

Metodika analýzy hudebně-interpretačního výkonu s podporou
The MemoVision Software

Brno 2023

Certifikovaná metodika k projektu „Paměť zvuku – evoluční principy interpretační tradice české hudby na příkladu děl Antonína Dvořáka a Bedřicha Smetany“

Matěj Ištvanek¹, Klára Hedvika Mühlová², Lubomír Spurný², Martin Mejzr³

- 1) Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií,
Ústav telekomunikací
Technická 3082/12, Královo Pole, 616 00 Brno
- 2) Masarykova univerzita, Filozofická fakulta, Ústav hudební vědy
Arna Nováka 1, 602 00 Brno-střed
- 3) Národní muzeum, České muzeum hudby
Karmelitská 388/2, 118 00 Praha 1 – Malá Strana

T A
Č R

Metodika analýzy hudebně-interpretačního výkonu s podporou The MemoVision Software byla vytvořena se státní podporou Technologické agentury ČR v rámci Programu Éta (projekt TL05000527 Paměť zvuku: evoluční principy interpretační tradice české hudby na příkladu děl Antonína Dvořáka a Bedřicha Smetany).

Obsah

Seznam obrázků	3
Seznam zkratek	5
Předmluva	6
1 Úvod	7
1.1 Cíle a kontext uplatnění metodiky	7
1.2 Přínos předkládané metodiky	10
1.3 Komu je metodika určena	10
1.3.1 Muzikologové a hudební teoretici	11
1.3.2 Hudební interpreti	11
1.3.3 Hudební pedagogové	11
1.3.4 Audio inženýři	12
1.3.5 Pracovníci paměťových, informačních a sdělovacích institucí ve veřejném a neziskovém sektoru	12
1.3.6 Pracovníci v soukromém sektoru	12
2 Hudebně-interpretační analýza	13
2.1 Definice hudebně-interpretační analýzy pro potřeby metodiky	13
2.2 Postupy hudebně-interpretační analýzy	14
2.2.1 Analýza bez podpory počítače	15
2.2.2 Analýza s podporou počítače	17
2.3 Propojení technických a hudebních cílů	18
3 Funkce softwaru pro hudební analýzu	19
3.1 Parametrisace	20
3.1.1 Dynamika	21
3.1.2 Barva a výška	22
3.1.3 Čas	23
3.2 Uživatelský vstup	24
3.2.1 Hudební nahrávky	24
3.2.2 Výběr reference	25
3.2.3 Anotace	25
3.3 Zpracování dat	29
3.3.1 Chroma reprezentace	29

3.3.2	Synchronizace nahrávek	29
3.3.3	Doplňující funkce	30
3.3.4	Metoda maximální relevance	31
3.4	Výstup softwaru	32
3.4.1	Vizualizace dat uživateli	32
4	Uživatelské prostředí softwaru	34
4.1	Moduly	34
4.1.1	Správce nahrávek	34
4.1.2	Výběr sekcí	37
4.1.3	Přehrávač interpretací	39
4.1.4	Vizualizace	40
5	Případová studie – konkrétní případ analýzy	44
5.1	Tvorba databáze	44
5.1.1	Klavírní díla	44
5.1.2	Smyčcové kvartety	46
5.2	Analýza smyčcových kvartetů	46
5.3	Analýza klavírních nahrávek	50
Závěr		53
Seznam literatury a zdrojů		54

Seznam obrázků

3.1	Přehled uživatelského rozhraní a momentálních funkcí softwaru.	20
3.2	Zobrazení rozdílů mezi parametrem RMS a subjektivní hlasitostí podle doporučení EBU R 128.	22
3.3	Dialogové okno s možnostmi vytvoření nové vrstvy pro analýzu hudební nahrávky v Sonic Visualiseru.	26
3.4	Příklad časově-frekvenční reprezentace, konkrétně logaritmicky filtrovaného spektrogramu na základě STFT. Vertikální fialové čáry představují časové pozice taktů dané nahrávky (anotace).	27
3.5	Poslední vertikální fialová čára představuje fyzický konec posledního taktu, který je pro výpočet délky a tampa každého taktu potřeba do anotace přidat. . .	27
3.6	Příklad strojového zápisu anotací: první sloupec označuje časovou značku po sobě jdoucích taktů, druhý sloupec představuje index (číselné označení) daného taktu.	28
3.7	Příklad zobrazení chroma reprezentace pro prvních deset sekund dvou různých interpretací Smetanova <i>Cibuličky</i>	30
3.8	Příklad zobrazení relevance jedné skupiny nahrávek vůči druhé. Světlá barva značí vyšší relevanci, a tedy vyšší rozdíly v tempu jedné skupiny vůči druhé v rámci daných taktů.	32
4.1	Uvítací stránka softwaru s možnostmi registrace uživatele a přihlášení pomocí uživatelského jména a hesla.	35
4.2	Výběr nebo založení a pojmenování nového projektu (session).	35
4.3	První modul <i>Track manager</i> – nahrávání audio nahrávek, anotací, metadat a možnost přiřazení metadat a binárních štítků a spuštění výpočtu parametrů a synchronizace.	36
4.4	Vyskakovací okno po dokončení výpočtu parametrů a synchronizace – systém upozorní, pokud došlo k detekci potenciálně chybné hudební struktury a označí množství úseků, jejich časovou značku a nahrávku, které se tento problém vůči referenci týká.	37
4.5	Přiřazování binárních štítků (labels) každé nahrávce.	38
4.6	Druhý modul softwaru <i>Section selector</i> . Nad linkou označující takty dané kompozice/nahrávky je znázorněna uživatelem upravená metrická struktura pro správný výpočet tempa.	38
4.7	Příklad barevné škály a zobrazení modulu <i>Interpretation player</i> pro několik nahrávek vybrané kompozice. Modré vertikální čáry označují časové pozice taktů v každé nahrávce podle volby rozsahu taktů na barevné škále.	39

4.8	Srovnání mnoha interpretací stejného díla v modulu <i>Interpretation player</i> . Červená vertikální čára označuje aktuální pozici přehrávání (zde začátek prvního taktu). Barevný pruh napravo od názvu každé nahrávky zobrazuje aktuální přiřazení interpretace do jedné z binárních tříd.	40
4.9	Výběr binárních štítků / třídy pro vizualizaci rozdílů pomocí barevné škály.	41
4.10	Výběr vlastních úseků na základě předchozího výběru.	41
4.11	Zobrazení potenciálně chybných synchronizovaných úseků s referenční nahrávkou. Uživatel by měl zkонтrolovat vizuálně, případně poslechem, zda tyto úseky odpovídají stejně hudební struktuře.	41
4.12	Zobrazení prvních deseti nejrelevantnějších taktů podle vybraného parametru a binárních štítků.	42
4.13	Příklad vizualizace RMS, hlasitosti a aktivační funkce detektoru dob pro dvě interpretace a zvolený úsek kompozice.	42
4.14	Vizualizace srovnání hlasitosti a tempa vybraných nahrávek a úseku kompozice pro jednotlivé takty.	43
4.15	Vizualizace grafu regrese (trendu) podle vybraného parametru: zde pro nahrávky před a po roku 2000.	43
4.16	Příklad srovnání aktivačních funkcí automatického detektoru dob na kratším hudebním úseku – Alban Berg Quartett a Pražákovo kvarteto.	43
5.1	Trvání celé věty <i>Smyčcového kvartetu č. 13 G dur</i> , op. 106, 3. věty, pro jednotlivá kvarteta naší databáze včetně binárního rozdělení dle původu (český–nečeský)	48
5.2	Trvání taktů 1–78 <i>Smyčcového kvartetu č. 13 G dur</i> , op. 106, 3. věty, pro jednotlivá kvarteta naší databáze včetně binárního rozdělení dle původu (český–nečeský)	48
5.3	Barevná škála pro vybrané takty Dvořákova <i>Smyčcového kvartetu č. 13 G dur</i> , op. 106, 3. věty; rozdíl mezi třídami český–nečeský.	49
5.4	Průměrné tempo a dynamika pro takty 79–127 a třídy český–nečeský u Dvořákova <i>Smyčcového kvartetu č. 13 G dur</i> , op. 106, 3. věty.	49
5.5	Tempo pro vybraná smyčcové kvarteta a takty 107–123 u Dvořákova <i>Smyčcového kvartetu č. 13 G dur</i> , op. 106, 3. věty.	50
5.6	Dynamika pro vybraná smyčcové kvarteta a takty 107–123 u Dvořákova <i>Smyčcového kvartetu č. 13 G dur</i> , op. 106, 3. věty.	50
5.7	Průměrné tempo a dynamika taktů 68–74 klavírních nahrávek Smetanovy <i>Sousedské</i> ; skupina českých (modře) a zahraničních (červeně) nahrávek.	51
5.8	Strojový notový zápis taktů 68–76 Smetanovy <i>Sousedské</i>	51
5.9	Délka trvání taktů 68–74 klavírních nahrávek Smetanovy <i>Sousedské</i> ; skupina českých (modře) a zahraničních (červeně) nahrávek.	52

Seznam zkratek

AAC	Advanced Audio Coding; jeden ze standardů ztrátové komprese zvuku
dB	decibel; poměrová fyzikální jednotka
DTW	Dynamic Time Warping; dynamické borcení časové osy
EBU	European Broadcasting Union; sdružení rozhlasových a televizních stanic
CHARM	Centre for the History and Analysis of Recorded Music
LUFS	Loudness Unit Full Scale; normalizovaná jednotka hlasitosti
MIDI	Musical Instrument Digital Interface; technický standard pro popis komunikace mezi zařízeními a specificky formát symbolických dat
MIR	Music Information Retrieval; získávání informací z hudebních signálů
MIREX	Music Information Retrieval Evaluation eXchange; komunita a iniciativa pro testování systémů MIR
MIT	Massachusetts Institute of Technology; druh tzv. svobodné licence, která vznikla na Massachusettském technologickém institutu
MPA	Music Performance Analysis; analýza interpretačního výkonu
mp3	MPEG, layer III; ztrátový formát kódování zvukové informace založený na kompresním algoritmu definovaném skupinou MPEG
MR	Maximum Relevance; metoda maximální relevance na základě tzv. společné informace
MrMsDTW	Memory Restricted Multi-scale Dynamic Time Warping; pokročilá varianta algoritmu dynamického borcení časové osy
MusicXML	Musical Extensible Markup Language; druh XML zápisu hudebních dat dle západního notačního systému
NM	Národní muzeum
NM-ČMH	Národní muzeum, České muzeum hudby
OGG	zkratka a přípona pro jeden z tzv. kontejnerů pro data
RMS	Root Mean Square; efektivní hodnota signálu
STFT	Short-Time Fourier Transform; krátkodobá Fourierova transformace
TA ČR	Technologická agentura České republiky
UI	User Interface; uživatelské rozhraní
VUT KInG	Kvalitní interní granty Vysokého učení technického v Brně
WAV	případně WAVE; zkratka a přípona pro zvukový formát Waveform audio file format
XML	Extensible Markup Language; druh specifického jazyka a formát souborů pro ukládání, přenos a rekonstrukci dat

Předmluva

Empiricky založený výzkum hudební interpretace je výrazně se formující a progresivní disciplínou současné muzikologie, v níž se uplatňují interdisciplinární postupy za využití analytických schopností aktuálně dostupných technologií. Téma, které podle tradičních muzikologických systematik středoevropského prostoru spadá pod disciplínu „teorie a dějiny hudební interpretace“, v sobě integruje vysoce komplexní souvislosti, které svým tematickým záběrem pokrývají jak základní ontologické otázky identity hudby jako entity, tak praktické aspekty sahající např. do oblasti hudební pedagogiky a k praktickým otázkám generačního předávání umění hry na hudební nástroj.

Možnosti analytického vytěžení tohoto tématu při zkoumání záznamů a dalších dokladů konkrétního hudebního provedení jsou pak odvozené od šíře výzkumného prismatu a povahy otázek, které si samotný výzkumník klade. Prezentovaná metodika pro výzkumnou práci v oblasti empirické analýzy hudebně-interpretačního výkonu do tohoto kontextu vstupuje formována souvislostmi času a místa, v nichž vznikala a k jejichž tradici částečně náleží. Kromě české či šířeji středoevropské výzkumné tradice reflekтуje metodika i nejnovější vývoj tématu na mezinárodní scéně, který je ve studované problematice aktuálním hybatelem hlavně ve smyslu integrace technologií do hudební analýzy. Metodika je úzce propojena s nově vyvinutým analytickým nástrojem The MemoVision Software, který následuje současný technologický pokrok ve zpracování informací z hudebních nahrávek se zaměřením na cíle hudebně-interpretační analýzy.

I přes úzkou vazbu metodiky na přidružený software je koncipována šířejí než pouhý manuál k softwaru. Metodika svým obsahem uvádí do výzkumné práce se softwarem, a to v kontextu historických a současných trendů hudebně-interpretační analýzy i v souvislosti s aktuálně přijímanými postupy empirické výzkumné práce obecně. V jednotlivých kapitolách jsou nejdříve nastíněny pricnipy hudebně-interpretační analýzy a následně popsány základní myšlenky technických metod zpracování hudebních nahrávek. Důraz je kladen na provázanost a aplikovatelnost moderních metod pomocí nově vytvořeného nástroje pro zpřístupnění, zrychlení a zvětšení repertoáru možností hudební analýzy. V poslední kapitole jsou prezentované možnosti demonstrovány na krátkých případových studiích. Přejeme čtenářům hodně inspirace v souvislosti se studováním pojednávaného tématu.

Autorský kolektiv:

Matěj Ištvanek, Klára Hedvika Mühlová, Lubomír Spurný a Martin Mejzr

1 Úvod

Následující kapitoly pojednávají o cílech a přínosu metodiky a jejím zařazení v komunitě výzkumníků – od muzikologů přes výzkumníky zabývající se hudebně-interpretační analýzou až po audio inženýry.

1.1 Cíle a kontext uplatnění metodiky

Problematika digitalizace hudebních médií je v kontextu současné společnosti velmi důležitým tématem. V této souvislosti je pochopitelný příchod inteligentních nástrojů pro zpracovávání a orientaci v datech, které svým rozsahem překračují individuální poznávací a rozhodovací kapacity člověka i celého týmu lidí. Tato východiska sdílí i v rámci této metodiky prezentovaný, nově vyvinutý analytický nástroj – The MemoVision Software.

Předkládaná metodika si klade za cíl seznámit s – v současném technologickém standardu koncipovanými – možnostmi empirické analýzy hudebně-interpretačního výkonu, jež jsou podpořeny softwarem The MemoVision. Charakter této analýzy lze s pomocí softwaru pojmostit jak v přístupu „bottom-up“ (od dat směrem k teorii), tak přístupem „top-down“ (od teorie směrem k datům). Software je určen jako podpůrný nástroj pro analýzu různých hudebních interpretací a metodika prezentuje možnosti a inspirace související s teoretickými a funkčními otázkami ohledně využití tohoto nástroje. Poslání metodiky je však širší než vymezený cíl. Metodika se zaměřuje i na rozbor problematiky subjektivity běžného intuitivního hodnocení hudební interpretace a zároveň přednáší možnosti hodnocení, které korespondují s možnostmi empirického výzkumu interpretace s využitím podpory počítače. Analýza hudebně-interpretačního výkonu (obor Music Performance Analysis, MPA) má široké možnosti využití v hudební oblasti, a to jak v rámci muzikologie, tak v rámci vlastní hudební nebo pedagogické praxe. V kontextu muzikologického výzkumu se s empiricky pojatou analýzou hudebně-interpretačního výkonu otevírají nové možnosti studia provozovací praxe, a to jak v perspektivě historické muzikologie, tak v systematické muzikologii. Provozovací praxe pak tvoří základní kámen zprostředkování hudby obecně a je rovněž základem pro utváření estetického standardu znějící hudby.

Poznání hudební interpretace souvisí také s poznáním zákonitostí hudby samé. Ta v sobě obsahuje celou řadu charakteristik týkajících se principů fyziologie slyšení, psychologie a dalších dimenzí lidského vnímání i schopnosti myslit a vyjadřovat se. Z tohoto pohledu jsou nové metody studia hudebně-interpretačního výkonu zásadním příspěvkem k poznání procesů lidského sdělování hudebními prostředky. V oblasti historického výzkumu přináší nové technologické možnosti – reprezentované pojednávaným analytickým nástrojem – významné prohloubení a rozšíření výzkumné perspektivy. Empirické studium mechanismů a principů vývoje

interpretační tradice z historického pohledu integruje celou řadu faktorů a dílčích kategorií. Pro tento výzkum jsou tak relevantní např. kontexty vzniku díla a přístupnost autorské představy o provedení díla, specifické estetické normy utvářející provedení, které souvisejí s žánrem a stylem skladby (tentototo aspekt přemostí uje vnitřní charakteristiky stavby kompozice a vnější okolnosti jejího zpřítomnění v interpretaci), dále kontexty pořízení zvukového záznamu provedení či návaznosti interpretace na vybraný stylový směr, proud nebo školu interpretační praxe. Zvuková nahrávka obsahuje celou řadu dalších informací než jen fyzikální veličiny, které nahrávku vytváří a teoreticky technicky plně popisují. Právě objektivní sledování a na základě faktů argumentované rozkrývání naznačených historických kontextů provedení díla a jeho záznamu v konkrétní nahrávce přináší novou transparentnost ohledně konstrukce vědeckého vysvětlení i procesu dostoupení analytických závěrů muzikologie. Nové technologie a podpůrné analytické nástroje umožňují doložit a zdůvodněně vyjádřit dříve jen subjektivně interpretované charakteristiky díla a jeho interpretace a posouvají tak muzikologické poznání na novou úroveň.

V rámci systematického směru hudebního výzkumu nabízí nový podpůrný analytický nástroj lepší porozumění vnitřním souvislostem a také pochopení vlivů vnitřní stavby díla na výsledný zvukový tvar jeho provedení. Tento hlubší a o data opřený vhled do strukturálních souvislostí anatomie skladebného celku a navazujících zákonitostí jeho zvukové realizace je však pouze základním východiskem a prvním krokem k vytěžení potenciálu této analýzy v kontextu všech dostupných disciplín systematické hudební vědy a jejích dalších interdisciplinárních přesahů.

Hudební praxe pak k naznačeným liniím zaměření systematického výzkumu (dílo, realizace díla, lidský faktor v kontextu zkoumaného díla) doplňuje pomyslnou „druhou stranu mince“ – zaměřuje se na technické a praktické aspekty provedení, aspekty studia a koncipování interpretačního pojetí skladby, hudební pedagogiku a didaktiku a v neposlední řadě konstituuje i vlastní výstavbu znějícího celku v čase. S tím souvisí i aplikační možnosti analytického softwaru v této oblasti. Představovaný analytický nástroj zde může být dobrým pomocníkem a atraktivním doplněním tradičních metod hudební pedagogiky a praktického přístupu k hudební interpretaci. Nástroj MemoVision Software umožňuje interpretům jednak sebereflexi a možnost postupné kultivace vlastního zvoleného interpretačního přístupu s oporou v softwaru, jednak mohou za pomoci softwaru studovat a exaktně porovnávat jiné interpretace. Všechny představené směry pro aplikaci nového analytického nástroje (tj. historická a systematická linie muzikologického výzkumu i hudební praxe) jsou obohaceny vlastnostmi softwaru. Software umožňuje rozšířit a zpřesnit poznávací a hodnotící kapacity hudebního analyтика. Tyto schopnosti vycházejí z principu poloautomatické a statistické analýzy, které jsou v rámci funkcí softwaru implementovány a tvoří jeho funkční základ. Metodika tak navazuje na současný technologický standard v analýze hudebně-interpretačního výkonu a nabízí aktuálně dostupné možnosti pro studium rozdílů různých interpretací stejného díla s podporou počítače.

Se zpřesněním analytických schopností výzkumníka či hudebního interpreta a rozšířením jeho poznávací (paměťové, sluchové i výkonnostní) kapacity souvisí i problematika dosažení určité základní úrovně standardizace hudebně-interpretační analýzy obecně. Zatímco dlouhé období v dějinách hudebně-interpretační analýzy bylo prostoupeno subjektivitou hodnotitele a ovlivněno subjektivním přístupem v hodnocení interpretační estetiky, empirické porovnávání interpretací nabízí možnost objektivního porovnání. I když tato nová objektivita směřuje, s rozvojem takto nastavené analýzy, k vytvoření specifických analytických kategorií pro zarámování určité konstelace parametrů hudebního provedení, neznamená v žádném případě nové technické řešení normativní nástroj pro rozhodování o tom, která interpretace je lepší než jiná. Takové tvrzení je vždy již interpretací naměřených výsledků podléhající nutně subjektivitě výzkumníka či hudebního interpreta pracujícího se softwarem. Tato metodika tedy není nástrojem či zprostředkovatelem hodnocení hudebních interpretací z hlediska jejich „dokonalosti“ či „autentičnosti“ (a tedy v ohledech, které jsou subjektivně hodnotově zabarvené), ale pojednává pouze o principech a možnostech, které do výzkumné práce přináší nové technologické nástroje, a zároveň upozorňuje na proměnu hudebně-interpretační analýzy jako celku právě těmito nástroji.

V perspektivě nových trendů a společenských změn představuje nástroj MemoVision jednu z cest, jak lze propojit lidské schopnosti s širšími možnostmi technického zpracování dat tak, aby bylo docíleno překonání lidských kapacit při zpracování velkých datových objemů. S tím se současně vynořují dosud nepoznané nové horizonty inspirované technickými možnostmi. Práce se softwarem MemoVision se zakládá na synergii poznávacích schopností lidského faktoru s možnostmi počítače a využívá zrychlení procesu získávání relevantní hudebních dat. K vlastní analýze hudebně-interpretačního výkonu s tímto novým nástrojem je však třeba přistupovat poučeně, s vědomím limitů a charakteristik, které tento přístup a použité metody mají.

