



Masarykova univerzita, Fakulta informatiky

Humanizace MIDI záznamu

Diplomová práce

Pavel Hamřík
29. 5. 2011

Zadání práce

Prohlášení

Prohlašuji, že tato bakalářská práce je mým samostatně vypracovaným, původním autorským dílem. Veškeré zdroje a prameny v práci použité nebo čerpané jsem řádně citoval a v závěru práce uvedl odkaz na příslušný zdroj.

Pavel Hamřík

Poděkování

Na tomto místě bych chtěl poděkovat zejména docentu Josefu Prokešovi nejen za to, že na sebe vzal úděl vedení této práce, ale zejména za jeho průběžnou podporu mé hudební angažovanosti.

Za ochotu při konzultacích hudebně teoretické stránky této práce děkuji doc. Milanu Polákovvi, doc. Romanu Novozámskému a prof. Václavu Kuntovi, kteří působí na Hudební fakultě Janáčkovy akademie múzických umění v Brně.

Technickou konzultaci ke zpracování MIDI souborů a možností jejich konverze poskytl pan Günter Nagler, autor aplikací GNMIDI, z nichž některé tato práce využívá.

Za nahrání hudebních podkladů do MIDI jakožto i perfektní a svědomitou přípravu na něj děkuji saxofonistovi a klarinetistovi Filharmonie Brno MgA. Lukáši Daňhelovi, Ph.D.

Shrnutí

Práce analyzuje MIDI záznam hry hráče na dechový nástroj, vyvozuje z něj závěry týkající se chybovosti charakteristické pro interpretaci hudební předlohy a zabývá se teoretickými možnostmi i praxí reprodukce této chybovosti na pro počítač srozumitelném přepisu této předlohy (opět MIDI data). Práce srovnává dosavadní metody a nástroje využívající omezených možností generování chyb pomocí pseudonáhody a nastiňuje či přímo realizuje jejich zlepšení na základě analýzy vstupního MIDI záznamu.

Klíčová slova

humanizace, hudba, interpretace, MIDI, dechové nástroje, pseudonáhoda, distribuce náhodné veličiny, VST

Obsah

Zadání práce.....	2
Prohlášení.....	3
Poděkování.....	3
Shrnutí.....	4
Klíčová slova	4
1 Úvod.....	8
1.1 Premisy práce	9
1.2 Cíle práce.....	9
1.2.1 Zhodnocení cílů práce.....	9
2 Humanizace	11
2.1 Definice humanizace	11
2.2 Definice kvantizace	11
2.3 Vymezení interpretace a humanizace.....	11
2.4 Proč humanizovat už MIDI?	12
2.4.1 MIDI jako záznam výsledků motorické činnosti interpreta	13
2.4.2 Oddělení od producenta zvuku	14
2.5 Barva tónu a jeho další kvality.....	14
2.6 Jaké parametry MIDI lze humanizací ovlivňovat?	15
2.6.1 Pozice začátku	15
2.6.2 Délka noty	15
2.6.3 Subtonální výška noty.....	16
2.6.4 Dynamika noty	16
3 Sběr dat	18
3.1 Kritéria pro výběr interpreta.....	18
3.2 Volba nástroje	18
3.2.1 Dechové nástroje.....	18
3.2.2 Volba konkrétního dechového nástroje	19
3.3 Výběr hudebního materiálu	19
3.3.1 Fritz Kroepsch: 167 Excercises for Daily use	20
3.3.2 Alexander Borodin: Polovecké tance.....	20
3.3.3 Bedřich Smetana: Má vlast – Šárka.....	20

3.3.4	Béla Bartók: Concerto of Orchestra – I. Introduzione.....	20
3.3.5	Leoš Janáček: Glagolská mše	20
3.3.6	Nikolai Rimsky-Korsakow: Capriccio espagnol.....	20
3.3.7	Petr Iljič Čajkovskij: Masky	21
3.4	Příprava interpreta.....	21
3.5	Záznam.....	21
3.5.1	Vybavení	21
3.5.2	Metronom a práce s tempem	22
3.5.3	Způsob vlastního nahrávání	22
3.6	Kontrolní přepis.....	22
4	Komparatistika a analýza.....	23
4.1	Získání čitelných dat	23
4.1.1	Export MIDI.....	23
4.1.2	Převod na text	23
4.1.3	Zpracování textového přepisu MIDI souborů	24
4.1.4	Výpočty a grafy v Excelu	25
4.2	Postup při analýze.....	25
4.2.1	Sestavení statistických tabulek.....	25
4.3	Pohled na data	28
4.4	Analýza odchylky začátku znění noty	28
4.4.1	Distribuce odchylek pozice začátku noty v čase	34
4.5	Analýza odchylky délky znění noty.....	36
4.5.1	Distribuce odchylek délky znění noty	40
4.6	Dynamické odchylky	40
4.6.1	Odchylky průměrné dynamiky not.....	41
4.6.2	Průběh dynamiky dlouhých not.....	43
4.7	Nástin humanizace ladění a intonace.....	45
5	Metody humanizace.....	47
5.1	Stávající metody.....	47
5.1.1	Humanizované parametry stávajících metod	47
5.2	Dílič nástroje vylepšení metody humanizace	48
5.2.1	Zpracování dle pozice noty v taktu.....	48
5.2.2	Detekce délky noty.....	48
5.2.3	Paměť kontextuálních událostí.....	48

5.2.4	Limitace změn parametrů	48
5.2.5	Volitelné křivky rozložení náhodné veličiny	49
5.2.6	Pokročilý generátor pseudonáhody	49
5.2.7	Determinismus.....	49
5.3	Otázka polyfonie	50
5.4	Otázka tempa.....	50
5.5	Rozšiřování kontextu.....	51
6	Aplikace.....	52
6.1	Standard VST Module Architecture	52
6.2	Analýza požadavků	53
6.3	Návrh aplikace.....	53
6.3.1	Parametry v MIDI.....	53
6.3.2	Determinismus a pseudonáhoda.....	54
6.3.3	Distribuční funkce a jejich implementace	54
6.4	Graphical user interface.....	56
6.4.1	Automatizace v DAW	57
6.5	Implementace	58
6.6	Licence	59
7	Závěr.....	60
7.1	Motivace	60
7.2	Zdokonalení humanizačních metod.....	60
7.3	Zhodnocení.....	61
	Rejstřík.....	62
	Použitá literatura	63
	Obsah elektronické přílohy.....	65

1 Úvod

Každý člověk disponuje určitou jedinečnou nebo alespoň specifickou sestavou kvalitativních vlastností. Jedinečnost každého člověka jej předurčuje reagovat na některé podněty jinak, než by na ty samé reagoval jiný jedinec. Stejně tak se dva různí lidé zachovají ve stejné situaci jinak – situace pro ně vlastně stejná není. Naproti tomu funkční počítačové programy jedinečné nejsou. Jsou naprogramovány tak, aby se dvě kopie též aplikace ve stejné situaci chovaly stejně. Žádoucí vlastností počítačového software je, aby byl přesný, sám od sebe nechyboval.

Software však na trh hudební produkce pronikl se značnou razancí. Díky cenově dostupnému hardwaru a uživatelsky přívětivému programovému vybavení si dnes počítače vydobýly místo coby nástroje hudební produkce jak v domácích podmínkách, tak v profesionálním prostředí.

Software pro hudební produkci se chová jako jakýkoliv jiný software – dělá přesně to, co mu zadáme, že má dělat. Nic méně, ale ani nic více. Hudba je však živoucí fenomén, plná zvrátů, obrátů a důmyslných interpretačních nuancí nebo i prosté nezáměrné chybovosti. Počítač můžeme přimět k tomu, aby respektoval hudební kompozici, poskytuje nám k tomu ostatně širokou paletu nástrojů. Mnohem složitější je ale přiblížit zvuk produkovaný na softwarové hudební platformě zvuku, jaký by vznikl při snímání živého hudebníka nebo hudebníků. Pomineme-li hudebně-akustické faktory, jakými jsou například přirozené doznívání ve skutečném akustickém prostředí, technologické aspekty (mikrofony a další vybavení) a různé možnosti realizace nahrávky, zůstává nám ještě jeden velmi významný aspekt – nuance ve hře živého interpreta.

Hráč na nástroj nebo zpěvák totiž pouze striktně nepřehrává dostupný kompoziční materiál, ale tento materiál studuje, vnímá a interpretuje – dává mu tvář a jiskru. Využívá svého citu, vědomí i obrátů, jimž se za léta své praxe jako interpret naučil. (1) Kromě toho člověk není schopen naprosto přesně ovládat své motorické funkce, dokonale sledovat a dodržovat například tempo, rytmus nebo využívat svého dokonalého hudebního sluchu, pokud jím vůbec disponuje, takže řada detailů, které se ve výsledné hře hráče nebo zpěvu zpěváka objeví, vzniká vlastně mimoděk a někdy jsou až nežádoucí. Z hlediska této práce jsou to ale právě tyto drobnosti, které odlišují hudbu produkovanou člověkem a strojově přesnou programovanou hudbu přehrávanou počítačem.

Hudbě hrané počítačem ale musíme rovněž přiznat některé výhody. Jak píše Daniel Forró ve svém článku o automatizaci, počítač zvládne v řadě ohledů mnohem více, než je v lidských silách. To je ostatně také jeden z důvodů, proč jich vůbec využíváme.

Elektronika tady má i další extenzní role – stroje dokáží zahrát mnohem rychleji, přesněji, čistěji, výš, hlouběji, hlasitěji... než lidští interpreti. (2)

Toho můžeme s výhodou využít, podaří-li se nám podchytit alespoň některá specifika lidskosti na poli hudební interpretace. Počítač pak může hrát ve svém nedostižném tempu či tonální

čistotě, ale přidáme mu schopnost lidsky (dále budeme využívat raději výrazu *pseudolidsky*) chybovat, čímž jej přiblížíme požadavkům na určitou živost nebo živoucnost naší hudební tvorby.

1.1 Premisy práce

Vstupními daty práce bude záznam hry kvalitního hudebního interpreta, protože, jak si vysvětlíme dále, jen u něj můžeme mluvit o skutečných, lidských nuancích hry, nikoliv pouze o chybách a technických nedostacích, které se ve hře projeví díky nedostatečnému zvládnutí nástroje.

Předpokládáme, že v materiálu najdeme určité charakteristické prvky, kdy se hráč bude v podobných situacích chovat podobným způsobem, nepřesnost určité techniky či figury bude ve srovnatelné situaci vždy víceméně stejná, tedy pro danou situaci charakteristická. (3) Tyto charakteristické body identifikujeme a pokusíme se je dát do souvislosti s hudebním kontextem, ve kterém se vyskytují.

Z toho vyplývá další předpoklad, totiž že identifikované charakteristiky bude možno strojově reprodukovat na základě počítačem provedené analýzy vstupního MIDI¹ proudu not, jakým může být například přepis notového záznamu nebo data vytvořená pomocí některého editoru MIDI.

1.2 Cíle práce

Cílem práce je napodobit lidské chování v hudební praxi provozované na počítači. Práce se omezuje na chování jednohlasých dechových nástrojů. Zaměření této části práce je na praktické využití metody skladateli realizujícími své (i jen dílčí, např. náslechové) práce pomocí virtuální instrumentace pomocí softwarového vybavení počítače.

Druhotným cílem je pak ukázat, že dosavadní metody používané pro humanizaci do MIDI zaznamenaného hudebního materiálu jsou z hlediska věrohodného napodobení hry živého hráče na dechový nástroj nedokonalé, a že v tomto směru existuje prostor pro jejich další zdokonalování a rozšiřování.

Dalším cílem je teoreticky postihnout problematiku humanizace na počítači komponované hudby, kdy bude zvláštní pozornost věnována MIDI záznamu dechových nástrojů. Je důležité, aby uživatel souvisejících postupů chápal, proč a kdy k nim saháme.

1.2.1 Zhodnocení cílů práce

Prvním krokem bude vyvození závěrů z analýzy materiálu zaznamenaného hudebníkem hrajícím na digitální dechový nástroj schopný přesného přenosu hráčova snažení do počítače. Jestliže bude náš předpoklad o charakteristickém chybování (viz část 1.1) správný, bude možné toto chování do určité míry reprodukovat při práci v prostředí hudebního software.

¹ MIDI (Musical Instrument Digital Interface) – standardizovaný protokol umožňující elektronickým hudebním nástrojům, počítačům a dalšímu vybavení společnou komunikaci a synchronizaci. (11)

Tuto reprodukci bude pro snadnou použitelnost provádět modul standardu VST Module Architecture firmy Steinberg resp. jeho aktuální verze VST 3.5.

2 Humanizace

2.1 Definice humanizace

Humanizace je v širším překladu polidštění, přiblížení se lidskému chování nebo lidským vlastnostem. (4) (5) V užším slova smyslu je pak termín humanizace rozuměn jako zanesení pro člověka charakteristické chybovosti do přesné, strojové, elektronické hudební produkce. To bude předmětem našeho zkoumání – jak přenést interpretační charakteristiky živého hráče do (nejčastěji) kvantizovaného či přímo programovaného (ve smyslu notace vytvořené v počítači např. v editoru MIDI) MIDI záznamu tak, abychom se výsledkem přiblížili uvěřitelnosti a příjemnosti produkce realizované člověkem, nebo se alespoň dostatečně vzdálili od produkce strojově přesné.

V rámci této práce budeme humanizací rozumět *záměrné vychýlení parametrů přesného a pro počítač srozumitelného notového předpisu takovým způsobem, že se jeho přehrání počítačem bude blížit hře živého interpreta.*

2.2 Definice kvantizace

Kvantizace je v našem případě zaokrouhlení parametrů MIDI záznamu na předem nastavený rastr. Jedná se o odstranění nepřesností v časech začátku tónu, jeho délce, výšce ladění, intenzitě a dalších.

Za kvantizovaný záznam můžeme považovat rovněž MIDI zapsané do počítače pomocí nehudbních prostředků, typicky myši či klávesnice, v nějakém MIDI editoru, jestliže jej záměrně nevytváříme nepřesný, ale dodržujeme implicitní přichytávání na rastr (což se týká zejména pozice začátku a konce tónu v čase).

Kvantizovaný bude také MIDI vygenerované nějakým notačním software na základě běžné partitury, protože každá nota má (až na malé výjimky) jasně danou dobu, kdy začíná a po jak dlouhou dobu zní, jakou má dynamiku atd.

2.3 Vymezení interpretace a humanizace

Nezbytné pro nás bude uvědomění si rozdílu mezi humanizací hudebního materiálu a jeho interpretací. Interpretací budeme rozumět širší nastudování skladby, uvědomění si souvislostí a vztahů mezi jejími jednotlivými stavebními prvky, a rovněž vědomé nerespektování některých dílčích aspektů předepsaných notovým zápisem (např. záměrná změna délky některých not kvůli budování nebo naopak uvolnění napětí). (6)

Rozpoznání systému vztahů jednotlivých složek hudebního stylu je zásadním činitelem toho, co nazýváme „rozuměním“ hudbě. (7)

Rozhodování v této oblasti totiž vychází z širší zkušenosti hráče s hudební interpretací, roli hraje zejména umělecká vyzrálost, hudební představitost, schopnost udržet vlastní osobitost a zároveň respektovat charakter díla, který mu vtiskl jeho skladatel. Nesmíme opomenout ani technickou zdatnost hudebníka, aby ho jeho případná nedostatečná dovednost nebrzdila v uchopení skladby tak, jak mu to ukládají výše zmíněné aspekty. Interpret se tak stává zprostředkovatelem hudebního díla mezi skladatelem a posluchačem. (8)

Interpretaci se pomocí počítače nebudeme pokoušet napodobovat. V našem případě totiž není motivací vyřadit lidský element z procesu hudební tvorby, lépe řečeno z její čistě nekompoziční stránky. Místo toho se setkáme s lidským zásahem a softwarovou humanizací na půli cesty. Od hudebníka (resp. tvůrce hudby) budeme očekávat zanesení nejzákladnějších interpretačních nuancí, jakou je kupříkladu změna dynamiky na úrovni jasně rozpoznatelných změn (např. piano – mezzopiano) nebo úprava délek not hudebního díla, na počítači pak necháme jemné odchýlení těchto parametrů v intervalech, které nenaruší skladatelský záměr, ale rozbijí strojovou strohost softwarové reprodukce díla.

Naší motivací pro takové řešení je totiž snaha vytvořit nástroj, který bude pro komponistu představovat jednoduchou, ale efektivní pomůcku pro narušení přílišné jednotnosti přehrávaného materiálu. Naopak nechceme vytvořit nástroj, který by o vlastní „vůli“ měnil skladatelský záměr do takové míry, že by bylo nutné vstupní MIDI data upravovat do nelogických konstrukcí jen proto, že se modul rozhodl určitou konstrukci interpretovat jiným způsobem, než jak ji vnímá tvůrce.

Zejména současná (i vážná) hudba je charakteristická snahou o narušení původních, tradičních struktur způsoby novými, neotřelými a neočekávanými. Šíře kontextu, který by bylo nutné analyzovat, by byla nepřehledná a nakonec bychom se smyčkou vrátili k původnímu problému – jak umožnit hudebnímu tvůrci realizovat něco nového, nečekaného, jestliže vycházíme z analýzy něčeho starého, již viděného a slyšeného.

Proto vytvoříme nástroj, který bude umožňovat kontrolované ovlivnění jednotlivých not i jejich celků na úrovni drobných změn jejich pozice, délky, ladění (na subtonální úrovni) a dynamiky. Skutečnou interpretaci, tedy vědomé ovlivnění hudebního díla a zásadní změny v jeho vyznění, ponecháme na realizátorovi. V dalším textu nebudeme již význam interpretace chápat v jeho širším smyslu, ale budeme jím rozumět zmíněné drobné nuance provázející hudebníkovu produkci, jež se budeme snažit softwarově napodobit.

2.4 Proč humanizovat už MIDI?

Přístupů k humanizaci hudebního materiálu nebo hudebního záznamu lze nalézt více. Jak se v dalším textu dozvíme, zvukový záznam hry živého hráče obsahuje určité artefakty, z nichž některé lze zanést i do existující nahrávky. Dnes je běžnou studiovou praxí „pomáhat“ zejména méně zdatným hudebníkům v kvalitě produkce tím, že se dovedným zásahem zvukaře, například stříhem, dorovná nedostatečná přesnost v čase nasazení tónu či úderu do perkusivního nástroje. Stejnou metodu lze úspěšně použít opačně – příliš dokonalou nahrávku můžeme takto záměrně degradovat a v některých případech tak dosáhnout posluchačsky zajímavějších výsledků, než jaké by se dostavily při poslechu materiálu neupraveného. Zde

máme na mysli spíše upravení nahrávky generované počítačem, ale ze zvukařské praxe jsou známy i případy hudebníků, jejichž projev je zejména pro sólovou hru příliš „suchý“.

Kromě střihačského zásahu přichází v úvahu také využití různých modulačních efektů, (9) které fakticky či pocitově upravují ladění tónu, jeho dynamický průběh (intenzita nasazení tónu, změna dynamiky v jeho průběhu a před jeho koncem a doznívání po jeho konci, dále například vytvoření efektu vibrata), řidčeji pak např. fázi zvukové vlny.

Tyto metody mají ovšem řadu společných nevýhod. V dnešní době, kdy je trh nasycen virtuálními nástroji² využívajícími matematické modely chování svých hardwarových vzorů, je škoda nevyužít jejich možností reagovat na chování předlohy, na jejímž základě svůj zvuk generují. Tou je již po mnoho let MIDI.

2.4.1 MIDI jako záznam výsledků motorické činnosti interpreta

MIDI totiž nemusí být zdaleka jen pro počítač čitelným a srozumitelným notovým záznamem určeným k reprodukci partitur. Již starší hardwarové syntezátory, jež neměly vlastní klaviaturu, komunikovaly se svými řídicími kontroléry právě pomocí protokolu MIDI (týkalo se to samozřejmě i syntezátorů, jež klaviatury měly, ale umožňovaly například odesílat data na ni hraná do dalšího zařízení nebo naopak přijímat řídicí povely z jiného zdroje). Klaviatury patřily mezi první tzv. MIDI kontroléry³, které umožnily přenášet aktivitu hráče ve formě jakýchkoli událostí, nikoliv pouze generovat a zaznamenávat výsledný zvuk nástroje. (10)

Dnes existují snímače, které jsou schopné převést hru na kytaru na digitální seznam událostí, jež se na nástroji odehrávají (začátek chvění struny, intenzita tohoto chvění, ladění struny a mnoho dalšího). Nejedná se tedy o záznam zvukový, ale záznam toho, co hráč na nástroji provádí, resp. toho, jaký to má na nástroj dopad. Kromě kytary jsou dnes dostupné kontroléry dechové, bicí i perkusivní, snímače na další strunné nástroje (jakými jsou např. housle, violoncella atp.), MIDI lze také generovat analýzou zvukových impulzů, např. hlasu.

Klíčovou myšlenkou je, že jsme schopni oddělit činnosti, které provádí hráč na daném nástroji, od výsledného zvukového projevu – dokážeme pořídit poměrně obsáhlý záznam pohybu prstů na klávkách dechového kontroléru, intenzitu dechu a další, aniž bychom tato data museli složitě získávat analýzou záznamu zvuku. Takový materiál pak můžeme analyzovat a matematicky a statisticky popsat. Z pohledu humanizace tak získáme představu o tom, jak se živý interpret v určitých hudebních situacích, figurách a obratech chová, a toto chování pak přenést na kvantizovaný záznam. To v ideálním případě provedeme pomocí specializovaného modulu pro naše vývojové prostředí.

² Virtuální nástroj – typicky softwarové vybavení počítače, které na základě (nejčastěji) MIDI vstupu generuje slyšitelný zvuk. V širším slova smyslu můžeme za virtuální nástroj považovat jakýkoliv generátor zvuku, jehož výstup je možné ovlivnit datovým vstupem, tedy např. i externí hardwarový syntezátor.

³ MIDI kontrolér – hardwarový nástroj často identický nebo velmi podobný nějakému hudebnímu nástroji, který sám o sobě negeneruje zvuk, ale hru hráče předává k dalšímu zpracování v podobě MIDI dat.

2.4.2 Oddělení od producenta zvuku

Nespornou výhodou je oddělení hudebního předpisu od výsledného zvuku. Nejenže hráč na piano není při použití MIDI klávesového kontroléru vázán pouze na klavírní rejstříky, ale zejména je zachována plná upravitelnost předloh výsledného hudebního materiálu. Můžeme se k nim kdykoliv vracet, části z nich kopírovat a přesouvat, využívat je jako zdroj partitur pro jiné nástroje v aranži nebo jen využít modernější virtuální nástroj, jenž je díky pokroku ve vývoji schopen reagovat citlivěji na některé nuance interpretovy hry, které byly starším, méně dokonalým nástrojem nedostatečně zohledněny nebo zcela opomenuty.

Zejména poslední uvedený faktor je z hlediska humanizace klíčový. Nabízí se zde polemika, zda by nebylo lepší modifikovat už výsledný zvuk v takové míře, že by dojem z humanizace byl lepší než ten ze zvuku generovaného virtuálním nástrojem na základě pouhého humanizovaného MIDI. Zde si ale musíme uvědomit předchozí argumentaci, kdy takové úpravy přímo ve zvuku jsou kromě stříhu značně obtížné a nedokonalé; uvážit musíme rovněž skutečnost, že už nyní existují vyspělé nástroje kombinující metody generování zvuku s matematickým modelováním tak, že výsledný zvuk je velmi věrný. I při použití kvalitního sampleru⁴ od něj můžeme očekávat poměrně věrnou nebo alespoň z tvůrčího hlediska zajímavou reakci na přes MIDI předané parametry. Lze navíc očekávat, že vývoj bude v tomto směru dále pokračovat.

