

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

Fakulta elektrotechniky
a komunikačních technologií

BAKALÁŘSKÁ PRÁCE



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH TECHNOLOGIÍ

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION

ÚSTAV TELEKOMUNIKACÍ

DEPARTMENT OF TELECOMMUNICATIONS

NÁSTROJ PRO SIMULTÁNNÍ PŘEHRÁVÁNÍ VÍCE INTERPRETACÍ JEDNÉ SKLADBY

A TOOL FOR SIMULTANEOUS PLAYBACK OF MULTIPLE COMPOSITION INTERPRETATIONS

BAKALÁŘSKÁ PRÁCE

BACHELOR'S THESIS

AUTOR PRÁCE

AUTHOR

Michael Švejcar

VEDOUCÍ PRÁCE

SUPERVISOR

Ing. Štěpán Miklánek

BRNO 2020

Bakalářská práce

bakalářský studijní program **Audio inženýrství**
specializace Zvuková produkce a nahrávání
Ústav telekomunikací

Student: Michael Švejcar

ID: 203749

Ročník: 3

Akademický rok: 2019/20

NÁZEV TÉMATU:

Nástroj pro simultánní přehrávání více interpretací jedné skladby

POKYNY PRO VYPRACOVÁNÍ:

Interpretace skladeb klasické hudby se liší ve změnách lokálního tempa a také v celkové délce trvání daného hudebního provedení. Student v rámci této práce navrhne a naprogramuje nástroj, který bude umožňovat simultánní přehrávání více interpretací jedné skladby. Nástroj by měl pracovat s libovolnou sadou interpretací skladby z žánru klasické hudby, přičemž nejmenší možný počet interpretací dané skladby je dvě. Uživatel by měl být schopen v reálném čase volit, která interpretace skladby je právě přehrávána. Přehrávání bude vhodné graficky znázornit kurzory na časových osách. Pod sebou budou zobrazeny jednotlivé časové osy daných interpretací. Předpokladem je využití metody dynamického borcení časové osy, která bude sloužit k synchronizaci jednotlivých interpretací.

DOPORUČENÁ LITERATURA:

[1] MÜLLER, Meinard. Fundamentals of music processing. New York, NY: Springer Berlin Heidelberg, 2015. ISBN 9783319219448.

[2] KONZ, Verena. Automated Methods for Audio-Based Music Analysis with Applications to Musicology. Saarland, 2012. Disertace. Saarland University.

Termín zadání: 3.2.2020

Termín odevzdání: 8.6.2020

Vedoucí práce: Ing. Štěpán Miklánek

doc. Ing. Jiří Schimmel, Ph.D.
předseda rady studijního programu

UPOZORNĚNÍ:

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

ABSTRAKT

Cílem této bakalářské práce bylo vytvoření programu umožňujícího simultánní přehrávání více interpretací jedné skladby a plynulého přepínání mezi nimi. K těmto účelům byl použit App Designer v prostředí MATLAB, jenž je určen k vývoji aplikací s grafickým uživatelským rozhraním. Pro vývoj aplikace bylo klíčové zejména použití dostupných toolboxů spolu s algoritmy pro výpočet víceúrovňového dynamického borcení časové osy. Výsledný přehrávač IntSwitcher umožňuje uživateli načíst dvě nahrávky interpretací jedné skladby. Ze vstupních souborů jsou nejprve vypočteny chromagramy, které jednotlivé nahrávky charakterizují z hlediska tónového vývoje v čase. Následně je na chromagramy aplikována metoda víceúrovňového dynamického borcení časové osy, jejímž výstupem je optimální cesta. Jedná se o matici, ve které jsou k sobě přiřazeny hudebně korespondující vzorky načtených audio souborů s rozlišením 50 ms. Dle této závislosti je během přehrávání určována korespondující časová pozice aktuálně neaktivní stopy, společně s pozicí jejího posuvníku. V případě přepnutí přehrávané nahrávky tedy začne druhá stopa hrát ve stejné části skladby, i pokud se tato část nachází v rozdílných časech jednotlivých nahrávek. Výsledná aplikace se jeví jako užitečný nástroj ke studování rozdílů mezi různými interpretacemi stejného hudebního díla.

KLÍČOVÁ SLOVA

chromagram, dynamické borcení časové osy, získávání informací z hudby, audio přehrávač, synchronizace, podobnost, interpretace

ABSTRACT

The purpose of this Bachelor's thesis was to create a piece of software which enables the user to simultaneously play back multiple interpretations of a musical piece and switch between them instantaneously. This was achieved using the App Designer in the MATLAB environment, which is intended for developing applications with graphical user interface. The key to the development of the application was especially the use of available toolboxes and algorithms for computing chromagrams and multiscale dynamic time warping. The final IntSwitcher player enables the user to load two recordings of interpretations of one song. Chromagrams which characterize the individual recordings in terms of tonal development over time are first calculated from the input files. After that, the multiscale dynamic time warping method is applied on the chromagrams, which outputs the warping path. The warping path in this case is a matrix, in which musically corresponding samples of loaded audio files are assigned together with the resolution of 50 ms. From this, the corresponding time position of currently inactive track is computed along with its slider position. If the user switches the currently played recording, the second track starts playing in the same part of composition, even if that part is at a different time in each of the individual recordings. The final software is an appropriate tool for studying differences between various interpretations of the same musical piece.

KEYWORDS

chromagram, dynamic time warping, music information retrieval, audio player, synchronization, similarity, interpretation

ŠVEJCAR, Michael. *Nástroj pro simultánní přehrávání více interpretací jedné skladby*. Brno, 2020, 43 s. Bakalářská práce. Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav telekomunikací. Vedoucí práce: Ing. Štěpán Miklánek

Vysázeno pomocí balíčku `thesis` verze 3.05; <http://latex.feec.vutbr.cz>

PROHLÁŠENÍ

Prohlašuji, že svou bakalářskou práci na téma „Nástroj pro simultánní přehrávání více interpretací jedné skladby“ jsem vypracoval samostatně pod vedením vedoucího bakalářské práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou všechny citovány v práci a uvedeny v seznamu literatury na konci práce.

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

Brno

.....

podpis autora

PODĚKOVÁNÍ

Rád bych poděkoval vedoucímu bakalářské práce panu Ing. Štěpánu Miklánkovi za odborné vedení, konzultace, trpělivost a podnětné návrhy k práci.

Obsah

Úvod	10
1 Zvuk a jeho základní vlastnosti	11
1.1 Rychlost šíření zvuku	11
1.2 Zvuková vlna	12
1.3 Amplituda	12
1.4 Frekvence	13
1.4.1 Alikvótní tóny	13
1.4.2 Barva zvuku	14
2 Digitalizace zvuku	15
2.1 Vzorkování	15
2.2 Kvantování	16
2.3 Finální digitální soubor	16
3 Proces synchronizace dvou nahrávek	17
3.1 Fourierova řada	17
3.2 Diskrétní Fourierova řada	18
3.3 Diskrétní Fourierova transformace	18
3.4 Rychlá Fourierova transformace	19
3.5 Spektrogram	19
3.6 Chromagram	20
3.6.1 Zjištění tónového obsahu	21
3.6.2 Zjištění chroma-tónového obsahu (Parametr CP)	21
3.6.3 Kompenzace logaritmického vnímání zvuku (parametr CLP)	21
3.6.4 Normalizace a kvantizace vektorů (parametr CENS)	22
3.6.5 Potlačení témbrových složek zvuku (parametr CRP)	22
3.7 Metoda dynamického borcení časové osy	22
3.7.1 Monotónnost	24
3.7.2 Kontinuita	24
3.7.3 Začátek a konec optimální cesty	24
3.7.4 Oblast optimální cesty	24
3.7.5 Omezení strmosti	24
3.7.6 Víceúrovňové dynamické borcení časové osy	25
4 Výsledky dosažené v zahraničí	26

5	Popis programu IntSwitcher	28
5.1	Informační panel	29
5.2	Tlačítka „Nahrát soubor“	29
5.3	Tlačítka „Analyzovat“	29
5.4	Tlačítka „Přehrát/pauza“	30
5.5	Časové posuvníky	31
5.6	Počítadla času	31
5.7	Globální tlačítka „Přehrát/pauza“	31
5.8	Globální tlačítka „Přepnout“	31
5.9	Globální tlačítka „Zastavit“	32
5.10	Posuvník „Hlasitost“	32
5.11	Menu programu	32
	5.11.1 Nastavení	32
	5.11.2 O programu	35
6	Testování programu IntSwitcher	36
	Závěr	37
	Literatura	38
	Seznam symbolů, veličin a zkratk	41
	Seznam příloh	42
A	Přiložené soubory	43

Seznam obrázků

1.1	Zvuková vlna.	12
1.2	Amplituda zvukové vlny.	12
1.3	Fundament a vyšší harmonické složky periodického signálu.	14
3.1	Spektrogram perkusivní smyčky	19
3.2	Chromagram	20
3.3	Přiřazení jednotlivých elementů posloupností pomocí DTW.	23
3.4	Optimální cesta mezi dvěma chromagramy.	23
4.1	International Audio Laboratories Erlangen: Interpretation Switcher	26
4.2	The Interpretation Switcher.	27
4.3	The Sheet Music Interface.	27
5.1	Schéma programu Intswitcher	28
5.2	Tlačítko nastavení	33
5.3	Velikost bufferu	34
5.4	Barva pozadí	34
5.5	Smazat analyzované chromagramy	35
5.6	O programu	35

Úvod

Při záznamu více nahrávek stejné skladby zahráných různými hudebními tělesy nebudou výsledné zvukové nahrávky nikdy totožné – budou se lišit zejména v celkové délce jejich trvání, v tempu, v dynamice přednesu, v instrumentaci i tónovém paletě použitých nástrojů.

Cílem této práce je vytvořit program audio přehrávače, který uživateli umožní přehrávat více interpretací jedné skladby a plynule mezi nimi přepínat v reálném čase tak, aby se zvuková stopa po přepnutí spustila od korespondující části skladby. Tento program by mohl být vhodným pomocníkem zejména pro učitele a studenty muzikologie, kde by sloužil ke kritickému poslechu více interpretací, jejich srovnávání a následnému posuzování významných či drobných rozdílů mezi nimi. V případě použití klasického audio přehrávače je toto znesnadněno nutností manuálního hledání stejné části skladby po přepnutí nahrávky.

V práci jsou nejprve popsány základní vlastnosti zvuku a procesy týkající se převodu zvukového záznamu do digitální podoby, na které je dále navázáno objasněním metod analýzy zvukových nahrávek z hlediska jejich frekvenčního a tónového vývoje v čase. Dále je popsána synchronizace interpretací pomocí metody dynamického borcení časové osy, na jejímž základě je postavena funkčnost přehrávače. Následuje kapitola věnovaná programu IntSwitcher, který je výsledkem praktické části této práce. V té jsou popsány jeho jednotlivé ovládací prvky a nastíněny klíčové algoritmy. K vývoji bylo z důvodu dostupnosti potřebných toolboxů zvoleno interaktivní programové prostředí MATLAB vytvořené společností Mathworks. Poslední kapitola je věnovaná testování funkčnosti výsledného programu.

