VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA INFORMAČNÍCH TECHNOLOGIÍ
FACULTY OF INFORMATION TECHNOLOGY

ÚSTAV POČÍTAČOVÉ GRAFIKY A MULTIMÉDIÍ
DEPARTMENT OF COMPUTER GRAPHICS AND MULTIMEDIA

ROZPOZNÁVANIE TÓNOV A AKORDOV NA ZARIADENIACH ANDROID

NOTES AND CHORDS RECOGNITION ON ANDROID DEVICES

BAKALÁŘSKÁ PRÁCE
BACHELOR'S THESIS

AUTOR PRÁCE — OLIVER KUCHAROVIČ
AUTHOR

VEDOUCÍ PRÁCE — Ing. SZŐKE IGOR, Ph.D.
SUPERVISOR

BRNO 2017

Vysoké učení technické v Brně - Fakulta informačních technologií

Ústav počítačové grafiky a multimédií Akademický rok 2016/2017

Zadání bakalářské práce

Řešitel: **Kucharovič Oliver**
Obor: Informační technologie
Téma: **Rozpoznávaní tónů a akordů na platformě Android**
Tone and Chord Recognition as an Android Application
Kategorie: Softwarové inženýrství

Pokyny:

1. Nastudujte základy implementace v Android OS. Analyzujte dostupné aplikace pro rozpoznání tónů a akordů.
2. Navrhněte a implementujte aplikaci pro automatický převod audia do notováho zápisu.
3. Aplikaci otestujte na vhodném vzorku beta-testerů.
4. Pokračujte v implementaci a změnách GUI pro co nejlepší UX. Získejte zpětnou vazbu od uživatelů. Zveřejněte aplikaci na Google Play.
5. Zhodnoťte výsledky a navrhněte směry dalšího vývoje.
6. Vyrobte A2 plakátek a cca 30 vteřinové video prezentující výsledky vaší práce.

Literatura:

- Podle pokynů školitele

Podrobné závazné pokyny pro vypracování bakalářské práce naleznete na adrese http://www.fit.vutbr.cz/info/szz/

Technická zpráva bakalářské práce musí obsahovat formulaci cíle, charakteristiku současného stavu, teoretická a odborná východiska řešených problémů a specifikaci etap (20 až 30% celkového rozsahu technické zprávy).

Student odevzdá v jednom výtisku technickou zprávu a v elektronické podobě zdrojový text technické zprávy, úplnou programovou dokumentaci a zdrojové texty programů. Informace v elektronické podobě budou uloženy na standardním nepřepisovatelném paměťovém médiu (CD-R, DVD-R, apod.), které bude vloženo do písemné zprávy tak, aby nemohlo dojít k jeho ztrátě při běžné manipulaci.

Vedoucí: **Szőke Igor, Ing., Ph.D.**, UPGM FIT VUT
Datum zadání: 1. listopadu 2016
Datum odevzdání: 17. května 2017

VYSOKÉ UČENÍ TECHNICKÉ V BRNE
Fakulta informačních technologií
Ústav počítačové grafiky a multimédií
612 66 Brno, Božetěchova 2

doc. Dr. Ing. Jan Černocký
vedoucí ústavu

Abstrakt

Cieľom tejto práce je rozpoznávanie tónov a akordov na ľubovoľnom hudobnom nástroji pomocou mikrofónu v mobilných zariadeniach Android. Aplikácia sníma alebo nahráva zvuky, ktoré analyzuje a zapisuje do zápisu podobnej hudobnej notovej osnove v reálnom čase. Práca popisuje problematiku rozpoznávania a podkladá riešenie danej problematiky pomocou rýchlej Fourierovej transformácie a analýzy spektra hlasitostí. Ďalej práca popisuje hudobnú teóriu a trénovanie jednotlivých tónov. Aplikácia bola podrobená užívateľským testom a testom na presnosť rozpoznávania. Záver obsahuje výsledky z testovania a štatistiky z publikácie na Google Play.

Abstract

The aim of this work is to recognize tones and chords on any musical instrument using the microphone on Android mobile devices. The app scans or records the sounds which are analyzed. It writes to a visual representation which is similar to the musical notation. The work describes the problems of recognition and the solution of the problem using fast Fourier transformation and analysis of the spectrum of volumes. Work describes the theory of music and the training of individual tones. The application was subjected to user tests and recognition accuracy tests. The conclusion includes tests results and statistics from the Google Play publication.

Kľúčové slová

Audio, rozpoznávanie, signál, akordy, tóny, spektrum, vzorkovanie, FFT, frekvencia, WAV, polyfónia, Android, mobilné zariadenia, hudobné nástroje, noty, reálny čas

Keywords

Audio, recognition, signal, chords, tones, spectrum, sampling, FFT, frequency, WAV, polyphony, Android, mobile devices, musical instruments, sheet music, real time

Citácia

KUCHAROVIČ, Oliver. *Rozpoznávanie tónov a akordov na zariadeniach Android.* Brno, 2017. Bakalářská práce. Vysoké učení technické v Brně, Fakulta informačních technologií. Vedoucí práce Szőke Igor.

Rozpoznávanie tónov a akordov na zariadeniach Android

Prehlásenie

Prehlasujem, že som túto bakalársku prácu vypracoval samostatne pod vedením pána Ing. Igora Szökeho, Ph.D. a uviedol som všetky literárne pramene a publikácie, z ktorých som čerpal.

. . . . . .
Oliver Kucharovič
18. mája 2017

Poďakovanie

Chcel by som poďakovať vedúcemu mojej práce, pánovi Ing. Igorovi Szökemu, Ph.D. za vedenie mojej práce, ochotu, venovaný čas a rady, ktoré mi poskytol. Ďalej chcem poďakovať každému, kto mi pomohol aplikáciu otestovať.

Obsah

1 Úvod **3**
1.1 Formulácia problému . . . 3
1.2 Rozvrhnutie práce . . . 4

2 Rozpoznávanie tónov a akordov z audio signálu **5**
2.1 Analýza metód na rozpoznávanie signálu . . . 5

3 Návrh aplikácie **8**
3.1 Analýza podobných aplikácii . . . 8
3.1.1 Testovanie funkcionality aplikácií . . . 8
3.1.2 Testovanie hlavných výhod a nevýhod aplikácií . . . 9
3.2 Problémy návrhu aplikácie . . . 10
3.2.1 Mikrofóny v zariadeniach . . . 10
3.2.2 Problémy snímania hlbokých tónov . . . 11
3.2.3 Detekcia rušivých zvukov okolia . . . 11
3.3 Cieľ a výhody návrhu aplikácie . . . 11

4 Prevod audio signálu na akordy **12**
4.1 Rýchla Fourierová transformácia . . . 12
4.2 Spektrum hlasitostí . . . 13
4.3 Hudobná teória . . . 14
4.3.1 Intervali . . . 14
4.3.2 Hudobné stupnice . . . 15
4.3.3 Akordy . . . 15
4.3.4 Akordické značky . . . 16
4.3.5 Enharmonická zámena . . . 16

5 Implementácia **17**
5.1 Snímanie audio signálu v reálnom čase . . . 17
5.1.1 Nastavenie nahrávania na zariadení . . . 17
5.1.2 Snímanie audio signálu . . . 18
5.1.3 Aplikácia Rýchlej Fourierovej transformácie . . . 18
5.1.4 Získanie spektra hlasitostí . . . 19
5.1.5 Algoritmus na nájdenie lokálnych maxím . . . 19
5.1.6 Získanie zahraných tónov . . . 20
5.1.7 krivka priepustnosti tónov . . . 21
5.1.8 Získanie a analýza dôležitých tónov . . . 21
5.1.9 Algoritmus určovanie akordov . . . 22

5.2 Trénovanie rozpoznávania tónov . . . 23
5.3 Detekcia šumu a hluku v okolí . . . 23
5.4 Reprezentácia audio signálu . . . 23
5.4.1 Audio reprezentácia . . . 24
5.4.2 Reprezentácia informácií po štarte aplikácie . . . 24
5.4.3 Reprezentácia počas nahrávania . . . 25
5.5 Užívateľské rozhranie . . . 26
5.5.1 Hlavná obrazovka aplikácie . . . 26
5.5.2 Snímanie hluku miestnosti . . . 27
5.5.3 Obrazovka nahrávania zvuku v reálnom čase . . . 27
5.5.4 Obrazovka prehrávanie hudobných ideí . . . 28
5.5.5 Výber uloženého hudobného záznamu . . . 29

6 Testovanie aplikácie **30**
6.1 Aktívne testovanie počas vývoja aplikácie . . . 30
6.2 Komplexné testovanie presnosti rozpoznávania . . . 30
6.2.1 Testovanie monofónneho audio záznamu . . . 31
6.2.2 Testovanie polyfónneho audio záznamu . . . 31
6.2.3 Testovanie aplikácie na hudobných skladbách . . . 32
6.2.4 Zhodnotenie výsledkov testovania . . . 33
6.3 Zhodnotenie užívateľských testov . . . 33
6.4 Zhodnotenie výsledkov z Google Play . . . 34

7 Záver **35**
7.1 Zhrnutie výsledkov . . . 35
7.2 Možnosti ďalšieho vývoja . . . 35

Prílohy **37**

A Snímky obrazovky aplikácie a obrázky **38**

B Obsah CD **41**

Kapitola 1

Úvod

Bakalárska práca sa zaoberá rozpoznávaním tónov a akordov pomocou zabudovaného mikrofónu v mobilných a tabletových zariadeniach, ktoré podporujú platformu Android. Aplikácia pracuje z hardwarovým mikrofónom v zariadení a sníma alebo nahráva zvuky z okolia, ktoré analyzuje a zapisuje do zápisu podobnej hudobnej notovej osnove. Aplikácia rozpoznáva tóny a akordy z hudobných nástrojov alebo zo spevu.
Hlavný cieľ práce je vytvorenie záznamu a vizuálnej reprezentácie pre nahranú ideu hudobníka, ku ktorému bude mať kedykoľvek prístup. Ďalej práca obsahuje jednoduché, prehľadné a zaujímavé užívateľské rozhranie, ktoré umožňuje danému hudobníkovi rýchlo a jednoducho nahrať hudobnú ideu. Práca je navrhnutá ako moderná rýchla mobilná aplikácia, ktorá sníma a rozpoznáva tóny a akordy v reálnom čase.

1.1 Formulácia problému

Získanie presných tónov a akordov zo zvuku nasnímaného cez mikrofón je zložitý a náročný proces. V dnešnej dobe tento problém stále nie je uzavretý, a preto vznikajú nové spôsoby a metódy. Práca funguje ako mobilná aplikácia, čo v dnešnej dobe má výhody dostupnosti či rozmanitosti a nevýhody malého menej kvalitného hardwaru a nedostatku výpočetnej sily oproti desktopovým aplikáciám. Práve kvôli dostupnosti je práca navrhnutá na mobilnú platformu Android.
Android je mobilná platforma na báze programovacieho jazyka JAVA. Takéto zariadenia sú momentálne jedny z najrozšírenejších na trhu a sú vhodné na prácu s multimédiami. Moderné mobilné zariadenia obsahujú celkom kvalitný hardware a schopnú výpočetnú silu, pomocou ktorej vznikajú stále novšie a výkonnejšie aplikácie nielen na prácu s multimédiami.
Bakalárska práca sa zaoberá so zvukom, ktorý je snímaný pomocou integrovaných mikrofónov, ktoré môže viesť k snímaniu okrem dôležitých informácií aj k snímaniu hluku a šumu z okolia. Tieto problémy sa nedajú dokonale odstrániť, pretože sa nejedná o káblové spojenie, kde sú zvukové údaje izolované od okolia. Taktiež mikrofóny mobilných zariadení sa nemôžu vyrovnať kvalitných mikrofónom určeným na nahrávanie hudby. Ale pre rýchle menej kvalitné nahrávanie, ktorého výsledky máme vlastne neustále pri sebe sú mobilné zariadenia vhodnou voľbou.
Zaznamenávaný zvuk sa zaznamenával z mikrofónu zariadenia, ktoré sa vhodne nakonfigurovalo. A požila sa metóda Rýchlej Fourierovej transformácie, ktorá je vhodná na prácu so signálmi v reálnom čase.

Na testovanie boli použité reálne hudobné nástroje a skladby prehrávané z kvalitného hudobného prehrávača. Ku skladbám existuje notový zápis, podľa ktorého sa rozpoznávanie testovalo. Reálne hrané etudy na hudobných nástrojoch boli hrané veľakrát, v rôznych oktávach a taktiež k nim existuje zápis na testovanie.

1.2 Rozvrhnutie práce

V kapitole 2 sú popísané rôzne metódy rozpoznávania tónov a akordov. Návrh aplikácie aj s analýzou podobných aplikácií obsahuje kapitola 3. Kapitola 4 sa venuje implementácií algoritmu, reprezentácie a užívateľského prostredia. Testovanie aplikácie je v kapitole 5 a na koniec v kapitole 6 je záver, kde sa nachádza zhrnutie výsledkov a možnosti ďalšieho vývoja.

Kapitola 2

Rozpoznávanie tónov a akordov z audio signálu

Akord je súzvuk minimálne troch hudobných tónov v jednom čase. Za hudobný tón je v akustike považovaný každý zvuk stálej frekvencie. Tento zvuk alebo signál je nutné zaznamenávať, aby sa s daným signálom dalo ďalej pracovať. Pomocou zaznamenania hudobného signálu sa v dnešnej dobe dokážu rozpoznávať jednotlivé zložky nasnímaného signálu a neskôr rozpoznať zahrané hudobné tóny a akordy v danom čase. Momentálne problematika rozpoznávania tónov a akordov nie je uzavretá a existuje mnoho spôsobov a metód ako sa o to pokúsiť.

2.1 Analýza metód na rozpoznávanie signálu

Množstvo spôsobov ako riešiť danú problematiku odlišujú rôzne parametre, ktoré sú pri každom návrhu systému rozličné. Medzi takéto parametre sa udáva napríklad výpočetná náročnosť, priemerná presnosť rozpoznaných tónov a akordov, čas trvania rozpoznávania atď. Tieto parametre súvisia s presnou špecifikáciou, ktorá odpovedá každému systému inak. Špecifikácia môže, ku príkladu obsahovať vlastnosti stroja alebo zariadenia, na ktorom bude metóda rozpoznávania implementovaná, ďalším príkladom môže byť použiteľnosť v reálnom svete, komu je aplikácia určená, aký čas výpočtu je prijateľný atď.

