



MULTIMÉDIA 2016/2017

ONLINE ANALÝZA ZVUKU

Bc. Tomáš Hudziec, xhudzi01

Bc. Matúš Dobrotka, xdobro12

BRNO, 7.5.2017

1 Zadání

Výsledkem projektu by měla být aplikace pracující jako ladička zvoleného hudebního nástroje (kytary, houslí apod.) Aplikace by měla v reálném čase vzorkovat vstupní zvukový signál a zjišťovat jeho frekvenci, indikovat ji a indikovat též nejbližší ladicí tón a odchylku. Aplikace by měla mít i možnost kalibrace na “standardní stupnici” i “relativní ladění” se zadaným/zahraným vztažným tónem. Pro snadné používání programu by aplikace měla mít intuitivní grafické uživatelské rozhraní. Program by měl umět též generovat/přehrávat referenční tóny.

2 Řešení

Jako laděný hudební nástroj jsme zvolili kytaru. Nicméně aplikace není závislá pouze na frekvencích kytarových strun a umožňuje tak ladit i jiné nástroje s tóny jiných frekvencí.

2.1 Technologie

Aplikace je implementována v jazyce Java s uživatelským rozhraním založeným na platformě JavaFX a je určena pro operační systém Windows. Toto omezení vyplývá z použití zvukového ovladače ASIO od firmy Steinberg [1], který je určen primárně pro Windows. Ovladač je určen pro profesionální práci s audiem a jeho výhodou je nízká latence.

Dále jsou aplikací použity následující knihovny:

- JAsioHost [2], rozhraní pro komunikaci s ovladačem ASIO v jazyce Java, nahrazuje dnes už zastaralé Java Sound API
- JTransforms [3], otevřená knihovna v jazyce Java, poskytuje rychlé implementace Fourierovy a dalších transformací

2.2 Vzorkování signálu

Signál z mikrofону je zvukovou kartou periodicky snímán a převáděn na číselné hodnoty. Ty jsou nahrány do vyrovnávací paměti (bufferu) a přes ovladač předány aplikaci. Výchozí velikost bufferu je 512 vzorků a lze ji měnit v nastavení ovladače.

Pro další analýzu (diskrétní kosinovou transformaci, zkr. DCT, viz dále) jsou vzorky signálu shromažďovány do většího kruhového bufferu, do kterého jsou vkládány prokládaně. Tento postup byl přejat z [4, s. 18]. Pro tuto aplikaci však postačuje pouze jeden buffer místo osmi používaných ve zmíněném zdroji. Velikost tohoto bufferu byla ponechána na 16384 položkách, kdežto vzorkovací frekvence byla nastavena na 8kHz, kvůli co nejmenší rozpoznávací odchylce tónu. Tato odchylka souvisí s rozlišitelností DCT a lze ji spočítat jako $8000 / 16384 = 0,488\text{Hz}$.

Velikost vzorku je 24 bitů (= 3 bajty), z čehož je možno vypočítat časový interval, ve kterém se periodicky provádí analýza signálu: $16384 \cdot 3 / (8 \cdot 8000) = 768\text{ms}$ + doba analýzy, což podle měření vychází přibližně na 1 sekundu.

2.3 Detekce základní frekvence

Nad navzorkovaným signálem s aplikovanou okenní funkcí se provede diskretní kosinová transformace, a z jejího výsledku se zjistí základní frekvence signálu. Detekovaná frekvence snímaného signálu se v aplikaci zobrazuje co 1 vteřinu.

2.3.1 Předzpracování okenní funkcí

Aby spektrum signálu počítané DCT nebylo zatíženo chybami, na signál se před transformací aplikuje okenní funkce. Byla zvolena Hannova okenní funkce [5]:

$$w(n) = 0.5 \left(1 - \cos\left(\frac{2\pi n}{N-1}\right) \right).$$

Každý vstupní vzorek $x(n)$ je pak vynásoben svým koeficientem $w(n)$ a výsledkem je výstupní vzorek $y(n)$. Ve výsledku budou sníženy hodnoty vzorků, které jsou blíže k okrajům.