1 Zvuk a jeho základní vlastnosti

Zvuk je definován jako mechanické vlnění, které se šíří v látkovém prostředí, a je schopné vyvolat smyslový vjem stimulací sluchových orgánů. Jinými slovy, jde o velmi rychlé změny atmosférického tlaku, které se formou vibrací přenášejí mezi jednotlivými molekulami. Tyto rychlé změny atmosférického tlaku nazýváme tlakem akustickým, který je na atmosférický tlak superponován. Věda, která se zabývá zvukem od jeho vzniku, přes šíření až po jeho subjektivní vnímání, se nazývá akustika [1].

Zvuk vzniká všude, kde dochází k vibracím, které se mohou ve svém prostředí dále šířit – například lusknutí prstů, drnknutí o strunu kytary i tření kol vlakové soupravy o koleje. Může být rovněž vytvořen umělým způsobem na syntezátoru a následně přehrán pomocí reproduktorů.

1.1 Rychlost šíření zvuku

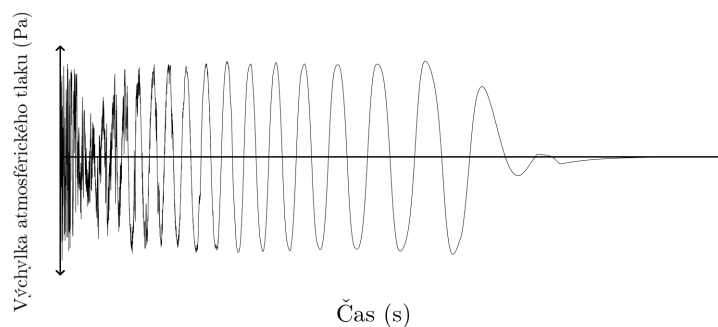
Rychlost šíření zvuku c (m/s) není konstantní, ale závisí na podmínkách prostředí, ve kterém se zvuk šíří. Při výpočtech, u kterých se předpokládá, že jde o šíření vzduchem a není třeba přespříliš velké přesnosti, se obvykle pracuje s hodnotou $c = 340$ m/s. Pokud je hodnota rychlosti šíření zvuku ve vzduchu potřeba vyjádřit přesněji, lze pro její výpočet využít vztah

$$c = 331,57 + 0,607\vartheta \quad (\text{m/s}), \quad (1.1)$$

kde ϑ je teplota ve °C.

1.2 Zvuková vlna

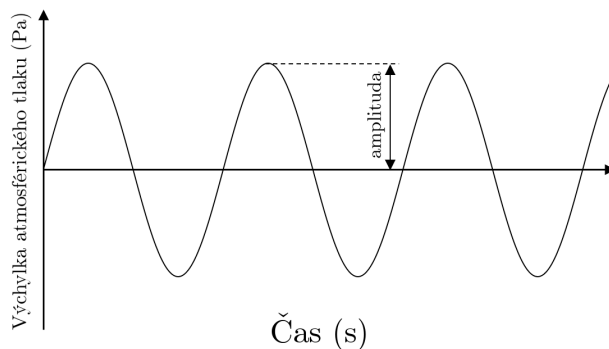
Zvuková vlna představuje průběh výchylky atmosférického tlaku v daném bodě prostoru v čase – jedná se tedy o grafické znázornění vývoje zvuku. Princip zvukové vlny je vhodné využít všude, kde je třeba analyzovat základní zvukové vlastnosti. Zejména digitální audio nahrávky bývají zapisovány ve formě navzorkované zvukové vlny. Proces digitalizace zvuku bude popsán v následujících kapitolách.



Obr. 1.1: Zvuková vlna.

1.3 Amplituda

Jedním ze základních parametrů popisujících zvukové vlny je amplituda, někdy také nazývaná rozkmit. Jedná se o maximální hodnotu tohoto signálu v čase, tudíž bod, ve kterém je výchylka atmosférického tlaku vlivem šíření zvuku nejvyšší nebo naopak nejnižší.



Obr. 1.2: Amplituda zvukové vlny.

1.4 Frekvence

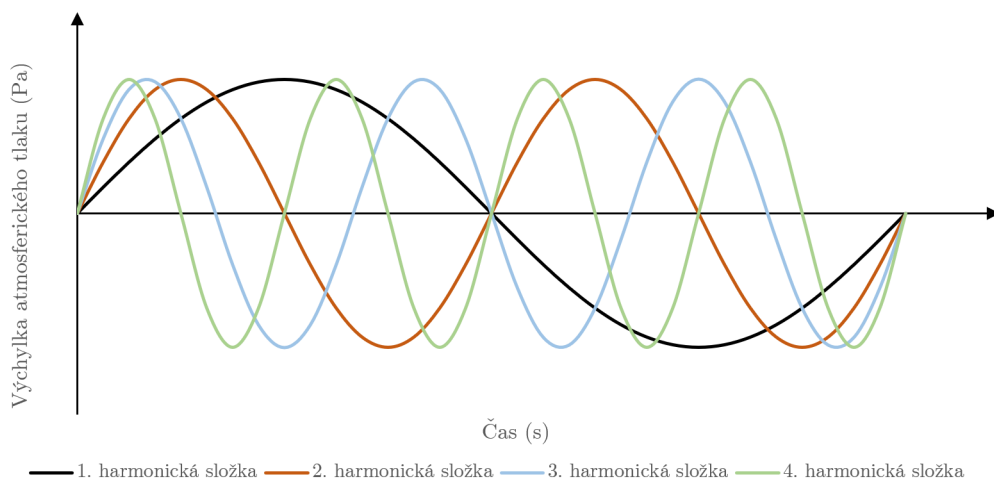
Dalším důležitým parametrem popisujícím zvukové vlny je frekvence, jinými slovy kmitočet. Obecně je definovaná jako počet opakování periodického děje za jednotku času, obvykle jednu sekundu. Její jednotkou je Hz (hertz) a převrácenou hodnotu frekvence nazýváme časovou periodou.

V případě zvuku frekvence určuje rychlost kolísání atmosférického tlaku. Slyšitelné spektrum frekvencí u člověka je přibližně 20 Hz až 20 kHz. Vzhledem k tomu, že sluchový orgán každého člověka je na různé frekvence jinak citlivý a spodní i horní hranice slyšitelnosti se mění s věkem i časem stráveném v hlučném prostředí, je tento rozsah u každého jednotlivce individuální. Frekvence nižší, než je spodní hranice slyšitelnosti, se nazývají infrazvuk, naopak frekvence vyšší než hranice horní slyšitelnosti nazýváme ultrazvuk.

Pokud je zvuková vlna periodická, jedná se o tón. Pomocí frekvence lze určit absolutní výšku tónu, oproti tomu subjektivně vnímaná výška tónu má jednotku mel a není s frekvencí vždy totožná [2].

1.4.1 Alikvótní tóny

V přírodě se zřídka objevují tóny pouze o jedné frekvenci. Ty lze vytvořit převážně umělým způsobem. Tóny přirozené se ovšem skládají ze součtu vícero kmitočtů. Pokud například hráč na kytaru zahraje tón A4 o frekvenci $f = 440$ Hz, rozezná se nejen tento kmitočet, nýbrž i kmitočty odpovídající jeho celočíselným násobkům, tudíž $f = 880$ Hz, $f = 1320$ Hz a tak dále, viz obr. 1.3. Nejnižší frekvence tohoto tónu se nazývá fundament. Ten určuje výšku celého tónu a jeho intenzita bývá zpravidla největší. Vyšší vzniklé frekvenční složky se nazývají jako alikvótní tóny, někdy také vyšší harmonické. Jejich intenzita zpravidla klesá s postupně narůstající frekvencí, alikvótní tóny pohybující se výše ve frekvenčním spektru jsou tedy méně výrazné, nežli alikvóty frekvenčně blíže fundamentu. Bylo dokázáno, že pokud v tónu chybí fundament, ale vyskytují se v něm jeho vyšší harmonické složky, fundament je i přesto možné slyšet, neboť lidské vnímání zvuku absenci fundamentu kompenzuje [3].



Obr. 1.3: Fundament a vyšší harmonické složky periodického signálu.

1.4.2 Barva zvuku

Pro zjištění informací o frekvenčním spektru signálu se používá rozklad pomocí Fourierovy transformace [4]. Podle toho, které alikvóty se při zaznění tónu rozezní a podle poměrů jejich hlasitostí se určuje barva zvuku, neboli tímbr. Díky barvě zvuku je možné rozeznat původ vzniku zvuku, v praxi je tím umožněno například rozeznat housle od fagotu, i pokud je na oba nástroje zahrán stejný tón.

Pro popis různých tímbrů se zažily pojmy často vycházející ze smyslových oblastí. Například tóny s vysokým obsahem vyšších harmonických se obvykle označují jako světlé či průzračné, naproti tomu tóny s nízkým obsahem alikvót bývají označovány jako tóny tupé či tmavé [5]. Používání různých barev zvuku má uplatnění zejména v hudbě, neboť jsou schopny ve člověku vyvolávat rozmanité pocity. Jejich vnímání je samozřejmě značně subjektivní a dá se objektivizovat jen do určité míry.

2 Digitalizace zvuku

Aby bylo možné se zvukem jakýmkoli způsobem pracovat na počítači, je třeba ho nejprve zaznamenat pomocí mikrofону, převést do digitální podoby a posléze zakódovat do vhodného souborového formátu. V následujících podkapitolách budou popsány procesy vzorkování a kvantování analogového zvukového signálu, které jsou při digitalizaci nevyhnutelné. K těmto účelům slouží analogově číslicový převodník, který je klíčovým prvkem zvukových rozhraní.

2.1 Vzorkování

První ze dvou fází transformace signálu spojitého na signál diskrétní je vzorkování. Analogový signál je spojitý v čase, tudíž v kterémkoli jeho časovém úseku se nachází nekonečný počet funkčních hodnot. Zápis takového signálu v digitální podobě by vyžadoval nekonečně velkou paměť i výpočetní výkon, proto je nezbytné jeho převedení na konečný počet bodů vyjádřených jejich funkčními hodnotami. Počet vzorků v časovém úseku vyjadřuje vzorkovací frekvence f_{vz} .

Dříve než začne být zpracováván signál vzorkován, je třeba z něj odstranit frekvenční složky, jejichž frekvence je vyšší, než polovina vzorkovacího kmitočtu. Tento fakt vychází z Shannonova teorému (někdy také nazývaného jako Nyquistův teorém, Kotělnikovův teorém či Shannonův-Nyquistův teorém), který říká, že pokud chceme docílit přesné rekonstrukce spojitého, frekvenčně omezeného harmonického signálu, je třeba tento signál vzorkovat s frekvencí, která je alespoň dvakrát vyšší, nežli nejvyšší frekvence signálu zpracovávaného [4]. Shannonův teorém lze tedy vyjádřit vztahem

$$f_{vz} > 2f_{\max} \quad (\text{Hz}), \quad (2.1)$$

kde f_{vz} je vzorkovací frekvence a f_{\max} nejvyšší frekvence zpracovávaného signálu.

