

Interaktivní portál pro efektivní multimediální exploraci

Vedoucí projektu

- RNDr. Jakub Lokoč, Ph.D.

Řešitelé

- Bc. Tomáš Grošup
- Bc. Přemysl Čech
- Bc. Lukáš Navrátil
- Bc. Jakub Kinšt
- Bc. Miroslav Macík

Motivace projektu

Cílem projektu je vytvořit interaktivní webový portál, který uživateli umožní exploraci multimediálních dat. V rámci SW projektu se předpokládá explorace pouze nad obrázky, systém by však měl být navrhován a implementován tak, aby bylo možné rozšíření pro jiný typ multimediálních dat s maximální znovupoužitelností společné logiky.

Multimediální explorace

Multimediální exploraci je chápáno plynulé a interaktivní grafické zobrazování pohledů (např. formou shluků podobných obrázků zobrazených ve 2D prostoru) na části multimediální databáze, které vznikají na základě interaktivního chování uživatele. K tomu je zapotřebí interaktivní grafické rozhraní, které uživateli umožní vykonávat explorační operace, např., přiblížení/oddálení pohledu na data (shluk jednoho typu obrázků nebo obecný náhled na data) nebo posunutí určitým směrem (více obrázků podobných těm vlevo). Posloupností takovýchto operací by se měl uživatel pohybovat (explorovat) multimediální kolekcí tak, aby směřoval k požadovanému výsledku.

Uživatelé projektu

Předpokládají se dvě cílové skupiny uživatelů:

- Běžní uživatelé - jejich typickým požadavkem je položení vyhledávacího dotazu a explorace v získaných kolekcích dat. Další možností je nahrání vlastní kolekce a provádět exploraci přímo nad těmito specifickými daty.
- Uživatelé provádějící vývoj podobnostních modelů - tito uživatelé budou systém využívat pro ladění existujících modelů (např. testování vlivu jednotlivých parametrů na výsledky vyhledávání) nebo pro testování nových komponent modelů.

Portál má být rovněž užitečný pro vznikající vědeckou spolupráci s univerzitou v Klagenfurtu, jejíž pracovníci pracují především na grafické stránce explorace, zatímco SIRET vyvíjí podobnostní modely, které potřebuje testovat a ladit.

Očekávaná funkcionálnita

Webový portál a explorace

Portál bude podporovat exploraci rozsáhlých statických kolekcí uložených na lokálním disku a menších dynamicky získaných kolekcí - opatřených např. z výsledků textového vyhledávání portálů třetích stran, jako je např. Flickr, na základě uživatelem specifikovaných klíčových slov. Multimediální data z portálů třetích stran budou systémem analyzována a budou vizualizována v naší explorační struktuře.

Webový portál bude podporovat zobrazování explorovaných dat pomocí fyzikálního modelu částic. Ten můžeme reprezentovat jako ohodnocený graf, kde vrcholy jsou množinou obrázků z datasetu vrácené na nějaký dotaz a hrany představují podobnosti mezi dvojicemi obrázků. Pro vykreslení grafu bude použit fyzikální model, kde vrcholy představují částice a ohodnocení hran představuje přitažlivou/odpudivou sílu mezi dvojicí vrcholů. Vrcholy exploračního grafu představující jednotlivé obrázky jsou pak umístěny na obrazovce na pozice částic, ve kterých se nacházejí po ustálení systému.

Podobnostní modely a deskriptory

Portál bude podporovat různé možnosti kombinace podobnostních modelů. Ty jsou založeny na deskriptorech dat, které představují předzpracované informace o datech, a podobnostních funkcích, které různými transformacemi a výpočty nad deskriptory vypočítávají podobnosti mezi prvky v kolekci. U běžných uživatelů se předpokládá využití přednastavených modelů, které by měly fungovat univerzálně, ovšem rozhraní systému umožní modely i modifikovat (nastavovat parametry, zaměňovat dílčí části modelů) a využít tak systém k vývoji a testování modelů.