V rámci péče o kulturní dědictví znamená vývoj nových nástrojů pro práci s velkým objemem dat perspektivu poznání hlubších, univerzálnějších principů kulturního dědictví a vzorců přítomných v datech, které nemusejí být běžnými lidskými poznávacími schopnostmi odhalitelné. Nové analytické nástroje tak přinášejí nové perspektivy, a tím i mohou narážet na definice základních pojmu a kategorií péče o kulturně-historické entity v širokém slova smyslu. Jedním z praktických příkladů těchto změn je možnost rychlého odhalení duplikátu na základě samotného hudebního obsahu v rámci velkého množství nahrávek jedné skladby – duplikátu, tedy dvou zcela totožných hudební interpretací prezentovaných pod různými popisujícími údaji.

Prezentovaný software i tato metodika tak mohou napomoci jak lepšímu porozumění a orientaci v hudbě, tak posluchačské sebereflexi. Současně mohou přinést i nové možnosti pro péči o kulturní dědictví v paměťových institucích. Využití metodiky a softwaru je možné všude tam, kde se pěstuje hudba – tedy od recepce přes analýzu a odborné studium až po vlastní provozování a uchovávání hudby.

1.2 Přínos předkládané metodiky

Tato metodika prezentuje empirický přístup ke zkoumání hudební interpretace a uvádí do českého kontextu pokročilé a částečně automatizované metody zpracování větších objemů dat k hudebně-interpretačnímu výkonu. Obecný přínos metodiky spočívá ve zvýšení přístupnosti nových technických postupů v oblasti zpracování signálů a strojového učení s aplikací na hudebně-analytickou praxí. Předkládaný software reflektuje dosavadní technické pokroky v automatizované analýze parametrů hudebních nahrávek a tradiční muzikologicky zaměřenou analýzu hudebně-interpretačního výkonu. Výrazně zmenšuje potřebný čas pro vytváření plně manuálních anotací k hudebním nahrávkám pro zachycení jejich charakteristik a srovnání vůči ostatním interpretacím. Metodika aplikuje technické metody analýzy a zasazuje je do širšího kontextu v současném paradigmatu obecně přijímaných výzkumných postupů. Software přináší pomyslnou syntézu počítačově podporovaných technik analýzy hudebně-interpretačního výkonu a metodika uvádí a rozebírá širší souvislosti s hudebně-interpretační praxí.

V kontextu České republiky je prezentovaná metodika a s ní úzce propojený software unikátním nástrojem, který však navazuje na starší a již téměř zapomenutou domácí výzkumnou tradici, v jejímž čele stojí badatelské a publikační počiny Milana Kuny a Miloše Bláhy ze 60. a 70. let 20. století. Jejich technické možnosti narázely na mnohé limitace, a kvůli tomu bylo velmi obtížné získat přesná data o interpretačních výkonech většího množství nahrávek. Na jejich zkoumání navazujeme, ale využíváme pokročilé metody zpracování signálů a strojového učení, které již mnohé limitace do určité míry překonaly. V duchu této tradice i pozdějších zahraničních výzkumů na podobná téma je hlavním analytickým prismatem prezentovaného přístupu srovnávací analýza s prvky smíšené výzkumné strategie kombinující kvalitativní a kvantitativní přístup. Hlavním přínosem této metodiky v muzikologickém kontextu je zejména zpřístupnění metod interdisciplinárního oboru Music Information Retrieval (MIR), stojícího na rozhraní počítačových věd, strojového učení, audio inženýrství a matematiky, muzikologické odborné veřejnosti. Metodika zasazuje vybrané analytické postupy do širšího kontextu a reflektuje muzikologický výzkumný proces od počátečních přípravných kroků výzkumu až po nastínění některých možností interpretace výsledků analýzy.

1.3 Komu je metodika určena

Metodika je určena především zájemcům z řad odborné veřejnosti. Z hlediska konkrétních cílových skupin, kterým je tato metodika určena především, lze zdůraznit následující skupiny:

- muzikologové a hudební teoretici,
- hudební interpreti,

- hudební pedagogové,
- audio inženýři,
- pracovníci paměťových, informačních a sdělovacích institucí, kteří přicházejí do kontaktu s hudebními záznamy, a to jak ve veřejném, tak neziskovém sektoru,
- pracovníci v soukromém sektoru.

Uplatnění metodiky samozřejmě není vyloučeno mimo tyto definované zájmové a profesní okruhy. Příjemcem metodiky mohou být i laičtí zájemci z řad široké veřejnosti. V následujících kapitolách jsou stručně nastíněny možná využití jednotlivými skupinami.

1.3.1 Muzikologové a hudební teoretici

Muzikologové mohou metodiku využít jak pro výzkum zaměřený historicky, tak systematicky. V rámci aplikované muzikologie lze prací se softwarem podpořit např. dramaturgickou práci muzikologa, tedy v rovině komparace a následného výběru preferovaného interpretačního přístupu ke dramaturgem zvolené skladbě. Software se tak může nepřímo podílet na edukaci koncertního publiku.

Software umožňuje exaktní a empirické porovnání vybraných interpretací zvolené skladby a jeho využití je, z hlediska konkrétních výzkumných témat, v rámci muzikologie univerzální napříč muzikologickými disciplínami. Rozsah možných výzkumných témat pokrývá oblast akustických vlastností hudební skladby, zasahuje prostor vysvětlení zákonitostí hudební struktury a jejich dopadu na znějící tvar a rozšiřuje možnosti vnímání hudební interpretace posluchačem.

1.3.2 Hudební interpreti

Hudební interpret využije software jak k porovnání stávajících dostupných provedení vybrané skladby, a to s možností jejich seřazení v čase, pokud tyto informace má k dispozici, tak i k porovnání vlastních nahrávek zachycujících formování vlastního interpretačního přístupu ke skladbě. Oba tyto typy komparací mohou interpretovi poskytnout vhled do interpretačního pojetí v rámci volby sousledností a kombinací dynamiky a zvoleného tempa jednotlivých úseků interpretované skladby.

1.3.3 Hudební pedagogové

Přínos softwaru pro hudební pedagogiku úzce souvisí s přínosy analytického nástroje pro hudební interpreti. V rámci hudební pedagogiky ve výuce hudebních dějin a hudební teorie může software poskytnout podporu pro názorné ukázky vybraných jevů souvisejících s vývojem interpretační tradice a upozornit na teoretické zákonitosti provedení. V oblasti hudební pedagogiky

zaměřené na výuku hry na hudební nástroj jsou benefity využití softwaru shodné s přínosy pro hudebního interpreta v rámci osvojování interpretované skladby.

1.3.4 Audio inženýři

Přínos metodiky pro audio inženýry je jednak v oblasti seznámení se základními postupy a strukturou muzikologického pohledu na analýzu hudebně-interpretačního výkonu, tak v technické oblasti – v seznámení se základními i pokročilejšími funkcemi, uživatelským prostředím i výzkumným uplatněním softwaru. Software využívá moderní techniky zpracování signálů, hudební synchronizace a umělé neuronové sítě pro detekci a následné porovnání hudebně zaměřených parametrů digitálních signálů. S audio inženýrstvím souvisí i v současnosti dynamicky se rozvíjející vědecká disciplína MIR se základem v počítačových vědách. V kontextu České republiky se jedná o velmi málo známou oblast vědy, metodika a software tedy přinášejí českým audio inženýrům a výzkumníkům určitý vhled i do této problematiky.

1.3.5 Pracovníci paměťových, informačních a sdělovacích institucí ve veřejném a neziskovém sektoru

Pro pracovníky institucí ve veřejném a neziskovém sektoru, kteří přicházejí do kontaktu s hudebními záznamy, je software užitečným nástrojem pro posouzení stylové orientace hudebního interpreta a inspirací pro tvorbu nových kategorií metadat souvisejících s novými možnostmi, jak dané hudební provedení exaktně popsat. Výhodnou pomůckou je v rámci softwaru i jeho funkce odhalování duplikátů ve velkém množství různých provedení téže skladby a detekce rozdílné hudební struktury vůči notovému předpisu.

1.3.6 Pracovníci v soukromém sektoru

Soukromý sektor může software využít podobně jako instituce ve veřejném a soukromém sektoru, např. pro zpracování a analýzu nahrávek v rámci velkých databází, a zahrnuje pracovníky v soukromém sektoru, konkrétně v organizacích, které se zabývají zvukem, hudbou a příbuznými obory nebo pracovníky v komerčních sdělovacích audio a audio-vizuálních médiích. S tím souvisí i možnosti netradičních popisů nahrávek a vizualizace třeba pro propagaci nových nahrávek a podporu jejich začlenění do kontinuity interpretační tradice (např. v územním pohledu – tj. z hlediska provenience nahrávky či rodiště, místa studií a osobních vlivů ve vývoji interpreta nebo ve stylovém pohledu).

2 Hudebně-interpretační analýza

Následující kapitola uvádí problematiku hudebně-interpretační analýzy jak v minulosti u nás, tak ve světě v rámci současných trendů včetně limitací, na které námi předkládaný přístup reaguje. Téma zahrnuje vzájemné propojení okruhu témat, která v základních konstitučních obrysech ohraničují problematiku analýzy hudebně-interpretačního výkonu jako specifické výzkumné disciplíny. Kapitola vychází od pracovního vymezení pojmu hudebně-interpretační analýzy, stručně seznamuje s vývojovými i aktuálními tendencemi tohoto směru bádání a nastínuje systematické, metodologické a procesní otázky vlastní výzkumné práce v daném tématu.

2.1 Definice hudebně-interpretační analýzy pro potřeby metodiky

Ivan Poledňák definuje ve druhém svazku známé třísvazkové publikace *Hudební věda* (1988) [1], strana 572, úkoly disciplíny „teorie a dějiny hudební interpretace“ následovně: „*Úkolem disciplíny, kterou zde nazýváme Teorie a dějiny hudební interpretace, je shrnovat, třídit a analyzovat dosavadní muzikologické poznatky jakkoli se týkající způsobů provozování hudby, zavádět metody vhodné k jejich zvláštnímu zkoumání a vypracovat na základě empirických faktů a jejich zobecňování pokud možno jednotnou teorii, v jejímž rámci by byly historické a etnické typy hudební interpretace vyloženy jako zvláštní případy globálního „dělání hudby“ člověkem jako společenským tvorem*“. Citované pracovní vymezení studované problematiky, vztážené na celý výzkumný záběr disciplíny „teorie a dějiny hudební interpretace“, již od doby svého vydání na stránkách uvedené publikace ztratilo, ve světle současného paradigmatu, něco ze své dobové aktuality. Nicméně, část citátu zůstává inspirativní i v době vzniku této metodiky.

Ambice konstituovat disciplínu „teorie a dějiny hudební interpretace“ s globální platností vychází z filosofických diskusí o existenci antropologických konstant v lidských kulturách, přičemž prezentované stanovisko tuto existenci přijímá a potvrzuje. V současné době tyto unifikaci tendence oslabují, a automatické přijetí předpokladů o stejnosti fenoménů napříč kulturami bývá zpochybňováno. Vytvoření jednotné koncepce kultury již není prioritou, a to především v kontextu postupující globalizace, ve které je pochopení role a individuality mizejících lokálních kulturních projevů velkým tématem. Naše metodika z hlediska těchto diskusí stojí pomyslně stranou bez vkládání hodnotových premis do výzkumného procesu ani do interpretace výsledků. Výhodou zastávaného přístupu je právě jeho primární zakotvenost v datech a uplatnění kritického myšlení pro konstrukci kategorií, která tato data vřazují do souvislostí a interpretují. Tento přístup nevylučuje bádání od teoretických či filosofických východisek, zakládá si však na argumentování a zdůvodňování na základě analyzovaného materiálu.

Poledňákův citát z výše uvedené publikace nabízí zajímavý a aktuální podnět pro odvození námi využívané definice hned ve své první části: „*Úkolem disciplíny [...] je shrnovat, třídit a analyzovat dosavadní muzikologické poznatky jakkoli se týkající způsobů provozování hudby, zavádět metody vhodné k jejich zvláštnímu zkoumání...*“. Citát totiž vybízí jak k přezkoumávání starších muzikologických poznatků o provozování hudby, tak ke konstrukci a zavádění nových metod pro vlastní dosahování nových odborných poznatků v úzkém kontaktu s hudebním provedením. V citátu je zároveň duální klasifikace na oblast dějin (1) a teorie (2), tedy na oblast historickou a ahistorickou, překročena akcentováním procesu vědecké práce. Toto propojení dvou oblastí v sobě teoretickou a historickou problematiku integruje jako kategorie stejné důležitosti a relevance v rámci poznávacího procesu.

Poznávací proces představuje ústřední bod i pro přístup, z něhož předkládaná metodika vychází. Pro zpřesnění pojetí hudebně-interpretační analýzy v rámci prezentovaného výzkumného přístupu tedy vymezujeme hudebně-interpretační analýzu jako systematický a vědecký (termín „vědecký“ zde akcentujeme zejména z hlediska badatelské etiky a vědeckého pojetí pravdy) poznávací proces, který jako svůj předmět pojímá komplexní, zdůvodněné, podložené a kritické zkoumání hudební interpretace coby akustické realizace hudebního díla, a to se všemi jejími souvislostmi a důsledky včetně předpokladů tohoto poznávání. Zvolený přístup není primárně kvantitativní, či kvalitativní, ale oba tyto přístupy účelně a funkčně kombinuje s využitím principu komparativní výzkumné strategie.

V duální systematizaci muzikologických disciplín, rozdelené na disciplíny systematické a historické, bychom námi proponovanou podobu analýzy hudebně-interpretačního výkonu zařadili do kategorie systematické s širokým interdisciplinárním přesahem. Zde je nutno poznamenat, že analýza hudebně-interpretačního výkonu není výsostnou doménou muzikologie, ale lze na ni nahlížet i pohledem příbuzných oborů. V Poledňákem využívané muzikologické sistematice, která exponuje kategorie „teorie“ a „dějiny“ bychom pak náš přístup zařadili primárně pod teoretické zkoumání, které však s dějinnou problematikou úzce souvisí.

Analytický proces je procesem poznávání, s tím, že jako jednu z hlavních premis poznávání reflekujeme výzkumný cíl. Tato metodika se otázkami cíle výzkumu zabývá spíše okrajově a ilustrativně a zmiňuje především místo a význam stanovení výzkumného cíle ve výzkumném a poznávacím procesu. Metodika tedy nechce být univerzální „kuchařkou“ či přímým návodem na realizaci výzkumu hudební interpretace, ale spíše nástrojem a inspiračním zdrojem, jak o této analýze uvažovat v kontextu výzkumníkem kladených otázek.

2.2 Postupy hudebně-interpretační analýzy

Starší analytickou praxí míníme vývojová období hudebně-interpretační analýzy primárně před rozvojem metod MIR a MPA, tedy primárně empirického přístupu kvantitativní povahy, který

v době vzniku této metodiky akcentuje problematiku automatizace hudebně-interpretační analýzy a automatické extrakce informací z hudebního materiálu. Toto starší období rozdělujeme na počátky hudebně-interpretační analýzy (ve smyslu podrobného, cíleného a ze zvukového záznamu vycházejícího analytického přístupu) bez podpory počítače a na období s převážně manuálním zpracováním zkoumaného hudebního vzorku za podpory počítače.

2.2.1 Analýza bez podpory počítače

Historický kontext analýzy hudebně-interpretačního výkonu bez použití počítače je nutné připomenout zejména s ohledem na analytické vlastnosti starších výzkumných modelů. Empirické studium hudební interpretace mělo před nástupem počítačů a dalších výpočetních technologií nejprve intuitivní charakter, později se rozvinuly i modely zaměřené na sledování hudebních parametrů pomocí jejich co nejpřesnějšího manuálního měření.

V pravém slova smyslu se jedná o analýzu (ve smyslu vědecké metody) jen v malém množství případů. Interpretace byla od doby vzniku hudební veřejnosti často předmětem hodnocení ze strany hudebních kritiků i zainteresovaných laiků. Tyto názory hudební veřejnosti z různých historických období mohou být bezpochyby zajímavým a důležitým podpůrným pramenem pro studium vývoje interpretačního přístupu v rámci daného hudebního díla či segmentu hudebního repertoáru, avšak jejich povaha je většinou subjektivní, navíc často ovlivněna dobovým estetickým ideálem. Rozbory hudební interpretace před nástupem moderních technologií se pak úzce provazují a překrývají s analýzou vlastní hudební struktury, resp. s analýzou hudební formy, harmonického průběhu ve skladbě, tematicko-motivickou analýzou atp.

Zásadními zdroji poznání v této problematice jsou odborné spisy věnované hudební interpretaci. Spisy orientované na pianistickou problematiku představují od konce 19. století reprezentativní publikace věnované hudební interpretaci. V zahraničí je příkladem francouzský spis Alfreda Cortota „Principes Rationnels de la Technique Pianistique“ (1928) [2]. V českých zemích může rané formování odborného přemýšlení o hudební interpretaci reprezentovat text Karla Hoffmeistera „Vývoj klavírní virtuzosity“, zápis autorovy přednášky publikovaný nejprve v časopise Hudební výchova v roce 1911 [3], a později i jako samostatná publikace. Nepominutelnou součástí vývoje studia tématu v Českých zemích je pak disertační práce Ludvíka Kundery „Příspěvky k analýze hudební reprodukce“ z roku 1927 [4], obhájená na Filosofické fakultě Brněnské univerzity. Kunderova práce usiluje o ustavení základních a objektivních analytických kategorií v analýze hudebně-interpretačního výkonu, přičemž tato analýza má především deskriptivní charakter.

Zásadní přínos v odborném řešení tématu představuje ve druhé čtvrtině 20. století v Českých zemích badatelská aktivita Jaroslava Zicha. Zich byl od roku 1937 zaměstnán v Československém rozhlasu a tato fáze kariérní dráhy ovlivnila jeho zájem o znějící hudební tvar. Od roku

1937 Zich publikuje k tématu hudební interpretace první studie. Výrazným knižním publikáčním počinem je pak spis „Prostředky výkonného hudebního umění“ (1959) [5]. Zich své poznatky v tématu následně syntetizuje v práci „Kapitoly a studie z hudební estetiky“ (1975) [6]. Jaroslav Zich problematiku systematizuje v kontextu trendů dobové hudební teorie. Interpretaci sleduje v parametru času a parametru zvuku. Na těchto základech pak staví komplexnější pojem „interpretacní pojetí“ a svou koncepci dokládá několika modelovými analýzami.

Na Jaroslava Zicha pak navazuje dobově progresivní a do jisté míry unikátní experimentální výzkum Milana Kuny a Miloše Bláhy. Autorská dvojice publikovala několik originálních studií, a později rovněž monografii věnovanou empirické analýze hudebně-interpretacního výkonu „Čas a hudba: k dramaturgi časových prostředků v hudebně interpretacním výkonu“ (1982) [7]. Tyto texty jsou výjimečné především úzkým sepjetím s hudební praxí a přímou spoluprací s hudebními interprety. Do výzkumu se promítá i dobová nahrávací technologie – Kuna a Bláha sami nahrávali interprety ve studiu na magnetofonové pásy a následně manuálně analyzovali tento zvukový záznam. Monografie „Čas a hudba“ na základě technicky jednoduchých avšak relativně nepřesných měření předkládá systém pro analýzu hudebně-interpretacního výkonu v časovém rozmezí. Kromě rozboru jednotlivých parametrů časového průběhu interpretace se autorská dvojice zaměřila i na tvůrčí aspekt a zavedla pojem „dramaturgie hudebně časových prostředků“, kam řadí např. relaci základních temp nebo dramaturgii základního tempa.

Na myšlenky a postupy představené Milanem Kunou a Milošem Bláhou plynule navazuje Oleg Podgorný ve své diplomové práci s názvem „Rozšířené metody interpretacní analýzy a některé zákonitosti hudební interpretace“ (vedoucí Josef Bek, obhájeno 1980 na FF UK) [8]. Podgorný předkládá aplikaci metody vyvinuté Kunou a Bláhou s větším zastoupením vizualizace průběhu analyzovaných úseků.

Na mezinárodní výzkumné scéně se empirický výzkum tématu začíná rozvíjet primárně až počátkem 80. let. Tyto práce stojící na počátku dalšího překotného vývoje v poznávání zákonitostí hudební interpretace reprezentuje např. John Rink. Relevantní publikací je např. Rinkem editovaný sborník „The practice of performance: studies in musical interpretation“ (1995) [9]. Publikace je strukturována deduktivně, od základních kategorií přes otázku struktury a významu v hudebním provedení až po procesní pohled na interpretaci. V tomto sborníku jsou zastoupeni badatelé, kteří se tématu v 90. letech 20. století věnovali, a publikace např. Nicholase Cooka předznamenávající další vývoj.

Za pomyslný vrchol výzkumných tendencí empirického zkoumání hudební interpretace zařazeného na manuálním měření akustických parametrů lze považovat četné analytické studie José-Antonia Bowena uveřejňované v prestižních mezinárodních muzikologických časopisech. Jedna z jeho nejvlivnějších publikací je „Duration and Flexibility: Techniques in the Analysis of Performance“ [10]. Bowen ve svých textech nahlíží na znějící hudební tvar z mnoha úhlů pohledu a analýzu jednoho díla pozvedá na novou úroveň četnými komparacemi.

