

Appendix A Resumé in Czech language

Tato práce se zabývá vyhledáváním obrázků a specifických objektů v obrazových databázích. Na vstupu uživatel zadá obrázek objektu resp. scény a vyhledávací stroj vrátí obrázky stejného objektu resp. scény z databáze. Teze se zaměřuje na *bag-of-words* přístup, který je jedním z nejefektivnějších pro tento typ úlohy. Specifický objekt může pokrývat pouze část obrázku nebo může být z části překrytý jiným objektem. Práce vylepšuje více částí standardních bag-of-words postupů.

Nová similaritní funkce je definovaná pro bag-of-words vyhledávání obrázků. Tato funkce je naučená bez učitele, oproti standardní metodě nevyžaduje extra paměťový prostor a je více diskriminabilní než eukleidovský L2 *soft assignment* nebo *Hamming embedding*. Navrhovaná similaritní funkce dosahuje na standardních databázích vyšší *mean average precision* než všechny dosud publikované výsledky v literatuře.

Jsou studovány účinky velmi jemné kvantizace u velkých vizuálních slovníků (až 64 milionů slov) a ukazuje se, že výsledky vyhledávače specifických objektů se zlepšují se zvyšujícím se množstvím slov. Toto pozorování je v rozporu s předešlými publikovanými výsledky. Dále ukazujeme, že s velikostí slovníku se zvyšuje rychlost *tf-idf* skórování.

Všechny *state-of-the-art* výsledky vyhledávačů publikované v literatuře byly dosaženy s použitím *query expansion* kroku, který zásadně vylepšuje kvalitu vyhledávání. Přinášíme tři modifikace automatické query expansion: (i) metodu předcházející selhání query expansion kroku vzniklou přítomností tzv. *confusers*, (ii) vylepšenou geometrickou verifikaci, která inkrementálně vytváří statistický model objektu s přibývajícimi verifikovanými obrázky a (iii) učení relevantního geometrického kontextu, který zásadně zlepší výsledky, pokud je využitý v query expansion.

Všechny tři vylepšení query expansion kroku byly testované na standardních databázích Paříž a Oxford, kde dosáhly *state-of-the-art* výsledky.

Nakonec byly formulovány nové úlohy vyhledávání. Ukázali jsme, že klasické uspořádání výsledků založené na podobnosti obrázků odpoví pouze jeden z možných dotazů uživatele. Místo vyhledávání nejvíce se podobajících obrázků navrhuje metody přibližování a oddalování, které zodpoví otázky “*Co je to?*” a “*Kde je to?*”.

Formulujeme další dvě úlohy: (i) pro každý pixel zadaného obrázku nalezni jeho maximální rozlišení v obrázcích dané databáze a (ii) pro každý pixel zadaného obrázku vrať četnost jeho výskytu v databázi. Tyto úlohy postavené na zoom-in a zoom-out metodách vyžadují dvě nové techniky: hierarchickou “query expansion” a verifikaci geometrické konzistence na nalezených obrázcích, která je dostatečně robustní, aby předešla odklonění se od původního objektu v průběhu vyhledávání. Experimenty ukazují, že navrhované metody dokáží najít překvapivě drobné detaily na testovaných obrázcích, a to i detaily jenom těžko vyditelné pouhým okem.