definován jako vážený průměr daného rozdělení. Pro náhodnou veličinu X s diskrétním rozdělením je definována jako:

$$E(X) = \sum_I s_i p_i, \quad (5.1)$$

kde $P[X = s_i] = p_i$ pro $i \in I$.

5.2 Rozptyl

Rozptyl je střední hodnota kvadrátů odchylek od střední hodnoty $E(X)$, vyjadřuje tedy variabilitu rozdělení souboru náhodných hodnot kolem její střední hodnoty. Bývá označován jako σ^2 nebo var . Výpočet rozptylu je definován vztahem:

$$\sigma^2 = \frac{1}{n} \sum_{i=1}^n (x_i - E(X))^2. \quad (5.2)$$

5.3 Směrodatná odchylka

Směrodatná odchylka značená řeckým písmem σ je v teorii pravděpodobnosti a statistice často používanou mírou statistické variability. Je definována vztahem:

$$\sigma = \sqrt{\sigma^2} = \sqrt{E((X - E(X))^2)}, \quad (5.3)$$

kde X je náhodná veličina, σ^2 její rozptyl a $E(X)$ její střední hodnota. Směrodatná odchylka vypovídá o tom, nakolik se od sebe navzájem typicky liší jednotlivé případy v souboru zkoumaných hodnot.

5.4 Medián

Medián je kvantil rozdělující statistický soubor na dvě stejně početné množiny. Bývá označován $Q_{0,5}$.

5.5 Percentil

Percentil rozděluje statistický soubor na setiny. K -tý percentil je označován jako $Q_{\frac{k}{100}}$.

5.6 Mezipercentilové rozpětí

IR (*Interpercentile range* - mezipercentilové rozpětí) udává rozdíl mezi 99. percentilem a 1. percentilem. Definováno je jako:

$$IR = P_{0,99} - P_{0,01}. \quad (5.4)$$

5.7 Kvartil

Tři kvartily rozdělují statistický soubor na čtvrtiny. Jako dolní kvartil - $Q_{0,25}$ je označováno 25 % prvků, jako horní kvartil - $Q_{0,75}$ je označováno 75 % prvků. Dolní a horní kvartily bývají také značeny jako Q_1 a Q_3 .

5.8 Mezikvartilové rozpětí

IQR (*Interquartile range* - mezikvartilové rozpětí) představuje rozdíl mezi třetím a prvním kvantilem. Reprezentuje tedy oblast hodnot, které má středních 50 % hodnot proměnné. Je definováno jako:

$$IQR = Q_3 - Q_1. \quad (5.5)$$

6 Synchronizace nahrávek

Byly vytvořeny celkem 3 databáze nahrávek české klasické komorní a orchestrální hudby viz kapitola 9. Mimo analýzu celých nahrávek jsou také analyzovány vybrané úseky skladeb zvlášť. Základním úkonem je tedy výběr stejného notového úseku ze všech interpretací. Výběr stejného notového úseku by mohl být proveden manuálně, konkrétně v podobě poslechu a vystříhnutí stejného úseku v libovolném editačním software. Pro účely této práce byl v prostředí MATLAB implementován nástroj, který umožňuje stejný notový úsek vyhledat a vystříhnout ze všech verzí skladby bez nutnosti časově náročného poslechu všech nahrávek. Tento nástroj využívá dynamické borcení časové osy (DTW) a je schopen provést stříh úseků s přesností 0,1 s. Na obrázku 6.1 je možné vidět znázornění procesu výběru stejného notového úseku z interpretací. Popis fungování tohoto nástroje bude popsán v následujících kapitolách.



Obr. 6.1: Proces výběru stejného notového úseku z interpretací

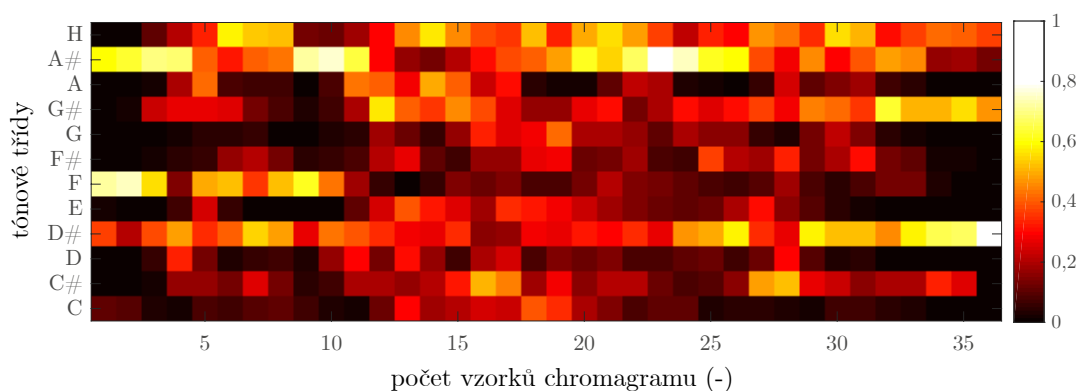
6.1 Chromagram

Pro úspěšnou synchronizaci interpretací pomocí dynamického borcení časové osy, je nutné definovat parametr, který vhodným způsobem popisuje hudební signál v čase. Parametr, který umožňuje provést robustní synchronizaci dvou nebo více interpretací dané skladby se nazývá *chromagram*. Tento parametr je nezávislý na změnách tempa nebo dynamiky daných interpretací a reprezentuje nahrávky na základě jejich tónového obsahu. Chromagram je možné vypočítat v různých rozlišení od hrubých po jemnější. V MATLABu je tento parametr je reprezentován maticí, která má 12 řádků, a počet sloupců se rovná počtu vzorků tohoto parametru. Řádky této matice odpovídají 12 tónům tzv. chromatické stupnice (tónové třídy). Chromatická stupnice obsahuje všechny tóny používané v západní hudbě, znázorněno v tabulce 6.1.

Tab. 6.1: Tónové třídy

1	2	3	4	5	6	7	8	9	10	11	12
C	C♯	D	D♯	E	F	F♯	G	G♯	A	A♯	H

Na obrázku 6.2 je možno vidět vizualizaci chromagramu nahrávky vypočítaného v prostředí MATLAB. Pomocí tohoto znázornění je možné pozorovat změny energie v jednotlivých tónových třídách podle toho, jak se měnil tónový obsah nahrávky v čase. Hodnoty chromagramu se pohybují v rozmezí od 0 do 1, přičemž hodnota 0 znamená, že se daný tón v nahrávce nevyskytuje, naopak hodnota 1 znamená maximální výskyt daného tónu z hlediska jeho energie. Tento druh normalizovaného chromagramu se nazývá CENS (Chroma Energy Normalized Statistics) a je hojně využíván v oblasti synchronizace nahrávek. Je využita implementace z Chroma Toolboxu pro MATLAB [15].



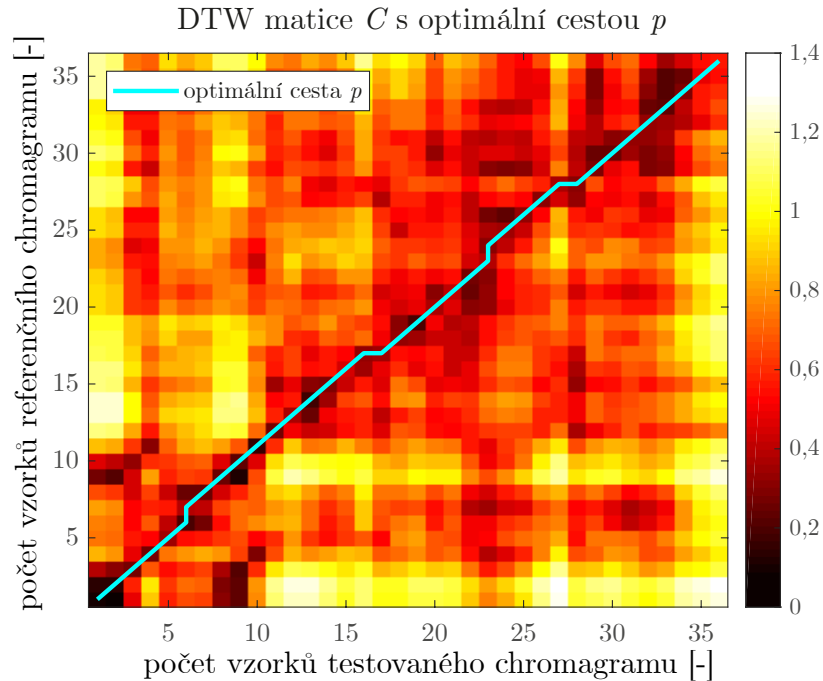
Obr. 6.2: Vizualizace chromagramu

Z důvodu časové náročnosti výpočtu chromagramů jsou tyto parametry po výpočtu uloženy ve formě .mat souborů používaných v prostředí MATLAB. Tento krok umožňuje provedení synchronizace interpretací a výběr stejného notového úseku pomocí načtení těchto souborů a využití dynamického borcení časové osy. Proces synchronizace nahrávek bude podrobněji popsán v následující kapitole.

6.2 Dynamické borcení časové osy

Dynamické borcení časové osy je známá technika zarovnávání (synchronizování) dvou parametrů o různých délkách. V případě této práce se jedná o synchronizaci dvou nebo více chromagramů na základě jednoho referenčního chromagramu. Jak již bylo zmíněno, chromagramy sledují tónový obsah nahrávek v čase. Všechny interpretace skladby se vyznačují různou dynamikou a tempem. Jediná společná vlastnost všech interpretací jedné skladby je téměř stejný vývoj tónového obsahu v čase. Synchronizace dvou chromagramů na nízkém rozlišení je vizualizována na obrázku 6.3.

Výsledkem synchronizace pomocí DTW je optimální cesta (*warping path*), která popisuje synchronizaci dvou chromagramů. Optimální cesta určuje, které vzorky testovaného chromagramu odpovídají vzorkům referenčního chromagramu z hlediska jejich časové posloupnosti.



Obr. 6.3: Vizualizace synchronizace dvou chromagramů

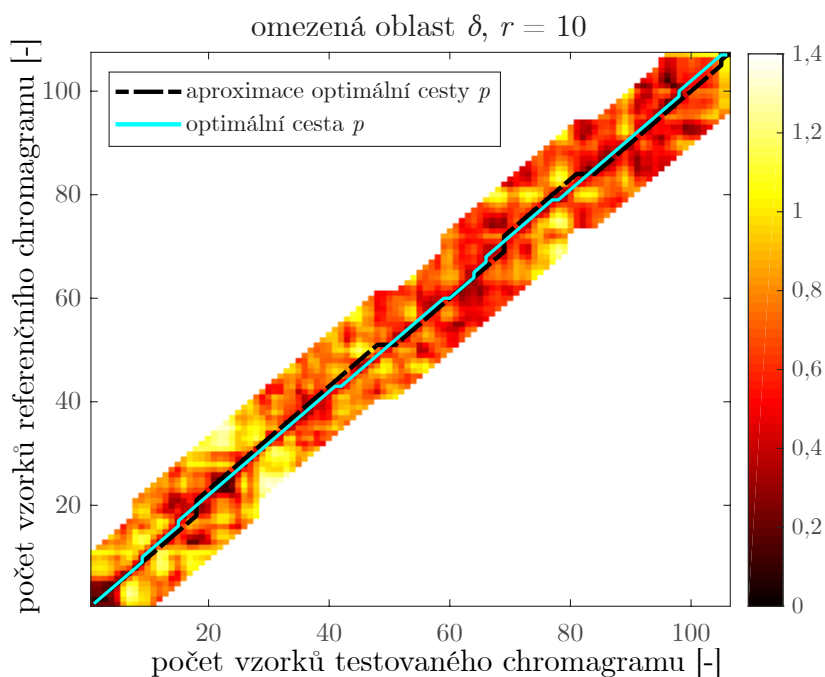
Optimální cesta p je v MATLABU vyjádřena jako matice o dvou sloupcích. První sloupec odpovídá vzorkům referenčního chromagramu a ve druhém sloupci jsou obsaženy příslušné vzorky testovaného chromagramu. Tímto postupem je získán přehled o tom, které vzorky testovaného chromagramu odpovídají vzorkům referenčního chromagramu na základě jejich tónového obsahu. Výše zobrazená synchronizace je provedena pomocí *Sakoe-Chiba* DTW algoritmu, konkrétně byla použita implementace `dynamicTimeWarpingSakoeChiba` z Audio Analysis Library Toolboxu pro MATLAB [7].

Ačkoli je použití jednoho z klasických algoritmů pro výpočet DTW funkční a dostatečně rychlé pro nízké rozlišení chromagramů, jak již bylo zmíněno cílem je dosažení synchronizace nahrávek s přesností 0,1 s. Chromagramy použité pro výslednou synchronizaci mají rozlišení 10 Hz, jinými slovy jedna sekunda nahrávky je reprezentována 10 vzorky chromagramu. Pro tento účel je vhodné použít víceúrovňové DTW neboli *Multi Scale Dynamic Time Warping*. Víceúrovňové DTW bylo v rámci této práce implementováno v MATLABu, postup vychází z publikací [23] a [14]. Implementace bude popsána v následující kapitole 6.3.

