Záznamový software pro ASR systém LASER

Josef Chlad

Vedoucí práce: Ing. Kamil Ekštein, Ph.D

Cíl práce

- SW pro pořizování korpusových nahrávek
- ukládání dat ve formátu WAV a PCM
- analýza, ořezání, normalizace a kontrola
- multiplatformní v rámci Win 32/64 a Linux
- GUI – zobrazení nahrávky, jednoduché ovládání
- správa pořízených nahrávek

Stávající řešení

- platforma Windows
- pouze v anglickém jazyce – Linguist
- chybějící normalizace
- nestandardní konfigurace pomocí .ini souboru
- neumí ukládat ve formátu WAV
- neobsahuje detekci případů, kdy mluvčí začne mluvit dříve než je spuštěn záznam, avšak v době začátku záznamu toto zvukovým signálem oznámí

Řešení

- výběr jazyka C++ / QT
- normalizace intenzity a délky šumu na okrajích nahrávky
- detekce Clippingu, špatného začátku/konce nahrávky
- kontrola dynamiky signálu
- konfigurace v programu a XML
- struktura rozložení GUI

Závěr a zhodnocení

- funkční software
- vhodný výběr programovacího jazyka
- dodrženy podmínky pro zpracovávání a ukládání signálu, vzhledu aplikace a možnosti nastavení

- nedostatečná modularita, oddělenost vrstev
- podrobnější uživatelská příručka

Otázka 1.

- Z jakého důvodu je možné klávesovou zkratkou ovládat pouze spouštění/pozastavení nahrávání a nikoliv také ostatní funkce aplikace?

Jelikož není v ComboBoxu pro nastavení velikosti písma zakázáno ruční zadání, a události stisku klávesy tedy odchytává tato komponenta.

Otázka 2.

- Při spuštění záznamu vykazuje aplikace značné zpoždění odezvy (cca 500 – 700 ms), čím je to podle Vás způsobeno a jak by bylo možné tento problém odstranit?

Může být způsobeno grafickými úpravami nad GUI, které jsou prováděny před samotným záznamem. Možností, jak zpoždění odstranit, může být také optimalizace na úrovni kompilace, tj. např. uložení často používaných proměnných do registru procesoru. V mých podmínkách takovéto zpoždění pozorováno nebylo.