2.2.2 Analýza s podporou počítače

Pro výzkum hudebně-interpretačního výkonu s podporou počítače představuje zásadní zlom zařazení *Centre for the History and Analysis of Recorded Music* (CHARM) v kooperaci univerzit Royal Holloway, University of London; King's College of London a University of Sheffield v roce 2004¹. Toto výzkumné centrum mělo ve svém portfoliu celou řadu aktivit, mimo vlastní výzkum hudebního provedení na hudebním záznamu např. diskografický projekt, digitalizaci hudebních médií, pořádání odborných konferencí a řešení dalších výzkumných projektů. Centrum díky grantové podpoře fungovalo do roku 2004, od roku 2009 s novým výzkumným programem a novým názvem „the AHRC Research Centre for Musical Performance as Creative Practice“.

V rámci CHARM působili kromě zakladatele Nicholase Cooka, který přispěl hudebně-interpretační analýze například knihou „Beyond the score: music as performance“ (2013) [11], též již zmíněný John Rink, a také výrazná osobnost analýzy hudebně-interpretačního výkonu, Eric Clarke. Jedním ze zásadních projektů centra byl pak projekt „Style, performance, and meaning in Chopin's Mazurkas“ (referovaný jako „project Mazurka“)². V tomto projektu byly výzkumníky z CHARM uplatněny program Sonic Visualiser³ [12] vyvinutý ve spolupráci s Centre for Digital Music na Queen Mary, University of London. Program Sonic Visualiser znamenal zásadní posun v počítačem podporované analýze hudebně-interpretačního výkonu hlavně ve smyslu vytváření ručních anotací. V našich případových studiích jsme ho také použili pro vytvoření anotací ve formě časových pozic pro tzv. referenční nahrávky každé kompozice použitého dataset (více v kapitole 3.2.3).

V aktuální interdisciplinární studii Alexandra Lercha „An Interdisciplinary Review of Music Performance Analysis“ (2020) [13] autor popisuje současný stav kombinace technického přístupu ke zpracování hudebních nahrávek a hudební analýzy. Tento přehledový článek uvádí různé přístupy k analýze hudebního výkonu, od měření parametrů interpretačního výkonu, zkoumání vztahů parametrů k jednání a záměrům interpreta až po percepční účinky těchto parametrů na posluchače a možnosti hodnocení hudebního výkonu. Nakonec pojednává o MPA ve vztahu k MIR a poukazuje na příležitosti pro spolupráci a budoucí interdisciplinární výzkum v obou oblastech. Právě na tato téma navazujeme také my předkládanou metodikou a softwarem.

¹<https://charm.rhul.ac.uk/index.html> (cit. 6. 12. 2023)

²<http://mazurka.org.uk/> (cit. 6. 12. 2023)

³<https://www.sonicvisualiser.org/> (cit. 6. 12. 2023)

2.3 Propojení technických a hudebních cílů

Hudebně je naším cílem rozbor hudebních interpretací ve smyslu analýzy fyzických audio nahrávek a srovnávání různých interpretačních přístupů a hudebních výkonů mezi sebou. Na rozdíl od většiny přístupů ale nezačínáme kvalitativní analýzou, nýbrž kvantitativním přístupem zahrnujícím sběr většího množství hudebních dat a získání tzv. anotací, které zachycují charakteristiky jednotlivých interpretací.

V MPA a digitální muzikologii je běžné, že pro analýzu rozdílů je použit malý počet specifických nahrávek nebo symbolických transkripcí (např. MIDI formát dat), které výzkumníci nejprve ručně anotují (vytvoří data ve formě například časových značek taktů, dob, tónů, logických celků apod.). Problémem je, že manuální třídění a vytváření dat je složitý a zdlouhavý proces. Na druhou stranu, plně automatické metody často nejsou dostatečně spolehlivé na to, abychom mohli považovat automaticky generované anotace (třeba odhadnuté lokální tempo interpretace) za „pravdivé“ například ve smyslu muzikologických závěrů. Proto se automatické metody příliš nevyužívají, pokud nejsou doprovázeny manuální kontrolou.

V oboru MIR však existují poloautomatické metody, které pro získání relevantních dat můžeme využít, protože dosahují vyšší přesnosti než automatické metody (a dostatečně přesné pro jejich aplikaci), a zároveň jsou méně časově náročné než manuální přístup. Jejich implementace však vyžaduje technické a programátorské znalosti, a tudíž není v hudebních kruzích běžná. Zde také leží náš přínos – v našem interdisciplinárním kolektivu jsme vytvořili software na základě kombinace moderních technických přístupů, jednoduššího uživatelského prostředí a hudebních cílů. Díky tomu jsme schopni analyzovat velké množství nahrávek a statisticky určit místa kompozice, kde se specifické nahrávky nejvíce liší od ostatních interpretací. Tento proces můžeme nazývat kvantitativní analýzou. Na základě vypočítaných rozdílů a možnosti přehrávat interpretace na zvolených místech a poslouchat jejich rozdíly software dále nabízí vizualizaci různých parametrů nahrávek, které do určité míry korespondují s expresivitou nebo hudebně-interpretačním výkonem.

Následující kapitoly nejprve rozebírají stručný popis metod extrakce informací z hudebních nahrávek orientovaný na analýzu interpretačního výkonu a následně uvádí jednotlivé moduly (části) předkládaného softwaru.

3 Funkce softwaru pro hudební analýzu

Tato kapitola popisuje funkce softwaru od technického zpracování až po uživatelskou interakci s reflektováním cílů hudební analýzy. Předkládaný text primárně neslouží k úplnému matematickému popisu – přestože detailní rozbor algoritmů a použitých metod je nad rámec této metodiky, je vhodné uvést základní informace použitých metod a postupy při zpracování dat pro účely hudební analýzy uvnitř samotného softwaru. Předložený software je založen na dvou předpokladech:

- zprostředkování moderních výpočetních metod komparativní hudebně-interpretační analýzy muzikologické obci, ale i libovolným uživatelům,
- otevřená politika ve smyslu licence, dostupnosti a zdrojového kódu.

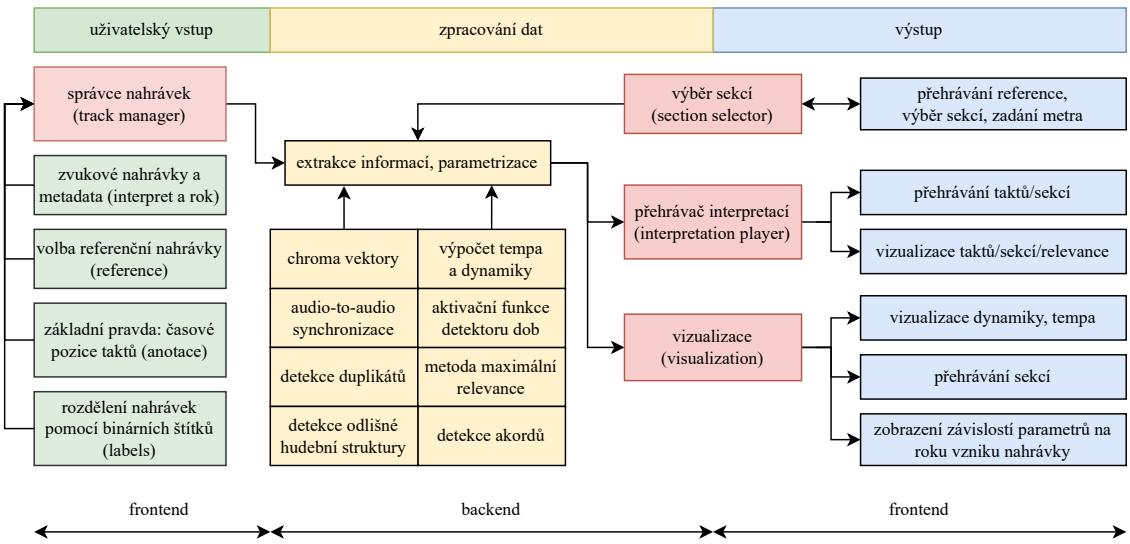
Software je zaměřen na relativně úzké využití pro analýzu interpretačního výkonu hudebních děl a specifické poloautomatické extrakci informací z hudebních nahrávek. Zároveň však poskytuje uživatelsky přívětivé prostředí pro širší komunitu uživatelů, kteří mohou software nebo třeba jen některého jeho funkce využívat. Uživatelská část (frontend) je založena na webovém prostředí a jazyce JavaScript s využitím Vue.js frameworku. Díky tomu není nutné software instalovat lokálně, ale lze k němu přistupovat přes webovou stránku na internetu. Serverové výpočetní prostředí (backend) používá jazyk Python a balíček Flask pro komunikaci s frontendem. Software je volně dostupný přes digitální repozitář GitHub¹ pro lokální instalaci a sdílení zdrojového kódu včetně případných aktualizací a zároveň přes samostatnou webovou stránku projektu pro snadnější přístup bez nutnosti lokální instalace. Odkaz na webovou stránku softwaru bude dostupný přes oficiální stránky projektu². Licence softwaru bude otevřená; předpokládá se MIT licence, ale přesný typ ještě není finální. Může totiž dojít ke změně kvůli balíčkům třetích stran a jejich licencím, které software musí reflektovat.

Výhodou softwaru je použití jazyku Python pro implementaci výpočetních metod. Přestože je obecně pomalejší než některé jiné varianty, ve zpracování dat a umělé inteligenci je momentálně nejpoužívanější. Nové metody a analýzy bývají často vyvíjeny a testovány právě v Pythonu, proto je jeho použití v našem softwaru výhodou. Software je totiž založen na modulech, jednotlivých částech, kdy každá obsahuje specifické funkcionality, které lze v budoucnu upravovat a přidávat podle vývoje výpočetních metod, algoritmů a možností komparativní hudební analýzy. Následující obrázek 3.1 představuje základní schéma funkcionalit softwaru.

Schéma obsahuje tři hlavní části – uživatelský vstup, zpracování dat a výstup. Formální rozdělení na tzv. frontend a backend odpovídá tomu, zda uživatel komunikuje s webovým prostředím přes webový prohlížeč (frontend), nebo jestli zpracování dat probíhá v pozadí na serveru

¹<https://github.com/stepanmk/memovision> (cit. 6. 12. 2023)

²<https://www.pametzvuku.cz/> (cit. 6. 12. 2023)



Obr. 3.1: Přehled uživatelského rozhraní a momentálních funkcí softwaru.

(backend). Zelené panely znázorňují volby uživatele (nahrávání nahrávek a příslušných metadata ve formě názvu interpreta a roku vzniku nahrávky, manuálních anotací ve formě časových pozic taktů, volba referenční nahrávky a přiřazení štítku pro každou nahrávku dle cílů analýzy), žluté panely použité algoritmy a metody extrakce nebo zpracování informací a modré panely ukazují výstupy softwaru ve formě vizualizace (grafů), numerických hodnot nebo možnosti přehrávání nahrávek, zvolených taktů a sekcí a sonifikace anotací. Následující kapitoly rozebírají motivaci pro volbu jednotlivých parametrů a popisují části schématu.

3.1 Parametrizace

Z předchozích studií zabývajících se interpretační analýzou i z obecné hudební intuice je patrné, že některé parametry hudebních nahrávek lze považovat za lepší ukazatele vývoje interpretace než jiné. V našem případě jsme se inspirovali konvenčními parametry pro komparativní analýzu interpretačního výkonu. Základní rozdělení parametrů popisující expresivitu interpretace je následující [13]:

- dynamika: jak se mění hlasitost ve frázích nebo celku, postprocessing,
- barva: instrumentace, techniky hry, akustické podmínky, postprocessing,
- výška: intonace, odchylky od notového předpisu, techniky (např. vibrato),
- čas (tempo): rytmická struktura, mikro struktura (začátky tónů, dob), globální tempo a lokální tempové odchylky nebo rubato.

Jednotlivé kategorie do určité míry charakterizují interpretaci, hudební přednes a styl hry. Tyto třídy parametrů se mohou dále dělit na konkrétní parametry, se kterými lze dále pracovat. Jejich výpočet a měření může být ale částečně zavádějící a pouze některé skutečně odráží danou interpretaci v absolutní i relativní mře. Problém totiž nastává v možnostech extrakce těchto parametrů. Přestože lze určit například pozici taktů, dob a tónů v hudební nahrávce automaticky, neznamená to, že daný parametr skutečně odpovídá vlastnostem hudebního výkonu. Všechny automatické metody mají chybovost, která znemožňuje „důkaznost“ prováděných analýz, protože ty mohou vycházet z nepřesných dat. V rámci MPA je tedy časté, že parametry jsou buďto tzv. anotovány (značeny nebo vytvářeny) experty ručně pro danou sadu hudebních nahrávek, nebo se využívají automatických metod. Jejich výsledky jsou následně zkонтrolovány a případně upraveny. Jiné parametry (hlasitost, výška tónů v čase) bychom mohli považovat do určité míry za přesné (avšak na základě zvolené metody), ale při vytváření závěrů z analýz je vždy nutné myslet na technické limitace, které mohou zkreslit výsledky testování nebo stanovené hypotézy. V předloženém softwaru na tento problém narážíme také, ale pomocí metody synchronizace (kapitola 3.3.2) výrazně zrychlujeme proces manuální anotace datasetu a zároveň redukujeme nutnost využití plně automatických metod. Tento poloautomatický přístup zlepšuje přesnost detekce, ale zároveň podléhá limitaci, protože potřebujeme alespoň jednu sadu ručních anotací pro získání zvolených časových parametrů (například pozice taktů ve všech nahrávkách).

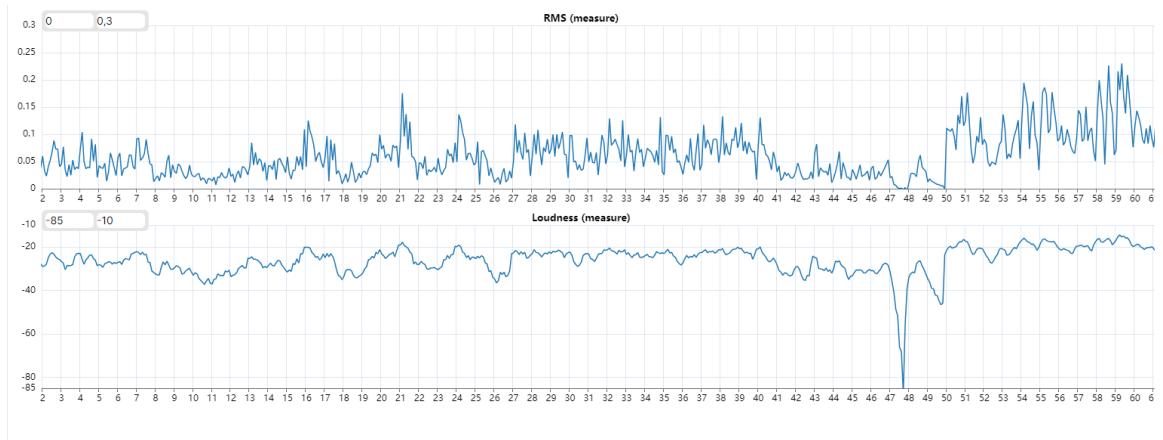
3.1.1 Dynamika

Dynamika může být sporným parametrem, jelikož při analýze různých nahrávek ve většině případů nemáme pod kontrolou nahrávací proces (výběr, vlastnosti a pozice mikrofonů při nahrávání) včetně použití kompresorů, limiterů a dalších efektů při postprodukci. Dalším problémem je kódování a kvalita samotných nahrávek (knihovny, muzea, internetové zdroje a databáze, soukromé sbírky...), fyzické nosiče (voskové válečky, gramofonové desky, CD...) a ztrátové digitální zvukové formáty (MP3, AAC...). Přestože můžeme mít databázi s jednotným formátem zvukových dat, pokud jsme je sami nenahrávaly a nevydávaly, nelze vždy zaručit, že neprošly různou konverzí a formátováním, které mohly změnit dynamický rozsah a kvalitu digitální nahrávky. Při výpočtu dynamiky mohou být výsledky nepřesné (reflektují dostupnou nahrávku, která nemusí odpovídat původní interpretaci nebo původnímu záznamu) v absolutní i relativní mře. Je pravděpodobné, že v dostupných nahrávkách klasické hudby se silná komprese dynamiky spíše nepoužívá. Přesto nelze brát absolutní hodnoty dynamiky jako směrodatné.

Ve většině studií, která extrahuje informaci o dynamice z hudebních nahrávek za účelem výzkumu rozdílů interpretací, je použito RMS (Root Mean Square) neboli efektivní hodnota signálu. Tu lze v případě digitálního signálu pro každý vzorek n signálu $x(n)$ určit jako:

$$x_{\text{RMS}} = \sqrt{\frac{1}{N} \sum_{n=0}^{N-1} x(n)^2},$$

kde N je celková délka analyzovaného rámce. V našem případě



Obr. 3.2: Zobrazení rozdílů mezi parametrem RMS a subjektivní hlasitostí podle doporučení EBU R 128.

provádíme segmentaci celé nahrávky na rámce o délce 2048 vzorků, avšak u každé studie se to může lišit (s minimálním dopadem na interpretaci výsledků). Pro více informací o programové implementaci viz výpočet v modulu librosa³. Tímto způsobem zjistíme efektivní hodnotu signálu, která do určité míry vypovídá o dynamice nahrávky, ale neodpovídá subjektivní hlasitosti. Pro účely porovnání změn dynamiky se řídíme doporučením EBU R 128⁴ a počítáme kromě standardního RMS také hlasitost (loudness). Pro více informací k výpočtu hlasitosti a její programové implementaci odkazujeme na [14].

Na obrázku 3.2 demonstrujeme rozdíl mezi RMS a hlasitostí u jedné z nahrávek datasetu. Celkový průběh obou křivek je velmi podobný, ale v hodnotách i vývoji hlavně kratších úseků se mohou lišit. Navíc na rozdíl od RMS, kde je jednotka ve významu dynamiky složitě interpretovatelná, jednotkou hlasitosti je decibel (dB), respektive v našem konkrétním případě tzv. Loudness Units relative to Full Scale (LUFS), který je vztažen k vlastnostem lidského slyšení (pomocí specifické filtrace) a lépe reflekтуje subjektivní vjem dynamiky. Je však patrné, že u výpočtu hlasitosti dochází k vyhlazení rychlých dynamických změn, celkový subjektivní dojem změny dynamiky je ale zachycen lépe.

3.1.2 Barva a výška

Barvou neboli témbrem rozumíme frekvenční složení zvuků nebo také celé nahrávky, které tvoří výsledný zvukový vjem. Barvu můžeme definovat jako vlastnost zvuku, díky které jsme schopni rozlišit hudební nástroje nebo obecně druh zvuku, pokud všechny ostatní vlastnosti jsou stejné. Například pokud zahrajeme určitý tón na klavír a stejný tón (výška a trvání tónu) na housle,

³<https://librosa.org/doc/main/generated/librosa.feature.rms.html> (cit. 6. 12. 2023)

⁴<https://tech.ebu.ch/docs/r/r128.pdf> (cit. 6. 12. 2023)

jsme pomocí frekvenční analýzy ve vnitřním uchu a mozku schopni rozeznat, že se tóny liší, případně identifikovat na základě zkušenosti, o které nástroje se jedná.

Frekvenční složení digitálního zvuku můžeme získat harmonickou analýzou, matematicky diskrétní Fourierovou transformací. V reálných aplikacích se využívá efektivního algoritmu rychlé Fourierovy transformace. Ta rozkládá signál v čase na jednotlivé frekvenční složky, ze kterých je signál složen – díky tomu jsme pro daný časový úsek schopni odhadnout výšku zahranych tónů a také barvu zvuku. Pokud transformaci aplikujeme na kratší časové úseky, címž získáváme tzv. krátkodobou Fourierovu transformaci (STFT), a výsledek zobrazíme, získáváme tzv. spektrogram neboli časově-frekvenční zobrazení odhadu výkonové spektrální hustoty. Při využití strojového učení a hlubokých neuronových sítí jsme například schopni zjistit výšku všech zahranych tónů klavíru v čase a libovolné polyfonické struktury s vysokou přesností (obecná automatická detekce výšky pro libovolnou polyfonní strukturu a instrumentaci je ale stále velmi složitá). Matematický popis transformace a rozbor pokročilých systémů automatické hudební transkripce jsou nad rámec této metodiky. Je však vhodné podotknout, že z frekvenčního složení můžeme zjistit více přínosných hudebních informací než z časového průběhu signálu, což koresponduje se skutečností, že frekvenční analýzu využívá i lidské ucho. Variantami STFT používanými v MIR jsou například chroma vektory (viz kapitola 3.3.1, časově-frekvenční reprezentace, kdy na svislé ose jsou frekvence přepočítány na půltóny, normalizovaná intenzita (z anglického *intensity*, nejedná se však o fyzickou intenzitu zvuku) je pak znázorněna barvou jednotlivých buněk matice (třetí osa reprezentace) a vyjadřuje, jak moc je daný půltón v čase obsažen v signálu. Podobné reprezentace mohou sloužit systémům pro detekci tónů, hudebních dob, taktů nebo harmonické struktury a akordů.

Barvu a frekvenční složení hudební nahrávky může ovlivňovat mnoho faktorů; mnohé z nich však nelze považovat za součást interpretace, ale spíše za artefakty zvukové produkce, možnosti a postprodukce. Ve spektru můžeme pozorovat, detekovat a následně porovnávat instrumentaci nebo třeba výrazové prvky typu vibrato. Pokud bychom počítali obecné parametry zvukových signálů, které se v MIR třeba pro klasifikaci žánrů nebo zvukových objektů používají (spektrální jasnost, kontrast apod.), naše výsledky by nemusely reflektovat obecné kvality interpretace, jako spíše použití digitálních efektů nebo ztrátové komprese v postprodukci. Protože barva může být sporným parametrem v posouzení rozdílů interpretací, v našem postupu využíváme časově-frekvenční reprezentace pouze pro tzv. synchronizaci, nikoliv pro výsledné vyhodnocování rozdílů.