6.3 Nástroj pro synchronizaci nahrávek

Tento nástroj využívá pro synchronizaci nahrávek víceúrovňové DTW. Kompletní synchronizaci dvou nahrávek pomocí chromagramů je možné popsat v následujících krocích.

1. Výpočet základní úrovně synchronizace pomocí klasického algoritmu DTW, v tomto případě je použit již zmíněný algoritmus *Sakoe-Chiba*. Jsou použity chromagramy s hrubým rozlišením 1/9 Hz. Je získána optimální cesta.
2. Je třikrát zvětšeno rozlišení chromagramů a znovu je vypočítáno DTW, tentokrát na omezené oblasti δ . Omezená oblast δ matice C je definována aproximovanou cestou z předchozího kroku a hodnotami matice ve všech směrech kolem aproximované cesty p . Parametr r definuje velikost oblasti kolem aproximované cesty p . Ve výsledku je získána optimální cesta p s vyšším rozlišením.
3. Předchozí krok se opakuje s rozlišením chromagramů 1 Hz, 1 sekunda nahrávky odpovídá jednomu vzorku chromagramu.
4. Poslední krok je proveden totožně jako kroky 2 a 3. Je použito konečné rozlišení 10 Hz, 1 sekunda nahrávky odpovídá 10 vzorkům chromagramu. Výsledkem je synchronizace dvou nahrávek s přesností 0,1 s.



Obr. 6.4: Vizualizace omezené oblasti δ

Pro názornost je na obrázku 6.4 vyobrazena omezená oblast δ , kde parametr r udává počet hodnot matice ve směrech nahoru, dolů, doleva a doprava kolem aproximované cesty p . Pouze tyto hodnoty jsou použity pro výpočet DTW. Ostatní

hodnoty matice nejsou při výpočtu DTW brány v potaz a je tím zajištěn mnohonásobně kratší čas synchronizace. Parametr r je vhodně zvolen tak, aby nedocházelo k chybnému výpočtu DTW.

V MATLABu byla naprogramována funkce s názvem `multiLevelDTW`, první krok je počítán pomocí funkce `dynamicTimeWarpingSakoeChiba` převzaté z Audio Analysis Library Toolboxu pro MATLAB. Zbylé kroky jsou již provedeny pomocí naprogramované funkce `restrictedDTW`. Tato funkce počítá DTW na omezené oblasti δ pomocí funkce `delta`. Z výše uvedeného postupu vyplývá, že kroky 2 až 4 probíhají iterativně s rostoucím rozlišením chromagramů. Cílem je postupné zpřesňování optimální cesty až na konečnou úroveň. Hlavní benefit víceúrovňového DTW spočívá zejména ve výrazném zkrácení času synchronizace dvou nahrávek. Tato skutečnost je znázorněna v tabulce 6.2.

Tab. 6.2: Porovnání DTW algoritmů

Skladba	DTW (s)	víceúrov. DTW (s)
Janáčkův 1. smyčcový kvartet, 4. věta	14,92	0,67
Janáčkův 2. smyčcový kvartet, 4. věta	40,51	1,33
Dvořákova 9. symfonie, 4. věta	82,41	2,83

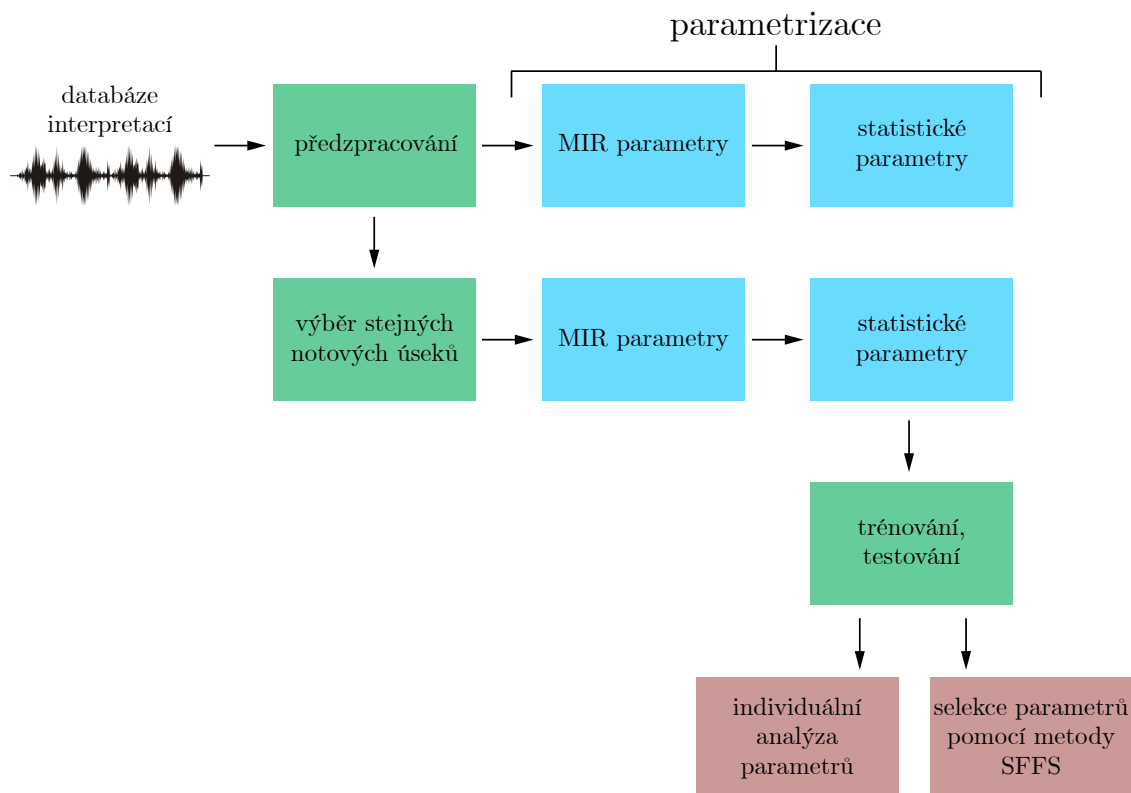
Ve výše zmíněné tabulce je uveden příklad průměrné doby trvání synchronizace vždy dvou interpretací jedné skladby pomocí klasického DTW a víceúrovňového DTW. Na první pohled je možné pozorovat razantní zkrácení doby synchronizace nahrávek. Skladby uvedené v tabulce jsou seřazeny od nejkratší po nejdelší. Nutno zmínit, že synchronizace a výsledná optimální cesta při použití víceúrovňového DTW je téměř identická ve srovnání s klasickým DTW algoritmem.

6.4 Výběr stejného notového úseku

V předchozí kapitole byl představen nástroj pro synchronizaci interpretací stejné skladby. Byly vypočítány chromagramy všech interpretací tří skladeb představených v kapitole 9. V prostředí MATLAB byla naprogramována funkce `batchDTW`, která slouží pro vystříhnutí stejného úseku ze všech interpretací. Funkce má pouze vstupní parametry `t1`, `t2`, `ref` a `sectionName`. Parametry `t1` a `t2` udávají počátek a konec vybraného notového úseku v referenční nahrávce v sekundách. Parametr `ref` udává číslo referenční nahrávky v databázi. Poslední parametr `sectionName` je textový řetězec typu string sloužící k pojmenování vybraného úseku. Po spuštění funkce uživatel je vyzván k zadání cesty k adresáři s vypočítanými chromagramy, poté je nutné zadat cestu k samotným nahrávkám. Je provedeno uložení vystřížených úseků

v adresáři s názvem odpovídajícím parametru `sectionName`, tento adresář je uložen v adresáři původní databáze nahrávek.

7 Vyhodnocovací systém



Obr. 7.1: Schéma vyhodnocovacího systému

Na obrázku 7.1 je znázorněno blokové schéma vyhodnocovacího systému, který bude použit pro klasifikaci nahrávek. V první části je nutné předzpracovat a sjednotit všechny nahrávky na jednotný formát. Výpočet parametrů bude proveden z celých nahrávek i vybraných notových úseků. Ze získaných MIR parametrů budou dále vypočítány statistické parametry. Statistické parametry jsou použity jako vstup pro klasifikační algoritmy. Všechny vypočítané statistické parametry jsou představeny v příloze B

Rozdělení nahrávek do kategorií je provedeno v podobě binární klasifikace, tedy databáze nahrávek jsou rozděleny na dvě kategorie. Konkrétně se jedná o kategorie Česko a svět. Celá statistická analýza je rozdělena na dvě části. První část je věnována individuální analýze parametrů. Tím je myšleno, že je klasifikace prováděna vždy pouze pomocí jednoho z vypočtených parametrů. Ve druhé části budou vybrány skupiny významných parametrů pomocí metody SFFS (*Sequential Floating Forward Search*). Bude zjištěno, jakou měrou se může přesnost klasifikace nahrávek zlepšit oproti předešlé individuální analýze parametrů.

8 Evaluace systému

V následujících podkapitolách budou nastíněny metody, pomocí kterých je možné provádět individuální klasifikaci jednotlivých statistických parametrů.

8.1 Klasifikační stromy

Klasifikační stromy nacházejí své využití v mnoha oborech. Jsou používány v disciplínách jako diagnóza, kognitivní věda, umělá inteligence nebo data mining. Hlavním úkolem klasifikačních stromů je vyprodukovat přesný klasifikátor.

Klasifikační strom klasifikuje instance pomocí třídění do stromu od jeho kořene až do jednotlivých listů, které poskytují klasifikaci instance. Každý uzel stromu specifikuje a testuje danou vlastnost instance a každá větev vycházející z daného uzlu koresponduje s možnou hodnotou této vlastnosti.

Klasifikace instance začíná v kořenovém uzlu stromu testováním vlastnosti, která je specifikována tímto uzlem. Poté se pokračuje směrem dolů po větvích stromu, kde větve odpovídají vlastnostem daného příkladu [25].

8.2 Křížová validace

Křížová validace je statistická metoda vyhodnocování a porovnávání učících se algoritmů rozdělením dat na dvě části. Jedna část je použita pro trénování a druhá pro testování daného modelu. Tato metoda je používána zejména v případě, kdy je potřeba vyhodnotit soubor, který nedisponuje velkým množstvím prvků.

Základní formou křížové validace je takzvaná *k-fold* křížová validace. Ostatní formy křížové validace jsou specifické případy *k-fold* křížové validace nebo vyžadují několik opakování.

Jedním z těchto případů může být metoda *leave-one-out*. V tomto případě *k-fold* křížové validace je k rovno počtu prvků v analyzovaném souboru. Jinými slovy v každé iteraci jsou téměř všechna data použita pro trénování a model je testován pomocí jednoho prvku [1].

8.3 Vyhodnocení algoritmů

V této kapitole jsou představeny analytické nástroje, pomocí kterých je možné zhodnotit učící se algoritmy a porovnat jejich klasifikační schopnosti.

8.3.1 Matice záměn

V oblasti strojového učení, zejména při statistické klasifikaci, je používána tzv. matice záměn neboli *confusion matrix*. Tato matice je znázorněna v podobě tabulky a umožňuje vizualizovat vlastnosti a kvalitu daného klasifikátoru při využití metody strojového učení, která je definována jako učení s učitelem (*supervised learning*) [22].

		Předpovězená kategorie	
		pozitivní	negativní
Skutečná kategorie	pozitivní	TP	FN
	negativní	FP	TN

Obr. 8.1: Grafické znázornění matice záměn

Na obrázku 8.1 je vyobrazena obecná matice záměn. Pole matice obsahují četnosti, kolikrát došlo na zkoumané datové množině k dané kombinaci skutečné a předpovězené hodnoty. Na hlavní diagonále jsou znázorněny případy správné klasifikace, na diagonále vedlejší pak případy klasifikace nesprávné. Jedná-li se o binární klasifikátor, jsou definovány 4 možné výsledky.

- True positive (TP): počet správně pozitivních nálezů
- False positive (FP): počet falešně pozitivních nálezů
- True negative (TN): počet správně negativních nálezů
- False negative (FN): počet falešně negativních nálezů

V následujících podkapitolách budou popsány některé vybrané metriky, které jsou běžně používány v rámci zhodnocení kvality a vlastností binárního klasifikátoru. Tyto metriky jsou vždy vypočteny z výše zmíněných možných výsledků.

8.3.2 Přesnost

Pomocí přesnosti (*accuracy*) může být hodnocena kvalita klasifikace provedená pomocí binárního klasifikátoru. Tato metrika může být značně zavádějící v případech nevyrovnaného zastoupení dvou klasifikačních tříd.

Pokud jsou brány v potaz výsledky z matice záměn, které byly zmíněny v předchozí kapitole 8.3.1, celková přesnost (ACC) je definována jako:

$$ACC = \frac{TP + TN}{TP + TN + FN + FP} \cdot 100 (\%). \quad (8.1)$$

Výsledkem této rovnice je procentuální vyjádření přesnosti (ACC).

