

# MULTIMEDIA BROWSER P.R.A.S.E.

**Jakub Kubalík**

Master Degree Programme, FIT BUT

E-mail: xkubal05@stud.fit.vutbr.cz

Supervised by: Petr Schwarz

E-mail: schwarzp@fit.vutbr.cz

## ABSTRACT

Article describes multi-modal browser which is being developed for browsing and searching audio-video archives. Main focus is given to description of architecture of whole system - primary to the system of components and messaging. My contribution and future directions are presented in conclusion.

## 1 ÚVOD

Prvně by bylo dobré se zmínit, co představovaný multimediální prohlížeč je a proč projekt vznikl. Jak už název napovídá, půjde o prohlížeč různých typů multimediálních dat. Nejedná se však o prohlížeč, jak ho obvykle známe, pro prohlížení určitého specifického typu multimediálních dokumentů. Prohlížeč umožňuje zobrazit více typů multimediálních dokumentů v jednom okně, popř. v jedné aplikaci s kterými je samozřejmě možné interagovat. Navíc je možné tyto dokumenty svázat určitým způsobem, obvykle časem tak, že se uživateli jeví jako jeden multimediální dokument. To umožňuje vytvářet jakési pokročilejší multimediální prezentace s možností přímé interakce s uživatelem.

Projekt vznikl hlavně pro potřeby výzkumné skupiny Speech působící pod záštitou Fakulty informačních technologií Vysokého učení technického v Brně. Výzkumná skupina pracuje na řadě projektů zabývajících se zpracováním řeči a přirozeného jazyka, kde výstup jejich snažení většinou nemá přímo grafickou podobu a je proto nutné jejich výsledky nějakým způsobem prezentovat.

Co se týče jména projektu, není zatím příliš ustálené a jednotné. Oficiální název projektu je P.R.A.S.E., který představuje zkrácení slov „Presentation As Synchronized Experience“. Název však svádí k zkrácené výslovnosti, proto jej v textu příliš nepoužívám. Jako vývojové jméno nejčastěji používáme MBrowser, což zkracuje „Multimedia Browser“ nebo také „Multi-modal Browser“.

## 2 ZÁKLADNÍ PARAMETRY

Nejprve by bylo dobré zmínit některé základní vlastnosti projektu. Vstupními požadavky při vývoji prohlížeče byly hlavně:

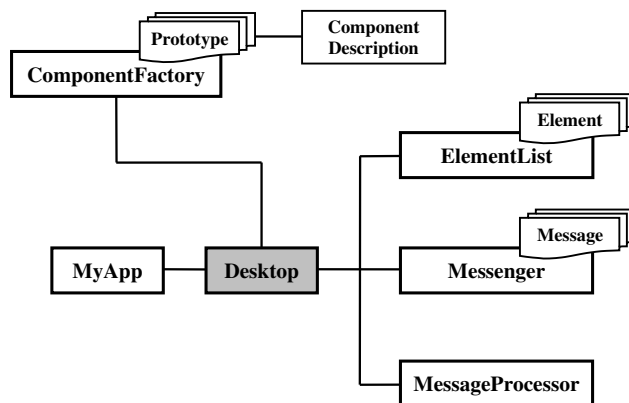
- Snadno konfigurovatelný prohlížeč.
- Snadno rozšiřitelný o nové technologie a nové funkce.

- Zpracování toku multimediálních dat (např. přehrání video či audio záznamů, dále zobrazování snímků prezentací případně různých prepisů řečových rozpoznávačů nebo jejich výstup).
- Možnost on-line distribuce dat (živé přednášky — zpracování výstupu rozpoznávačů v reálném čase).
- Vyhledávač a vývojové nástroje s otevřenou architekturou a otevřenými zdrojovými kódy.

Pro usnadnění konfigurace je na většině vhodných míst prohlížeče využit formát XML s jednoduchou strukturou, tj. hlavní konfigurace prohlížeče, jednotlivých komponent i vstupní data jako např. textové prepisy, specifikace sekvence snímků prezentace nebo třeba popisky časové osy — to vše je definované ve formátu XML. Snadnou rozšiřitelnost prohlížeče umožňuje v celku jednoduché komponentové rozhraní na kterém je celá aplikace založena. To však na druhou stranu vyžadovalo propracovanější jádro. Z hlediska otevřenosti a přenositelnosti je aplikace vyvíjena v jazyce C++ pro standardní GCC překladač a pro grafické rozhraní je použita knihovna wxWidgets, která je také platformově přenositelná a z programátorského hlediska vhodná pro tento projekt. Některé požadavky však stále nejsou splněny, ale základní struktura je natolik univerzální, aby nebyl problém chybějící funkce časem doplnit.

