

## **Jan Melechovský – Klasifikace hudebních žánrů pomocí MFCC**

Automatická klasifikace hudebních žánrů doznala za poslední dvě až tři desetiletí značné pozornosti, a to především díky požadavkům hudebních databází, aplikací, i některých profesionálních masteringových studií. Důvodem je mimo jiné i to, že k dosažení co nejlepšího mixu a masteringu daného žánru je třeba adekvátního přístupu a mnohdy se stává, že podklady pro zvukového inženýra neobsahují dostatečné informace a ne každý interpret uvádí hudební žánr, který prezentuje. Je více než pravděpodobné, že vývoj v této oblasti bude podobným tempem pokračovat i nadále a detekce hudebních žánrů se tak bude neustále zdokonalovat, bude nabírat na větší komplexnosti parametrů a jejich klasifikací a v závislosti na tom by mělo být dosahováno vyšší úrovně úspěšnosti výstupu. Z výše uvedených důvodů kvituji zájem o obeznámení a studii dílčích aplikací, které by v budoucnu ke zmíněné komplexní a velmi přesné detekci žánrů mohly přispět.

Důkazem takového zájmu je i práce p. Jana Melechovského týkající se klasifikace hudebních žánrů s využitím spektrálního parametru MFCC.

Předem bych rád uvedl, že z písemné práce je jasně vidět, že je míra obeznámení diplomanta s problematikou velmi uspokojivá, a dokazuje se tak patřičná zodpovědnost a pečlivost v přístupu k danému tématu. Struktura je přehledná, prezentace a výčet jednotlivých parametrů a metod možných pro použití při klasifikaci je dostačující. Jak již bylo zmíněno, autor v práci využívá spektrálního parametru MFCC (Mel-Frequency Cepstral Coefficients). Detekovaný signál, tedy určitá píseň/skladba je rozdělena do segmentů, které jsou dále štěpeny do minisegmentů a pro tyto minisegmenty je vypočteno několik MFCC. Následně jsou koeficienty zpracovány pomocí aritmetického průměru a směrodatné odchylky a k finálnímu rozdělení do osmi žánrových kategorií je užito metody k-NN (k-Nearest Neighbors). Podobného přístupu využívá ve své bakalářské práci také p. Jaroslav Bartoš (*Detekce hudebních žánrů pro účely masteringu gramofonových desek*, Praha 2016), se kterou p. Melechovský výsledky své práce srovnává. Pozitivním zjištěním je, že v práci p. Melechovského je dosaženo vyšší úspěšnosti, a to i při použití většího počtu žánrů. Velmi sympatické je i rozdělení klasifikace do několika testů (kap. 4, oddíl 4.4 – 4.10), z čehož je jasně zřetelný zájem autora o důkladné otestování zvoleného systému. Neméně přínosný je v oddílu 4.10 pokus o zrychlení algoritmu, a to z důvodu časové úspory – v případě většího penza vstupních dat je tato úspora velice vítána. Při úspěšnosti 67,2% (což je o pouhých 1,6% méně, než u pomalého režimu) je použití rychlého algoritmu velmi efektivní. Zvolená metoda klasifikace hudebních žánrů tak přináší zajímavé výsledky a systém detekce značně vylepšuje.

(pokr. na další str.)

.....  
podpis oponenta

Nyní bych rád uvedl poznámky k jednotlivým úsekům práce, které by mohly vést k zamyšlení, či dostát bližšího objasnění:

1. V úvodu se uvádí, že „*obsahuje-li hudba nepřírozené skoky, může dojít k poškození desky, v horším případě k poškození řezacího zařízení.*“ **Co je myšleno pod pojmem „nepřírozené skoky“?** Dále, pokud autor zmiňuje tato rizika, měl by být obeznámen s jejich konkrétním projevem – tedy k jakému poškození desky dochází a jak by mohlo být poškozeno řezací zařízení. (Nutno podotknout, že řezací zařízení bývají doslova „prošpikovávána“ řadou pojistek a záchraných systémů, takže pravděpodobnost jejich poškození v souvislosti s přenášeným signálem je velmi malá.)
2. V kapitole 2 (oddíl 2.2 MFCC) je v druhém odst. uvedeno, že „*Mel-frekvence je škála frekvence*“. Nejedná se zrovna o povedený češtinářský obrat – možná by bylo lepší použít sousloví „škála frekvencí“, a ještě lépe „frekvenční škála“. Není to však nijak závažný nedostatek, pouze doporučení do budoucna. V určitých situacích by to mohlo znamenat nezřetelný význam takového vyjádření, a tím i devaluaci informativní hodnoty.
3. V témže oddílu (2.2 MFCC) na str. 19 je uvedena „*KL transformace*“. Bylo by vhodné ji taktéž uvést v seznamu použitých zkratk.
4. V oddílu 2.3.2 autor zmiňuje parametr „*Spectral Rolloff*“. **Jeho popis není zcela zřetelný** a proto bych poprosil o ústní doplnění, aby bylo více jasné, o jaký parametr se jedná a co vyjadřuje. Navíc v případě, že je zmiňován parametr „*rolloff*“, nabízí se i **uvedení parametru „spectral brightness**“. Do budoucna doporučuji pro úplnou přehlednost k popisu takových parametrů přikládat i jejich grafické znázornění ve spektrálním zobrazení.
5. V kapitole 3 jsou zmíněny dvě metody klasifikace – metoda s učitelem a metoda bez učitele. Každá metoda má určitou charakteristiku a hodí se pro odlišný přístup ke klasifikaci získaných parametrů, tedy i k docílení lehce odlišných výsledků. Výhod a nevýhod obou metod se můžeme dovtípit z tvrzení z následujících odstavců, avšak pro lepší přehlednost by bylo dobré u jejich uvedení na začátku kapitoly **výhody a nevýhody obou metod krátce popsat**.
6. Obrázky v příloze, které graficky znázorňují klasifikaci jednotlivých žánrů jsou velmi zajímavé a hodilo by se více takových (minimálně od každého žánru jeden). Zároveň by nebylo špatné zvolit takové písně, jejichž forma je jasně vidět z vlnového průběhu, a pak by bylo více zřejmé, jak se klasifikace jednotlivých žánrů mění v závislosti na strukturálním oddílu skladby (sloka, refrén, bridge, atd.) – je jasné, že u některých žánrů tohle není jednoduše proveditelné. Vzhledem k tomu, že se jedná o grafický výstup se zajímavou informační hodnotou, bylo by také dobré pro lepší přehlednost tato okna s grafickým výstupem horizontálně rozšířit.

Závěrem bych rád vyzdvihl cenný přínos zvolené metody a přístupu ke klasifikaci parametrů. Výsledky dosažené tímto postupem jsou více než uspokojivé. Zadáni práce bylo splněno a diplomant dosáhl signifikantního vylepšení systému klasifikace hudebních žánrů, zároveň byla nastíněna i možná vylepšení implementované metody. Drobné nedostatky se v práci tu a tam objevují (především v přehlednosti vyjádření u některých termínů), ovšem informativní hodnotu této práce nijak závažně neovlivňují.

Vzhledem k výše uvedenému navrhuji klasifikaci práce **stupněm B (velmi dobře)**.