8.3.3 Senzitivita a specificita

Senzitivita, jinými názvy také *true positive rate*, *recall*, udává, kolik procent vzorků z pozitivní kategorie bylo správně klasifikováno jako pozitivních. Senzitivita (SEN) je definována jako:

$$\text{SEN} = \frac{\text{TP}}{\text{TP} + \text{FN}} \cdot 100 (\%). \quad (8.2)$$

Specificita naopak udává, kolik procent vzorků z negativní kategorie bylo správně klasifikováno jako negativních. Specificita (SPE) je definována pomocí vztahu:

$$\text{SPE} = \frac{\text{TN}}{\text{TN} + \text{FP}} \cdot 100 (\%). \quad (8.3)$$

8.3.4 Precision

Hodnota precision (*positive predictive value*) je definována jako:

$$\text{PPV} = \frac{\text{TP}}{\text{TP} + \text{FP}} \cdot 100 (\%), \quad (8.4)$$

kde TP je počet skutečně pozitivních vzorků a FP je počet falešně pozitivních vzorků. Oproti senzitivitě, která procentuálně vyjadřuje počet správně zařazených pozitivních vzorků, hodnota PPV vyjadřuje procentuální počet vzorků, které byly klasifikovány jako pozitivní.

8.3.5 F1 Score

V oblasti statistické analýzy je také často používáno vyjádření kvality klasifikátoru, které má název F_1 Score [24]. Metrika F_1 Score je definována pomocí vztahu:

$$F_1 = \frac{2\text{TP}}{2\text{TP} + \text{FP} + \text{FN}}. \quad (8.5)$$

Tato metrika bere v potaz hodnoty precision (PPV) a recall (SEN), které jsou použity pro výpočet výsledné hodnoty F_1 . Tato hodnota je vypočtena jako harmonický průměr dvou výše zmíněných metrik precision a recall. Hodnota F_1 se pohybuje v rozmezí 0 až 1. Pokud je hodnota F_1 rovna 1, hodnoty metrik precision a recall jsou ideální.

8.3.6 Matthewsův korelační koeficient

Matthewsův korelační koeficient je používán v oblasti strojového učení pro určení míry kvality binárního klasifikátoru a byl poprvé představen biochemikem Brianem W. Matthewsem v roce 1975 [12].

Výpočet tohoto koeficientu je dán vztahem:

$$\text{MCC} = \frac{\text{TP} \cdot \text{TN} - \text{FP} \cdot \text{FN}}{\sqrt{(\text{TP} + \text{FP})(\text{TP} + \text{FN})(\text{TN} + \text{FP})(\text{TN} + \text{FN})}}. \quad (8.6)$$

Matthewsův korelační koeficient je považován za vyrovnanou metriku, která může být použita i v případě, pokud jsou klasifikované třídy velmi nevyrovnané [2].

Jedná se o korelační koeficient, pomocí kterého je sledována míra korelace mezi skutečnými a předpovězenými vzorky. Hodnota je pohybuje v intervalu od -1 do 1 . Je-li hodnota rovna 1 , jedná se o bezchybnou predikci. Pomocí nulové hodnoty je reprezentována ne lepší než náhodná predikce. Hodnotou -1 je označena totální neshoda mezi skutečnými hodnotami a predikcí.

9 Databáze nahrávek

Tato kapitola je věnována popisu databází nahrávek, které byly podrobeny statistické analýze za účelem jejich klasifikace do dvou kategorií Česko a svět. Byly vybrány celkem 3 skladby z oblasti české klasické hudby. Dvě z vybraných skladeb spadají do jednoho časového i stylového období a zároveň spadají do oblasti komorní hudby. Jedná se o dva smyčcové kvartety českého skladatele Leoše Janáčka, 1. smyčcový kvartet „Kreutzer Sonata“ a 2. smyčcový kvartet s názvem „Listy Důvěrné“. Třetí vybranou skladbou je 9. symfonie „Z nového světa“ skladatele Antonína Dvořáka. Tato skladba byla vybrána pro porovnání schopností klasifikátoru rozlišit dané interpretace orchestrální skladby v porovnání se skladbami komorního charakteru.

V tabulce 9.1 je znázorněno zastoupení jednotlivých interpretací v rámci jednotlivých skladeb. Celkový počet interpretací je omezen zejména dostupností samotných nahrávek v digitalizované podobě a jedná se o hlavní limitující faktor samotné statistické analýzy. V ideálním případě by měly být české a světové nahrávky v rámci všech skladeb přibližně stejně zastoupené. Je tedy nutné zdůraznit, že databáze dostupné k analýze jsou značně nevyvážené v rámci rozdělení interpretací na české a světové nahrávky. Tato skutečnost ale vychází již ze samotného rozdělení nahrávek do výše zmíněných kategorií. Bude tedy nutné nevyváženost databází zohlednit při zhodnocování výsledků analýzy.

Tab. 9.1: Informace o jednotlivých databázích nahrávek

skladba	počet	české	světové
Janáčkův 1. smyčcový kvartet, 4. věta	52	14	38
Janáčkův 2. smyčcový kvartet, 4. věta	51	16	35
4. věta Dvořákovy 9. symfonie, původ dirigenta	57	16	41
4. věta Dvořákovy 9. symfonie, původ orchestru	57	11	46

Je nutno poznamenat, že v případě Dvořákovy 9. symfonie „Z nového světa“ je provedeno dvojí rozdělení. Kategorizace v případě této nahrávky není jednoznačná. V prvním případě jsou nahrávky rozděleny podle původu dirigenta, v druhém případě je provedeno dělení nahrávek podle původu orchestru. Úkolem je zjistit, která z klasifikací bude vykazovat lepší výsledky. Ovšem i v tomto případě je nutné připomenout, že je tato klasifikace značně omezena nevyvážeností nahrávek v rámci dvou skupin. Konkrétně v případě Dvořákovy 9. symfonie je rozdělení nejméně ideální a opět bude brán zřetel na tento předpoklad v rámci evaluace konečných výsledků analýzy.

V kapitole 7 byl popsán obecný postup při analýze. Ve skutečnosti budou pomocí vytvořeného nástroje, který byl zmíněn v kapitole 6.4, z každé skladby vybrány dva

významně odlišné úseky v rámci všech interpretací. Odlišnost úseků je posuzována zejména na základě kontrastní úrovně dynamiky. Dále budou analyzovány všechny interpretace skladeb v jejich celkové délce trvání pro porovnání a kontrolu úspěšnosti klasifikace oproti vybraným úsekům.

Ve výsledku bude tedy provedena klasifikace každé databáze nahrávek třikrát. Ve dvou případech budou použity vybrané notové úseky, v případě třetím budou klasifikovány interpretace jako celky. V následujících kapitolách budou představeny všechny výsledky klasifikace v rámci tří databází nahrávek. Výsledky jsou rozděleny do dvou kapitol, které nesou názvy Individuální analýza parametrů a Selektce parametrů pomocí SFFS.

10 Individuální analýza parametrů

Tato kapitola je věnována první části statistické analýzy provedené v rámci této diplomové práce. Při individuální analýze parametrů je použit vždy jeden parametr z 320 vypočítaných statistických parametrů a je následně využit při trénování a testování klasifikačního algoritmu. Provedením tohoto procesu je zjištěna schopnost klasifikovat nahrávky do kategorií pouze pomocí jednoho parametru. V každé dílčí části analýzy bylo vybráno vždy pouze 5 nejvýznamnějších parametrů, které při klasifikaci vykázaly nejlepší vlastnosti. Všem nahrávkám z kategorie Česko byla přiřazena hodnota „1“, jinými slovy, jedná se o pozitivní vzorky. Nahrávky z kategorie svět mají přiřazeny hodnotu „0“, jedná se tedy naopak o vzorky negativní. Seznam všech vypočítaných statistických parametrů je dostupný v podobě přílohy B.

10.1 1. Janáčkův smyčkový kvartet

Jak již bylo zmíněno v kapitole 9, ze všech nahrávek byly vybrány dva kontrastní notové úseky, které byly podrobeny parametrizaci. Dále byla také provedena parametrizace nahrávek v celkové délce trvání.

Tab. 10.1: 1. Janáčkův smyčkový kvartet, individuální analýza parametrů, notový úsek A

Parametr	ACC	SEN	SPE	PPV	F ₁	MCC
mfcc12 (IQR)	71,04 %	69,52 %	72,54 %	52,19 %	0,55	0,39
spectralKurtosis (99p)	70,31 %	69,64 %	71,55 %	49,11 %	0,54	0,37
mfcc1 (ir)	70,29 %	68,24 %	72,28 %	48,34 %	0,52	0,36
eventDensity (Q1)	70,12 %	57,55 %	74,32 %	49,93 %	0,49	0,30
mfcc8 (1p)	68,85 %	54,52 %	74,82 %	47,65 %	0,46	0,27

ACC – přesnost klasifikace, SEN – senzitivita, SPE – specificita, PPV – positive predictive value, F₁ Score, MCC – Matthewsův korelační koeficient

V případě notového úseku A 1. Janáčkovy smyčkového kvartetu vykázal parametr mfcc12 (IQR) nejlepší vlastnosti z celé množiny vypočtených statistických parametrů. Bylo dosaženo přesnosti klasifikace ACC = 71,04 %. Klíčovou vlastností klasifikátoru je v tomto případě taktéž vysoká hodnota senzitivity, a to z důvodu procentuálního vyjádření správně zařazených nahrávek českého původu. Parametr mfcc12 (IQR) dosáhl druhé největší hodnoty senzitivity SEN = 69,52 %, nepatrně větší hodnotu senzitivity SEN = 69,64 % měl parametr spectralKurtosis (99p). Celkově nelze výsledky dosažené v rámci výše zmíněného notového pokládat za výrazně významné. Nutno dodat, že první notový úsek byl záměrně vybrán tak, aby se vyznačoval nižší průměrnou úrovní dynamiky.

Notový úsek B 1. Janáčkova smyčcového kvartetu byl vybrán tak, aby byl co nejvíce kontrastní oproti předchozímu úseku s nízkou dynamikou.

Tab. 10.2: 1. Janáčkův smyčcový kvartet, individuální analýza parametrů, notový úsek B

Parametr	ACC	SEN	SPE	PPV	F ₁	MCC
mfcc9 (mean)	84,68 %	74,39 %	88,86 %	76,36 %	0,71	0,63
mfcc3 (Q3)	81,76 %	79,56 %	83,21 %	67,71 %	0,69	0,59
mfcc9 (Q1)	81,54 %	78,50 %	83,08 %	65,92 %	0,67	0,58
mfcc9 (Q3)	80,73 %	69,08 %	86,72 %	70,36 %	0,64	0,55
mfcc8 (Q3)	72,67 %	62,33 %	76,44 %	54,48 %	0,53	0,37

ACC – přesnost klasifikace, SEN – senzitivita, SPE – specifická, PPV – positive predictive value, F₁ Score, MCC – Matthewsův korelační koeficient

Z tabulky 10.1 jsou na první pohled patrné dvě skutečnosti. Celkově bylo dosaženo větší přesnosti klasifikace, řádově o více než 10 % oproti předchozímu notovému úseku. Druhou skutečností je, že v případě notového úseku B byly bez výjimky nejvíce úspěšné statistické parametry vypočtené z Melovských křivkových koeficientů. Parametr mfcc9 (mean) vykázal významnou hodnotu přesnosti klasifikace ACC = 84,68 % a taktéž dosáhl vysokých hodnot F₁ Score F₁ = 0,71 a Matthewsova korelačního koeficientu MCC = 0,63. Za významnou se dá taktéž považovat hodnota senzitivity parametru mfcc3 (Q3) SEN = 79,56 % a hodnota senzitivity parametru mfcc9 (Q1) SEN = 78,50 %.

Tab. 10.3: 1. Janáčkův smyčcový kvartet, individuální analýza parametrů, celé nahrávky

Parametr	ACC	SEN	SPE	PPV	F ₁	MCC
mfcc9 (median)	77,88 %	79,29 %	77,35 %	60,07 %	0,63	0,53
mfcc8 (median)	75,91 %	75,69 %	77,12 %	59,79 %	0,61	0,49
brightness (var)	74,05 %	62,12 %	79,43 %	55,11 %	0,53	0,39
brightness (std)	73,72 %	59,69 %	79,37 %	54,91 %	0,52	0,37
mfcc9 (mean)	73,07 %	68,75 %	75,41 %	55,17 %	0,56	0,41

ACC – přesnost klasifikace, SEN – senzitivita, SPE – specifická, PPV – positive predictive value, F₁ Score, MCC – Matthewsův korelační koeficient

Parametr mfcc9 (median) vykázal nejlepší vlastnosti při klasifikaci celých nahrávek 1. Janáčkova smyčcového kvartetu. Hodnota přesnosti klasifikace tohoto parametru byla ACC = 77,88 %. Tento parametr zároveň vykázal významnou hodnotu senzitivity SEN = 79,29 %. Jsou-li zhodnoceny výsledky individuální analýzy parametrů v případě 1. Janáčkova smyčcového kvartetu, nejlépe byly klasifikovány notové

úseky s vysokou úrovní dynamiky. Poté následovaly celé nahrávky, přičemž nejhůře dopadla klasifikace notových úseků s nízkou úrovní dynamiky.