Pravdivostný model Probalistic Latent Component Analysis

Analýza sígnálu pomocou pravdivostného modelu Probalistic Latent Component Analysis [1] je moderná metóda, ktorá je perspektívna, pretože je úspešne schopná rozpoznávania a stále sa rozširuje. Metóda rozpoznáva tóny a akordy po vytvorení hudobného záznamu a je schopná celé spektrum rozložiť na samostatné zdroje zvuku. Pre rozpoznávanie tónov je nutné natrénovať hodnoty. Presnosť tejto metódy je približne 80 %, rozpoznávanie neprebieha v reálnom čase.
Výhody metódy sú v presnosti rozpoznávania tónov a akordov, množstva literatúry a metóda sa neustále vedecky rozširuje. Nevýhody sú, že rozpoznávanie neprebieha v reálnom čase, čas výpočtu je pomalší a nutnosti trénovania hodnôt pre rozpoznávanie.

Porovnávanie signálu s referenčnými dátami

Jedná sa o metódu [2], ktorá rozoberie signál na zložky a pracuje s týmito zložkami. Zložky sa môžu rozoberať, napríklad pomocou rýchlej Fourierovej transformácie. Dané zložky sa podrobia skalárnemu súčinu s referenčnými akordmi. Súčin sa ďalej normalizuje podľa počtu tónov. Normalizovaný súčin sa ďalej porovnáva a výsledkom sú reálne zahrané akordy. Daný spôsob je vhodný pre jeden vybraný hudobný nástroj, podľa ktorého sa zvolili referenčné akordy, rozpoznávanie môže prebiehať v reálnom čase.
Výhody danej metódy spočívajú v rýchlosti rozpoznávania akordov, menej náročnej implementácii a možnosti pracovania v reálnom čase. Medzi nevýhody patrí, zameranie iba na jeden špecifický hudobný nástroj a nutnosť dodať referenčné akordy.

Rozpoznávanie tónov pomocou neurónových sietí

Táto metóda sa úspešne využíva pri rozpoznávaní reči. Neurónové siete [3] sú náročnejšie na implementáciu, výpočetne náročné, nutnosť trénovania zložiek. Ich presnosť rozpoznávania tónov je vysoká, ale závisí od počtu zavedených neurónov a počtu vstupov. Metóda sa musí učiť od určitej dátovej sady, jedná sa o štatistickú metódu.
Výhody metódy sú v presnosti rozpoznávania tónov a akordov, možnosť zvolenia počtu neurónov, použiteľnosti v rôznych smeroch rozpoznávania a množstva literatúry. Nevýhody metódy sú v nutnosti učiť algoritmus od určitej dátovej sady, zložitejšia implementácia a algoritmus je výpočetne náročný.

Rozpoznávanie akordov podľa analýzy spektra hlasitostí

Jedná sa o metódu, ktorá snímaný audio signál dekomponuje na zložky pomocou rýchlej Fourierovej transformácie [5] [7], ktorá sa neskôr normalizuje. Normalizovaný signál odpovedá spektru hlasitostí v jednom čase nasnímania. Pomocou analýzy tohto spektra a frekvenčne závislej krivky sa získajú zahrané tóny. Metóda je menej náročná na implementáciu, výpočetne sa jedná o menej náročný algoritmus. Rozpoznávanie môže prebiehať v reálnom čase. Presnosť danej metódy nie je tak vysoká ako presnosť pravdivostného modelu.
Výhody metódy sú v rýchlosti rozpoznávania akordov a tónov, možnosť rozpoznávania v reálnom čase a nízkej výpočetnej náročnosti. Nevýhody danej metódy sú v nižšej presnosti rozpoznávania a, že je potreba trénovať hodnoty.
Metódu rozpoznávania akordov podľa analýzy spektra hlasitostí som použil v práci. Dôvody sú nasledovné:

1. Nízka výpočetná náročnosť
2. Možnosť rozpoznávania v reálnom čase
3. Rýchlosť algoritmu
4. Využitie na ľubovoľný hudobný nástroj
5. Možnosť prispôsobenia algoritmu pre zariadenia Android
6. Dostupnosť vhodných knižníc pre danú metódu

Zariadenia Android nemajú taký výkon, ako napríklad stolný počítač, preto bolo nutné zvoliť metódu, ktorá je výpočetne nenáročná. Jedným z parametrov návrhu aplikácie je rýchle rozpoznávanie v reálnom čase (bližšie popísané v kapitole 3), ktoré metóda ponúka. Mobilná aplikácia by mala byť dostupná pre širokú verejnosť, preto funkčnosť rozpoznávania na ľubovoľnom hudobnom nástroji je veľmi dôležitá.

Kapitola 3

Návrh aplikácie

Táto kapitola analyzuje a vyhodnocuje podobné aplikácie, ukáže na problémy pri návrhu aplikácie a objasní celkový návrh.
Aby bolo vôbec možné navrhnúť aplikáciu, ktorá bude niečím zaujímavá a pre užívateľa lákavá je nutné preskúmať podobné aplikácie, ktoré sa zaoberajú podobným problémom. Cieľom bolo aplikácie vyskúšať a otestovať podľa vhodných kritérií, ktoré do vývoja objasnili a vytýčili ciele, na ktoré sa aplikácia špecifikovala.
Ďalej sú v tejto kapitole objasnené problémy, ktoré pri návrhu aplikácie nastali. Tieto problémy súvisia buď s hardwarom v mobilných zariadeniach, alebo sa vyskytujú všeobecne pri tvorbe audio záznamu.
V sekcii 3.3 je objasnený návrh aplikácie, jej výhody a nevýhody a pre koho je práca určená.

3.1 Analýza podobných aplikácii

Cieľom bolo nájsť podobné dostupné aplikácie na Google Play, ktoré analyzujú alebo rozpoznávajú prijímaný audio signál. Na aplikáciách sa sledovalo ich užívateľské rozhranie, funkcionalita, presnosť, rýchlosť a výhody či nevýhody. Tieto vysledované údaje pomohli vytýčiť presnejšie ciele pre výsledný produkt a boli inšpiráciou k vytvoreniu aplikácie, ktorá bude obsahovať čo najviac výhod.
Vhodný výsledok z analýzy podobných aplikácii bude aplikácia ktorá nebude kópiou inej, už dostupnej aplikácie na Google Play, ale aby bola niečím iná a pre užívateľa bola zaujímavá. Vybrali sa vhodné kritéria na testovanie, podľa ktorých sa aplikácie budú hodnotiť.

3.1.1 Testovanie funkcionality aplikácií

Vybrané aplikácie boli testované a sledovala sa ich funkcionalita, ktorá je kľúčová k docieleniu vhodného návrhu aplikácie. Medzi testované parametre funkcionality patria:

1. Podpora polyfónie – snímanie viac tónov v jednom čase
2. Možnosť nahrávania – zaznamenávanie audio signálu do ľubovoľnej reprezentácie dát, jedná sa o zvukovú a písomnú reprezentáciu a uloženie danej reprezentácie pre neskoršie použitie
3. Detekcia v reálnom čase – analyzovanie a reprezentovanie audio signálu do písomnej reprezentácie v reálnom čase

4. Vkladanie audio súboru zo zariadenia – možnosť vkladať audio súbory na analýzu do aplikácie

Zaznamenané údaje sú reprezentované v tabuľke 3.1.

Názov	Podpora polyfónie	Možnosť nahrávania	Detekcia v reálnom čase	Vkladanie audio súboru zo zariadenia
AnySong Chord Recognition	-	-	Nie	Áno
Chord detector	Áno	Nie	Áno	Nie
hum on!	Nie	Áno	Áno	Nie
Chord detector ZAX	Áno	Nie	Áno	Nie
Mystage.fm note recognition	Nie	Áno	Nie	Nie
MusicTrancription lite	Áno	Nie	Nie	Áno
Guitar music analyzer	Áno	Nie	Nie	Nie
Music Translator (naša)	Áno	Áno	Áno	Nie

Tabuľka 3.1: Funkcionalita aplikácií

Z tabuľky 3.1 vyplýva, že z uvedených 7 aplikácii podporuje polyfóniu 5, možnosť vytvárania záznamu dokážu 2, podporu detekcie v reálnom čase dokážu 3 a vkladať súbor do aplikácie vedia iba 2. Dôležitým faktorom ale je, že z uvedených aplikácii ani jedna nepodporuje prvé tri kritéria súčasne, čiže podporu polyfónie, nahrávanie a detekcia v reálnom čase. Dokonca ani jedna z vybraných nepodporuje súčasne prvé dva parametre testovania. Cieľom bolo, aby aplikácia podporovala polyfóniu a monofóniu, aby bolo možné vytvárať záznam, reprezentovať a znovu otvoriť a aby aplikácia dokázala vykonávať detekciu v reálnom čase.

3.1.2 Testovanie hlavných výhod a nevýhod aplikácií

V tomto testovaní bolo hlavným cieľom vyhodnotiť hlavné výhody a nevýhody testovaných aplikácií. Tieto parametre boli testované z pohľadu užívateľa. Použili sa užívateľské testy ako:

1. Rýchlosť uskutočnenia analýzy audio záznamu od štartu aplikácie
2. Presnosť analýzy audio záznamu na jednoduchý audio záznam
3. Podpora vkladania audio formátov
4. Koľko funkcií aplikácia obsahuje
5. Prehľadnosť a čitateľnosť užívateľského rozhrania
6. Všeobecná použiteľnosť aplikácie
7. Čitateľnosť písomnej reprezentácie audio dát
8. Množstvo informácii z audio záznamu
9. Platená verzia

Zaznamenané údaje sú reprezentované v tabuľke 3.2.

Názov	Hlavné výhody	Hlavné nevýhody
AnySong Chord Recognition	-	Málo podporovaných formátov
Chord detector	Presný, rýchly	Málo funkcií
hum on!	Prehľadné rozhranie	Málo funkcií
Chord detector ZAX	Množstvo informácií	Nefunkčná, Nepresná
Mystage.fm note recognition	-	Pomalá detekcia, Nečitateľné dáta
MusicTrancription lite	Množsvo funkcií	Bez akordov, Platená verzia
Guitar music analyzer	Rýchle	Nepresná, Málo informácií

Tabuľka 3.2: Hlavné výhody a nevýhody aplikácií

Z tabuľky 3.2 vyplýva, že každá testovaná aplikácia má nevýhody, ktoré sú nežiaduce v návrhu aplikácie. Tieto nevýhody môžu užívateľa odradiť od používania danej aplikácie. Naopak niekoľko aplikácií obsahovalo aj výhody, s ktorými sa práca inšpirovala pri vývoji. Snahou bolo výhody testovaných aplikácii nekopírovať, ale vytvoriť aplikáciu, ktorá bude konzistentná a držať sa istej idei.

3.2 Problémy návrhu aplikácie

Rozhodnutím, že aplikácia bude na mobilných zariadeniach Android [10] je nutnosťou rátať s problémami, ktoré sa vyskytnú. Jeden z týchto problémov je, že sa musia používať integrované mikrofóny v hardware Android zariadení. Ďalší takýto problém je vo výpočetnej sile týchto zariadení. Android zariadení je veľmi veľa, líšia sa vo výrobcoch, cene a roku výroby. Ďalšie problémy návrhu súvisia so všeobecnými problémami, ktoré sa vyskytujú pri záznamu audio dát cez mikrofón, sú to napríklad problémy nežiaduceho šumu v miestnosti alebo problémy snímania vyšších harmonických tónov[1].
Predpokladom je, že aplikácia bude pracovať lepšie na novších a kvalitnejších zariadeniach. Problémy budú menšie, nahrávanie kvalitnejšie a detekcia presnejšia vďaka výkonnejšiemu hardwaru a novšiemu, rýchlejšiemu softwaru.

3.2.1 Mikrofóny v zariadeniach

Kvalitné štúdiové mikrofóny [12] majú frekvenčný rozsah približne od 20 Hz do 20 kHz, citlivosť mikrofónu sa pohybuje asi od 115 do 132 db/mW a obsahujú rôzne filtre a funkcie proti skresleniu. Ich cena sa pohybuje od 200 – 1000 eur. Dostupnejšie a bežnejšie dynamické mikrofóny majú frekvenčný rozsah menší asi od 50 Hz do 15 kHz a nižšiu citlivosť. Cena kvalitného dynamického mikrofónu začína asi od 50 eur. Nie všetky dynamické mikrofóny sú ale určené na nahrávanie hudobného nástroja. Viac informácii ohľadom štúdiových mikrofónov je na internetovej stránke výrobcov[2].
Na porovnanie, mikrofóny v zariadeniach Android závisia od výrobcu. Ich veľkosť je menšia a nie sú určené na prácu s hudobnými nástrojmi či so spevom. Ich frekvenčný rozsah je menší a v lepšom prípade začína od 100 Hz, čo je päťkrát horšie oproti štúdiovým mikrofónom určeným na nahrávanie hudobných nástrojov. Bohužiaľ je veľmi náročné zistiť presné

[1] Jedná sa o problém, kedy popri základnom tóne znie spoločne ďalší čiastkový tón (niekedy sa tiež nazýva tento tón alikvótny)

[2] Pre príklad uvádzame stránku výrobcu mikrofónov Shure http://www.shure.eu/products/microphones

hodnoty, kvôli veľkému počtu zariadení. Samozrejme veľký dôraz sa kladie na cenu zariadenia a predpokladom je, že novšie zariadenia budú obsahovať stále kvalitnejší hardware a lepší mikrofón.

