

Posudek oponenta diplomové práce

Název práce: Statistical Natural Language Processing Methods in Music Notation Analysis

Řešitel: Bc. Jindřich Libovický

Vedoucí: Mgr. Nino Peterek Ph.D.

Obsah práce

Diplomová práce Jindřicha Libovického se zabývá extrakcí melodie ze zvukových souborů použitím metod spojených se zpracováním přirozeného jazyka.

Je rozdělena na tři části. První část je teoretická, věnuje se porovnávání hudby a přirozeného jazyka z několika různých pohledů. Zmiňuje množství souvisejících prací, včetně těch, jejichž snahou je hudbu automaticky zpracovávat. Druhá část se zabývá zpracováním melodie. Za tímto účelem bylo třeba vytvořit korpus melodií z dostupných MIDI souborů, ve kterých bylo nutné určit melodii hlavní. V práci je popsáno několik různých metod pro nalezení hlavní melodie a pro vytvoření korpusu byla nakonec použita kombinace dvou z nich. Další kapitola se věnuje modelování melodie pomocí n-gramového modelu. Notový zápis je pro modelování třeba přizpůsobit: místo tónů pracuje model s intervaly, místo délek pak s celočíselnými poměry. Několik variant modelů melodie bylo natrénováno pomocí toolkitu SRILM, součástí práce je i ukázka automaticky vygenerované melodie. Třetí část se zabývá extrakcí melodie přímo ze zvukového signálu. Zkouší se zde několik metod detekce základního tónu ze zvukového spektrogramu. Detekovaná melodie je následně převedena do notového zápisu. Na základě všech hypotéz ze spektrogramu a natrénovaného modelu melodie se určí nejpravděpodobnější posloupnost not. Následuje evaluace, diskuse a závěr. Práce má 66 stránek a je psána anglicky.

Hodnocení

Práce je přehledně strukturovaná a psaná dobře srozumitelnou angličtinou. Na několika místech je však příliš stručná a není jasné, co autor přesně zamýšlel. V kapitole 3.2.2 například nejsou příliš jasné definice skóre s_1 , s_2 a s a také několik dalších používaných veličin není přesně v textu vysvětleno. O tom, co znamenají názvy modelů v tabulce 3.4 jsem se například dozvěděl až v příloze práce. O stručnosti svědčí i fakt, že bez rešeršní části a příloh má práce pouze 27 stránek. V přílohách je možná až příliš velké množství velmi podobných tabulek plných těžko srovnatelných čísel. Na druhou stranu, v rešeršní části je velmi pěkně zpracovaný přehled o historii zpracování hudby a o její příbuznosti s přirozeným jazykem. Kladně hodnotím i přehlednou diskusi a závěr práce. Zdůraznil bych i velký podíl odvedené programovací práce. Jindřich všechny algoritmy pro zpracování signálu a extrakce melodie ze zvukového souboru naimplementoval sám, součástí programu je i grafický výstup, na kterém je přehledně vidět spektrum zvukového souboru, nalezené melodie i odvozených not.

Velmi hodnotným výstupem práce je korpus melodií, který obsahuje více než 400 hodin melodií rozdělených podle žánrů. Takový korpus je ideálním zdrojem pro další experimenty týkající se modelování hudby.

Otázky a náměty

- 1) Používal jste pro modelování tedy pouze desetigramy? Měřil jste něco jako “out-of-vocabulary”, čili množství destigramů v testovacích datech, které nebyly v datech trénovacích? Nezkoušel jste experimentovat i s nižšími n-gramy?
- 2) Nebylo by výhodné se při modelování melodie zaměřit více na takty? Ve vygenerovaných melodiích je vidět, že jdou noty často přes hranice taktů, což není moc přirozené.
- 3) Přijde mi, že vygenerovaná melodie má problémy s vracením se k počáteční délce i výšce not. To je důsledkem toho, že se modelují pouze intervaly a poměry délek not. Možná by pomohla kombinace s modelem, který by modeloval absolutní výšku i délku not na n-gramech nižšího řádu. Další výzkum by mohl směřovat k vytváření modelů pro danou tóninu a rytmus. To by ovšem znamenalo jejich rozpoznávání v trénovacím korpusu.
- 4) U algoritmu “Simple Peak Detection” uvádíte, že často požadovanou melodii vůbec nezaznamená. Jak k tomu může dojít? Z jeho definice mi připadá, že by tato metoda měla pokrýt všechno. I z obrázku je to patrné.
- 5) Algoritmy pro detekci základního tónu často rozpoznávají frekvence nižší než 100Hz, což jsou z hlediska hlavní melodie poměrně nízké tóny (nižší než velké G). Zkoušel jste tyto nízké frekvence odfiltrovat?

Závěr

Diplomová práce Jindřicha Libovického splnila vytyčený úkol prozkoumat možnosti zpracování hudebních melodií a popisuje mnoho zajímavých experimentů. Jednoznačně ji **doporučuji** k obhajobě.

V Praze dne 10. května 2013

David Mareček
Ústav formální a aplikované lingvistiky
MFF UK