10.2 2. Janáčkův smyčcový kvartet

Stejně jako v případě 1. Janáčkova smyčcového kvartetu byly vybrány dva kontrastní úseky na základě úrovně dynamiky. Jako první byla provedena klasifikace notového úseku s nízkou úrovní dynamiky.

Tab. 10.4: 2. Janáčkův smyčcový kvartet, individuální analýza parametrů, notový úsek A

Parametr	ACC	SEN	SPE	PPV	F ₁	MCC
spectralFlatness (ir)	78,46 %	71,29 %	82,34 %	67,83 %	0,64	0,52
spectralFlatness (std)	72,49 %	61,04 %	77,58 %	56,96 %	0,54	0,37
spectralFlatness (var)	71,95 %	63,02 %	77,22 %	56,37 %	0,55	0,38
decayLeap (ir)	70,73 %	67,07 %	73,04 %	55,52 %	0,55	0,38
decayTime (1p)	70,40 %	65,71 %	72,91 %	54,25 %	0,54	0,37

ACC – přesnost klasifikace, SEN – senzitivita, SPE – specificita, PPV – positive predictive value, F₁ Score, MCC – Matthewsův korelační koeficient

Významnou hodnotu přesnosti klasifikace vykázal parametr spectralFlatness (ir) ACC = 78,46%. Za zmínku také stojí hodnoty senzitivity a specificity tohoto parametru – SEN = 71,29%, SPE = 82,34%.

Tab. 10.5: 2. Janáčkův smyčcový kvartet, individuální analýza parametrů, notový úsek B

Parametr	ACC	SEN	SPE	PPV	F ₁	MCC
brightness (ir)	88,84 %	90,31 %	88,34 %	78,49 %	0,81	0,75
brightness (IQR)	85,37 %	81,97 %	86,82 %	75,77 %	0,74	0,66
mfcc1 (IQR)	84,21 %	84,85 %	84,52 %	72,13 %	0,75	0,66
spectralSpread (std)	80,62 %	82,97 %	80,21 %	66,89 %	0,70	0,59
brightness (std)	80,47 %	82,84 %	79,48 %	67,77 %	0,70	0,60

ACC – přesnost klasifikace, SEN – senzitivita, SPE – specificita, PPV – positive predictive value, F₁ Score, MCC – Matthewsův korelační koeficient

Při analýze druhého notového úseku bylo dosaženo až o 10% větší přesnosti klasifikace. Podobné zlepšení přesnosti klasifikace bylo zjištěno i v případě nahrávek 1. Janáčkova smyčcového kvartetu. Je tedy možné spekulovat o vlivu vysoké úrovně dynamiky vybraných částí nahrávek na celkovou úspěšnost klasifikace. Parametr brightness (ir) dosáhl významné hodnoty přesnosti klasifikace ACC = 88,84%, a zároveň hodnota senzitivity tohoto parametru, přesáhla hranici 90%.

Tab. 10.6: 2. Janáčkův smyčcový kvartet, individuální analýza parametrů, celé nahrávky

Parametr	ACC	SEN	SPE	PPV	F ₁	MCC
mfcc1 (var)	91,68 %	89,23 %	92,75 %	86,05 %	0,84	0,80
mfcc1 (std)	91,40 %	89,84 %	92,09 %	85,27 %	0,85	0,80
mfcc1 (ir)	83,28 %	83,56 %	84,01 %	69,63 %	0,72	0,63
spectralSkewness (Q3)	77,45 %	70,76 %	80,33 %	65,41 %	0,64	0,50
mfcc1I (QR)	75,80 %	84,67 %	71,99 %	58,30 %	0,66	0,53

ACC – přesnost klasifikace, SEN – senzitivita, SPE – specificita, PPV – positive predictive value, F₁ Score, MCC – Matthewsův korelační koeficient

Nejllepších výsledků bylo dosaženo při analýze celých nahrávek 2. Janáčkovy smyčcového kvartetu, kde dva z parametrů dosáhly přesnosti klasifikace přesahující hranici 90 %. V tabulce 10.6 je možno vidět, že až na výjimku v podobě parametru spectralSkewness (Q3), opět dominují parametry vypočtené z Melovských kepst-rálních koeficientů. Významné přesnosti klasifikace ACC = 91,68 % bylo dosaženo v případě parametru mfcc1 (var), přičemž je tento výsledek podpořen i vysokou hodnotou senzitivity SEN = 89,23 %.

10.3 Dvořáková 9. symfonie, rozdělení dle dirigenta

Nahrávky Dvořákovy 9. symfonie byly nejprve rozděleny podle původu dirigentů. Opět byly vybrány dva úseky k analýze a zároveň byly parametrizovány skladby jako celky.

Tab. 10.7: Dvořáková 9. symfonie, individuální analýza parametrů, notový úsek A, klasifikace podle dirigentů

Parametr	ACC	SEN	SPE	PPV	F ₁	MCC
decayTime (99p)	73,40 %	65,91 %	76,46 %	53,94 %	0,55	0,40
decayTime (ir)	72,53 %	66,08 %	75,93 %	53,71 %	0,54	0,39
mfcc11 (mean)	71,58 %	57,23 %	78,31 %	51,19 %	0,49	0,33
rollOff (ir)	70,62 %	57,13 %	75,95 %	51,81 %	0,50	0,31
mfcc11 (median)	69,85 %	49,03 %	78,03 %	46,13 %	0,43	0,26

ACC – přesnost klasifikace, SEN – senzitivita, SPE – specificita, PPV – positive predictive value, F₁ Score, MCC – Matthewsův korelační koeficient

Při analýze prvního notového úseku nebylo dosaženo zvláště významných výsledků klasifikace. Parametr decayTime (99p) dosáhl přesnosti klasifikace ACC = 73,40 %. Nejvyšší hodnotu senzitivity SEN = 66,08 % vykázal parametr decayTime (ir).

Tab. 10.8: Dvořákova 9. symfonie, individuální analýza parametrů, notový úsek B, klasifikace podle dirigentů

Parametr	ACC	SEN	SPE	PPV	F ₁	MCC
mfcc1 (Q3)	73,44 %	69,36 %	74,88 %	52,18 %	0,56	0,41
tempo (median)	72,08 %	72,66 %	71,92 %	50,33 %	0,55	0,40
attackTime (std)	69,90 %	68,17 %	70,26 %	47,94 %	0,52	0,35
mfcc8 (ir)	69,28 %	63,86 %	72,06 %	46,10 %	0,49	0,32
attackTime (var)	68,79 %	68,84 %	69,00 %	47,63 %	0,52	0,34

ACC – přesnost klasifikace, SEN – senzitivita, SPE – specificita, PPV – positive predictive value, F₁ Score, MCC – Matthewsův korelační koeficient

Je-li porovnán druhý vybraný notový úsek s předchozím, výsledky klasifikace vykazují přibližně stejný charakter. Pouze parametr tempo (median) dosáhl vyšší hodnoty senzitivity SEN = 72,66 %.

Tab. 10.9: Dvořákova 9. symfonie, individuální analýza parametrů, celé nahrávky, klasifikace podle dirigentů

Parametr	ACC	SEN	SPE	PPV	F ₁	MCC
attackSlope (ir)	72,09 %	53,27 %	79,48 %	55,45 %	0,49	0,33
attackSlope (99p)	72,05 %	52,30 %	80,43 %	54,73 %	0,47	0,32
attackLeap (99p)	71,95 %	50,39 %	79,94 %	54,36 %	0,47	0,31
spectralFlatness (var)	71,54 %	61,85 %	75,44 %	52,36 %	0,52	0,35
attackLeap (ir)	71,54 %	51,88 %	79,06 %	52,79 %	0,47	0,31

ACC – přesnost klasifikace, SEN – senzitivita, SPE – specificita, PPV – positive predictive value, F₁ Score, MCC – Matthewsův korelační koeficient

Nakonec byly analyzovány nahrávky v celé délce trvání. Je-li brána v potaz přesnost klasifikace nahrávek, výsledky jsou srovnatelné s oběma vybranými notovými úseky. Přesnost klasifikace se u všech parametrů pohybuje na hranici 70 % a nejedná se o nikterak významné výsledky, zvláště je-li bráno v potaz vysoce nevyvážené zastoupení nahrávek.

10.4 Dvořáková 9. symfonie, rozdělení dle orchestru

Tato část analýzy byla provedena totožně jako část předcházející. Jediným rozdílem bylo rozdělení databáze nahrávek podle původu orchestru.

Tab. 10.10: Dvořákova 9. symfonie, individuální analýza parametrů, notový úsek A, klasifikace podle orchestrů

Parametr	ACC	SEN	SPE	PPV	F ₁	MCC
mfcc3 (mean)	74,97 %	58,86 %	79,60 %	45,23 %	0,45	0,34
tempo (var)	74,17 %	72,62 %	75,07 %	44,22 %	0,50	0,39
mfcc1 (Q1)	74,05 %	62,62 %	78,14 %	43,19 %	0,45	0,35
tempo (std)	73,66 %	70,04 %	75,21 %	42,93 %	0,48	0,37
attackSlope (var)	73,35 %	61,37 %	77,01 %	43,02 %	0,45	0,33

ACC – přesnost klasifikace, SEN – senzitivita, SPE – specificita, PPV – positive predictive value, F₁ Score, MCC – Matthewsův korelační koeficient

Celkově bylo možné pozorovat mírný nárůst přesnosti klasifikace oproti rozdělení podle původu dirigentů. V případě prvního notového úseku bude zmíněn pouze parametr tempo (var), u kterého dosáhla hodnota přesnosti klasifikace ACC = 74,17 %, přičemž senzitivita tohoto parametru byla SEN = 72,62 %.

Tab. 10.11: Dvořákova 9. symfonie, individuální analýza parametrů, notový úsek B, klasifikace podle orchestrů

Parametr	ACC	SEN	SPE	PPV	F ₁	MCC
brightness (99p)	78,67 %	66,05 %	82,48 %	51,74 %	0,52	0,43
mfcc7 (99p)	75,92 %	55,36 %	81,30 %	46,20 %	0,44	0,33
mfcc1 (median)	74,75 %	78,60 %	74,79 %	44,72 %	0,53	0,43
mfcc1 (Q1)	73,48 %	67,19 %	75,70 %	42,62 %	0,46	0,35
mfcc4 (99p)	72,64 %	66,76 %	74,65 %	41,30 %	0,45	0,33

ACC – přesnost klasifikace, SEN – senzitivita, SPE – specificita, PPV – positive predictive value, F₁ Score, MCC – Matthewsův korelační koeficient

V tabulce 10.11 je možno vidět výsledky individuální analýzy parametrů druhého vybraného notového úseku Dvořákovy 9. symfonie. Parametr brightness (99p) vykázal přesnost klasifikace ACC = 78,67 %. Největší senzitivity klasifikace dosáhl parametr mfcc1 (median) – SEN = 78,60 %.

Tab. 10.12: Dvořákova 9. symfonie, individuální analýza parametrů, celé nahrávky, klasifikace podle orchestrů

Parametr	ACC	SEN	SPE	PPV	F ₁	MCC
rollOff (std)	76,99 %	55,46 %	83,18 %	45,53 %	0,44	0,33
rollOff (var)	76,95 %	56,27 %	82,49 %	45,12 %	0,44	0,33
mfcc5 (Q3)	76,64 %	68,63 %	78,73 %	48,70 %	0,51	0,41
tempo (median)	74,83 %	64,34 %	78,40 %	45,34 %	0,47	0,36
spectralSpread (median)	74,65 %	57,94 %	79,65 %	47,65 %	0,45	0,34

ACC – přesnost klasifikace, SEN – senzitivita, SPE – specificita, PPV – positive predictive value, F₁ Score, MCC – Matthewsův korelační koeficient

V rámci analýzy celých skladeb opět nebylo dosaženo výrazně dobrých výsledků. Přesnost klasifikace ani v jednom z případů nepřesáhla hranici 80 %.

10.5 Zhodnocení individuální analýzy parametrů

Jsou-li zhodnoceny všechny výsledky individuální analýzy parametrů, lze vyvodit několik závěrů.

Oba analyzované smyčcové kvartety vykazují lepší přesnost klasifikace oproti nahrávkám symfonického orchestru. Zatímco u orchestrálních nahrávek nebyla dosažena přesnost klasifikace více jak 80 %, u obou smyčcových kvartetů byla tato hranice překonána. V případě 2. smyčcového kvartetu byla dosažena hodnota klasifikace přesahující hranici 90 %.

Je-li brána v potaz přesnost klasifikace v rámci dvou smyčcových kvartetů, byly dosaženy srovnatelné výsledky. Z této skutečnosti vyplynulo, že i v rámci skladeb jednoho autora (Leoše Janáčka) bylo dosaženo obdobných výsledků.

Zároveň je možné tvrdit, že při rozdělení orchestrálních nahrávek podle původu orchestru bylo dosaženo nepatrně lepších výsledků v porovnání s rozdělením podle původu dirigenta. Tato skutečnost však není úplně jednoznačná, protože použitá databáze orchestrálních nahrávek nemá ani zdaleka ideálně rozložené zastoupení zkoumaných kategorií.

