

Image retrieval - Podobnost obrázků (SIFT)

Jaroslav Hrách a Petr Pejša

Popis projektu

Cílem projektu je vytvoření aplikace, která měří podobnost vstupního obrázku s databází obrázků na základě analýzy extrahovaných SIFT deskriptorů.

SIFT (Scale-Invariant Feature Transform) je technika extrahování lokálních vlastností z obrázku. Výstupem aplikace metody SIFT na obrázek je množina deskriptorů popisující důležité body na obrázku.

Způsob řešení

Aplikace se skládá ze dvou částí – z grafického uživatelského rozhraní (GUI) a z výpočetní části.

Jako GUI je zde webová stránka vytvořena za pomoci frontendového frameworku **Bootstrap**, který umožňuje snadnou manipulaci s **HTML** a **CSS**. Funkčnost a zpracování jednotlivých modulů aplikace je zajištěno přes **PHP**, ve kterém se moduly používají jednoduše přes shell. Dále je přes PHP měřen čas běhu aplikace a všech jejích modulů. V neposlední řadě byl při návrhu kladen důraz na dobré UX – aplikace je jednoduchá a snadno použitelná.

Výpočetní část je rozdělena na tři moduly: ASIFT, Cluster a SQFD.

Modul **ASIFT** slouží k extrahování deskriptorů za pomoci ASIFT knihovny, která byla upravena pro naše potřeby. Původní knihovna zpracovává dva obrázky a hledá mezi nimi shodné body a ještě k nim vytvoří obrázek s grafickým znázorněním. V naší aplikaci jsme si vystačili pouze s extrahováním deskriptorů z jednoho obrázku a následném uložení do textového souboru.

Modul **Cluster** z extrahovaných deskriptorů uložených v textovém souboru vytvoří clustery na základě euklidovské vzdálenosti bodů. Body jsou ve stejném clusteru, jestliže je jejich vzájemná vzdálenost menší než ϵ (po experimentálním testování je jako nejlepší varianta považováno $\epsilon=3$). Následně pro každý cluster nalezneme centrální bod, který bude cluster reprezentovat, zároveň zachováme data jako je počet bodů v clusteru, zahodíme deskriptory ostatních bodů.

Clusterování je nedílnou součástí naší aplikace, poněvadž knihovna ASIFT generuje příliš mnoho bodů. Zpracování takto ohromného počtu bodů je časově náročné, proto je vhodnější snížit jejich počet. Díky tomu se urychlí doba výpočtu pro následující modul SQFD.

Asymptotická složitost clusterování je v nejhorším případě $O(N^2)$, kde N je počet bodů vygenerovaných knihovnou ASIFT.

Po clusterování přichází na řadu algoritmus **SQFD (Signature Quadratic Form Distance)**.

Algoritmus SQFD se provede pro každý obrázek v databázi – jeho úkolem je porovnat dva obrázky pomocí clusterů uložených v textových souborech a vypočítat tak míru jejich podobnosti. Ve zkratce funguje SQFD následovně: vytvoří vektor $WqWo$, který je složen z vektoru Wq a $-Wo$ zřetězením. Vektor Wq resp. Wo je N -složkový pro obrázek s N clusterů, každá složka vektoru je charakterizována jako poměr počtu bodů v clusteru oproti celkovému počtu bodů. Dále vytvořím matici A_{fx} , která je vytvořena následujícím vzorcem $A_{i,j} = \frac{1}{1+L2(ci,cj)}$ kde ci představuje i -tou pozici vektorem $WqWo$ a $L2$ je funkce na výpočet euklidovské vzdálenosti.

Výsledná míra podobnosti je vypočtena $SQFD = \sqrt{WqWo * A_{fx} * WqWo^T}$.

Asymptotická složitost SQFD je v nejhorším případě $O(N*M)$, kde N je počet clusterů jednoho porovnávaného obrázku a M je počet clusterů druhého obrázku, zároveň paměťová složitost je $O(N*M)$.

Příklad výstupu


Na níže přiloženém obrázku je ukázka funkce programu. V horní části stránky je zobrazen vstupní obrázek, ke kterému hledáme obrázky podobné. Pod ním jsou obrázky z databáze seřazené dle skóre podobnosti vypočteným algoritmem SQFD. Pod všemi obrázky jsou data o trvání výpočtu.

Z příkladu výstupu je jasně vidět, že obrázky Pražského hradu jsou na prvních pozicích se skóre 0,51 – 0,55. Další obrázky, které se liší, mají skóre od 0,59 do 1,65.