3.2.2 Problémy snímania hlbokých tónov

Z toho vyplýva, že spodný frekvenčný rozsah mikrofónov v Android zariadeniach je približne 100 Hz a najhlbší tón na štandardne ladenej basgitare je *E1*, ktorý ma frekvenciu približne 40 Hz. Ďalej mužský nízky spev môže začínať asi od 50 Hz. Takže mikrofón s rozsahom od 100 Hz nebude frekvencie pod hranicou snímať a môže nasnímať vyšší harmonický tón. V tomto prípade je ideálna možnosť nasnímania vyššej oktávy tónu z vlnenia, a to pri 50 Hz je oktáva 100 Hz, ktorú by už mohol mikrofón zachytiť, táto možnosť sa nemusí podariť v každom prípade. Toto vlnenie teda môže zapríčiniť nepresnosti v hlbokých frekvenciách, napríklad nasnímaním iného tónu ako oktávy.
Tento problém sa nedá riešiť jednoducho, momentálne návrh aplikácie závisí od kvality použitého mikrofónu v zariadeniach.

3.2.3 Detekcia rušivých zvukov okolia

Jedným z najväčších problémov nahrávania audio dát cez mikrofón je snímanie spolu s dôležitými dátami aj dáta, ktoré sú nepotrebné, šum z miestnosti alebo okolia kde prebieha nahrávanie a nežiaduce zvuky. Nežiaduce zvuky musí používateľ znížiť na minimum, aby efektivita nahrávania bola čo najvyššia. Medzi šum okolia sa môžu rátať, napríklad šumivý zvuk zo zosilňovača, šum zo snímačov u hudobných nástrojov, ktoré sa produkujú do miestnosti a hudobník ich nechce mať na hudobnom zázname.
Kvalitné mikrofóny dokážu tieto problémy znížiť na minimum. Android zariadenia nemajú k dispozícii takéto príslušenstvo, takže je ideálne tento ruch a šum nezahrnúť do výslednej detekcie audio signálu a počítať s ním pri návrhu aplikácie a tak urobiť rozpoznávanie tónov a akordov čo najpresnejšie a najrýchlejšie.

3.3 Cieľ a výhody návrhu aplikácie

Cieľom aplikácie a jej výhody návrhu sú v rýchlom a jednoduchom nahrávaní v reálnom čase a opakovanom prehrávaní hudobných ideí hudobníka, ktoré mu analyzuje a rozpozná program do nôt a akordov a zapíše do zápisu. Zápis a zvuk sa dá uložiť do pamäte zariadenia a hudobný záznam spolu so zápisom dokáže hudobník hocikedy otvoriť. V ideálnom prípade hudobník nemusí mať so sebou vlastnoručne napísaný hudobný nápad, ale postačí mu táto mobilná aplikácia. Tieto uložené nápady má užívateľ neustále k dispozícii vo svojom mobilnom zariadení.
Na druhej strane návrh aplikácie obsahuje aj vyššie zmienené nevýhody, ktoré sa snaží zredukovať na minimum.
Návrh vychádza z podobných aplikácii, ktoré sú dostupné. Cieľom je navrhnúť aplikáciu, ktorej podobná obdoba nie je ešte na Google Play, tým pádom dielo bude špecificky určené pre presnú klientelu všetkých hudobníkov a užívateľov, ktorí majú záujem o hudbu a chcú komponovať vlastné nápady alebo sa vzdelávať v hudbe a aplikácia im môže uľahčiť prácu a prepis hudobného diela či hudobného nápadu do čitateľnej textovej reprezentácie.

Kapitola 4

Prevod audio signálu na akordy

Najdôležitejšou časťou aplikácie je proces rozpoznávania akordov a tónov z prijímaného audio signálu cez vstavaný mikrofón v zariadení Android [10]. Jedná sa o výpočetne komplikovanejší proces, ktorý je ešte komplikovanejší kvôli menšiemu výkonu procesorov v použitých zariadeniach, zmienených v návrhu aplikácie. Cieľom je, aby algoritmus zaťažoval mobilné zariadenia čo najmenej, to znamená, aby sa zariadenia nezasekávali v procese výpočtu, aby sa výdrž batérie čo najmenej znižovala a aby sa zariadenie dalo bežne používať popri používaní aplikácie, napríklad na prijímanie hovorov, prijímanie textových správ atď.
Tento proces prevodu sa dá vykonávať viacerými spôsobmi. Práca dbá na rýchlosť a presnosť rozpoznávania. Ďalší parameter je rozpoznávanie akordov a tónov v reálnom čase. V tomto prípade sa javí ako vhodný algoritmus rýchla Fourierová transformácia, ktorá je vhodná na prácu s diskrétnym signálom v reálnom čase.
Táto kapitola obsahuje popis signálov rýchlej Fourierovej transformácie, rozbor spektra hlasitostí, rozbor zvuku a dôležitú hudobnú teóriu pre rozpoznanie tónov a akordov.

4.1 Rýchla Fourierová transformácia

Rýchla Fourierová transformácia (ďalej FFT[1]) je rýchla verzia diskrétnej Fourierovej transformácia (ďalej DFT). DFT [8] teda prevádza diskrétny signál v časovej oblasti do oblasti frekvenčnej. Transformácia sa uplatňuje v spracovávaniu vstupného signálu v reálnom čase. Dáta z nahrávania máme v podobe číslicového signálu, takže pre N vzorkou stačí počítať prvých *N/2* hodnôt dvojstranného spektra, a tie potom pre získanie jednostranného spektra vynásobiť dvoma, pretože ostatných *N/2* hodnôt sú veličiny k nim komplexne združené a amplitúdy spektra sú dané moduly týchto hodnôt.
FFT je spôsob výpočtu Fourierovej transformácie, ktorým sa získajú rovnaké výsledky, ale rýchlejšie. DFT potrebuje *0(N*N)* operácií, ale FFT iba *0(N log N)*. Najznámejší FFT algoritmus výpočtu predstavili v roku 1965 J. W. Cooley a J. W. Tukey [4]. V súčasnej dobe sa dá FFT používať pre ľubovoľný počet prvkov. Vzorkovací signál sa rozdelí na dve vzorky, párne a nepárne vzorky. Potom sa algoritmus dá popísať ako:

$$F_n = \sum_{k=0}^{N-1} f_k \cdot e^{-j\frac{2\pi}{N}k_n}, n = 0, 1, ..., N-1 \tag{4.1}$$

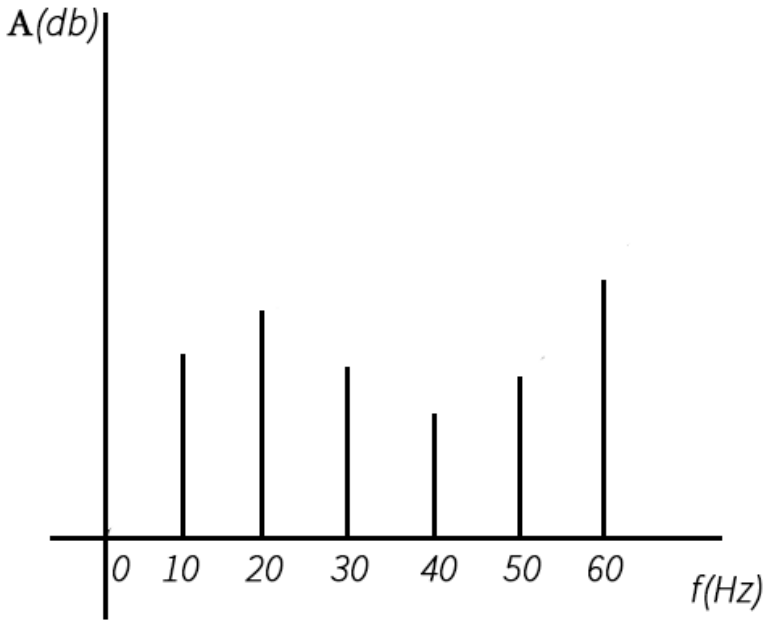
[1] FFT z anglického jazyka *Fast Fourier transformation*

4.2 Spektrum hlasitostí

Spektrum hlasitostí je vlastne 2D zobrazenie hodnôt hlasitostí v diskrétnom čase. Výsledkom FFT je pole imaginárnych a reálnych čísiel, ktoré reprezentujú prvky, ktoré boli nasnímané. Na získanie diskrétneho spektra hlasitostí nasnímaných frekvencií sa musel použiť vzorec na normalizáciu FFT, ktorý je:

$$S = 10 \cdot log_{10}(realPart^2 + imaginaryPart^2) \tag{4.2}$$

Zo vzorca 4.2 sa dá vyvodiť, že každú imaginárnu a reálnu zložku vzorec prevádza na hodnoty, ktoré sú schopné porovnávania, a tak je možné zisťovať ďalšie informácie o nasnímanom signáli, ako napríklad zistenia najväčšieho indexu v danom poli.



Obr. 4.1: Spektrum hlasitostí

Spektrum hlasitostí je známe a obsahuje dôležité údaje. Z obrázku 4.1 možno vidieť, že na ose x sú znázornené indexy v diskrétnom čase, ktoré reprezentujú frekvenciu a na ose y sú decibely alebo iné váhové hodnoty, ktoré určujú silu danej frekvenciu v daný čas.
Zo spektra hlasitostí nie je problém získať najhlasnejšiu frekvenciu. Táto frekvencia sa zaradila ako najsilnejšiu frekvenciu a podľa nej sa dokázali zistiť aj ďalšie zahrané tóny v rovnaký čas.

4.3 Hudobná teória

Hudobná teória [6] [13] je nevyhnutná na vytváranie akordov zo snímaného audio signálu. Akordy sa určovali z nameraných tónov. Určenie akordu vychádzalo s najhlbšieho nasnímaného tónu, teda hlasu bas. Podľa basu sa ďalej určovali ostatné akordické tóny pomocou intervalov.

4.3.1 Intervali

Interval v hudobnej teórií znamená vzdialenosť medzi dvoma tónmi. V práci sa jednalo o harmonické zisťovanie intervalov, pretože tóny boli znejúce v rovnakom čase. Najmenší interval v európskej hudbe je poltón napr. tóny C a $C\#$.
Názvy intervalov sú odvodené od toho, koľko stupňov stupnice sú jednotlivé tóny od seba vzdialené. Názvy sú číslovky z latinského jazyka. Vzostupne sú to príma, sekunda, tercia, kvarta, kvinta, sexta, septima a oktáva. Príma a oktáva sú rovnaké tóny napr. C a C, ktoré sú vzdialené o jednu oktávu. Ďalej bolo nutné pracovať z 12 intervalmi a nie iba z ôsmimi, ktoré sme vymenovali vyššie, preto sa pridávajú ďalšie, bližšie určené intervaly a to sú napríklad malá a veľká tercia. Tabuľka 4.1 pre prehľadnosť intervalov. Pomocou intervalov sa mohol určiť hraný akord v danom čase.

Stupeň	Meno intervalu	Počet poltónov	Zákl. tón: C	Pomer frek.
1	čistá príma	0	C - C	1:1
2	malá sekunda	1	C - Db	16:15
2	veľká sekunda	2	C - D	9:8
3	malá tercia	3	C - Eb	6:5
3	veľká tercia	4	C - E	5:4
4	čistá kvarta	5	C - F	4:3
5	čistá kvinta	7	C - G	3:2
6	malá sexta	8	C - Ab	8:5
6	veľká sexta	9	C - A	5:3
7	malá septima	10	C - Bb	16:9
7	veľká septima	11	C - H	15:8
8	čistá oktáva	12	C - C	2:1

Tabuľka 4.1: Prehľad intervalov

Pri zaznamenávaní akordov sa používajú aj vyššie stupne ako napr. nóna, decíma atď. Nóna je vlastne sekunda od základného tónu, ktorá je o jednu oktávu vyššie. Toto značenie sa využíva pri znázorňovaní akordov (nepíše sa *C2* ale *C9*).

4.3.2 Hudobné stupnice

Stupnica je rad tónov usporiadaných podľa určitých pravidiel v rozmedzí jednej oktávy. Pravidlá sa týkajú podľa počtu tónov v jednej oktáve a vzdialenosti medzi nimi. Najmenší počet tónov, ktoré môže stupnica obsahovať je 5, vtedy sa jedná o pentatonickú stupnicu (penta ako päť a tonika ako základný tón). Tóny v stupnici sú vzdialené buď poltón, celý tón alebo sú vzdialené viac ako celý tón.
Stupnica sa určuje od základného tónu, to znamená, že je nutné určiť jej ďalšie stupne a tenzie od toniky, teda základného tónu. Stupnica *C dur* obsahuje tóny *C, D, E, F, G, A, H, C*. Z toho sa dá ľahko určiť, aké intervaly stupnica obsahuje od základného tónu. Napríklad medzi tónom *C* a *D* sú 2 poltóny alebo 1 celý tón, takže stupnica obsahuje veľkú sekundu. Medzi tónom *C* a *E* sú 4 poltóny, teda stupnica má veľkú terciu. Takto sa získavajú potrebné stupne.

4.3.3 Akordy

Akord je súzvuk najmenej 3 tónov. Podľa počtu tónov, ktoré akord obsahuje delíme akordy na trojzvuky, štvorzvuky alebo viaczvuky. Medzi základné akordy patrí kvintakord, ktorý sa skladá z dvoch tercií, napr. tóny *C, E* a *G*. Tón *C* má veľkú terciu tón *E* a tón *E* má veľkú terciu *G*, medzi tónmi *C* a *G* je interval kvinta. Jedná sa o durový kvintakord, pretože akord obsahuje od basu C interval veľkú terciu a interval čistú kvintu, jedná sa o súzvuk 1.,3. a 5. stupňa. Týmto spôsobom sa dajú určiť kvintakordy, ako durový, molový, zmenšený a zväčšený kvintakord.
Kvintakord je možné hrať aj v jeho obratoch, to znamená, že stupne budú vymenené. Ak je spodným tónom tercia, ide o prvý obrat, sextakord a ako je spodným tónom kvinta, ide o druhý obrat, kvartsextakord.
Pri štvorzvukoch sa pridáva do akordy siedmy stupeň, septima, jedná sa tak o septakord. Pri tónoch *C, E, G* a *H* je skladba, tón *C* má veľkú terciu tón *E*, tón *E* má veľkú terciu *G* a tón *G* má veľkú terciu *H*, medzi tónmi *C* a *G* je interval kvinta a medzi tónmi *C* a *H* je interval veľká septima, z toho vyplýva durový septakord. Pri pridaní siedmeho stupňa sa vyskytuje viac kombinácií ako u kvitakordov. Akordy sa v práci zapisujú podľa akordických značiek.
Septakordy je možné hrať v obratoch podobne ako kvintakordy, ich teoretický zápis je taktiež veľmi podobný. Základný prehľad akordov je popísaný v tabuľke 4.2. Pri viac zvukoch je vhodné dopisovať ďalšie stupne k akordickej značke. Keď daný súzvuk tónov bude obsahovať napríklad malú sextu, tak sa k akordickej značke dopíše b6, čo presnejšie značí hudobníkom, ktoré stupnice sú vhodné použiť na daný akord.