11 Selekcce parametrů pomocí SFFS

V této kapitole budou představeny výsledky dosažené za pomoci metody SFFS, která je použita k selekci skupin významných parametrů. Z celkové množiny 320 vypočítaných statistických parametrů byly pomocí této metody selektovány pouze ty parametry, které jsou relevantní při řešení problematiky klasifikace do kategorií, které byly již výše představeny. Hlavní úkolem je vybrat podmnožinu z celkového počtu vypočtených parametrů, která zlepšuje kvality klasifikátoru oproti pouhé individuální analýze parametrů. Pro řešení tohoto problému byla zvolena metoda SFFS (*Sequential Floating Forward Selection*) [13]. Selekcce parametrů je důležitý krok pro zredukování dimensionality prostoru parametrů a je tak získán přehled o relevanci daných parametrů v rámci statistické analýzy. Pro tyto účely byla zvolena implementace v prostředí MATLAB dostupná z [4].

11.1 1. Janáčkův smyčcový kvartet

Stejně jako při individuální analýze parametrů byly použity totožné notové úseky a nahrávky jako celky.

Tab. 11.1: 1. Janáčkův smyčcový kvartet, selekcce SFFS, notový úsek A

n	ACC	SEN	SPE	PPV	F ₁	MCC
1	72,04 %	71,47 %	72,27 %	51,56 %	0,60	0,41
2	75,46 %	69,07 %	78,05 %	56,92 %	0,62	0,45
3	77,00 %	75,60 %	77,57 %	58,04 %	0,65	0,50
4	76,54 %	76,53 %	76,54 %	57,29 %	0,65	0,50
5	77,19 %	80,00 %	76,05 %	57,88 %	0,67	0,52

n – počet parametrů, ACC – přesnost klasifikace, SEN – senzitivita, SPE – specificita, PPV – positive predictive value, F₁ Score, MCC – Matthewsův korelační koeficient

V případě notového úseku A došlo ke zlepšení přesnosti klasifikace o přibližně 7 % oproti individuální analýze parametrů. Ve výsledku bylo dosaženo hodnoty přesnosti klasifikace ACC = 77,19 %, a to za použití kombinace 5 parametrů.

Tab. 11.2: 1. Janáčkův smyčcový kvartet, selekce SFFS, notový úsek B

n	ACC	SEN	SPE	PPV	F ₁	MCC
1	84,12 %	74,13 %	88,16 %	71,98 %	0,73	0,62
2	91,35 %	86,13 %	93,46 %	84,60 %	0,85	0,79
3	93,23 %	86,40 %	96,00 %	90,01 %	0,88	0,83
4	93,90 %	87,20 %	95,62 %	89,44 %	0,88	0,84
5	93,35 %	86,67 %	96,05 %	90,16 %	0,88	0,84

n – počet parametrů, ACC – přesnost klasifikace, SEN – senzitivita, SPE – specifická, PPV – positive predictive value, F₁ Score, MCC – Matthewsův korelační koeficient

Při selekci parametrů v rámci druhého notového úseku byla překonána hranice 90 % přesnosti klasifikace již za použití 2 parametrů. Výsledná hodnota přesnosti klasifikace při kombinaci 5 parametrů byla ACC = 93,35 %.

Tab. 11.3: 1. Janáčkův smyčcový kvartet, selekce SFFS, celé nahrávky

n	ACC	SEN	SPE	PPV	F ₁	MCC
1	77,31 %	79,07 %	76,59 %	57,86 %	0,67	0,52
2	87,85 %	92,27 %	86,05 %	72,83 %	0,81	0,74
3	90,69 %	92,80 %	89,84 %	78,93 %	0,85	0,79
4	91,12 %	95,20 %	89,46 %	78,77 %	0,86	0,81

n – počet parametrů, ACC – přesnost klasifikace, SEN – senzitivita, SPE – specifická, PPV – positive predictive value, F₁ Score, MCC – Matthewsův korelační koeficient

Jako poslední byla provedena selekce parametrů pro klasifikaci celých nahrávek. Maximální hodnota přesnosti klasifikace byla ACC = 91,12 %, přičemž je tento výsledek podpořen významnou hodnotou senzitivity SEN = 95,20 %. Tato hodnota se dá považovat za významnou zejména díky tomu, že díky ní lze procentuálně vyjádřit kolik českých nahrávek bylo správně klasifikováno.

11.2 2. Janáčkův smyčcový kvartet

Tab. 11.4: 2. Janáčkův smyčcový kvartet, selekce SFFS, notový úsek A

n	ACC	SEN	SPE	PPV	F ₁	MCC
1	79,06 %	71,75 %	82,40 %	65,58 %	0,68	0,53
2	82,35 %	78,75 %	84,00 %	69,38 %	0,74	0,61
3	85,10 %	84,50 %	85,37 %	72,66 %	0,78	0,67
4	85,33 %	85,13 %	85,43 %	72,91 %	0,78	0,68

n – počet parametrů, ACC – přesnost klasifikace, SEN – senzitivita, SPE – specifická, PPV – positive predictive value, F₁ Score, MCC – Matthewsův korelační koeficient

V případě notového úseku A 2. smyčcového kvartetu Leoše Janáčka došlo ke zlepšení přesnosti klasifikace o přibližně 7 % oproti individuální analýze parametrů, přičemž finální hodnota přesnosti klasifikace byla ACC = 85,33 %.

Tab. 11.5: 2. Janáčkův smyčcový kvartet, selekce SFFS, notový úsek B

n	ACC	SEN	SPE	PPV	F ₁	MCC
1	88,39 %	89,13 %	88,06 %	77,43 %	0,83	0,75
2	94,94 %	95,13 %	94,86 %	89,45 %	0,92	0,89
3	95,88 %	94,25 %	96,63 %	92,77 %	0,93	0,91
4	96,35 %	95,63 %	96,69 %	92,97 %	0,94	0,92

n – počet parametrů, ACC – přesnost klasifikace, SEN – senzitivita, SPE – specifická, PPV – positive predictive value, F₁ Score, MCC – Matthewsův korelační koeficient

Je-li brán v potaz notový úsek B 2. smyčcového kvartetu, přesnost klasifikace překročila hranici 90 % již při kombinaci dvou parametrů, konkrétně ACC = 94,94 %. Hodnota senzitivity při kombinaci dvou parametrů taktéž dosáhla významné hodnoty – SEN = 95,13 %.

Tab. 11.6: 2. Janáčkův smyčcový kvartet, selekce SFFS, celé nahrávky

n	ACC	SEN	SPE	PPV	F ₁	MCC
1	91,84 %	90,00 %	92,69 %	84,97 %	0,87	0,81
2	94,90 %	92,50 %	96,00 %	91,49 %	0,92	0,88
3	95,53 %	93,63 %	96,40 %	92,31 %	0,93	0,90
4	95,61 %	93,63 %	96,51 %	92,54 %	0,93	0,90

n – počet parametrů, ACC – přesnost klasifikace, SEN – senzitivita, SPE – specifická, PPV – positive predictive value, F₁ Score, MCC – Matthewsův korelační koeficient

Selekce parametrů při klasifikaci celých nahrávek se dá označit za směrodatnou zejména díky tomu, že nedošlo k subjektivnímu výběru notových úseků. Stejně jako v případě 1. smyčcového kvartetu bylo dosaženo významné hodnoty přesnosti klasifikace $ACC = 94,90\%$, a to za použití pouze dvou parametrů. Při této kombinaci bylo zároveň dosaženo i vysoké hodnoty senzitivity $SEN = 92,50\%$.

11.3 Dvořáková 9. symfonie, rozdělení dle dirigenta

Tab. 11.7: 9. Dvořáková symfonie, selekce SFFS, klasifikace podle dirigentů, notový úsek A

n	ACC	SEN	SPE	PPV	F ₁	MCC
1	73,61 %	66,25 %	76,49 %	52,84 %	0,59	0,40
2	85,44 %	75,75 %	89,22 %	74,44 %	0,75	0,65
3	85,54 %	82,75 %	86,63 %	71,34 %	0,76	0,67
4	85,93 %	82,25 %	87,37 %	72,17 %	0,77	0,67
5	86,63 %	82,50 %	88,24 %	73,42 %	0,78	0,68

n – počet parametrů, ACC – přesnost klasifikace, SEN – senzitivita, SPE – specifická, PPV – positive predictive value, F₁ Score, MCC – Matthewsův korelační koeficient

V případě prvního notového úseku Dvořákovy 9. symfonie došlo ke zlepšení celkové přesnosti klasifikace přibližně o 13 % oproti individuální analýze parametrů. Nejen v tomto případě však není hodnota přesnosti klasifikace nejdůležitějším ukazatelem zejména díky nevybalancované databázi nahrávek. Tato skutečnost je potvrzena i hodnotou senzitivity $SEN = 82,50\%$.

Tab. 11.8: 9. Dvořáková symfonie, selekce SFFS, klasifikace podle dirigentů, notový úsek B

n	ACC	SEN	SPE	PPV	F ₁	MCC
1	75,58 %	76,50 %	75,22 %	54,92 %	0,64	0,48
2	78,00 %	77,75 %	78,10 %	58,22 %	0,66	0,52
3	78,14 %	72,25 %	80,44 %	59,38 %	0,65	0,50
4	78,63 %	75,75 %	79,76 %	59,87 %	0,67	0,52

n – počet parametrů, ACC – přesnost klasifikace, SEN – senzitivita, SPE – specifická, PPV – positive predictive value, F₁ Score, MCC – Matthewsův korelační koeficient

V tabulce 11.8 je možno vidět, že při selekci parametrů pro klasifikaci druhého notového úseku nebylo dosaženo výrazného zlepšení vlastností klasifikace.

Tab. 11.9: 9.Dvořákova symfonie, selekce SFFS, klasifikace podle dirigentů, celé nahrávky

n	ACC	SEN	SPE	PPV	F ₁	MCC
1	79,44 %	76,13 %	80,73 %	61,14 %	0,68	0,54
2	81,16 %	85,88 %	79,32 %	62,16 %	0,72	0,60
3	83,89 %	88,63 %	82,05 %	65,91 %	0,76	0,66
4	86,28 %	96,50 %	82,29 %	68,05 %	0,80	0,72

n – počet parametrů, ACC – přesnost klasifikace, SEN – senzitivita, SPE – specifická, PPV – positive predictive value, F₁ Score, MCC – Matthewsův korelační koeficient

Naopak je nutno zmínit, že při klasifikaci celých nahrávek bylo díky selekci parametrů dosaženo významné hodnoty senzitivity SEN = 96,50 %, tedy je potenciálně možné pozorovat jistý vliv původu dirigenta na výsledek klasifikace nahrávek.

11.4 Dvořáková 9. symfonie, rozdělení dle orchestru

Tab. 11.10: 9.Dvořákova symfonie, selekce SFFS, klasifikace podle orchestrů, notový úsek A

n	ACC	SEN	SPE	PPV	F ₁	MCC
1	74,39 %	71,67 %	75,11 %	43,42 %	0,54	0,40
2	79,93 %	68,00 %	83,11 %	52,11 %	0,59	0,47
3	82,39 %	67,17 %	86,44 %	57,69 %	0,62	0,51
4	82,81 %	67,67 %	86,84 %	58,63 %	0,63	0,52

n – počet parametrů, ACC – přesnost klasifikace, SEN – senzitivita, SPE – specifická, PPV – positive predictive value, F₁ Score, MCC – Matthewsův korelační koeficient

Při rozdělení orchestrální nahrávek dle původu orchestru bylo zastoupení kategorií nejméně ideální. Během analýzy prvního notového úseku nebylo dosaženo příliš významných výsledků klasifikace.

Tab. 11.11: 9.Dvořákova symfonie, selekce SFFS, klasifikace podle orchestrů, notový úsek B

n	ACC	SEN	SPE	PPV	F ₁	MCC
1	79,16 %	65,67 %	82,76 %	50,76 %	0,57	0,44
2	83,30 %	71,33 %	86,49 %	59,00 %	0,64	0,54
3	85,44 %	74,17 %	88,44 %	63,13 %	0,68	0,59
4	86,42 %	78,67 %	88,49 %	64,75 %	0,71	0,63

n – počet parametrů, ACC – přesnost klasifikace, SEN – senzitivita, SPE – specifická, PPV – positive predictive value, F₁ Score, MCC – Matthewsův korelační koeficient

I v případě druhého notového úseku nebylo dosaženo příliš příznivých výsledků. Bylo dosaženo hodnoty přesnosti klasifikace $ACC = 86,42\%$, nicméně senzitivita nepřesáhla hranici 80% . Dosažená přesnost klasifikace byla ovlivněna především větším počtem správně zařazených světových nahrávek.