2.5 Barva tónu a jeho další kvality

Ve spojení se zvukem jakéhokoliv nástroje nemůžeme opomenout jeho významný kvalitativní faktor, kterým je tzv. barva. Barvu zvuku resp. tónu můžeme charakterizovat jako frekvenční rozložení zvuku nástroje ve zvukovém spektru. Podíl na ní má množství a frekvence vyšších harmonických složek, které vznikají vedle základního (hráčem žádaného) tónu (někdy také uváděné jako *aliquótní tóny*).

Barva zvuku souvisí s řadou hráčských technik, například silou úhozu či buzení dechem, kdy lze očekávat nárůst vyšších harmonických složek a to nejen v amplitudě, ale i ve spektru jako takovém. Na druhou stranu MIDI standard General MIDI⁵ nepředepisuje o barvě tónu, (11) navíc tolik specifické pro ten který nástroj. Tyto informace jsou schopné pojmout standardy některé z nástaveb General MIDI, (12) jejich podpora v dnes používaných virtuálních nástrojích je však zanedbatelná.

Ač je barva vedle frekvence, délky a amplitudy z fyzikálního hlediska nedílnou složkou každého tónu a tedy by se dalo očekávat, že jí budeme v rámci zkoumání humanizačních mechanismů věnovat značkou pozornost, z pohledu humanizace MIDI dat je to informace naprosto nepodstatná. Barvu tónu a její reprodukci ponecháme na vhodném virtuálním nástroji.

⁴ Sampler – generátor zvuku, který nevyužívá zvukové syntézy, ale přehrává vzorky zvuků získané z nahrávek či jen obecně zvukových záznamů.

⁵ General MIDI – v současnosti nejpoužívanější MIDI standard

2.6 Jaké parametry MIDI lze humanizací ovlivňovat?

General MIDI je poměrně široký standard a při hře na MIDI kontrolér se kromě zjevné informace o začátku a konci noty předává na MIDI datový výstup celá řada dalších parametrů. Teoreticky bychom samozřejmě mohli uvažovat humanizaci všech z nich, nicméně některé slouží jen jako pomocné parametry nebo neodpovídají vlastnostem nástrojů, z nichž kontroléry svou podobu čerpají. Pro účely humanizace, tedy nápodoby skutečných hráčských vlastností, má pro nás smysl zabývat se jen těmi z nich, které přenášejí informace nějakým způsobem charakterizující hráčskou činnost na nástroj skutečný. V širším pohledu bychom mohli vzít jako referenci hru na MIDI kontrolér jako takový a nedívat se na ni například jako na hru na saxofon, nicméně tímto směrem bychom zbytečně rozšířili kontext našeho zkoumání na úkor skutečně podstatných informací předávaných MIDI kontrolérem.

Dále si uvedeme, které jsou nejdůležitější parametry, jejichž modifikací se budeme zabývat, a nastíníme si, jakým způsobem bude mít smysl je ovlivňovat. Detaily těchto způsobů budou samozřejmě vycházet z vlastní analýzy záznamu reprodukce etud hudebníka a dostaneme se k nim podrobně v kapitole 4.

2.6.1 Pozice začátku

Umístění začátku znění noty v čase je pravděpodobně nejvýraznějším faktorem ovlivňujícím vnímání toku hudebního materiálu posluchačem. Je totiž přímo svázáno s rytmotvornou funkcí, kdy rytmus, resp. celá tzv. *kinetická složka*⁶, je jedním z nejdůležitějších prostředků hudebního výrazu. (13)

Jak uvidíme v analytické části této práce, téměř žádná nota není hrána přesně v časové pozici, která odpovídá času jí předepsanému v notové předloze. Míra variace tohoto parametru charakteristická pro lidskou interpretaci se navíc liší dle kontextu, v jakém se ta která nota vyskytuje. Není dána jen tempem (kde si například i laik dovede představit, že v určitých rychlostech už hráč nestihá tak přesně měnit pozice prstů, nasazení dechu atp.), ale i pozicí v taktu, náročností okolních partů či jednotlivých sousedních not.

Ovlivněním parametru pozice začátku znění MIDI noty se nám snadno podaří narušit nejhůře vnímaný strojový efekt slyšitelný při přehrávání úplně kvantizovaného záznamu – začátky not mají z definice kvantizace přesně stejné časy začátků relativně ke zvolenému rastru. Ve spojení s konstantní dynamikou se nám k poslechu dostává interpretačně „mrtvý materiál“.

Pozice začátku noty je z pohledu MIDI dat moment, kdy MIDI kontrolér obdrží budicí signál (v případě dechového kontroléru EWI je to sepnutí proudem vzduchu v trubici za náústkem) a na svůj MIDI výstup předá začátek MIDI události Note On společně s její výškou a dalšími parametry.

2.6.2 Délka noty

Délku noty můžeme snadno zaměnit za pozici konce znění noty, tedy za okamžik, kdy kontrolér přestane dostávat budicí signál a ukončí MIDI událost hrané noty. To by však platilo jen

⁶ Kinetická složka – rytmická, tempová a metrická složka. (13) (Dalšími složkami jsou pak melodická a harmonická.)

v okamžiku, kdy by nástup noty byl zcela přesný (na rastr). Ve chvíli, kdy interpret začíná tón mimo tento rastr, může být taková záměna značně zavádějící.

Délka není tak významným faktorem ovlivňujícím subjektivní pocit z vnímání humanizované skladby. Do jisté míry zasahuje do rytmické funkce přesnosti nástupu noty, ale do jejího významu zasahuje řada dalších faktorů včetně toho, že nástup tónu je jak fyzikálně (dechem hráče) tak i pocitově (změna ladění, setkání se s očekávanou dobou, změna zvuku z ticha) více akcentován. Svou roli na druhou stranu sehraje ve chvíli, kdy se budeme zabývat rychlými sledy krátkých, značně rytmizujících not – například v případě arpeggií, staccatových běhů či jednoduše velmi rytmizovaných partů s důrazem na vysokou přesnost. Posluchač zde bude vnímat výpadky na úrovni nižší, než je pouze očekávaná délka noty vyjádřená v dobách a tím předpovězená hudebním kontextem.

2.6.3 Subtonální výška noty

Pro dechové nástroje jsou charakteristické určité odchylky od přesného ladění. Je to dáno nutností doladit výšku tónu dechem bez nějaké motorické rutiny, jakou si vypracují například hráči na bezpražcové smyčcové nástroje (housle, violoncella atp.). Z dechových nástrojů má nejvyšší potenciál znít přesně trombon, neboť změna tónu je v něm realizována použitím snižce, kde je zmíněnou motorickou rutinu možné vyvinout.

Jestliže se tedy budeme zabývat humanizací průběhu tónu dechového nástroje, má smysl se věnovat změně tzv. mikroladění⁷ během jeho znění. (14) Dechový kontrolér EWI naneštěstí nedisponuje možností snímat drobné odchylky v ladění, neboť výška tónu je získávána výhradně z hráčova prstokladu. V tomto případě se tedy odkážeme na vhodnou literaturu, která změnu ladění tónu v jeho průběhu popisuje.

2.6.4 Dynamika⁸ noty

Dynamika se ve spojení s pozicí začátku podílí na vyznění tónu nejvíce. Už ze samotného pojmu dynamiky vyplývá, že jeho výraznost přesně odpovídá hlasitosti, s jakou se ozve. (15) Při humanizaci bude klíčovým problémem vhodné narušení konstantní dynamiky jednotlivých tónů hudebního materiálu.

Dechové nástroje patří mezi nástroje kontinuálně buzené, protože zvuk v nich vzniká průchodem vzduchu tělem nástroje a zaniká záhy po jeho skončení. Na rozdíl od např. nástrojů bicích perkusivních, kdy nás zajímá zejména dynamika buzení tónu (jeho průběh nebývá až na výjimky hráčem ovlivněn), v případě dechových nástrojů se můžeme dívat na tón během celé délky jeho znění. Kromě dynamiky nasazení tónu budeme pozorovat také změnu dynamiky během celého buzení tónu změnou síly proudu vzduchu, který hráč vdechováním do nástroje vytváří.

Lze očekávat, že nasazení tónu bude mít o něco vyšší dynamiku, (16) ačkoliv její změna bude mít v krátkém čase vzestupnou tendenci – dech nebude mít ihned maximální výšku. Dále bude tón v dynamice o něco klesat a začne kolísat v závislosti na tom, jak vyrovnaně bude interpret schopen vdechovat do nástroje další vzduch.

⁷ Mikroladění – ladění na úrovni menší než nejmenší hudební interval, což je v západní hudbě půltón.

⁸ Dynamika – změna síly či hlasitosti tónů či částí skladby. V případě hudebních nástrojů je obvykle spjata se silou úhozu, intenzitou buzení struny či plátku atp.

Tón po konci buzení dechem poměrně rychle zaniká. (17) Zbytkové rezonance jsou dané fyzikálními vlastnostmi toho kterého nástroje a jsou spíše věcí virtuální instrumentace, aby se věrohodným dozníváním případně zabývala. V analytické části se ale podíváme i na to, jakým způsobem hráč ukončuje nasazení dechu a jak strmý je jeho sestup.

3 Sběr dat

Výzkumná část práce se zabývá získáním vhodných vstupních dat pro jejich následnou analýzu (viz kapitola 4 na straně 23). Je částí především praktickou. Na teoretické rovině bylo nutné především vybrat vhodné dechové nástroje, jim určené partitury a interpreta, který bude schopen tyto partitury adekvátně nastudovat a nahrát. Výstupem této části je zisk několika nahrávek stejných skladeb nebo jejich částí tak, aby bylo v analytické části dostatek materiálu k rozhodnutí, které hráčské obraty jsou interpretačně charakteristické. Paralelně s nimi jsme ještě vytvořili přesný přepis studovaných partitur tak, aby byly nástroje pro hudební komparatistiku schopné srovnání nahrávek interpreta s kvantizovanými přepisy.

Syrová (neupravená) vstupní data jsou součástí elektronické přílohy diplomové práce jako soubory MIDI.

3.1 Kritéria pro výběr interpreta

Záměrem práce bylo postihnout interpretaci možnosti hráčů různých úrovní technické i interpretační zdatnosti. Z konzultací s prof. Václavem Kuntem však vyplynulo, že o skutečné schopnosti interpretace odpovídající hudebně estetickým nárokům, která bude navíc nějakým způsobem charakteristická, můžeme hovořit až u kvalitních, ve svém nástroji dobře vzdělaných interpretů. To platí zejména v případě, že chceme dosáhnout vytvoření seriózní nahrávky určité hudebně kulturní úrovně – není pro nás totiž žádoucí vytvářet náhodné, kolísající a technicky nedostatečné party dechových nástrojů.

Z tohoto důvodu jsme se nakonec přiklonili k vyhledání zdatného, interpretačně vyzrálého hráče na dechové nástroje. Nezbytným požadavkem bylo absolutorium ve zvoleném nástroji nebo vysokoškolské vzdělání příslušného oboru. Zároveň jsme hledali aktivního a v ideálním případě i profesionálního hráče – tedy takového, pro kterého je hra na dechový nástroj denní rutinou. Hráč musel být navíc schopen cvičením překonat drobné odlišnosti ve způsobu hry na svůj nástroj a na dechový kontrolér EWI.

Jen tak můžeme považovat vstupní data pro analýzu za relevantní k pozdější implementaci vylepšené humanizační metody vzniklé na jejich základě a to takovým způsobem, že užití této metody bude dostatečně obecně platné a praktické nasazení nebude zbytečně úzce profilované.

3.2 Volba nástroje

3.2.1 Dechové nástroje

Dechové nástroje (aerofony) pracují na principu rozechvívání vzduchu v dutině nástroje. Dělí se na nástroje žesťové, kdy je vzduch rozechvíván rty (lesní roh, trubka, pozoun, tuba), nástroje dřevěné (klarinet, saxofon, fagot, flétna), u nichž je chvění způsobeno působením vzduchu na plátek resp. třením o hranu nástroje (flétna).

Pro úplnost musíme ještě zmínit existenci těch nástrojů, které jsou sice rozeznávány proudem vzduchu, ale ten do nich není vhnán tělem (plicemi), ale jiným způsobem. Zástupci těchto nástrojů jsou například akordeon, harmonium či varhany.

V rámci měření interpretačních technik jsme byli omezení na možnosti nástroje EWI USB firmy AKAI Professional. Způsob hraní na něj se nejvíce blíží hře na nástroje dřevěné, konkrétně flétnu, či saxofon, pro něž disponuje i softwarovou změnou prstokladu (aktivace a deaktivace některých snímacích dotykových ploch a reakce na určité vzory kontaktu s nimi).

3.2.2 Volba konkrétního dechového nástroje

Na základě konzultací s prof. Václavem Kuntem a doc. Romanem Novozámským jsme dospěli k závěru, že nejvhodnějším nástrojem z hlediska širě vzhledu do palety dechových nástrojů bude klarinet nebo saxofon.

Problém klarinetu v souvislosti s naším záznamovým kontrolérem nicméně spočívá v přefukování do duodecimy⁹ nikoliv do oktávy, jak je tomu například u saxofonu a hlavně u nástroje EWI. Ten není schopen rozpoznat přefuk podle intenzity proudu vzduchu (pouze vydá na datový výstup maximální velikost kontroléru CC-2 – Breath Control), ale ke změně oktáv dochází posunem palce po oktávových transpozičních válečcích na zadní části nástroje. Z tohoto důvodu padla konečná volba na saxofon coby nástroj nejpodobnější snímacímu zařízení, které jsme měli k dispozici.

Saxofon a klarinet jsou si navíc velmi podobné co do hraného materiálu, etudy pro klarinet není pro saxofon třeba transponovat a z technického hlediska tak bylo možné vybrat i některé party určené pro klarinet a zahrát je saxofonovými hmaty na kontroléru EWI. Proč nám tato extrapolace nevadí, si uvedeme dále v části 3.4 Příprava interpreta.

3.3 Výběr hudebního materiálu

Jestliže chceme analyzovat paletu základních hudebních situací a technik, budeme potřebovat hudební díla nebo jejich části, které hru těmito technikami obsahují. Zároveň si musíme stanovit reálné limity, abychom vhodně omezili kontext, ve kterém se budeme při analýze pohybovat tak, abychom dospěli k relevantním výsledkům a nebyli jen zavaleni množstvím materiálu, v němž bude obtížné najít jakékoliv souvislosti. (18)

Ve spolupráci s interpretem jsme vybrali etudy a části skutečných hudebních děl určených ke koncertní produkci. Partitury jsou různé interpretační náročnosti, snadné skladby jsme se rozhodli vypustit. Výběr tedy kolísá od skladeb středně technicky náročných až po ty velmi náročné.

Jednotlivé partitury si přiblížíme v následujícím textu, stejně tak si uvedeme, na co se chceme v té které zaměřit. Partitury jsou přílohou této práce.

⁹ Interval složený z oktávy a kvinty, tedy 19 půltónů v rovnoměrně temperovaném ladění.

3.3.1 Fritz Kroepsch: 167 Exercises for Daily use

Jedná se o sérii tří etud technického charakteru střední až vyšší náročnosti. Pro záznam jsme vybrali etudy 1., 6. a 9.

První etuda v 4/4 taktu a šestnáctinových notách nám umožní zkoumat rytmickou přesnost ve vyšším tempu. Cvičení je navíc hráno technickou staccato¹⁰, kdy bude tón vzhledem k tempu ukončován přiložením jazyka na náústek, (17) nikoliv pouhým přerušením dechu.

Šestá etuda je hrána ve svazovaných sextolách vždy po dvanácti šestnáctinových notách s drobnou změnou dynamiky (zvýšení) na konci.

Devátá etuda s dvaatřicetinovými notami je velmi náročná na přesnost a můžeme v ní dobře vysledovat chování zejména přesnosti nasazení tónu. Protože se často nemění výška hrané noty, můžeme izolovat dynamický průběh od chyb vzniklých vlivem jiných faktorů, například právě změnou hmatu.

3.3.2 Alexander Borodin: Polovecké tance

V tomto případě sledujeme klarinetový part jako část většího hudebního celku. Pozornost můžeme věnovat občasnému výskytu akcentu marcato¹¹ a techniky staccato. Part se v závěru více rytmitizuje a předkládá interpretačně náročnější figury, zároveň zahrnuje dlouhé noty (půlové i celé), na nichž můžeme zkoumat dynamiku průběhu delšího tónu, navíc v běžném toku skladby, což bude jeho nejpřirozenější výskyt.

3.3.3 Bedřich Smetana: Má vlast – Šárka

Jedná se o krátké, velmi expresivní sólo ze Smetanova díla. Part kolísá v dynamice a uvádí řadu zajímavých akcentů a krátkých přechodových tónů, u nichž se dá očekávat velká míra interpretačního „ulehčení“ kvůli posluchačsky přívětivému vyznění daného místa, toto ulehčení bude ale na druhou stranu pro hudebníka charakteristické.

3.3.4 Béla Bartók: Concerto of Orchestra – I. Introduzione

Vláčné sólo, předepsané dolce, ukazuje hru v klidném tempu se řadou svazovaných not, navíc v méně obvyklém 3/8 taktu. Kolísání v dynamice je malé, dobře bude vidět průběh tónu svázaného přes tři předposlední takty.

3.3.5 Leoš Janáček: Glagolská mše

Dvě sóla z Janáčkovy Glagolské mše ukazují interpretovu práci v pomalejším tempu, ve výškách a rovněž práci s dynamikou v průběhu partu.

3.3.6 Nikolai Rimsky-Korsakow: Capriccio espagnol

Rychlejší sóla v krátkých notách s využitím melodické ozdoby – trylky¹². Jedná se o technicky náročnou pasáž, méně zkušení hráči od hraní trylky upouštějí. Začátek druhého sólového partu pak provází zajímavé střídání hry legato¹³ a staccato.

¹⁰ Staccato – krátce, odtrženě. (33)

¹¹ Marcato – se zdůrazňováním, důrazně. (33)

¹² Trylek – tón se po určité době rychle střídá s obvykle vyšším tónem.

¹³ Legato – vázaně. (33)

3.3.7 Petr Iljič Čajkovskij: Masky

Další expresivní sólo obsahující především rytmické změny a rychlé nástupy, dále můžeme sledovat průběh ozdob a změny způsobu hraní při přechodu z rychlých dvaatřicetinových běhů do táhlých, dlouhých tónů.

3.4 Příprava interpreta

K interpretaci jsme navázali spoluprací s prvním klarinetistou Filharmonie Brno MgA. Lukášem Daňhelem, Ph.D. Vedle klarinetu je i zkušeným hráčem na saxofon se zájmem o digitální dechové nástroje. Doktor Daňhel měl po výběru etud nástroj k dispozici k dvoutýdennímu cvičení, které bylo dostatečné ke zvládnutí jak určitých odchylek v ovládnutí nástroje oproti saxofonu (místo přefuku oktávové válečky a pevné dotykové ovládnutí místo klapky saxofonu, které pod prsty mírně uhýbají), tak i nastudování vybraných etud na úrovni běžné interpretace.

Vzhledem k tomu, že u interpreta se jednalo o pouhou nutnost technicky si osvojit nový nástroj podobný těm, na které je zvyklý hrát, nebyla nijak dotčena jeho schopnost a dovednost hudební interpretace jako takové. V opačném případě by nám nahrávky poskytly relevantní výsledky jen stěží.

Ozvučení kontroléru v rámci přípravy bylo zajišťováno virtuálním nástrojem ARIA Instrument Engine od firmy Garittan, který je dodáván přímo pro kontrolér EWI a obsahuje kvalitní knihovny instrumentálních rejstříků většiny běžných dechových nástrojů. Ke cvičení byly pro věrohodnost odezvy využívány klarinet (Eb) a saxofon (většinou sopránový).

3.5 Záznam

3.5.1 Vybavení

Záznam interpretace jsme provedli na výkonném 64bitovém systému s operačním systémem Microsoft Windows 7. Latenci¹⁴ běžného zvukového softwarového vybavení systému Windows jsme kompenzovali použitím externí zvukové karty Konnekt 8 firmy TC Electronic a ovladačů ASIO. Ovladači uváděná latence byla 5,2 ms na vstupu i na výstupu, tedy 10,4 ms celkem; interpret byl s reálnou rychlostí odezvy systému spokojen, v pohodlné hře ho dle jeho slov nijak neomezovala.

Data z MIDI kontroléru byla dále vedena do DAW¹⁵ Cubase 5 firmy Steinberg a zaznamenána bez jakýchkoliv úprav či vyřazování pravděpodobně neužitečných parametrů přímo na vstupu. V hostitelské aplikaci byl k ozvučení použit stejný software, který měl k dispozici interpret při cvičení; samotný zvuk byl ale užitečný pouze pro odposlech interpretem, k analýze MIDI materiálu je zvuk toho kterého nástroje irelevantní.

¹⁴ Latenci v tomto případě myslíme zpoždění mezi hráčovou akcí provedenou na nástroji a ozevem zvuku z výstupního zvukového zařízení, která je při využití virtuální instrumentace běžná u nespécializovaných a méně výkonných počítačů.

¹⁵ DAW – Digital Audio Workstation. Hostitelská aplikace pro další práci s MIDI, zvukem a zejména virtuálními nástroji (např. na bázi samplování nebo syntézy).

3.5.2 Metronom a práce s tempem

Pro typickou sólovou nebo i orchestrální hru je běžné určité kolísání v práci s časem hry, (19) v našem případě jsme si ale něco takového nemohli dovolit – získat přesné odečty odchylek jednotlivých parametrů by bylo nemožné. Proto byly vybrané party zaznamenány tzv. na klik, interpretovi tedy zněl ještě metronom a udával předepsané tempo skladby, kterého se měl hráč nutně držet.

3.5.3 Způsob vlastního nahrávání

Abychom mohli z nahrávek vyvodit nějaké relevantní statistické závěry, bylo nutné každý part nahrát více než jednou. Vzhledem k časové vytíženosti interpreta jsme byli nuceni nahrávat všechn materiál v jediném časovém bloku a naplánovat tedy nahrávání tak, aby nedocházelo ke zbytečnému zkreslení důsledkem únavy, vyčerpání či prostě přílišného se ponoření do interpretovaného partu a ztráty určitého hráčského nadhledu, který v nahrávkách požadujeme.

Pro analýzu pro nás navíc není klíčové dívat se mnohokrát na tu samou notu, řada situací je konzistentní s jinými výskyty v té samé skladbě, některé etudy jsou dokonce celé zaměřené na jeden konkrétní jev, takže i při relativně malém počtu nahrávek můžeme z analýzy zobecnit závěry o chování hráče v průběhu skladby.

Každou skladbu nebo její část jsme se proto rozhodli nahrát čtyřikrát až pětkrát. Pět cyklů jsme volili spíše u kratších segmentů, u delších byla patrná určitá únava a „přecvičenost“ interpreta již po čtvrtém cyklu, některá data proto musela být vyřazena. Již při čtyřech cyklech je navíc dobře vidět, zda je určitá interpretační nuance pro hráče charakteristická a má proto smysl se jí z hlediska humanizace kvantizovaného MIDI materiálu hlouběji zabývat.

Mezi jednotlivými cykly jsme dělali přestávky pro vydýchání, údržbu nástroje nebo krátké procvičení obtížného partu podle požadavků hráče.

3.6 Kontrolní přepis

Vedle vlastního záznamu hráčovy interpretace vybraných skladeb bylo nutné vytvořit přesný přepis jejich partitur. Tento přepis je pro naše potřeby volně zaměnitelný za kvantizovaný záznam, neboť většina parametrů je konstantní nebo partiturovou determinovaná.

Jako předlohy byly využity originální partitury hudebních vydavatelství, nejedná se tedy o žádné učebnicové ani jiné neúplné či jakkoliv editované přepisy. Samotný přepis byl proveden v notačním software Finale firmy MakeMusic. Z něj byla vedle samotné notace (ta je pro představu o vybraných partech přílohou této práce) exportována MIDI data jako standardní MIDI soubory.

4 Komparatistika a analýza

Tato část práce se věnuje srovnání dat získaných způsoby popsanými v kapitole 3 a jejich analýze s cílem vyhledat v materiálu nuance, jež by bylo možné označit za interpretačně charakteristické.