2.3.2 Diskretní kosinová transformace

Diskretní kosinová transformace (DCT) je operace, která převádí signál z časové do frekvenční oblasti. Je příbuzná známější diskretní Fourierově transformaci (DFT), ale na rozdíl od ní produkuje pouze reálné koeficienty. To je důvodem, proč byla zvolena pro tuto aplikaci, není totiž potřeba komplexní čísla přepočítávat na jejich reálné velikosti a ušetří se výpočetní výkon.

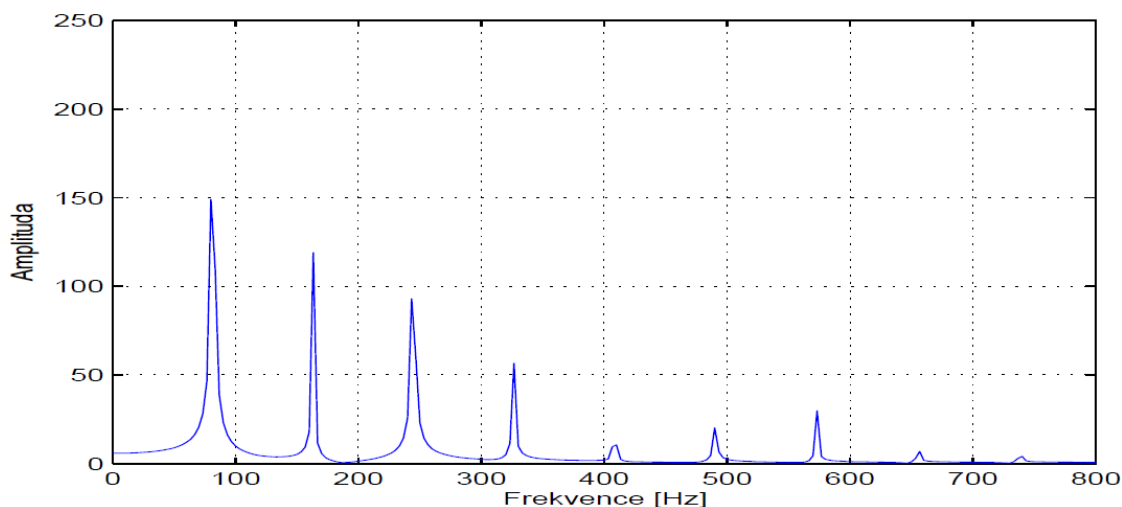
Výsledný koeficient DCT je dán součtem všech vstupních koeficientů vynásobených kosinem

$$X_k = \sum_{n=0}^{N-1} x_n \cos\left[\frac{\pi}{N} \left(n + \frac{1}{2}\right) k\right] \quad k = 0, \dots, N-1.$$

členu určujícího pořadí každého koeficientu:

2.3.3 Analýza spektra

Výsledkem DCT je frekvenční spektrum vstupního signálu. Je to závislost energie na frekvenci. Každý signál lze složit z více kosinových vln o různých frekvencích a spektrum nás informuje,



Spektrum nejhrubší kytarové struny E2 (82,41Hz), obrázek převzat z [4, s. 11].

jakou amplitudu má dílčí kosinová vlna o dané frekvenci.

Zvuk kytarové struny lze také složit z mnoha kosinových vln o různých frekvencích. Ve spektru je pak více nenulových hodnot, které je třeba zkoumat, viz obrázek.

První špička ve spektru je základní frekvencí tónu v časové oblasti. Další špičky jsou frekvence tzv. (vyšších) harmonických, jsou to celočíselné násobky základní frekvence, na kterých současně kmitá kytarová struna.