Jak podobnostních funkcí, tak deskriptorů obrázků existuje velké množství. V rámci projektu bude do systému implementován podobnostní model založený na technice Bag of Features, přičemž se předpokládá navíc implementace alespoň dvou "state of the art" rozšíření tohoto modelu (některé z množiny Hamming Embedding, VLAD, Fisher Kernel nebo Asymmetric Hamming Embedding případně nově se objevivších). Předpokládá se vlastní implementace modelů, nelze však předem vyloučit částečné nebo úplné použití knihoven třetích stran.

Předpokládá se implementace alespoň dvou různých deskriptorů a to nízkodimenzionálních deskriptorů (např. PCT) a některých vysokodimenzionálních deskriptorů z množiny SIFT, SURF nebo cSIFT. Zvláště pro extrakci deskriptorů existuje velké množství knihoven, pravděpodobně se jejich extrakce nebude implementovat znovu a použijí se externí knihovny.

Explorační a indexační struktury

Dle typu dat bude portál využívat paměťové (spočítané ihned) i perzistentní (předpočítané) explorační datové struktury. Ty se mohou pro menší kolekce budovat např. z matice vzdáleností, pro rozsáhlé kolekce je potřeba je konstruovat levnějšími způsoby.

Pro potřeby rychlé a efektivní multimediální explorace je nutné použít metrické indexační struktury. Bude využita existující implementace některých datových struktur (např. PM-Tree) z existující knihovny SOL napsané v C++ pro vědecké potřeby naší katedry. Bude implementován wrapper této knihovny do C# a dále tato knihovna bude upravována pro explorační potřeby projektu. Indexy se budou používat pro typické operace jako jsou dotazy na nalezení k nejbližších sousedů nebo vzdálenostní dotazy, avšak dále struktury budou upraveny pro potřeby hierarchického procházení a explorace za pomoci fyzikálního modelu částic. Budou potřeba nové dotazy typu zanoření do struktury, vynoření ven, posun pohledu do všech směrů a další. Tyto dotazy je třeba zanalyzovat, dodefinovat a doimplementovat. Dále budou implementovány specifické indexační struktury (např. M-Index) pro tyto potřeby přímo v prostředí .NET. Na indexech se také budou simulovat různé podobnostní modely založené na odlišných deskriptorech, metrických funkcích či klíčových vlastnostech zkoumaných objektů, kdy jednotlivé indexy se mohou chovat rozdílně v závislosti na různém nastavení parametrů.

Anotace dat a testování modelů

Pro potřeby testování (případně trénování) podobnostních modelů bude vytvořena hra pro mobilní platformu Android (optimalizována především pro tablety), díky které bude možné pomocí hry pro jednoho i více hráčů anotovat obrázky ve statických kolekcích. Pro potřeby automatizovaného měření efektivity vyhledávání bude připravena a otestována i sada benchmark testů (rychlost nalezení obrázku, kvalita klastrování danou podobnostní strukturou).

Platforma

Systém poběží jako webová aplikace, front-end bude napsán v HTML5, Javascriptu a TypeScriptu, serverová část bude fungovat v ASP.NET MVC při použití jazyka C# . V případě potřeby mohou být některé knihovny vyžadující vysoký výkon napsané v C++. Android aplikace bude implementována pomocí standardních knihoven Android SDK v jazyce Java.

Rozdělení činností

Tomáš Grošup – analýza a návrh podobnostních struktur pro statické i dynamické datové kolekce, návrh celkové architektury portálu.

Přemysl Čech – analýza a implementace perzistentních podobnostních struktur, správa indexačních datových struktur, signaturové podobnostní modely.

Lukáš Navrátil – implementace, indexování a trénování pokročilejších metrických a nemetrických podobnostních modelů. Návrh obecných rozhraní pro komunikaci mezi komponentami portálu.

Jakub Kinšt – hra pro tablety (i telefony) se systémem Android určená pro „crowdsourcové“ vytváření anotovaných obrázkových kolekcí a následné zpracování anotovaných obrázků.

Miroslav Macík – implementace konfigurovatelné klientské části pro standardní webové prohlížeče, benchmark scénáře pro měření efektivity portálu, sběr, analýza a vyhodnocení naměřených dat.