Tab. 11.12: 9. Dvořákova symfonie, selekce SFFS, klasifikace podle orchestrů, celé nahrávky

n	ACC	SEN	SPE	PPV	F ₁	MCC
1	70,46 %	74,67 %	69,33 %	39,29 %	0,51	0,37
2	77,68 %	68,50 %	80,13 %	48,68 %	0,56	0,44
3	81,16 %	67,50 %	84,80 %	55,25 %	0,60	0,49
4	81,54 %	69,67 %	84,71 %	55,83 %	0,62	0,51

n – počet parametrů, ACC – přesnost klasifikace, SEN – senzitivita, SPE – specifita, PPV – positive predictive value, F₁ Score, MCC – Matthewsův korelační koeficient

Podobně jako v případě druhého notového úseku hodnoty přesnosti klasifikace přesahující 80% bylo dosaženo zejména vlivem nevybalancované databáze, tudíž relevance této části analýzy není nikterak významná.

11.5 Zhodnocení selekce parametrů pomocí SFFS

Již při individuální analýze parametrů byly zaznamenány významné výsledky týkající se klasifikace nahrávek dvou smyčcových kvartetů stejného autora. Při selekci parametrů došlo ke zlepšení přesnosti klasifikace u obou kvartetů nad hranici 90% . Nutno podotknout, že se tak stalo i při analýze celých nahrávek a kombinaci malého počtu parametrů.

Naproti tomu během klasifikace nahrávek 9. Dvořákovy symfonie již nebylo dosaženo obdobně dobrých výsledků. Výsledky se výrazně nelišily při obou zvolených rozděleních. Nutno však zdůraznit, že v případě rozdělení podle původu dirigentů bylo dosaženo významné hodnoty senzitivity $SEN = 96,50\%$ v rámci klasifikace celých nahrávek při kombinaci 4 parametrů.

Celkově bylo možné pozorovat, že lepší vlastnosti klasifikace se projeví zejména při klasifikaci komorní hudby, v tomto případě smyčcových kvartetů. Klasifikace orchestrálních nahrávek se nedá považovat za tolik úspěšnou, což by mohlo být ovlivněno jednak rozložením samotné databáze, a i samotným charakterem nahrávek. Dále také může být kvalita této klasifikace ovlivněna nejednoznačným rozdělením na kategorie Česko a svět.

12 Závěr

První kapitola byla věnována shrnutí informací o oblasti Music Information Retrieval a bylo provedeno krátké shrnutí historie této vědní oblasti.

V druhé kapitole byly shrnuty poznatky o základních vlastnostech hudebního signálu z hlediska časové a frekvenční oblasti. Byly definovány pojmy, kterými jsou hodnoceny vlastnosti hudebních signálů. Dále byly shrnuty informace o barvě zvuku, dynamice a tempu hudby a hudebních nahrávek. V této kapitole byla také provedena definice některých pojmů používaných v hudební teorii pro popis hudby.

Dále byl v kapitole zpracování hudebního signálu proveden popis kroků, které je nutné provést, než je provedena samotná parametrizace. Tyto kroky jsou shrnuty v podkapitole předzpracování hudebního signálu, jejich provedení je klíčové za účelem správného parametrizování hudebního signálu.

Následovala kapitola s názvem Parametry obsažené v MIR Toolboxu, která obsahuje popis všech vypočítaných parametrů pomocí MIR Toolboxu pro MATLAB. Parametry jsou roztrženy do tří kategorií, konkrétně parametry popisující dynamiku hudebního signálu, parametry popisující barvu zvuku a parametry popisující rytmickou strukturu hudebního signálu.

Kapitola vybrané statistické parametry obsahuje definice běžně používaných statistických parametrů. Jsou představeny vybrané statistické parametry využití v rámci předběžné analýzy.

Další kapitola nese název Synchronizace nahrávek. V této kapitole je provedena definice chromagramu, což je parametr popisující tónový obsah nahrávky v čase. Je definován způsob, jakým je možné využít chromagramy pro synchronizaci nahrávek pomocí dynamického borcení časové osy. V této kapitole byl dále představen nástroj pro synchronizaci nahrávek, který využívá tzv. víceúrovňového DTW. Nástroj pro synchronizaci nahrávek byl implementován v prostředí MATLAB a byl úspěšně použit k výběru stejných notových úseků ze všech interpretací.

Následující část práce byla věnována návrhu vyhodnocovacího systému, pomocí kterého byly nahrávky klasifikovány do dvou kategorií Česko a svět. Zde bylo dále nastíněno rozdělení provedené statistické analýzy na dvě části, konkrétně se jednalo o individuální analýzu parametrů a selekci parametrů pomocí metody SFFS.

V kapitole individuální analýza parametrů byly shrnuty výsledky klasifikace nahrávek za použití vždy pouze jednoho parametru. V každé dílčí části analýzy bylo představeno 5 parametrů, pomocí kterých bylo dosaženo nejlepší kvality klasifikace. Analyzovány byly dva smyčcové kvartety a jedna orchestrální skladba. Je možno tvrdit, že při analýze smyčcových kvartetů bylo dosaženo celkově lepší kvality klasifikace oproti nahrávkám orchestrálním. Jsou-li shrnuty dosažené výsledky, parametry, které popisují spektrální složení a barvu zvuku vykazaly lepší vlastnosti než ostatní

skupiny parametrů. Převážná většina parametrů s významnými klasifikačními vlastnostmi popisuje spektrální složení analyzovaných nahrávek. U obou analyzovaných smyčcových kvartetů bylo i v případě individuální analýzy parametrů dosaženo vysoké přesnosti klasifikace. Konkrétně pomocí parametru mfcc9 (mean) bylo dosaženo přesnosti klasifikace ACC = 84,68 %, dále potom parametr mfcc1(var) vykázal přesnost klasifikace ACC = 91,68 %. Významné výsledky přesnosti klasifikace byly v případě smyčcových kvartetů dále podpořeny i ostatními vypočítanými metrikami. V případě analýzy 9. Dvořákovy symfonie byl při rozdělení podle dirigentů nejvýznamnější parametr mfcc1 (Q3), který dosáhl přesnosti klasifikace ACC = 73,44 %. Při rozdělení podle orchestrů vykázal nejlepší vlastnosti při klasifikaci parametr brightness (99p) s hodnotou přesnosti ACC = 78,67 %.

Po individuální analýze parametrů následovala selekce parametrů pomocí metody SFFS. U všech třech skladeb bylo dosaženo větší přesnosti klasifikace, než tomu bylo v případě individuální analýzy parametrů. Opět i v tomto případě bylo dosaženo výrazně lepších výsledků při analýze obou smyčcových kvartetů. Kvalita klasifikace byla podpořena i vysokou hodnotou senzitivity. Hodnota senzitivity byla směrodatným ukazatelem zejména kvůli tomu, že byl získán přehled o tom, kolik českých nahrávek bylo správně zařazeno. I při kombinaci nízkého počtu parametrů bylo dosaženo významných výsledků, a to u obou smyčcových kvartetů. Při klasifikaci nahrávek orchestrálních bylo opět dosaženo poněkud horší kvality klasifikace. Při analýze 1. Janáčkova kvartetu bylo dosaženo přesnosti klasifikace ACC = 91,11 % za kombinace 4 parametrů. Již při kombinaci 2 parametrů bylo u 2. Janáčkova kvartetu dosaženo přesnosti klasifikace ACC = 94,49 %. Významný výsledek klasifikace při analýze 9. Dvořákovy symfonie bylo dosažení hodnoty senzitivity SEN = 96,50 %, a to s rozdělením databáze podle dirigentů. Nejlepší výsledek při rozdělení podle orchestrů byla hodnota přesnosti klasifikace ACC = 86,42 %.

Jsou-li shrnuty všechny výsledky analýzy, je možno dojít k několika závěrům. Kvalita klasifikace je závislá na charakteru samotných hudebních nahrávek a bylo dosaženo větší přesnosti klasifikace v oblasti komorní hudby. Tato skutečnost je také ovlivněna ne zcela jasnou kategorizací v případě orchestrálních nahrávek. Je-li brán v potaz původ orchestru a původ dirigenta, objevuje se víc proměnných, které mohou ovlivnit celkové výsledky klasifikace. Rozdělení orchestrálních nahrávek tedy není ani zdaleka tak jednoznačné, jako tomu bylo v případě nahrávek smyčcových kvartetů.

Nakonec je nutno zmínit neideální rozdělení všech databází, které bylo ovlivněno zejména charakterem vybraných klasifikačních skupin. Takto natrénované klasifikační modely nejsou reálně použitelné ke klasifikaci dalších nahrávek, zejména díky velmi omezenému množství dostupných interpretací. Výsledky by měly sloužit jako ilustrace možnosti sledovat míru odlišnosti českých nahrávek oproti nahrávkám svě-

tovým.

I přes veškeré negativní dopady lze některé z dosažených výsledků považovat za směrodatné díky tomu, že byla vyloučena náhodná klasifikace dvou rozdílných komorních skladeb jednoho autora.

Pokud by mělo být dosaženo výsledků o větší vypovídající hodnotě, databáze by musely být výrazně početnější, s lepším rozdělením klasifikačních skupin.

Literatura

- [1] ARLOT, S., ALAIN C. *A survey of cross-validation procedures for model selection*. Statistics Surveys [online]. 2010, 4, 40-79 [cit. 2019-05-12]. DOI: 10.1214/09-SS054. ISSN 1935-7516. Dostupné z: <http://projecteuclid.org/euclid.ssu/1268143839>
- [2] BOUGHORBEL, S. *Optimal classifier for imbalanced data using Matthews Correlation Coefficient metric*. PLOS ONE [online]. 2017, 12(6) [cit. 2019-05-09]. DOI: 10.1371/journal.pone.0177678. ISSN 1932-6203. Dostupné z: <https://dx.plos.org/10.1371/journal.pone.0177678>
- [3] GALÁŽ Z., MEKYSKA J., SMÉKAL Z. *Classification analysis tool* [online.] [cit. 2018-12-06] Dostupné z URL http://splab.cz/wp-content/uploads/2016/08/Classification_analysis_tool.zip.
- [4] GALÁŽ, Z. *Sequential floating feature selection tool* [online.] [cit. 2019-04-05] Dostupné z URL http://splab.cz/wp-content/uploads/2015/06/sffs_analysis.zip.
- [5] GOUYON, F. *On The Use of Zero-crossing Rate For An Application Of Classification of Percussive Sounds In: Proceedings of the COST G-6 conference on the digital audio effects (DAFX-00)* [online]. 2000. [cit. 2018-11-09]. Dostupné z URL: <http://www.mtg.upf.edu/files/publications/dafx00-gouyon.pdf>
- [6] CHANG-HSING, L. *Automatic Music Genre Classification Based on Modulation Spectral Analysis of Spectral and Cepstral Features* [online]. 2009 [cit. 2018-11-09]. Dostupné z URL: <http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=4895319>.
- [7] GIANNAKOPOULOS, T. *Introduction to audio analysis: a MATLAB approach*. ISBN 9780080993881.
- [8] KNEES, P., SCHEDL M. *Music similarity and retrieval: an introduction to audio and web-based strategies*. 1. New York, NY: Springer Berlin Heidelberg, 2016. ISBN 9783662497203.
- [9] LARTILLOT, O. *Estimating Tempo and Metrical Features by Tracking the Whole Metrical Hierarchy*. [online]. 2013 [cit. 2018-11-09]. Dostupné z URL: <https://jyx.jyu.fi/bitstream/handle/123456789/41612/Olivier%20Lartillot%20-%20Estimating%20Tempo%20And%20Metrical%20Features%20By%20Tracking%20The%20Whole%20Metrical%20Hierarchy.pdf?sequence=1&isAllowed=y>

- [10] LARTILLOT, O. *MIRtoolbox 1.7.1 User's Manual* [online]. 2018 [cit. 2018-11-09]. Dostupné z URL: <https://www.jyu.fi/hytk/fi/laitokset/mutku/en/research/materials/mirtoolbox/manual1-7-1.pdf>
- [11] LARTILLOT, O. *Multi-feature modeling of pulse clarity: design, validation and optimization* [online]. 2008 [cit. 2018-11-09]. Dostupné z URL: <https://pdfs.semanticscholar.org/c4b6/25978b8e4bc26019e2ec57b867b563d73f5e.pdf>
- [12] MATTHEWS, B.W. *Comparison of the predicted and observed secondary structure of T₄ phage lysozyme. Biochimica et Biophysica Acta (BBA) - Protein Structure* [online]. 1975, 405(2), 442-451 [cit. 2019-05-09]. DOI: 10.1016/0005-2795(75)90109-9. ISSN 00052795. Dostupné z: <https://linkinghub.elsevier.com/retrieve/pii/0005279575901099>
- [13] MOLINA, L.C. *Feature selection algorithms: a survey and experimental evaluation. In: 2002 IEEE International Conference on Data Mining, 2002. Proceedings* [online]. IEEE Comput. Soc, 2002, s. 306-313 [cit. 2019-05-02]. DOI: 10.1109/ICDM.2002.1183917. ISBN 0-7695-1754-4. Dostupné z: <http://ieeexplore.ieee.org/document/1183917/>
- [14] MÜLLER, M. *An Efficient Multiscale Approach to Audio Synchronization* [online]. 2006. [cit. 2018-12-08] Dostupné z URL: http://ismir2006.ismir.net/PAPERS/ISMIR0615_Paper.pdf.
- [15] MÜLLER, M. *Chroma Toolbox: MATLAB Implementations for Extracting Variants of Chroma-based audio features* [online]. 2011. [cit. 2018-12-08] Dostupné z URL: <https://pdfs.semanticscholar.org/31cf/b91158031038ad16218997cc9724a9ab5c66.pdf>.
- [16] NALINI, N.J. *Music emotion recognition: The combined evidence of MFCC and residual phase 2015* [online]. [cit. 2018-11-09]. Dostupné z URL: http://ac.els-cdn.com/S1110866515000419/1-s2.0-S1110866515000419-main.pdf?_tid=57e426b8-317f-11e7-9b7a-00000aacb360&acdnat=1493981046_c460ac6ad452557d692ed704c525f74e.
- [17] NEUŽIL, J., BENKO, M. *Skripta základů hudební teorie*. Praha, 2010. Dostupné také z: http://download.kjj.cz/pub/vyuka/2010/kl/skripta_kl.pdf. Skriptum. Konzervatoř Jaroslava Ježka.
- [18] NYMOEN, K. *Validating attack phase descriptors obtained by the Timbre Toolbox and MIR Toolbox* [online]. 2017 [cit. 2018-11-09]. Dostupné z URL: <https://www.duo.uio.no/bitstream/handle/10852/56057/AttackTime-SMC2017.pdf?sequence=1>