Akord	Intervaly	Značenie akordu	Tóny
Dur	1 3 5	C	C E G
Mol	1 b3 5	Cm, C-, Cmin	C Eb G
Zväčšený	1 3 #5	Caug, C+	C E G#
Zmenšený	1 b3 b5	Cdim	C Eb Gb
Dur7	1 3 5 7	Cmaj7	C E G B
Dominantný	1 3 5 b7	C7	C E G Bb
Mol7	1 b3 5 b7	Cm7, C-7	C Eb G Bb
Zmenšený 7	1 b3 b5 bb7	Cdim7	C Eb Gb Bbb

Tabuľka 4.2: Prehľad základných akordov

4.3.4 Akordické značky

V programe a hudobnej teórií sa akordy označujú značkami. Dôvodom je skrátený zápis, ktorý urýchľuje čítanie hudby z listu hudobníkovi pri prednese hudobného zápisu. Akordické značky sú medzinárodne zaužívané, slúžia ako univerzálny dorozumievajúci jazyk pre hudobníkov.
Aplikácia používa značku pri znázornení napr. molového septakordu *D* zápis, *D-7*, kde značka mínus značí, že sa jedná o akord molový a číslo sedem značí, že sa jedná o septakord.

4.3.5 Enharmonická zámena

V hudobnej teórií je možné zapísať tón viacerými spôsobmi, napríklad tón *C#* je možné zapísať *Db*, alebo tón *D* je možné zapísať ako *D* alebo *Cx* alebo *Ebb*. Z toho vyplýva, že *#* nám zvýši hodnotu o poltón, *b* zníži hodnotu o poltón, *x* zvýši hodnotu o celý tón a *bb* zníži hodnotu o celý tón. Enharmonická zámena sa používa na jasnejšie značenie akordov a stupníc.
V aplikácii je použité značenie znižovanie o poltón, teda *b*[2].

[2]Toto značenie vychádza z kvintového a kvartového kruhu, označenie *b* vychádza z kvartového kruhu a bolo prvý krát použíté v hudobnom žánri jazz v Amerike

Kapitola 5

Implementácia

Pri zaznamenávaní zvuku bola daná informácia prenášaná po bitoch s ktorou sa dalo uspokojujúco pracovať. Nad bitovým poľom bola použitá rýchla Fourierová transformácia, ktorá sa normalizovala a výsledkom bolo spektrum hlasitostí nasnímaných frekvencií. Na spektrum bol použitý vlastný algoritmus na nájdenie lokálnych maxím a nakoniec sa toto spektrum analyzovalo a získavalo dôležité údaje. Získala sa frekvenčne závislá, ktorá sa v práci nazýva ako krivka priepustnosti tónov. Nakoniec sa pracovalo s nasnímanými tónmi a určili sa akordy, ktoré sa vkladali do reprezentácii na zobrazenie pre užívateľa.
V tejto kapitole sa rozoberie implementácia algoritmu snímania audio signálu v reálnom čase a prevod tohto signálu na rozpoznané tóny a akordy taktiež v reálnom čase. Sekcia 5.1. V sekcii 5.2 je znázornené ako prebiehalo trénovanie tónov, podľa ktorého sa určovali zahrané tóny. Detekcia šumu a hluku v okolí je popísaná v sekcii 5.3. V kapitole je taktiež rozobraná reprezentácia audio signálu a jeho formáty, sekcia 5.4 a záver kapitoly obsahuje popis užívateľského rozhrania v sekcii 5.5.

5.1 Snímanie audio signálu v reálnom čase

Aplikácia sa snaží o analýzu a prepis snímaného audio signálu v reálnom čase. Po zapnutí aplikácie užívateľ môže hneď sledovať rozpoznávanie tónov a akordov. Je to vhodné na naladenie hudobného nástroja alebo zistenie či hlasitosť jeho hudobného nástroja je dostatočná alebo naopak nie je moc hlasná. V samotnom nahrávaní svojho hudobného nápadu taktiež dokáže sledovať správnosť nahrávania a detekcie v reálnom čase. Reálny čas snímania mu môže uľahčiť a zrýchliť nahrávanie svojich ideí.
Cieľom aplikácie je dosiahnuť rýchlosť a jednoduchosť v používaní aplikácie, toto boli jedny z hlavných parametrov aplikácie.
V tejto sekcii sa vysvetlí daný prevod audio signálu do rozpoznaných tónov a akordov, ktorý sa opakuje vždy pre nové nasnímané audio dáta v reálnom čase.

5.1.1 Nastavenie nahrávania na zariadení

V prvom rade bolo nutné nastaviť vhodný typ nahrávania u Android zariadení [10] [11]. Dôležitým aspektom bol výkon zariadení. V Ideálnom prípade by snímanie audio signálu bolo čo najmenej náročné pre zariadenia, tým pádom rýchlosť nahrávania a detekcia bude rýchlejšia a presnejšia. Pre vhodné nastavenie snímania audio signálu v aplikácii boli zvolené nasledujúce parametre:

- SAMPLERATE – 11250 Hz
- RECORDING CHANNELS – MONO (1)
- AUDIO ENCODING – ENCODING PCM 16 BIT

Vzorkovacia frekvencia

Vzorkovacia frekvencia [12] alebo (sample rate) je počet vzoriek audio signálu za sekundu meraný v Hz. Štandardná kvalita je práve 44100, je to takzvaná kvalita CD. Táto kvalita je preto zaužívaná, pretože obsahuje šírku snímaného frekvenčného pásma (bandwidth), ktorý odpovedá počuteľnosti ľudského ucha, čo predstavuje rozsah až do 22 kHz. V práci sa jedná o menej kvalitné nahrávanie, pretože vybraná vzorkovacia frekvencia je iba 11250 Hz. Dôvody zníženia vzorkovacej frekvencie sú znížené náročnosti nahrávania (úspora času pre užívateľa a úspora zdrojov zariadenia) a vstavané mikrofóny v Android zariadeniach nedosahujú nahrávacie kvality, ktoré by dokázalo zaznamenávať zvuk v tak veľkej šírke pásma

Nahrávací kanál

Vybranie nahrávacieho kanálu [12] (channels) monofónnej zvukovej reprezentácie (ďalej MONO), bolo nutné pre presnejšie rozpoznávanie tónov a akordov, pretože nahrávané bity boli usporiadanejšie ako u stereofónnej zvukovej reprezentácie. Výhody zvolenia MONO sú v jednoduchšom použití, menej zaťažuje zariadenie a výsledný audio súbor je menší. Stereo nahrávanie je síce kvalitnejšie, ale pri nahrávaní zo vstavaného mikrofónu to nie je veľký rozdiel.

Pulzne kódová modulácia

Pulzná kódová modulácia [14] (ďalej PCM) je metóda používaná na digitálne reprezentovanie vzorkovaných analógových signálov. V PCM sa amplitúda analógového signálu pravidelne odoberá v rovnomerných intervaloch a každá vzorka sa kvantuje na najbližšiu hodnotu v rozsahu digitálnych krokov. U práci zvolené PCM 16 bit je vhodné s MONO nahrávaním a dostatočne zachováva kvalitu nahrávania a šetrí zdroje mobilného zariadenia.

5.1.2 Snímanie audio signálu

Po nastavení zariadenia sa dokázalo získavať bitové pole, ktoré obsahovalo surové dáta z audio signálu. Tieto dáta sa získavajú počas celého nahrávania audio signálu v reálnom čase. Dané dáta sa dokážu okamžite analyzovať a zaznamenať do reprezentácie, ktoré závisí podľa výkonu mobilného zariadenia (LG Nexus každých 100 – 300 ms).
Na danú analýzu bitového poľa v reálnom čase sa použila rýchla Fourierova transformácia. Snímanie audio signálu sa pozastaví, ak užívateľ zatvorí alebo odloží aplikáciu na pozadie telefónu. Týmto aplikácia zabraňuje zbytočnému čerpaniu zdrojov mobilného zariadenia.

5.1.3 Aplikácia Rýchlej Fourierovej transformácie

Bitové pole sa podarilo analyzovať v reálnom čase, spôsobom rozdelenia bitového poľa na imaginárnu a reálnu časť. Tým pádom sa vedelo aplikovať FFT na daný komplexný signál [9]. Pomocou FFT sa získal signál obsahujúci transformovanú imaginárnu a reálnu časť. Zo

signálu bolo potrebné získať spektrum hlasitostí, teda absolútny signál. Vzorkovací signál sa rozdelí na dve vzorky, párne a nepárne vzorky. Potom ide algoritmus popísať ako:

$$F_n = \sum_{k=0}^{\frac{N}{2}-1} (e_n \cdot e^{-j\frac{2\pi}{N}2k_n} + o_n \cdot e^{-j\frac{2\pi}{N}2k_n+1}) = \sum_{k=0}^{\frac{N}{2}-1} e_n \cdot e^{-j\frac{2\pi}{N}2k_n} + e^{-j\frac{2\pi}{N}k_n} \sum_{k=0}^{\frac{N}{2}-1} o_n \cdot e^{-j\frac{2\pi}{N}k_n} \quad (5.1)$$

Z výrazu 5.1, že transformácia obsahujúca N bodov sa rozdelí na dve transformácie obsahujúce *N/2* bodov. Rovnakým spôsobom sa bude pokračovať ďalej smerom do ďalších delení, až výsledná postupnosť obsahuje iba 2 prvky. Počet operácií sa znížil z *0(N*N)* na *O(N log N)*.
Tento algoritmus sa vykonával tak rýchlo ako bolo možné, podľa výkonu zariadenia, čo je zmienené vyššie. Práve preto bolo nutné zvoliť metódu, ktorá je dostatočne rýchla. Keďže veľkosť poľa je známy počet vďaka šírke pásma, ktorý vyplýva zo vzorkovacej frekvencie, je daná FFT vhodná metóda. FFT má logaritmickú zložitosť a v práci sa počíta približne iba so 400 vzorkami. Logaritmická zložitosť je oveľa úspornejšia na zdroje v tomto prípade oproti, napríklad exponenciálnej zložitosti, ktorá sa vyskytuje napríklad u algoritmu DFT.

Dĺžka a presah okna

Počet frekvenčných pásiem sa zvolil dĺžkou okna, ktorá je 512. Z tohto čísla sa vo výsledku rátalo iba 399 vzoriek, pretože vzorky nad hodnotu 403 reprezentovali veľmi vysokú frekvenciu, ktorá približne odpovedala tónu *D9*. Vzorky pod hodnotou 4 boli naopak veľmi hlboké, jednalo sa o frekvencie pod 40 Hz. Tieto hlboké frekvencie aplikácia nedokáže rozoznať kvôli nízkej citlivosti mikrofónov v zariadeniach Android.
Presah okna v algoritme nie je implementovaný, pretože sa jedná o rozpoznávanie tónov a akordov v reálnom čase a implementácia presahu okna by výrazne spomalila rozpoznávanie a mohla by spôsobovať chyby. V aplikácii sa snímajú a podliehajú detekcii dáta podľa výkonu zariadenia, priemerne sa jedná o hodnoty od 30 do 70 milisekúnd.

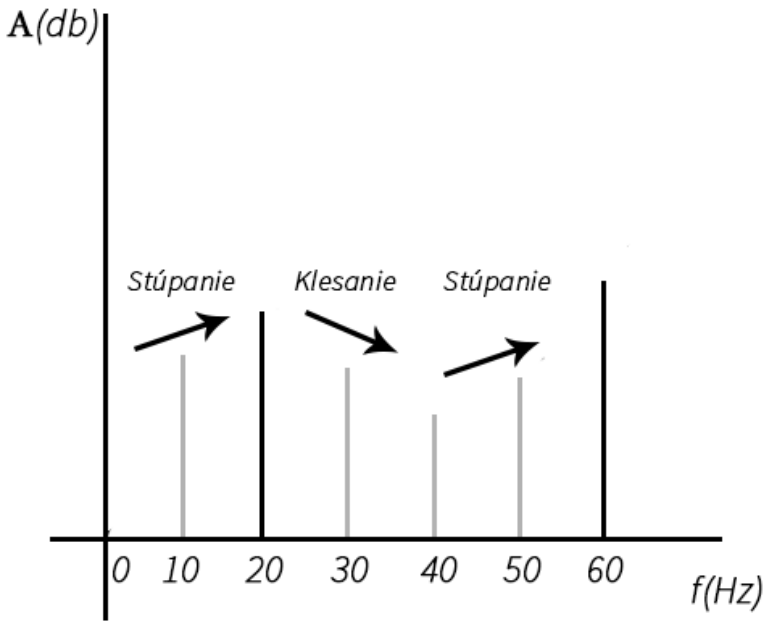
5.1.4 Získanie spektra hlasitostí

Získaný signál z FFT bolo nutné prvok po prvku normalizovať podľa vzorca 4.2, aby signál neobsahoval imaginárnu a reálnu zložku. Absolútny signál sa teda získal po normalizácii všetkých prvkov, ktoré sa nachádzali v signáli z FFT.
Absolútny signál reprezentoval spektrum hlasitostí pre dané indexy, z ktorého sa dokázali získať potrebné údaje (nasnímané tóny v danom časovom okamžiku) na rozpoznávanie tónov a akordov.
Absolútny signál je usporiadané pole, ktoré obsahuje na každom indexe energiu. Indexy reprezentujú vzorky, inak frekvencie z nahrávania. S danými indexami sa nepracovalo ideálne, pretože daný index hneď nereprezentoval frekvenciu (frekvencia by sa musela získať prevodom) a už vôbec nie hudobný tón. Indexy boli zoradené od najhlbšej frekvencie po najvyššiu. Rozhodnutím na záver bolo, že index sa nebude brať ako frekvencia nasnímaného audio signálu, ale ako celočíselná hodnota, s ktorou sa bude pracovať.