4.1 Získání čitelných dat

4.1.1 Export MIDI

Prvním krokem bylo odstranění hodnoty kontroléru Aftertouch, protože jeho hodnoty se duplikovaly s hodnotami kontroléru CC-2 Breath Control. Oba disponují stejným rozlišením 0–127, (11) takže nebyl důvod využívat kontrolér Aftertouch k tomu, k čemu jsme ve skutečnosti chtěli využít parametr ukládající sílu dechu – a tím je zjištění dynamického průběhu tónu. Pro další zpracování bylo mnohem přehlednější vyřadit jeden ze zdvojených parametrů již v tomto kroku.

Dále byl všechen zaznamenaný materiál exportován do samostatných MIDI souborů. Pro co největší přiblížení nahrávek a přepisu byl i tento převeden v prostředí DAW Cubase 5 a znovu exportován pro použití v dalších bodech procesu.

4.1.2 Převod na text

Nahraná i předepsaná MIDI data jsou v prostředí Cubase dobře viditelná, při dostatečném přiblížení je možné pouhým okem pozorovat jemné rozprostření nepřesností kolem udávaného rastru – přesně jevy, jimž chceme věnovat více pozornosti.

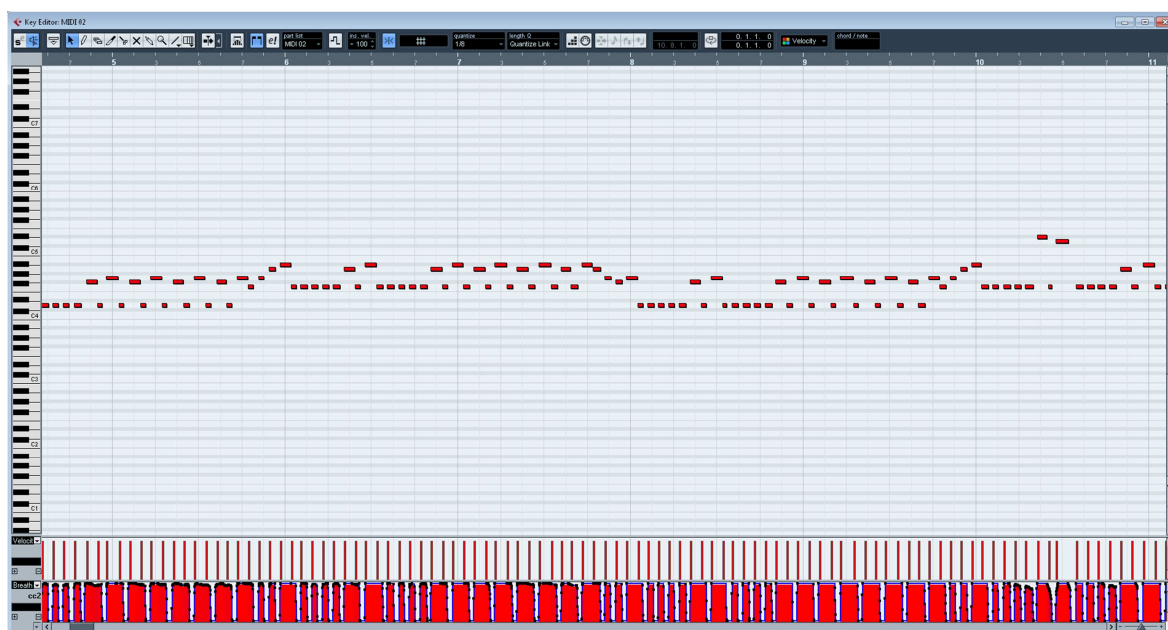
Na druhou stranu je v této podobě nemožné odečíst přesně odchylky parametrů od rastru (viz Obrázek 4.1). Pro tyto účely bylo nutné jednotlivé MIDI soubory převést do textové podoby. K tomu jsme využili volně šiřitelné nástroje z rodiny GNMIDI¹⁶. Po několika testovacích bězích nad předem známými krátkými (obsahujícími malé množství datových MIDI zpráv) shledali jako nejvhodnější nástroj MIDI Note, který na výstup dává podrobný výpis každé události v MIDI souboru tak, jak v něm za sebou následují.

Program je jako celá General MIDI norma poměrně starý, je určený ještě pro 16bitové počítače vybavené operačním systémem Microsoft DOS. Pro jeho emulaci jsme použili prostředí open source projektu DOSBox¹⁷. V něm bylo možné řádkový výstup nástroje MIDI Note přeměňovat do textového souboru stejným způsobem jako v původním MS-DOS.

Výsledkem tohoto kroku byl tedy textový soubor obsahující sekvenční výpis zpráv (událostí) obsažených ve vstupním MIDI souboru formátovaný takovým způsobem, že na každém řádku se nacházela jedna zpráva a jí příslušející parametry.

¹⁶ Nástroje GNMIDI jsou dostupné na adrese <http://www.gnmidi.com/>.

¹⁷ Projekt DOSBox je dispozici na adrese <http://www.dosbox.com/>.



Obrázek 4.1 – MIDI data zobrazená v editoru typu piano roll

4.1.3 Zpracování textového přepisu MIDI souborů

Výpočty prováděné nad MIDI daty budeme provádět v programu Microsoft Excel 2010. Aby bylo možné tato data efektivně vkládat do tohoto tabulkového procesoru, bylo třeba textový výstup nástroje MIDI Note znovu převést.

Dalším důvodem bylo strojové převedení hodnot parametru CC-2 Breath Control na parametr Velocity, který v MIDI uchovává informaci o dynamické noty. (11) Parametr Velocity totiž kontrolér EWI předává implicitně konstantní v hodnotě 120 (což je v hudební terminologii přibližně fortissimo possibile¹⁸ (11)), takže noty jako takové nástupní dynamiku nemění. Co v praxi způsobuje kýženou změnu dynamiky, je právě interpretace parametru CC-2 Breath Control nástrojem pro přehrávání MIDI dat jako parametru dynamiku obsluhujícího. Pro určení nástupní hodnoty dynamiky konkrétní noty jsme použili vážený průměr hodnot změn parametru CC-2 Breath Control, ke kterým docházelo během celé délky noty. Tato hodnota pro základní představu o dynamice kratší noty díky vlivu délky noty na posluchačovu percepci dynamické změny postačí. (20)

Pro analýzu průběhu dynamiky dlouhých not jsme promítli změnu parametru CC-2 Breath Control na osu s rozlišením na milisekundy, takže budeme moci později vykreslit detailní křivku průběhu dynamiky kontroléru. Nejprve jsme vybrali vhodné noty, tedy noty s délkou alespoň jednu dobu. Zajímala nás pak poslední hodnota změny parametru CC-2 Breath Control těsně předcházející zprávě Note On a všechny ostatní změny parametru až do konce délky jejího znění (včetně). Tam, kde to bylo možné, jsme zaznamenali i průběh dechu do úplného upadnutí na nulovou hodnotu (tzn. noty následované pomlčkou nebo na konci partu).

¹⁸ Fortissimo possibile – tak silně, jak je to jen možné. (33)

Pro tento převod jsme naprogramovali jednoduchý parser v jazyce PHP¹⁹ (kvůli rychlosti vkládání a odběru dat je proces realizován jako sada jednoduchých webových aplikací rozlišených podle požadované podoby výstupních dat; navíc odpadají u takovéhoho proprietárního řešení zbytečné komplikace s datovými typy). Parser ze vstupního pole přečte obsah textového souboru, vynechá všechny řádky obsahující irelevantní informace a na výstup předá pro Excel srozumitelný formát hodnot oddělených tabulátory, který lze do tabulkového procesoru přímo vložit.

4.1.4 Výpočty a grafy v Excelu

Ostatní operace byly provedeny v prostředí Excel, které umožňovalo rychlé výpočty nad vloženými hodnotami a vykreslení grafů ilustrujících průběh odchylek při interpretaci hudební předlohy u jednotlivých not.

4.2 Postup při analýze

Pro všechny skladby jsme vypracovali obsáhlé tabulky umožňující přehledné a rychlé sestavení získaných dat do statisticky vhodných celků, z nichž jsme následně vypracovali grafy poskytující vizuální východisko pro efektivní identifikaci charakteristických artefaktů interpretovy hry.

4.2.1 Sestavení statistických tabulek

Vypracovali jsme dva druhy tabulek – první zkoumá pouze dynamický průběh dlouhých not, druhý pak všechny ostatní parametry (vyjma mikroladění, jak jsme si objasnili dříve).

Tabulka prvního typu má ve sloupcích vyneseny jednotlivé noty, řádky tabulky pak představují milisekundy průběhu jejich znění. Protože zkoumané noty se délkou pohybovaly přibližně v rozsahu od 1,7 sekundy do 3 sekund, vytvořili jsme ještě paralelní měřítko počítající průměr z každých deseti řádků, abychom získali přehlednější křivku v grafu. Zde nebylo nutné použít průměr vážený, protože parser, ze kterého jsme vstupní data přebírali, vyplňuje i ty časové jednotky, které žádnou změnu dynamiky neobsahovaly (použitím poslední předchozí známé hodnoty), řada tedy nevytváří žádné časové nesouvislosti, které by bylo případně nutné zohlednit. Pro každou notu jsme z pěti nahrávek téže skladby vypočetli rovněž průměr parametru (a také jeho průměrované kroky) ze všech hodnot na témže řádku tabulky – to znamená v tomtéž čase.

Na základě vypočítaných hodnot jsme pak vytvořili graf *Průběh dynamiky* tónů v čase (horizontální osa). Dynamika, která je v případě dechového kontroléru EWI obsluhovaná parametrem CC-2 Breath Control, je vyjádřena jednotkami z rozsahu 1–127 a můžeme ji tedy volně zaměnit za parametr Velocity, který vyjadřuje dynamiku not v obecném (ve smyslu na dechové nástroje nespecializovaném) MIDI. V těchto jednotkách jsou udávány hodnoty vertikální osy grafu. Abychom zachovali přehlednost grafů, vytvořili jsme pro každou nahrávku graf zvlášť, nicméně ve všech uvádíme průměrnou hodnotu, díky čemuž můžeme pozorovat, jak se jednotlivé křivky liší.

¹⁹ Zdrojový kód parseru je přílohou diplomové práce.

Time	T(stp)	AVAvgd	AV Vel	1AVAvgd	1AV	2AVAvgd	2AV	1:1Avgd	1:1Vel	1:2Avgd	1:2Vel
-2			44		58		44		54		70
-1			59		58		69		54		70
0			59		58		69		54		70
1			58		57		68		54		67
2			58		57		68		54		69
3			58		56		68		54		69
4			58		56		68		54		69
5	5	57	57	56	56	66	67	54	54	67	67
6			57		57		67		54		67
7			57		57		67		54		67

Tabulka 4.1 – Příklad části tabulky prvního typu

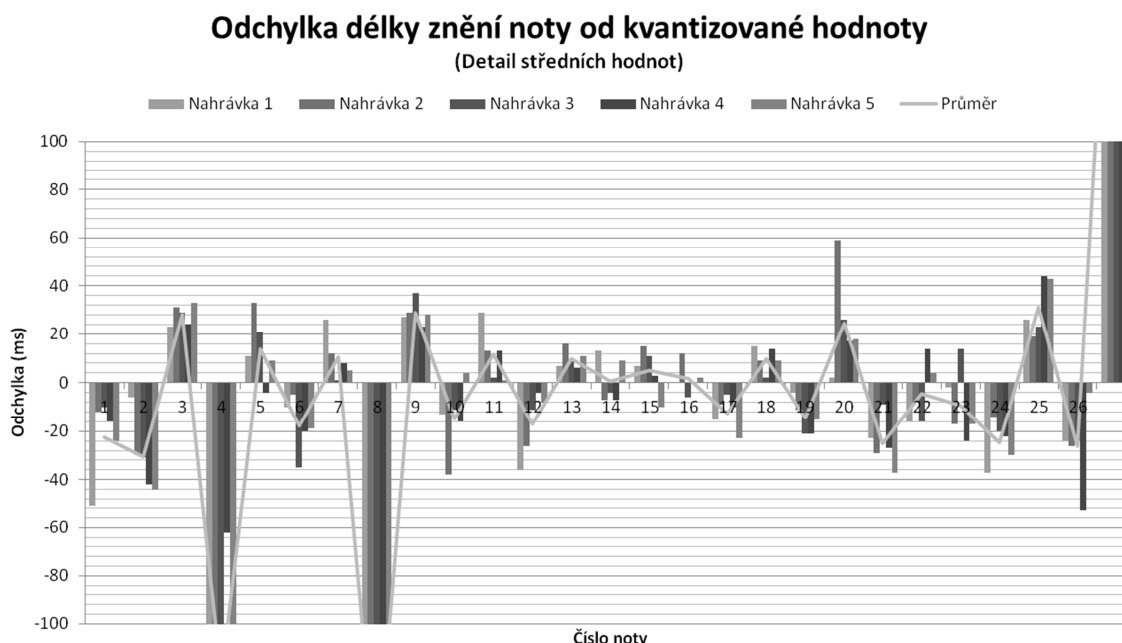
Tabulka druhého typu (viz Tabulka 4.1) je podstatně složitější vzhledem k zatížení na množství sledovaných parametrů. Je rozdělená do sedmi hlavních sekcí. V první sledujeme čas začátku a délku kvantizované noty, ve druhé části počítáme průměrné hodnoty odchylky začátku, dynamiky, délky a odchylky délky. V následujících pěti sekcích pak uvádíme pro každou z čtyř až pěti nahrávek téže skladby čas začátku znění noty, jeho odchylku od kvantizované hodnoty uvedené v části jedna, délku noty a opět její odchylku od kvantizované hodnoty a nakonec dynamiku a její odchylku od průměrné dynamiky vypočítané z dynamiky téže noty ve všech verzích nahrávky. Každý řádek tabulky odpovídá jedné notě, její pořadové číslo a výškové určení je uvedeno v první části. Jestliže některá nota nebyla v jedné z nahrávek zastoupena (interpret ji nezahrál nebo byly její parametry příliš extrémní, byla vyřazena. Stejně tak jsme vyřadili noty, které interpret hrál nad rozsah požadovaný partiturou (s výjimkou průběhu trylků).

QNote	QStrt	QLth	QName	AVStrt	AVVel	AVLth	AVLth±	1Strt	1Strt±	1Lth	1Lth±	1Vel	1Vel±
1	2 057	341	g#4	-13	50	318	-23	2 057	0	290	-51	41	-9
2	2 399	342	c#5	-22	50	311	-31	2 361	-38	336	-6	46	-4
3	2 742	342	d#5	-52	53	370	28	2 699	-43	365	23	51	-2
4	3 085	2 056	c5	-23	58	1 933	-123	3 066	-19	1 913	-143	60	2
5	5 142	342	b4	-75	53	356	14	5 054	-88	353	11	64	11
6	5 485	342	f#4	-60	62	324	-18	5 408	-77	332	-10	67	5
7	5 828	341	g#4	-73	70	351	10	5 742	-86	367	26	63	-7
8	6 171	2 056	c5	-52	51	1 890	-166	6 117	-54	1 786	-270	45	-6
9	8 228	341	g#4	-93	59	370	29	8 139	-89	368	27	51	-8
10	8 571	341	c#5	-54	63	326	-15	8 514	-57	328	-13	63	0

Tabulka 4.2 – Příklad části tabulky druhého typu

Z této tabulky jsme byli dále schopni sestavit a nechat vykreslit šestici grafů. První graf *Odchylka v přesnosti nástupu tónu* je sloupcového typu a ukazuje přesnost nástupu noty v čase, kdy je na vertikální ose vynesena odchylka začátku znění noty v nahrávce v milisekundách (všechny grafy mají na horizontální ose vyneseno číslo noty). Druhý (sloupcový) graf *Průběh délky znění noty* ukazuje absolutní délky not v milisekundách (vertikální osa), abychom si

udělali představu, jak dlouhé noty srovnáváme. S ním souvisí třetí (sloupcový) graf *Odchylka délky znění noty od kvantizované hodnoty*, který délku znění noty převádí z absolutní veličiny na relativní – na odchylku vzhledem k délce kvantizované noty (resp. délce dané partiturou). Sloupcový graf je proložen křivkou průměru těchto odchylek napříč všemi nahrávkami, aby byla patrná změna odchylky (zda roste, klesá, či se nemění). Protože odchylky v délce znění mají značný rozptyl, vytvořili jsme ještě čtvrtý doplňující graf s podtitulem *Detail středních hodnot* (viz Graf 4.1), který je identický s grafem třetím, ale jeho vertikální osa má pevně stanovené minimum a maximum tak, abychom mohli detailně prozkoumat střední hodnoty odchylek, které bychom z původního grafu kvůli rozlišení odečetli jen těžko. Pátý graf *Průběh parametru Velocity* je liniového typu a ukazuje hodnoty parametru Velocity nad jednotlivými notami vstupních dat a jejich změny. Doplňkově zahrnuje křivku ukazující průměr z vynášených hodnot. Šestým (opět sloupcovým) grafem *Odchylka parametru Velocity od průměrné hodnoty* si přibližujeme rozdíl parametru oproti průměrné hodnotě téže noty ve všech analyzovaných nahrávkách.



Graf 4.1 – Příklad kombinovaného grafu

Z tabulky druhého typu jsme dále vypočítali rozložení odchylek jednotlivých parametrů a uložili je do třetí, pomocné tabulky. V případě odchylek chceme na horizontální osu grafu později vynést škálu od -200 do 200 ms, v případě odchylek parametru Velocity pak uvažovat rozsah od -40 do 40 jednotek. V obou případech se rozsahy ukázaly být dostatečné, větší škála by navíc vedla k nepoužitelným výsledkům za hranici běžné hráčské chyby – nepůjde tedy o hledané interpretační nuance. Dále jsme provedli krokování škály, abychom získali přehlednější výstupní graf. V případě časových jednotek jsme počítali výskyty odchylek spadající do kroků o velikosti 10 ms, v případě parametru Velocity to byly dvě jednotky. Do takto zkonstruovaného rastru jsme uvedli spočtenou četnost výskytu jednotlivých odchylek v parametrech – vytvořili jsme histogram.

Všechny materiály odkazované v této části práce jsou k nalezení v příloze. Zde se budeme zabývat jen těmi, které zřetelně ilustrují zkoumané jevy. Analyzovány byly ale všechny vstupní podklady, i ty zde nezmíněné.

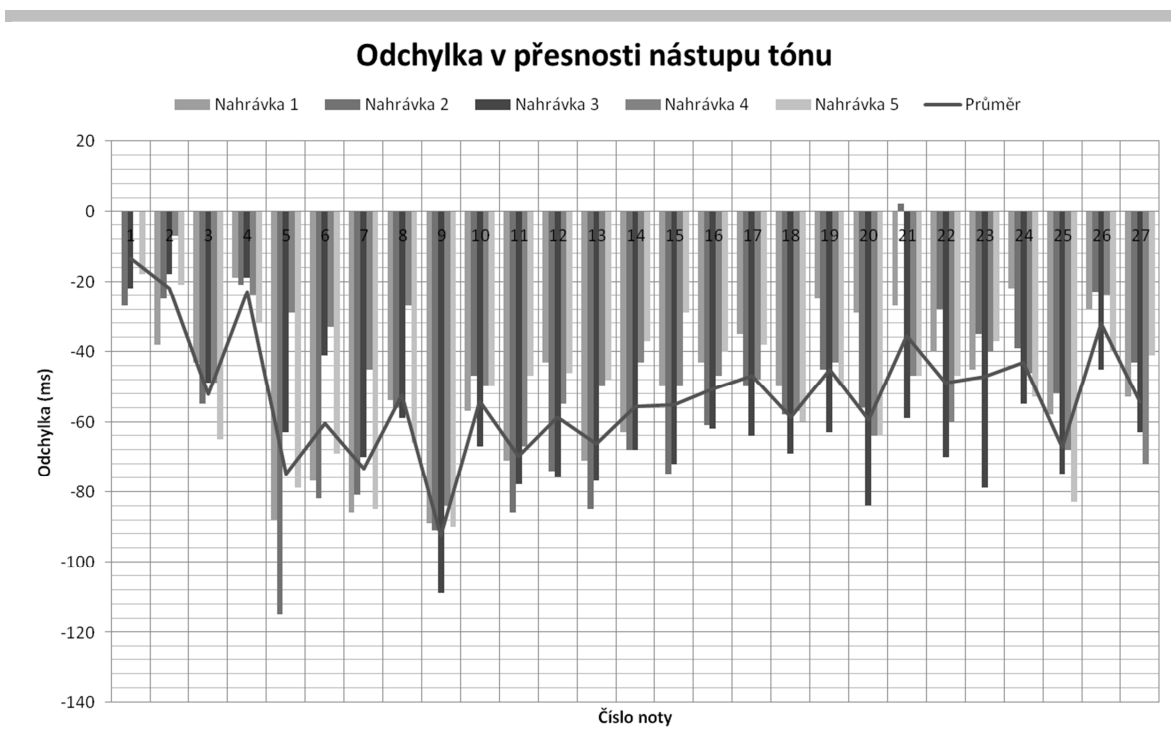
4.3 Pohled na data

Na základě grafických i numerických reprezentací zpracovaných dat se nyní pokusíme najít artefakty interpretačně charakteristických nuancí hry doktora Daňhela.

Připomeňme si zde, že každý vybraný part byl zaznamenán většinou v pěti různých nahrávkách, které jsou pro účely analýzy postaveny vedle sebe. Už ze zběžného pohledu na grafy vykreslující různé pohledy na tato vstupní data je zřejmé, že interpret se ve všech pěti nahrávkách choval na stejných místech velmi podobným způsobem. V rámci naší motivace k vylepšení dosavadních metod humanizace je to ideální demonstrace nutnosti použít více než pseudonáhody k vytvoření člověku podobné „pozitivní“ chybovosti.

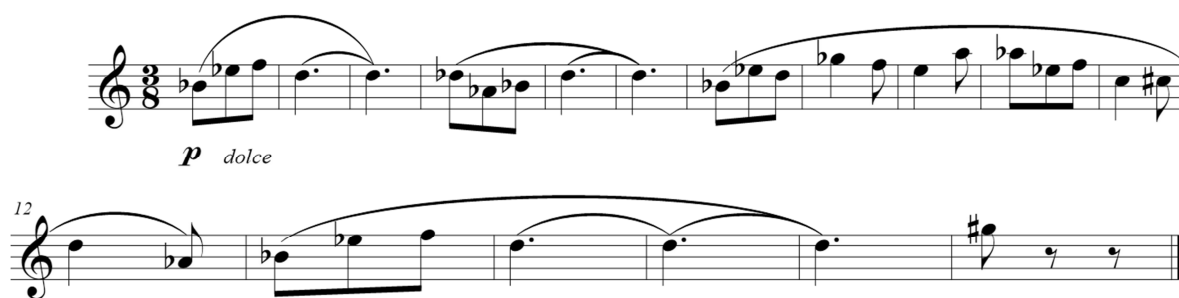
4.4 Analýza odchylky začátku znění noty

Základním východiskem zde pro nás bude graf *Odchylka v přesnosti nástupu tónu*, nezbytně uvedený v souvislost s partiturou skladby. Budeme totiž hledat charakteristickou chybovost v závislosti na hudebním kontextu, ve kterém zkoumaný jev identifikujeme.



