

Multimédia - MUL 2010 / 2011

Automatický nástroj pro digitalizaci MC záznamů

Petr Mrázek
Miroslav Dvořák

xmraze03@stud.fit.vutbr.cz
xdvora11@stud.fit.vutbr.cz

Cíl projektu

Cílem projektu je vytvoření co možná nejjednoduššího nástroje, který uživateli umožní digitalizaci analogového zvukového záznamu na MC nosiči do formátu WAV nebo MP3. Důraz je přitom kladen na co možná nejjednodušší použití a snadné ovládání. Cílovou skupinou jsou uživatelé, kteří nedisponují potřebným know-how, pro použití komplexnějších aplikací (Soud Forge, Gold wave, Audacity aj.), případně nemají čas manuálně zpracovávat záznamy a hledají co možná nejautomatizovanější nástroj.

Jak nástroj funguje

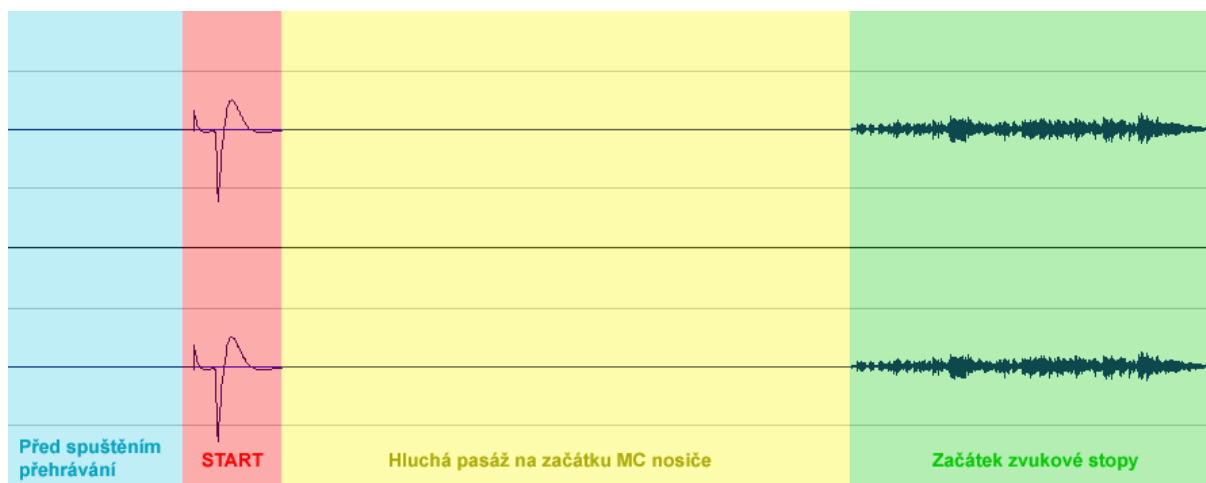
Předpokladem je vlastnictví přehrávače MC nosičů se sluchátkovým výstupem, zvuková karta schopná zachytávat zvuk a propojovací line-in kabel (jack 3,5 samec -> jack 3,5 samec). Obsluha nástroje je velmi jednoduchá, po spuštění aplikace stiskem jediného tlačítka se „odjistí“ nahrávání a pak již aplikace pouze čeká, až jí budou přehrávány záznamy - sama detekuje, kdy nahrávání spustit a kdy ukončit. Prací uživatele je jeden stisk tlačítka a následné přehrávání „kazet“ aplikací. Po dokončení těchto úkonů najde ve vybraném adresáři výsledné zvukové soubory (WAV / MP3).

Co nástroj poskytuje

1. automatická detekce začátku a konce strany kazety
2. automatické ukládání jednotlivých záznamů ve formátu WAV / MP3
3. normalizace záznamů*
4. odstranění hluchých pasáží na začátku a konci záznamů*

Řešení některých popsanych funkcí

- **automatická detekce startu záznamu** – každé spuštění přehrávání na magnetofonovém přehrávači spočívá v přiložení čtecí hlavy na pásku – toto je doprovázeno velice specifickým projevem ve výstupním / vstupním signálu, tzv. lupancem – viz obrázek.