5.1.5 Algoritmus na nájdenie lokálnych maxím

Na získané spektrum hlasitostí bol aplikovaný vlastný algoritmus na získanie lokálnych maxím. Tento algoritmus funguje veľmi jednoducho, začal od najhlbšieho tónu a sledoval či

ďalší index je hlasnejší alebo tichší. Ak ďalší index bol hlasnejší, tak si uložil daný index a kontroloval indexy ďalej, až kým nenarazil na index, ktorý je tichší, toto sa nazvalo stúpaním v hlasitosti ku zahranému tónu. Tým pádom sa našiel lokálny vrchol, ktorý sa nachádza jeden index pred spomínaným tichším indexom. Od prvého uloženého indexu sa indexy nulovali až po nájdený vrchol, ten si algoritmus nechal nenulový a znovu si prvý tichší index uložil a algoritmus pokračoval v hľadaní hlasnejšieho indexu ako indexu pred ním, toto hľadanie bolo nazvané klesaním. Po nájdení hlasnejšieho indexu algoritmus pokračoval ako na začiatku. Ak úplne prvý index bol hlasnejší ako nasledujúci index, tak sa nenuloval prvý index a pokračovalo sa ďalej klesaním. V podstate sa algoritmus orientoval podľa klesania hlasitostí a stúpania hlasitostí a nuloval všetky indexy okrem bodu zlomu, a to index pred označením klesania, teda nájdeného lokálneho maxima. Presnosť algoritmu je približne 85 %, ale nikdy nedochádza v algoritme k nulovaniu hodnoty, ktorá má byť braná ako vrchol. Grafické znázornenie algoritmu je ukázané na obrázku 5.1. Z obrázku vyplýva, že iba hodnoty na 20 a 60 Hz (tmavšie hodnoty) budú po algoritme nenulové.



Obr. 5.1: Získanie lokálnych maxím

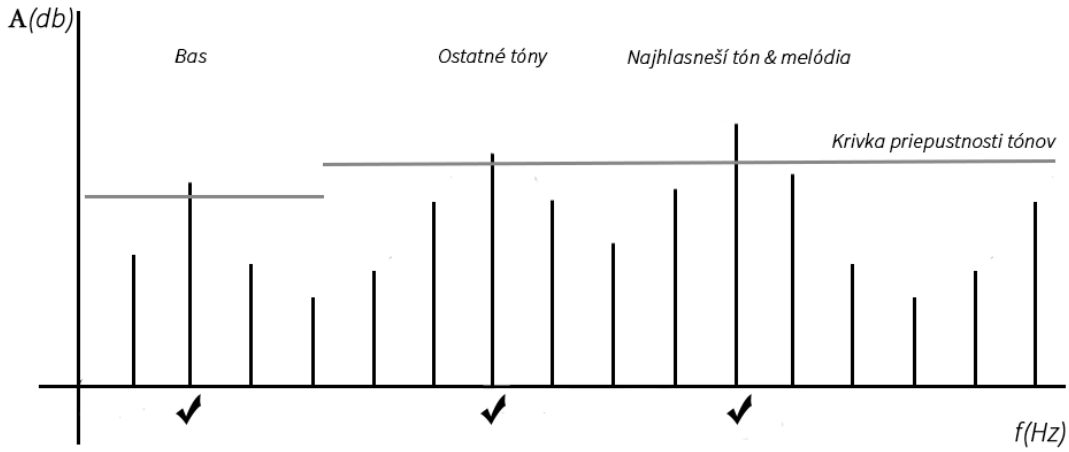
5.1.6 Získanie zahraných tónov

Na získanie hraných tónov bolo dôležité správne pracovať so spektrom hlasitostí, ktoré obsahovalo len vrcholy. Zo spektra sa dokáže zistiť, ktoré indexy majú aké veľké hodnoty, tieto hodnoty reprezentujú hlasitosť.
Najhlasnejší index je index s najväčšou hodnotou v spektre hlasitostí. Táto najhlasnejšia hodnota určuje najväčšiu hlasitosť celého spektra v aktuálnom čase a algoritmus sa orientuje podľa tejto hodnoty na získanie aj ostatných zahraných tónov v tomto čase.

Na získanie aj ostatných zahraných tónov, ktoré boli tichšie sa aplikovala takzvaná krivka priepustnosti tónov.

5.1.7 krivka priepustnosti tónov

Krivka priepustnosti tónov (ďalej len krivka) pracuje veľmi jednoducho a mení sa stále podľa novo získaného nasnímaného spektra hlasitostí v danom čase.
Krivka fungovala nasledovne, a to, že získaná najhlasnejšia hodnota reprezentovala hlasitosť, ktorou sa v danom časovom okamžiku približne hralo. Vyplynulo, že ďalej treba hľadať indexy s hodnotou, ktorá bude približne rovnaká. Krivka sa teda definovala na hodnotu o 10 % nižšiu ako najhlasnejšie nasnímaná hodnota v danom čase. Postupom času a aktívnym testovaním aplikácie počas vývoja sa zistilo, že v hlbokých frekvenciách zariadenie nedokáže nasnímať takú vysokú hlasitosť ako vo vyšších frekvenciách, a preto sa musela krivka v hlbokých frekvenciách znížiť o 15 % oproti najhlasnejšej nasnímanej hodnote.
Krivka teda neprepustí indexy na ďalšiu analýzu, dokým hlasitosť daného indexu neprekročí krivku. V basoch sa jednalo teda o 15 % a v stredových a vysokých frekvenciách o 10 % z hodnoty najhlasnejšieho nasnímaného indexu. Indexy, ktoré sa cez túto krivku dostali sa neskôr presne analyzovali a získali sa rozpoznané tóny a akordy, teraz sa budú označovať ako dôležité dáta. Krivka aj s označenými prepustenými tónmi pod osou x je znázornená na obrázku 5.2.



Obr. 5.2: Krivka priepustnosti tónov

5.1.8 Získanie a analýza dôležitých tónov

Dôležité dáta boli zoradené a uložené do poľa. Pre určenie nasnímaného akordu v danom čase sa získané indexy musia analyzovať a rozdeliť na niekoľko skupín. Tieto skupiny sú nasledovné.

- Najhlasnejší index
- Melódia

- Bas
- Ostatné nasnímané indexy

Pomocou týchto kategórií sa mohli určiť zahraný akord v reálnom čase. Z obrázku 5.2 sa dá všimnúť tieto dôležité tóny.

Najhlasnejší index

Táto kategória reprezentuje najhlasnejší nasnímaný index. Kategória a jej index sa ráta do výsledného rozpoznania akordu. Táto informácia je taktiež dôležitá pre užívateľa, a preto sa táto hodnota zobrazuje v aplikácii (v aplikácii sa to zobrazuje už ako tón, nie ako index).

Melódia

Melódia určovala najvyššie nasnímaný index z pohľadu frekvencie. Jedná sa o index podľa, ktorého určujeme frekvenčne najvyššie položenú linku. Podľa tejto linky sa vie hudobník orientovať, čo bolo zahrané frekvenčne najvyššie, to sú často témy skladieb, spev alebo sólo jedného hudobného nástroja.
Táto kategória sa zahrňuje do rozpoznania akordu a užívateľovi sa znázorňuje do reprezentácie, ako veľmi dôležitý tón, ktorý bol rozpoznaný z pôvodného audio signálu.

Bas

Kategória bas určovala frekvenčne najhlbšie nasnímaný index. Tento index bol veľmi dôležitý, pretože akord sa skladal od basu, teda od tohto indexu a užívateľovi ho aplikácia taktiež znázorňuje samostatne (v aplikácii sa bas zobrazuje už ako tón, nie ako index).

Ostatné nasnímané indexy

Ostatné nasnímané indexy sa nachádzajú v poli a určujú ostatné stupne, tenzie akordu. Zahrnula sa táto kategória na rozpoznanie akordov a užívateľ taktiež vidí jednotlivé tóny, podobne znázornené ako bas.

5.1.9 Algoritmus určovanie akordov

Zo zahraných kategorizovaných tónov sa mohol určiť výsledný akord, zahraný v danom čase. Postupovalo sa tak, že na začiatku sa zistil najspodnejší hraný tón, teda bas a či sa nasnímali aj iné tóny. Ak toto odpovedalo pravde, tak sa získali hodnoty iných tónov k basu. Tieto hodnoty pomohli určiť ostatné tenzie, aké akord obsahuje, napríklad terciu, kvintu, sekundu atď [6].
Tóny v programe majú hodnotu od 1 po 12 [13], čo zodpovedá počtu poltónov v európskej hudbe od primy po prvú vyššiu oktávu. Tón C má hodnotu 1, tón Db hodnotu 2, atď. Po určení hodnoty tónov sa dané tóny porovnávali k basu a uložili sa výsledné rozdiely. Vznikla vzdialenosť počtu poltónov daného tónu oproti basu, týmto spôsobom sa našli hudobné intervaly, ktoré sú popísané a znázornené v tabuľke 4.1.
Z uložených počtov sa najprv zisťovali tercie, ak akord obsahoval obe tercie, jednalo sa o akord s označením *7b9* alebo *7+9*. Ak akord obsahoval iba veľkú terciu predpokladal sa durový akord, ak iba malú terciu predpokladal sa molový akord. Ďalej sa zisťovala septima, potom kvinta, ďalej sexta a nakoniec sekunda. Ak zahrané tóny neobsahovali žiadnu terciu,

predpokladali sa sus akordy alebo obraty akordov, kde sa dopisovali iba stupne, podobne ako u durových či molových akordoch. Prehľad základných akordov, ich tóny a tenzie sú znázornené v tabuľke 4.2.

5.2 Trénovanie rozpoznávania tónov

Na získanie tónu z nasnímaných indexov sa potrebovali natrénovať tóny, ktoré reprezentujú jednotlivé indexy.
Aby bolo možné pracovať s inou hodnotou ako s presnou frekvenciou, tak sa pre každý tón natrénovali jeho celočíselné hodnoty. Trénovanie prebiehalo na online generátore frekvencií[1]. Rátalo sa s ladením 440 Hz pre tón *A4*.
Každý tón má približný rozsah frekvencií, napríklad pre tón *A4* je to od 428 Hz do 452 Hz a všetky indexy, ktoré spadali do tohto intervalu sa označili ako tón *A4*. Takto sa postupovalo od tónu *E1* až po *C8* a natrénovali sa tóny pre aplikáciu. Generovaním frekvencie a snímaním audio signálu v aplikácii sme získali dané indexy. Na záver sa indexy a tóny skontrolovali podľa elektrického klavíru a upravili rozsah indexov pre dané tóny.
Nasnímané a kategorizované indexy v reálnom čase sa pomocou týchto tónových intervalov prevádzali na už reálne hudobné tóny.

5.3 Detekcia šumu a hluku v okolí

V aplikácii je cieľom oddeliť dôležité dáta od nedôležitých. Nedôležité dáta sa vyskytujú v šume a ruchu okolia, práve toto okolie sa sníma a oddelia sa dané dáta.
Pri zapnutí aplikácie sa automaticky začne snímať zvuk z okolia po dobu 3 sekúnd, užívateľ by nemal púšťať žiadne zvuky a ani hrať na hudobný nástroj. Počas týchto sekúnd sa spustí detekcia hlasitosti okolia, ktorá prebehne toľko krát koľko stihne za daný časový interval (nastavené 3 sekundy snímania šumu). Najvyššia hlasitosť sa vždy ukladá po dobu tohto snímania. Po ukončení tejto detekcie sa vytvorí priemer týchto hlasitostí a vznikne tzv. prah snímania zvuku.
Prah snímania zvuku určuje najnižšiu možnú hranicu, ktorú musí prekonať zvuk, aby sa sním počítalo a bol zaznamenaný medzi dôležité dáta a nie ako šum a hluk z okolia. Tým pádom sa vyradili nežiaduce signály a do výpočtu akordov a tónov sa dostali iba dôležité nasnímané dáta.

5.4 Reprezentácia audio signálu

V aplikácii sa reprezentoval audio signál troma spôsobmi. Prvý spôsob je po štarte aplikácie a vypisuje informácie z audio signálu v reálnom čase a nezachováva si históriu. Druhý spôsob je po nastavení nahrávania a zaznamenané informácie sú zaznamenávané počas celej doby nahrávania a v reálnom čase vypisované do objektu a posledný spôsob je vo forme audio záznamu.
V programe je cieľom, aby sa textové reprezentácie čo najviac podobali notovému zápisu a aby použité značky odpovedali hudobnej teórii. Je nutné audio záznam reprezentovať vo vhodnom audio formáte a v postačujúcej kvalite.