### 3 STAVBA PROHLÍŽEČE

Samotné jádro aplikace jsem sice nevyvinul, ale čas od času provádím nějaké úpravy a hlavně je nutné jej popsat pro pochopení základní struktury a funkce prohlížeče. Celá aplikace se skládá z mnoha částí, ale celá architektura je stavěna na základě dvou hlavních částí nebo lépe řečeno systémech. Jsou to komponentový systém a systém zasílání zpráv. Detailněji je architektura zachycena na následujícím obrázku.



**Obrázek 1:** Architektura multimediálního prohlížeče

Základem celé struktury je část Desktop, která zajišťuje komunikaci všech ostatních částí a spravuje a přiděluje systémové prostředky. Základem komponentového systému je část ElementList, která spravuje instance jednotlivých právě používaných komponent. Další neméně důležitou částí komponentového systému je část ComponentFactory, ta obsahuje prototypy všech možných komponent, tj. popis a reference na jejichž základě je možné vytvářet instance jednotlivých komponent. Poslední dvě významné části jsou Messenger a MessageProcessor, které představují systém zasílání zpráv. Část Messenger přijímá zprávy od jednotlivých komponent, filtruje je a rozesílá dalším komponentám. Část MessageProcessor je úzce spjata s předchozí částí a zajišťuje hlavně zpracování systémových zpráv, ty umožňují např. změnu konfigurace a vzhledu za běhu aplikace.

## 4 FUNKCE PROHLÍŽEČE

Jak je zřejmé z popisu architektury prohlížeče, základním konceptem aplikace jsou komponenty komunikující prostřednictvím zasílání zpráv. Rozesílané zprávy jsou samozřejmě ve formátu XML a jejich parametry si v podstatě určují komunikující komponenty. V některých případech jsme však komunikaci sjednotili specifikací tvaru některých typů zpráv, aby nedocházelo k nejednoznačnostem — např. synchronizační zprávy mají přesně danou strukturu aby bylo možné průběhy jednotlivých multimediálních toků časově sjednotit. Použití formátu XML i v případě reálné komunikace by se mohlo zdát neoptimální, ale v případě interakce s uživatelem je rychlost dostatečná a zpracovávání textové podoby zpráv nepůsobí žádné problémy.

Základní funkcí prohlížeče je umožnit uživateli prohlížet multimediální obsah, což mu umožňují jednotlivé komponenty. Vývoj komponent byla hlavní část mojí práce — příkladem takových komponent je komponenta HtmlBrowser pro zobrazování jednoduchého html obsahu nebo přímo z ní vycházející komponenta Transcript zobrazující textový přepis audio/video záznamu, dalšími komponentami jsou např. komponenta pro zobrazení prezentace v podobě sekvence obrázku (Cardslides) nebo zobrazení časové osy s odlišením jednotlivých mluvčích (TimeAxis) a několik dalších vizuálních i nevizuálních komponent. Jednotlivé komponenty samy zajišťují interakci s uživatelem a komunikaci s ostatními komponentami provádějí už prostřednictvím systému zasílání zpráv.

Navíc vzhled a funkce spuštěné aplikace není pevně stanoven, ale je právě definován v hlavní konfiguraci prohlížeče, která specifikuje které komponenty se z celé palety vyberou a rozmístí se na obrazovku, popř. do okna. Navíc komponenty je možné přidávat/odebírat za běhu. To z prohlížeče dělá opravdu multi-modální systém.

## 5 ZÁVĚR

Prohlížeč zatím není prakticky nasazen, jelikož je ve fázi intenzivního vývoje a ještě není schopen úplně uspokojit požadavky výzkumné skupiny. V budoucnu by však měl nahradit doposud používaný multimediální prohlížeč švýcarského výzkumného institutu IDIAP nazvaný JFerret, který je součástí jejich projektu MultiModal Media Fileserver. Multimediální prohlížeč P.R.A.S.E. se prohlížeči JFerret velice podobá s tím rozdílem, že JFerret je napsaný v programovacím jazyce java, což značně omezuje jeho výkon a navíc se jeho vývoj pozastavil a v některých situacích přestává postačovat.

Náš prohlížeč si však díky své konfigurovatelnosti a vysoké modalitě klade i jiný cíl a to pomoci studentům a dalším uživatelům. Studentům by měl konkrétně pomáhat při studiu — v současnosti se z téměř všech přednášek probíhajících na naší fakultě pořizuje záznam, ukládá na velkokapacitní servery a uložení a velká část z nich jich je přímo přístupných studentům přes webové rozhraní, popř. přes RTSP protokol jako VoD (Video on Demand). Vytvořit textové přepisy jednotlivých přednášek není problém a prohlížeč by pak představoval systém pro rychlý přístup k informacím ve formě video záznamu. Student by například při přípravě na zkoušku měl možnost rychlého vyhledání fráze nebo problému, kterému nerozumí a dostal by odkaz případně přímo úsek záznamu přednášky, která o dané problematice pojednává.

## REFERENCE

- [1] Kubalík, J.: Multimediální prohlížeč pro přednášky - Bakalářská práce, Brno, 2007