Úsek signálu podbarvený červeně je jednoznačně odlišitelný od okolního průběhu (modrá a žlutá část). Proto stačí po „odjištění“ aplikace sledovat hodnotu energie signálu za určitý časový úsek (pár desítek ms) – což není nic jiného než prosté sečtení absolutních hodnot jednotlivých sampleů v daném časovém okně. Podnětem pro začátek zaznamenávání zachytávaného zvuku je náhlé zvýšení energie signálu oproti předchozím oknům. Projev zastavení nahrávání, kdy hlava odskočí od pásky, je srovnatelný s projevem přiložení hlavy a jeho detekce probíhá stejně – jedná se o rychlý nárůst hodnoty energie signálu v porovnání s energií v okolí, která je téměř nulová.

- **Normalizace zvuku** – ve výsledku znamená upravení hladiny zvuku na „normální“ mez. Pro tuto proceduru existují 2 důvody, jeden čistě praktický a druhý vyplývá z metodiky zpracování analogového signálu na jeho diskrétní digitální reprezentaci. Praktickým důvodem je myšleno, že při přehrávání několika skladeb je nepříjemné, když mezi nimi kolísá hlasitost. Tím druhým důvodem je, že zaznamenaný zvuk je zaznamenaný v určitém „rozlišení“ – to se skládá ze vzorkovací frekvence a počtu bitů, na nichž je reprezentována hodnota jednoho vzorku (sample). Normalizací se tedy snažíme dosáhnout stejné hlasitosti u všech nahrávek a také plně využít možnosti formátu, ve kterém zaznamenáváme. V praxi se tedy jedná o průchod nahrávky, nalezení maximálních hodnot, kterých tento signál dosahuje – hodnoty sampleů - a následného vypočítání koeficientu, jehož aplikací na jednotlivé vzorky signálu dosáhneme chtěných hodnot. Zde je vhodné podotknout, že správně nastavená hladina zvuku při nahrávání je velmi kritická pro kvalitu výsledné nahrávky. Pokud je signál přebuzený, jeho přesahující hodnoty jsou ořezány na MAX hodnotu, které vzorek může dosáhnout, naopak pokud je příliš tichý, nevyužíváme plně možností záznamového formátu a následná normalizace sice upraví hladinu hlasitosti, ale výsledný efekt by se dal přirovnat ke zvětšování rozlišení u obrázků v malém rozlišení – obrázek se sice zvětší, ale jeho kvalita není taková, jako by byl v daném rozlišení původně zachycen.
- **Odstranění hluchých pasáží ze začátku a konce záznamu** – postup je velmi zřejmý z obrázku, znázorňujícího začátek skladby. Na začátku záznamu máme spouštěcí lupanec, pak signál s relativně nízkou energií a na začátku zvukové stopy energie prudce vzroste, zde je možné všechny sampley, které se vyskytují před zvýšením energie signálu (začátek zvukové stopy) zahodit. Mírně obrácený postup je aplikován i pro ořezání koncové hluché pasáže.

Implementace

Aplikace byla vyvinuta v jazyce C++ za pomoci knihoven Qt4, mp3lame a sndfile. Jedná se o klasickou Qt GUI aplikaci. Nahrávání je řešeno pomocí QtMultimedia. Zvuk je nahráván do dočasných souborů ve formátu signed 16bit raw PCM. Po ukončení nahrávání jsou tyto soubory zpracovány do finální podoby - buď klasický WAV, nebo MP3. Zpracování probíhá v samostatném vlákně (jejich počet není omezen, ale v praxi nikdy není více jak jedno aktivní). Bylo by triviální přidat podporu formátů OGG Vorbis a FLAC, protože použité knihovny s nimi umí pracovat.

Překlad je řešen pomocí překladového systému CMake a funguje pod OS Windows (s překladačem MinGW32 a Qt4 SDK 4.7.3) a Linux (GCC, Qt4 4.7.x, libsndfile, libmp3lame). Zdrojový kód odevzdaný do IS kvůli velikosti neobsahuje knihovny pro Windows.