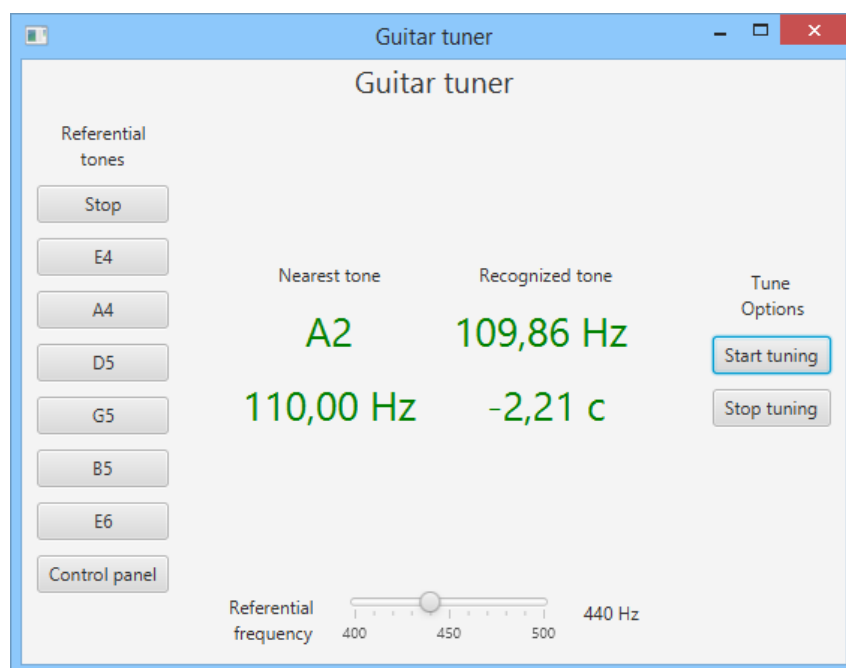
Pro zjištění základní frekvence tónu lze použít více přístupů. Nejjednodušším je výběr největší hodnoty ze spektra. Musí ale platit předpoklad, že základní tón má největší energii. To však není vždy pravdou, někdy má největší energii první nebo dokonce druhá harmonická.

Proto se zavádí metoda zvaná Harmonic Product Spectrum (HPS) [6], což je součin prvních n harmonických, ze kterého se pak vybírá maximum. Metoda byla v této aplikaci vyzkoušena, ale paradoxně nepřinesla lepší výsledky. Jako maximum byly častěji označovány právě násobky základní frekvence tónu.

Dalším pokusem byla modifikace metody HPS, kde byla operace součinu nahrazena součtem. Tato úprava vykazovala přibližně stejné výsledky jako metoda jednoduchého maxima, a proto bylo použito jednoduché maximum, které je výpočetně nejméně náročné.

2.4 Uživatelské rozhraní

Náš program je tvořen jedním oknem, které obsahuje všechny potřebné informace, jak lze vidět na následujícím obrázku.



Na levé straně tohoto okna se nachází seznam tlačítek pro generování referenčních tónů představujících jednotlivé struny kytary pro ladění podle sluchu. Protože hluboké tóny odpovídající tlustým strunám kytary nejsou dobře z reproduktoru slyšet, všechny tóny byly přeneseny o 2 oktávy výše. Tento přesun nevadí pro ladění, tóny posunuté o oktávu jsou ve znění velmi podobné a dá se podle nich ladit. Výsledné tóny tedy jsou E4, A4, D5, G5, B5, E6 v pořadí od nejtlustší struny až po nejtenčí. Kromě těchto tlačítek se zde nacházejí ještě 2 tlačítka - tlačítko *Stop* pro ukončení přehrávání tónu a tlačítko *Control panel* pro zobrazení nastavení použitého ovladače.

Pravá strana okna obsahuje pouze 2 tlačítka - tlačítko *Start tuning*, kterým lze spustit ladění nástroje a tlačítko *Stop tuning*, které ukončí ladící mód aplikace.