[SIFT](#) [Hlavní stránka](#) [Obrázky obrázků](#)


Image retrieval - podobnost obrázků (SIFT)

Vstupní obrázek:




Vybrané episkop: 3


Podobné obrázky




Skóre: 0.0511088831




Skóre: 0.055384169




Skóre: 0.0950797565




Skóre: 0.0501571556




Skóre: 0.0526221827




Skóre: 0.063332274



Skóre: 0.0729275509



Skóre: 0.1111732844



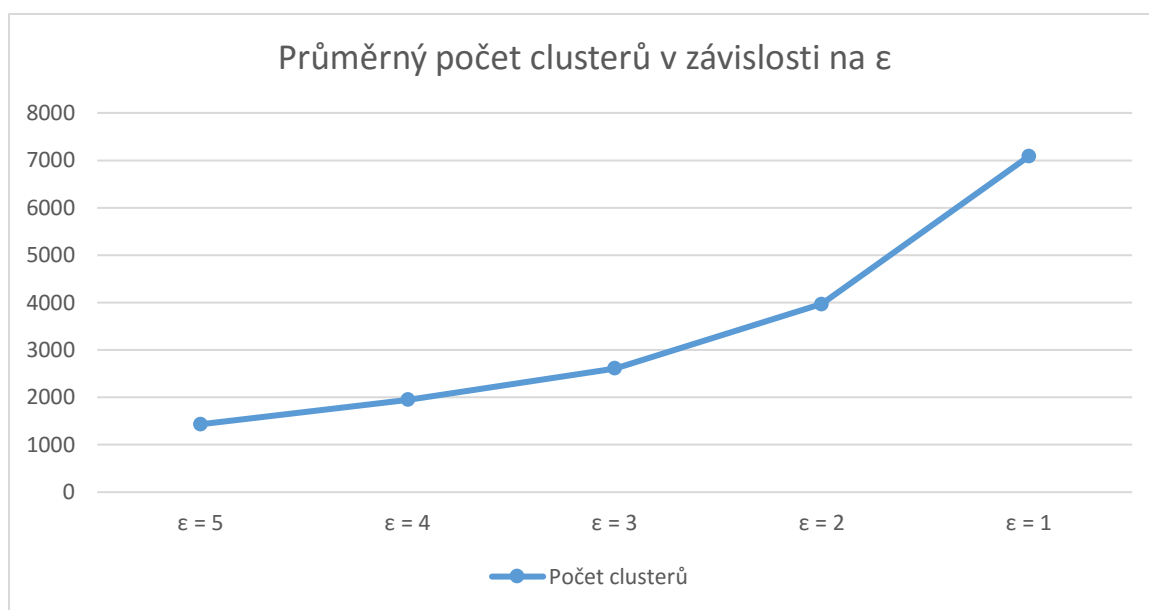
Skóre: 0.1056445322

Doba výpočtu ASIFT: 25.90s
Doba výpočtu cluster: 79.13s
Doba výpočtu SQFD: 63.78s
Celková doba výpočtu: 164.81s