Pokud by Shannonův teorém nebyl při vzorkování dodržen, docházelo by k nená vratnému jevu nazývanému aliasing, vlivem kterého by byl výsledný navzorkovaný signál oproti signálu původnímu značně zkreslen. V takovém případě se ve spektru výsledného signálu objeví frekvenční složky, které v originální nahrávce nebyly přítomny. V praxi se obvykle z důvodu vytvoření určité rezervy a snížení nároků na strmost antialiasingových filtrů používá vzorkovací kmitočet o něco vyšší, než je dvojnásobek nejvyšší frekvence zpracovávaného signálu. Například u záznamu na CD se používá vzorkovací frekvence 44,1 kHz i přesto, že průměrný člověk se zdravým sluchem slyší pouze do cca 20 kHz.

2.2 Kvantování

Protože počítače umí zapisovat čísla pouze s omezenou přesností určenou počtem bitů, kterým je číslo vyjádřené, je třeba hodnoty jednotlivých vzorků signálu zaokrouhlit. Tento proces se nazývá kvantování. Během kvantování jsou hodnoty vzorků signálu porovnávány s referenčními hodnotami, někdy také nazývanými jako kvantizační hladiny, přičemž hodnota každého vzorku je zaokrouhlena na její nejbližší kvantizační hladinu. Protože je výsledný signál zpracováván ve dvojkové soustavě, počet kvantizačních hladin analogově číslicových převodníků je roven hodnotě 2^N , kde N odpovídá počtu bitů bitové hloubky nakvantovaného signálu.

Čím menší počet kvantizačních hladin je při kvantování použit, tím dochází k většímu kvantizačnímu zkreslení, neboť míra zaokrouhlení jednotlivých hodnot je větší. Vždy se tedy jedná o ztrátový proces, během kterého dochází k nenávratným změnám. Vlivem kvantizačního zkreslení vzniká kvantizační šum, který je roven odchylkám nakvantovaných hodnot od původních funkčních hodnot signálu. Míra vzniklého kvantizačního šumu se obvykle vyjadřuje v decibelech jako SNR – „signal to noise ratio“. Jedná se o odstup užitečného signálu od šumu, který udává i výsledný dynamický rozsah nahrávky [6].

Kvantizační šum, který obvykle zní jako nepříjemné zkreslení, se do jisté míry dá maskovat procesem zvaným dithering. Ten je vhodné použít všude, kde dochází ke snižování bitové hloubky nahrávky. Vlivem ditheringu bývá do nahrávky před kvantizací přimíchán nízkoúrovňový šum, který způsobí, že výsledný kvantizační šum je více náhodný, více rozptýlený ve frekvenčním spektru a méně závislý na frekvenčním obsahu kvantované nahrávky. Nevýhodou tohoto procesu je drobné zvýšení prahu šumu. Existuje několik metod ditheringu, od jednodušších až po více pokročilé, u kterých je frekvenční spektrum šumu filtrováno dle psychoakustických modelů [7].

2.3 Finální digitální soubor

Program IntSwitcher, jenž je výsledným produktem praktické části této práce, pracuje s digitálními nahrávkami zakódovanými pomocí pulzně kódové modulace (PCM) do bezztrátového audio formátu `.wav`.

3 Proces synchronizace dvou nahrávek

Synchronizace časových os dvou nahrávek s rozdílnými globálními i lokálními tempy a různou dynamikou je komplexní proces využívající několikero různých algoritmů. V této kapitole budou popsány jednotlivé procesy, které výsledné synchronizaci předcházejí. Jako první je nezbytné jednotlivé nahrávky vhodným způsobem analyzovat z hlediska jejich frekvenčního, případně i tónového vývoje v čase.

3.1 Fourierova řada

Fourierova řada, pojmenovaná po francouzském matematikovi a fyzikovi Josephu Fourierovi, vychází z poznatku, že jakýkoli periodický signál lze rozložit na sumu nekonečného počtu signálů tvaru sinus či kosinus s vhodně zvolenými frekvencemi, amplitudami a počátečními fázemi.

Tvarů Fourierovy řady existuje několik, přičemž první tvar Fourierovy řady se vypočítá dle

$$s(t) = \frac{a_0}{2} + \sum_{k=1}^{\infty} a_k \cos(k\omega_1 t) + b_k \sin(k\omega_1 t). \quad (3.1)$$

$$a_0 = \frac{2}{T_1} \int_0^{T_1} s(t) dt, \quad (3.2)$$

$$a_k = \frac{2}{T_1} \int_0^{T_1} s(t) \cos(k\omega_1 t) dt, \quad (3.3)$$

$$b_k = \frac{2}{T_1} \int_0^{T_1} s(t) \sin(k\omega_1 t) dt, \quad (3.4)$$

kde $k = 1, 2, 3, 4, \dots$, $\omega_1 = \frac{2\pi}{T_1}$ je úhlový kmitočet a T_1 perioda periodického signálu $s(t)$.

3.2 Diskrétní Fourierova řada

Zatímco použitím klasické Fourierovy řady je rozkládán spojitý signál, pomocí diskrétní Fourierovy řady je rozkládán signál nespojitý s omezeným počtem vzorků. Oblast jejího využití je tedy zejména ve zpracování digitálních audio signálů, například v prostředí MATLAB.

Pomocí diskrétní Fourierovy řady je periodické posloupnosti $\tilde{s}[n]$ s periodou N přiřazen obraz $\tilde{S}[k]$, periodická posloupnost s periodou N :

$$\tilde{S}[k] = \sum_{n=0}^{N-1} \tilde{s}[n] e^{-jk \frac{2\pi}{N} n}, \quad k = 0, 1, \dots, N-1, \quad (3.5)$$

kde vlnovka vyjadřuje periodicitu [4].

3.3 Diskrétní Fourierova transformace

Diskrétní Fourierova transformace (dále jen DFT) vychází z principů Fourierovy řady a používá se pro zjištění spektrálního obsahu signálu s konečným počtem vzorků. Tento signál lze vyjádřit jako konečnou posloupnost $s[n]$ délky N , které je pomocí DFT přiřazen obraz $S[k]$, jenž je také konečnou posloupností s délkou N .

Při výpočtu se nejprve posloupnosti $s[n]$ s délkou N přiřadí periodická posloupnost $\tilde{s}[n]$ s periodou N :

$$\tilde{s}[n] = s[\text{mod}_n(n)], \quad (3.6)$$

pomocí které se určí obraz diskrétní Fourierovy řady – periodická posloupnost $\tilde{S}[k]$ dle:

$$\tilde{S}[k] = \sum_{n=0}^{N-1} \tilde{s}[n] e^{-jk \frac{2\pi}{N} n}, \quad k = 0, 1, \dots, N-1, \quad (3.7)$$

ze které se vybere jedna perioda o délce N :

$$S[k] = R_N[k] \cdot \tilde{S}[k]. \quad (3.8)$$

3.4 Rychlá Fourierova transformace

Rychlá Fourierova transformace (dále jen FFT) je algoritmus umožňující rychlý výpočet DFT. Zatímco výpočet diskrétní posloupnosti o N vzorcích pomocí DFT vyžaduje řádově N^2 operací, k výpočtu té samé posloupnosti pomocí FFT je třeba řádově pouze $2/N$ operací.

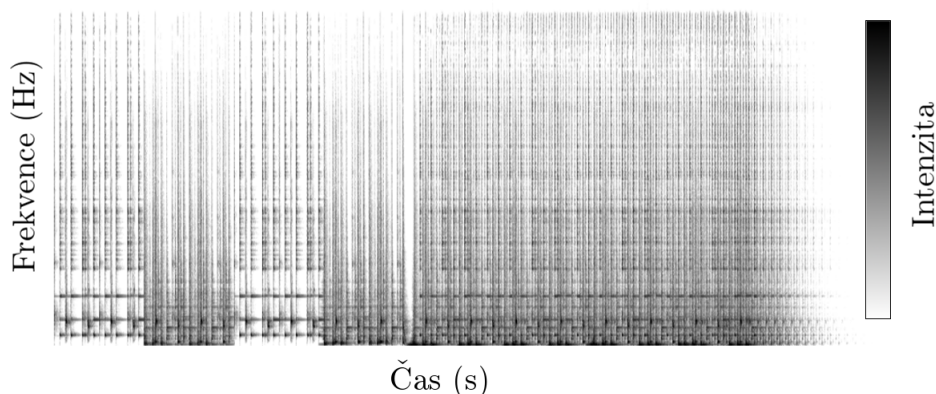
Metod pro výpočet FFT existuje několik, přičemž nejznámější a nejpoužívanější je „*Cooley-Tukey algoritmus*“ [10], který byl popularizován již v šedesátých letech minulého století [4, 8, 9].

3.5 Spektrogram

Spektrogram je jedním ze základních nástrojů pro frekvenčně-časovou analýzu zvuku. Jedná se o vizuální reprezentaci frekvenčního spektra v čase, která může být znázorněna dvourozměrně či třírozměrně. V prvním případě reprezentuje svislá osa frekvenci, vodorovná osa čas a intenzita je znázorněna barevným odstínem. V případě trojrozměrného zobrazení spektrogramu vyjadřuje intenzitu třetí osa.

Výpočet spektrogramu bývá nejčastěji realizován pomocí FFT. Digitální zvukový záznam se v časové oblasti rozdělí do krátkých rámců, pro které je následně pomocí FFT vypočítané frekvenční spektrum. Jednotlivé segmenty jsou poté ve výsledném spektrogramu seřazeny vedle sebe tak, že každý z nich formou sloupce pixelů znázorňuje frekvenční spektrum daného časového úseku.

Volba délky časového okna závisí na potřebné přesnosti výsledného spektrogramu – kratší okno znamená lepší rozlišení času, ale horší rozlišení frekvence, naopak delší časové okno způsobí větší rozlišení frekvence, ale menší rozlišení času [11].



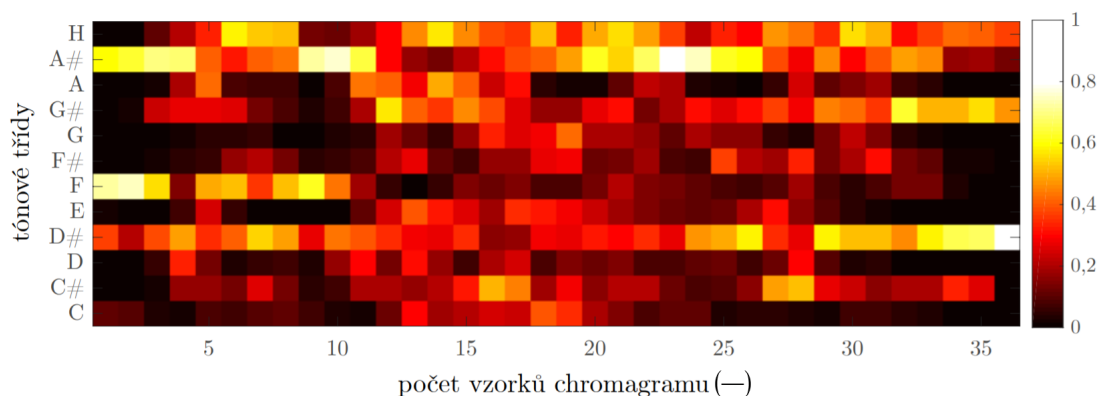
Obr. 3.1: Spektrogram perkusivní smyčky.