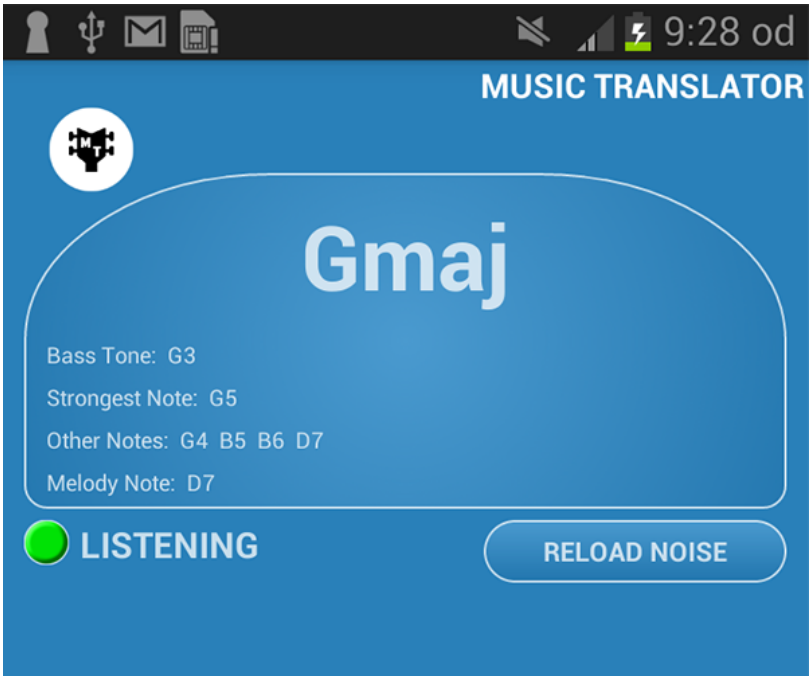
[1] Generátor je dostupný na http://www.szynalski.com/tone-generator/

5.4.1 Audio reprezentácia

Dôležitým aspektom pre užívateľa je zaznamenávanie jeho nápadov vo forme audio záznamu, ktoré si vie hocikedy prehrať. Android zariadenia ponúkajú audio zaznamenávanie v rôznych formátoch. Dôležité je, aby formát zaznamenaného audio záznamu bol čo najpoužívanejší a bolo možné ho otvoriť bez problémov aj mimo aplikáciu, napríklad v počítači.
V aplikácii sa rozhodlo používať triedu *AudioRecord*, ktorá umožnila získavať bitové pole, ktoré bolo vhodné na rozpoznávanie tónov a akordov, ale podporovaný audio formát (raw) nezodpovedal požiadavkám. Preto sa rozhodlo bitové pole prekonvertovať na vhodný formát, a to Waveform audio file format (ďalej wav). Konvertovanie prebiehalo pri ukončení režimu nahrávania audio záznamu. Bolo nutné v práci vytvoriť hlavičku pre wav, pretože priame konvertovanie Android neponúkal. Hlavičku bolo nutné nakonfigurovať presne a obsahovala údaje v bitoch, ktoré sa museli nastaviť podľa nakonfigurovaného zariadenia pre nahrávanie, ktoré sa popísali vyššie v podsekcii 5.1.1. Daná hlavička sa vždy pripojila na začiatok uloženému bitovému poľu po nahrávaní audio dát.
Aplikácia ukladá bitové pole iba v prípade, keď sa jedná o užívateľom zvolené nahrávanie jeho hudobnej idei. Je zbytočné plytvať zdrojmi mobilného zariadenia a ukladať audio dáta mimo zvoleného nahrávacieho režimu, preto sa snímanie zastavuje, napríklad pri minimalizovaní aplikácie.

5.4.2 Reprezentácia informácií po štarte aplikácie

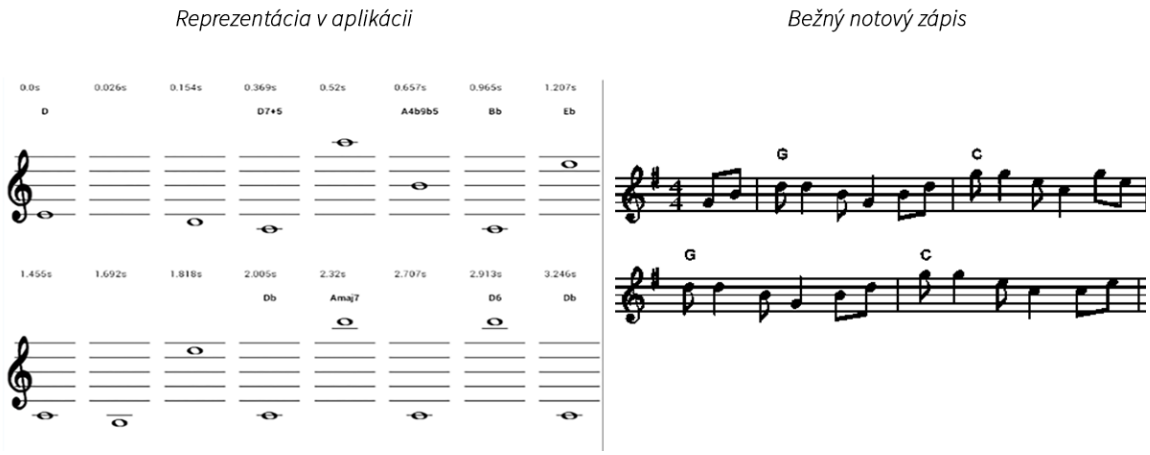
Táto reprezentácia slúži na rýchle zobrazenie informácií z audio signálu v reálnom čase hneď po štarte aplikácie, táto reprezentácia je graficky znázornená na obrázku 5.3. Medzi zobrazované informácie patria: zahraný akord, najhlasnejší tón (Strongest Note), frekvenčne najvyšší (Melody Note) a najnižší tón (Bass Tone), teda melódia a bas a posledná informácia sú ostatné nasnímané akordy (Other Notes). V reprezentácii sa neuchováva história, takže sa jedná len o rýchlu analýzu audio signálu. Táto reprezentácia je vhodná napríklad na doladenie hudobného nástroja alebo na zistenie hlasitosti ruchu v miestnosti.



Obr. 5.3: Reprezentácia informácii po štarte aplikácie

5.4.3 Reprezentácia počas nahrávania

Daná reprezentácia obsahuje zápis, ktorý je čo najviac podobný notovému zápisu. Notový zápis ako aj reprezentácia v aplikácii obsahuje linky, kľúč. V reprezentácii sa použil husľový kľúč, ktorý je univerzálnejší ako napríklad basový kľúč. Na linky sa umiestňujú tóny. V aplikácii sa nerozlišuje dĺžka nôt, takže sa používali iba celé noty. Dĺžku zahraných tónov reprezentuje časový údaj, ktorý je znázornený pri každom objekte.
Reprezentácia obsahuje históriu zahraných tónov a akordov po dobu nahrávania a údaje sa vykresľujú a vypisujú v reálnom čase. Medzi zobrazené informácie patrí akord v danom čase, najvyšší tón zaznamenaný na notovej osnove a čas v sekundách, ktorý odpovedá, kedy boli informácie nasnímané od začiatku nahrávania.
Po ukončení nahrávania sa reprezentácia nevymaže, ale užívateľ môže s danou reprezentáciou pracovať alebo ju analyzovať a uložiť či zahodiť.
Porovnanie bežného notového zápisu [13] a našej reprezentácie je vidieť na obrázku 5.4. Daná reprezentácia v práci nepoužíva dĺžky nôt, užívateľ sa orientuje podľa času nad každým objektom. Ďalej aplikácia nerozpoznáva tempo nahrávaného hudobného nápadu, a preto nezobrazuje taktové čiary.



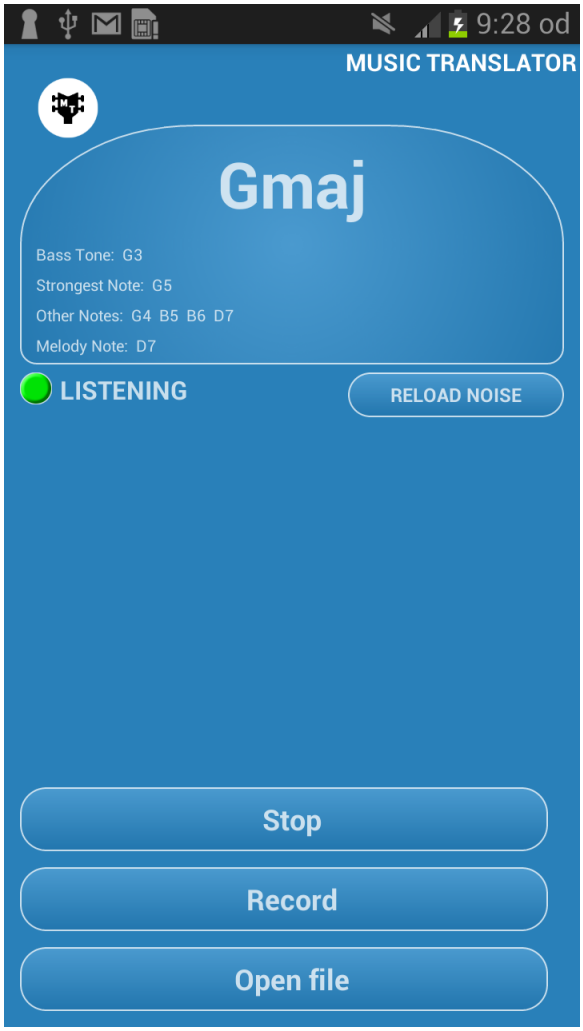
Obr. 5.4: Reprezentácia počas nahrávania a notový zápis

5.5 Užívateľské rozhranie

Pri tvorbe mobilnej aplikácie, ktorá zobrazuje dáta čo najrýchlejšie je nutné vytvoriť ľahko ovládateľné, čo najviac informačné a intuitívne užívateľské rozhranie.
Cieľom je zvoliť jednoduché rozhranie, ktoré bude minimalistické a bude obsahovať iba dôležité dáta, ktoré budú vždy rýchlo dostupné pre hudobníka. Ďalšie kritérium je v nezaťažovaní zariadenia, ktoré sa touto cestou dá dosiahnuť.

5.5.1 Hlavná obrazovka aplikácie

Užívateľ na hlavnej obrazovke vidí vrchný panel, kde sa mu zobrazujú analyzované zvuky z mikrofónu zariadenia. Informácie obsahujú posledný nasnímaný akord a tóny v reálnom čase. V spodnej časti obrazovky má ľahko dostupné tlačidlá. Na obrázku 5.5 je vidieť grafické rozhranie hlavnej obrazovky.



Obr. 5.5: Hlavná obrazovka

5.5.2 Snímanie hluku miestnosti

Po zapnutí aplikácie sa zobrazí užívateľovi obrazovka so snímaním hluku miestnosti, užívateľ by nemal hrať dokým táto obrazovka je viditeľná. Táto funkcia slúži na detekciu hlasitosti zvuku v miestnosti, ktorý nie je pre neskoršiu detekciu dôležitý. Po ukončení tejto funkcie si algoritmus zapamätá priemernú hlasitosť, ktorá bola nameraná a pri neskoršom analyzovaní zvukov bude algoritmus počítať iba s hodnotami, ktoré sa dostali nad túto priemernú hlasitosť.

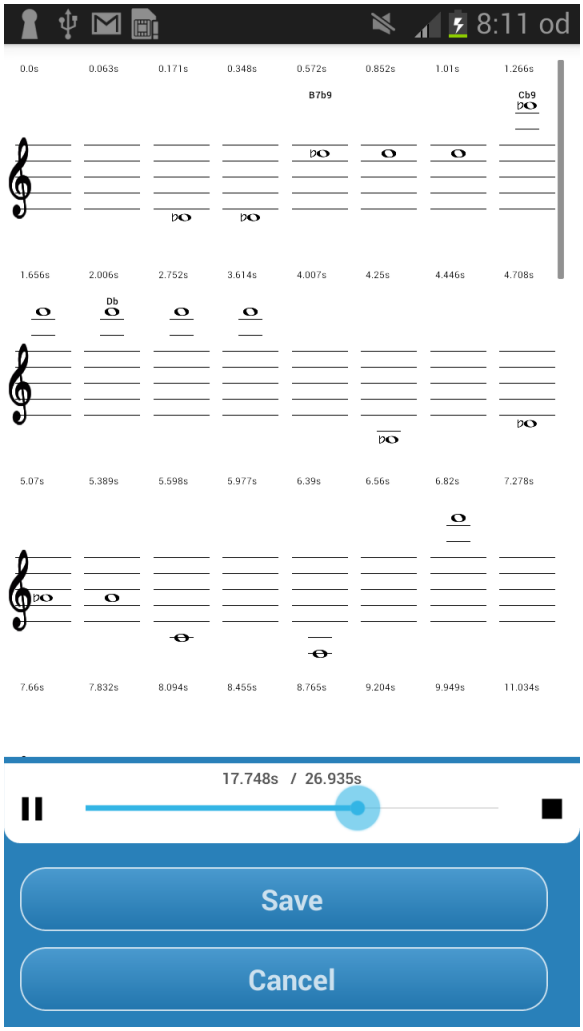
5.5.3 Obrazovka nahrávania zvuku v reálnom čase

Obrazovka je aktívna po kliknutí tlačidla nahrávanie v hlavnej obrazovke. Táto obrazovka obsahuje vrchný panel rovnaký ako u hlavnej obrazovky, stredný panel s hodnotami už zaznamenaných a analyzovaných akordov a tónov v zápise podobnej notovej osnovy z hudobnej teórie. Na spodnej časti obrazovky sú tlačidlá, ktorými vie užívateľ spustiť, pozastaviť alebo ukončiť nahrávanie. Obrazovka slúži na nahrávanie hudobných ideí a je hlavnou funkciou aplikácie. Po stlačení tlačidla na ukončenie nahrávania sa aplikácia buď vráti späť na

hlavnú obrazovku (ak nahrávanie neprebehlo) alebo pokračuje na obrazovku prehrávania hudobných ideí.

5.5.4 Obrazovka prehrávanie hudobných ideí

Daná obrazovka môže byť zobrazená dvoma spôsobmi. Prvý spôsob je zobrazenie obrazovky po nahrávaní zvuku v reálnom čase, kde obrazovka obsahuje zápis práve nahraného zvuku do zápisu podobného notovej osnove na väčšej časti obrazovky, ďalej obsahuje prehrávací panel s tlačidlami (spusti, pauza, stop), text s časom prehrávaného súboru, grafickou a dotykovou časovou osou prehrávania. Táto obrazovka obsahuje ešte tlačidlá na uloženie hudobnej idei alebo zrušenie a zahodenie tejto hudobnej idei. Po uložení hudobnej idei sa uloží zvukový záznam do súboru (.wav) a hudobný zápis do textovej reprezentácie JSON a uloží sa taktiež do súboru (.txt). Užívateľ si vyberie meno ako sa daný nápad bude volať a pod takým menom sa súbory uložia do pamäte zariadenia. Druhý spôsob je po otvorení uloženého hudobného nápadu z obrazovky výber uloženého hudobného záznamu. Oba spôsoby sú skoro identické, druhý spôsob obsahuje namiesto možnosti uloženia záznamu (keďže záznam je už uložený) text s názvom súboru.
Obrázok 5.6 znázorňuje prvý spôsob prehrávania hudobnej idei.