3.1.3 Čas

Posledním hlavním parametrem hudebních interpretací z pohledu zvukové nahrávky je čas (v literatuře se objevuje primárně pojem *timing*, tedy spíše „časování“, nikoliv *time* „čas“; přesto

pro naše účely a kontext metodiky překládáme tento výraz jako „čas“). Ten hraje klíčovou roli při zjišťování rozdílů mezi přednesy interpretů. To je dáno několika faktory. Zaprvé, globální tempo nahrávky nebo větších hudebních celků a lokální rytmické odchylky ve smyslu dob, taktů nebo hudebních motivů jsou důležitým faktorem interpretace a zadruhé, tempo a časová informace obecně jsou méně ovlivněny zvukovým řetězcem, od nahrávání až po produkci, a v podstatě nezáleží na kvalitě záznamu ve smyslu digitálního formátu nebo ztrátové komprese.

Časová informace by teda měla být v každém případě uchována. I zde však existují výjimky a případy, které bychom neměli opomenout. Jde hlavně o externí faktory. Pokud máme velmi starou nahrávku, například uchovanou na voskovém válečku, může nastat problém s její reprodukcí (změna tempa skladby při nestabilní rychlosti otáčení válečku při záznamu nebo přehrávání). Limitace analogových médií může být problematická také kvůli omezenému fyzickému místu pro záznam (například omezená délka záznamu podle vlastnosti fyzického nosiče), kdy interpreti mohli být nuteni hrát rychleji, jen aby stihli skladbu nahrát na dané médium. Dalším problémem může být porovnávání záznamů natočených v dedikovaných nahrávacích studiích, kdy je možné záznam nahrát několikrát a posléze vybrat nejlepší provedení nebo provádět stříhy, včetně záznamů z živých koncertů. Jelikož v naší analýze porovnáváme nahrávky i staršího data a kombinujeme studiové a živé záznamy, nemůžeme všechny faktory vyloučit.

3.2 Uživatelský vstup

3.2.1 Hudební nahrávky

Při práci se softwarem se v prvé řadě předpokládá, že uživatel má k dispozici nahrávky, které chce analyzovat. Nahrávky mohou být v libovolném standardním audio formátu (WAV, MP3, OGG, ACC atp.) a jejich časová délka, případně velikost, ovlivní do určité míry rychlosť zpracování. Délka i velikost jsou teoreticky neomezené, doporučujeme však velmi dlouhé nahrávky (více jak 30 minut) rozdělit na několik kratších a provést hudební analýzu pro každý úsek zvlášť. Nahrávky by měly obsahovat stejný hudební materiál ve smyslu skladby/kompozice. Pokud některé nahrávky obsahují potlesk nebo nejsou stříženy přesně vůči referenční nahrávce, může to způsobit odchylky v synchronizaci nebo bude uživatel upozorněn na různou hudební strukturu některých nahrávek, což lze v modulu Přehrávač interpretací manuálně zkонтrolovat. Uživatel může sporné sekce nahrávky posléze poslechnout a zjistit, zda došlo pouze k nepřesnostem, ale nahrávky se ve skutečnosti kompozičně neliší. Pro účely projektu a případové studie jsme spolupracovali s pracovníky Národního muzea v Praze, kteří vytvořili a nyní spravují databázi klavírních nahrávek Bedřicha Smetany a Antonína Dvořáka (viz kapitola 5.1) pro účely navazujících případových studií.

3.2.2 Výběr reference

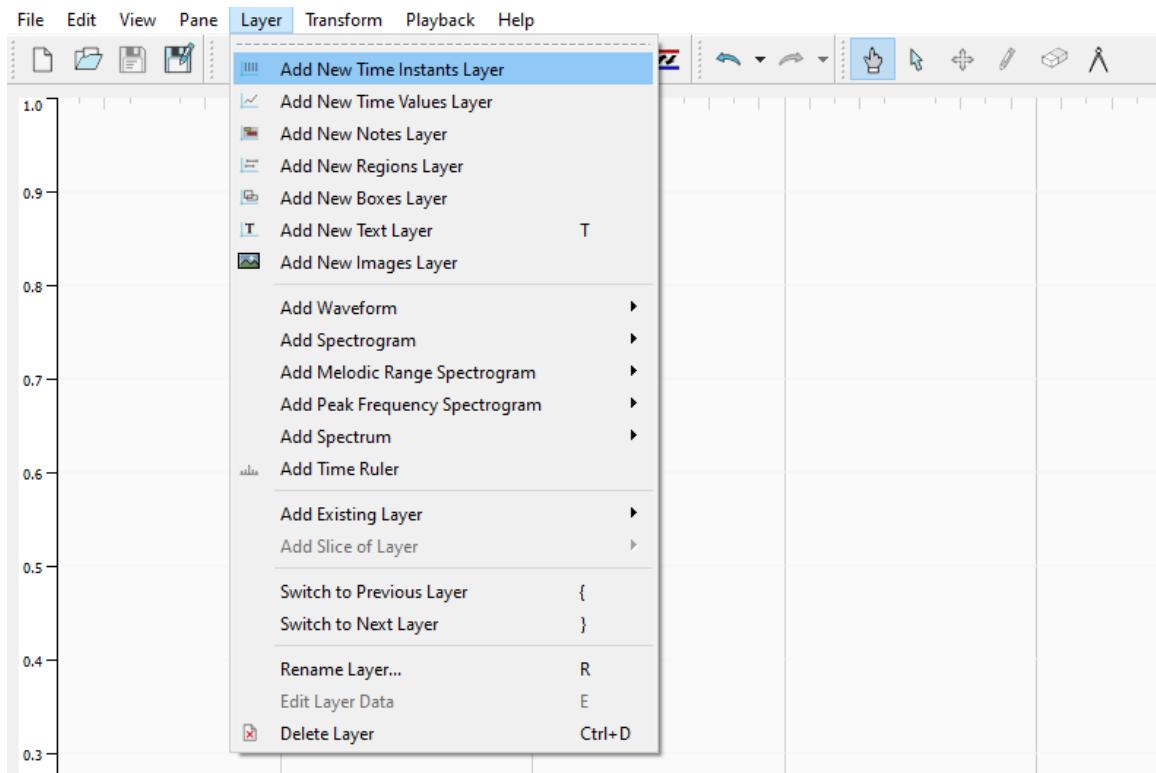
Referenční nahrávka (nahrávka zvolená pro manuální anotaci a následný přenos informací na ostatní interpretace) by měla splňovat některá kritéria, aby byla vhodnou referencí pro synchronizační proces. Neměla by se příliš odlišovat od ostatních nahrávek ve smyslu výrazných lokálních změn v tempu a měla by mít dobrou audio kvalitu. Pochopitelně jsou tyto charakteristiky relativní a sporné, synchronizační proces by ale měl být dostatečně robustní – většinou záleží více na konkrétní kompozici (harmonický a tempový vývoj) spíše než na dané interpretaci. Podmínkou je, aby byly časové anotace k dispozici alespoň k jedné interpretaci ze sady analyzovaných nahrávek. Tuto interpretaci pak můžeme označit za referenční pro synchronizační proces. V následující kapitole je popsán proces vytvoření manuálních anotací.

3.2.3 Anotace

Pro využití analyzačních funkcí softwaru a rozdelení nahrávek do přehledné struktury je nutné mít ještě před samotným začátkem analýzy k dispozici časové anotace ve formě pozic taktů pro referenční nahrávku. Námi navrhovaný software obchází potřebu automatické detekce pozic taktů pomocí audio-to-audio synchronizace (viz kapitola synchronizace nahrávek) a jedné ruční anotace, tzv. základní pravdy. Prvním krokem je vlastní výběr referenční nahrávky, pro kterou uživatel musí mít k dispozici časové pozice taktů. Ty může vytvořit například v již zmíněném Sonic Visualiseru.

Sonic Visualiser

Sonic Visualiser [12], známý software v oboru výpočetní muzikologie, komparativní analýzy a analýzy interpretačního výkonu, který je vhodný hlavně pro vytváření a následnou vizualizaci ručních anotací, disponuje velmi omezenou extrakcí a vyhodnocením informací. Pokud uživatel vytváří ruční anotace, je velmi přehledný a pohodlný na ovládání. Pokud je naším cílem automatická analýza nízko- a vysokoúrovňových parametrů (tóny, doby, takty, akordy), spoléhá na přídavné Vamp moduly (plugins), které již dnes nejsou příliš aktuální, respektive ne-reflektují nejnovější metody a přístupy k detekci parametrů. V Sonicu není zabudovaný systém, který by využíval synchronizační metodu pro přenos anotací z jedné referenční nahrávky na ostatní cílové nahrávky (částečnou funkcionalitu synchronizace můžeme najít v přidruženém programu Sonic Lineup), ani nevyužívá strojové učení nebo umělé neuronové sítě. Detekční metody nejsou ve většině případů vhodné ani pro poloautomatickou analýzu, kdy využíváme výsledků automatického systému s následnou ruční korekcí, protože i v době jejich vzniku nedosahovali dostatečné přesnosti.

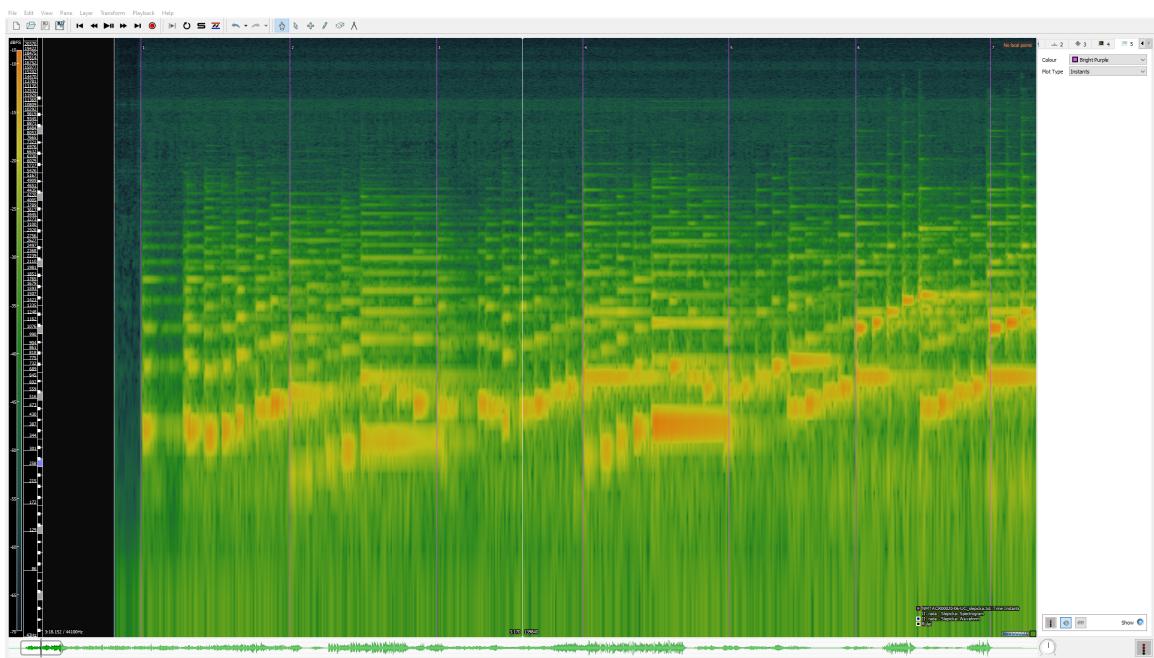


Obr. 3.3: Dialogové okno s možnostmi vytvoření nové vrstvy pro analýzu hudební nahrávky v Sonic Visualiseru.

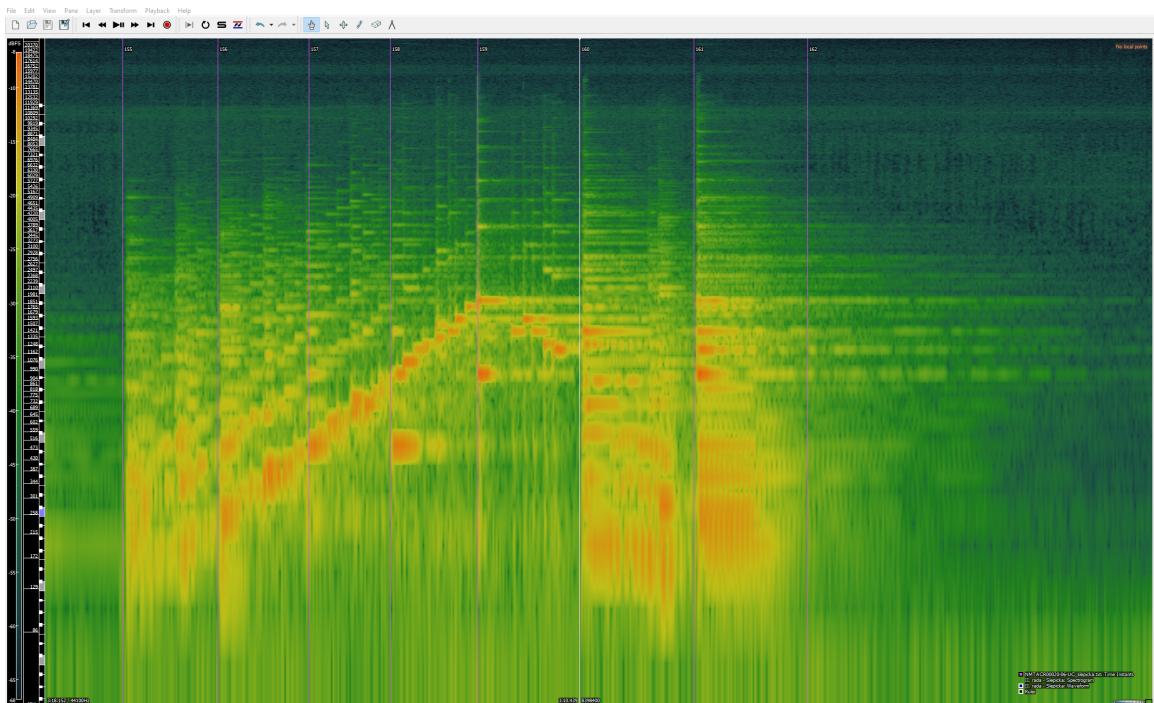
Po nahrání zvolené nahrávky do Sonic Visualiseru může uživatel vytvořit novou vrstvu (viz obrázek 3.3), přehrát nahrávku a pomocí opakovaného stlačení příslušné klávesy (v základním nastavení se jedná o Enter nebo středník) vytvářet časové značky, které následně při zpětném poslechu upraví na správnou pozici.

Sonic Visualiser má také zabudovanou časově-frekvenční vizualizaci, například v podobě spektrogramu v logaritmickém měřítku, která může dopomoci k přesnějšímu zarovnání začátků taktů k nástupu tónů taktu (viz obrázek 3.4).

Každý začátek taktu však neobsahuje tón a v některých kompozicích nebo vybraných paritech může dojít k situaci, kdy je potřeba označit několik taktů po sobě, přičemž ani jeden neobsahuje žádnou hudební událost. V tom případě doporučujeme rozdělit volné takty rovnoměrně nebo pravidelně podle tempa předchozích taktů. V budoucnu lze metodu synchronizace (která je důvodem pro vytvoření ruční anotace referenční nahrávky) doplnit kotevními body, které označují úsek nahrávky, pro které má mít synchronizace čistě lineární průběh. Konec každého taktu je dán začátkem dalšího, což však neplatí pro poslední takt. Proto poté, co jsou všechny takty v Soniku označeny, je potřeba přidat poslední anotaci ve formě časové značky konce posledního taktu (viz obrázek 3.5).



Obr. 3.4: Příklad časově-frekvenční reprezentace, konkrétně logaritmicky filtrovaného spektrogramu na základě STFT. Vertikální fialové čáry představují časové pozice taktů dané nahrávky (anotace).



Obr. 3.5: Poslední vertikální fialová čára představuje fyzický konec posledního taktu, který je pro výpočet délky a tampa každého taktu potřeba do anotace přidat.

Soubor	Úpravy	Formát	Zobrazení	Nápočeda
1.123265306	1			
1.848163265	2			
2.540408163	3			
3.151020408	4			
3.853061224	5			
4.491201814	6			
5.100408163	7			
5.721791383	8			
6.364444444	9			
7.028435374	10			
7.644557823	11			
8.244807256	12			
8.848979592	13			
9.476099773	14			
10.050612245	15			
10.685328798	16			
11.464489796	17			
12.143673469	18			
12.695510204	19			
13.280000000	20			

Obr. 3.6: Příklad strojového zápisu anotací: první sloupec označuje časovou značku po sobě jdoucích taktů, druhý sloupek představuje index (číselné označení) daného taktu.

Nakonec uživatel exportuje anotace do textového souboru tak, aby mohly být dále použity pro analýzu. Podle standardu MIR⁵ a MIREX komunity⁶ by měly být časové značky/anotace/pozice uchovány ve formátu sloupcového vektoru (každý řádek obsahuje jednu hodnotu) vzestupně od prvního taktu po poslední. Nejčastěji se používá standardní textový soubor typu .txt nebo .csv, někdy se jako druhá hodnota na daném řádku uvádí index taktu (viz obrázek 3.6).

Přenos anotací

Anotace slouží pro následný přenos na všechny ostatní interpretace. Díky tomu jsme schopni na hrávky segmentovat na jednotlivé takty a logické úseky. Uživatel se může jednoduše navigovat v hudebním materiálu a vizualizace parametrů může mít společnou relativní osu, kde jednotkou jsou takty místo klasického času. Proces vytváření manuálních anotací může být urychlen pomocí automatických metod, kdy je nejprve využit automatický detektor, a následně je výsledek opraven expertem/uživatelem. Právě k těmto účelům slouží mimo jiné Vamp pluginy pro Sonic Visualiser, ale protože nedosahují dostatečné přesnosti, úprava automatické detekce je často časově náročnější než vytvoření anotací zcela manuálně. Na základě limitací klasických metrik MIR pro porovnávání automatických detektorů v případě muzikologické analýzy byla implementována evaluační metrika pro simulaci vhodnosti detekčního systému pro manuální opravu výsledných anotací [15]. V naší dřívější studii jsme tuto metriku použili ke srovnání automatického state-of-the-art přístupu automatické detekce taktů se synchronizací [16]. Z výsledků

⁵<https://www.ismir.net> (cit. 6. 12. 2023)

⁶https://www.music-ir.org/mirex/wiki/MIREX_HOME (cit. 6. 12. 2023)

je patrné, že automatický systém ve většině případů nedosahuje dobrých výsledků na rozdíl od synchronizace, která je i přes své limitace vhodnějším přístupem pro komparativní hudební analýzu.

3.3 Zpracování dat

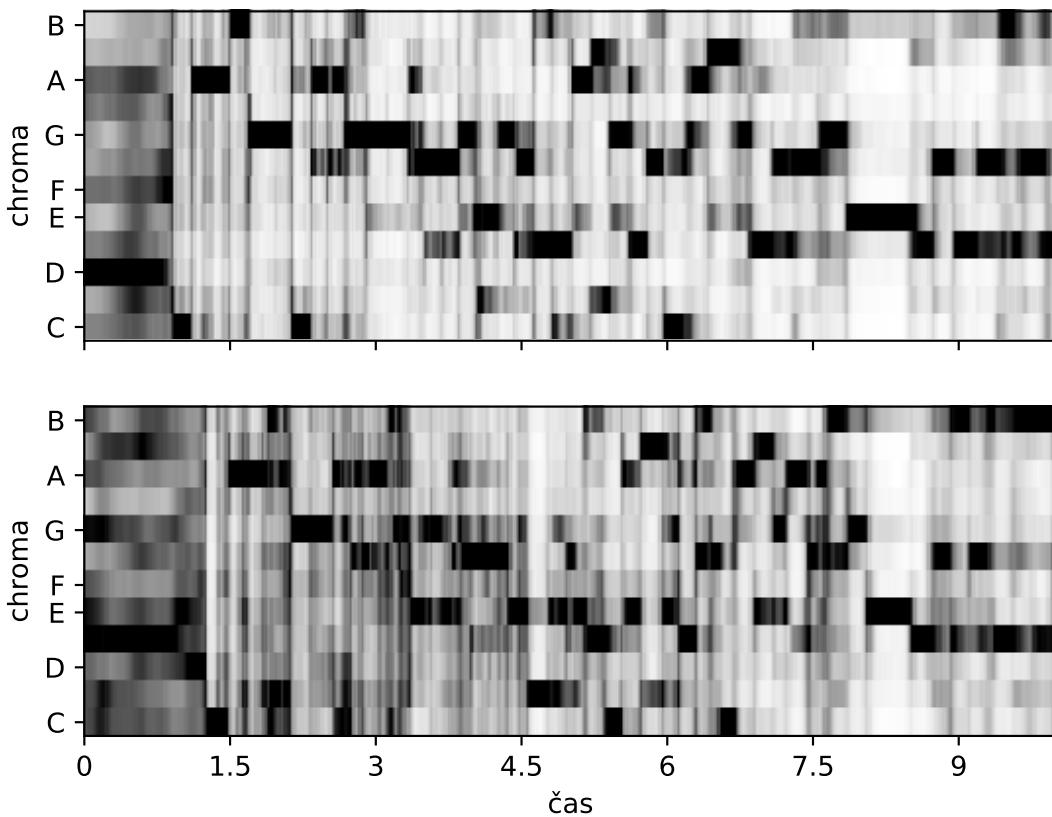
3.3.1 Chroma reprezentace

V kapitole 3.1.2 byla zmíněna časově-frekvenční reprezentace získaná pomocí STFT. Pro účely dalšího zpracování využíváme tzv. chroma vektory. Jedná se o reprezentaci odvozenou z STFT, kdy na svislé ose získáváme místo frekvenčních pásem pásma půltónová (12 hodnot podle půltónů v rámci jedné oktávy) [17]. Hodnoty jsou kumulativní, pásmo pro tón c odpovídá všem oktávám tónu c dohromady, které signál v daném místě obsahuje. Příklad chroma reprezentace je uveden na obrázku 3.7. Detaily a implementace chroma vektorů jsou uvedeny ve zdrojích [18] a [19]. Tento druh zobrazení má velkou výhodu – pokud předpokládáme, že máme k dispozici různé interpretace stejné skladby, chroma vektory by se teoreticky měly lišit jen v čase, ale melodická/harmonická struktura by měla být zachována, což lze pozorovat na výše zmíněném obrázku. Díky tomuto předpokladu můžeme využít v následujícím kroku synchronizaci.