3.6 Chromagram

Jako vhodné řešení pro analýzu tónového obsahu hudebních děl se jeví vytvoření a následné použití tzv. chromagramů, jež reprezentují jednotlivé záznamy na základě jejich tónového vývoje v čase, nezávisle na dynamice či změnách lokálního tempa. Princip chromagramů vychází z poznatku, že člověk vnímá tóny, jejichž frekvence se liší o oktávu, jako tóny dokonale souznějící. V hudební teorii se název těchto tónů vyjadřuje stejným písmenem, lišícím se pouze o přídatné jméno určující jejich polohu (velké C (65,4 Hz), malé C (130,8 Hz)...). Díky tomuto lze výšku tónu popsat dvěma způsoby. Prvním způsobem je zápis pomocí frekvence první harmonické složky neboli fundamentu, druhá možnost je zapsání tónu pomocí vlastnosti chroma, která udává, o který tón se jedná, nezávisle na jeho poloze. Západní hudba vychází z temperovaného ladění, kde je oktáva rozdělena na dvanáct půltónů (půltón = nejmenší vzdálenost mezi dvěma tóny na klavíru). Tato stupnice je nazývána stupnicí chromatickou. Chromagram tedy popisuje vývoj celkové energie jednotlivých tónů chromatické stupnice v čase. Součet dílčích energií jednoho tónu ve všech oktávách je roven hodnotě energie jednoho z dvanácti atributů vektoru v daném časovém intervalu.

Chromagramy jsou využívány jak v oblastech synchronizace a srovnávání hudebních děl, tak i pro analýzu a následné studování struktury skladeb. Dále se používají pro identifikaci názvu skladeb, čehož lze využít například u cover verzí k rozpoznání, o kterou píseň se jedná [12].

Chromagramy lze vytvořit různými způsoby, přičemž se nedá obecně říci, který ze způsobů je nejlepší, neboť vhodnost jejich použití závisí na vstupním materiálu i na jejich následném zpracování. V následujících podkapitolách bude uveden způsob získání jednotlivých parametrů potřebných k sestavení chromagramu pomocí implementace z Chroma Toolboxu pro MATLAB.



Obr. 3.2: Chromagram.

3.6.1 Zjištění tónového obsahu

Prvním krokem pro vytvoření chromagramu je zjištění tónového obsahu nahrávky. Toho je docíleno tak, že se zvukový signál pomocí sady filtrů rozdělí do 88 frekvenčních subpásem, jejichž střední frekvence odpovídají frekvencím výšek tónů A0 až C8. Vzhledem k tomu, že ne každá nahrávka je laděná naprosto stejně (obvykle se ladí podle $A4 = 440$ Hz), zavádí se do procesu také algoritmy, které dokáží případné frekvenční posuny rozpoznat a vykompenzovat posunem středních frekvencí filtrů pro jednotlivá subpásma. Aby bylo dosaženo dostatečně velkého spektrálního rozlišení pro spodní frekvence, je třeba použít buďto nízký vzorkovací kmitočet, nebo dlouhé časové okno. K dosažení dostatečně velkého spektrálního rozlišení v celém slyšitelném pásmu je použit vzorkovací kmitočet $f_{vz} = 22,05$ kHz pro vysoké kmitočty, $f_{vz} = 4,41$ kHz pro střední frekvence a $f_{vz} = 882$ Hz pro kmitočty nízké [15]. Následně se pro každé z 88 subpásem vypočítají jejich efektivní hodnoty s fixní velikostí časových oken, které mají 50% překryv. Výsledkem je zjištění charakteru krátkodobé energie každého z 88 frekvenčních subpásem.

3.6.2 Zjištění chroma-tónového obsahu (Parametr CP)

V dalším kroku se zjistí tónový obsah nahrávky vyjádřený pomocí vlastností chroma. Toho lze dosáhnout tak, že se u stejných časových segmentů nahrávky sečtou efektivní hodnoty subpásem se stejnými chroma vlastnostmi. Jinými slovy, jsou sečteny všechny efektivní hodnoty náležící subpásmům, které odpovídají stejným notám, lišícím se pouze o oktávu nebo její celočíselné násobky. Díky tomuto se pro každý časový segment získá dvanácti dimenzionální vektor $x = (x(1), x(2), \dots, x(12))^T$, kde $x(1)$ odpovídá chroma vlastnosti vyjadřující notu C, $x(2)$ odpovídá chroma vlastnosti vyjadřující notu C# (cis) apod.

3.6.3 Kompenzace logaritmického vnímání zvuku (parametr CLP)

Vzhledem k tomu, že závislost lidmi vnímané hlasitosti zvuku na jeho intenzitě není lineární, ale přibližně logaritmická [14], je vhodné do výpočtu chroma parametrů zahrnout logaritmickou amplitudovou kompresi. Každá hodnota energie e náležící jednotlivým segmentům je nahrazena za hodnotu $\log(\eta \cdot e + 1)$, kde η je vhodná pozitivní konstanta.

3.6.4 Normalizace a kvantizace vektorů (parametr CENS)

Přidáním míry abstrakce použitím statistiky na energetické rozložení jednotlivých chroma pásem se docílí zjištěním parametru CENS, který je velmi užitečný ve srovnávání a rozpoznávání nahrávek. Při výpočtu parametru CENS je nejdříve každý chroma vektor normalizován podle ℓ^1 -normy, poté se aplikuje kvantizace s vhodně zvolenými prahy. V následujícím kroku se tyto vlastnosti dále vyhlazují pomocí okna délky $w \in \mathbb{N}$ a podvzorkují činitelem d . Výsledné parametry jsou normalizované podle ℓ^2 -normy a označované jako CENS_d^w .

3.6.5 Potlačení témbrových složek zvuku (parametr CRP)

K potlačení vlivu témbrových složek jednotlivých tónů nahrávky na výsledný chromagram lze využít parametru CRP. Toho je dosaženo tak, že se po zjištění tónového obsahu a aplikaci kompenzace logaritmického vnímání zvuku (parametr CLP) výsledná zlogaritmovaná tónová reprezentace transformuje pomocí diskrétní kosinové transformace. V dalším kroku se odstraní získané spodní tónově-frekvenční kepstrální koeficienty (PFCCs), neboť bylo vyzorováno, že souvisejí s barvou zvuku [16, 17]. Horní tónově-frekvenční kepstrální koeficienty se ponechají netknuty. Po následné aplikaci zpětné diskrétní kosinové transformace se na jednotlivé vektory vyjadřující efektivní hodnoty časových segmentů subpásem použije metoda zjištění chroma-tónového obsahu (parametr CP), díky které se získají výsledné dvanáctidimenzionální chroma vektory. Tato metoda vykazuje větší odolnost vůči změnám instrumentace a různým poměrům vyšších harmonických složek jednotlivých tónů nahrávky na výsledný chromagram [19, 18].

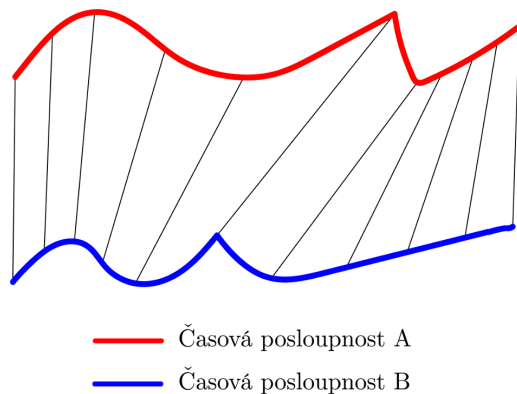
3.7 Metoda dynamického borcení časové osy

V předchozích kapitolách bylo nastíněno, jak mohou být různé interpretace skladeb pomocí chroma vlastností převedeny do chromagramů, díky kterým se dají dále analyzovat a porovnávat mezi sebou. Aby bylo možné jednotlivé interpretace synchronizovat z hlediska jejich časových os, je nutné použít techniku známou jako „dynamické borcení časové osy“ (anglicky *dynamic time warping*, zkráceně DTW). V současnosti bývá tato metoda využívána v různých odvětvích. Poprvé ji v padesátých letech minulého století vynalezl a použil americký matematik Richard Ernest Bellman [20], ovšem k jejímu většímu rozšíření došlo až v letech sedmdesátých, kdy se začala používat k automatickému rozpoznávání slov ve zvukových nahrávkách [21]. Tuto metodu lze ovšem efektivně využít téměř kdekoli, kde je potřeba zjistit podobnost časových posloupností dvou signálů.

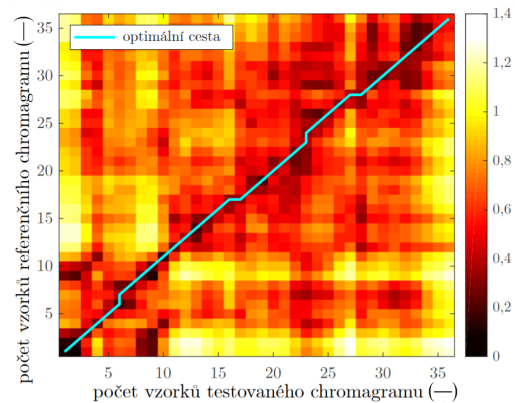
Cílem DTW je porovnávání dvou posloupností – v případě této práce se jedná o již výše zmíněné chromagramy. Obecně lze tedy první posloupnost označit jako $X = (x_1, x_2, \dots, x_N)$, kde $N \in \mathbb{N}$ odpovídá délce posloupnosti a druhou posloupnost jako $Y = (y_1, y_2, \dots, y_M)$ s délkou $M \in \mathbb{N}$. Elementy x_n a y_n jsou v tomto případě chroma vektory. X může být tedy posloupnost chroma vektorů první interpretace skladby, zatímco Y posloupnost chroma vektorů druhé interpretace skladby.

Protože dané dvě interpretace skladby nikdy nebudou zahrány v naprosto identickém tempu, jejich odpovídající posloupnosti se budou lišit délkou i přes to, že spolu obě nahrávky hudebně korespondují. Uvažujme tedy, že posloupnost X odpovídající první interpretaci skladby má délku $N = 10$, zatímco druhá posloupnost Y korespondující s druhou interpretací skladby má délku $M = 13$. Cílem DTW je kompenzovat rozdíly v lokálních tempech dvou nahrávek tím, že se k sobě přiřadí hudebně korespondující elementy, jejichž časy mohou být různé. Toho je docíleno tak, že se některé elementy posloupnosti buď přeskočí, nebo použijí vícrát než jednou. Například elementu x_3 mohou být přiřazeny dva elementy y_4 a y_5 a naopak.