Obr. 5.6: Obrazovka prehrávania hudobnej idei

5.5.5 Výber uloženého hudobného záznamu

Do obrazovky sa užívateľ dostane po stlačení tlačidla otvoriť súbor na hlavnej obrazovke aplikácie. Obrazovka načíta uložené súbory vytvorené aplikáciou (hudobné idey alebo skladby, ktoré si užívateľ zaznamenal) z pamäte zariadenia a zobrazí ich do listu. List obsahuje jednotlivé hudobné nápady, ktoré sú zoradené podľa dátumu nahrania. Ďalej jednotlivé zložky obsahujú názov (meno pod akým bol súbor uložený), dátum uloženia súboru a dĺžku hudobnej idei. Po kliknutí na jednotlivé zložky sa zobrazí obrazovka na prehrávanie hudobných ideí, ktorá bude obsahovať hudobný zápis a zvuk zo súborov pod rovnakým menom.

Kapitola 6

Testovanie aplikácie

Testovanie prebiehalo v dvoch fázach. Prvá fáza znamenala aktívne testovanie počas vývoja aplikácie, pomocou online generátora frekvencií a neustáleho testovania pomocou hrania na hudobných nástrojoch, ktoré sa zaznamenávalo a porovnávalo. Druhá fáza alebo komplexné testovanie presnosti bolo zamerané na väčšie a komplexnejšie testy. Tieto testy prebiehali na hudobných nástrojoch a prehrávaním skladieb a etúd cez kvalitný audio systém.
Testovanie prebiehalo taktiež na rôznych mobilných zariadeniach, aby výsledky testov boli čo najpresnejšie a vypovedali čo najviac. Veľa závisí od vstavaného mikrofónu a výkonu v mobilnom zariadení ako je popísané v sekcii 3.2. Na testovanie sa použili rôzne mobilné zariadenia (Samsung Note1, Samsung Galaxy S3, LG Nexus).
V kapitole na začiatku je rozobrané aktívne testovanie, ktoré prebiehalo počas vývoja aplikácie, opis hlavných testov na presnosť rozpoznávania tónov a akordov v sekcii 6.2 a na záver sú vyhodnotené užívateľské testy v sekcii 6.3.

6.1 Aktívne testovanie počas vývoja aplikácie

Aktívne testovanie počas vývoja aplikácie bolo veľmi dôležité a nutné pre predbežné výsledky pri vývoji ohľadom presnosti rozpoznávania tónov a akordov z audio signálu.
Ako prvé sa zvolil kvalitný generátor frekvencií, ktorý je dostupný online. Zistil sa rozsah tónov, ako je popísané v kapitole 5.2. Neskôr sa tieto hodnoty porovnávali s reálnych hraním na hudobných nástrojoch a dolaďovali nepresnosti.
Tento proces sa opakoval vždy po zmenení alebo vylepšení algoritmu na rozpoznávanie tónov a jeho výsledky sa zaznamenávali a porovnávali. Na záver vývoja aplikácie sa začali vykonávať väčšie a komplexnejšie testy, ktoré sú zobrazené v tabuľkách nižšie.

6.2 Komplexné testovanie presnosti rozpoznávania

Komplexné testovanie presnosti pozostávalo z troch veľkých testov, test monofónneho záznamu, test polyfónneho záznamu, test hudobných skladieb a etúd. Audio záznam bol zaobstaraný na pianíne Petrof a basových gitarách Yamaha trb-1006 a Fender jazz-bass. Audio záznam bol zaobstaraný v bežných podmienkach neodhlučnenej izby, ako sa aplikácia bude najčastejšie používať.
V testovaní sa zisťovala percentuálna presnosť hraných tónov alebo akordov a priemernú veľkosť chyby. Hodnoty sa získali opakovaným hraním stupníc alebo akordov alebo hraním skladieb z audio systému. Po každom zázname sa zaznamenali nasnímané hodnoty. Po do-

končení opakovacích testov sa hodnoty spriemerovali a zapísali do tabuliek.
Prvé dva typy testov na presnosť daných tónov alebo akordov sa získali porovnaním audio záznamu a reprezentácie záznamu v aplikácii. Testy sa vykonávali v rôznych oktávach, z toho sa dá dedukovať presnosť v určitých frekvenčných pásmach.
Posledné testy prebiehali ako porovnávanie audio nahrávky z notového zápisu skladby. Cieľom bolo si zvoliť vhodné skladby. Porovnávalo sa či je tón alebo akord zahraný presne podľa referenčnej nahrávky alebo partitúry, pri každej odchýlke oproti pôvodnej partitúry bol tón či akord zaznamenaný ako chyba. Po testoch na presnosť sa dané chyby a zhody spočítali a vypočítala sa percentuálna úspešnosť zhody.
Veľkosť chyby sa merala zaznamenávaním o koľko poltónov sa líšil zahraný tón od zaznamenaného tónu v našej aplikácii. Výsledná priemerná veľkosť chyby sa získala ako priemer zaznamenaných chybových poltónov a počtu chýb. Do tohto výsledku sa nepripočítavali zhody.

6.2.1 Testovanie monofónneho audio záznamu

Tento test pozostával s opakovaním hraním stupníc v rôznych hudobných kľúčoch a oktávach. Stupnice boli hrané v rôznych tempách a na rôznych hudobných nástrojoch a mobilných zariadeniach. Výsledky testu monofónneho audio záznamu sú reprezentované nižšie.

Hudobný nástroj	Hraná oktáva	Priemerná presnosť hraných tónov	Priemerná veľkosť chyby
Basová gitara	C1 - C2	31,25%	3 poltóny
Basová gitara	C2 - C3	68,75%	1.8 poltónu
Klavír	C3 - C4	78,5%	1,2 poltónu
Klavír	C4 - C5	72,5%	1,35 poltónu
Klavír	C5 - C6	71%	1,4 poltónu

Tabuľka 6.1: Presnosti a chyby monofónneho audio záznamu

Z tabuľky 6.1 vyplýva, že aplikácia má problémy rozpoznávať tóny v hlbokých frekvenciách, môžu za to menej kvalitné mikrofóny v zariadeniach Android, tento problém sme popísali v návrhu aplikácie v sekcii 3.2.2. Ďalej z tabuľky sa dá dedukovať obstojná presnosť pri klavíry, priemerná veľkosť chyby nie je veľká vo vyšších frekvenciách, v podstate sa jedná o drobné nepresnosti, ktoré si dokáže hudobník sám uvedomiť pri vypočutí nahraného hudobného záznamu.

6.2.2 Testovanie polyfónneho audio záznamu

Tento test pozostával s opakovaným hraním základných akordov v rôznych hudobných kľúčoch. Akordy boli hrané v rôznych tempách a na rôznych zariadeniach. Výsledky testu polyfónneho audio záznamu sú reprezentované nižšie.

Hudobný nástroj	Hraná oktáva	Priemerná presnosť hraných akordov
Basová gitara	C2 - C3	10,7%
Klavír	C2 - C3	25,6%
Klavír	C3 - C4	21,8%
Klavír	C4 - C5	20,2%
Klavír	C5 - C6	22%

Tabuľka 6.2: Presnosti a chyby polyfónneho audio záznamu

Z tabuľky 6.2 sa dajú dedukovať podobné chyby a nepresnosti ako v teste monofónneho audio záznamu. Namerané nepresnosti pri snímaní akordov majú priemerne iba 22 %, chyby pri akordoch nešli zmerať tak ako veľkosť chyby v poltónoch pri monofónnom zázname. Nízku presnosť pri rozpoznávaní akordov sa dá ospravedlniť nasnímaním správnych tónov a niektoré tóny boli nasnímané o oktávu vyššie alebo nižšie, takže sa stávalo, že sa akord skladal od basu, ktorý reálne nebol zahraný najhlbšie. Napríklad, že reálne zahraný akord sa skladal z tónov *C, E, G*, kde tón *C* predstavoval bas, z toho vyplýva akord *C dur*. Aplikácia ale nasnímala tóny v poradí *G, C, E*, tón *G* predstavoval bas, a teda výsledný akord bol *G6*.

6.2.3 Testovanie aplikácie na hudobných skladbách

V poslednom rade sa aplikácia testovala na hudobných skladbách a etudách, ktoré boli hrané cez kvalitný audio prehrávač v neodhlučnenej miestnosti. Skladby sme vyberali podľa rôznych hudobných inštrumentov, tempa skladby a náročnosti danej skladby. Výsledné vybrané namerané hodnoty sa nachádzajú v tabuľke 6.3.

Autor skladby	Názov skladby	Hudobný inštrument	Priemerná presnosť hraných akordov	Priemerná presnosť hraných tónov	Priemerná veľkosť chyby
N. Adderley	Work Song	Trúbka	Monofónny záznam	55%	2,5 poltóny
J. S. Bach	Prelude in G major	Čelo	Monofónny záznam	33%	3 poltóny
L. v. Beethoven	Moonlight sonata	Klavír	20%	60%	2 poltóny

Tabuľka 6.3: Test hudobných skladieb

Z tabuľky 6.3 vyplýva, že hudobný nástroj trúbka má lepšie výsledky ako čelo, ktoré produkuje oveľa hlbšie tóny ako trúbka, ktoré sa horšie rozpoznávajú. V hudobnej skladbe od J. S. Bacha sa nachádzajú rýchle pasáže, s ktorými mala aplikácia problémy. Preto je daná presnosť tónov iba približne 33 %. N. Adderleyho skladba má veľké intervalové skoky, ktoré aplikácia celkom obstojne rozpoznala, ale pri rozpoznaní krátkych rovnakých vysokých tónov mala problémy a snímala ich ako jeden nepretržitý dlhý tón. U klavírnej skladby od L. v. Beethovena sa získala presnosť približne 60 %, čo je obstojný výsledok. Skladba je v miernom tempe a hrajú sa prevažne dlhé tóny, ktoré aplikácia dokáže celkom schopne rozpoznávať.
Z tohoto testu sa môže dedukovať, že u každej testovanej skladby by hudobník nemal mať problém určiť základný tón skladby a tóninu skladby čo je veľmi dôležité. Bohužiaľ výsledky sú približné (asi na 5 % zaokrúhlenie), kvôli rozdielnej reprezentácii rozpoznaných tónov a akordov oproti hudobnej notovej osnove.

6.2.4 Zhodnotenie výsledkov testovania

Aplikácia dokáže užívateľovi zaznamenať jednoduchú hudobnú skladbu alebo jeho hudobný nápad na presnosť v rozmedzí 50 % - 75 % s drobnými odchýlkami o pár poltónov (1,2 - 2 poltónov). V zložitejších hudobných skladbách aplikácia rozpoznáva tóny a akordy orientačne, ale užívateľ vie veľmi ľahko zistiť tóninu a základný tón skladby.
Zistilo sa, že okrem hlbokých frekvencií má aplikácia problém so snímaním hudobných nápadov vo vysokom tempe, táto nepresnosť môže prameniť z menej výkonných mobilných zariadení.
Pri návrhu aplikácie sa rátalo s niektorými problémami, testy nám ukázali, že aplikácia dokáže pracovať a môže užívateľovi byť nápomocná a zjednoduší mu prácu s hudobnou skladbou alebo jeho hudobnou ideou, ktorú chce nahrať a nejako textovo zaznamenať.

6.3 Zhodnotenie užívateľských testov

Po uvedení aplikácie na Google Play sa oslovila skupina ľudí, ktorá si stiahla aplikáciu a bola ochotná ju otestovať a na záver vyplniť dotazník[1], ktorý obsahoval niekoľko otázok rozdelených do skupín. Skupiny boli zamerané na zistenie hudobnej znalosti užívateľa, funkčnosť aplikácie, jej dizajn a výsledné hodnotenie.
Dotazník zodpovedalo asi 10 ľudí, jednalo sa o profesionálnych, amatérskych hudobníkov a užívateľov, ktorí hrávajú na hudobný nástroj iba občas. Z oslovených ľudí sa všetci vyznali minimálne v základoch hudobnej teórie, z čoho sa môže dedukovať, že aplikácii rozumeli. Väčšina oslovených užívateľov vlastní zariadenie Android, z ktorých 90 % ho využíva každý deň. Ani jedno testované zariadenia Android nie je staršie ako 2 roky. Z toho vyplýva, že aplikáciu by vedeli využívať denne.
Každému užívateľovi fungovala aplikácia, z toho 80 % označilo rozpoznávanie tónov a akordov ako čiastočne nepresné a zbytok užívateľov označilo možnosť veľmi nepresné u akordov. Rýchlosť rozpoznávania tónov a akordov aplikácie bola ohodnotená na 80 % ako rýchla a na 20 % ako uspokojujúca. Dvaja užívatelia podotkli chybu pri vypnutí a zapnutí aplikácie, že rozpoznávanie neprebiehalo a museli aplikáciu znovu vypnúť a zapnúť. Test funkcionality zaznamenal obstojné výsledky, aplikácia rozpoznáva akordy obstojne a tóny veľmi dobre, rýchlosť rozpoznávania je ohodnotená veľmi dobre a vyskytujúce chyby sú zanedbateľné.
Ohľadom dizajnu a užívateľského rozhrania sa zistilo, že 80 % oslovených pochopilo aplikáciu hneď a 20 % do pár minút. Všetci označili, že dizajn aplikácie sa im páči a na záver podotkli len dvaja užívatelia minoritné pripomienky. Z toho vyplýva, že aplikácia je jednoduchá na pochopenie a má lákavé a pekné užívateľské rozhranie.
Posledné testy sa zaoberali zhrnutím a všeobecnými otázkami. 60 % užívateľov označilo aplikáciu za plne využiteľnú na každodennú prácu a zbytok označilo možnosť využívania s obmedzením. Otázka ohľadom zakúpenia aplikácie za jedno euro bola zodpovedaná na 40 % za zakúpenie, 40 % odpovedala možno a 20 % by si aplikáciu nezakúpilo. Ohľadom nových zmien a vylepšení v aplikácii sa vyjadrili užívatelia na 80 % pozitívne a 20 % podotklo nepresnosti v rozpoznávaní akordov. Návrhy na vylepšenia aplikácie boli minoritné a pravdepodobne sa zahrnú do budúceho vývoja. Užívatelia ohodnotili na 60 % aplikáciu ako výborný nápad, ktorú by využívali každý deň, 40 % označili aplikáciu ako dobrú a nápomocnú a možno by ju využívali a zbytok označil prácu za dobrý nápad a aplikáciu by využíval len po zdokonalení funkcionality.