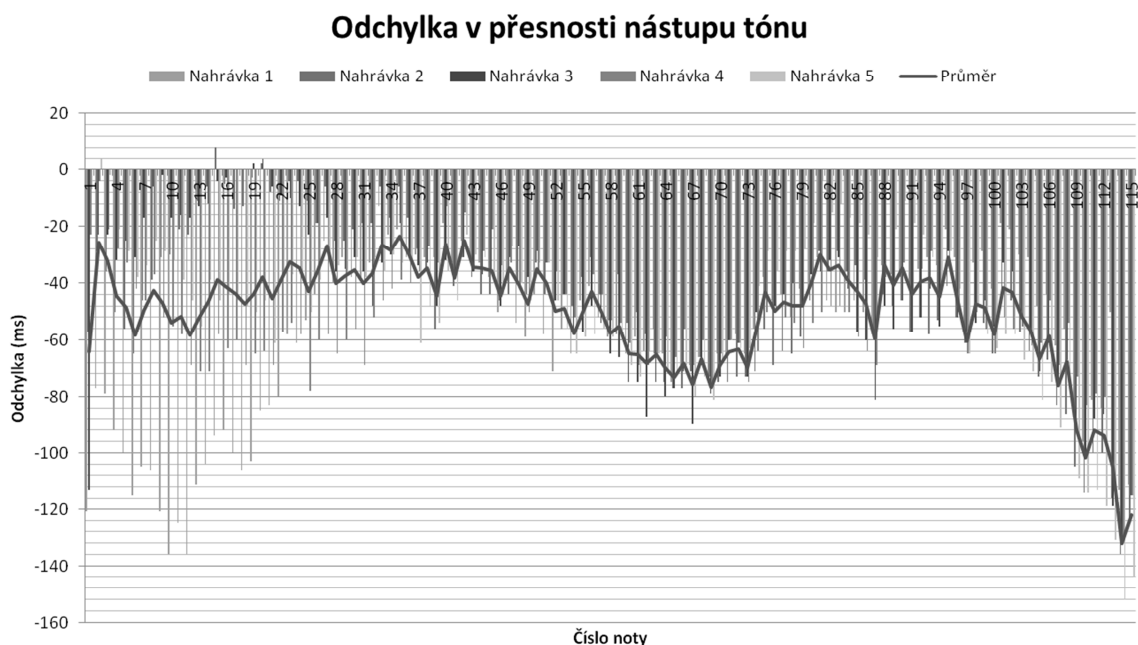
Graf 4.2 – Part skladby Concerto for Orchestra

Vezmeme-li si pro začátek Graf 4.2, pak z vychýlení sloupců do záporné části vertikální osy vidíme, že hráč v nasazení tónu přesné doby spíše předbíhá. Přesto vidíme jasné snížení v některých místech křivky, nevelké absolutně, ale znatelné relativně k aktuální úrovni předcházení se přesnému času (musíme si stále uvědomovat, že sledujeme souvislou hru proudu not, takže je možné, že hráč se nám může v čase krátkodobě vychýlit vpřed nebo vzad a během celého tohoto úseku bude odchylka kontinuálně ovlivněna tímto vychýlením, než se interpret zase „setká“ s přesně udávanými dobami).



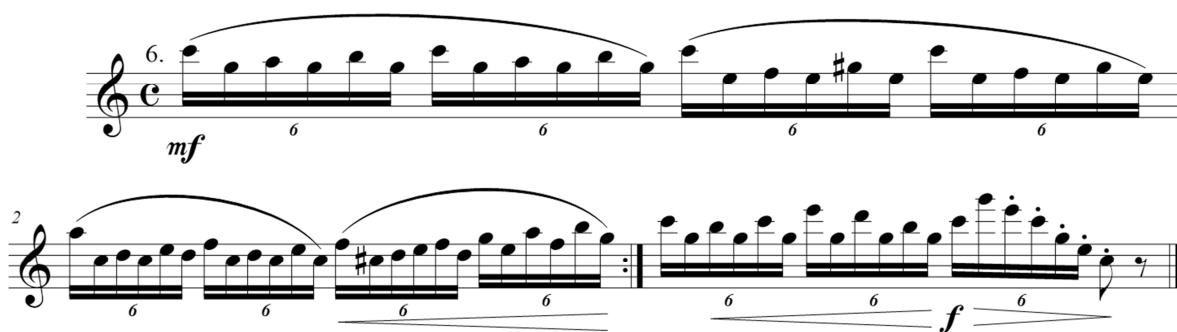
Partitura 4.1 – Béla Bartók: Concerto for Orchestra

Podíváme-li se nyní do partitury (viz Partitura 4.1; noty jsou číslovány sekvenčně zleva doprava, ligatury jsou samozřejmě považovány za jeden tón), zjistíme, že snížené odchylky téměř pravidelně odpovídají notám s počátky na celé době. Větší odchylky jsou naopak patrné v případě rychlých sledů krátkých not (zde osminové vzestupy). Interpret také pravidelně zvyšuje svou nepřesnost poté, co zakončuje některý dlouhý tón – toto je dobře vidět například na notách č. 26 a 27.



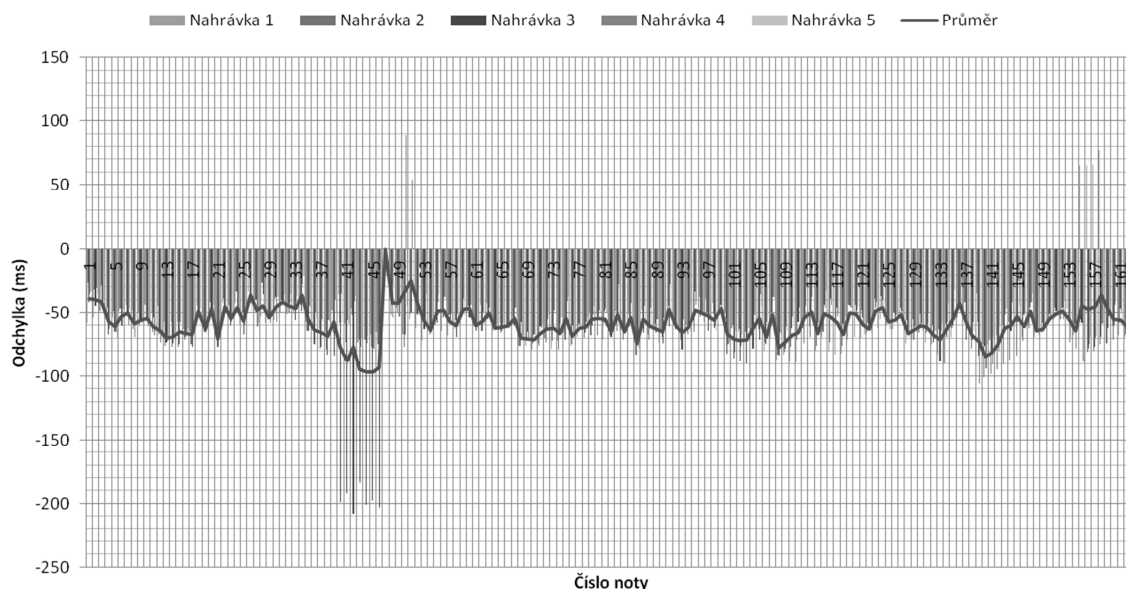
Graf 4.3 – Etuda 6 ze 167 Exercises for Daily Use

V Graf 4.3 vidíme sled 115 not stejné délky (kromě poslední noty cvičení). Etuda je ideální pro zkoumání rytmických odchylek, protože je přímo určena pro technické cvičení časové přesnosti. Skládá se ze šestnáctinových sextol v čtyřčtvrtečním taktu, první nota sextoly je navíc vždy mírně akcentována. Z hlediska výše popsaného jevu nás tedy budou zajímat noty č. 1, 7, 13, ... V tomto případě ovšem pozorujeme jev opačný – interpret svou nepřesnost v případech not scházejících se s celými dobami taktu (a to je u šestnáctinových sextol v 4/4 taktu právě a pouze každá šestá nota) naopak jemně zvyšuje. To si ale můžeme vysvětlit také složitějším přehmatem, kterého bylo při přechodu zapotřebí. Na tomto grafu je dobře vidět jev popsáný v předchozím odstavci – a sice že interpret může dočasně vybočit z uloženého přesného tempa (odchylka je průběžně vyšší) a zase se k němu postupně vrátit.



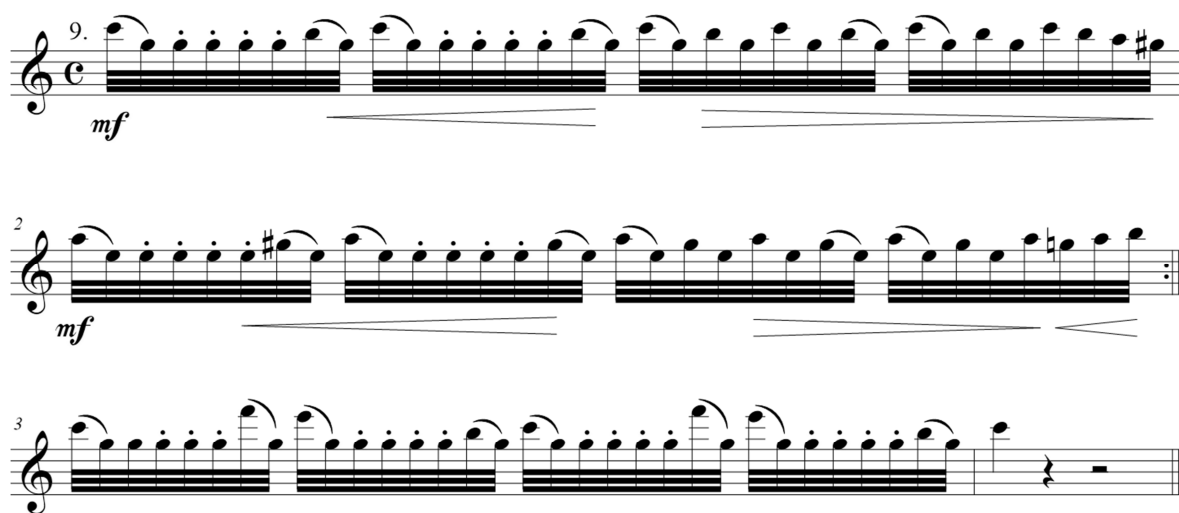
Partitura 4.2 – Fritz Kroepsch: Etuda 6 ze 167 Exercises for Daily Use

Odchylka v přesnosti nástupu tónu



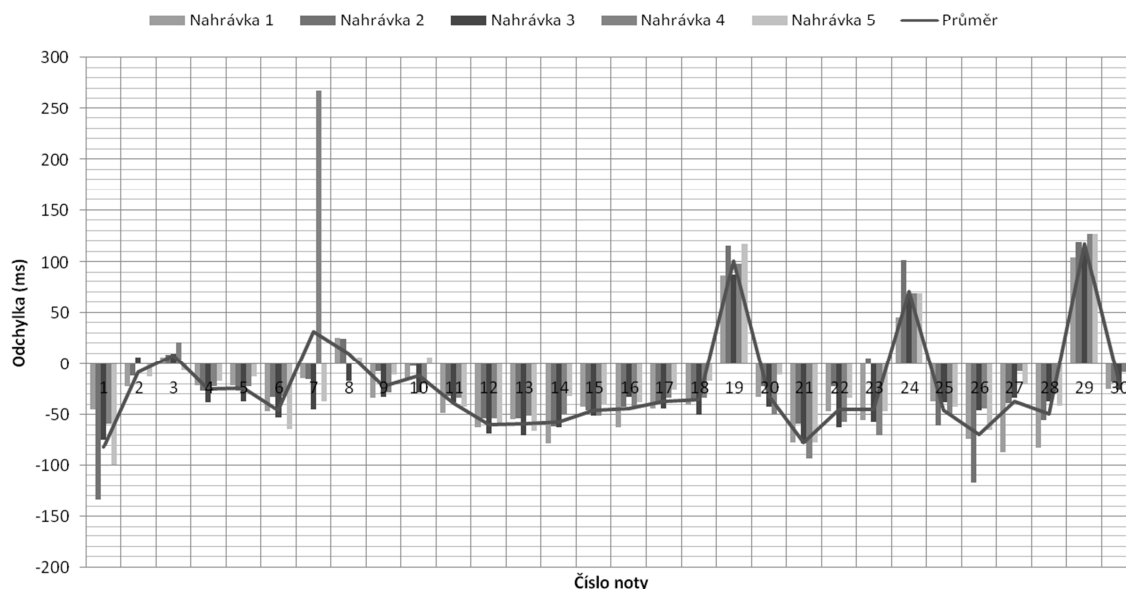
Graf 4.4 – Etuda 9 ze 167 Exercises for Daily Use

Ve cvičení 9 z téže sady jako předchozí ukázka (Graf 4.4) naopak zase pozorujeme úkaz podporující tezi o zvyšující se přesnosti v celých dobách. Výkyvy jsou sice minimální, ale zato poměrně přesně kopírují místa, kde se dvaatřicetinové pravidelné sledy setkávají s čtyřčtvtečným taktem. Protože cvičení je z velké části tvořeno staccaty na jednom tónu, interpret se může více soustředit na nasazení tónů akcentovaných, kterými v tomto případě budou terciové a kvartové vzestupy.



Partitura 4.3 – Fritz Kroepsch: Etuda 9 ze 167 Exercises for Daily Use

Odchylka v přesnosti nástupu tónu



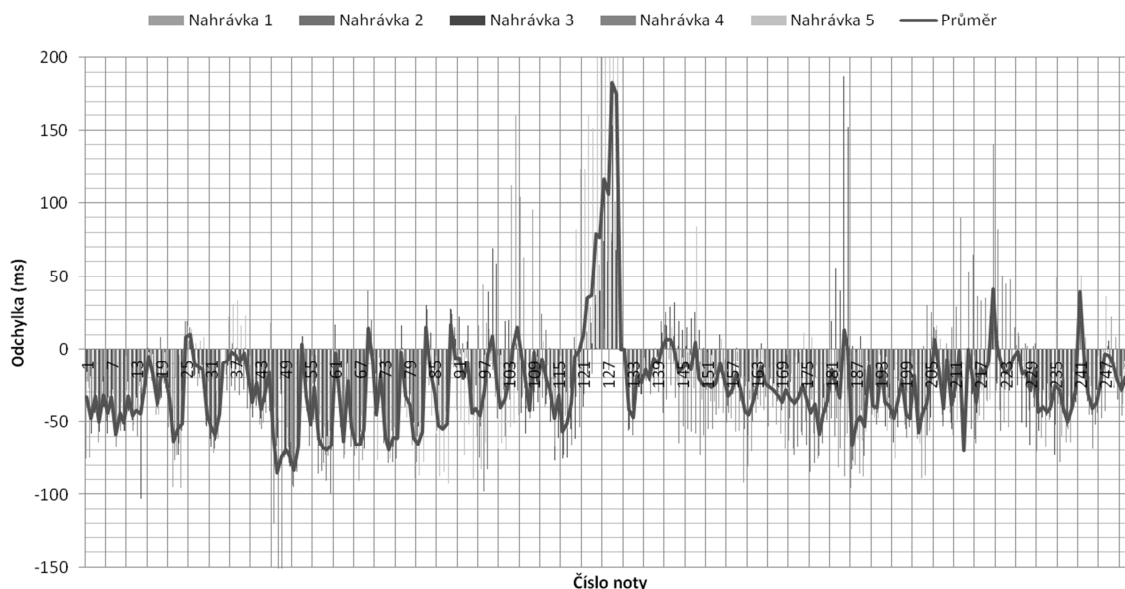
Graf 4.5 – Část skladby Šárka z cyklu Má vlast

Zajímavým potvrzením tohoto jevu je i expresivní sólový part ze Smetanovy Šárky (Graf 4.5). Vzhledem k charakteru partu je jasné, že zde sledujeme interpretační odchylky v užším slova smyslu, tedy záměrné nedostatky v přesnosti. Skladba rovněž obsahuje řadu tónů, které přecházejí mezi takty. Povšimneme si ale např. kromě noty č. 12 také not č. 20, 25 a 30. Jsou to noty, které následují za poměrně hodně zpožděnými krátkými šestnáctinami, interpret se ale přesto rychle vrací do rytmu, a to dokonce s poměrně malou absolutní odchylkou. Tyto noty také začínají na prvních dobách svých taktů.



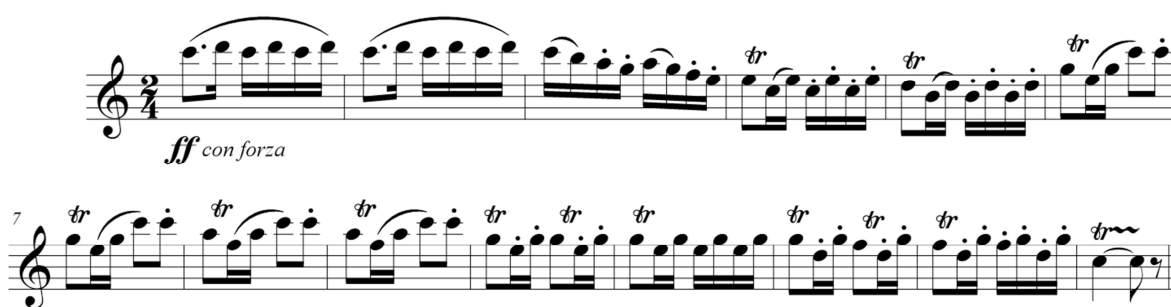
Partitura 4.4 – Bedřich Smetana: Část skladby Šárka z cyklu Má vlast

Odchylka v přesnosti nástupu tónu



Graf 4.6 – Části skladby Alborado z Capriccio espagnol

A konečně nejlépe lze demonstrováný jev vidět na dvou vybraných částech skladby Alborado z díla Capriccio espagnol. I přesto, že se jedná o technicky poměrně dost náročné party (časté trylky následované krátkými notami, staccatové proklady v rychlém tempu a pravidelném šestnáctinovém běhu, občasné proložení triolou, svazování not těsně před přechodem do staccata), interpret se pravidelně vrací až do téměř nulové časové odchylky. O dynamičnosti partu svědčí velké skoky mezi místy, kde jsou odchylky opravdu malé, a těmi, kde se hráč dostává až kolem 50 ms před rastr. Pro vysvětlení dodejme, že silné zpomalení v (horizontálně) střední části grafu pak mapuje místo, kde se jeden a půl doby trvajícím trylkem ukončuje první vybraný part a kde si interpret pochopitelně dovoluje značně individualistické zacházení s tóny.



Partitura 4.5 – Nikolai Rimsky-Korsakov: První části skladby Alborado z Capriccio espagnol

Na uvedených příkladech jsme si ukázali, že pozice noty v taktu znatelně ovlivňuje míru, s jakou bychom měli notu humanizovat. Ve spojení s učebnicovým nárokem na zvýšenou přesnost počátečních a dále celých dob v taktu (17) vidíme, že jestliže u celých dob budeme

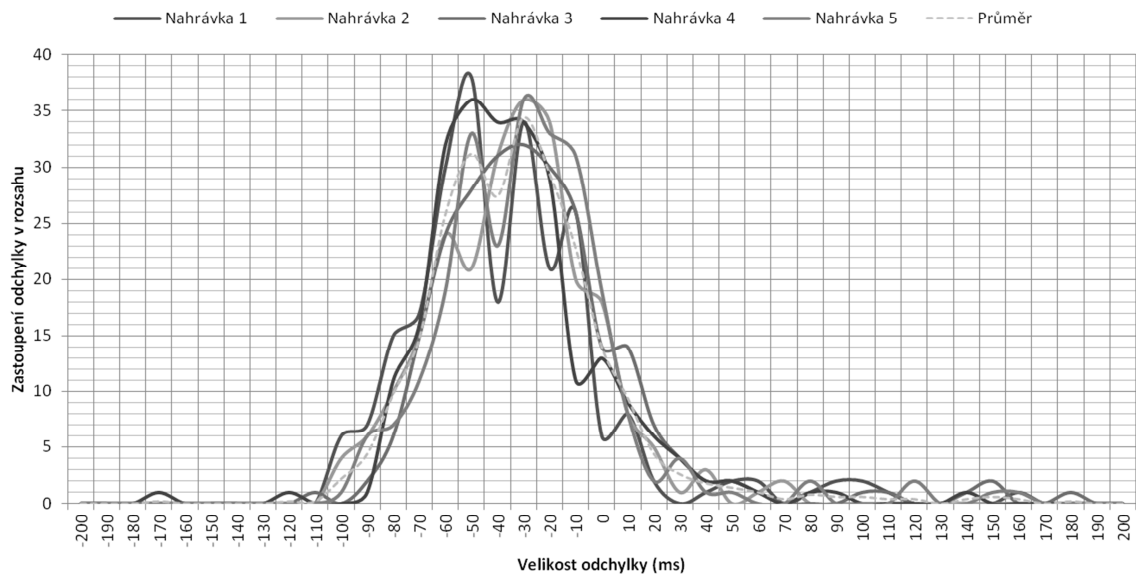
volit odchylky nižší, budeme si moci dovolit i poměrně necharakteristickou humanizaci na notách ostatních.

4.4.1 Distribuce odchylek pozice začátku noty v čase

Jestliže vezmeme v potaz informace o tempu skladeb a prozkoumáme Graf 4.2 až Graf 4.6, vidíme, že ač se tempa skladeb značně různí, maxima odchylek nasazení tónů se pohybují kolem 50 ms, minima pak nabývají hodnot přibližně polovičních. Jen ve výjimečném případě hraje interpret s pravidelným (tzn. výskyt hodnot má nějaký vzor) nulových vychýlením oproti očekávané přesnosti dané kvantizovanými hodnotami délek not. Tato skutečnost je zajímavá zejména proto, že z ní můžeme usuzovat nezávislost průměrné hladiny odchylky na tempu skladby. To pouze zvyšuje obtížnost hraného materiálu a ve spojení s dalšími prvky technické náročnosti jen vytváří větší rozpětí mezi notami hranými s důrazem na přesnost a těmi ostatními, což později v humanizačních metodách zohledníme.

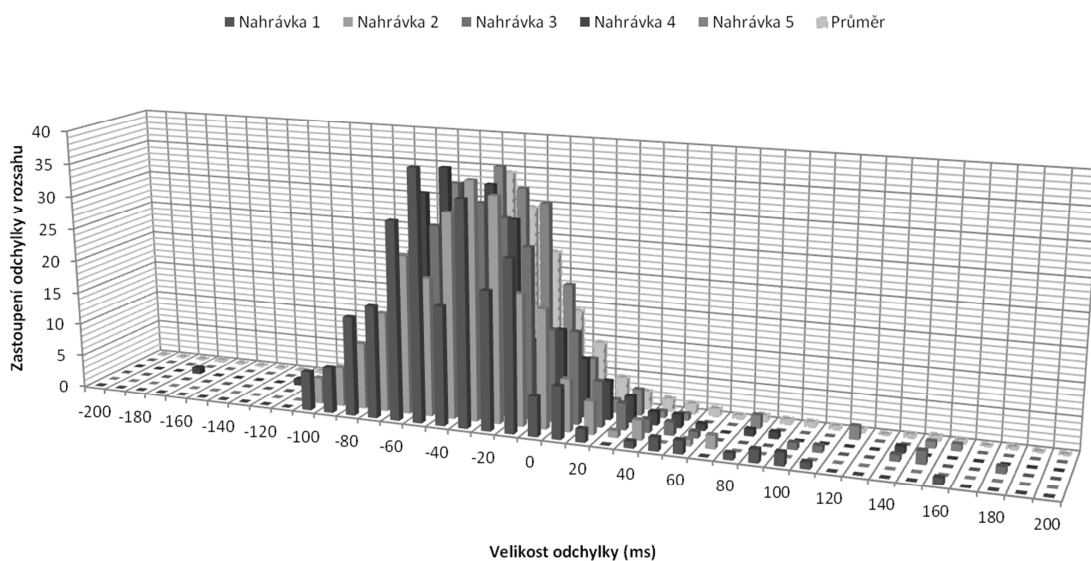
Důvodem této konzistentní odchylky bude pravděpodobně také snaha kompenzovat dobu, kdy nástroj není co do kvality tónu v optimálním provozu. Nástroj potřebuje nějaký čas, než se sloupec vzduchu, jehož vibrace vytváří samotný tón, po začátku vdechování ustálí. Interpret navíc není schopen vpravit skokově do nástroje potřebné množství vzduchu – chvíli trvá, než se z nulového proudění dostane na požadovanou hodnotu. Kompenzací této latence se hráč přiblíží situaci, v níž se tón v okamžiku, kdy dosáhne optimální kvality, kryje s okamžikem svého skutečného (resp. skladbou požadovaného) začátku. Toto vysvětlení podporuje také graf *Rozložení odchylek délky tónů* využívaný dále v části 4.5 – střední hodnota křivky představující histogram přesnosti ukončení tónů se pohybuje kolem 0 ms, nikoliv kolem -50 ms jako v případě nasazení tónu. Zvuk totiž po přerušení buzení nástroje vdechovaným vzduchem zaniká velmi rychle, takže žádná kompenzace v tomto případě není potřebná. (21) Předsunutí začátku tónu o 50 ms je tedy činnost vědomá nebo alespoň záměrná a naučená v případě, že ji interpret aplikuje podvědomě.