Algoritmus DTW tedy přiřadí jednotlivé korespondující časové segmenty dvou posloupností k sobě, čímž je získána optimální cesta (v anglickém jazyce nazývaná jako *warping path*). V Matlabu je optimální cesta vyjádřena formou matice o dvou sloupcích, přičemž první sloupec odpovídá jednotlivým vzorkům referenčního chromagramu a druhý sloupec vzorkům chromagramu testovaného. Tímto způsobem se zjistí, které časové segmenty jednotlivých nahrávek sobě hudebně korespondují, čehož lze dále využít pro tvorbu algoritmu přehrávače interpretací [22].



Obr. 3.3: Přiřazení jednotlivých elementů posloupností pomocí DTW.



Obr. 3.4: Optimální cesta mezi dvěma chromagramy.

3.7.1 Monotónnost

U algoritmu DTW je vhodné stanovit určitá pravidla omezující tvar a oblast výsledné optimální cesty. Tím se sníží pravděpodobnost chybné synchronizace jednotlivých vzorků posloupností a zredukuje se nároky na výpočetní výkon.

Základní vlastnost, kterou musí optimální cesta splňovat, je monotónnost. Monotónnost zaručuje, že optimální cesta nejde zpět v čase, což by mělo za následek opakování některých vzorků po delších časových úsecích.

3.7.2 Kontinuita

Další vlastnost, kterou musí optimální cesta splňovat, je kontinuita. Tak je zaručeno, že výsledná cesta je spojitá, tudíž nepřeskakuje dopředu v čase a nevynechává žádné úseky posloupností.

3.7.3 Začátek a konec optimální cesty

Aby docházelo k synchronizaci obou celých posloupností a ne pouze k synchronizaci části jedné posloupnosti k posloupnosti druhé či synchronizaci pouze částí u obou posloupností, je nutné vymezit pozice začátku a konce optimální cesty. Toho je docíleno tak, že optimální cesta mezi dvěma posloupnostmi začíná v levém dolním rohu a končí v pravém horním rohu její celé možné oblasti, viz obr. 3.3.

3.7.4 Oblast optimální cesty

Vzhledem k tomu, že DTW obvykle porovnává posloupnosti, u kterých lze předpokládat, že jsou velmi podobné (v případě praktické části této práce se jedná pouze o drobné rozdíly v dynamice a tempech interpretací), je vhodné zajistit, aby se optimální cesta příliš nevzdalovala od diagonály. To by způsobovalo nerovnoměrné rozložení přiřazených bodů v časech posloupností.

3.7.5 Omezení strmosti

Pokud by některé části výsledné optimální cesty byly příliš strmé či naopak téměř vodorovné, docházelo by k přiřazování krátkých částí první posloupnosti k dlouhým částem posloupnosti druhé, což by v extrémním případě mohlo znamenat i přiřazení několika stovek vzorků jedné posloupnosti jednomu vzorku druhé posloupnosti. Pro zamezení tohoto jevu je vhodné omezit strmost optimální cesty. Pokud by tedy došlo k přiřazení několika vzorků pouze jednomu vzorku posloupnosti po sobě, algoritmus by byl donucen přejít k vzorku dalšímu [23].

3.7.6 Víceúrovňové dynamické borcení časové osy

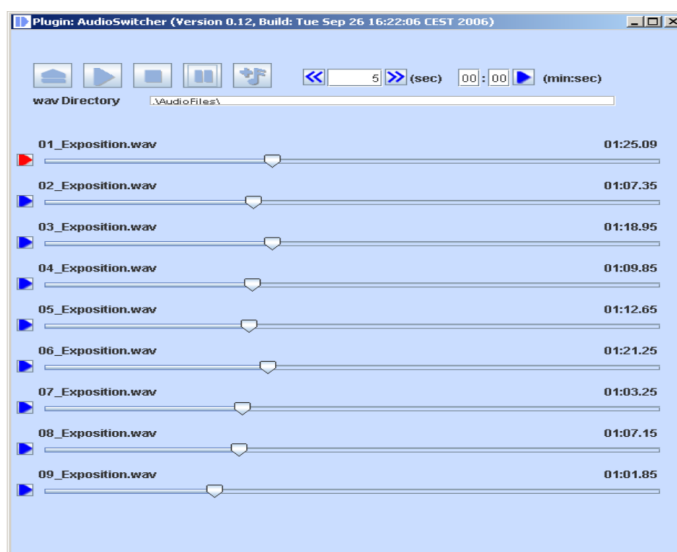
Vzhledem k tomu, že při návrhu a použití jedné z výše zmíněných technik pro omezení optimální cesty je optimální cesta stále neznámá, často bývá těžké najít správný kompromis mezi výběrem co nejmenší oblasti pro tuto cestu, čímž se zaručí nižší nároky na výpočetní výkon i zamezí vzniku chybných přiřazení, ale zároveň výběrem oblasti dostatečně veliké, aby v ní výsledná optimální cesta po výpočtu doopravdy ležela. Jako vhodné řešení se jeví použití algoritmu, který vymezení oblast cesty v závislosti na vstupním materiálu, na rozdíl od výše uvedených technik, které jsou na zpracovávaných datech nezávislé. Tato metoda se nazývá víceúrovňové dynamické borcení časové osy (anglicky *multiscale DTW*) a její princip spočívá v tom, že se optimální cesta nejdříve vypočítá v malém rozlišení a na základě pozice a tvaru takto vypočítané cesty se v dalším kroku počítá cesta s větším rozlišením, ale v menší vymezené oblasti. Při výpočtu optimální cesty mezi dvěma datově objemnými posloupnostmi lze tento krok opakovat vícekrát, přičemž v posledním kroku se vždy výsledná optimální cesta počítá s rozlišením nejvyšším.

Při návrhu a použití algoritmu pro tuto metodu je nutné brát v potaz, že při zmenšování rozlišení chromagramů mohou být důležité informace o tónovém vývoji zkresleny, či dokonce ztraceny. To může způsobit chybný výpočet cesty, což znamená špatné vymezení oblasti pro výpočet na další úrovni. Vzhledem k tomu, že k nenávratné ztrátě dat dochází při zmenšování rozlišení pokaždé, je nutné najít nejmenší možné rozlišení, které stále ještě vede k relevantním výsledkům. Pro kompenzaci chybovosti způsobené tímto principem je vhodné při výpočtu každé další úrovně vymezenou oblast rozšířit do všech směrů i přes to, že se tím zvýší nároky na výpočetní výkon a čas [22].

4 Výsledky dosažené v zahraničí

Německému výzkumnému centru „*International Audio Laboratories Erlangen*“ se podařilo vyvinout nástroj „*Interpretation Switcher*“ (obr. 4.1), který umožňuje uživateli vybírat si mezi několika nahrávkami stejné skladby [24].

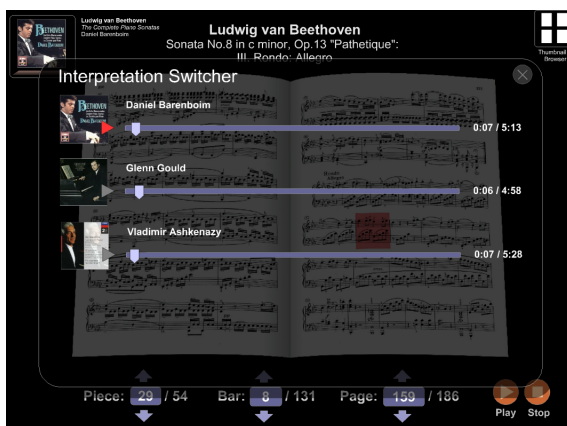
Každá nahrávka je zde reprezentovaná grafickým posuvníkem, který zobrazuje aktuálně přehrávanou pozici v čase. Posuvník aktuálně přehrávané nahrávky má barvu červenou, ostatní posuvníky barvu modrou. Zatímco časová pozice posuvníku aktuálně přehrávané nahrávky se vždy pohybuje konstantní rychlostí, pozice posuvníků ostatních nahrávek se mění relativně vzhledem ke změnám tempa vůči nahrávce aktuálně přehrávané. Posuvník aktuálně přehrávané nahrávky tudíž slouží jako reference, neboť podle jeho pozice jsou určovány i pozice ostatních posuvníků tak, aby korespondovaly stejné části hudebního díla. Po přepnutí aktuálně přehrávané nahrávky na nahrávku jinou začne nově přehrávaná nahrávka hrát v čase, který odpovídá stejné části skladby, jako byla přehrávaná před přepnutím. Nově přehrávaná nahrávka se tak stane opět nahrávkou referenční a posuvníky ostatních nahrávek se přizpůsobí vzhledem k ní.



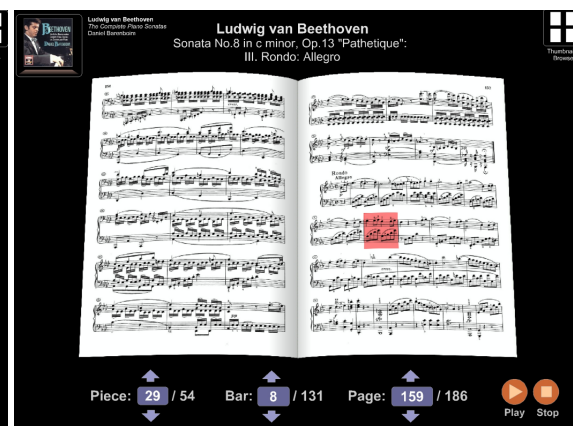
Obr. 4.1: International Audio Laboratories Erlangen: Interpretation Switcher.

Velmi podobného výsledku se podařilo dosáhnout roku 2008 skupině vývojářů z vícero německých univerzit [25]. Tento software mezi sebou také dokáže synchronizovat více nahrávek stejné skladby jako je tomu v případě výše zmíněného *Interpretation Switcheru*, je ale navíc doplněn o funkci synchronizace audio nahrávek s naskenovaným notovým záznamem dané skladby.

Uživatel si může zvolit, zda chce přepínat mezi jednotlivými interpretacemi, nebo sledovat v notovém záznamu zvýrazněné noty korespondující s aktuálně přehrávanou pozicí skladby v čase.



Obr. 4.2: The Interpretation Switcher.



Obr. 4.3: The Sheet Music Interface.

Oba znázorněné přehrávače fungují na stejném principu – nahrávky jsou nejdříve analyzovány a pomocí chroma parametrů jsou vytvořeny odpovídající chromagramy, ze kterých se následně metodou DTW zjistí jejich korespondující časové úseky. V případě synchronizace audio nahrávky s notovým zápisem je notový zápis také transformován do chromagramu, čehož je docíleno analýzou pomocí technik známých jako „*Optical music recognition*“ (zkráceně OMR) [26].