3.3.2 Synchronizace nahrávek

Synchronizace v kontextu hudebních nahrávek a oboru MIR je metoda tzv. časového zarovnání, kdy naším cílem je zjistit, které vzorky jedné nahrávky odpovídají vzorkům druhé nahrávky. Jedná se o přiřazení hodnot jedné reprezentace k hodnotám druhé, odpovídající reprezentaci. Reprezentace se mohou lišit v čase, předpokládá se však nějaký společný parametr na základě zvolené reprezentace, v našem případě harmonická struktura zjištěná pomocí chroma vektorů. V rámci navrhovaného softwaru je základem synchronizace metoda DTW (Dynamic Time Warping neboli dynamické borcení časové osy), respektive její pokročilá varianta MrMsDTW (Multiscale memory-restricted DTW) [20] dostupná přes balíček synctoolbox⁷. Výsledkem je křivka definující optimální borcení času pro kombinaci daných dvou nahrávek (reference a cílové nahrávky). Pomocí této křivky můžeme časové anotace referenční nahrávky interpolovat na všechny cílové nahrávky datasetu. Tím získáme informaci o taktech pro všechny nahrávky, které sdílí harmonickou strukturu. Synchronizačních technik existuje více, v našem případě využíváme tzv. audio-to-audio synchronizaci na bázi DTW, kdy používáme synchronizaci dvou hudebních nahrávek (resp. jejich upravených časově-frekvenčních reprezentací). Pokud bychom

⁷<https://github.com/meinardmueller/synctoolbox> (cit. 6. 12. 2023)



Obr. 3.7: Příklad zobrazení chroma reprezentace pro prvních deset sekund dvou různých interpretací Smetanovy *Cibulicky*.

měli k dispozici notový zápis ve strojové podobě (například MIDI, XML nebo MusicXML soubor z notačního softwaru), mohli bychom extrahouvat pozice taktů a aplikovat tzv. score-to-audio synchronizaci, a tím obejít nutnost manuálních anotací. V rámci komparativní hudební analýzy v oboru MPA jsou manuální anotace běžné a často je časově náročnější vytvořit strojový notový zápis než manuální anotace pozic taktů. Záleží na cíli analýzy a na dostupnosti dat. Je to však funkce, která by v budoucí verzi navrhovaného softwaru mohla být k dispozici.

3.3.3 Doplňující funkce

Kromě synchronizace v softwaru probíhá výpočet odhadu ladění, detekce odlišné hudební struktury a detekce duplikátních nahrávek. Kontroluje se, zda uživatel nevědomě nenahrál duplikáty, které by pro analýzu byly nadbytečné a mohly by analýzu zkreslit. Nahrávky se porovnávají na základě zmíněných chroma vektorů převedených do rastrových obrázků ve stupních šedi a odečtu hodnot pixelů všech kombinací nahrávek. Pokud je jedna nahrávka kratší, posune se při výpočtu po pixelech vůči druhé. Díky tomu jsme schopni zachytit i duplikáty, kdy jedna

nahrávka je pouze kratším výsekem jiné nahrávky (ale jedná se o stejnou interpretaci). Tato funkce vznikla také samostatně v rámci našeho projektu TA ČR pro účely Moravské zemské knihovny⁸.

Nelze vždy zaručit, že uživatel nahraje pouze interpretace, které následují stejný notový zápis. Proto je v softwaru implementována metoda, která na základě vlastností křivky z MrMsDTW metody porovnává, zda se referenční a cílová nahrávka shoduje v harmonickém obsahu. Výsledkem je upozornění uživatele, že některé kombinace nahrávek nemusí mít stejnou strukturu, a označení času nepřesnosti. V modulu přehrávače interpretací pak uživatel může ručně zkontovalovat, zda systém určil rozdílná místa správně. Zjištění nepřesnosti není triviální a v momentální implementaci může docházet k případům, kdy výpočet není přesný nebo nezachytí všechny rozdílné úseky, proto doporučujeme manuální zběžnou kontrolu. Metoda je více popsána v naší předchozí studii [21].

3.3.4 Metoda maximální relevance

Poté, co došlo k synchronizaci všech párů nahrávek (reference vs. všechny ostatní), vytvoříme tzv. data matici, která obsahuje délku trvání každého taktu v sekundách všech nahrávek. Každá nahrávka navíc obsahuje dříve zvolený binární štítek (label), díky kterému spadá do jedné ze dvou kategorií/skupin. Můžeme tedy získat například skupiny český–nečeský, nahrávka pořízena *před rokem 2000 – po roku 2000 včetně* atp. podle vstupu uživatele na začátku analýzy v prvním modulu správce nahrávek. Následně aplikujeme metodu maximální relevance (MR), pomocí které zjistíme největší rozdíly mezi dříve zvolenými dvěma třídami nahrávek. Pro matematický popis metody odkazujeme čtenáře na [22], pro programovou implementaci pak na digitální repozitář implementace, kterou jsme použili v našem softwaru⁹. V obou případech se jedná o popis tzv. Maximum-relevance minimum-redundancy metody, která kromě relevance počítá i s redundancí, tedy nadbytečností. Tu lze využít například pro trénování klasifikátorů na bázi strojového učení, kde její aplikace dává smysl a výsledná redukce dat může zvýšit přesnost klasifikace [23]. V našem případě by byla její implementace nevhodná, protože naším cílem je zjistit pouze relevantní parametry pro rozdelení binárních tříd, nikoliv určení parametrů, které jsou relevantní a zároveň nekorelované vůči sobě. Z celé metody jsme tedy vybrali jen část s výpočtem relevance a přepočítali ji na relevantní relevanci v kontextu všech taktů dané skladby (její hodnoty se pohybují mezi 0 a 1) pro lepší čitelnost výsledků.

⁸<https://github.com/xistva02/AudioDuplicateFinder> (cit. 6. 12. 2023)

⁹<https://github.com/smazzanti/mrmmr> (cit. 6. 12. 2023)

3.4 Výstup softwaru

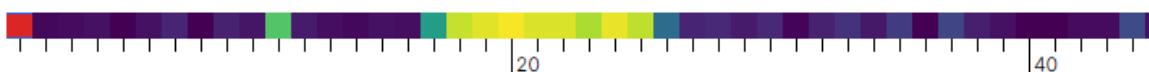
3.4.1 Vizualizace dat uživateli

Výsledkem celého výpočetního procesu v současné verzi softwaru je:

- délka každého taktu všech dostupných (nahraných) interpretací,
- segmentace každé nahrávky na takty,
- výpočet tempa (rozlišení jedna hodnota pro jeden takt),
- výpočet dynamiky (10 hodnot na jeden takt),
- možnost přehrání libovolných taktů/sekcí,
- vizualizace relevance všech taktu na základě vybraných parametrů a binárních štítků,
- vizualizace tempa a dynamiky pro zvolené takty/sekcí a interpretace,
- vizualizace odhadu akordické struktury,
- vizualizace délky zvolených sekcí interpretací vůči roku vzniku nahrávek,
- vizualizace dalších parametrů podle současných možností softwaru¹⁰.

Základní výpočet zahrnuje několik parametrů reflektující dynamiku a tempo. Zajímavé jsou také modely neuronových sítí, které provádí automatickou transkripci klavírních nahrávek na tzv. piano rolls, tedy vizualizace zahraných tónů (počátků a konců) v čase včetně odhadu použití pedálu. Ty mohou přinést nové srovnávací možnosti pro zjištění rozdílů například v přesnosti nástupu každého tónu v akordu, sekvenci apod. Ukázky piano rollu dvou interpretací stejného úseku skladby je možné prozkoumat například na digitálním repozitáři *Piano Transcription* od bytedance¹¹.

Jedním z hlavních přínosů softwaru je možnost využít metodu MR pro libovolně zvolené binární štítky (dvě separátní kategorie nahrávek podle vlastního výběru) s následným libovolným zobrazením parametrů při simultánním poslechu rozdílů interpretací. Můžeme zvolit dvě třídy podle charakteristik nahrávek/interpretací (původ, hudební škola, věk, gender atp.) nebo analyzovat rozdíly jedné nahrávky vůči všem ostatním. V tom případě má daná nahrávka jeden štítek a všechny ostatní štítek druhý. Příklad vizualizace relevance taktu použité v softwaru je uveden na obrázku 3.8.



Obr. 3.8: Příklad zobrazení relevance jedné skupiny nahrávek vůči druhé. Světlá barva značí vyšší relevanci, a tedy vyšší rozdíly v tempu jedné skupiny vůči druhé v rámci daných taktu.

¹⁰Další vizualizace závisí na možnostech softwaru, které se v budoucnu mohou měnit.

¹¹https://github.com/bytedance/piano_transcription (cit. 6. 12. 2023)

Na základě barevných rozdílů (tmavá modrá barva odpovídá nízké relativní relevanci, světlá zelená až žlutá pak vysoké relativní relevanci) vidíme, které takty nebo sekce nahrávky jsou relevantnější a u kterých není větší rozdíl v kontextu rozdílů zvolených tříd. Červeně je zvýrazněn aktuálně zvolený nebo přehrávaný takt (více v kapitole 4.1.3). Tato informace je přínosná, protože právě zobrazení relevance nám může poskytnout prvotní odhad, na které části skladby bychom se při zkoumání rozdílů interpretací měli zaměřit.

V posledním modulu *Visualization* pak můžeme vybrat zvolené úseky a vizualizovat rozdíl i v ostatních parametrech. Stejným způsobem můžeme interpretace přehrávat a rychle přepínat mezi více nahrávkami pro okamžitou sluchovou zpětnou vazbu. Parametry jsou uvedeny v synchronizované časové osy a dále přepočítány na tzv. relativní osu a převzorkovány na 10 hodnot pro každý takt, kdy osa x obsahuje takty namísto klasického času. Díky tomu lze sledovat vývoj parametrů uvnitř struktury nahrávky (a odkazovat se v rámci analýzy zároveň na notový zápis) a okamžitě porovnávat jednotlivé interpretace. V momentální verzi softwaru lze sledovat vývoj dynamiky pomocí efektivní hodnoty (RMS) a hlasitosti (loudness) na základě EBU R 128, aktivační funkci detektoru dob na základě námi natrénované neuronové sítě, odhad akordické struktury pomocí volně dostupné neuronové sítě a tempo a délku trvání podle synchronizovaných taktů.

4 Uživatelské prostředí softwaru

Tato kapitola poskytuje základní orientaci v uživatelském prostředí softwaru, popisuje jednotlivé funkční prvky a schematicky naznačuje jednotlivé funkce dostupné v softwaru. Tato část metodiky bude úzce propojena s uživatelskou příručkou k softwaru (dostupnou v digitálním repozitáři softwaru¹). Je nutné podotknout, že funkce a vizuální stránka softwaru prezentované v metodice odpovídají současné verzi při datu vydání této metodiky, v budoucnu se však mohou měnit a nemusí uvedeným příkladem odpovídat. Jedná se o vizuální podobu některých modulů nebo nové, přídavné funkce vzniklé na základě zpětné vazby uživatelů, výzkumníků nebo potřeb analýzy s vývojem nových metod.

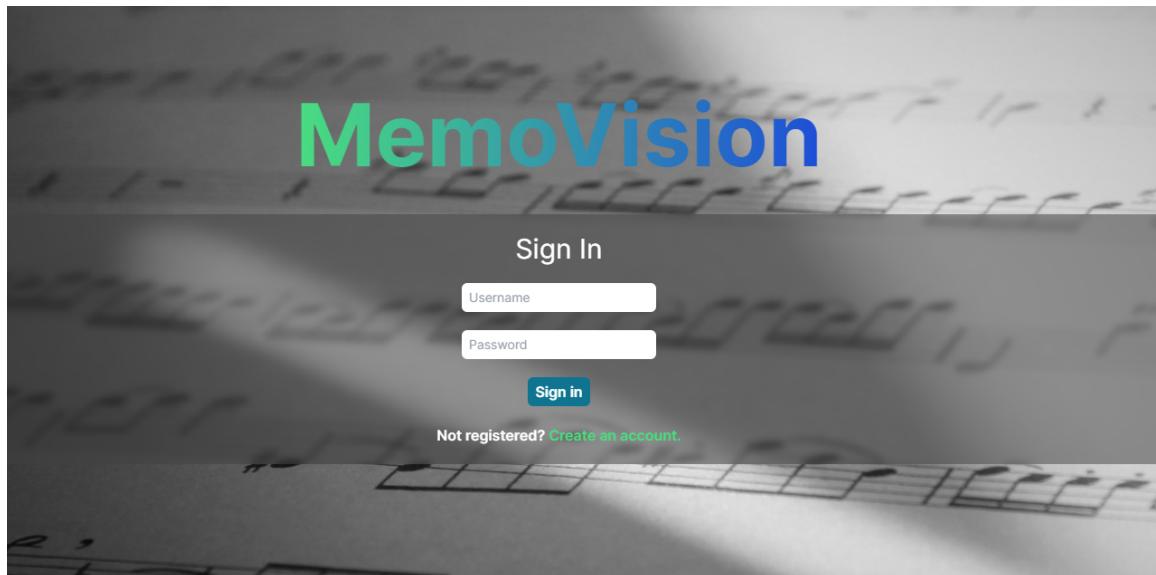
Vizuální stránka softwaru je založena na webovém prostředí, čemuž odpovídá celé uživatelské rozhraní (UI). Uživatel je nejprve vyzván k registraci – vytvoří účet na základě e-mailu a hesla (viz obrázek 4.1). Následně si vytvoří tzv. session (novou instanci projektu) nebo si vybere dříve vytvořenou (viz obrázek 4.2). Uživatel může spouštět jednotlivé funkce a celkovou navigaci v softwaru pomocí myši nebo klávesových zkratek. Software je koncipován modulárně, jeho hlavní funkce jsou rozděleny do jednotlivých modulů, které zprostředkovávají nahrávání dat, výpočty, přehrávání a vizualizaci hudebních nahrávek a úseků. Jde o architekturu umožňující jednodušší rozšíření funkcí v budoucnu. Pokud se do softwaru přidají nové metody nebo druh analýzy, nemusí se měnit základní funkcionality, pouze se přidá modul, ve kterém proběhne separátní výpočet a vizualizace dat. Uživatel může mezi moduly přepínat, případně se vracet na předchozí moduly a měnit referenční nahrávku, přidávat data, nebo upravovat předdefinované sekce (více v dalších kapitolách).

4.1 Moduly

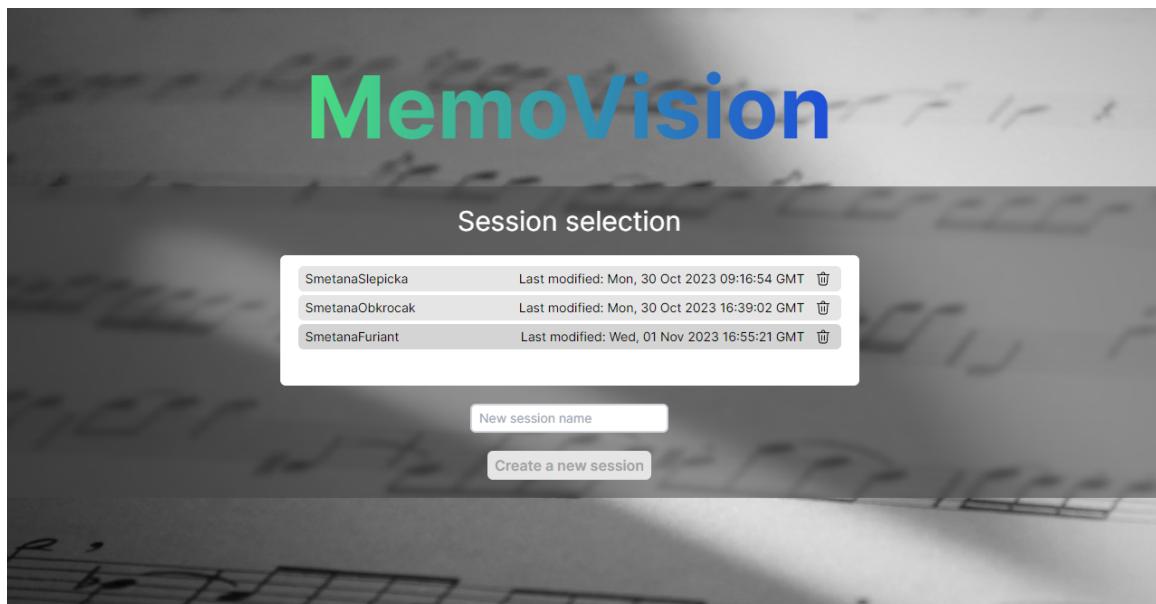
4.1.1 Správce nahrávek

Poté, co máme nahrávky lokálně ve formě audio souborů k dispozici, můžeme je nahrát v prvním modulu *Track manager* (Správce nahrávek), který je zaměřen na nahrání vstupních dat – hudebních nahrávek, metadat a pozic taktů pro referenční nahrávku v textovém .txt nebo .csv souboru ve standardním formátu. Standardním formátem se rozumí sloupcový vektor časových pozic taktů v sekundách ve vzestupném pořadí (každá hodnota je oddělena novým řádkem) nebo dva sloupcové vektory oddělené tabulátorem, kdy je k předchozímu formátu přidán druhý sloupec vyjadřující číselné označení taktu ve vzestupném pořadí (viz kapitola 3.2.3). Uživatel nahraje do programu nahrávky, anotace ve formě časových pozic taktů a označí referenční nahrávku, ke které má k dispozici anotaci ve formě časových pozic taktů. Následně pomocí

¹<https://github.com/stepanmk/memovision> (cit. 6. 12. 2023)



Obr. 4.1: Uvítací stránka softwaru s možnostmi registrace uživatele a přihlášení pomocí uživatelského jména a hesla.



Obr. 4.2: Výběr nebo založení a pojmenování nového projektu (session).

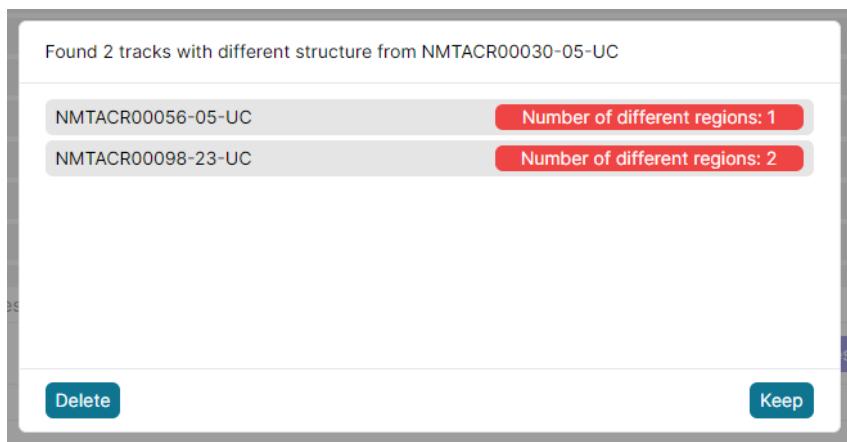
tlačítka Process all tracks (zpracovat všechny nahrávky) provede výpočet chroma vektorů, aktivační funkce detektoru dob, audio-to-audio synchronizace, nalezení duplikátů, převod anotací, zjištění rozdílné hudební struktury, výpočet relevance taktů a odhad dynamiky a tempa (obrázek 4.3).



Obr. 4.3: První modul *Track manager* – nahrávání audio nahrávek, anotací, metadat a možnost přiřazení metadat a binárních štítků a spuštění výpočtu parametrů a synchronizace.

Tento proces může být výpočetně náročný podle délky a počtu nahrávek. Systém uživatele upozorní, pokud došlo k synchronizační chybě včetně informace, o které nahrávky se vůči referenci jedná (viz obrázek 4.4). Uživatel může nahrávky vymazat (Delete) nebo ponechat (Keep). Doporučujeme nahrávky vždy ponechávat a v modulu *Interpretation player* nahrávky vůči referenci zkontovalovat. Může se totiž stát, že software chybně vyhodnotil neshodná místa nebo se nahrávky liší pouze v některých taktech, což nemusí zkreslit analýzu jiných částí kompozice.

