

Oponentský posudek na disertační práci

Autorka práce: Petra Budíková

Název práce: Towards Large-Scale Multi-Modal Image Search

Téma práce – multimodální podobnostní vyhledávání – je aktuální a velmi komplexní. Většina prací v oblasti Multimedia Retrieval řeší nějaký specifický podproblém, ať už je to „databázový úkol“ (typicky nějaká indexační technika urychlující vyhledávání), anebo úkol „kvalitativního modelování v doméně“ (konkrétní model extrakce vlastností a podobnostní míry vedoucí ke zlepšení přesnosti vyhledávání). Málo prací se však zabývá nejdůležitější otázkou, a to použitelností podobnostního vyhledávání v konkurenci tradičních dotazovacích paradigmat (full-textové vyhledávání, strukturované vyhledávání). Oblast content-based multimedia retrieval, ačkoliv je na scéně již dvě dekády, se totiž dosud neuplatnila ve větším rozsahu, než jako akademické implementace, či jako pokusné projekty velkých vyhledávačů (např. Google Images). Ve velké míře je to kvůli zjednodušujícím předpokladům, které konečnému uživateli (resp. databázovému programátorovi) nabízejí vesměs jednoduché dotazy příkladem, navíc v rámci jednoho podobnostního modelu (v práci zvané jako single-modality přístupy). V této práci se autorka zabývá komplexním pojetím podobnostního vyhledávání, a to více-modálním vyhledáváním agregujícím více dílčích podobnostních/rankovacích modelů v rámci dotazu, pro které navíc navrhuje podporu v podobě dotazovacího jazyka vycházejícího z SQL. Kromě popisu více-modálního přístupu pak autorka prezentuje experimentální studie (webové prostředí a automatická anotace), ve které jsou zmíněné přístupy použity.

Výsledky práce a hodnocení:

Autorka se podílela na návrhu nových přístupů rychlého více-modálního podobnostního vyhledávání ve velkých obrazových kolekcích. Jedním z nich je aproximovaný distribuovaný Threshold algoritmus, dalším potom jiná aproximativní technika využívající postprocessing (ranking). Samostatným výsledkem je potom návrh nového dotazovacího jazyka rozšiřujícího SQL o konstrukce vhodné pro podobnostní dotazování.

Kromě zmíněných původních přínosů je dalším výsledkem práce velmi kvalitní rešerše mapující toto odvětví výzkumu podobnostního vyhledávání. V rámci rešerše autorka přehledně a logicky zmapovala/kategorizovala velmi různorodou rodinu přístupů k podobnostnímu vyhledávání. Kategorizace by mohla usnadnit orientaci vědcům pracujícím v této oblasti, jelikož se zde taxonomie díky bouřlivému rozvoji dosud neustálila.

Výsledky představené v této práci byly publikovány ve sbornících několika mezinárodních konferencí, významným výsledkem je pak článek přijatý do časopisu IJMDM. Toto penzum publikací je dostatečné a reprezentativní v rámci výsledků představených v disertační práci.

Struktura práce:

Po úvodu do problematiky vyhledávání v rozsáhlých obrazových databázích autorka formuluje cíle práce. Ve třetí kapitole se zabývá jedno-modálními přístupy a ve čtvrté potom obsáhle popisuje více-modální přístupy k vyhledávání v obrazových databázích. Pátá kapitola

rozpracovává více-modální paradigma v kontextu metrického modelu indexování, mimo jiné autorka popisuje distribuované techniky zpracování více-modálních dotazů, postprocessing a také dotazovací jazyk SimSeQL. V šesté kapitole autorka prezentuje výsledky komplexního benchmarku na webových obrazových datech a v sedmé kapitole pak aplikace v automatické anotaci a klasifikaci obrázků. Osmá kapitola práci shrnuje a vytyčuje směry budoucího výzkumu.

Připomínky k textu:

- Na str. 15 se neformálně definuje termín „modality“. Není zde uvedena citace, odkud autorka čerpá. Dosud jsem měl tento termín spojen s trošku jiným konceptem, a to jako „způsob“, resp. paradigma vyhledávání. Tedy ne jako datový typ s podobnostním modelem, vše pracující pod paradigmatickým query-by-example (jak se zde používá), ale spíše celé paradigma získávání informace. Například dotazování je jedna modalita (způsob získávání informace), browsing další, filtrace/doporučení ještě jiná, apod. Zde užívaná definice není chybná, jen by mně zajímal původ.
- Na str. 30 bych ocenil citace k early/late fusion.
- Zaměňování termínů feature/deskriptor lehce mate, obecně se feature chápe jako dílčí vlastnost (kterých může mít objekt více) a deskriptor jako jeden složený popisovač objektu.

Jazyková a grafická úroveň:

Práce je na nadprůměrné jazykové úrovni. Sem tam jsem našel překlep neodhalitelný spellcheckerem (např. záměna are – as), ovšem v míře zcela minimální.

Otázka k obhajobě:

Bylo nutné zavádět zcela nový dotazovací jazyk, když autoři [13] již podobné rozšíření SQL zavedli? V [13] je navíc užito pouze „legacy“ rozšíření jazyka SQL, tj. v objektově-relační databázi lze podobnostní konstrukce použít přímo ve standardním příkazu SELECT, nemusí být tedy implementován zcela nový databázový stroj. Bylo motivací k zavedení nového jazyka jeho specifický překlad do „strojového kódu“ platformy MUFIN?

Závěry:

Autorka jednoznačně prokázala schopnosti potřebné k samostatné tvůrčí vědecké práci, přičemž práce samotná splňuje standardní požadavky kladené na disertační práce v oboru informatika. Doporučuji práci k obhajobě před odbornou komisí a navrhuji hodnocení A - výborně. Vzhledem k rozsahu a ucelenosti doporučuji práci k ocenění děkanem FI MU jako vynikající.

V Praze dne 4.3. 2013


Doc. RNDr. Tomáš Skopal, Ph.D.
MFF UK