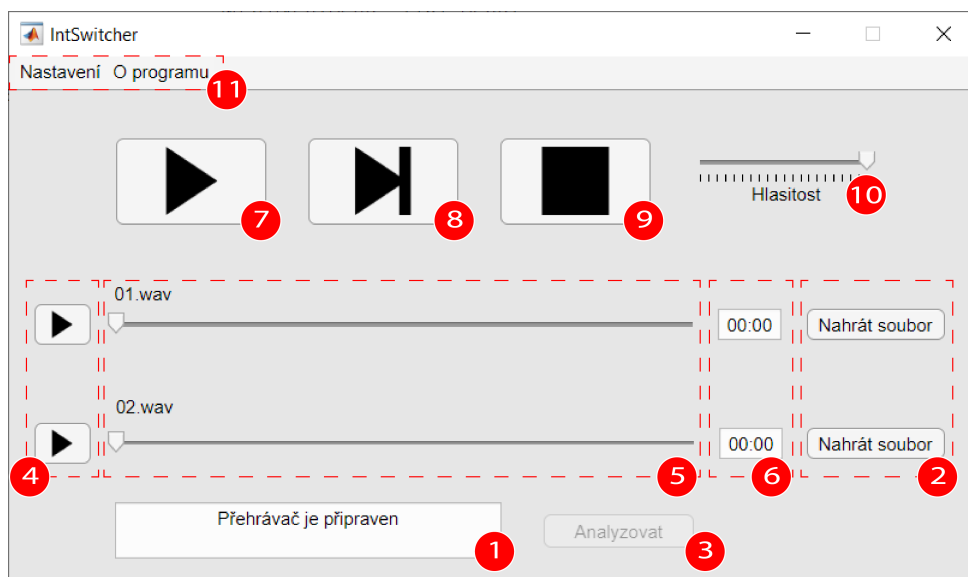
Cílem praktické části této práce popsané v následujících kapitolách je navrhnout a naprogramovat nástroj, který bude synchronizovat nahrávky interpretací a umožní uživateli mezi nimi v reálném čase přepínat, podobně jako toho bylo dosaženo v nástrojích uvedených v této kapitole.

5 Popis programu IntSwitcher

Výsledkem praktické části této bakalářské práce je program „Intswitcher“ realizovaný v Matlabu. Tento program uživateli umožňuje přehrávat dvě různé interpretace jedné skladby a přepínat mezi nimi v reálném čase tak, aby uživatel po přepnutí nemusel v druhé nahrávce manuálně hledat stejnou část dané skladby, jako která hrála v nahrávce před přepnutím. Program, na rozdíl od klasického hudebního přehrávače, rozpozná, která část skladby je zrovna přehrávaná a druhá stopa po přepnutí začne hrát v čase korespondujícím.

Software byl vyvinut v Matlabu verze R2019b, při jeho otevření ve starších verzích nemusí fungovat správně či vůbec. Dále je pro správnou funkcionalitu programu nutné mít v Matlabu nainstalovaný DSP System Toolbox a jako výchozí pracovní složku zvolit celý adresář s názvem `IntSwitcher`, včetně všech jeho podsložek. Podsložka `data` obsahuje ikony programu, uložené nastavení programu, zdrojové kódy pro výpočet chromagramů a DTW, podokna programu a složku pro ukládání již analyzovaných chromagramů. Soubor samotného programu se jmenuje `IntSwitcher.mlapp` a nachází se v nejvyšší úrovni již zmiňovaného adresáře `IntSwitcher`, který lze nalézt v příloze této práce.

V následujících podkapitolách budou popsány jednotlivé ovládací prvky tohoto programu, včetně popisu některých jeho klíčových algoritmů. Číslo podkapitoly vždy odpovídá číslu ovládacího prvku na schématu programu, viz obr. 5.1.



Obr. 5.1: Schéma programu IntSwitcher.

5.1 Informační panel

Pro snadnou interakci uživatele s programem obsahuje program informační panel, který uživatele vyzývá k další akci, informuje o aktuálním stavu programu či případně vypisuje chybové hlášky. Například po zapnutí programu se zde objeví hláška „Nahraj soubory k přehrávání“, uživatel je tudíž ihned naveden ke vhodným krokům pro realizaci přehrávání, aniž by byl nucen číst tuto dokumentaci.

5.2 Tlačítka „Nahrát soubor“

Pro načtení audio souborů do přehrávače pro jejich pozdější přehrávání slouží dvojice tlačítek s popisem „Nahrát soubor“. Po stisku jednoho z těchto dvou tlačítek se otevře okno, ve kterém je uživatel vyzván k výběru daného audio souboru ve formátu `.wav`. Tyto soubory je možné načítat z libovolných adresářů na disku. O úspěšném načtení audio souboru do programu je uživatel informován pomocí informačního panelu a název načteného souboru se objeví nad posuvníkem korespondujícím s danou stopou, viz podkapitola 5.5. Pokud uživatel načte soubor v jiném formátu, na informačním panelu se naopak vypíše chybová hláška, a je nezbytné tento celý krok zopakovat. Aby bylo možné přehrávač používat, je nutné do programu načíst obě vzájemně si korespondující audio stopy, respektive dvě nahrávky interpretací jedné skladby. Vzorkovací kmitočty těchto dvou nahrávek nemusí být totožné.

5.3 Tlačítko „Analyzovat“

Po načtení obou audio stop do přehrávače je třeba nahrávky analyzovat stiskem tohoto tlačítka. Při stisknutí tlačítka po načtení pouze jedné nebo žádné nahrávky vypíše informační panel chybovou hlášku, která uživatele vyzve k načtení obou audio stop. Pokud je vše v pořádku, program po stisknutí tlačítka začne provádět analýzu a uživatel je v jejím průběhu informačním panelem informován o aktuálně prováděném kroku. Během analýzy přehrávač nejprve vytvoří chromagramy obou nahrávek, z nichž následně vypočítá optimální cestu, která je používána pro plynulé přepínání mezi nahrávkami, pro výpočet aktuálních pozic posuvníků i pro zobrazení časových údajů přehrávaných nahrávek. S optimální cestou program pracuje ve formě dvouřádkové matice, přičemž každý řádek reprezentuje jednu stopu a hodnoty buněk v jednom sloupci představují vzájemně si odpovídající časové body nahrávek. Časové rozlišení těchto bodů je 50 ms. Analyzované chromagramy se ukládají do adresáře `IntSwitcher/data/chromagramy` do souborů `.mat`, názvy těchto souborů odpovídají názvům korespondujících audio souborů na disku. V těchto souborech `.mat` je kromě matice daného chromagramu uložena také cesta k odpovídajícímu audio

souboru na disku, která slouží ke kontrole, zda-li se jedná o správný chromagram. Pokud program zjistí, že uložený chromagram náleží jinému audio souboru, který má sice stejný název, ale je uložen v jiném adresáři, chromagram nebude ze souboru načten ale proběhne jeho výpočet. Pokud uložený chromagram náleží stejnému souboru, jako který byl uživatelem načten do přehrávače, chromagram se ze souboru načte a není třeba jej znovu počítat, čímž se urychlí proces analýzy nahrávek.

Po úspěšném skončení analýzy se na informačním panelu vypíše hláška „Přehrávač je připraven“, tlačítka analýzy se deaktivuje a tlačítka pro přehrávání audio stop se aktivují. V tuto chvíli je již možné program používat k přehrávání a porovnávání interpretací. Pokud by uživatel do jedné z audio stop nahrál nový soubor, tlačítka přehrávače se opět deaktivují, tlačítka analýzy se aktivuje a je třeba opět provést analýzu s novým audio souborem. V takovém případě bude probíhat výpočet chromagramu pouze u nově načtené nahrávky.

5.4 Tlačítka „Přehrát/pauza“

Jsou-li do přehrávače načteny obě audio stopy, byla provedena analýza a na informačním panelu se zobrazila hláška „Přehrávač je připraven“, je možné pomocí těchto tlačítek spustit jednu z načtených stop. Po spuštění přehrávání se ikona tlačítka změní na ikonu pauzy, což signalizuje změnu funkce tohoto tlačítka – pokud je přehrávač pozastaven, tlačítka slouží ke spuštění přehrávání, pokud přehrávač hraje, tlačítka slouží k pozastavení přehrávané stopy. Třetí funkcí těchto tlačítek je plynulé přepínání mezi stopami. Toho lze docílit stisknutím tlačítka „Přehrát/pauza“ u právě nepřehrávané stopy.

Přehrávání nahrávek je realizované pomocí `while` smyčky, v jejíž každé iteraci dochází k načtení takového počtu vzorků ze souboru, který odpovídá nastavené velikosti bufferu, viz podkapitola 5.11.1. Tyto vzorky jsou dále dosazované do zvukové karty, kde dochází k D/A převodu. Z prvního načteného vzorku v každém cyklu je počítána aktuální časová pozice dané audio stopy vydělením čísla udávajícího pozici daného vzorku v audio souboru hodnotou vzorkovacího kmitočtu f_{vz} . Tím je získána aktuální časová pozice přehrávané stopy v sekundách, pomocí které je z matice optimální cesty zjištěna korespondující časová pozice druhé nahrávky. Ta je poté opět vynásobena vzorkovacím kmitočtem f_{vz} , čímž se zjistí pozice aktuálně korespondujícího vzorku druhé stopy. Tato hodnota je v programu dočasně uložena a pokud dojde k přepnutí aktuálně přehrávané nahrávky uživatelem, přehrávač začne hrát druhou stopu od tohoto vypočítaného vzorku. Tím je docíleno plynulého přepínání mezi stopami tak, aby po přepnutí druhá nahrávka hrála ve stejné části skladby, i pokud jsou tyto stejné části skladby v jiných časech přehrávaných interpretací.

5.5 Časové posuvníky

Každé audio stopě v přehrávači náleží jeden časový posuvník. U aktuálně přehrávané nahrávky posuvník zobrazuje momentální pozici v čase, v případě posuvníku náležícímu druhé stopě, než která je aktuálně přehrávaná, zobrazuje posuvník pozici korespondující. Tahem posuvníku myši je možné přehrávané stopy v čase přetáčet. Nad posuvníkem je vždy zobrazen název odpovídajícího audio souboru.

Pozice posuvníků se určují z aktuálně přehrávaného vzorku. V případě momentálně neaktivní stopy je pozice určena z korespondujícího vzorku zjištěného z optimální cesty. Začátek posuvníku odpovídá prvnímu vzorku nahrávky, konečný bod posuvníku naopak náleží poslednímu vzorku dané audio stopy.

5.6 Počítadla času

Vpravo od posuvníků se nacházejí počítadla času, které uživatele informují o aktuální časové pozici přehrávané nahrávky a korespondující časové pozici druhé stopy.

Aktuální časová pozice je, podobně jako u posuvníků, vypočtena z aktuálně přehrávaného vzorku, v tomto případě vyděleného vzorkovacím kmitočtem f_{vz} . Získaný časový údaj v sekundách je poté funkcí v Matlabu převeden do formátu mm:ss, který je na počítadlech času zobrazen.

5.7 Globální tlačítko „Přehrát/pauza“

Toto tlačítko má stejnou funkci jako tlačítko „Přehrát/pauza“ vlevo od posuvníků daných stop, rozdíl je v tom, že s ním lze ovládat obě nahrávky, přičemž tlačítko vždy ovládá právě tu nahrávku, která je zrovna přehrávaná, nebo byla přehrávaná jako poslední. Pokud je přehrávač neaktivní, tlačítko slouží ke spuštění přehrávání, pokud program aktuálně jednu ze stop přehrává, tlačítko slouží k jejímu pozastavení.