V následujícím modulu má možnost si místa chybné synchronizace zobrazit a přehrát. V momentálním nastavení je detektor rozdílné struktury poměrně netolerantní a může označit i místa, kde nedošlo k chybě, pouze se nahrávky výrazně liší v lokálním tempu (detekuje rozdílnou strukturu při zhruba 3x vyšším tempu jedné nahrávky vůči druhé v rámci krátkých časových intervalů). Stejně tak mohou nastat situace, kdy se hudební obsah nahrávek lokálně liší (repetice několika taktů), ale systém rozdíl nezaznamená, jelikož synchronizace může mít v případě nenalezení shodných úseků lineární charakter. Kromě manuální kontroly potenciálně neshodných úseků dále doporučujeme vizuálně zkontovalovat začátky a konce nahrávek včetně očividných chyb v synchronizaci na základě příliš rychlé sekvence taktů za sebou.



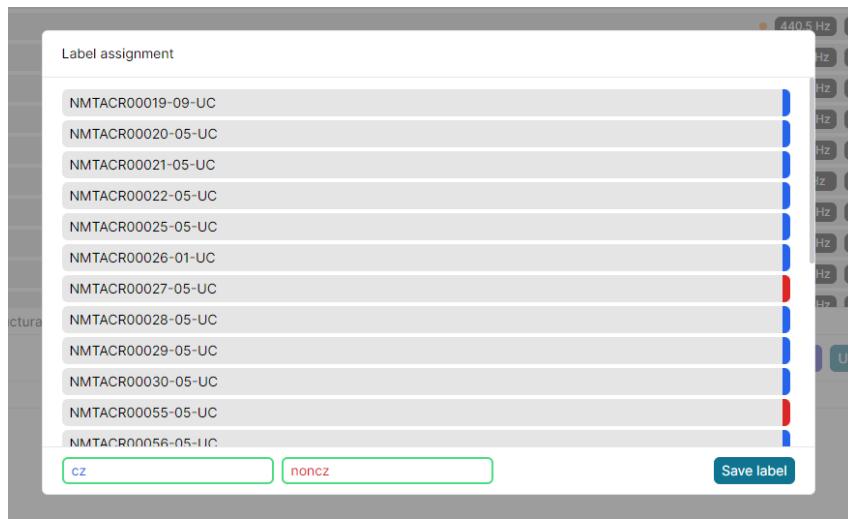
Obr. 4.4: Vyskakovací okno po dokončení výpočtu parametrů a synchronizace – systém upozorní, pokud došlo k detekci potenciálně chybné hudební struktury a označí množství úseků, jejich časovou značku a nahrávku, které se tento problém vůči referenci týká.

Dále může označit tzv. binární štítek (label) pro každou nahrávku pomocí tlačítka *Assign labels* (přiřadit štítky), který zařazuje danou interpretaci do jedné ze dvou možných kategorií podle povahy analýzy. Může se jednat například o kategorii „hudební nahrávka s rokem provedení před 2000“ vs. „hudební nahrávka po roce 2000“ nebo binární rozdělení podle původu interpretů (viz obrázek 4.5). Jedinou podmínkou je, že každá interpretace musí být označena jedním z možných štítků (interně je jedno, o jaký druh štítku se jedná, software zpracovává jen binární hodnoty štítků – jedna kategorie je označena jako 0 a druhá jako 1 bez návaznosti na jejich obsah). Důležitá je jen navzájem se vylučující povaha štítků a existence této informace pro každou nahranou interpretaci. Software navíc automaticky vypočítá scénáře, kdy každá nahrávka je označena jako jedna kategorie a všechny ostatní nahrávky jako kategorie druhá. Toho se následně využívá v následujícím modulu, kdy uživatel může porovnávat rozdíly mezi libovolnou nahrávkou a všemi ostatními. Uživatel pak může štítky měnit podle potřeby a povahy analýzy. Pro některé funkce softwaru (jako třeba vizualizace dynamiky a tempa) štítky nejsou potřeba.

4.1.2 Výběr sekcí

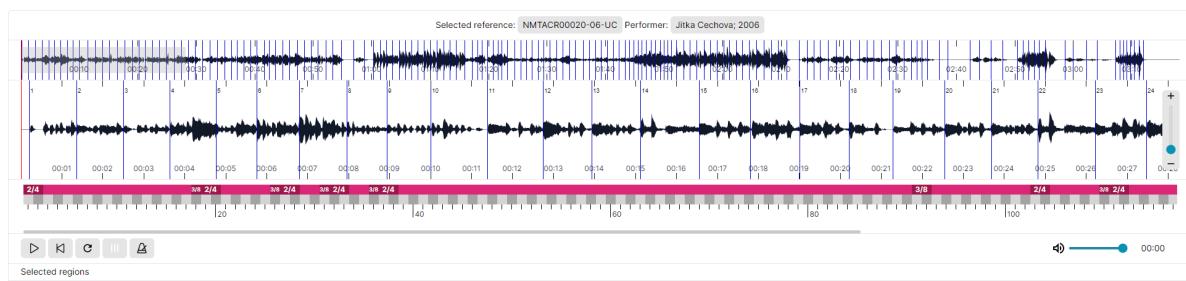
Druhý modul *Section selector* (výběr sekcí/regionů) slouží k první orientaci v hudebním obsahu nahrávek. Uživateli je představeno zobrazení referenční nahrávky včetně pozic taktů a jejich číselné pořadí². Uživatel může poslouchat nahrávku, přiblížovat libovolné sekce, zapínat/vypínat

²Je nutno podotknout, že číselné označení taktů nemusí nutně korespondovat s opravdovým označením v novém zápisu, jelikož software pracuje s anotacemi jako celými čísly. Není možné začínat předtaktím a označit první takt například 0,5, což se v oborech MIR a MPA používá.



Obr. 4.5: Přiřazování binárních štítků (labels) každé nahrávce.

klik v časových okamžicích taktů, měnit jejich zobrazení a zvyšovat hlasitost nahrávky i kliku. Důležitou součástí tohoto modulu je výběr tzv. regionů (sekcí). Jedná se o libovolné sekce referenční nahrávky (definované rozmezím taktů), které uživatele zajímají. Počet sekcí je teoreticky neomezený a jejich význam spočívá v podrobnější analýze a přehlednějším zobrazení vypočítaných dat pro zvolené úseky. Uživatel dále může upravit metrický předpis částí skladby podle notového předpisu. Pokud tak neučiní, software automaticky předpokládá metrický zápis 4/4 na celé skladbě. Manuální korekce není nutná, ale při výpočtu tempa nebude absolutní hodnota odpovídat realitě (smysluplné budou pouze rozdíly mezi interpretacemi), proto je vhodné všechny metrické změny zvýraznit (viz obrázek 4.6).



Obr. 4.6: Druhý modul softwaru *Section selector*. Nad linkou označující takty dané kompozice/nahrávky je znázorněna uživatelem upravená metrická struktura pro správný výpočet tempa.

4.1.3 Přehrávač interpretací

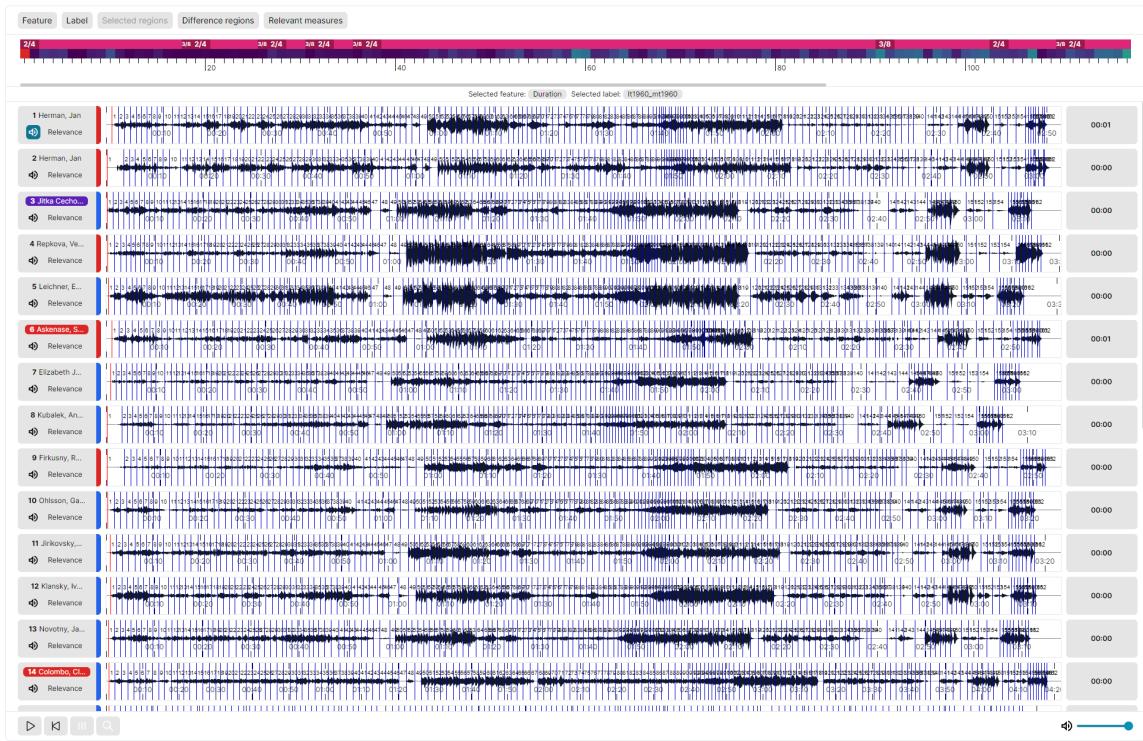
Třetí modul *Interpretation player* (Přehrávač interpretací) je zaměřen na přehrávání všech interpretací, kontrolu správné struktury interpretací, vizualizaci vypočítané relevance taktů a zobrazení dalších informací. V horní části nad nahrávkami je zobrazení tzv. relevantních taktů pomocí barevné škály. Díky metodě MR je z matice délky taktů a rozptylu dynamiky vypočítána relevance každého taktu vůči zadaným štítkům (viz kapitola 3.3.4). Relevance taktů je přepočítána na relativní relevanci a následně znázorněna číslem (0 až 1), pozicí v nahrávce a barevnou škálou. Informace je zobrazena jak graficky, tak numericky pomocí tlačítka *Relevant measures* v horní liště. Libovolné úseky kompozice lze označit pomocí přetažení myši přes vícero taktů uvnitř barevné škály (obrázek 4.7).



Obr. 4.7: Příklad barevné škály a zobrazení modulu *Interpretation player* pro několik nahrávek vybrané kompozice. Modré vertikální čáry označují časové pozice taktů v každé nahrávce podle volby rozsahu taktů na barevné škále.

V prostřední části obrazovky modulu můžeme vidět rozdíly mezi časovou délkou jednotlivých taktů a přehrávat libovolné interpretace (obrázek 4.8). Uživatel si může zvolit nahrávku, která se bude aktivně přehrávat (indikace pomocí zvýrazněného obrázku reproduktoru v levém panelu a lineární změny červené vertikální čáry označující aktuální pozici). Ostatní nahrávky se přehrávají pasivně (vertikální čára se mění nelineárně v závislosti na hudebním obsahu, nikoliv času dané nahrávky). Mezi nahrávkami lze libovolně přepínat, měnit nahrávku, která se aktuálně přehrává, a pohled pro snadnější orientaci.

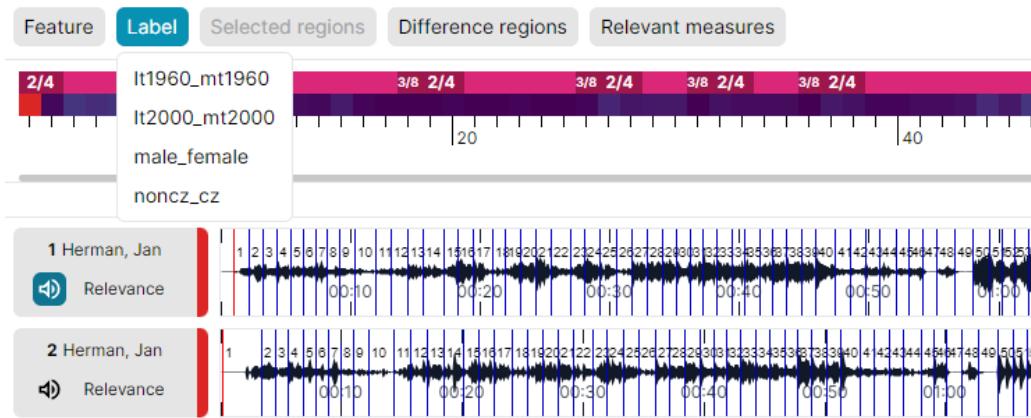
Horní panel umožňuje následující akce: vybrat parametr, podle kterého se počítá relevance a dříve vytvořené štítky (obrázek 4.9), zvýraznit úseky chybné synchronizace (potenciálně odlišný hudební obsah), vybrat definované regiony (obrázek 4.10) a vizualizovat taky s nejvyšší relevancí pro danou kombinaci parametrů a štítků (obrázky 4.11 a 4.12). V levé části lze měnit přehrávanou interpretaci a také označit nahrávku, pro kterou má být zobrazena relevance ve scénáři jedna nahrávka vs. ostatní nahrávky.



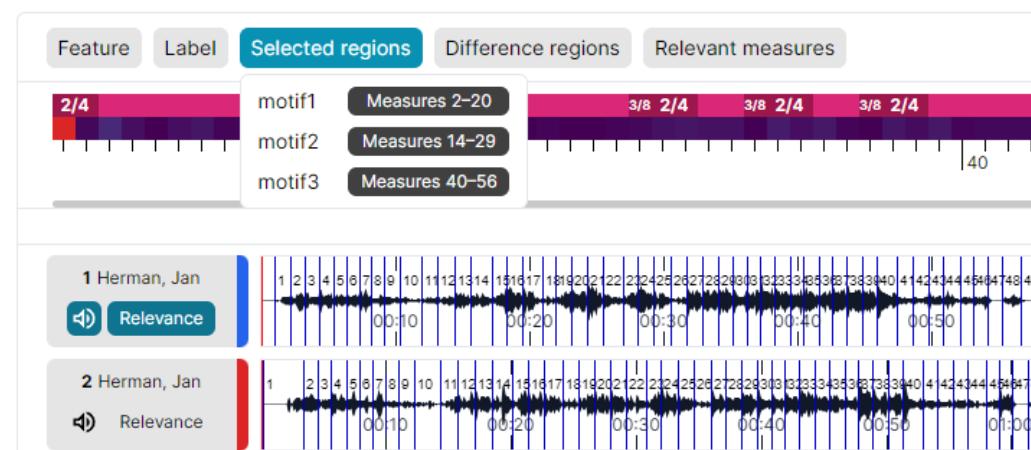
Obr. 4.8: Srovnání mnoha interpretací stejného díla v modulu *Interpretation player*. Červená vertikální čára označuje aktuální pozici přehrávání (zde začátek prvního taktu). Barevný pruh od názvu každé nahrávky zobrazuje aktuální přiřazení interpretace do jedné z binárních tříd.

4.1.4 Vizualizace

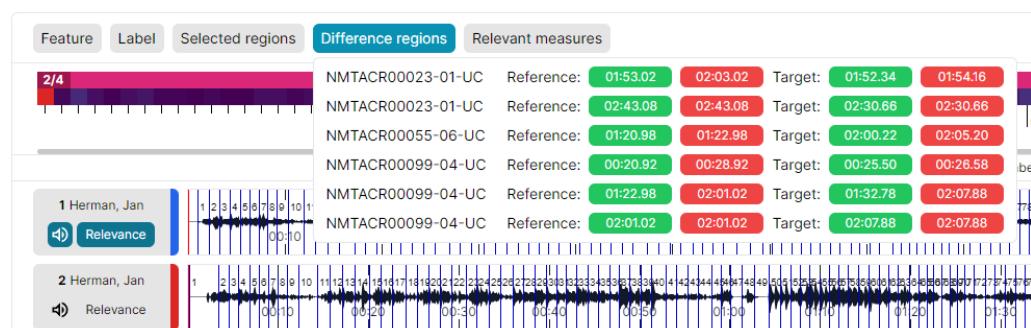
Modul *Visualization* (Vizualizace) umožňuje vizualizaci vypočítaných parametrů. Uživatel může volit nahrávky, u kterých chce parametry zobrazit na levé straně modulu. Stejně jako v předchozím modulu, na horní straně může vybírat úseky taktů nebo předdefinované regiony. Zde je dostupná také volba jednotlivých parametrů (buďto parametry každé nahrávky v čase, nebo statistické srovnání parametrů všech zvolených nahrávek pro osu taktů). Uživatel může porovnávat parametry libovolných interpretací vůči sobě a zároveň přehrávat zvolenou nahrávku včetně vertikálního ukazatele aktuální časové pozice v nahrávce a parametrech (obrázky 4.13 a 4.14). Na obrázku 4.15 je zachycena délka trvání celých nahrávek a jejich rozřazení do kontextu roku nahrávky a zvolené binární třídy, v tomto případě nahrávky před a po roku 2000. Třídu nahrávek a zvolený analyzovaný úsek lze libovolně měnit, a tím dosáhnout variabilní analýzy dle našich výzkumných otázek a cílů. Pro doplnění přidáváme vizualizace tzv. aktivační funkce z námi natrénuvaném detektoru dob (obrázek 4.16). Křivka má lokální maxima (špičky), které by měly znamenat výskyt doby v daném časovém okamžiku.



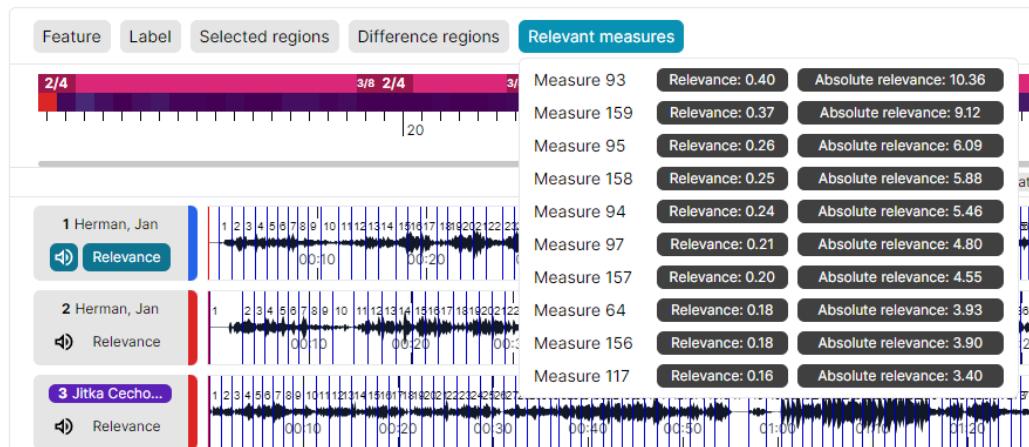
Obr. 4.9: Výběr binárních štítků / třídy pro vizualizaci rozdílů pomocí barevné škály.



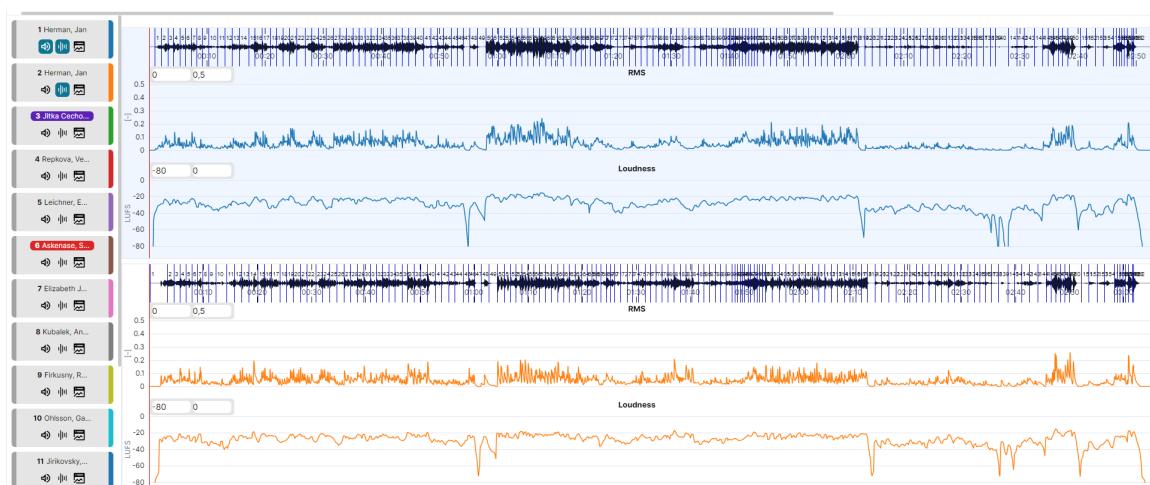
Obr. 4.10: Výběr vlastních úseků na základě předchozího výběru.



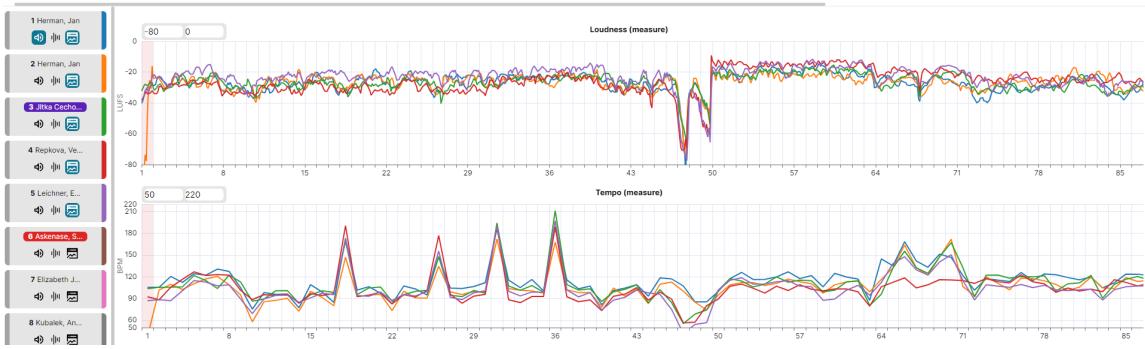
Obr. 4.11: Zobrazení potenciálně chybných synchronizovaných úseků s referenční nahrávkou. Uživatel by měl zkontovalovat vizuálně, případně poslechem, zda tyto úseky odpovídají stejné hudební struktuře.