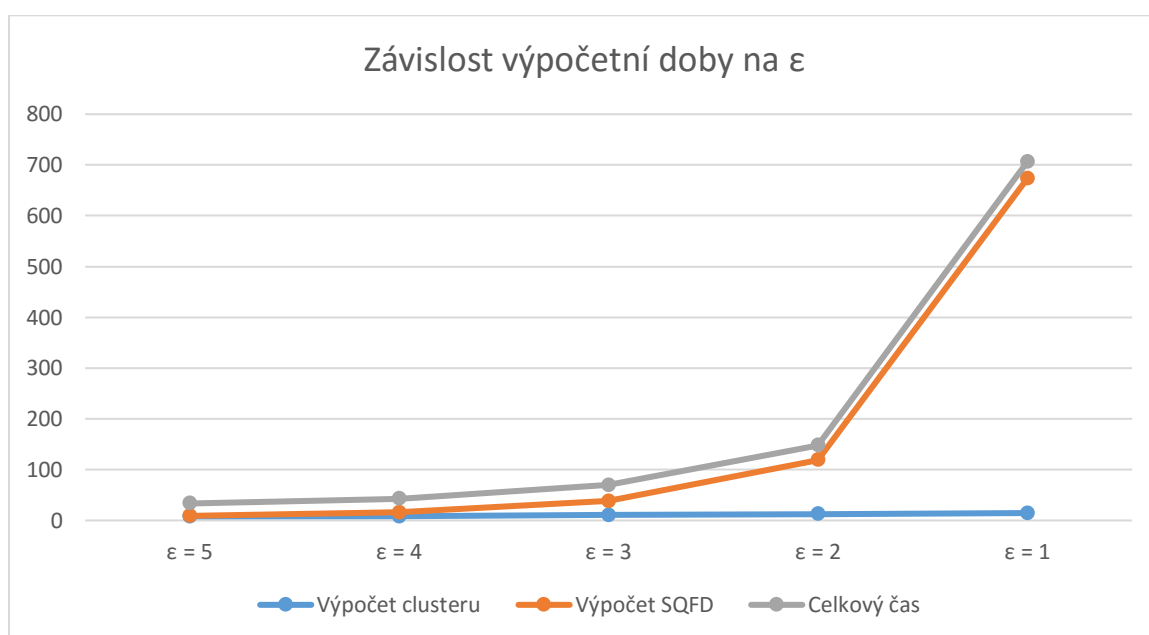
© Jaroslav Hrdlička, Petr Pojta

Experimentální sekce

Výpočetní doba hledání podobného obrázku závisí na zvoleném ϵ , které určuje maximální vzdálenost mezi body, které mají ještě patřit do stejného clusteru. Na následujícím grafu je zachycen průměrný počet vytvořených clusterů v závislosti na zvoleném parametru ϵ .



Pro malé ϵ se vytvoří velké množství clusterů, které obsahují méně bodů. Následný výpočet podobnosti algoritmem SQFD je pomalý díky asymptotické složitosti $O(N \cdot M)$, kde N je počet clusterů jednoho porovnávaného obrázku a M je počet clusterů druhého obrázku. Z pohledu výpočetní doby je ideální mít menší počet clusterů, tak aby clusterly obsahovaly více bodů.



Pro větší ϵ se body budou více shlukovat k sobě a vytvoří se mohutnější clustery. Těchto clusterů bude menší počet a budou obsahovat více bodů. Pokud se ϵ zvolí až příliš velké je možné že se clustery vytvoří až moc obsáhlé, jejich případné porovnávání nebude odpovídající.

Je tedy vhodné najít ϵ takové, aby výpočetní doba byla co nejkratší a zároveň aby byly výsledky relevantní. Nejprve tedy zopakujme co ϵ vlastně znamená. Hodnota parametru ϵ je maximální vzdálenost mezi body pro spojení do clusteru. Pokud jsou tedy body blíže než zvolené ϵ budou spolu v jednom clusteru. Z experimentálního pozorování jsme zjistili, že pokud je clusterů méně než 500, výsledky hledání nejsou příliš dobré. Naopak pokud je clusterů příliš mnoho např. více než 10 000 je výpočetní doba neúnosná, zároveň paměťová složitost je $O(N \cdot M)$.

V naší aplikaci jsme se tedy rozhodli omezit výběr ϵ na hodnoty 1, 2, 3, 4 a 5, u kterých jsme dosahovali dobrých výsledků jak z pohledu výpočetní doby, tak z kvality výsledků.

Diskuse

Nejlepších výsledků jsme v rámci našeho testování dosahovali při porovnávání obrázku oproti výřezu z tohoto obrázku. Skóre takového obrázku je řádově lepší než u ostatních obrázků z databáze.

Zároveň u velice rozmanitých obrázků dochází k vygenerování obrovského množství klíčových bodů knihovnou ASIFT. Při následném vytvoření clusterů a porovnání se lehce zamění za jiný podobně rozmanitý obrázek. Metoda určování podobnosti SQFD je totiž založená pouze na poloze clusteru a počtu bodů uvnitř.

Porovnávání pouze podobných obrázků je většinou dobré a podobné obrázky se řadí na první místa. Někdy se na první místa dostanou obrázky jiné.

Závěr

Cílem projektu bylo vytvoření aplikace, která měří podobnost vstupního obrázku s databází obrázků na základě analýzy extrahovaných SIFT deskriptorů.

Z extrahovaných deskriptorů jsme vytvořili clustery pomocí euklidovské vzdálenosti mezi nimi. V těchto clusterech jsme našli centrální prvek, který reprezentuje cluster.

Porovnání obrázků bylo řešeno algoritmem SQFD, který porovná clustery dvou obrázků a vypočte podobnostní míru.

Podle podobnostní míry vstupního obrázku oproti obrázkům z databáze seřadíme obrázky v databázi tak, aby nejpodobnější obrázky byli první.

Výsledky, které dostáváme, jsou dobré díky omezení výběru hodnoty ϵ pouze na (1, 2, 3, 4, 5). Jiné hodnoty ϵ buď snižují kvalitu výsledku, nebo výrazně prodlužují dobu běhu.