5.8 Globální tlačítko „Přepnout“

Toto tlačítko slouží k přepínání aktuálně přehrávané stopy na stopu druhou, nehledě na to, která ze dvou nahrávek právě hraje. Používáním tohoto tlačítka lze mezi stopami přepínat rychleji po sobě, než při použití tlačítek „Přehrát/pauza“ vlevo od posuvníků, protože není potřeba přesouvat kurzor z jednoho tlačítka na druhé a zpět.

5.9 Globální tlačítko „Zastavit“

Tímto tlačítkem lze zastavit přehrávání. Po jeho stisknutí se pozice obou posuvníků vrátí na začátek, počítadla času se vynulují a po opětovném spuštění jedné ze dvou stop začne daná stopa hrát od začátku.

5.10 Posuvník „Hlasitost“

Tento posuvník slouží k úpravě hlasitosti přehrávání. V pozici krajně vlevo je přehrávání kompletně ztišeno, v pozici vpravo je hlasitost naopak maximální.

Změna hlasitosti je implementovaná pomocí násobení vzorků, které se načítají ze souboru a posílají do zvukové karty, hodnotami v rozsahu od 0 do 1. Hodnota, kterou jsou vzorky násobeny, závisí na nastavené poloze posuvníku. Vzhledem k tomu, že lidmi vnímaná závislost hlasitosti zvuku na jeho intenzitě není lineární, ale přibližně logaritmická, je vhodné implementovat posuvník sloužící ke změně hlasitosti tak, aby byl tento jev částečně vykompenzován. Toho lze docílit použitím exponenciální funkce na závislost mezi polohou posuvníku a hodnotou, kterou jsou dané vzorky násobeny. Při použití lineární funkce pro zesílení je subjektivně vnímaný rozsah zesílení první poloviny posuvníku větší, než rozsah druhé poloviny posuvníku. Vhodnost tvaru použité exponenciální funkce závisí na dynamickém rozsahu přehrávané nahrávky i hlasitosti přehrávání, plynulost zesilování bude tedy vždy pouze přibližná. Bylo otestováno několik tvarů exponenciálních funkcí pro zesilování, přičemž jako nejvhodnější a nejplynuleji znějící tvar funkce pro změnu hlasitosti u testovaných nahrávek byla vybrána a použita závislost

$$y = 3,1623 \cdot 10^{-3} \cdot e^{5,757 \cdot x}, \quad (5.1)$$

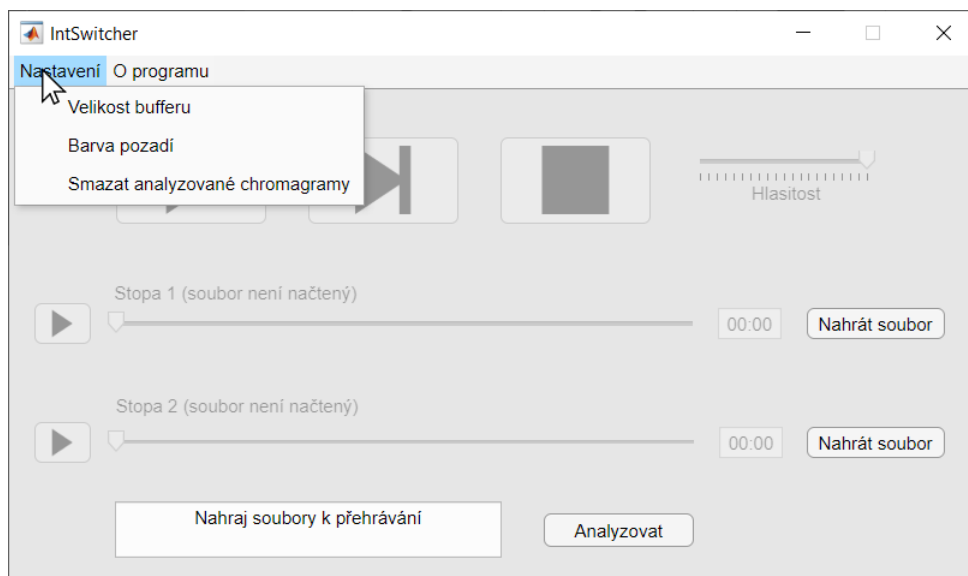
kde y je hodnota, kterou jsou násobeny vzorky přehrávané nahrávky a x je aktuální hodnota posuvníku v rozsahu 0 až 1. Tento tvar funkce kompenzuje logaritmické vnímání zvuku člověkem pro dynamický rozsah 50 dB [27].

5.11 Menu programu

Horní lišta programu fungující jako menu obsahuje dvě tlačítka, jejichž funkcionalita je popsána v následujících podkapitolách.

5.11.1 Nastavení

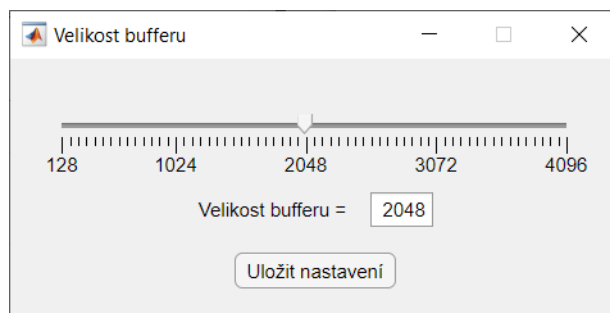
Po kliknutí na toto tlačítko se otevře nabídka, která uživateli umožňuje zvolit jednu ze tří možností nastavení programu.



Obr. 5.2: Tlačítko nastavení.

Velikost bufferu

První ze tří možností nastavení programu je nastavení velikosti bufferu. Ten lze nastavit v rozsahu 128 až 4096 vzorků, výchozí nastavení programu je 2048 vzorků. Nižší velikosti bufferu způsobují lepší odezvu programu, ale přehrávání může být nestabilní. Nastavení lze provést buďto posuvníkem, nebo zadáním číselné hodnoty do pole vedle textu „Velikost bufferu =“. Velikost bufferu lze měnit pouze v případě, že přehrávač aktuálně nepřehrává žádnou z načtených stop, v opačné situaci je na informačním panelu vypsána chybová hláška a změna hodnoty se neuloží. V případě úspěšného přenastavení velikosti bufferu je po stisknutí tlačítka „Uložit nastavení“ okno zavřeno a uživatel je o provedené akci informován informačním panelem. Nastavená hodnota velikosti bufferu se ukládá do souboru `IntSwitcher/data/app_nastaveni.mat`, ze kterého je po vypnutí a opětovném zapnutí programu načtena.



Obr. 5.3: Velikost bufferu.

Barva pozadí

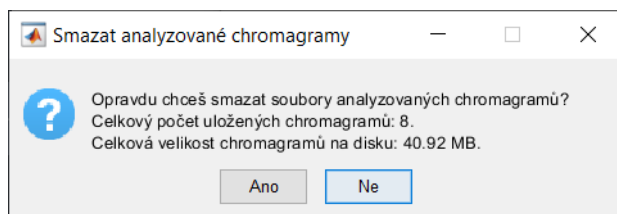
Další možností nastavení programu je barva pozadí. Uživatel si zde může přizpůsobit vizuální stránku programu výběrem jedné z předpřipravených barev či namícháním vlastní barvy. O změně barvy je uživatel informován informačním panelem a nastavená barva pozadí se společně s nastavenou hodnotou bufferu ukládá do souboru `IntSwitcher/data/app_nastaveni.mat`, ze kterého je při zapínání programu načtena.



Obr. 5.4: Barva pozadí.

Smazat analyzované chromagramy

Pomocí této funkce je možné vymazat uložené chromagramy nahrávek z adresáře IntSwitcher/data/chromagramy, které byly analyzovány v minulosti. Uživatel je v tomto okně informován o celkovém počtu uložených chromagramů a jejich celkové velikosti na disku a dále se program uživatele zeptá, zda-li si opravdu přeje tyto soubory vymazat. V případě úspěšného smazání uložených chromagramů je uživatel o provedené akci informován informačním panelem.



Obr. 5.5: Smazat analyzované chromagramy.

5.11.2 O programu

V případě kliknutí na tlačítko „O programu“ se otevře okno informující uživatele o autorovi programu, účelu jeho vzniku i období, ve kterém byl vyvíjen.



Obr. 5.6: O programu.

6 Testování programu IntSwitcher

Program byl otestován na sadě nahrávek interpretací dvou skladeb uvedených v tabulkách níže. V těchto tabulkách jsou zaznamenány změřené časové hodnoty délek trvání výpočtu chromagramů jednotlivých nahrávek i výpočtů *optimálních cest* mezi dvojicemi nahrávek. K tomuto testování byl použit laptop s procesorem Intel Core i7-4700HQ, velikostí operační paměti 12 GB a operačním systémem Windows 10.

Dále byla testována správnost chodu přehrávače a přepínání mezi dvěma audio stopami. Program správně rozpoznal korespondující časové body nahrávek a při přepínání mezi nimi skutečně přepínal do stejných úseků daných skladeb. Program také úspěšně ukládal i načítal již analyzované chromagramy i nastavení uživatele. Při přepínání stop lze ovšem občas slyšet krátké lupnutí, neboť přehrávač na stopy neaplikuje fade-in ani fade-out. Dalším drobným nedostatkem programu je nutnost použití relativně velkého bufferu pro dosažení plynulého přehrávání, v porovnání s obdobným audio softwarem. To je způsobeno implementací programu v Matlabu.

Tab. 6.1: Dvořákova 9. symfonie, 4. věta

číslo nahrávky	délka nahrávky (mm:ss)	doba trvání výpočtu chromagramu (s)	doba trvání výpočtu warping path (s)
1	9:49	12,3	
2	10:24	12,9	3,2
3	9:54	12,6	
4	11:43	14,4	4,1
5	12:41	15,8	
6	10:36	13,0	4,1

Tab. 6.2: Janáčkův 1. smyčcový kvartet, 4. věta

číslo nahrávky	délka nahrávky (mm:ss)	doba trvání výpočtu chromagramu (s)	doba trvání výpočtu warping path (s)
1	5:31	6,6	
2	5:58	7,6	1,4
3	5:16	6,4	
4	5:18	6,6	1,2
5	5:57	7,5	
6	5:28	6,7	1,3

Závěr

První kapitola teoretické části této bakalářské práce byla věnována základním vlastnostem zvuku, jako je rychlost jeho šíření, frekvence či barva, čímž byly definovány základní znalosti nutné k pochopení kapitol navazujících.

V následující kapitole nesoucí název „Digitalizace zvuku“ byl popsán proces záznamu zvuku do digitální podoby, jehož provedení je nezbytné pro vytvoření audio souborů, se kterými přehrávač interpretací pracuje. U části týkající se vzorkování byl popsán zejména Shannonův teorém, který je nutno dodržet, aby nedocházelo k výrazné ztrátě kvality vlivem aliasingu. V podkapitole „Kvantování“ byl popsán vliv počtu kvantizačních hladin na intenzitu vzniklého kvantizačního šumu a byl nastíněn proces ditheringu, kterým lze do jisté míry tento šum maskovat.