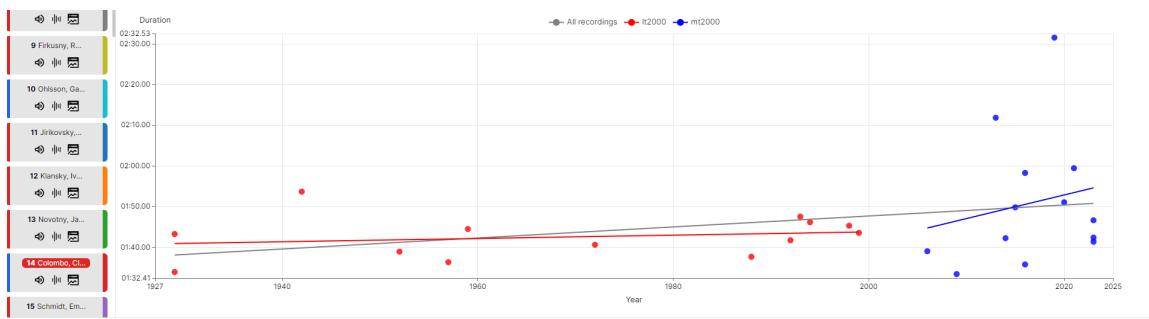
Obr. 4.12: Zobrazení prvních deseti nejrelevantnějších taktů podle vybraného parametru a binárních štítků.



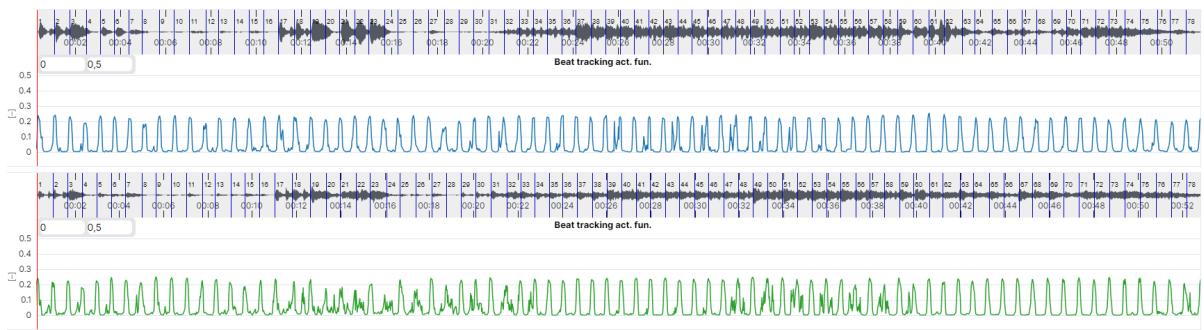
Obr. 4.13: Příklad vizualizace RMS, hlasitosti a aktivační funkce detektoru dob pro dvě interpretace a zvolený úsek kompozice.



Obr. 4.14: Vizualizace srovnání hlasitosti a tempa vybraných nahrávek a úseku kompozice pro jednotlivé takty.



Obr. 4.15: Vizualizace grafu regrese (trendu) podle vybraného parametru: zde pro nahrávky před a po roku 2000.



Obr. 4.16: Příklad srovnání aktivačních funkcí automatického detektoru dob na kratším hudebním úseku – Alban Berg Quartett a Pražákovo kvarteto.

5 Případová studie – konkrétní případ analýzy

V případové studii je na hudebním materiálu ukázána modelová hudebně-interpretační analýza s podporou softwaru. Analýza je strohá a slouží pro ukázku možností diskutované metody a samotného softwaru. Na začátek (kapitola 5.1) je uveden postup tvorby specifické databáze klavírních nahrávek a smyčcových kvartetů, kterou využíváme. Pro každou kompozici jsme vytvořili jednu sadu anotací pro referenční nahrávku (na jejímž výběru příliš nezáleží) ve formě pozic taktů. Ta slouží pro přenos pozic taktů na všechny ostatní nahrávky databáze pro danou skladbu (viz kapitola 3.2.3). Následně komentujeme výsledky a demonstrujeme potenciál poloautomatické hudebně-interpretační analýzy (kapitola 5.2).

5.1 Tvorba databáze

5.1.1 Klavírní díla

Tvorba databáze nahrávek klavírních skladeb sestávala celkem ze tří fází. V první přípravné fázi byla činnost zaměřena na definování relevantních kategorií potřebných metadat a návrh struktury databáze dvořákovsko-smetanovských skladeb. Za tímto účelem byl proveden sběr tzv. testovacích nahrávek. Jako reprezentativní (a početně dostatečně široký) vzorek byla zvolena Chopinova *Nokturna*. K dohledanému penzu nahrávek (celkem 215) byla pořízena tabulka relevantních popisných metadat v předem stanoveném rozsahu, syntaxi a struktuře. Na tomto základě pak byla navržena základní struktura databáze tak, aby byla kompatibilní se standardizovaným knihovnickým popisem zvukových dokumentů za účelem možnosti budoucího exportu či migrace dat. Na základě vlastních nahrávek byla tak určena pole vztahující se k technickým a administrativním metadatům. Nahrávky byly získány z různorodých zdrojů:

- domovské fondy NM,
- digitální archiv partnerské instituce, Supraphon a.s.,
- akvizice původních zvukových nosičů (gramofonové desky, CD),
- soukromých archivů sběratelů,
- nahrávky dostupné open-access v rámci hudebních databází,
- nahrávky dostupné na dalších webových či streamovacích platformách.

Na tomto heterogenním vzorku dat byly rozvinuty základní postupy práce s digitálními objekty, navazující na standardizované a unifikované postupy zavedené v praxi Zvukové laboratoře Národního muzea (NM). V prvé řadě byl k danému zvukovému nosiči nebo tematickému celku (tj. digitálnímu albu, kompliaci nebo jejich části obsahující relevantní skladby) připojen unikátní identifikátor. Ten svou skladbou kombinuje abecední a číselný popis a vypovídá o tom,

kdo provádí digitalizaci, o sbírku jaké instituce se jedná (případně pod jakým projektem digitalizace probíhá), jaká je unikátní pozice daného nosiče v konkrétní evidenční řadě a jaké je pořadí nahrávky v rámci nosiče/nosičů. Příkladem je kód NMTACR00021-01-UC, přičemž obsahově se identifikátor skládá z následujících:

- NM = digitální nahrávky získává Národní muzeum,
- TACR = sběr a digitalizace probíhá v rámci projektu TA ČR,
- 00021 = označuje unikátní číselnou pozici dimenzovanou pro 99 999 pozic,
- 01 = první skladba v rámci nosiče (v případě nutnosti propojit více nosičů je to kombinace dvou pozic, např. 01-01),
- UC = zařazena uživatelské kopie (user copy).

Pro pozdější akvizici tematických nahrávek zaměřených na vybrané dvořákovské a smetanovské klavírní cykly byly použity původní identifikátory zavedené v praxi NM. Důvodem bylo zařazení již dříve digitalizovaných nahrávek (např. gramofonových desek, které byly zpracovány v souvislosti s projekty VISK7; např. NMVISME-3753100152-01-01-UC) nebo zvukových nosičů, jimž byl identifikátor přidělen v rámci interní digitalizační strategie v Českém muzeu hudby (např. NMCMH02695-MF3639-01-UC). Tyto unikátní identifikátory obsahují oproti výše představené struktuře navíc i odkaz na původní identifikační číslo (inventární nebo evidenční číslo; např. MF-3639).

V dalších dvou fázích byla provedena celková rešerše, akvizice a digitalizace nahrávek již se zaměřením na klavírní dílo Bedřicha Smetany a Antonína Dvořáka. Byly dodržovány výše naznačené postupy, přičemž byl kladen důraz na dostupnost zvukových nosičů a nahrávek v domovských institucích (především ve sbírkových fonitech NM-ČMH) a v digitálním archivu partnerské instituce, tj. Supraphon a.s. Jako finální výzkumný vzorek byly vybrány dva rozsáhlejší klavírní cykly obou skladatelů – *České tance I a II*, op. 21 (1877, 1879) od Bedřicha Smetany a *Poetické nálady* op. 85, B161 (1889) od Antonína Dvořáka.

V rámci Smetanova cyklu *České tance*, bylo získáno celkem 308 nahrávek, kdy v celkovém korpusu dat převažují interpretace českých klavíristů a klavíristek (celkem 218). Více než třetina nahrávek (126) pochází ze souborných vydání tohoto klavírního cyklu. Z jednotlivých skladeb má v databázi největší početní zastoupení *Hulán* (celkem 40 různých snímků), následován skladbami *polka a-moll* (34) a *Slepice* (26). Na časové ose lze nahrávky sledovat v tomto početním složení: do roku 1950 (24), 1950–2000 (138), po roce 2000 (146). Nejstarší jsou snímky několika skladeb z konce dvacátých let 20. století (např. snímky Jana Heřmana či Josefa Jiránka pro Pathé z roku 1929).

V případě Dvořákova cyklu *Poetické nálady* byl zařazen méně obsáhlý soubor nahrávek (celkem 210). Oproti *Českým tancům* zde převažují interpretace zahraničních klavíristů a klavíristek (164). Získané zvukové snímky jsou u Dvořákova cyklu *Poetické nálady* jsou z hlediska

doby vzniku mladší, nejsou zde zahrnuty žádné nahrávky před rokem 1950, a pro další období je rozložení následující: 1950–2000 (62), po roce 2000 (148). Nejstarším zvukovým celkem v databázi je souborné vydání cyklu v interpretaci Radoslava Kvapila nahrané roku 1968.

5.1.2 Smyčcové kvartety

Připravený software lze využít rovněž u jiných žánrových a druhových typů či u odlišných nástrojových obsazení. V našem případě jsme volili také smyčcové kvartety, tedy skladby rozsáhléji s širším nástrojovým obsazením. Pro tento případ byla vytvořena databáze nahrávek z kvartetní tvorby českých skladatelů. Její součástí jsou záznamy skladeb Bedřicha Smetany, Antonína Dvořáka a Leoše Janáčka v celkovém počtu 1315 nahrávek. Databáze byla vytvořena pro účely projektu VUT KInG „Identification of the Czech origin of digital music recordings using machine learning“, ze kterého vznikla také publikace specificky analyzující vybrané smyčcové kvartety [21], konkrétně smyčcové kvartety Bedřicha Smetany (*Smyčcové kvartety č. 1 a č. 2*), Antonína Dvořáka (*Smyčcové kvartety č. 12, č. 13 a č. 14*) a Leoše Janáčka (*Smyčcové kvartety č. 1 a č. 2*).

5.2 Analýza smyčcových kvartetů

Důležitým rozhodnutím, které vedlo preferenci domácí kvartetní tvorby, byl rovněž fakt, že tato díla umožňují sestavit snad nejrozsáhlejší databázi zvukových nahrávek různých interpretací. Rovněž z pohledu dějin interpretační praxe domácí hudby se jedná o nezanedbatelnou oblast. Česká kvartetní škola má svou historii a reprezentativní řada zvukových záznamů je toho dokladem. Je jednou z mála domácích uměleckých činností, u které lze sledovat kontinuální vývoj a výraznou interpretační tradici, se vsemi případnými cézurami a odchylkami. Pro jednu z případových studií jsme vybrali Dvořákův *Smyčcový kvartet č. 13 G dur*, op. 106 (1895), na které demonstrujeme popsané metody analýzy a software MemoVision.

Při posouzení interpretace obecně platí, že předpokladem jejího správného poslechu je pochopení struktury díla. Podobně jako interpret, musíme rovněž my rozlišovat mezi významnějšími motivickými a méně důležitými figurativními pasážemi (nesmí přitom být podceněn význam detailu). Přednes a formální analýza nesplývají v jedno. Jestliže se formální analýza zabývá hierarchickou strukturou díla, interpretační analýza hledá logiku zvukového ztvárnění a zprostředkovává posluchači představu formy. Provazuje strukturně významné části, vyhledává a artikuluje kontrastní plochy a nachází vztah k jiným zvukovým realizacím. Různé druhy grafických ztvárnění posuzovaných interpretací nám lépe pomáhají těmto krokům porozumět. V naší předchozí studii analýzy smyčcových kvartetů [21] byl na základě strojového učení

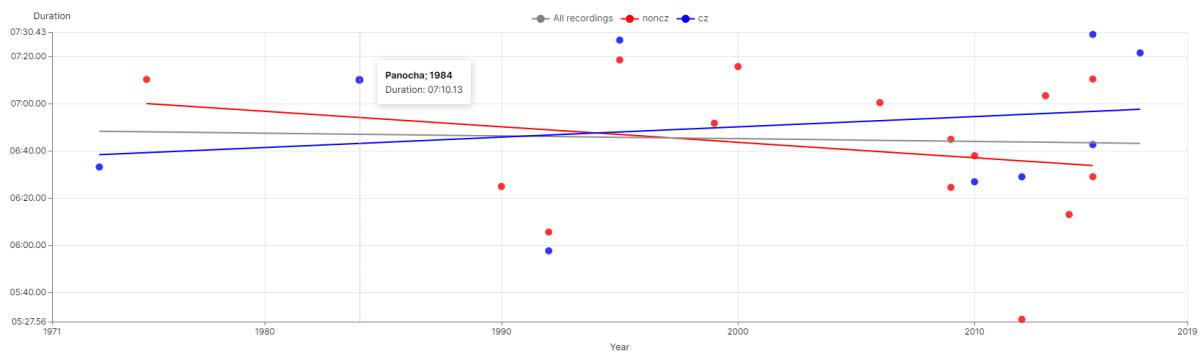
vytvořen klasifikátor, jenž dokáže odlišit nahrávky českých interpretů od zahraničních u vybraných smyčcových kvartetů s relativně vysokou přesností. V některých případech dokáže differencovat rovněž náhodný výběr nahrávek na základě trvání jednotlivých taktů a statistického vyhodnocení. Pokud se zaměříme na náhodnou množinu nahrávek vůči jiné náhodné, navzájem se vylučující množině, můžeme často nalézt společné faktory typické pro každou ze zvolených náhodných množin, a tím je navzájem statisticky odlišit.

Významným pomocníkem při kvantitativní klasifikaci nahrávek je barevné zobrazení relevance taktů průběhu interpretace (viz kapitola 3.4.1 a obrázek 3.8). Takový typ vizualizace prezentuje barevnou škálu (od tmavě fialové přes zelenou až po žlutou), která rychle odhalí potenciálně zajímavá místa ve smyslu rozdílů mezi zvolenými skupinami nahrávek. V našem případě jsme volili binární dvojici *český–nečeský* interpret (spíše než národnost interpretů tato kategorie sleduje příslušnosti k interpretační škole). Přes významnou výpovědní hodnotu grafu je nezbytné upozornit, že prezentovaný typ zobrazení je v mnoha ohledech obecný a hodí se pro objektivní vyhodnocení jen omezených aspektů interpretace. Taková statistika se může výrazně změnit, pokud bychom například do databáze zahrnuli školní či příležitostné výkony anebo pokud na základě obecného zjištění budou interpreti do detailu studovat tempo příslušných míst a přizpůsobovat se, at' již vědomě, či nevědomě pojetí druhé skupiny. Není tedy záznamem individuálního výkonu, ale staví na předem určených binárních štítcích. V uvedeném případě jde o dvojici *český–nečeský*, při vymezení jiných kategorií by bylo barevné rozložení jiné.

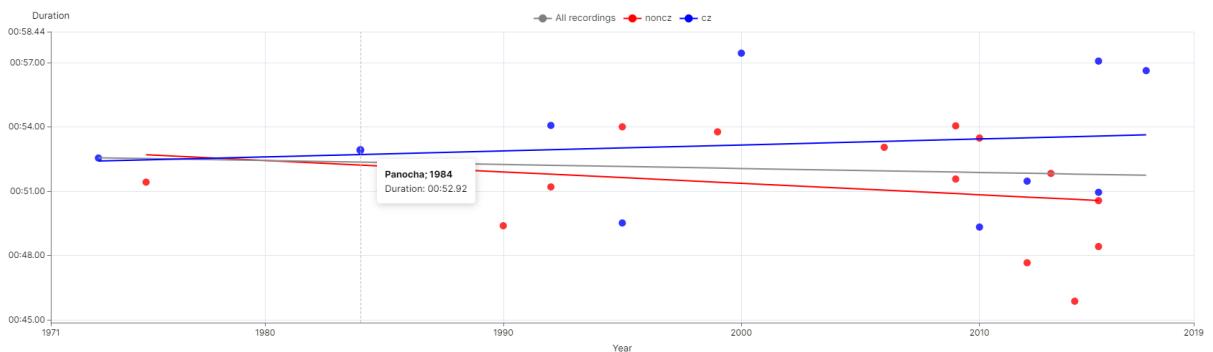
Dvořákův smyčcový kvartet, jenž byl dokončen nedlouho po návratu ze Spojených států na sklonku roku 1895, patří mezi autorova vrcholná díla a celosvětově patří k nejoceňovanějším autorovým skladbám. „Kvartetní symfonie“ jak je někdy kvartet pro svou délku, harmonickou bohatost a propracovanost vedení hlasů nazýván, klade maximální důraz na souhru interpretů. Třídlná 3. věta (Molto vivace) je pak výrazově rozdělena na krajní „úsečné“ až „vzdorovité“ části a střední lyrický oddíl. Časová mapa zobrazuje trvání prvních 78 taktů uvedené části. Na svislé ose je znázorněn čas v sekundách, vodorovná osa ukazuje dataci vzniku nahrávky, přímka uvnitř grafu představuje průměrnou hodnotu trvání úseku u českých (modře) a zahraničních interpretů (červeně) v rámci všech nahrávek naší databáze, které následovaly stejnou hudební strukturu (vůči referenci nedošlo k opakování taktů nebo k chybějící části). Šedá přímka je trend trvání úseku všech nahrávek dohromady. Zde je nutné podotknout, že se jedná o proložení dat pomocí přímky, předpokládáme lineární závislost dat, což ale nelze zaručit – tento předpoklad je často nesprávný, navíc nemáme k dispozici dostatečné množství dat, aby mohli závislost statisticky prokázat. Navíc prokládáme přímku přes celý graf i v letech, kdy žádné nahrávky nemáme. V rámci MPA a oblasti podobných zobrazení je to běžná praxe, musíme však tyto limitace znát při interpretaci výsledků.

Na následujícím obrázku 5.1 je zachycen graf závislosti trvání celé skladby Dvořákova *Smyčcového kvartetu č. 13 G dur*, op. 106, 3. věty, pro nahrávky naší databáze. Modře jsou

reprezentovány české nahrávky, červeně zahraniční. Přímky pak zachycují trend dat, který silně závisí na dostupných nahrávkách a přidání nebo odebrání několika nahrávek může celkový trend výrazně ovlivnit, proto je pouze orientační. Tímto se zároveň pokusíme porozumět tvrzením o dobových tendencích ke zrychlování či zpomalování interpretace. Průkazné výsledky těchto proměn je složité interpretovat při srovnání durata celých vět. Tempo kratších, významově uzavřených celků poměřovaných k trvání celku má většinou vyšší interpretační význam. Způsob vycházející z měření velkých částí či celých vět a výpočet průměrného tempa celé skladby mírá menší význam, neboť takové měření nic nevypovídá o průběhu interpretace jako takové. Obrázek 5.2 ukazuje stejný kvartet a interpret, ale tentokrát je vybrán pouze kratší úsek (takty 1–78).



Obr. 5.1: Trvání celé věty *Smyčcového kvartetu č. 13 G dur*, op. 106, 3. věty, pro jednotlivá kvarteta naší databáze včetně binárního rozdělení dle původu (český–nečeský)



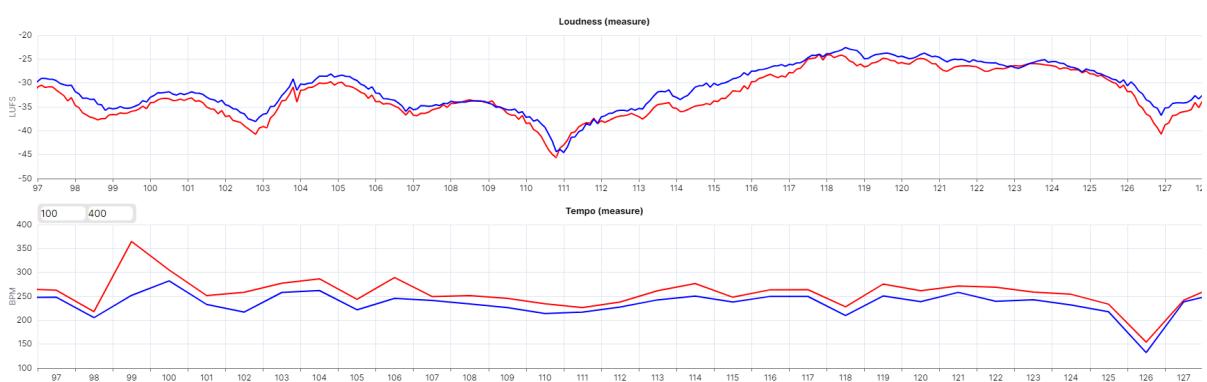
Obr. 5.2: Trvání taktů 1–78 *Smyčcového kvartetu č. 13 G dur*, op. 106, 3. věty, pro jednotlivá kvarteta naší databáze včetně binárního rozdělení dle původu (český–nečeský)

Příklad ukazuje, že snaha po zrychlování interpretace některých vět Dvořákových symfonii, která je dobře slyšitelná především u amerických orchestrů, není v kvartetních skladbách tolik patrná. Domácí soubory (modrá přímka) jdou navíc cestou opačnou a „klasicky“ pojímaný

Dvořák ke zrychlování nevybízí. U Dvořáka je zřejmá snaha hrát úvodní část skladby pomaleji (tento trend platí pro celou větu). Pohled na barevnou škálu (obrázek 5.3) nás upozorňuje na výraznější tempové rozdíly obou skupin v taktech 117–123, které potvrzují rozdílný trend pomalejší interpretace českými soubory rovněž v dalším průběhu skladby. Pro lepší zachycení celé situace jsou v následujícím obrázku 5.4 znázorněny průměrné tempové a dynamické rozdíly mezi českými a zahraničními interprety platné pro širší úsek (takty 97–127). Na rozdíl od grafu tempa, kde dostáváme rozlišení jedna hodnota pro jeden takt, graf dynamiky zobrazuje změny i uvnitř taktů pomocí 10 hodnot na jeden takt.