- [19] PAMPALK, E. *Content-based Organization and Visualization of Music Archives*, In: *Proceedings of ACM Multimedia* Juan-les-Pins, France. 2002. [cit. 2018-11-09]. Dostupné z URL: http://www.ofai.at/~elias.pampalk/publications/pam_mm02.pdf
- [20] PEETERS, G. *The Timbre Toolbox: Extracting audio descriptors from musical signals* 2011. [online]. [cit. 2018-11-10]. Dostupné z URL: https://www.mcgill.ca/mpcl/files/mpcl/peeters_2011_jasa.pdf
- [21] REUTER, C. *Karl Erich Schumann's principles of timbre as a helpful tool in stream segregation research*. Springer, Berlin, Heidelberg, 1996. ISBN 978-3-540-63526-0
- [22] RUSSELL, S. *Artificial intelligence: a modern approach*. 3rd ed. Upper Saddle River: Prentice Hall, c2010. ISBN 0-13-604259-7.
- [23] SALVADOR, S. *FastDTW: Toward Accurate Dynamic Time Warping in Linear Time and Space* 2004. [online]. [cit. 2018-12-06] Dostupné z URL: <https://pdfs.semanticscholar.org/05a2/0cde15e172fc82f32774dd0cf4fe5827cad2.pdf>
- [24] SASAKI, Y. *The truth of the F-measure* [online]. 2007, 2007, s.1–5 [cit. 2019-05-01]. Dostupné z: <https://www.toyota-ti.ac.jp/Lab/Denshi/COIN/people/yutaka.sasaki/F-measure-YS-260ct07.pdf>
- [25] SHALEV-SHWARTZ, S. *Understanding machine learning: from theory to algorithms*. New York, NY, USA: Cambridge University Press, 2014. ISBN 978-1-107-05713-5.
- [26] SYROVÝ, V. *Hudební akustika* Akademie múzických umění v Praze, 2013 ISBN 978-80-7331-297-8
- [27] TZANETAKIS, G. *Automatic Musical Genre Classification Of Audio Signals* [online]. s. 1–4. [cit. 2018-11-09]. Dostupné z URL: <http://ismir2001.ismir.net/pdf/tzanetakis.pdf>

Seznam symbolů, veličin a zkratek

ACC	Accuracy – přesnost klasifikace
BPM	Beats Per Minute – počet úderů za minutu
CENS	Chroma Energy Normalized Statistics
DTW	Dynamic Time Warping – Dynamické borcení časové osy
FFT	Fast Fourier Transform – Rychlá Fourierova transformace
MCC	Matthews Correlation Coefficient – Matthewsův korelační koeficient
MFCC	Mel Frequency Cepstral Coefficient - Melovský cepstrální koeficient
MIR	Music Information Retrieval
PPV	Positive Predictive Value
SEN	Sensitivita
SFFS	Sequential Floating Forward Search
SPE	Specifita

Seznam příloh

A Přiložené soubory	63
B Seznam vypočítaných parametrů	64

A Příložené soubory

Příložený soubor s názvem „dp-priloha.zip“ obsahuje 2 adresáře a soubor s hlavním dokumentem práce ve formátu .pdf. První adresář s názvem „MATLAB“ obsahuje všechny funkce použité při výpočtu parametrů z audio nahrávek. Adresář „MATLAB“ dále obsahuje podadresář s názvem „DTW“, ve kterém jsou obsaženy naprogramované funkce použité pro synchronizaci nahrávek. Podadresář „compFunctions“ v adresáři „MATLAB“ obsahuje veškeré funkce, které byly vytvořeny pro výpočet parametrů z nahrávek. Druhý adresář s názvem „MATLAB“ obsahuje výsledky individuální analýzy parametrů a selekce parametrů ve formě souborů typu .xlsx a .csv.

Obsah archivu „dp-priloha.zip“:

- dp-miklanek-stepan.pdf
- MATLAB
 - DTW
 - compFunctions
- vysledky
 - individualni-analyza
 - selekce-parametru

B Seznam vypočítaných parametrů

č.	název parametru	význam parametru
1	attackLeap (mean)	střední hodnota skoku náběhu
2	attackLeap (median)	medián skoku náběhu
3	attackLeap (std)	směrodatná odchylka skoku náběhu
4	attackLeap (var)	rozptyl skoku náběhu
5	attackLeap (Q1)	1. kvartil skoku náběhu
6	attackLeap (Q3)	2. kvartil skoku náběhu
7	attackLeap (IQR)	mezikvart. rozsah skoku náběhu
8	attackLeap (1p)	1. percentil skoku náběhu
9	attackLeap (99p)	99. percentil skoku náběhu
10	attackLeap (ir)	meziperc. rozsah skoku náběhu
11	attackSlope (mean)	střední hodnota sklonu náběhu
12	attackSlope (median)	medián sklonu náběhu
13	attackSlope (std)	směrodat. odchylka sklonu náběhu
14	attackSlope (var)	rozptyl sklonu náběhu
15	attackSlope (Q1)	1. kvartil sklonu náběhu
16	attackSlope (Q3)	2. kvartil sklonu náběhu
17	attackSlope (IQR)	mezikvart. rozsah sklonu náběhu
18	attackSlope (1p)	1. percentil sklonu náběhu
19	attackSlope (99p)	99. percentil sklonu náběhu
20	attackSlope (ir)	meziperc. rozsah sklonu náběhu
21	attackTime (mean)	střední hodnota doby náběhu
22	attackTime (median)	medián doby náběhu
23	attackTime (std)	směrodatná odchylka doby náběhu
24	attackTime (var)	rozptyl doby náběhu
25	attackTime (Q1)	1. kvartil doby náběhu
26	attackTime (Q3)	2. kvartil doby náběhu
27	attackTime (IQR)	mezikvartilový rozsah doby náběhu
28	attackTime (1p)	1. percentil doby náběhu
29	attackTime (99p)	99. percentil doby náběhu
30	attackTime (ir)	meziperc. rozsah doby náběhu
31	brightness (mean)	střední hodnota světlosti
32	brightness (median)	medián světlosti
33	brightness (std)	směrodatná odchylka světlosti
34	brightness (var)	rozptyl světlosti
35	brightness (Q1)	1. kvartil světlosti

36	brightness (Q3)	2. kvartil světlosti
37	brightness (IQR)	mezikvartilový rozsah světlosti
38	brightness (1p)	1. percentil světlosti
39	brightness (99p)	99. percentil světlosti
40	brightness (ir)	mezipercentilový rozsah světlosti
41	decayLeap (mean)	střední hodnota skoku poklesu
42	decayLeap (median)	medián skoku poklesu
43	decayLeap (std)	směrodatná odchylka skoku poklesu
44	decayLeap (var)	rozptyl skoku poklesu
45	decayLeap (Q1)	1. kvartil skoku poklesu
46	decayLeap (Q3)	2. kvartil skoku poklesu
47	decayLeap (IQR)	mezikvart. rozsah skoku poklesu
48	decayLeap (1p)	1. percentil skoku poklesu
49	decayLeap (99p)	99. percentil skoku poklesu
50	decayLeap (ir)	meziperc. rozsah skoku poklesu
51	decaySlope (mean)	střední hodnota sklonu poklesu
52	decaySlope (median)	medián sklonu poklesu
53	decaySlope (std)	směrodat. odchylka sklonu poklesu
54	decaySlope (var)	rozptyl sklonu poklesu
55	decaySlope (Q1)	1. kvartil sklonu poklesu
56	decaySlope (Q3)	2. kvartil sklonu poklesu
57	decaySlope (IQR)	mezikvart. rozsah sklonu poklesu
58	decaySlope (1p)	1. percentil sklonu poklesu
59	decaySlope (99p)	99. percentil sklonu poklesu
60	decaySlope (ir)	meziperc. rozsah sklonu poklesu
61	decayTime (mean)	střední hodnota doby poklesu
62	decayTime (median)	medián doby poklesu
63	decayTime (std)	směrodatná odchylka doby poklesu
64	decayTime (var)	rozptyl doby poklesu
65	decayTime (Q1)	1. kvartil doby poklesu
66	decayTime (Q3)	2. kvartil doby poklesu
67	decayTime (IQR)	mezikvart.rozsah doby poklesu
68	decayTime (1p)	1. percentil doby poklesu
69	decayTime (99p)	99. percentil doby poklesu
70	decayTime (ir)	meziperc. rozsah doby poklesu
71	duration (mean)	střední hodnota doby trvání
72	duration (median)	medián doby trvání
73	duration (std)	směrodat. odchylka doby trvání
74	duration (var)	rozptyl doby trvání

75	duration (Q1)	1. kvartil doby trvání
76	duration (Q3)	2. kvartil doby trvání
77	duration (IQR)	mezikvart. rozsah doby trvání
78	duration (1p)	1. percentil doby trvání
79	duration (99p)	99. percentil doby trvání
80	duration (ir)	meziperc. rozsah doby trvání
81	eventDensity (mean)	střední hodnota hustoty událostí
82	eventDensity (median)	medián hustoty událostí
83	eventDensity (std)	směrodat. odchylka hustoty událostí
84	eventDensity (var)	rozptyl hustoty událostí
85	eventDensity (Q1)	1. kvartil hustoty událostí
86	eventDensity (Q3)	2. kvartil hustoty událostí
87	eventDensity (IQR)	mezikvart. rozsah hustoty událostí
88	eventDensity (1p)	1. percentil hustoty událostí
89	eventDensity (99p)	99. percentil hustoty událostí
90	eventDensity (ir)	meziperc. rozsah hustoty událostí
91	mfcc1 (mean)	střední hodnota 1. MFCC
92	mfcc1 (median)	medián 1. MFCC
93	mfcc1 (std)	směrodatná odchylka 1. MFCC
94	mfcc1 (var)	rozptyl 1. MFCC
95	mfcc1 (Q1)	1. kvartil 1. MFCC
96	mfcc1 (Q3)	2. kvartil 1. MFCC
97	mfcc1 (IQR)	mezikvartilový rozsah 1. MFCC
98	mfcc1 (1p)	1. percentil 1. MFCC
99	mfcc1 (99p)	99. percentil 1. MFCC
100	mfcc1 (ir)	mezipercentilový rozsah 1. MFCC
101	mfcc2 (mean)	střední hodnota 2. MFCC
102	mfcc2 (median)	medián 2. MFCC
103	mfcc2 (std)	směrodatná odchylka 2. MFCC
104	mfcc2 (var)	rozptyl 2. MFCC
105	mfcc2 (Q1)	1. kvartil 2. MFCC
106	mfcc2 (Q3)	2. kvartil 2. MFCC
107	mfcc2 (IQR)	mezikvartilový rozsah 2. MFCC
108	mfcc2 (1p)	1. percentil 2. MFCC
109	mfcc2 (99p)	99. percentil 2. MFCC
110	mfcc2 (ir)	mezipercentilový rozsah 2. MFCC
111	mfcc3 (mean)	střední hodnota 3. MFCC
112	mfcc3 (median)	medián 3. MFCC
113	mfcc3 (std)	směrodatná odchylka 3. MFCC