Další kapitola se nejprve věnuje metodám analýzy zvuku z hlediska jejich frekvenčního vývoje v čase, jako je Fourierova řada, diskrétní Fourierova řada, DFT a FFT, které se používají pro generování spektrogramů a chromagramů, jež jsou v této kapitole následně popsány. Na závěr je pozornost věnována algoritmu DTW, jeho principu a podmínkám, které je vhodné stanovit pro jeho optimalizaci.

Dále byly uvedeny dva již existující přehrávače, jejichž funkcionality je podobná výslednému programu praktické části této práce.

Kapitola „Popis programu IntSwitcher“ obsahuje schéma ovládacích prvků programu realizovaného v rámci praktické části této bakalářské práce a vysvětluje funkce těchto jednotlivých prvků spolu s nastíněním klíčových algoritmů programu.

V poslední kapitole jsou uvedeny výsledky testování programu a seznamy skladeb, na kterých byl program testován.

Vytvořením programu „IntSwitcher“ byly úspěšně splněny body zadání této práce. Kvůli implementaci programu v Matlabu ovšem nebylo možné přehrávač ideálně optimalizovat, co se týče nároků na výpočetní výkon zařízení. Zlepšení těchto vlastností by bylo realizovatelné přepsáním programu do programovacího jazyka, který je na real-timeové aplikace určen, např. C++ nebo Java. Dále se zde nabízí příležitost k dalšímu vylepšení přehrávače, kupříkladu implementace možnosti přehrávání více než dvou interpretací či zobrazování zvukových vln v reálném čase během přehrávání.

Literatura

- [1] KINSLER, Lawrence E., Austin R. Frey, Alan B. COPPENS a James V. SANDERS. *Fundamentals of acoustics*. 4th ed. New York: Wiley. ISBN 0-471-84789-5.
- [2] STEVENS, S. S., J. VOLKMANN a E. B. NEWMAN. A Scale for the Measurement of the Psychological Magnitude Pitch. Cambridge, Massachusetts, 1937. Harvardova univerzita.
- [3] MATSUOKA, T. a K. ITO. Perception of missing fundamental and consideration on its characteristics. In: Proceedings of the 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (IEEE Cat.No.03CH37439). IEEE, 2003, s. 2059-2062. DOI: 10.1109/IEMBS.2003.1280141. ISBN 0-7803-7789-3. Dostupné také z: <http://ieeexplore.ieee.org/document/1280141/>
- [4] SMĚKAL, Zdeněk. *Analýza signálů a soustav - BASS*. Purkyňova 118, 612 00 Brno: Vysoké učení technické v Brně, 2012. ISBN 978-80-214-4453-9.
- [5] DOLAN, Emily I. a Alexander REHDING. Describing Sound: The Cognitive Linguistics of Timbre. The Oxford handbook of timbre. New York: Oxford University Press, 2018, 2018-06-07. DOI: 10.1093/oxfordhb/9780190637224.013.14.
- [6] PELGROM, Marcel J.M. *Analog-to-digital conversion*. New York: Springer, c2010. ISBN 978-90-481-8887-1.
- [7] WAGDY, Mahmoud F. Effect of various dither forms on quantization errors of ideal A/D converters. IEEE, 1989.
- [8] HECKBERT, Paul. *Fourier Transforms and the Fast Fourier Transform (FFT) Algorithm*. 1995.
- [9] Kozický, Claudio. *Porovnání algoritmů pro rychlou Fourierovu transformaci*. Bakalářská práce. Praha: České vysoké učení technické v Praze, Fakulta informačních technologií, 2017.
- [10] COOLEY, James W. a John W. TUKEY. An Algorithm for the Machine Calculation of Complex Fourier Series. *Mathematics of Computation*. 1965, **19**(90). DOI: 10.2307/2003354. ISSN 00255718. Dostupné také z: <https://www.jstor.org/stable/2003354?origin=crossref>

- [11] Classic Spectrograms. SMITH, Julius O. *Spectral Audio Signal Processing* [on line]. 2011 edition. [cit. 2019-10-31]. Dostupné z: https://www.dsprelated.com/freebooks/sasp/Classic_Spectrograms.html
- [12] XIAO, Chuan. Cover song identification using an enhanced chroma over a binary classifier based similarity measurement framework. In: *2012 International Conference on Systems and Informatics (ICSAI2012)*. Yantai: IEEE, 2012, 2012, s. 2170-2176. DOI: 10.1109/ICSAI.2012.6223482. ISBN 978-1-4673-0199-2. Dostupné také z: <http://ieeexplore.ieee.org/document/6223482/>
- [13] MIKLÁNEK, Štěpán *Určení místa původu hudebních interpretací české komorní a orchestrální hudby za pomoci technik Music Information Retrieval*. Technická 3058/10, 616 00 Brno, 2019. Diplomová práce. Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií. Vedoucí práce Tomáš Kiska.
- [14] Loudness. FASTL, Hugo a Eberhard ZWICKER. *Psychoacoustics: facts and models*. 3rd. ed. New York: Springer, c2007, s. 203-233. ISBN 3-540-23159-5.
- [15] MÜLLER, Meinard. *Information retrieval for music and motion*. New York: Springer, 2007. ISBN 35-407-4047-3.
- [16] TERASAWA, H., M. SLANEY a J. BERGER. The thirteen colors of timbre. In: *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, 2005*. IEEE, 2005, s. 323-326. DOI: 10.1109/ASPAA.2005.1540234. ISBN 0-7803-9154-3. Dostupné také z: <http://ieeexplore.ieee.org/document/1540234/>
- [17] AUCOUTURIER, Jean-Julien a Francois PACHET. *Improving Timbre Similarity: How high's the sky?*. 2004.
- [18] MÜLLER, M. *Chroma Toolbox: MATLAB Implementations for Extracting Variants of Chroma-based audio features* [online]. 2001. [cit. 2019-11-20] Dostupné z URL: <https://pdfs.semanticscholar.org/31cf/b91158031038ad16218997cc9724a9ab5c66.pdf>
- [19] MULLER, Meinard a Sebastian EWERT. Towards Timbre-Invariant Audio Features for Harmony-Based Music. In: *IEEE Transactions on Audio, Speech, and Language Processing* [online], 2010, **18**(3), s. 649-662 [cit. 2019-11-20]. DOI: 10.1109/TASL.2010.2041394. ISSN 1558-7916. Dostupné z: <http://ieeexplore.ieee.org/document/5410051/>

- [20] BELLMAN, Richard E. *Dynamic Programming*. 6th printing. New Jersey: Princeton University Press, 1957. ISBN 0-691-07951-X.
- [21] SAKOE, H. a S.CHIBA. Dynamic programming algorithm optimization for spoken word recognition. In: *IEEE Transactions on Acoustics, Speech, and Signal Processing* [online]. 1978, **26**(1), s.43-49 [cit. 2019-11-20]. DOI: 10.1109/TASSP.1978.1163055. ISSN 0096-3518. Dostupné z: <http://ieeexplore.ieee.org/document/1163055/>
- [22] MÜLLER, Meinard. *Fundamentals of music processing: Audio, Analysis, Algorithms, Applications*. New York, NY: Springer Berlin Heidelberg, 2015. ISBN 978-3-319-21944-8.
- [23] TSIPORKOVA, Elena. *Dynamic Time Warping Algorithm for Gene Expression Time Series* [online]. [cit. 2019-11-22]. Dostupné z: <http://www.psb.ugent.be/cbd/papers/gentxwarper/DTWAlgorithm.ppt>
- [24] KONZ, Verena. *Automated Methods for Audio-Based Music Analysis with Applications to Musicology*. 66123 Saarbrücken, Germany, 2012. Disertační práce. Saarland University. Vedoucí práce Prof. Dr. Meinard Müller.
- [25] DAMM, David, Christian FREMEREY, Frank KURTH, Meinard MÜLLER a Michael CLAUSEN. Multimodal presentation and browsing of music. In: *Proceedings of the 10th international conference on Multimodal interfaces - IMCI '08*. New York, New York, USA: ACM Press, 2008, 2008. DOI: 10.1145/1452392.1452436. ISBN 9781605581989. Dostupné také z: <http://portal.acm.org/citation.cfm?doid=1452392.1452436>
- [26] LI, Chuanzhen, Jiaqi ZHAO, Juanjuan CAI, Hui WANG a Huaichang DU. Optical Music Notes Recognition for Printed Music Score. In: *2018 11th International Symposium on Computational Intelligence and Design (ISCID)*. IEEE, 2018, 2018, s. 285-288. DOI: 10.1109/ISCID.2018.00071. ISBN 978-1-5386-8527-3. Dostupné také z: <https://ieeexplore.ieee.org/document/8695563/>
- [27] THOMAS, Alexander. Programming Volume Controls: a.k.a. Notice to Programmers of Audio Software and Hardware. *Dr. Lex' Site* [online]. [cit. 2020-05-19]. Dostupné z: <https://www.dr-lex.be/info-stuff/volumecontrols.html>

Seznam symbolů, veličin a zkratk

CD	Compact Disc – Kompaktní disk
DFT	Diskrétní Fourierova Transformace
DTW	Dynamic Time Warping – Dynamické borcení časové osy
f	Frekvence
FFT	Fast Fourier Transform – Rychlá fourierova transformace
f_{vz}	Vzorkovací kmitočet
MIDI	Musical Instrument Digital Interface
OMR	Optical Music Recognition
PCM	Pulse Code Modulation – Pulzně kódová modulace
PFCCs	Pitch-Frequency Cepstral Coefficients – Tónově-frekvenční kepstrální koeficienty
SNR	Signal to Noise Ratio – Odstup užitečného signálu od šumu

Seznam příloh

A Přiložené soubory

43

A Příložené soubory

Příložený soubor s názvem `bp-priloha.zip` obsahuje tento dokument práce ve formátu `.pdf` a složku `IntSwitcher`. Tato složka obsahuje soubor výsledného programu s názvem `IntSwitcher.mlapp`. Dále se v tomto adresáři nachází složka `data` obsahující podsložky se všemi datovými soubory potřebnými ke správnému chodu přehrávače a soubor `app_nastaveni.mat`, ve kterém jsou uloženy uživatelské preference nastavení programu. V podsložce `grafika` se nacházejí všechny obrázky a ikony, se kterými program pracuje, složka `chromagramy` slouží k ukládání chromagramů již analyzovaných nahrávek, složka `okna` obsahuje aplikace, které se otevírají při zvolení jedné z možností menu programu a složka `source` obsahuje funkce potřebné k výpočtům chromagramů a víceúrovňového DTW.

Obsah archivu `bp-priloha.zip`:

- `bp-svejcar-michael.pdf`
- `IntSwitcher`
 - `IntSwitcher.mlapp`
 - `data`
 - * `grafika`
 - * `chromagramy`
 - * `okna`
 - * `source`
 - * `app_nastaveni.mat`