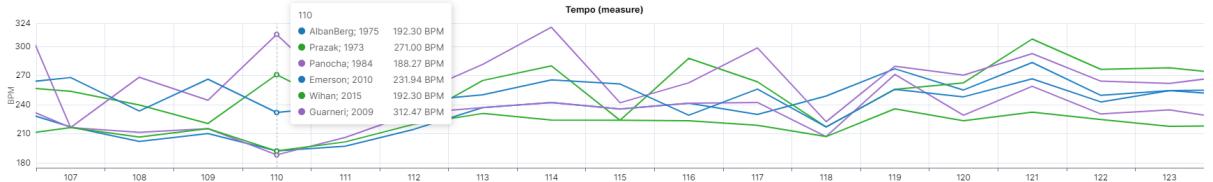


Obr. 5.3: Barevná škála pro vybrané takty Dvořákovova *Smyčcového kvartetu č. 13 G dur*, op. 106, 3. věty; rozdíl mezi třídami český–nečeský.

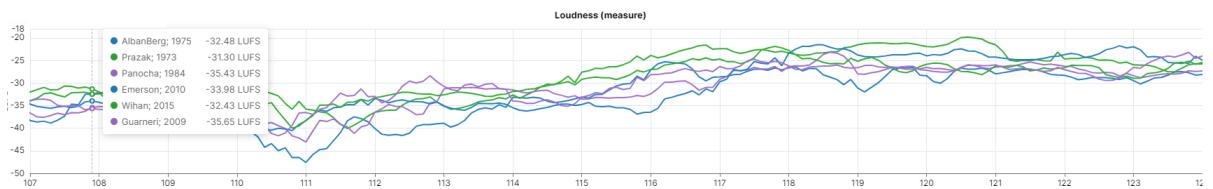


Obr. 5.4: Průměrné tempo a dynamika pro takty 79–127 a třídy český–nečeský u Dvořákovova *Smyčcového kvartetu č. 13 G dur*, op. 106, 3. věty.

Pro další zobrazení jsme vybrali dobově rozdílné nahrávky šesti souborů, tří zahraničních a tří českých: Pražské kvarteto (1973), Alban Berg Quartett (1975), Panochovo kvarteto (1984), Guarneri Quartett (2009), Emerson Quartett (2010) a Wihanovo kvarteto (2015). Obrázek 5.5 představuje vývoj tempa a obrázek 5.6 hlasitost/dynamiku. Rozpětí taktů 107–123 je z pohledu formy mezihrou mezi nástupy tématu střední části věty. Tento celek je formálně uzavřen taktem 126. Sestupný a vzestupný pohyb artikulovaný duolami je některými soubory frázován po dvojicích (111–112, 113–114, 115–116, ...) zohledňující vyšší celek 111–118, 119–126.



Obr. 5.5: Tempo pro vybraná smyčcové kvarteta a takty 107–123 u Dvořákova *Smyčcového kvartetu č. 13 G dur*, op. 106, 3. věty.



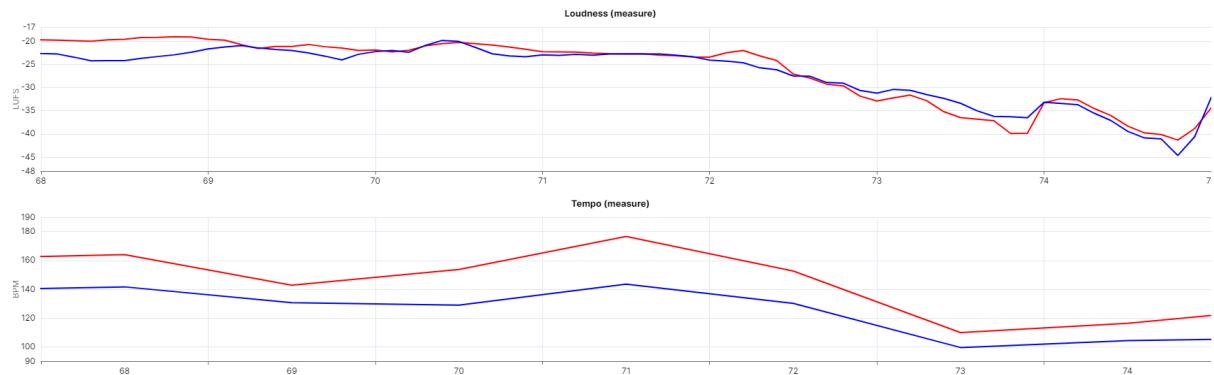
Obr. 5.6: Dynamika pro vybraná smyčcové kvarteta a takty 107–123 u Dvořákova *Smyčcového kvartetu č. 13 G dur*, op. 106, 3. věty.

5.3 Analýza klavírních nahrávek

Smetanův klavírní cyklus *České tance* patří k základním pilířům české klavírní literatury. Proto se zdají být tyto skladby důležitým pramenem zpracování pro software MemoVision, neboť vše, co může při výkladu technického zpracování zvukových nahrávek působit nově a komplikované, je dokumentováno na repertoáru dostatečně známém. Na rozdíl od kvartetních kompozic je struktura klavírních skladeb daleko snazší pro běžné čtení. Jistou nevýhodou je fakt, že *České tance* nepatří k běžným repertoárovým skladbám zahraničních pianistů a rovněž celkový počet dosažitelných nahrávek není tak velký. Také množství nahrávek jednotlivých skladeb obou cyklů se liší, neboť pouze část z nich je součástí kompletu *Českých tanců (I, II.)*. Podobně jako v případě smyčcových kvartetů pracujeme s binární třídou *český-nečeský* (zahraniční).

Pro následující výklad jsme vybrali následující mikrostrukturu (oblast taktů 69–72 s přesahem k taktů 73 a 74). Na obrázku 5.7 je zachyceno průměrné tempo a dynamika tříd *český–nečeský*. Modrá křivka odpovídá českým nahrávkám, červená pak zahraničním. Ze zobrazení je patrné, že zahraniční pianisté hrají takty pomaleji, takt 71 dokonce výrazně pomaleji, než čeští. Takt 70 můžeme považovat za dovršení gradace, jež má svůj počátek v předchozích šestnácti taktech. Obsahuje hned dvě sforzata: mohutný bas a přiznávku na dominantě, v diskantu finální znění nosného motivu *d-e-c* v nejvyšší poloze celé gradace. Motiv začíná na třetí době, proto jeho tempo ovlivňuje oba sledované takty, 69 i 70. Obrázek 5.8 uvádí korespondující strojový notový zápis (bez doplňujícího značení). Smetana ovšem dynamiku stupňuje až do akcentu v taktu 72; akcent je označen šipkou), což lze realizovat jen při dodržení vydavatelem uvedené

pedalizace. Akcent můžeme vnímat i jako tempový vrchol, v kontextu následujícího pokynu *diminuendo molto* v originálním notovém zápisu.



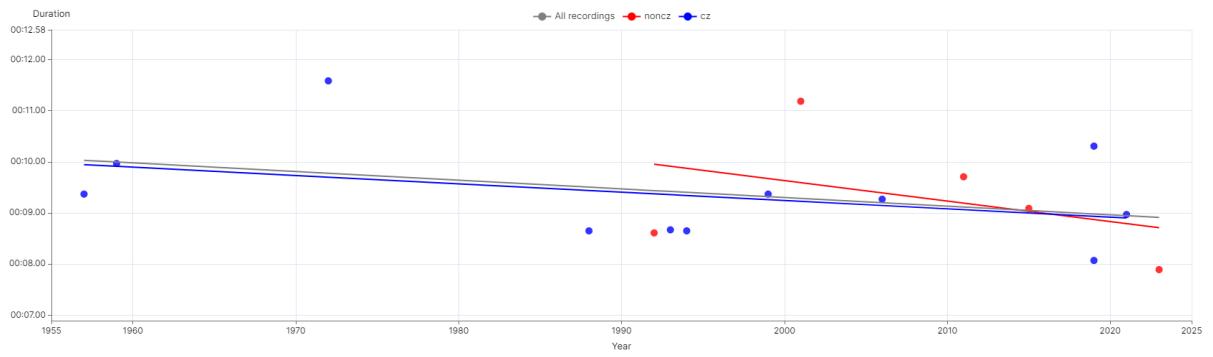
Obr. 5.7: Průměrné tempo a dynamika taktů 68–74 klavírních nahrávek Smetanovy *Sousedské*; skupina českých (modře) a zahraničních (červeně) nahrávek.



Obr. 5.8: Strojový notový zápis taktů 68–76 Smetanovy *Sousedské*.

Obrázek 5.9 ukazuje rozložení nahrávek v rámci let podle časové délky zahraného úseku. Z uvedeného grafu je patrná tendence ke zrychlování interpretace. Zkreslujícím faktorem ovšem může být poměrně malý počet nahrávek. Tendence k výraznějšímu zrychlování u zahraničních interpretů může být již zmiňované sólistické pojetí Smetanových klavírních skladeb.

S pomocí odhadnutého tempa a analýzy všech aspektů jednotlivých nahrávek (korelace tempových a dynamických tendencí, tempo rubato v rámci vrstev, frází, dvoutaktí a motivů, úhlové nuance, pedalizace) docházíme k závěru, že existují dva základní interpretační pohledy na



Obr. 5.9: Délka trvání taktů 68–74 klavírních nahrávek Smetanovy *Sousedské*; skupina českých (modře) a zahraničních (červeně) nahrávek.

vrchol v taktu 70 resp. 72. Reprezentativní počet českých pianistů pojímá takt 70 včetně předtaktí jako vrchol instrumentační, z hlediska tektoniky průběžný, subjektivně ještě jako součást tanečního veselí. Většina zahraničních interpretů hraje podle našich analýz vrchol v taktu 70 rétoričtěji, šířejí a více zatěžkaně, finální *crescendo* motivu pojímá spíše „beethovenovsky“.

Důvodem rozdílů může být odlišná imaginární instrumentace. Češi se od dětství setkávají se Smetanovou hudbou, přinejmenším *Má vlast* a *Prodaná nevěsta* patří k základnímu korpusu národní hudby. Široká znalost Smetanovy hudby tak snáze přesáhuje ryze klavírní pojetí. Plynulost tempové křivky tomu odpovídá, zatímco náhlé rétorické rozšíření u zahraničních interpretů prozrazuje spíše představu sólistickou. Z hlediska „harmonické chemie“, řečeno slovy Claudia Debussyho, se na sforzatované dominantě (t. 70) hudba sice „otevírá“, střední triolová vrstva však nadále pokračuje ve svém směřování k akcentu v taktu 72, a teprve v „koloraturním“ vzetupu (*molto diminuendo*, t. 72, 73) uvozuje nové téma dalšího dílu skladby. V tomto formovém švu (t. 69–74) se setkávají zrychlená varianta tance s pomalejší. Součástí interpretační tradice je tyto nepsané tempové vztahy ctít. Nelze přitom předpokládat, že by klavírní interpreti běžně disponovali speciálními znalostmi v etnomuzikologické a taneční oblasti, ale pravděpodobně tempové zákonitosti buďto intuitivně vycítí, nebo je empiricky přebírají v liniích interpret–posluchač a učitel–žák.

Závěr

Předložená metodika uvádí do problematiky hudebně-interpretační analýzy a prezentuje nové možnosti analýzy na základě softwaru MemoVision. V kapitole 1 a 2 jsou nejprve představeny cíle a kontext metodiky a následně popsána stručná historie hudebně-interpretační analýzy. Vzhledem k charakteru projektu nejsou tyto pasáže rozpracovány vyčerpávajícím způsobem, proto v textu odkazujeme na další publikace a autory, kteří se zkoumanou problematikou zabývají podrobněji. Kapitola 3 probírá parametry hudebních nahrávek, které lze strojově získat a zároveň jsou relevantní pro výzkum a porovnávání interpretačních výkonů. Rozebírá také druh a formát vstupních dat, které uživatel pro analýzu potřebuje, včetně nastínění hlavních metod softwaru. Kapitola 4 popisuje jednotlivé části (moduly) softwaru a prezentuje, jak lze dosáhnout cílů hudebně-interpretační analýzy pomocí nástroje MemoVision. V závěrečné části (kapitola 5), obsahující dvě krátké případové studie, představujeme konkrétní možnosti využití softwaru včetně procesu získání hudebních dat. Pro tuto příležitost jsme vybrali repertoár domácích skladatelů (Bedřich Smetana a Antonín Dvořák) s odlišným nástrojovým obsazením (smyčcový kvartet a klavír).

Nejen zájemci z řad interpretů či hudebních teoretiků, ale také širší odborná i laická veřejnost může ocenit možnost zobrazit a porovnat mnohé interpretace a zaujmout k nim kvalifikované stanovisko. Vizualizace průběhu tempa a dynamiky a jejich synchronizace s ostatními zvukovými záznamy téžé skladby umožňuje srovnání interpretací na základě předem vybraných charakteristik. Ve vizualizaci dat lze nacházet výraz a ospravedlňovat jej jako interpretační styl doby. Lze tak určit proměny ve hře jedince či generací v odlišných časových perspektivách, pokud získáme reprezentativní vzorek nahrávek a správně využijeme potenciál funkcí softwaru – například relevanci taktů pro dvě odlišné skupiny nahrávek na základě kvalitativního rozdělení. Můžeme se rovněž zamýšlet nad pravdivostí tvrzení o adekvátním a jediném správném podání národního repertoáru.

Je zřejmé, že v době nástupu nových technologií v IT a pronikání počítačové vědy do jednotlivých společenských a humanitních oborů nemůže zůstat ani muzikologie stát stranou. V tomto ohledu nabízí nástroj MemoVision reálnou možnost propojení oboru s širšími možnostmi zpracování zvukových nahrávek. K samotné analýze interpretačního výkonu musíme ovšem přistupovat poučeně, s vědomím limitů, které použité výpočetní metody mají. V rámci péče o kulturní dědictví představuje tento nástroj perspektivu hlubšího poznání, jež zůstává běžným lidských schopnostem ukryta. Proměna vůči minulým přístupům s sebou přináší další možnosti a snahu po novém definování základních pojmu z oblasti péče o kulturní statky. Představený software však nemusí být pouze pomocníkem při hudební analýze, další možností je jeho použití pro rychlé odhalení duplikátů, tedy totožných hudebních interpretací původně prezentovaných pod různými metadaty, nebo rychlou orientaci v různých verzích téžé hudební skladby.

Seznam literatury a zdrojů

- [1] LÉBL, Vladimír a POLEDŇÁK, Ivan. *Hudební věda: historie a teorie oboru, jeho světový a český vývoj: celostátní vysokoškolská učebnice pro studující filozofických fakult univerzit, pedagogických fakult a akademíí muzických umění*. Díl 2., část 1., Disciplíny hudební vědy. Učebnice pro vysoké školy. Praha: Státní pedagogické nakladatelství, 1988.
- [2] CORTOT, Alfred. *Principes Rationnels de la Technique Pianistique*, Francie, Paříž: Salabert, 1928.
- [3] HOFFMEISTER, Karel. *Vývoj klavírní virtuosity*. Praha: Knihovna Hudební výchovy, s. 53, 1911.
- [4] KUNDERA, Ludvík. *Příspěvky k analýze hudební reprodukce*. Disertační práce. Brno: Filosofická fakulta, 1927.
- [5] ZICH, Jaroslav. *Prostředky výkonného hudebního umění*. Hudební rozpravy. Praha: Státní nakladatelství krásné literatury, 1959.
- [6] ZICH, Jaroslav. *Kapitoly a studie z hudební estetiky*. Praha: Supraphon, 1975.
- [7] KUNA, Milan a BLÁHA, Miloš. *Čas a hudba: k dramaturgi časových prostředků v hudebně interpretačním výkonu*. Studie ČSAV. Praha: Academia, 1982.
- [8] PODGORNYÝ, Oleg. *Rozšířené metody interpretační analýzy a některé zákonitosti hudební interpretace*. Diplomová práce. Vedoucím práce byl Josef Bek. Praha, FF UK, 1980.
- [9] RINK, John. *The practice of performance: studies in musical interpretation*. London, UK: Cambridge University Press, Royal Holloway, 1995.
- [10] Bowen, Jose A. Tempo, Duration and Flexibility: Techniques in the Analysis of Performance. *Journal of Musicological Research*, roč. 16, č. 2, 1996. s. 1–47.
- [11] COOK, Nicholas. *Beyond the score: music as performance*. Oxford: Oxford University Press, 2013. s. 480. ISBN 978-01-993-5740-6.
- [12] CANNAM, Chris, LANDONE, Christian, SANDLER, Mark B. a BELLO, Juan Pablo. *The Sonic Visualiser: A Visualisation Platform for Semantic Descriptors from Musical Signals*. In Proceedings of the International Society for Music Information Retrieval Conference (ISMIR 2006). Victoria, Canada, 2006. s. 324–327.

- [13] LERCH, Alexander, ARTHUR, Claire, PATI, Ashis a GURURANI, Siddharth. An Interdisciplinary Review of Music Performance Analysis. [online] *Transactions of International Society for Music Information Retrieval (TISMIR)*, roč. 3, č. 1, s. 221–245, 2020. Dostupné z: <https://doi.org/10.5334/tismir.53>
- [14] STEINMETZ, Christian J. Steinmetz a REISS, Joshua D. pyloudnorm: *A simple yet flexible loudness meter in python*. [online] 150th AES Convention, Queen Mary University of London, London, UK. 2021. Dostupné z: <https://www.eecs.qmul.ac.uk/~josh/documents/2021/21076.pdf>
- [15] PINTO, António, DOMINGUES, Inês a DAVIES, Matthew E. P. *Shift If You Can: Counting and Visualising Correction Operations for Beat Tracking Evaluation*. In Extended Abstracts for the Late-Breaking Demo Session of the 21st Int. Society for Music Information Retrieval Conf. (ISMIR 2020), 2020. [online] Dostupné z: https://program.ismir2020.net/lbd_421.html
- [16] IŠTVÁNEK, Matěj a MIKLÁNEK, Štěpán. *Exploring the Possibilities of Automated Annotation of Classical Music with Abrupt Tempo Changes*. In Proceedings II of the 28th Conference STUDENT EEICT 2022 Selected Papers. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, 2022. s. 286–290. ISBN: 978-80-214-6030-0.
- [17] MÜLLER, Meinard. *Information Retrieval for Music and Motion*. Springer, 2007. ISBN 978-3-540-74047-6. DOI 10.1007/978-3-540-74048-3.
- [18] MÜLLER, Meinard, KURTH, Frank a CLAUSEN, Michael. *Chroma-Based Statistical Audio Features for Audio Matching*. In Proceedings of the Workshop on Applications of Signal Processing (WASPAA), USA, s. 275–278, 2005.
- [19] MÜLLER, Meinard a EWERT, Sebastian. *Chroma Toolbox: MATLAB implementations for extracting variants of chroma-based audio features*. In Proceedings of the International Conference on Music Information Retrieval (ISMIR 2011), Miami, Florida, USA, s. 215–220, 2011.
- [20] PRÄTZLICH, Thomas, DRIEGER, Jonathan a MÜLLER, Meinard. *Memory-restricted multiscale dynamic time warping*. In Proceedings of the 41st IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2016), Shanghai, China, s. 20–25, 2016.

- [21] IŠTVÁNEK, Matěj, MIKLÁNEK, Štěpán a SPURNÝ, Lubomír. Classification of interpretation differences in string quartets based on the origin of performers. *Applied Sciences*, roč. 3, č. 6, s. 185–205, 2023. [online] Dostupné z: <https://doi.org/10.3390/app13063603>
- [22] PENG, Hanchuan, LONG, Fuhui a DING, Chris. Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Trans. Pattern Anal. Mach. Intell.*, roč. 27, č. 8, s. 1226–1238, 2005. [online]. Dostupné z: <https://doi.org/10.1109/TPAMI.2005.159>
- [23] ZHAO, Zhenyu, ANAND, Radhika a WANG, Mallory. *Maximum relevance and minimum redundancy feature selection methods for a marketing machine learning platform*. In Proceedings of the 6th IEEE International Conference on Data Science and Advanced Analytics (DSAA 2019), Washington DC, USA, 2019.