Kompletní zdrojový kód projektu - i s knihovnami - je možné získat z webu github:

<https://github.com/peterix/tapedeck>

Sestavení (Linux):

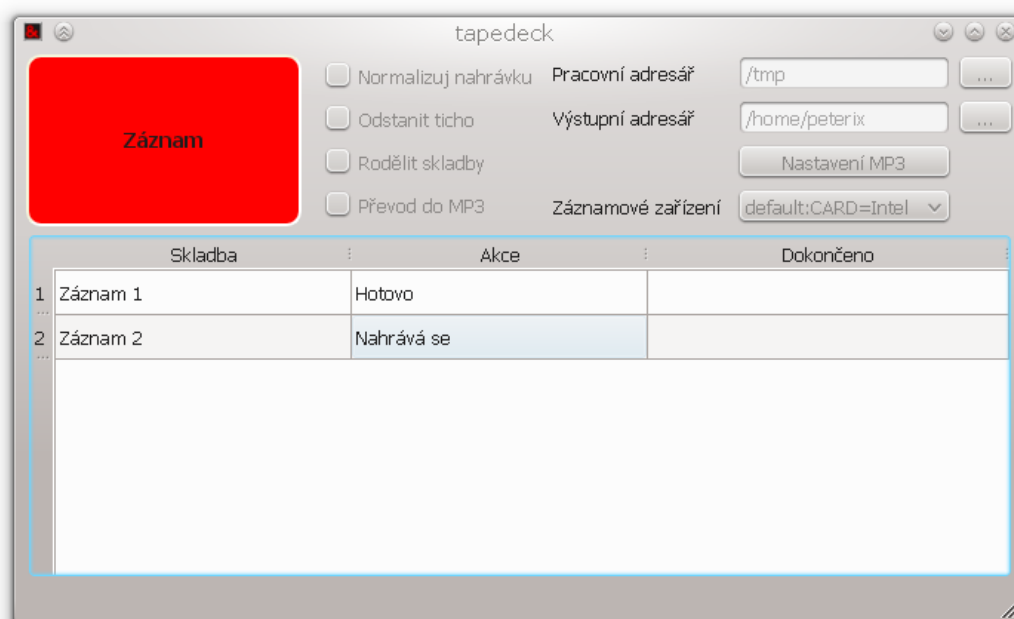
Je potřeba mít knihovny Qt 4.7, libsndfile a lame3 v systému. Pak stačí spustit následující:

```
cd build
cmake ..
make
```

Program je pak uložen ve složce output.

GUI

Jednoduché uživatelské rozhraní, v základu stačí spustit jediným tlačítkem „záznam“. Existuje samozřejmě také možnost nastavit cesty pro ukládání jak dočasných, tak výsledných souborů. Nastavit kvalitu MP3 komprese a další.



Rozšíření dostupná ve finální verzi (mimo-školní projekt)

- Filtrace šumu – jelikož aplikace je schopna získat vzorek jak šumu vznikajícího v propojení (úsek před stiskem tlačítka play na přehrávači), tak vzorek šumu vznikajícího na čtecí hlavě (úsek mezi stiskem tlačítka play a začátkem zvukové stopy), je možné implementovat filtraci šumu.
- Detekce přebuzení – aplikace může během záznamu kontrolovat hodnoty, které zapisuje a v případě přebuzení nahrávky (ořezávání samplů na hodnoty MAX), je možné uživatele upozornit na příliš silný vstup a požádat o opakování nahrávání. Stejným způsobem by mohla být detekována, po určité době nahrávání, i příliš nízká úroveň vstupního signálu a uživatel by byl na tuto skutečnost upozorněn a v zájmu zvýšení kvality nahrávky, požádán o její opakování se zesílením vstupu.
- Rozdělení skladeb – aplikace je prvoplánovitě vyvíjena pro digitalizaci záznamů setů, kde je záznam velmi pravděpodobně kontinuálního charakteru. Ovšem počítá se s rozšířením i pro posluchače ostatních hudebních žánrů a tedy dělení skladeb pomocí detekce mezer mezi nimi.