Spodní část okna obsahuje číselný posuvník pro nastavení požadované hodnoty referenční frekvence, pomocí které se počítá relativní ladění/generování tónů. Výchozí hodnota této frekvence je nastavena na 440 Hz, což představuje standardní frekvenci tónu A4 (komorní A).

Ve středu okna jsou asi nejdůležitější informace celé aplikace. Můžeme tady vidět dva sloupce - *Nearest tone* a *Recognized tone*. V prvním sloupci se zobrazují informace o nejbližším tónu, resp. půltónu vzhledem ke vstupnímu rozpoznávanému tónu. Jsou to název a frekvence nejbližšího tónu vzhledem ke zvolené referenční frekvenci.

Ve druhém sloupci se nachází frekvence rozpoznávaného vstupního tónu a její odchylka od nejbližšího tónu. Tato odchylka může být kladná nebo záporná - podle toho, zda je rozpoznávaný tón vyšší nebo nižší než jeho nejbližší tón. Odchylka je relativní a je uvedena v centech, což je setina půltónu. Pokud je odchylka dostatečně malá (menší než 10 centů), všechny tyto informace se zobrazí zeleně, jinak červeně.

2.5 Generování tónů

Aplikace umí generovat referenční tóny, které představují jednotlivé struny kytary. Pro každý tón je generován signál, jehož předpis je $\sin(2\pi \frac{f}{F_s})$, přičemž f je frekvence daného tónu vzhledem k dané referenční frekvenci a F_s je vzorkovací frekvence.

V následující tabulce jsou uvedeny frekvence jednotlivých “kytarových” tónů vzhledem k referenční frekvenci 440Hz [7].

Tón	E2	A2	D3	G3	B3	E4
Frekvence	82.41 Hz	110 Hz	146.8 Hz	196 Hz	246.9 Hz	329.6 Hz

2.6 Relativní ladění

Základní stupnici tvoří dohromady 12 půltónů, což představuje 1 oktávu. Pro lepší představu a přehled jsou tyto půltóny uvedeny v následující tabulce.[8]

Půltón	C	C#	D	D#	E	F	F#	G	G#	A	A#	B
--------	---	----	---	----	---	---	----	---	----	---	----	---

Náš projekt podporuje ladění až pro 8 oktáv tónů, tedy od C0 až po B7. Tyto jednotlivé tóny se od sebe liší svou frekvencí, kterou je možné vypočítat následujícím vztahem [9]:

$$f = \sqrt[12]{2^x} \cdot f_{ref}$$

přičemž f_{REF} je referenční frekvence tónu A4 a x je počet půltónů, které jsou mezi daným tónem a referenčním tónem A4.

Pokud bychom neuvažovali relativní ladění, znamenalo by to, že referenční frekvence bude vždycky 440 Hz. Vzhledem k tomu, že relativní ladění uvažujeme, referenční frekvence se může měnit, když uživatel změní tuto hodnotu pomocí posuvníku. V tom případě se automaticky přepočítají frekvence všech tónů a ve výsledku jsou zobrazeny hodnoty s ohledem na zvolenou referenční frekvenci. To samé platí při generování tónů. Tón je generován automaticky vzhledem na aktuální zvolenou referenční frekvenci.

2.7 Rozdělení práce

- **Tomáš Hudziec (vedoucí týmu)** - implementace mechanismu ladění (vzorkování vstupního signálu, výpočet frekvenčního spektra pomocí DCT, detekce základní frekvence prostřednictvím prostého maxima a metody Harmonic Product Spectrum), generování referenčního tónu
- **Matúš Dobrotka** - implementace relativního ladění a relativního generování tónů na základě referenční frekvence, generování referenčního tónu, grafické uživatelské rozhraní

3 Obsluha programu

Před prvním spuštěním programu je potřeba nainstalovat zvukový ovladač ASIO4ALL [10]. Samotná obsluha programu je jednoduchá. Když uživatel bude chtít ladit svou kytaru, tak po spuštění programu musí stisknout tlačítko *Start tuning*. Program indikuje zelenou barvou stav, kdy je vstupní tón dostatečně naladěný a uživatel tak může pokračovat laděním dalšího tónu. Červená barva indikuje opak, tedy vstupní tón je třeba ještě lépe doladit. Ukončit ladění lze stiskem tlačítka *Stop tuning*.