Rozložení odchylek přesnosti nástupu tónu



Graf 4.7 – Části skladby Alborado (Capriccio espagnol)

Rozložení odchylek přesnosti nástupu tónu



Graf 4.8 – Části skladby Alborado (Capriccio espagnol) v prostorovém zobrazení

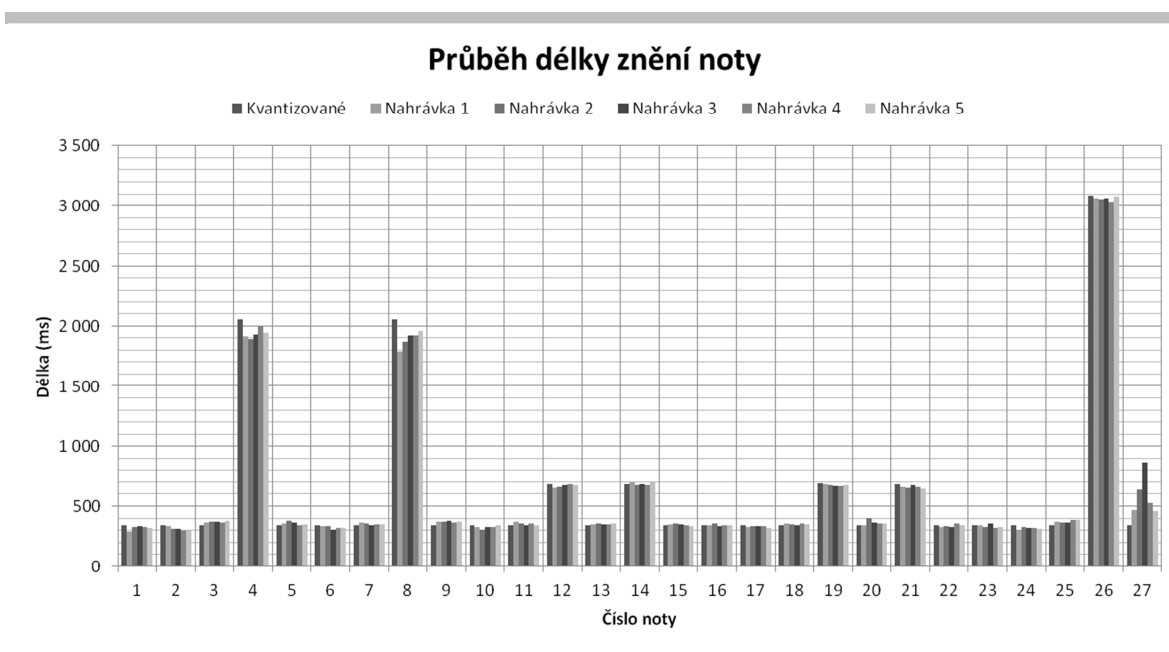
Z pohledu na Graf 4.7 (resp. Graf 4.8) můžeme usoudit, že rozložení (ve statistice označované jako hustota četnosti) odchylek vytváří tvar, ve kterém můžeme s trochou tolerance rozpoznat známou *Gaussovu křivku*. Není to jen proto, že Gaussova křivka, graf *normálního rozložení* (hustoty) pravděpodobnosti je z *Centrální limitní věty* rozložením limitním, (22) ale zejména proto, že normální rozdělení je charakteristické pro výskyt znaků a jejich odchylek v přírodě. (23) To nevyjímá samozřejmě ani lidi, a to na mnoha úrovních – genetické, znalost-

ní, dovednostní atp. Skutečnost, že histogram chybovosti doktora Daňhela odpovídá Gaussově křivce, jednak validuje korektnost našeho měření ve smyslu použitých metod, a jednak u humanizace vynucuje využití jiné distribuce náhodné veličiny využívající generátor pseudonáhody, než je pouhé její lineární rozdělení.

4.5 Analýza odchylky délky znění noty

V tomto případě sledujeme nepřesnost ukončení znění noty vzhledem k délce, která je jí předepsaná partiturou. K základní orientaci v délkách not nám slouží graf *Průběh délky znění noty*, díky němuž můžeme snadno seskupit noty dle jejich délky, neboť lze očekávat, že ta bude mít vedle notového okolí toho kterého tónu značný vliv na přesnost ukončení. Tím myslíme, že rytmizace u kratších not bude přesnější, zatímco dlouhé noty bude mít interpret tendenci subjektivně podhodnotit nebo nadhodnotit – bude mít pocit, že např. nota zní už „příliš dlouho“, a proto ji předčasně ukončí a naopak.

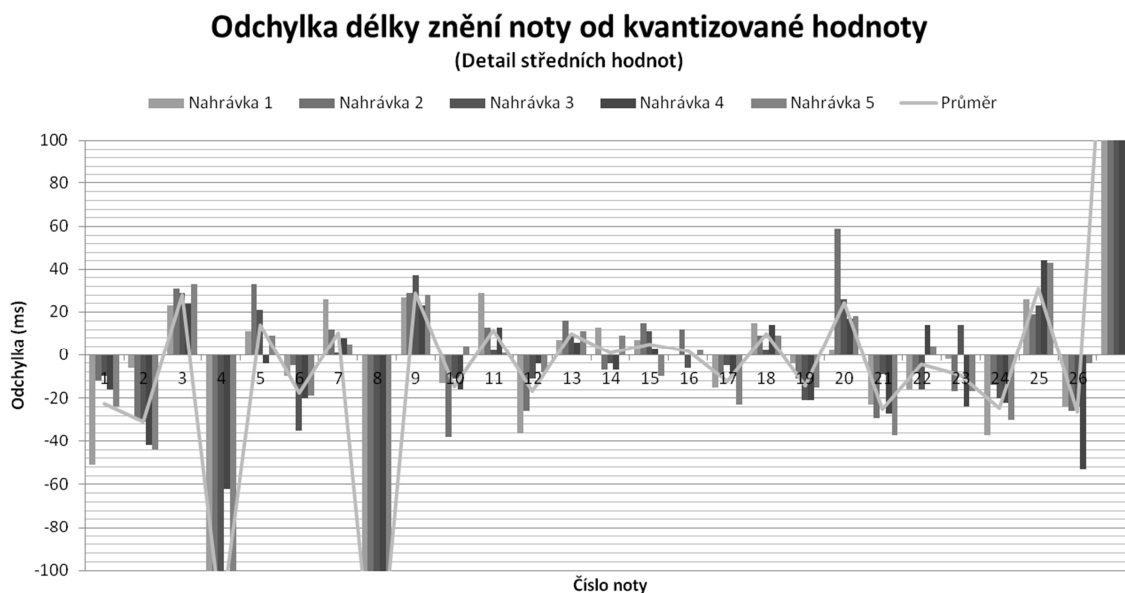
Skutečnou velikost odchylky od kvantizované hodnoty prozkoumáme v grafu *Odchylka délky znění noty od kvantizované hodnoty* či v jeho na střední hodnoty zaměřeném doplňujícím grafu.



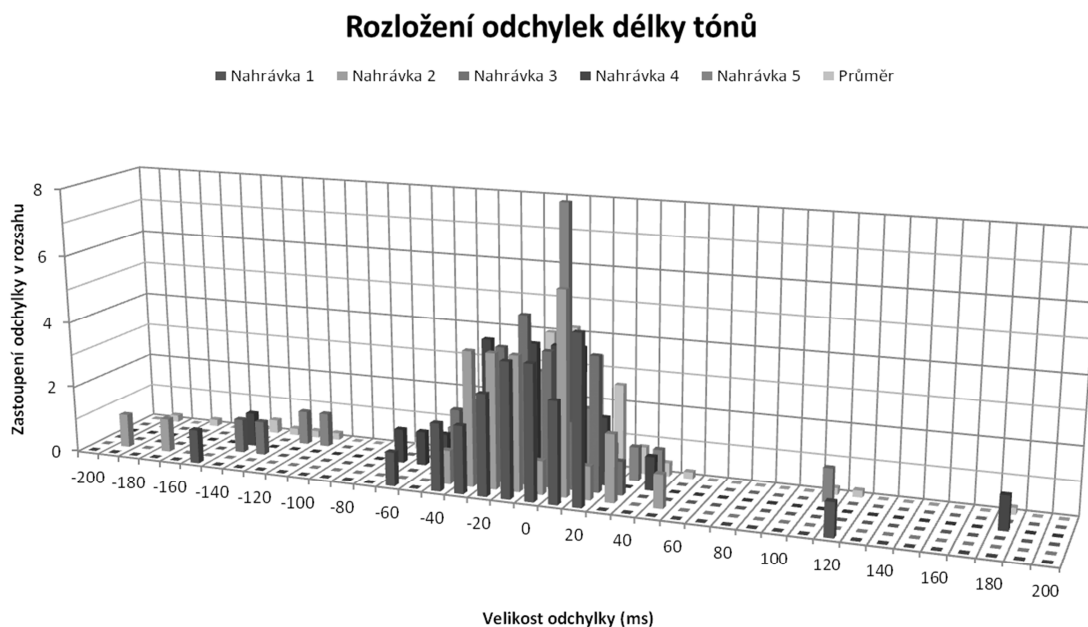
Graf 4.9 – Část skladby Concerto for Orchestra

Spojíme-li si Graf 4.9, Graf 4.10 a Partitura 4.1, pozorujeme, že průměr absolutní hodnoty průměrné nepřesnosti ukončení osminových not je zřejmě nižší, než je tomu u not čtvrtových či delších. Největších odchylek dosahuje interpret v případě not č. 4, 8, 18 a 21, které představují tóny znějící tři čtvrtové doby v případě prvních dvou not a jednu dobu v případě druhých dvou. Ač je rozptýl rozložení odchylek přesnosti délky noty přibližně stejný jako u odchylek přesnosti pozice jejich začátku, vidíme, že střední hodnota je posunuta až k nule (tedy k absolutní přesnosti). Ten samý jev pozorujeme také u partů z Poloveckých tanců (zde

odkazujeme z prostorových důvodů pouze elektronickou přílohu, vybraná část je značně dlouhá a v tištěné podobě bychom z grafu mnoho nevyčetli).



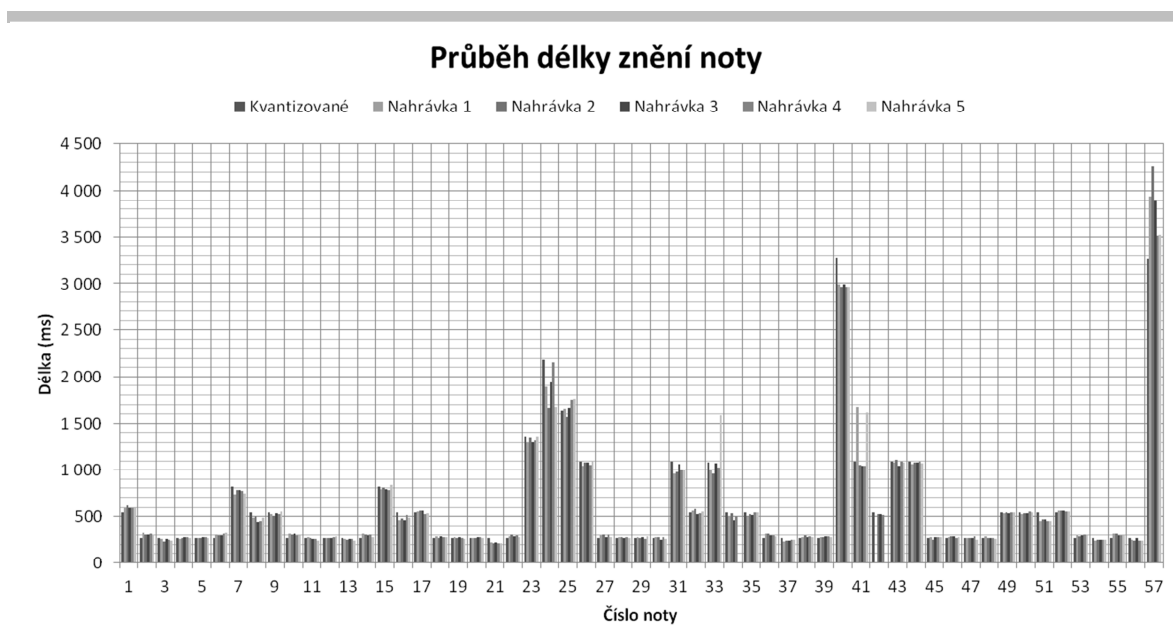
Graf 4.10 – Část skladby Concerto for Orchestra



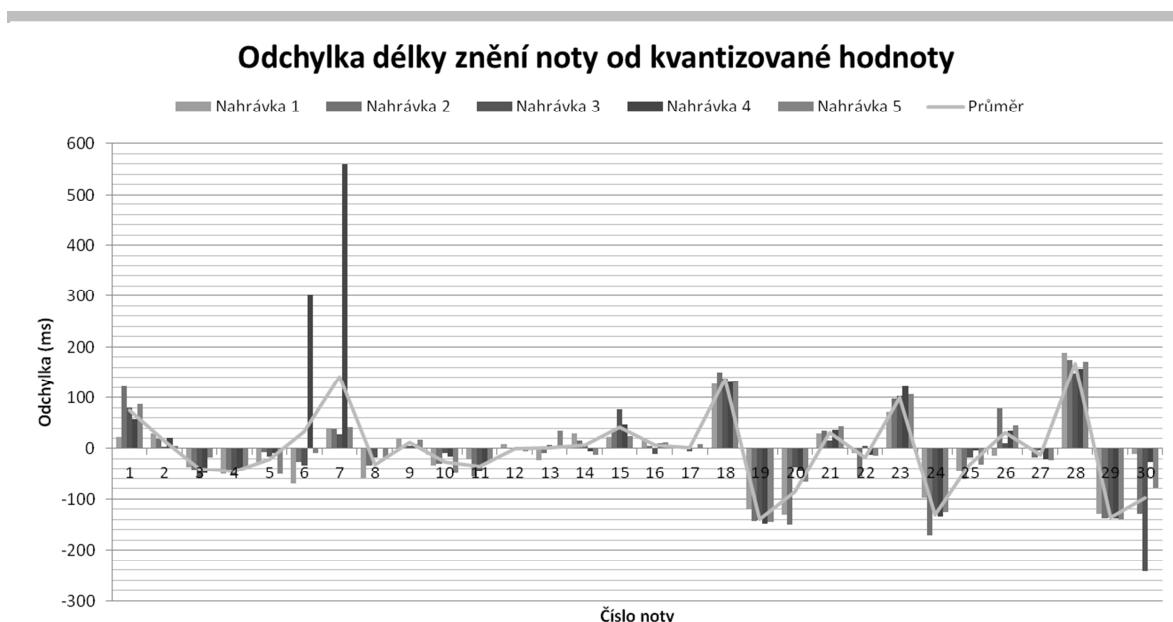
Graf 4.11 – Část skladby Concerto for Orchestra

Ve vyňaté části Smetanovy Šárky má interpret tendenci podhodnotit délku osamocených šestnáctinových not (noty č. 19, 24 a 29). Právě tyto noty mají ale zpožděný začátek vzhledem k partiturou předepsanému rastru, takže se interpret zřejmě jen snaží „dohnat“ čas, který při

nástupu noty ztratil. Zaměříme-li se ale na okolí těchto not, vidíme také (viz Graf 4.5), že předcházející i následující osminová nota je v začátku přesná (relativně k přesnosti ostatních not v okolí) a její délka opět kolísá k nadbytečnému prodloužení. Zde vysvětlení zatím přenecháme na jevu popsanému v předchozím odstavci.



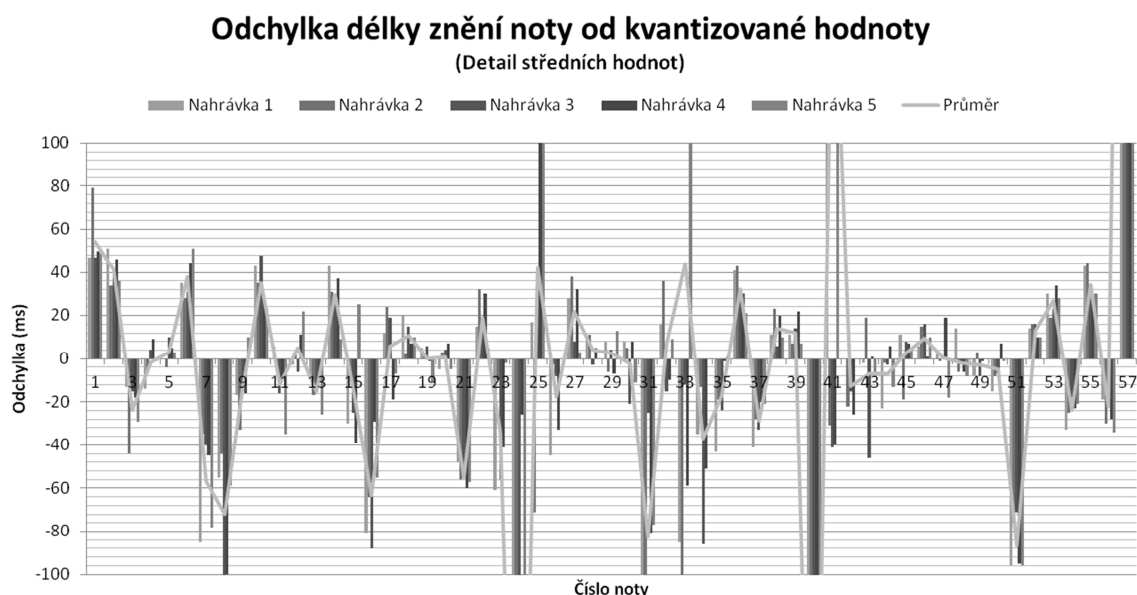
Graf 4.12 – Část skladby Šárka z cyklu Má vlast



Graf 4.13 – Část skladby Šárka z cyklu Má vlast

Zajímavá je i skupina not č. 12–17, které sice začínají všechny o 50–70 ms dříve a počet dob, po které znějí, je nestejný, ale jejich délka je naopak velmi přesná ve všech pěti nahrávkách.

Akcent značený v partituře na notě č. 15 se projevil prodloužením délky jejího znění průměrně o 41 ms.



Graf 4.14 – Části skladby Glagolská mše

V grafu odchylek délky znění noty Glagolské mše (viz Graf 4.14) vidíme dva typy událostí. První ukazuje, že u dlouhých not (zde máme na mysli noty čtvrt'ové a delší) má hráč stále tendenci přetěžovat délku. Naopak u kratších, více rytmizujících osminových struktur, je patrná kompenzace předčasného nástupu tónů.

p dolce

mf *espress.*

p *accel.* *cresc.*

Partitura 4.6 – Vybrané části skladby Glagolská mše

Ať už prodlužování či zkracování délky znění not je znakem interpretace v širším slova smyslu či ne, je jasné, že délku znění noty můžeme spojit s odchylkami pozice jejího začátku. Dalším kritériem bude relativní délka noty, kdy krátké a dlouhé noty mají odlišnou charakteristiku ovlivněnou již zmíněným „přeceňováním“ nebo naopak „podceňováním“.

4.5.1 Distribuce odchylek délky znění noty

U všech zkoumaných partů byla střední hodnota odchylky délky not skutečně v nule nebo jen mírně (10 ms) pod ní. Rozptyl veličiny pak zpravidla odpovídal rozptylu distribuce odchylky nasazení tónu. Výjimku tvoří etuda 6 z Exercises for Daily Use a Polovecké tance, kde je rozptyl menší. To je patrně způsobeno zaměřením cvičení i části skladby na rytmiku a také skutečností, že oba party obsahují noty v minimálním psaném délkovém rozpětí – u partu z Poloveckých tanců jsou to převažující osminové a šestnáctinové noty, které jsou navíc pravidelně seskupovány do celků tvořených pouze těmito notami, u druhého partu jsou to pak samé šestnáctinové sextoly. Hráč si tedy vytvoří návyk na určitou délku noty a tu snáze dodržuje s menším vlivem toho, s jakou odchylkou notu začíná.

4.6 Dynamické odchylky

Dynamické odchylky jsou z hlediska humanizace asi nejzajímavější částí. Pouhé rozmělnění dynamiky může být zejména v komplexních aranžích dostatečné pro rozbití strojového charakteru hraných not a pomůže také v případě, že si nemůžeme dovolit měnit pozici začátku noty (například chceme spojit zvuk dvou bicích souprav – z nepřesností by se vytvářel nežádoucí efekt podobný echu²⁰).

Zkoumání nepřesností v interpretaci dynamiky not jsme zkoumali dvojím způsobem – jako pouhou jednu hodnotu přiřazenou každé notě v partu (váženým průměrem ze všech změn parametru CC-2 Breath Control během jejího znění, viz oddíl 4.1.3 na straně 24) a u dlouhých not rovněž jako detailní průběh dynamiky po celou dobu znění tónu.

K dispozici jsou sice zdroje, které předepisují určité hodnoty parametru Velocity slovním vyjádřením dynamiky v notovém zápise (piano, forte atd.), ale tento převod není nijak standardizován a i různý notační software exportuje do MIDI dynamické hladiny různě. V součinnosti s v notách značenými crescendo a decrescendo²¹ by bylo velmi obtížné a nereprezentativní vnímat odchylku dynamiky not ve srovnání s takto převedenou partiturou. Z tohoto důvodu jsme jako referenční hladinu zvolili průměrnou dynamiku ze všech čtyř až pěti nahrávek téhož partu. Z tohoto důvodu nás zde bude zajímat absolutní hodnota odchylek od průměru, tzn. budeme hledat místa, kde interpret od průměru hodně kolísá a kde naopak málo.

Východisky pro tuto část analýzy jsou grafy *Průběh parametru Velocity* a *Odchylka parametru Velocity od průměrné hodnoty*. Distribuci odchylky ukazuje graf *Rozložení odchylek parametru*

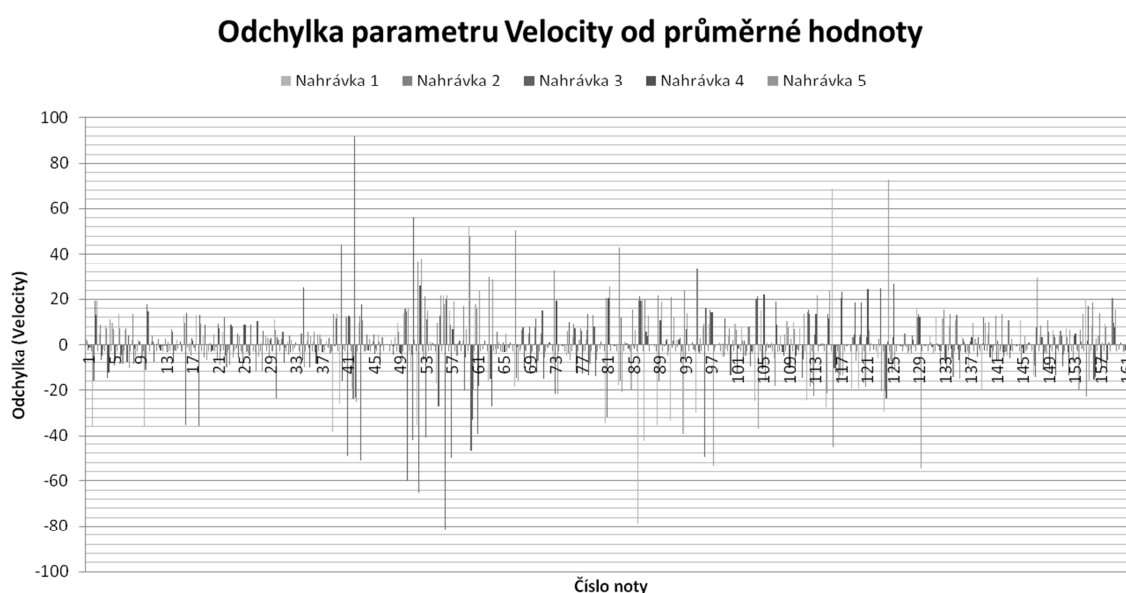
²⁰ Echo – Efekt podobný ozvěně: Za základním tónem se ozve jeho kopie, typicky v jiné dynamické hladině a s nějakým zpožděním. Echo je tedy efekt z rodiny Delay (zpožďovací efekt). (9)

²¹ Crescendo – s rostoucí silou, silněji. Decrescendo – s ubývajícím silou, stále slaběji. (33)

Velocity od průměrné hodnoty. Už po letném pohledu je i tento graf možné identifikovat s grafem normálního rozdělení, což opět ukazuje korektnost přístupu ve zkoumání odchylek.