114	mfcc3 (var)	rozptyl 3. MFCC
115	mfcc3 (Q1)	1. kvartil 3. MFCC
116	mfcc3 (Q3)	2. kvartil 3. MFCC
117	mfcc3 (IQR)	mezikvartilový rozsah 3. MFCC
118	mfcc3 (1p)	1. percentil 3. MFCC
119	mfcc3 (99p)	99. percentil 3. MFCC
120	mfcc3 (ir)	mezipercentilový rozsah 3. MFCC
121	mfcc4 (mean)	střední hodnota 4. MFCC
122	mfcc4 (median)	medián 4. MFCC
123	mfcc4 (std)	směrodatná odchylka 4. MFCC
124	mfcc4 (var)	rozptyl 4. MFCC
125	mfcc4 (Q1)	1. kvartil 4. MFCC
126	mfcc4 (Q3)	2. kvartil 4. MFCC
127	mfcc4 (IQR)	mezikvartilový rozsah 4. MFCC
128	mfcc4 (1p)	1. percentil 4. MFCC
129	mfcc4 (99p)	99. percentil 4. MFCC
130	mfcc4 (ir)	mezipercentilový rozsah 4. MFCC
131	mfcc5 (mean)	střední hodnota 5. MFCC
132	mfcc5 (median)	medián 5. MFCC
133	mfcc5 (std)	směrodatná odchylka 5. MFCC
134	mfcc5 (var)	rozptyl 5. MFCC
135	mfcc5 (Q1)	1. kvartil 5. MFCC
136	mfcc5 (Q3)	2. kvartil 5. MFCC
137	mfcc5 (IQR)	mezikvartilový rozsah 5. MFCC
138	mfcc5 (1p)	1. percentil 5. MFCC
139	mfcc5 (99p)	99. percentil 5. MFCC
140	mfcc5 (ir)	mezipercentilový rozsah 5. MFCC
141	mfcc6 (mean)	střední hodnota 6. MFCC
142	mfcc6 (median)	medián 6. MFCC
143	mfcc6 (std)	směrodatná odchylka 6. MFCC
144	mfcc6 (var)	rozptyl 6. MFCC
145	mfcc6 (Q1)	1. kvartil 6. MFCC
146	mfcc6 (Q3)	2. kvartil 6. MFCC
147	mfcc6 (IQR)	mezikvartilový rozsah 6. MFCC
148	mfcc6 (1p)	1. percentil 6. MFCC
149	mfcc6 (99p)	99. percentil 6. MFCC
150	mfcc6 (ir)	mezipercentilový rozsah 6. MFCC
151	mfcc7 (mean)	střední hodnota 7. MFCC
152	mfcc7 (median)	medián 7. MFCC

153	mfcc7 (std)	směrodatná odchylka 7. MFCC
154	mfcc7 (var)	rozptyl 7. MFCC
155	mfcc7 (Q1)	1. kvartil 7. MFCC
156	mfcc7 (Q3)	2. kvartil 7. MFCC
157	mfcc7 (IQR)	mezikvartilový rozsah 7. MFCC
158	mfcc7 (1p)	1. percentil 7. MFCC
159	mfcc7 (99p)	99. percentil 7. MFCC
160	mfcc7 (ir)	mezipercentilový rozsah 7. MFCC
161	mfcc8 (mean)	střední hodnota 8. MFCC
162	mfcc8 (median)	medián 8. MFCC
163	mfcc8 (std)	směrodatná odchylka 8. MFCC
164	mfcc8 (var)	rozptyl 8. MFCC
165	mfcc8 (Q1)	1. kvartil 8. MFCC
166	mfcc8 (Q3)	2. kvartil 8. MFCC
167	mfcc8 (IQR)	mezikvartilový rozsah 8. MFCC
168	mfcc8 (1p)	1. percentil 8. MFCC
169	mfcc8 (99p)	99. percentil 8. MFCC
170	mfcc8 (ir)	mezipercentilový rozsah 8. MFCC
171	mfcc9 (mean)	střední hodnota 9. MFCC
172	mfcc9 (median)	medián 9. MFCC
173	mfcc9 (std)	směrodatná odchylka 9. MFCC
174	mfcc9 (var)	rozptyl 9. MFCC
175	mfcc9 (Q1)	1. kvartil 9. MFCC
176	mfcc9 (Q3)	2. kvartil 9. MFCC
177	mfcc9 (IQR)	mezikvartilový rozsah 9. MFCC
178	mfcc9 (1p)	1. percentil 9. MFCC
179	mfcc9 (99p)	99. percentil 9. MFCC
180	mfcc9 (ir)	mezipercentilový rozsah 9. MFCC
181	mfcc10 (mean)	střední hodnota 10. MFCC
182	mfcc10 (median)	medián 10. MFCC
183	mfcc10 (std)	směrodatná odchylka 10. MFCC
184	mfcc10 (var)	rozptyl 10. MFCC
185	mfcc10 (Q1)	1. kvartil 10. MFCC
186	mfcc10 (Q3)	2. kvartil 10. MFCC
187	mfcc10 (IQR)	mezikvartilový rozsah 10. MFCC
188	mfcc10 (1p)	1. percentil 10. MFCC
189	mfcc10 (99p)	99. percentil 10. MFCC
190	mfcc10 (ir)	mezipercentilový rozsah 10. MFCC
191	mfcc11 (mean)	střední hodnota 11. MFCC

192	mfcc11 (median)	medián 11. MFCC
193	mfcc11 (std)	směrodatná odchylka 11. MFCC
194	mfcc11 (var)	rozptyl 11. MFCC
195	mfcc11 (Q1)	1. kvartil 11. MFCC
196	mfcc11 (Q3)	2. kvartil 11. MFCC
197	mfcc11 (IQR)	mezikvartilový rozsah 11. MFCC
198	mfcc11 (1p)	1. percentil 11. MFCC
199	mfcc11 (99p)	99. percentil 11. MFCC
200	mfcc11 (ir)	mezipercentilový rozsah 11. MFCC
201	mfcc12 (mean)	střední hodnota 12. MFCC
202	mfcc12 (median)	medián 12. MFCC
203	mfcc12 (std)	směrodatná odchylka 12. MFCC
204	mfcc12 (var)	rozptyl 12. MFCC
205	mfcc12 (Q1)	1. kvartil 12. MFCC
206	mfcc12 (Q3)	2. kvartil 12. MFCC
207	mfcc12 (IQR)	mezikvartilový rozsah 12. MFCC
208	mfcc12 (1p)	1. percentil 12. MFCC
209	mfcc12 (99p)	99. percentil 12. MFCC
210	mfcc12 (ir)	mezipercentilový rozsah 12. MFCC
211	mfcc13 (mean)	střední hodnota 13. MFCC
212	mfcc13 (median)	medián 13. MFCC
213	mfcc13 (std)	směrodatná odchylka 13. MFCC
214	mfcc13 (var)	rozptyl 13. MFCC
215	mfcc13 (Q1)	1. kvartil 13. MFCC
216	mfcc13 (Q3)	2. kvartil 13. MFCC
217	mfcc13 (IQR)	mezikvartilový rozsah 13. MFCC
218	mfcc13 (1p)	1. percentil 13. MFCC
219	mfcc13 (99p)	99. percentil 13. MFCC
220	mfcc13 (ir)	mezipercentilový rozsah 13. MFCC
221	rms (mean)	střední hodnota efektivní hodnoty
222	rms (median)	medián efektivní hodnoty
223	rms (std)	směr. odchylka efektivní hodnoty
224	rms (var)	rozptyl efektivní hodnoty
225	rms (Q1)	1. kvartil efektivní hodnoty
226	rms (Q3)	2. kvartil efektivní hodnoty
227	rms (IQR)	mezikvart.rozsah efektivní hodnoty
228	rms (1p)	1. percentil efektivní hodnoty
229	rms (99p)	99. percentil efektivní hodnoty
230	rms (ir)	meziperc. rozsah efektivní hodnoty

231	rollOff (mean)	stř. hodnota poklesu spektr. energie
232	rollOff (median)	medián poklesu spektr. energie
233	rollOff (std)	směr. odchylka poklesu spektr. energie
234	rollOff (var)	rozptyl poklesu spektr. energie
235	rollOff (Q1)	1. kvartil poklesu spektr. energie
236	rollOff (Q3)	2. kvartil poklesu spektr. energie
237	rollOff (IQR)	mezikvart. rozsah pokl. spektr. energie
238	rollOff (1p)	1. perc. poklesu spektr. energie
239	rollOff (99p)	99. perc. poklesu spektr. energie
240	rollOff (ir)	meziperc. rozsah pokl. spektr. energie
241	roughness (mean)	střední hodnota drsnosti zvuku
242	roughness (median)	medián drsnosti zvuku
243	roughness (std)	směrodatná odchylka drsnosti zvuku
244	roughness (var)	rozptyl drsnosti zvuku
245	roughness (Q1)	1. kvartil drsnosti zvuku
246	roughness (Q3)	2. kvartil drsnosti zvuku
247	roughness (IQR)	mezikvart. rozsah drsnosti zvuku
248	roughness (1p)	1. percentil drsnosti zvuku
249	roughness (99p)	99. percentil drsnosti zvuku
250	roughness (ir)	meziperc. rozsah drsnosti zvuku
251	spectralCentroid (mean)	střední hodnota spektr. centroidu
252	spectralCentroid (median)	medián spektrálního centroidu
253	spectralCentroid (std)	směrodat.odchylka spektr. centroidu
254	spectralCentroid (var)	rozptyl spektrálního centroidu
255	spectralCentroid (Q1)	1. kvartil spektrálního centroidu
256	spectralCentroid (Q3)	2. kvartil spektrálního centroidu
257	spectralCentroid (IQR)	mezikvart. rozsah spektr. centroidu
258	spectralCentroid (1p)	1. percentil spektrálního centroidu
259	spectralCentroid (99p)	99. percentil spektrálního centroidu
260	spectralCentroid (ir)	meziperc. rozsah spektr. centroidu
261	spectralFlatness (mean)	střední hodnota plochosti spektra
262	spectralFlatness (median)	medián plochosti spektra
263	spectralFlatness (std)	směr. odchylka plochosti spektra
264	spectralFlatness (var)	rozptyl plochosti spektra
265	spectralFlatness (Q1)	1. kvartil plochosti spektra
266	spectralFlatness (Q3)	2. kvartil plochosti spektra
267	spectralFlatness (IQR)	mezikvar. rozsah plochosti spektra
268	spectralFlatness (1p)	1. percentil plochosti spektra
269	spectralFlatness (99p)	99. percentil plochosti spektra

270	spectralFlatness (ir)	meziperc. rozsah plochosti spektra
271	spectralKurtosis (mean)	střední hodnota špičatosti spektra
272	spectralKurtosis (median)	medián špičatosti spektra
273	spectralKurtosis (std)	směr. odchylka špičatosti spektra
274	spectralKurtosis (var)	rozptyl špičatosti spektra
275	spectralKurtosis (Q1)	1. kvartil špičatosti spektra
276	spectralKurtosis (Q3)	2. kvartil špičatosti spektra
277	spectralKurtosis (IQR)	mezikvart. rozsah špičatosti spektra
278	spectralKurtosis (1p)	1. percentil špičatosti spektra
279	spectralKurtosis (99p)	99. percentil špičatosti spektra
280	spectralKurtosis (ir)	meziperc. rozsah špičatosti spektra
281	spectralSkewness (mean)	střední hodnota šikmosti spektra
282	spectralSkewness (median)	medián šikmosti spektra
283	spectralSkewness (std)	směrodatná odchylka šikmosti spektra
284	spectralSkewness (var)	rozptyl šikmosti spektra
285	spectralSkewness (Q1)	1. kvartil šikmosti spektra
286	spectralSkewness (Q3)	2. kvartil šikmosti spektra
287	spectralSkewness (IQR)	mezikvart. rozsah šikmosti spektra
288	spectralSkewness (1p)	1. percentil šikmosti spektra
289	spectralSkewness (99p)	99. percentil šikmosti spektra
290	spectralSkewness (ir)	meziperc. rozsah šikmosti spektra
291	spectralSpread (mean)	střední hodnota šířky spektra
292	spectralSpread (median)	medián šířky spektra
293	spectralSpread (std)	směrodatná odchylka šířky spektra
294	spectralSpread (var)	rozptyl šířky spektra
295	spectralSpread (Q1)	1. kvartil šířky spektra
296	spectralSpread (Q3)	2. kvartil šířky spektra
297	spectralSpread (IQR)	mezikvartilový rozsah šířky spektra
298	spectralSpread (1p)	1. percentil šířky spektra
299	spectralSpread (99p)	99. percentil šířky spektra
300	spectralSpread (ir)	meziperc. rozsah tempa
301	tempo (mean)	střední hodnota tempa
302	tempo (median)	medián tempa
303	tempo (std)	směrodatná odchylka tempa
304	tempo (var)	rozptyl tempa
305	tempo (Q1)	1. kvartil tempa
306	tempo (Q3)	2. kvartil tempa
307	tempo (IQR)	mezikvartilový rozsah tempa
308	tempo (1p)	1. percentil tempa

309	tempo (99p)	99. percentil tempa
310	tempo (ir)	meziperc. rozsah tempa
311	zcr (mean)	střední hodnota počtu průch. nul. úr.
312	zcr (median)	medián počtu průch. nulovou úr.
313	zcr (std)	směrodat. odchylka počtu průch. nul. úr.
314	zcr (var)	rozptyl počtu průchodů nul. úr.
315	zcr (Q1)	1. kvartil počtu průch. nul. úr.
316	zcr (Q3)	2. kvartil počtu průch. nul. úr.
317	zcr (IQR)	mezikvart. rozsah počtu průch. nul. úr.
318	zcr (1p)	1. percentil počtu průch. nul. úr.
319	zcr (99p)	99. percentil počtu průch. nul. úr.
320	zcr (ir)	meziperc. rozsah počtu průch. nul. úr.