[1] Dotazník bol vytvorený ako Google formulár (Google Form), dostupný na https://docs.google.com/forms

Z užívateľských testov sa zistilo, že aplikácia je rýchlo pochopiteľná, rýchla, funkčná, schopná rozpoznávania tónov a akordov, aplikácia sa môže používať každý deň a môže profesionálnym i amatérskym hudobníkom uľahčiť prácu s ich hudobnými ideami a nápadmi.

6.4 Zhodnotenie výsledkov z Google Play

Momentálne je aplikácia publikovaná na Google Play približne tri týždne. Prvých 20 dní bola aplikácia dostupná na stiahnutie zadarmo a zbytok času je dostupná na stiahnutie za 0.59 eur. Za obdobie si aplikáciu stiahlo 46 užívateľov, z tohto počtu si aplikáciu zakúpili dvaja užívatelia. Aplikáciu odinštalovalo 18 ľudí a 26 ľudí ju aktívne využíva. Za toto obdobie nastal iba jeden pád aplikácie. Momentálne má aplikácia priemerne hodnotenie na 5 bodov, čo je maximálny možný počet. Aplikáciu ohodnotilo 10 užívateľov.
Pri veľkom poklese sťahovania odstránim cenu za stiahnutie aplikácie a aplikácia bude znovu dostupná zadarmo. Momentálny počet užívateľov je malý, ale aplikácia je zameraná na úzku klientelu. Dôležité aspekty sú, že aplikácia je aktívne využívaná v 26 zariadeniach, a že aplikácia nezaznamenáva chyby a pády, okrem jedného prípadu, ktorý sa už nezopakoval.
Do budúcna je možné aplikáciu vylepšiť a aktualizovať verziu na Google Play. Aktualizácia môže súvisieť s výsledkami užívateľských testov a taktiež samozrejme s vylepšeným, presnejším alebo rýchlejším algoritmom na rozpoznávanie tónov a akordov.

Kapitola 7

Záver

7.1 Zhrnutie výsledkov

V práci som sa venoval vytvoreniu aplikácie na rozpoznávanie tónov a akordov pre mobilnú platformu Android. Pri návrhu aplikácie som študoval rôzne metódy prevodu audio signálu na rozpoznané tóny a akordy. Vybral som metódu FFT, ktorá spĺňala kritéria rýchlosti rozpoznávania a šetrí obmedzené zdroje u mobilného zariadenia. Natrénoval som indexy tónov, ktoré sa nachádzali v spektre hlasitostí po vypočítaní FFT. Spektrum hlasitostí som normalizoval a vyhľadal lokálne maximá. Trénovanie prebiehalo na generátore frekvencií a na záver sa nepresnosti vyladili na hudobných nástrojoch. Indexy tónov sa podľa natrénovaných informácii previedli na reálne tóny. Reálne tóny sa previedli na akordy, ktoré odpovedajú hudobnej teórii.
Testovanie prebiehalo cez reálne hudobné inštrumenty a cez kvalitný audio systém. Údaje sa porovnávali s danou partitúrou k danej etude alebo hudobnej skladbe. Výsledné hodnoty testovania na presnosť rozpoznávania sú v sekcii 6.2. Prácu som podrobil aj užívateľským testom, ktoré sú popísané v sekcii 6.3.
Mobilná aplikácia má ľahko pochopiteľné užívateľské rozhranie, ktoré zobrazuje dané rozpoznané tóny a ďalšie informácie o snímanom audio signáli. Práca taktiež umožňuje hudobné nápady ukladať a hocikedy otvárať v audio a grafickom formáte. Grafický formát sa podobá notovej osnove z hudobnej teórie pre ľahšiu čitateľnosť hudobníkovi a používateľovi.

7.2 Možnosti ďalšieho vývoja

Výsledky testovania na presnosť rozpoznávania tónov a akordov odhalili mnoho nedostatkov. Tieto nedostatky by sa dali postupom času eliminovať a robiť algoritmus stále presnejším. Možnosti zlepšenia by mohli byť v hlbokých frekvenciách, podpore viac nástrojov a rýchlejšom či stabilnejšom algoritme. Do budúcna verím, že Android zariadenia budú obsahovať výkonnejší hardware, čo značne ovplyvní celú aplikáciu vrátane samotného rozpoznávania tónov a akordov.
Z pohľadu užívateľských testov by bolo možné zapracovať na nových funkciách aplikácie, stabilnejšieho grafického rozhrania, lepších informácií o nahrávaní a otváraní súborov. Ďalej do budúcna by som chcel pridať novú funkciu, a to vkladanie súborov z telefónu, ktoré neboli vytvorené v našej aplikácii.

Literatúra

[1] BEDNAŘÍK, Jan. *Převod záznamu piana z wav do MIDI.* BRNO, 2013. Bakalárska práca. Vysoké učení technické v Brne. Vedoucí práce Ing. Igor Szőke, Ph.D.

[2] SOUDEK, Ondřej. *Rozpoznání akordů.* Brno, 2015. Bakalárska práca. Vysoké učení technické v Brne. Vedoucí práce Doc. Dr. Ing. Jan Černocký.

[3] BISHOP, Christopher M. *Neural networks for pattern recognition.* New York: Oxford University Press, 1995. ISBN 978-0198538646.

[4] KARBAN, Pavel. *Rýchla Fourierova transformace.* Plzeň, 2011. Dostupné tiež z: http://home.zcu.cz/~panek50/TEVS/12_prednaska.pdf.

[5] OSGOOD, Brad. *The Fourier Transform and its Applications.* Stanford, 2007. Dostupné tiež z: https://see.stanford.edu/materials/lsoftaee261/book-fall-07.pdf.

[6] LEVINE, Mark. *The Jazz theory book.* Petaluma, Calif: Sher Music, 1995. ISBN 18-832-1704-0.

[7] BLAHUT, Richard E. *Fast algorithms for signal processing.* Oxford: Oxford University Press, 2010. ISBN 978-0521190497.

[8] SMITH, Julius O. *Mathematics of the discrete fourier transform (DFT): with audio applications.* 2nd ed. s.l.: W3K Publishing, 2007. ISBN 978-0-9745607-4-8.

[9] Weisstein, Eric W. *Fast Fourier Transform.* Dostupné tiež z: http://mathworld.wolfram.com/FastFourierTransform.html.

[10] SMITH, Dave a Jeff. FRIESEN. *Android recipes: a problem-solution approach.* Second edition. New York: APRESS, 2012. ISBN 978-1-4302-4614-5.

[11] JACKSON, Wallace. *Android apps for absolute beginners.* 2nd edition. New York: APRESS, 2012. ISBN 978-1-4302-4788-3.

[12] PAWERA, Norbert. *Practical Recording 1: Microphones.* London: Division of Music Sales Limited, 2010. ISBN 978–0–85712–245–2.

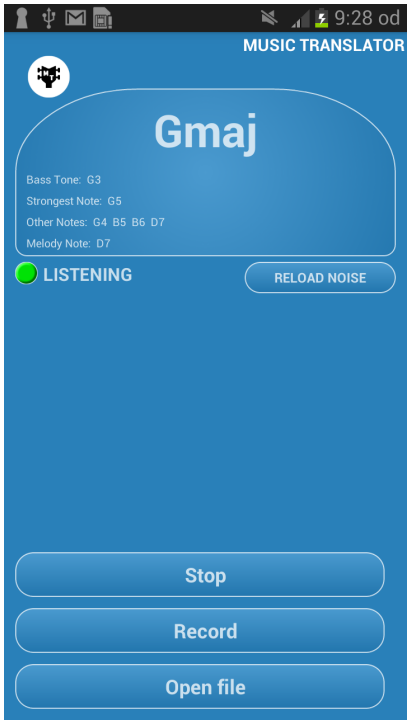
[13] KRUPPA, Dominik. *Program na výuku hudobnej harmónie.* Brno, 2011. Diplomová práca. Masarykova univerzita.

[14] WAGGENER, William N. *Pulse code modulation techniques: with applications in communications and data recording.* New York: Van Nostrand Reinhold, c1995. ISBN 04-420-1436-8.

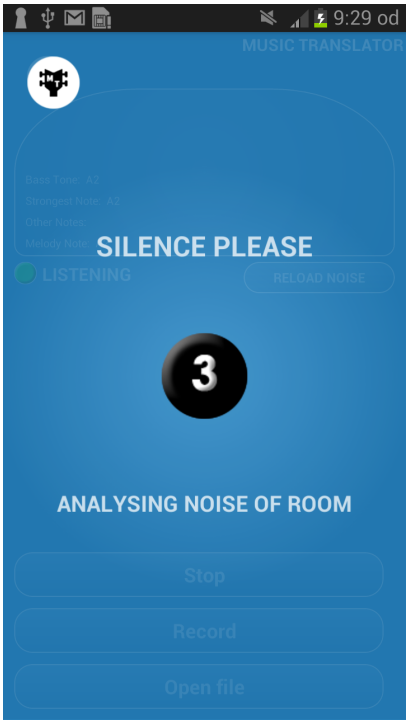
Prílohy

Príloha A

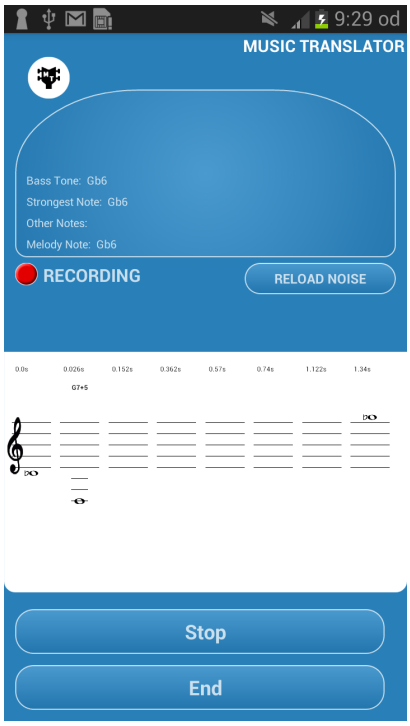
Snímky obrazovky aplikácie a obrázky



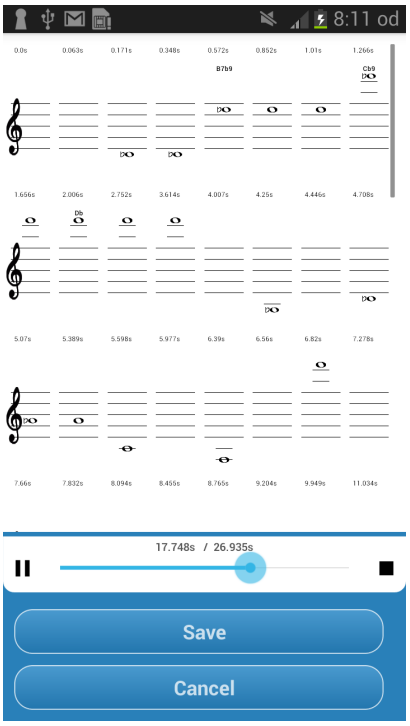
Obr. A.1: Hlavná obrazovka



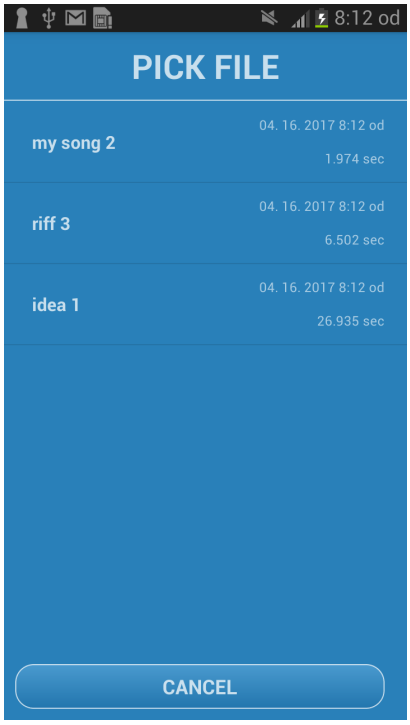
Obr. A.2: Obrazovka snímania hluku miestnosti



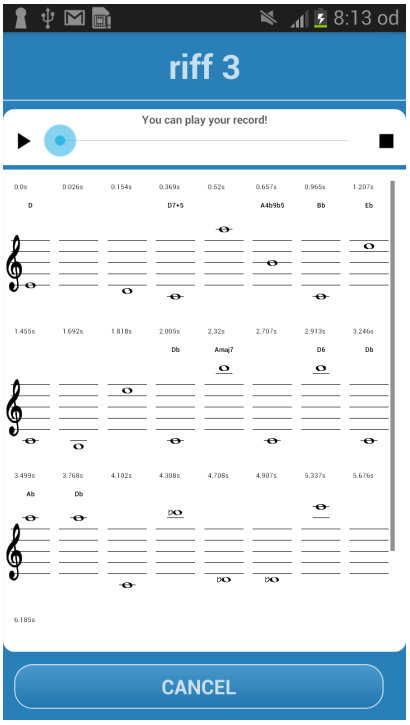
Obr. A.3: Nahrávanie zvuku v reálnom čase



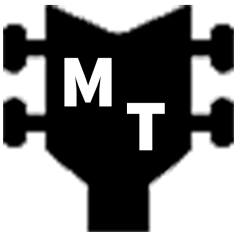
Obr. A.4: Prehrávanie hudobnej idei po ukončení nahrávania



Obr. A.5: Obrazovka uložených záznamov



Obr. A.6: Otvorenie uloženého záznamu



Obr. A.7: Ikona aplikácie

Obr. A.8: QR kód aplikácie

Príloha B

Obsah CD

Na priloženom CD sa nachádzajú nasledovné adresáre a súbory:

- **dotaznik** – PDF súbor užívateľského testu
- **MusicTranslator** – adresár pre Android Studio, obsahujúci zdrojové kódy a prostriedky aplikácie
- **doc** - adresár, obsahujúci túto prácu
- **texts** - adresár so zdrojovými kódmi práce (LaTeX)
- **video** - adresár s krátkym videom, demonštrujúcim aplikáciu
- **readme** - textový súbor s informáciami o CD a práci