Nakonec poznamenejme, že dechový kontrolér EWI disponuje pouze standardními 128 kroky parametru CC-2 Breath Control. Interpret může dechem poměrně snadno dosáhnout limitní hodnoty 127, která se už dále nemůže zvětšovat bez ohledu na to, jestli je hráč schopen vdechnout do nástroje s ještě větší intenzitou. Některé party jsou tak zatížené na noty znějící v maximální dynamice a v důsledku toho vykazují minimální odchylku. K této skutečnosti musíme v analýze přihlídnout, dlouhodobé setrvávání parametru CC-2 na maximu snadno odhalíme z grafu *Průběh parametru Velocity*.

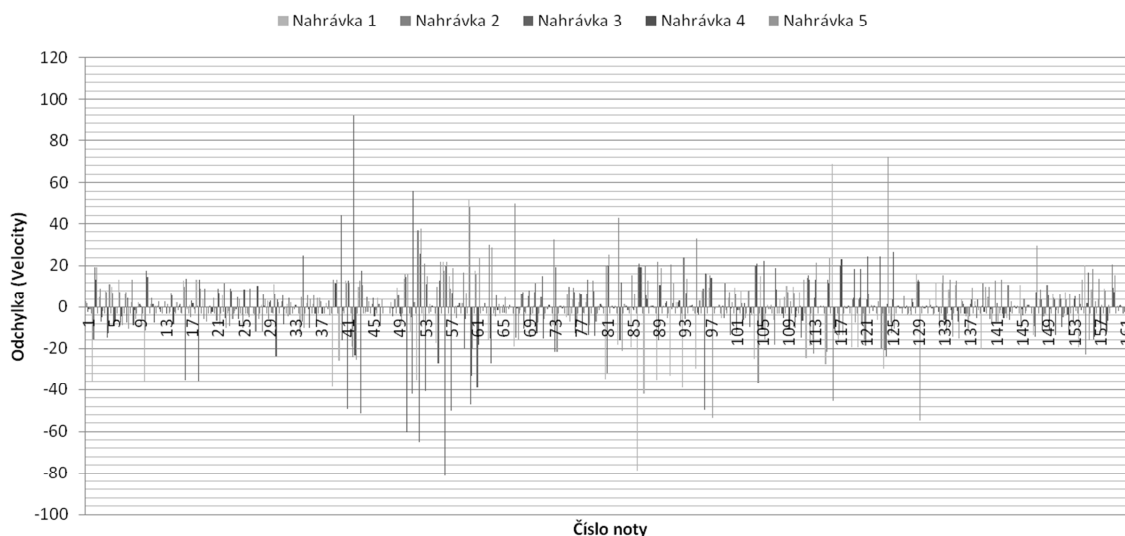
4.6.1 Odchylky průměrné dynamiky not



Graf 4.15 – Etuda 6 z Exercises for Daily Use

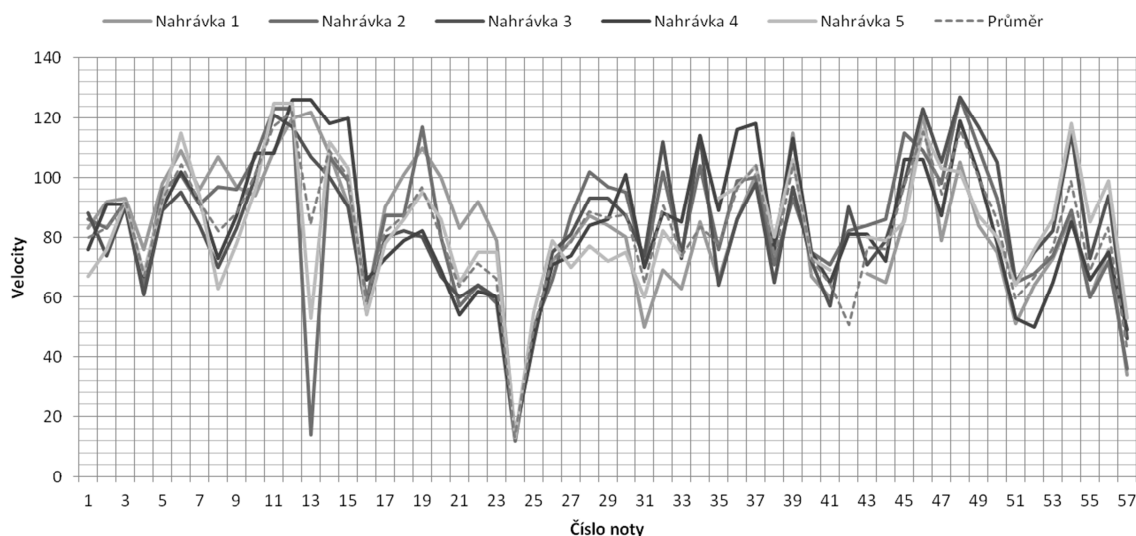
Na grafu (Graf 4.15) odchylky parametru Velocity etudy 6 vidíme značné absolutní rozpětí odchylek od průměru (0 na vertikální ose). Vzhledem k tomu, že se jedná o cvičení sestávající výhradně ze šestnáctinových triol, ukazuje nám to, že i při stejné délce noty interpret dynamicky kolísá, ač to po něm není partiturou vyžadováno. Totéž vyplývá z grafů etudy 9 (Graf 4.16) a Poloveckých tanců (vzhledem k rozsahu opět viz elektronická příloha). V obou cvičeních (etudy 6 a 9) není navíc hráč nucen zásadním způsobem měnit tóninu. Etuda 9 je nicméně obohacena o hru staccato a větší rozvlnění dynamiky je tak způsobeno spíše užitím této techniky.

Odchylka parametru Velocity od průměrné hodnoty



Graf 4.16 – Etuda 9 z Exercises for Daily Use

Průběh parametru Velocity (z váženého průměru změny CC-2)



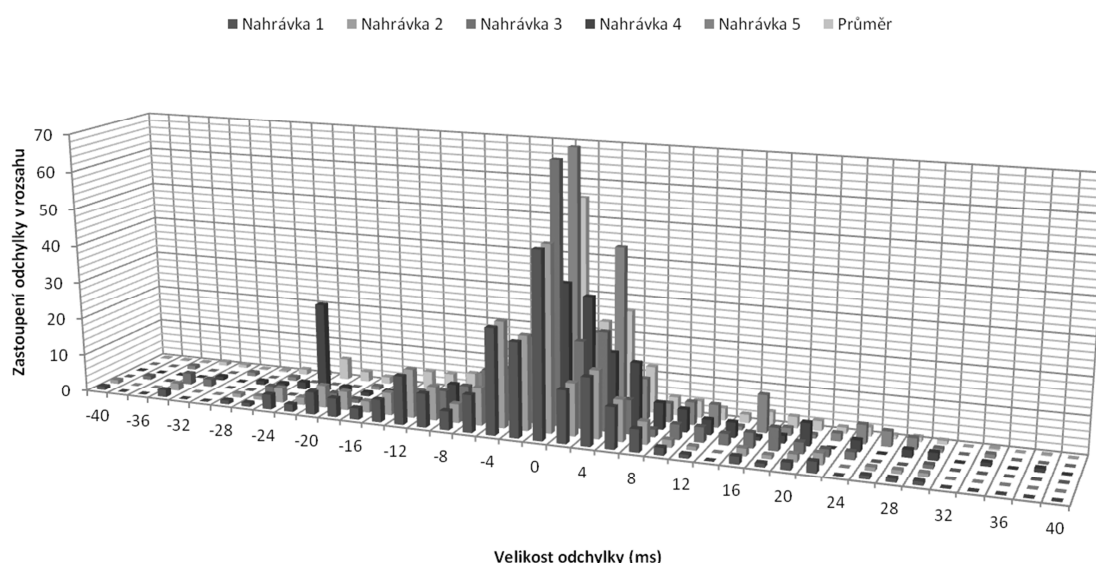
Graf 4.17 – Část skladby Glagolská mše

Zajímavější poznatky můžeme najít v partu z Janáčkovy Glagolské mše. V grafu *Průběh parametru Velocity* této skladby (Graf 4.1) je patrné větší setkávání dynamických křivek (což znamená minimální rozdíl) jednotlivých nahrávek v notách delších, osamocených, kdy má hráč čas vyrovnat dynamický průběh noty během jejího znění.

Ostatní skladby však nedávají ani takto nejednoznačné výsledky a musíme na tomto místě připustit, že co do průměrné dynamické hodnoty hraných not jsme nevypozorovali žádnou

charakteristickou pravidelnost, kterou bychom v modulu pro humanizaci mohli použít. Práce s dynamikou je jednou z esenciálních interpretačních metod (19), která dává navíc hráči zpětnou vazbu v řádu několikanásobně větším, než je tomu u drobných alterací rytmiky, času. Je tedy nasnadě, že velmi záleží na uvědomělé, byť na základě hudebního vzdělání a zejména citu vytvořené, volbě učiněné interpretem. U většiny skladeb se rozptýl distribuce odchylky dynamiky pohybuje kolem 10 jednotek (viz např. Graf 4.18) parametru Velocity (připomeňme si rozsah parametru 0–127 jednotek), což není příliš velká změna.

Rozložení odchylek parametru Velocity od průměrné hodnoty



Graf 4.18 – Části skladby Capriccio Espagnol

Tato informace je pro nás užitečná s ohledem na pozdější volbu optimálního rozptylu distribuce náhodné veličiny získávané z generátoru pseudonáhodných čísel. Čísla kolem 10 jednotek se objevují v souvislosti s humanizací dynamiky i v literatuře. (24) (25)

4.6.2 Průběh dynamiky dlouhých not

V rámci zkoumání dynamického průběhu dlouhých not jsme si v části 4.1.3 na straně 24 vybrali skupinu vhodných not, kterou jsme vyňali z jejich původního kontextu a vytvořili pro ně zcela zvláštní tabulku. Tyto noty jsme jako obvykle srovnávali v pěti verzích nahrávek.

Výsledkem je pětice grafů udávající jasný výsledek. Ten je nápadně podobný dynamickým obálkám – jakýmsi hrubým ekvivalentům skutečného dynamického průběhu znění tónu²² – uváděným pro dechové nástroje v literatuře. (3) (26)

Každý tón lze z pohledu dynamické obálky rozdělit a čtyři fáze: nástup (Attack), útlum (Decay), podržení (Sustain), uvolnění (Release). První dvě fáze popisují vznik zvuku, než se

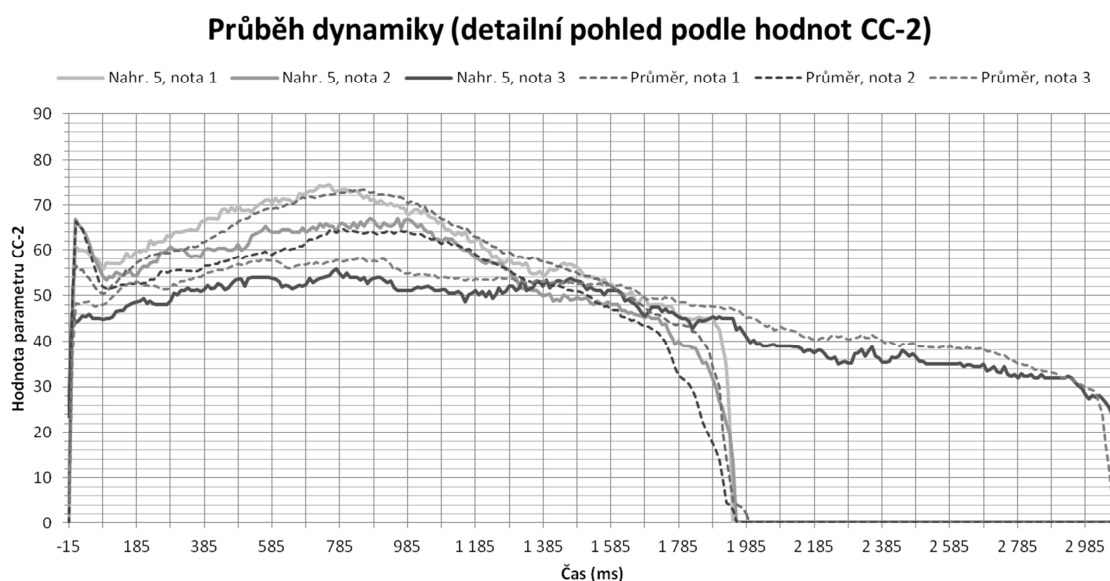
²² Dynamické obálky jsou známy rovněž pod pojmem ADSR (Attack, Decay, Sustain, Release) obálky.

dostane do fáze třetí, ve které zvuk setrvává v různých artefakty tolik nepoznamenané podobě. Třetí fáze odpovídá době, kdy je tón stále ještě buzen, ale již ustálen. Čtvrtá fáze pak shrnuje chování tónu po skončení buzení. Popis dynamické obálky pro nástroje bez kontinuálního buzení (různé perkusivní nástroje nebo strunné nástroje brnkací či hrané prstovými technikami) je nad rámec této práce, dechové nástroje ale sledují definici dynamické obálky téměř ukázkově.

Je však nutné si uvědomit určitý rozdíl mezi dynamickou obálkou zvuku celého dechového nástroje a obálkou pouhé síly dechu. Z fyzikálních procesů, které v dechovém nástroji zvuk vytváří, je zřejmé, že tyto obálky budou mít velmi podobný průběh, obálka síly dechu však není nijak modifikována např. rezonancí ozvučných částí nástroje atp.

Uved'me ještě, že zvolené tóny nejsou poznamenané žádnou speciální technikou, například pozvolným nástupem nebo dynamickými nuancemi při hře staccato, frulato apod.

Podívejme se nyní na graf páte nahrávky dlouhých not z Concerto for Orchestra (Graf 4.1). Dynamika rychle stoupá na maximální hodnotu tónu, která se později ještě objeví ve fázi podržení tónu. Nástup dechu na maximální dynamiku trvá přibližně 20 ms. Ihned po dosažení maxima dynamika dechu rychle padá o 10 jednotek níže. Sestup je spíše lineární, pozvolný a trvá asi 75 milisekund. Délka obou těchto fází je nezávislá na délce noty a je velmi přesně stejná ve všech pěti nahrávkách.



Graf 4.19 – Vybrané dlouhé noty z části skladby díla Concerto for Orchestra

Z hlediska humanizace bude nejzajímavější částí fáze držení tónu. Dynamika v ní nejdříve roste, až dosáhne maximální hodnoty na konci fáze nástupu tónu nebo ji mírně přesáhne. Poté začne zvolna klesat až do hladiny, ve které fázi držení zahájila. Celý průběh má tvar široké paraboly. Tato změna je zcela nepochybně dána získanou (naučenou) interpretovou dovedností, ale pro nás není problém ji konzistentně reprodukovat, což je pro věrohodné znění virtuálního dechového nástroje žádoucí. Zajímavé je, že u delší noty dochází

k maximalizaci dynamiky třetí fáze ve stejném absolutním čase (od začátku tónu) jako u not kratších. Tento čas odpovídá další celé době v taktu. Po celou dobu fáze držení tónu jsou patrné drobné odchylky v řádu 2–6 jednotek a toto chvění má frekvenci v řádu 0,01 Hz, pro lidské ucho tedy téměř nerozpoznatelná časová změna. Přesto toto chvění může působit na ucho na jiné než analytické rovině – zatímco u delších časů by si posluchač všiml kolísavé dynamiky a poslouchal by efekt podobný nebo dokonce identický s vibratem, v tomto případě je chvění natolik rychlé, že může přidávat na intenzitě nebo sytosti tónu, uvažujeme-li ovšem reálný akustický nástroj. Ověření této hypotézy je však nad rámec této práce a spadá spíše do oboru hudební akustiky a následně psychologie. Na druhou stranu reprodukce tohoto chvění nepředstavuje implementační problém.

Poslední fází je uvolnění. Začíná okamžikem, kdy interpret přestává vdechovat vzduch do trubice nástroje a proudění samovolně ustává. Tento sestup trvá díky neschopnosti člověka ustat s dechem z plného proudění okamžitě asi 200 ms. Setrvačnost vzduchu v nástroji má na zánik dynamiky zanedbatelný vliv, neboť odpor trubice je natolik vysoký, že po vědomém ukončení vdechování hráčem proudění končí téměř okamžitě. Zajímavou technikou je ukončení proudění vzduchu přiložením jazyka na náústek. V tom případě je čas uvolnění minimální. (27)

4.7 Nástin humanizace ladění a intonace

O dechových nástrojích obecně platí, že je u nich problém dosáhnout absolutní přesnosti v ladění. Pro ilustraci si uvedeme některé faktory, které ladění u dechového nástroje ovlivňují: teplota nástroje a vdechovaného vzduchu, dynamická intenzita vdechování, přesné naladění nástroje jako takového (délka zvukovodu), přesnost trubic mezi ventily u nástrojů žesťových. Většinu z nich ještě ovlivňuje změna výšky komorního A, protože dechy jsou konstruovány pro určitou frekvenci rezonance v nástroji a délce vln je přizpůsobena jak velikost nástroje, tak intervaly mezi klapkami a ventily – dechy tedy není možné dokonale přizpůsobit všem laděním. (17) (28) (29)

Jedinou výjimkou z většiny intonačních problémů je trombon (někdy také nazýván pozoun), který díky vybavenosti snížcem umožňuje plynule měnit délku vzduchové trubice v nástroji a tím i výšku ladění s potenciálně podobnou precizností, s jakou to zvládají například smyčce. Trombon je také označován za nejlépe ladící dechový nástroj. (30)

Samostatnou kapitolou je pak přesnost a stabilita nátisku na nátrubek u nástrojů žesťových, na plátek v případě nástrojů jedno- a dvouplátkových a nebo náústek u fléten (zde myslíme zejména flétnu příčnou a pikolu, kde hraje v ladění roli nasazení nástroje k ústům). (31)

Vzhledem k možnostem nástroje EWI, které jsme si již shrnuli dříve, ovšem nemáme možnost odchylky v ladění přesně změřit. I při použití rychlé automatické ladičky těžko můžeme vzorkovat v čase s dostatečnou rychlostí a precizností. Jestliže bychom naopak po interpretovi chtěli hrát tóny v dostatečné délce, narušili bychom zase hudební kontext a naměřené údaje by nebyly příliš relevantní.

Pro orientaci jsme nicméně získali nahrávky právě trombonových partů kvalifikace studentů Hudební fakulty JAMU pro účast na Pražském jaře. Z pěti sedmiminutových nahrávek různých

studentů hry na trombon jsme pomocí rychlé elektronické ladičky zjistili přibližnou odchylku ± 10 centů²³. U čtyř studentů byla patrná tendence intonovat mírně nad správnou hladinu, pouze u jednoho tomu bylo obráceně. Díky přílišné hrubosti měření však musíme tyto údaje uvažovat s rezervou.

V literatuře je u techniky hry vibratem uváděna maximální přípustná odchylka 50 centů, což je čtvrt tónu (v západní hudbě stále ještě méně než nejmenší možná tónová změna – uvažujeme nejrozšířenější rovnoměrně temperované ladění). (32) Taková hodnota je patrně přijatelná jen díky kolísání na obě strany přesné výšky tónu, průměr vibrata tedy vlastně stále leží přibližně v čistém tónu. Pohybujeme se ale stále v řádu, který jsme zevrubným pozorováním stanovili z nahrávek trombonů. U nich se určitá oscilace rovněž objevovala, ale není možné rozhodnout, zda to bylo dáno neschopností ladičky přesně identifikovat hraný tón a jeho odchylku, nebo k ní skutečně docházelo – to je pravděpodobné, nicméně touto metodou neprokazatelné.

Budeme se tedy spokojit s tím, že umožníme uživateli našeho modulu ovlivňovat i parametr výšky subtonálního ladění, kdy nejvyšší věrohodná hodnota se bude pohybovat kolem zmíněných ± 10 centů od přesné frekvence.

²³ Cent – každý tón resp. půltón se výškově dělí na 100 centů. Tedy C zvýšené o 100 centů zní stejně jako C#.

5 Metody humanizace

Tato kapitola se věnuje teorii stávajících metod používaných k humanizaci kvantizovaného MIDI záznamu a jejich vylepšením na základě výzkumu a analýzy provedené v rámci této práce.

5.1 Stávající metody

Stávající metody humanizace využívají téměř výhradně pseudonáhody a lineární a výjimečně normální distribuce náhodné veličiny. Existuje jen velmi málo modulů umožňujících alespoň volbu křivky rozložení náhodné veličiny, někdy naopak s výhodou možnosti vytvoření vlastní křivky pomocí kontrolních bodů.

Nevýhodou všech těchto metod je zejména nevěrohodnost generovaných odchylek, parametry jsou ovlivňovány mimo kontakt s realitou interpretačních schémat a neopírají se o poznatky skutečné hry hráče. I kdybychom měli k dispozici sofistikovaný nástroj umožňující dokonalou kontrolu rozložení náhodné veličiny a tím pádem i odchylek humanizovaného parametru, nemohli bychom vzít v potaz další zjištěné faktory nezbytné pro smysluplnou humanizaci – za všechny uveďme rozdíl v míře chybovosti na začátku taktu, na celých dobách a na notách hraných mimo ně. Neztracujeme na tomto místě použití pseudonáhody jako takové, budeme se jí zabývat i v rámci vylepšení humanizačních metod.

Jak jsme si v analytické části ukázali, nezbytným aspektem je ale kontrola nad změnou parametrů podle pozice v taktu, ve kterém se dané figury nachází. Primitivní formu této metody používají některé humanizační nástroje komerčních DAW (například Apple Logic Studio a funkce Swing obměňující odchylku na lichých a suchých dobách).

Motivací pro takové řešení je umožnit vyšší míru zanesení chybovosti při zachování použitelnosti. Náhoda při velké míře vytváří prosté chyby, ne uvěřitelnou „lidskost“. Obsáhlejší zdůvodnění si podáme přímo v části 5.2.1 věnované této metodě.

5.1.1 Humanizované parametry stávajících metod

Paleta cílových parametrů se v případě stávajících metod neliší od těch, které budeme využívat my. Jedná se tedy o modifikace postavení not v čase, změny jejich délky, dynamické odchylky a odchylky v ladění. Ovlivnění délky not se používá minimálně, stejně tak se příliš nevyužívá změn v ladění, některé metody navíc mění ladění po celých tónech ne v intervalech mikroladění – v takovém případě můžeme hovořit o humanizaci jen těžko. Nezahrnutí subtonálních odchylek je ale poměrně velká nedostatek, protože odchylky v ladění vedle posunu fáze vytváří dojem objemu hudebního tělesa – například dechové sekce v orchestru.