Po zastavení ladění může uživatel přehrávat libovolný z referenčních tónů. Pro ukončení přehrávání referenčního tónu je třeba stisknout tlačítko *Stop*. Pro změnu nastavení zvukového ovladače slouží tlačítko *Control panel*.

Pomocí posuvníku lze nastavovat frekvenci referenčního tónu v rozsahu 400 - 500 Hz pro relativní ladění.

4 Zhodnocení výsledků

Následující tabulka ukazuje vyhodnocení přesnosti pro nylonové a kovové struny dvou kytar.

	ref	kov	odch	nylon	odch
E4	329,63	328,61	-5,51	328,86	-4,16
H3	246,94	248,54	10,90	247,8	5,86
G3	196,00	391,36	-3,73	390,87	-6,58
D3	146,83	143,55	-39,80	143,55	-39,80
A2	110,00	109,38	-10,04	109,86	-2,27
E2	82,41	79,83	-55,78	328,61	-1,26

Odchyly jsou uvedeny v centech, což je setina půltónu. Žlutě jsou uvedeny hodnoty, kde byl změřená frekvence tónu dvojnásobná (první harmonická), a oranžově čtyřnásobná (druhá harmonická). Do šesti (nebo i deseti) centů je přesnost ladění dostačující, u některých tónů toho bylo dosaženo.

Lepšího odfiltrování vyšších harmonických by se dalo dosáhnout fungující implementací metody HPS.

5 Závěr

Projekt nám přinesl zkušenosti v oblasti (online) zpracování a analýzy zvukového signálu. Kromě toho jsme se dozvěděli i nové informace z teorie hudby. Program by mohl být základem pokročilejší aplikace, která by se mohla orientovat například na rozpoznávání kytarových akordů, nebo na ladění více druhů hudebních nástrojů se specifickějšími požadavky.

Literatura

- [1] Steinberg: Technologies [online].
Dostupný z WWW: <<https://www.steinberg.net/en/company/technologies.html>>
- [2] Martin Roth: JAsioHost [online].
Dostupný z WWW: <<https://github.com/mhroth/jasiohost>>
- [3] Piotr Wendykier: JTransforms [online].
Dostupný z WWW: <<https://github.com/wendykierp/JTransforms>>
- [4] Ondřej Soudek: Rozpoznávání akordů, bakalářská práce, Brno, FIT VUT v Brně, 2015.
Dostupný z WWW: <https://www.vutbr.cz/studium/zaverecne-prace?zp_id=88462>
- [5] Wikipedia: Hann function [online].
Dostupný z WWW: <https://en.wikipedia.org/wiki/Hann_function>
- [6] Tamara Smyth: Harmonic Product Spectrum (HPS) [online].
Dostupný z WWW:
<http://musicweb.ucsd.edu/~trsmyth/analysis/Harmonic_Product_Spectrum.html>
- [7] Seventh String Ltd.: Note frequencies [online].
Dostupný z WWW: <<https://www.seventhstring.com/resources/notefrequencies.html>>
- [8] Flutopedia: Pitch to Frequency Calculator [online].
Dostupný z WWW: <http://www.flutopedia.com/pitch_to_frequency.htm>
- [9] Michigan Technological University: Physics of music, 2015 [online].
Dostupný z WWW: <<http://www.phy.mtu.edu/~suits/NoteFreqCalcs.html>>
- [10] Michael Tippach: ASIO4ALL – Universal ASIO Driver For WDM Audio [online].
Dostupný z WWW: <<http://asio4all.com>>