5.2 Dílčí nástroje vylepšení metody humanizace

5.2.1 Zpracování dle pozice noty v taktu

Tato metoda řeší problém změny humanizační míry na základě pozice modifikované noty v taktu. Chceme uživateli poskytnout jednoduchý nástroj, kterým by zvolil míru humanizace na první době taktu, na dobách celých a teoreticky i na kterékoliv jiné menší době, v rozlišení můžeme snadno jít až např. do rastru čtyřiašedesátinových not, což je nejjemnější rozlišení standardu General MIDI. (11) Myšlenkou je umožnit uživateli zvolit menší míru vychýlení např. u dob počátečních, přízvuchných, akcentovaných a naopak větší míru u těch, které nejsou z hlediska sehrání se s metrem či ostatními nástrojovými party v aranži tak podstatné a můžeme u nich připustit větší změny.

Po praktické stránce pak celým dobám budeme přiřazovat menší odchylky času začátku a délky, jak vyplývá z výsledků analýzy v předchozí kapitole. Porovnávací rastr bude vhodné zvolit podle jmenovatele taktu – tedy např. v případě 3/8 taktu je to 8 a můžeme předpokládat, že komponista skladbu uvedl v takovém taktu právě proto, že se v něm bude lépe orientovaná. V tomto příkladu má tedy smysl uvažovat za celou dobu nikoliv notu čtvrtovou, ale osminovou.

5.2.2 Detekce délky noty

V části 4.6 jsme si ukázali smysluplnost rozdělení humanizačního mechanismu podle délky modifikovaných not. V daném případě šlo o rozdílné zpracování dynamiky not dlouhých (v řádu čtvrtových not a delších) a krátkých. Není samozřejmě důvod, proč bychom nemohli využít poznatky získané o průběhu dynamiky delších not i na noty krátké, první dvě fáze dynamické obálky proběhnou v čase velmi krátkém a v závislosti na délce noty se neliší.

Na druhou stranu můžeme distinkci dle délky využít jiným způsobem – v analýze odchylek délky znění tónu přesnost ukončení s rostoucí délkou noty klesala.

5.2.3 Paměť kontextuálních událostí

Zajímavé vylepšení, se kterým stávající metody nepočítají, je paměť hudebního kontextu. V základní variantě by k detekci většiny jevů popsaných v analytické části stačilo, aby si modul ukládal informace vždy o jedné notě těsně předcházející té právě humanizované.

Využití takových informací je nasnadě – z analýzy vyplynuly poznatky vážící se na určitou podobu hudebního okolí. Bez znalosti tohoto okolí bychom jen těžko věrohodně reprodukovali zachycené jevy.

Paměťová banka může jít ale ještě dále. Nejenže může uchovávat informace o předešlých notách, ale také k nim se vážící informace o humanizačními algoritmy provedených úpravách. V případě humanizace délky noty jsme např. pozorovali korelaci mezi odchylkami času začátku a délky znění noty.

5.2.4 Limitace změn parametrů

Užitečným nástrojem bude využití limitace výsledných odchylek. Vlastně supluje zalomení křivky distribuce náhodné veličiny strmě k nule na stranách grafu, které se navíc u většiny humanizačně myslitelných distribucí nevyskytuje, ale představuje efektivní prostředek uživa-

telské kreativity. Navíc jde o velmi parametr, jehož funkce je srozumitelná a poměrně transparentní.

Princip tkví v tom, že uživateli umožníme určit hranici, kterou odchylky humanizovaných parametrů nepřekročí. Horní a spodní práh bude užitečné volit zvlášť nebo tento parametr uvést v součinnost s volitelným posunem střední hodnoty distribuce náhodné veličiny.

5.2.5 Volitelné křivky rozložení náhodné veličiny

Už víme, že při rozhodování budeme využívat generátoru pseudonáhodných čísel. Základní algoritmus generátoru pseudonáhodných čísel však vybírá čísla ze zadaného rozsahu s lineární distribucí – výskyt každého čísla z rozsahu je stejně pravděpodobný, což je nakonec jeden ze základních požadavků na takový generátor kladených. Jak jsme si ale ukázali v kapitole 4, graf distribuce odchylek měřených veličin spíše odpovídá tvaru Gaussovy křivky. Bude tedy nutné algoritmus rozšířit tak, aby distribuce veličin jím generovaných měla graf distribuce odpovídající této křivce.

Kromě normálního rozdělení může uživatel požadovat i jiné distribuční křivky, takže implementujeme i některá další rozložení, jejichž aplikace bude podávat zajímavé výsledky.

5.2.6 Pokročilý generátor pseudonáhody

Při posuzování kvality generátoru pseudonáhodných čísel hledíme zejména na co nejrovnoměrnější distribuci generovaných hodnot v oboru hodnot a také na jejich vzájemné korelace, které v ideálním případě neexistují, v praxi se však u všech generátorů nějaké vyskytují. Dalším kritériem je tzv. perioda opakování²⁴, která by měla být co nejvyšší. V případě implementace generátoru do VST modulu a jeho běhu v reálném čase nás zajímá také složitost (v konečném důsledku efektivita).

Naším účelům nejlépe poslouží generátor *Tausworthe* od Pierra L'Ecuyera implementovaný v knihovně *GNU Scientific Library* (tu budeme dále používat pro generování náhodné veličiny v určité distribuci). Perioda tohoto generátoru je 2^{88} (přibližně 10^{26}) a pracuje i na 64bitových systémech, což je v současnosti jeho velkou předností.

5.2.7 Determinismus

Humanizace modulem bude deterministická – to znamená, že pro identická vstupní data bude vytvořen identický humanizovaný obraz. Nechceme uživateli upřít možnost vytvářet různé verze humanizovaného materiálu na týchž MIDI datech, ale naopak mu umožnit nechat souznít více nástrojů současně se stejnými odchylkami. Taková funkce může mít zejména v současné elektroakustické hudbě své opodstatnění – již dříve jsme zmínili možnost spojení zvuku dvou bicích souprav. Obecně se bude jednat o případy, kdy potřebujeme svázat více různých zdrojů zvuku do úzkého celku, aniž by mezi jednotlivými částmi výsledného spojení vznikalo nežádoucí pnutí a artefakty (např. již zmíněný echo efekt).

Abychom zároveň zachovali flexibilitu řešení, umožníme uživateli ovlivnit hodnotu, která inicializuje generátor pseudonáhodných čísel. Tato hodnota se někdy také nazývá *seed*. Fun-

²⁴ Perioda opakování – hodnota vyjadřující množství pseudonáhodných hodnot, které dokáže generátor vytvořit, než se řada začne opakovat.

guje přesně očekávaným způsobem – jestliže iniciujeme dvě instance stejného generátoru pseudonáhody stejným parametrem seed, budou obě předávat na výstup přesně tytéž hodnoty v tomtéž pořadí. Naopak změnu tohoto argumentu můžeme využít ve chvíli, kdy jsme s nastavením parametrů modulu spokojeni, ale nelíbí se nám nějaké konkrétní místo, kde humanizační proces nezafungoval optimálně nebo esteticky kvalitně. V takové situaci pouhou změnou argumentu s velkou pravděpodobností přisoudíme danému místu úplně jiné humanizační vlastnosti. Po nějakém počtu změn hodnoty seed obvykle dosáhneme požadovaného výsledku.

5.3 Otázka polyfonie²⁵

Zatím jsme se vůbec nezabývali otázkou řešení polyfonie. Důvod je prostý – dechové nástroje jsou monofonní, umožňují hrát nanejvýš jeden tón současně. Pouze dovední hráči zvládající moderní kreativní techniky hry na některé dechové nástroje (jedná se o nástroje plátkové) dovedou využít jejich specifické architektury k polyfonní hře. V drtivé většině kompozic se však s takovým chováním dechových nástrojů nepočítá. Námi analyzované vybrané části skladeb rovněž obsahují pouze jednohlasé dechové party.

V zásadě bychom situaci mohli snadno řešit prostým omezením polyfonie na jeden tón, což je obvykle ten, který se jako poslední objeví na vstupu modulu. Na druhou stranu je v kompetenci uživatele, aby zvážil, zda požaduje věrohodnou práci s dechovými nástroji, nebo sáhne raději po polyfonní aranži, kterou nechá naším modulem procházet. Bez limitéru polyfonie se bude modul chovat pro polyfonní noty stejně jako pro jakýkoliv jiný monofonní part – jestliže se na vstup dostanou v jednu chvíli dvě noty začínající ve stejný okamžik nebo po určité době souznějící, budou zpracovány v pořadí příchodu MIDI zprávy každá zvlášť. Omezení leží pouze v tom, že modul nebude polyfonií okolí zohledňovat coby parametr ovlivňující humanizaci.

Díky tomu, že nebudeme polyfonii omezovat, zachováváme zároveň použitelnost i mimo původně zamýšlenou oblast nasazení (viz část 5.5).

5.4 Otázka tempa

Jak v úvodu, tak v analytické části jsme si nastínili možnost, že výskyt a podoba interpretačních artefaktů se mění v závislosti na tempu. V případě analýzy průběhu dlouhých not jsme si toto tvrzení částečně vyvrátili, avšak u jiných cílů humanizace se může změna v závislosti na tempu chovat odlišně.

I když dáme zapravdu premise, že v rychlejším tempu bude nepřesnost vyšší zejména kvůli nedostatku času na kvalitativně obstojnou technickou práci s nástrojem a v pomalejším tomu bude naopak (a na druhou stranu v pomalém tempu se může vyskytnout dříve zmíněný problém s nadhodnocováním nebo podhodnocováním metra), bylo by zbytečné tuto funkci samostatně implementovat – jen by zatěžovala uživatele dalšími parametry. Ten totiž může běžné parametry tempu přizpůsobit ručně dle svého uvážení a i jestliže dojde během skladby

²⁵ Polyfonie – znění více not současně.

ke změně tempa, automatizační křivky zpřístupní přizpůsobení humanizačního modulu nové hodnotě v kterémkoliv místě kompozice. Kontrolu poměru humanizace v závislosti na tempu skladby ponecháme tedy v rukou uživatele.

5.5 Rozšiřování kontextu

Ačkoliv uvedené metody vychází z analýzy provedené na datech získaných z interpretace vzorových etud pro dechové nástroje, samotný modul budeme koncipovat mnohem otevřeněji, takže některé jeho části budou obecně využitelné pro kterýkoliv nástroj a budou podávat věrohodné výsledky. Zde máme na mysli zejména odchylky dynamiky a pozice začátku.

Ukázali jsme si, že distribuce měřených odchylek se blíží Gaussově křivce, která bude spolehlivě popisovat například rozložení chyb v rytmozaci na perkuse. Ačkoliv je modul určen pro zpracování dechových partů a odeslání humanizovaných MIDI dat do virtuálního nástroje s dechovými rejstříky, za jistých okolností může podat optimální výkon i na partiturách určených libovolnému jinému, v MIDI zachytitelnému nástroji. Hlavním argumentem zde totiž není pouhá schopnost rytmických a dynamických změn, ale součinnost s ostatními pomocnými metodami, kterými konkurenční software nedisponuje (pozice vzhledem k celým dobám, analýza délky noty).

Konkrétní míru i formu nasazení modulu na MIDI záznam nedechových nástrojů ponecháme plně na uživateli. Hlavním ukazatelem je konečně uspokojivost výsledků pro komponistu/producenta, potažmo posluchače. Komponista tak může na základě těchto kritérií nejlépe zvážit, zda mu zvolené parametry poskytují uspokojivý výstup, což by při použití modulu ve vhodném nastavení na druhou stranu mělo být možné.

6 Aplikace

Praktickým výstupem této práce je aplikace určená pro hostitelské aplikace typu DAW, v nichž bude provádět automatizovanou humanizaci MIDI dat v souladu s uživatelem nastavenými parametry. Aplikace bude respektováním charakteristických interpretačních artefaktů hráče na dechový nástroj identifikovaných v analytické části práce rozšiřovat paletu stávajících implementovaných humanizačních metod a posouvat možnosti humanizace MIDI záznamu dále.

V této kapitole si přiblížíme základní úskalí návrhu modulu, jejich řešení a postup implementace v prostředí VST 3.5.

6.1 Standard VST Module Architecture

Standard VST Module Architecture (dále jen VST MA) je produktem německé firmy Steinberg, která je rovněž významným producentem softwarových DAW, efektových modulů, ale také hardwarových zařízení. Kromě vlastní řady produktů (za všechny jmenujme alespoň Cubase a Nuendo) firma podporuje i vývojáře externích modulů pro své hostitelské aplikace. Aby byly tyto moduly přenositelné mezi různými systémy a často i platformami, byl vytvořen standard VST²⁶. Ten má řadu rodin a verzí. Pro vývoj zvukových efektů a procesorů se používá VST standard. Jeho základ tvoří právě VST MA, jejíž součástí je modul pro přímou práci s MIDI. Jeho podpora je bohužel omezena na produkty firmy Steinberg a ačkoliv i ostatní výrobci podporují standard VST jako takový, hostování MIDI efektů VST MA v aplikacích jiných výrobců často není možná.

Proto budeme modul vyvíjet jako virtuální nástroj VSTi²⁷, který nebude generovat zvuk, ale přijímat a odesílat MIDI události. Použitím standardního směrování datových toků v libovolné aplikaci podporující standard VST můžeme z jedné stopy určené pro zpracování událostí odeslat MIDI data určená k humanizaci, jinou stopou stejného typu pak zhumanizovaná data z modulu přijmout a hned odeslat do dalšího VST nástroje, který už bude generovat zvukový signál.

Volba tohoto standardu tedy leží v udržení multiplatformové nasaditelnosti modulu. Jeho využití tak tak nebude limitováno na produkty firmy Steinberg – standard VST je v současné době nejrozšířenější na platformě Microsoft Windows a lze pro něj vytvořit zapouzdřující aplikaci, která zajistí funkčnost i v rámci standardu Audio Unit pro systémy Apple Logic na Mac OS, nemluvě o přímé podpoře VST aplikacemi jako Ableton Live distribuovanými pro Windows i Mac OS.

²⁶ VST – Virtual Studio Technology

²⁷ VSTi – i znamená Instrument, využívá se pro odlišení virtuálních nástrojů a zvukových efektů, které oba patří do standardu VST.

Pro implementaci používáme software development kit VST 3.5²⁸. Ten implementuje základní třídy pro nasazení v prostředí podporujícím VST MA standard a obsahuje také potřebnou dokumentaci. Standard je nativně implementován v prostředí jazyka C++ a ač existují neoficiální alternativy i v jiných jazycích, budeme modul vyvíjet právě v něm.

6.2 Analýza požadavků

Základní požadavky na fungování humanizačního modulu jsme si již v shrnuli v části 5.2. Problematiku funkčnosti humanizačního modulu si rozvedeme dále v části návrhu aplikace, kde je přímo spojíme s návrhy jejich řešení.

6.3 Návrh aplikace

6.3.1 Parametry v MIDI

Hostitelská DAW aplikace bude našemu modulu předávat MIDI data v podobě zpráv. Modul zprávy přijme, zpracuje a některé z nich modifikuje. Všechny pak předá na výstup. Pro zcela běžné řetězení více MIDI efektů za sebou je to nezbytné – skutečnost, že náš modul některé parametry neošetřuje, neznamená, že to nebude dělat některý jiný v řetězci.

Note On je základní zprávou přenášející informaci o začátku noty. Obsahuje dva datové byty s informací o síle stisku (Velocity) a výšce noty. Následuje status byte, který identifikuje zprávu právě jako Note On a obsahuje informaci o čísle kanálu. Posunutím této zprávy v čase budeme humanizovat pozici začátku znění noty. Změnou odesílané hodnoty bytu s informací Velocity humanizujeme základní dynamiku tónu. Zpráva Note On s hodnotou Velocity 0 odpovídá také informaci o ukončení znění noty (nulová hlasitost).

Note Off je téměř identická s Note On a přenáší zřejmě informaci o konci znění noty. Volitelně také obsahuje informaci o intenzitě uvolnění – Release Velocity. Tuto informaci ale dechový kontrolér EWI nepřenáší, její simulaci se tedy nemá smysl zabývat. Modifikací času, ve kterém odešleme na výstup přijatou zprávu Note Off, ovlivníme délku znění noty.

MIDI Tuning Messages implementuje Steinberg přímo do svého VST 3.5 SDK ve formě prostého parametru události Note On. Alternativou k metodě využívající ladicích zpráv je použití ovladače Pitch Bend, to má však zásadní nevýhodu – umožňuje změnit pouze globální ladění, takže ovlivní všechny v daný okamžik znějící tóny. V monofonním kontextu toto nepředstavuje problém, podobná situace se nemůže objevit, ale při polyfonním materiálu musí uživatel vyvstalou situaci nějak ošetřit. Implementace řešení změny mikroladění pomocí MIDI Tuning zpráv je rovněž standardní cestou, která tuto nevýhodu zcela odstraňuje, korekce ladění se váže vždy na právě zpracovávanou notu.

²⁸ SDK VST 3.5 je dostupný na adrese <http://www.steinberg.net/en/company/developer.html>.

6.3.2 Determinismus a pseudonáhoda

V práci budeme využívat knihovnu *GNU Scientific Library*²⁹ (dále jen GSL), která nám usnadní implementaci generování čísel z různých rozsahů na základě odpovídajících rozložení.

V části 5.2.7 jsme si vysvětlili důležitost determinismu používaných algoritmů a celého systému. Zavedeme si proto proměnnou *seed*, kterou budeme inicializovat generátor pseudonáhodných čísel pomocí funkce `gsl_rng_set`.

```
void gsl_rng_set (const gsl_rng * r, unsigned long int s)
```

Generátor inicializovaný stejnou hodnotou *seed* bude generovat pokaždé stejnou řadu čísel. Výhodou použití GSL je možnost generovat náhodná čísla přímo respektující požadovanou distribuci, například pomocí funkce `gsl_ran_gaussian_ziggurat`, která využívá algoritmus Ziggurat³⁰ pro získání náhodné veličiny podle Gaussovy distribuce.

Proměnnou *seed* umožníme pro pohodlnost uživatele generovat rovněž náhodně běžným způsobem za použití funkce `time`. Proměnná bude vypsána do uživateli přístupného pole, takže pokud požaduje použití nějakého konkrétního *seedu*, jehož hodnotu si v minulosti poznamenal, může ji využít místo té pseudonáhodně vygenerované.

6.3.3 Distribuční funkce a jejich implementace

Poměrně složitou algoritmizaci distribučních a inverzních distribučních funkcí ponecháme na implementacích z GSL. Knihovna zahrnuje několik desítek nejpoužívanějších distribucí včetně moderních optimalizovaných algoritmů. Ačkoliv je pro nás kvalita náhodných veličin podstatným faktorem, nechceme, aby pouhý MIDI modul měl příliš vysoké nároky na hardware, zvláště ve chvíli, kdy vezmeme v úvahu běh více jeho instancí současně na různých stopách v hostitelské aplikaci například při humanizaci více různých nástrojů dechové sekce.

K implementaci jsme vybrali šestici distribucí, jejichž odpovídající GSL funkce vracejí náhodnou veličinu v humanizačně zajímavých rozděleních.

Gaussova distribuce – nebo také normální rozdělení je distribuce vyskytující se v kontextu popisu odchylek náhodných jevů v přírodě nejčastěji. Graf hustoty normálního rozdělení má známý zvonovitý tvar tzv. *Gaussovy křivky*. Hustota pravděpodobnosti Gaussovy distribuce je dána vztahem:

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

Z knihovny GSL budeme využívat funkci `gsl_ran_gaussian_ziggurat`, jejímž jedinou proměnnou (kromě generátoru pseudonáhody *r*; dále ji již nebudeme zmiňovat) σ , předávající funkci požadovaný rozptyl náhodné veličiny.

```
double gsl_ran_gaussian_ziggurat (const gsl_rng * r, double sigma)
```

²⁹ Knihovna je distribuována pod licencí GNU General Public License a je dostupná na adrese <http://www.gnu.org/software/gsl/>.

³⁰ Algoritmus Ziggurat – v současnosti nejrychlejší algoritmus generování náhodných veličin odpovídajících normální distribuci. Autory jsou George Marsaglia a Wai Wan Tsang.

Studentova distribuce – někdy také nazývaná *t-distribuce*. Je podobná normálnímu rozdělení, ale křivka jejího grafu není tolik strmá – hodnoty náhodné veličiny mají v porovnání s normálním rozdělením větší tendenci padat do extrémů. To je také důvodem jejího výběru k implementaci. Hustota pravděpodobnosti studentova rozdělení je dána vztahem:

$$f(t) = \frac{\Gamma \frac{r+1}{2}}{\sqrt{\pi r} \cdot \Gamma \cdot \frac{r}{2} \cdot \left(\frac{1+t^2}{r}\right)^{\frac{n+1}{2}}}$$

kde $r = n - 1$ je počet stupňů volnosti, $-\infty < t < \infty$, $\Gamma(z)$ je funkce gama.

GSL funkce vracející náhodnou veličinu ze studentovy distribuce má (kromě generátoru pseudonáhody) jediný parametr, kterým je *stupeň volnosti* nu .

```
double gsl_rng_tdist (const gsl_rng * r, double nu)
```

Obor hodnot této funkce (interval jejích návratových hodnot) je jako u všech těchto symetrických distribucí $-\infty < x < \infty$.

Laplaceova distribuce – zajímavá charakteristika Laplaceovy nebo tzv. dvojité exponenciální distribuce je zlom křivky grafu distribuce v její střední hodnotě (kde se setkávají dvě exponenciály). Hustota pravděpodobnosti je:

$$f(x) = \frac{1}{2a} e^{-\frac{|x|}{a}}$$

kde $x \in R$, $a > 0$. Parametrem příslušné funkce z knihovny GSL je *šířka* a .

```
double gsl_rng_laplace (const gsl_rng * r, double a)
```

Plochá (uniformní) distribuce – tato distribuce vznikne v situaci, kdy všechny hodnoty náhodné veličiny v určitém intervalu mají stejnou pravděpodobnost a pravděpodobnost výskytu všech ostatních je nulová.

$$f(x) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & \text{jinak} \end{cases}$$

Námi hledaným křivkám se nijak nepodobá, přesto je využívána některými stávajícími humanizačními nástroji a za určitých okolností můžeme jejím využitím dosáhnout zajímavých výsledků. GSL funkce požaduje dva parametry pro definici intervalu, protože je ale v aplikaci budeme obě řídit jediným parametrem, položíme $b = -a$.

```
double gsl_rng_flat (const gsl_rng * r, double a, double b)
```

Beta distribuce – tuto distribuce budou uživatelé volit, jestliže budou hledat experimentálnější výsledky běhu humanizačního algoritmu. V naší implementaci modifikuje parametry a , b tak, aby vždy ležely v intervalu $[0; 1]$, kdy $a + b = 1$, pak totiž funkce `gsl_rng_beta` vrací především extrémní hodnoty. Hustota pravděpodobnosti beta distribuce je určena vztahem:

$$f(x) = \frac{x^{a-1}(1-x)^{b-1}}{\int_0^1 t^{a-1}(1-t)^{b-1} dt}$$

Prototyp GSL implementace pak je:

```
■ gsl_ran_beta (const gsl_rng * r, double a, double b)
```

a opět nám bude stačit jediný uživatelem modifikovaný parametr p , protože položíme:

$$a = \left(\frac{3}{5}p + \frac{1}{5}\right) \text{ a } b = 1 - a.$$

Obor hodnot této funkce ale leží pouze v intervalu $[0; 1]$, takže návratovou hodnotu normalizujeme využitím vztahu $x = 2x - 1$.

Levyho vychýlená alfa-stabilní distribuce – je rozdělení, které v určité parametrizaci odpovídá normální distribuci. My ale využijeme těch situací, kdy se křivka vychyluje od jedné strany střední hodnoty a na druhé se četnost strmě blíží nule.

Její funkce rozložení pravděpodobnosti nemůže být analyticky zapsána, můžeme ale zapsat její obecnou charakteristickou funkci:

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \varphi(t) e^{-ixt} dt$$

GSL implementace funkce vracející náhodnou veličinu v tomto rozložení požaduje tři parametry: *velikost* c , *exponent* α a *vychýlení* β .

```
■ double gsl_ran_levy_skew (const gsl_rng * r, double c,
    double alpha, double beta)
```

V naší implementaci položíme:

$$c = 1, \alpha = \frac{7}{10} \left| p - \frac{1}{2} \right| \text{ a } \beta = 2p - 1 \text{ pro } p \in [0; 1]$$

čímž dosáhneme současného vychýlení při změně velikosti odchylky za použití jediného parametru p a možnosti měnit vychýlení na obě strany od střední hodnoty.

6.4 Graphical user interface

Většina hostitelských aplikací pro VST moduly implementuje implicitní *graphical user interface* (tj. grafické uživatelské rozhraní; dále jen běžné označení GUI), jestliže modul neobsahuje nějaké vlastní. Pro usnadnění práce uživatele jsme se rozhodli vytvořit vlastní, názornější GUI, odpovídající požadavkům na efektivní a přehledné ovládání modulu.

Pro zpřístupnění modulu širší veřejnosti byla logickou volbou jazyka uživatelského rozhraní angličtina.

GUI budeme koncipovat jako strohé, nerušivé, funkční. Přesto by si mělo zachovat vizuální atraktivitu a současný vzhled. Jako název jsme zvolili jednoduše *Humanizer*, protože je z něj

okamžitě jasné, jaká je funkce modulu. Podtitul *Based on Real Performances* pak má za cíl na první pohled odlišit produkt od konkurenčních aplikací.

Pro zlepšení ovladatelnosti nasadíme vzájemně provázanou dvojici ovládacích prvků – otočné ovladače podobné tradičním otočným potenciometrům hudebního hardwaru a editovatelný číselník zobrazující přesnou číselnou hodnotu toho kterého ovladače. Uživatel tak může experimentovat s posunem parametrů a zároveň je přesně doladovat, jestliže mu práce s otočným ovladačem nebude precizností dostačovat.

Celková velikost okna GUI nesmí být i přes množství zpřístupněných parametrů příliš velká, důraz tedy klademe i na kompaktnost řešení. Výsledný návrh GUI viz Obrázek 6.1.



Obrázek 6.1 – Grafické uživatelské rozhraní humanizačního modulu

6.4.1 Automatizace v DAW

Všechny měnitelné parametry modulu budou díky implementaci v rozhraní VST MA řiditelné tzv. automatizací. Automatizační křivky jsou grafy zachycující průběh nějakého parametru v čase, typicky čase skladby. Tak, jako může zvukař měnit např. dynamickou úroveň nástroje v mixu na koncertě, může uživatel DAW měnit libovolný pro automatizaci exportovaný parametr. Výhodou automatizace v prostředí hostitelských aplikací je možnost jejich záznamu a dokonce jejich programování předem. (9) Uživatel tak může měnit i mnohem více parametrů, než by fyzicky zvládl v reálném čase, a často i velmi netradičním způsobem – například je velmi obtížné provést skokovou změnu fyzickým tahovým ovladačem, za použití automatizace toto je dílem jediného tahu myši předem.

Uživatel tak dostane pod kontrolu všechny parametry ovlivňující humanizační algoritmy používané v modulu. Některé části skladby tedy mohou být humanizované odlišným způsobem než jiné bez nutnosti vytvářet pomocné stopy a stříhat MIDI data do bloků, které by bylo

nezbytné mezi pomocné stopy distribuovat. Z modulu se tím stává silný nástroj, jehož funkcionality nebude nutné obcházet ani v jeho mezních možnostech – snadno ji přizpůsobíme aktuálnímu kontextu skladby.

6.5 Implementace

Modul Humanizer je implementován v jazyce C++, což je jazyk, ve kterém je napsaná současná verze SDK *Virtual Studio Technology 3.5*. SDK poskytuje základní třídy, pro integraci vyvíjeného modulu do prostředí hostitelské aplikace standard VST podporující. SDK obsahuje předpřipravená řešení pro prostředí Microsoft Visual Studio, ve kterém byl modul také vytvořen. Humanizer byl odvozen od vzorového projektu efektového modulu AGain, jenž je distribuován společně s vývojářským balíčkem.

Aplikace je koncipována jako virtuální nástroj. Kvůli kompatibilitě stále implementuje výstupní zvukovou sběrnici, což některé hostitelské aplikace požadují, ale ve skutečnosti pouze vstupní datovou sběrnici přijímá události, zpracovává je a předává zpět na výstupní datovou sběrnici hostitelské aplikaci.

Modul je kompilován jako dynamická knihovna (DLL) pro prostředí Windows, ač je přípona souboru v souladu se specifikací VST 3.5 vedena jako vst3. Součástí práce jsou i zdrojové soubory, které umožňují převod aplikace i pro běh na jiných systémech (existuje například Audio Unit³¹ můstek pro platformu Mac³²). Výsledný modul pak v hostitelské aplikaci běží jako dvě vlákna, z nichž jedno realizuje samotnou činnost modulu a druhé obsluhuje grafické uživatelské rozhraní. Druhé vlákno lze podle potřeby ukončit a znovu inicializovat podle toho, jak uživatel uzavírá či otevírá okno GUI v aplikaci.

V rámci SDK byly přidány třídy zastávající operace s parametry modulu a jejich automatizací, konkrétně šlo o přidání parametrů pro procenta, časový posun, dynamickou změnu, nastavení distribuční křivky a změnu ladění. Nad rámec vývojového balíčku (třídy nejsou odvozeny od existujících rodičovských tříd v SDK) jsme implementovali dvě třídy realizující identifikaci vlastností noty a pokročilé generování náhodných veličin, obě jsou vedeny v samostatných souborech.

Aplikace využívá část knihovny GNU Scientific Library. Licence, pod níž je distribuovaná (GNU General Public License), umožňuje vyjmutí částí knihovny z celého balíku a její autoři to kvůli odlehčení výsledných aplikací doporučují. Knihovní funkce zahrnují různé druhy sofistikovaných matematických operací a algoritmů. V našem případě využíváme moduly pro generování pseudonáhody (standardní implementace metody rand v C++ může být v závislosti na systému, pro nějž je modul zkompilován, nedostatečná) a moduly pro operace s distribucemi náhodných veličin. Moduly, které nebyly pro kompilaci aplikace s požadovanými knihovními funkcemi nezbytné, byly odstraněny.

³¹ Audio Unit – standardu VST architektonicky podobný formát vyvíjený společností Apple.

³² Viz například <http://code.google.com/p/symbiosis-au-vst/> nebo <http://www.fxexpansion.com/>.

6.6 Licence

Protože modul Humanizer využívá součásti šířené pod licencí GNU General Public License³³, je také výsledný softwarový produkt uvolněn pod touto licencí. Předpokládáme, že dostatečným zveřejněním zdrojových kódů, které tato licence požaduje, bude umístění všech částí práce v systému *Theses.cz* resp. veřejné části elektronického archivu závěrečných prací v rámci *Informačního systému Masarykovy univerzity*.

Licence firmy Steinberg na použití SDK VST 3.5 dovoluje jeho využití pro vývoj komerčních i nekomerčních aplikací a pokud nedojde k redistribuci modifikovaného SDK jako takového bez souhlasu poskytovatele licence, není tato konfliktní s GNU GPL.

³³ V celém znění na <http://www.gnu.org/licenses/gpl-3.0.txt>.

7 Závěr

7.1 Motivace

Humanizace MIDI záznamu je opomíjené téma, které paradoxně zaostává za současným trendem vývoje aplikací, které umožňují stále věrnější napodobení zvuku reálných nástrojů, dozvuku reálných prostor, syntézy reálného hlasu a dalších. Filmovou hudbu dnes nahrává symfonický orchestr jen u těch nejdražších projektů, protože možnosti digitální reprodukce, samplování, matematického modelování chování nástrojů i hardwarových zařízení pokročily do takové míry, že lze získat věrohodný, živý zvuk, aniž by jeho produkce zahrnovala jiného živého člověka, než je komponista a zvukař.

Přitom všechny tyto nástroje stále používají primitivní, ale doposud nepřekonaný komunikační protokol uchováající a předávající řídicí informace, na jejichž základě teprve inženýrsky vyšperkované a odladěné virtuální nástroje generují zvuk. Je to právě MIDI, jehož vznik se datuje do 80. let minulého století, které stále dostatečně precizně zprostředkovává komunikační kanál mezi instrumentalistou a jeho nástrojem.

MIDI dávno vykročilo z pouhého nástroje pro dorozumění se mezi hardwarovými zařízeními v reálném čase, dnes je hojně používáno pro uchování hudební informace přibližně na úrovni tradiční partitury. MIDI už nemusíme jen nahrávat, ale i přímo zapisovat podobně, jako to komponista dělá s notami. Proto je důležité poskytnout mu efektivní nástroje pro přiblížení jeho ručně psané MIDI partitury nahrávce živého interpreta. Při objemu not, který je potřeba zpracovat, by bylo téměř nemyslitelné provádět humanizační úpravy ručně, ale metody automatické modifikace MIDI záznamu jsou v současnosti stále nedokonalé.

Analýza ve výzkumné části této práce identifikuje interpretační artefakty charakteristické pro hru skutečného instrumentalisty dechových nástrojů – saxofonu a klarinetu. Na vybraných částech skladeb známých všem hráčům na tyto nástroje (jedná se o „učebnicové příklady“) jsme si ukázali základní nuance v chování hráče v toku hudebního materiálu a uvedli je v souvislost s hudebním kontextem, v němž se objevují.

Zdůrazněním těchto artefaktů v textu práce a jejich reprodukcí pomocí modulu standardu VST 3.5 s prostým názvem Humanizer jsme se pokusili ukázat, že vnesení větší věrohodnosti či estetické ocenitelnosti netkví v co nejpřesnější a nejvěrnější reprodukci tónu dokonalých houslí, ale jde ruku v ruce s MIDI partiturou, bez níž bychom neslyšeli ani hlásku, a na práci s níž je toho ještě hodně co zlepšit.

7.2 Zdokonalení humanizačních metod

Díky statistické analýze interpretační dovednosti hráče na dechový nástroj jsme mohli navrhnout a představit dílčí nástroje, s jejichž pomocí lze s úspěchem zdokonalit stávající humanizační algoritmy používané v mnoha současných softwarových produkčních stanicích. Není to však pouhé nasazení těchto nástrojů, ale zejména jejich součinnost, která ve spojení

s citlivou parametrizací dokáže značně zpřesnit nasazení pseudonáhodně modifikovaných hodnot MIDI dat tak, aby došlo ke kýžené reprodukci interpretačně charakteristických artefaktů při udržení uživatelské přístupnosti a jednoduchosti ovládání.

Vezmeme-li v úvahu stávající metody generování odchylek, které pouze výjimečně využívají volitelné distribuce náhodných veličin, v některých parametrech zaostávají zcela (uved'me například změny mikroladění) a nadto humanizují bez jakéhokoliv zohlednění kontextu modifikované noty, vidíme, že je naše řešení v mnohém předčí. Volitelná míra odchylky od kvantizovaných dat na základě informací o pozici noty v taktu nebo o její délce, které navíc řešení bez zásahu uživatele vyhledává, nám dovoluje nastavit nižší míru ovlivnění na akcentovaných notách, kde by byla velká nepřesnost nežádoucí, a naopak vyšší tam, kde si ji dovolit můžeme.

Tyto změny posouvají počítačem prováděnou humanizaci do oblasti efektivních a účinných řešení problému věrohodné reprodukce kvantizovaného MIDI záznamu.

7.3 Zhodnocení

Teoretická část této práce nám odhaluje, jakým způsobem vědomě přistupovat k humanizaci kvantizovaného materiálu a kde leží hlavní body, které bychom v tomto procesu měli vzít v potaz. Práce s hudebním materiálem je nakonec tvůrčí proces od první myšlenky k poslednímu souboru mp3 na internetu. Nechceme tedy hudební tvůrce svazovat konvencemi, ale poukázat na nuance, jež se ve hře živých interpretů objevují a které zásadně odlišují interpretaci umělcem disponujícím duševním výkonem od počítače disponujícího zejména výkonem výpočetním. Námi předkládaný materiál leží někde na pomezí hudebního aranžérství a zvukového inženýrství a může být užitečný jak komponistům realizujícím svou tvorbu s pomocí počítače, tak i zvukařům pomáhajícím skladatelům a producentům dokončit jejich hudební díla.

Po praktické stránce dostávají hudebníci k dispozici nástroj na efektivní a všestrannou humanizaci MIDI záznamu. Ač vychází z monofonní interpretace partů určených dechovým nástrojům a bude pro takový hudební materiál fungovat nejlépe, možnosti jeho zásahu do MIDI dat přes něj procházejících jsou natolik univerzální a zároveň elegantně manipulovatelné, že dobře poslouží i pro modifikaci partů určených jiným, než dechovým nástrojům.

Celým tímto textem prochází neustálé explicitní i podtextové srovnávání lidské činnosti s prací počítačů. V dnešní době jsme se ale už dostali do prostoru, ve kterém mohou oba existovat v syntéze svých možností a vzájemnou kompenzací nedostatků toho druhého dosahovat překvapivě efektních výsledků za nízkou cenu invenční i výpočetní. Teoretická část diplomové práce ve spojení se svým praktickým výstupem v podobě VSTi modulu Humanizer právě v tomto duchu přibližuje strojovou hudební reprodukci počítače skutečné interpretaci.

Rejstřík

- algoritmus, 48, 49, 54, 57, 58, 60
- aplikace, 8, 21, 25, 52, 53, 58
- artefakt, 12, 25, 28, 44, 49, 52, 60, 61
- automatizace, 8, 57, 58
- cent, 46
- DAW, 21, 23, 47, 52, 53, 57
- determinismus, 49, 54
- distribuce, 34, 36, 40, 43, 47, 48, 49, 51, 54, 55, 56, 58, 61
- dynamika, 11, 12, 13, 15, 16, 20, 23, 24, 25, 26, 33, 40, 41, 43, 44, 45, 47, 48, 51, 53, 58
- efekt (proces), 13, 40, 45, 49, 52, 53, 58
- EWI, 15, 16, 18, 19, 21, 24, 25, 41, 45, 53
- Gaussova distribuce, 35, 41, 47, 49, 54, 55, 56
- generování, 13, 14, 36, 43, 47, 49, 52, 54, 58, 60, 61
- GNU Scientific Library, 54, 56
- humanizace, 11, 12, 13, 15, 16, 18, 22, 28, 34, 36, 40, 43, 44, 45, 47, 48, 49, 50, 52, 53, 54, 55, 57, 60, 61
- Humanizer, 56, 58, 59, 60, 61
- chybovost, 8, 9, 11, 20, 27, 28, 36, 47, 51
- interpretace, 8, 11, 12, 14, 15, 18, 19, 20, 21, 22, 23, 25, 26, 28, 32, 40, 43, 47, 51, 52, 60, 61
- kvantizace, 11, 13, 15, 18, 22, 26, 27, 34, 36, 47, 61
- metoda, 9, 12, 13, 14, 18, 28, 34, 43, 46, 47, 48, 51, 52, 53, 58, 60, 61
- MIDI, 3, 9, 11, 12, 13, 14, 15, 18, 21, 22, 23, 24, 25, 40, 47, 48, 49, 50, 51, 52, 53, 54, 57, 60, 61, 65
- mikroladění, 12, 16, 25, 46, 53, 61
- modelování, 13, 14, 60
- modul, 10, 12, 13, 43, 46, 47, 48, 49, 51, 52, 53, 54, 57, 58, 59, 60, 61
- náhodná veličina, 36, 40, 43, 47, 48, 49, 54, 55, 56, 58, 61
- nástroj dechový, 9, 13, 16, 18, 19, 21, 25, 43, 44, 45, 50, 52, 53, 60, 61
- nuance, 8, 9, 12, 14, 22, 23, 27, 28, 44, 60, 61
- odchylka, 16, 21, 22, 23, 25, 26, 27, 28, 29, 30, 32, 33, 34, 35, 36, 39, 40, 41, 43, 45, 46, 47, 48, 49, 51, 54, 56, 61
- partitura, 11, 13, 14, 18, 19, 22, 26, 27, 28, 29, 36, 37, 40, 41, 51, 60
- pseudonáhoda, 28, 36, 43, 47, 49, 54, 55, 58, 61
- rytmus, 8, 15, 16, 20, 21, 30, 32, 36, 39, 40, 43, 51
- SDK, 53, 58, 59
- skladatel, 9, 12, 61
- Steinberg, 10, 21, 52, 53, 59
- uživatel, 8, 9, 46, 48, 49, 50, 51, 52, 53, 54
- Velocity, 24, 25, 27, 41, 42, 43, 53
- Virtual Studio Technology, 10, 49, 52, 53, 56, 58, 59, 60, 61
- virtuální nástroj, 9, 13, 14, 17, 21, 44, 51, 52, 58, 60

Použitá literatura

1. **Poledňák, Ivan.** *Hudba jako problém estetiky*. Praha : Nakladatelství Karolinum, 2006. ISBN 80-246-1215-1.
2. **Forró, Daniel.** Automatizace hudby. *Muzikus.cz*. [Online] MUZIKUS, 30. 11 2002. [Citace: 23. 4 2011.] <http://www.muzikus.cz/pro-muzikanty-clanky/Automatizace-hudby~30~listopad~2002/>.
3. **Kopecký, Pavel.** *Základy elektronického zvuku a jeho kreativní zpracování*. Praha : Akademie múzických umění v Praze, 2008. ISBN 978-80-7331-121-6.
4. **Porter, Noah.** *Webster's Revised Unabridged Dictionary*. Springfield, Massachusetts, USA : G & C. Merriam Co., 1913.
5. **Houghton Mifflin Harcourt.** *The American Heritage® Dictionary*. Boston, Massachusetts, USA : Houghton Mifflin Harcourt, 2000. ISBN-13 978-0395825174.
6. **Holas, Milan.** *Hudební nadání*. Praha : Hudební fakulta AMU v Praze, 1994. ISBN 80-85883-007.
7. **Lissa, Zofia.** *Nové studie z hudební estetiky*. Praha : Supraphon, 1982. cnb000856496.
8. **Hons, Miloš.** *Hudební analýza a hudební myšlení*. Ústí nad Labem : Katedra pedagogiky Pedagogické fakulty Univerzity J. E. Purkyně v Ústí nad Labem, 2009.
9. **Vlachý, Václav.** *Praxe zvukové techniky*. Praha : Nakladatelství Muzikus, 2008. ISBN 978-80-86253-46-5.
10. **Urban, Ondřej.** *Instrumentář elektroakustické hudby*. Praha : Akademie múzických umění v Praze, 2007. ISBN 978-80-7331-115-5.
11. **Guérin, Robert.** *Velká kniha MIDI*. Brno : Computer Press, 2004. ISBN 80-7226-985-2.
12. **Jirsák, Martin.** Seri(ál)ová MIDI komunikace. *Muzikus*. Měsíčník, 2010, Sv. XXVI, 04.
13. **Kulka, Jiří.** *Psychologie umění*. Praha : Grada Publishing, a.s., 2008. ISBN 978-80-247-2329-7.
14. **Wilkinson, Scott.** Microtuning. *Electronic Musician*. [Online] 1. 2 2002. [Citace: 2011. 5 3.] http://emusician.com/mag/emusic_microtuning/.
15. **Syrový, Václav.** *Hudební akustika*. Praha : Akademie múzických umění v Praze, 2003. ISBN 80-7331-901-2.
16. —. *Technické základy elektroakustické hudby*. Praha : Akademie múzických umění v Praze, 1990.

17. **Kratochvíl, Jiří.** *Praktická metodika hry na klarinet 1.* Praha : Hudební fakulta Akademie múzických umění v Praze, 1996. ISBN 80-85883-14-7.
18. **Baisa, Vít.** *Zpracování dat z MID souboru pro potřeby hudební komparatistiky.* Brno : Fakulta informatiky Masarykovy univerzity, 2007. http://is.muni.cz/th/139654/fi_b/.
19. **Kratochvíl, Jiří.** *Praktická metodika hry na klarinet 2.* Praha : Hudební fakulta Akademie múzických umění v Praze, 1996. ISBN 80-85883-15-5.
20. **Franěk, Marek.** *Hudební psychologie.* Praha : Nakladatelství Karolinum, 2007. ISBN 978-80-246-0965-2.
21. **Emmerson, Simon.** *Living Electronic Music.* Aldeshot : Ashgate Publishing Limited, 2007. ISBN 978-0-7546-5546-6.
22. **Budíková, Marie.** *Statistika.* Brno : Masarykova univerzita v Brně, 2004.
23. **Zvárová, Jana.** *Biomedicínská statistika I.* Praha : Karolinum, 2002. ISBN 80-7184-786-0.
24. **Huber, David Miles.** *The MIDI Manual.* Oxford : Elsevier, 2007. ISBN 978-0-240-80798-0.
25. **Truesdell, Cliff.** *Mastering Digital Audio Production.* Indianapolis : Wiley Publishing, Inc., 2007. ISBN 978-0-470-10259-6.
26. **Fritz, Ian.** Wind Control of Modular Subtractive Synthesis. *The Electronic Sound-House.* [Online] 1998-2011. [Citace: 17. 5 2001.] http://home.comcast.net/~ijfritz/sy_close.htm.
27. **Mark C. Ely, Amy E. Van Deuren.** *Wind Talk for Woodwinds.* Cary, NC : Oxford University Press, 2009. ISBN 978-0195329254.
28. **Rehfeldt, Phillip.** *New Directions for Clarinet.* Berkley, Los Angeles, London : University of California Press, 1994. ISBN 978-0-520-03379-5.
29. **Šetek, Petr J.** Ladění a intonace. *Petr J. Šetek's Flute Pages.* [Online] 2010. [Citace: 17. 5 2011.] <http://home.c2i.net/petrs/flute/ladeni.htm>.
30. **Hejda, Miloslav.** Pozoun (Některé problémy vývoje, stavby a hry). *Klub trombonistů a tubistů.* [Online] [Citace: 15. 5 2011.] <http://home.tiscali.cz/trombonsfera/vyvoj.html>.
31. **Bartolozzi, Bruno.** *New Sounds for Woodwind.* New York, Toronto : Oxford University Press, 1967. ISBN 978-0-193-18607-1.
32. **Františák, Igor.** *Nové výrazové prostředky moderní techniky hry na klarinet.* Ostrava : Pedagogická fakulta Ostravské univerzity v Ostravě, 2005.
33. **Válek, Jiří.** *Italské hudební názvosloví.* Praha – Litomyšl : Nakladatelství Paseka, 2007. ISBN 978-80-7185-852-2.

Obsah elektronické přílohy

Partitury vybraných skladeb

Formát: Portable Document Format (PDF), Portable Network Graphics (PNG)

Umístění: datové CD, /partitury/

Analýza nahrávek

Formát: Sešit Excel

Umístění: datové CD, /midi_analyza.xlsx

VST3 modul Humanizer

(kompilace pro MS Windows)

Formát: Dynamic Link Library (DLL, zde VST3)

Umístění: datové CD, /aplikace/Humanizer/

Modul Humanizer

(zdrojové soubory a projekt Visual Studio)

Formát: C++, Visual Studio

Umístění: datové CD, /aplikace/zdroje/

Předvolby modulu Humanizer

Formát: VST Preset (vstpreset)

Umístění: datové CD, /aplikace/Humanizer/presets/

Manuál k modulu Humanizer

Formát: Portable Document Format (PDF)

Umístění: datové CD, /aplikace/Humanizer/

Vybrané virtuálně ozvučené etudy

(Ke každé skladbě přísluší čtyři soubory: quantized – kvantizovaná, standard – upravená běžnou humanizací, advanced – upravená použitím modulu Humanizer, real – nahrávka produkce doktora Daňhela.)

Formát: MPEG Layer 3 (MP3)

Umístění: datové CD, /ukazky/

Nahrávky produkce doktora Daňhela

Formát: soubory MIDI (MID)

Umístění: datové CD, /nahravky/

Převodní parsery textových výpisů z MIDI souborů

Formát: PHP

Umístění: datové CD, /parsery/