

Západočeská univerzita v Plzni
Fakulta aplikovaných věd
Katedra informatiky a výpočetní techniky

Bakalářská práce

Automatická anotace obrázků

Místo této strany bude
zadání práce.

Prohlášení

Prohlašuji, že jsem bakalářskou práci vypracovala samostatně a výhradně s použitím citovaných pramenů.

V Plzni dne 22. ledna 2017

Kateřina Kratochvílová

Abstract

The text of the abstract (in English). It contains the English translation of the thesis title and a short description of the thesis.

Abstrakt

V této bakalářské práci se zabýváme automatickou anotací obrázků..

Obsah

1	Úvod	6
2	JEC	7
2.1	Příznaky	7
2.2	Vzdálenosti	7
2.3	Přenesení klíčových slov	7
2.4	Vyhodnocení	8
3	Testovací databáze	9
4	Návrh systému	10
5	Použité programové prostředky / Požadavky na software	11
5.1	OpenCV	11
6	Výsledek	12

1 Úvod

V dnešní době, kdy je svět přesycen obrázky v digitální podobě, není vůbec snadné nalézt obrázek zobrazující požadovaný obsah. Naneštěstí počítače nedokáží vnímat obraz jako lidé, vnímají totiž obrazy jako sérii binárních informací. Přitom počítače a jejich práce s obrazy by se dala využít v mnoha oborech jako je lékařství nebo doprava. Na základě toho vyplývá na povrch problém jak spravovat digitální obrázky a efektivně mezi nimi vyhledávat. Prostřednictvím klíčových slov přiřazených k obrázkům se dá problém vyhledávání zjednodušit. Přiřazení klíčových slov probíhá pomocí procesu automatické anotace obrázků, kdy za pomoci trénovací množiny, ze které program natrénujeme, je k obrázku přiřazen jeden nebo více slov které vyjadřují jeho obsah. Automatická anotace obrázku je tedy proces, ve kterém jsou k obrázku automaticky přiřazena metadata, která obsahují klíčová slova (například les, strom).

Problematika porovnávání obrázků, po které následuje anotace, se dá rozložit na dva menší problémy. První je získání informace z trénovací množiny a druhý je jak rozhodovat jestli jsou si obrázky skutečně podobné a nakolik.

Cílem práce je navrhnout a implementovat software umožňující automatickou anotaci obrázků. Popsat konkrétní metodu.

2 JEC

Existuje skupina základních metod pro obrázkovou anotaci, která je postavena na hypotéze, že na základě podobnosti vzhledu obrázku jsou podílově přiřazena klíčová slova. K obrázkové anotaci se přistupuje jako k procesu přenášení klíčových slov od nejbližších sousedů. Struktura sousedů je konstruovaná použitím jednoduchých low-level obrázkových příznaků.

2.1 Příznaky

Barva a textura jsou považovány za dva nejdůležitější nízko-úrovňové příznaky pro obrázkovou reprezentaci. Nejběžnější barevné deskriptory jsou základem hrubého histogram, který je často využíván s obrázkovým srovnáním a indexovým schématem, primárně z důvodu jejich efektivnosti a snadného výpočtu. Obrázková textura je běžně zachycena s Wavelet vlastností. V části Gabor a Haar wavelets bylo prokázáno, že je velmi efektivní při vytváření rozptýlených diskriminačních obrázkových rysů. Omezení vlivu a sklon k individuálním funkcím, a maximalizování množství obsažených informací vybereme pro práci pár jednoduchých a snadno vypočítatelných funkcí.

2.2 Vzdálenosti

asda

2.3 Přenesení klíčových slov

Pro přenesení klíčových slov používáme metodu, kdy přeneseme n klíčových slov k dotazovanému obrázku \tilde{I} od K nejbližších sousedů v trénovací sadě. Mějme $I_i, i = 1, \dots, K$, tyto K nejbližší sousedy seřadíme podle vzrůstající vzdálenosti (tzn. že I_1 je nejvíce podobný obrázek). Počet klíčových slov k danému I_i je označen jako $|I_i|$. Dále jsou popsány jednotlivé kroky algoritmu na přenesení klíčových slov.

1. Seřadíme klíčová slova z I_1 podle jejich frekvence v trénovacích datech.
2. Z $|I_1|$ klíčových slov z I_1 přeneseme n nejvýše umístěná klíčová slova do dotazovaného \tilde{I} . Když $|I_1| < n$ pokračujte na krok 3.

3. Seřad klíčová slova sousedů od I_2 do I_K podle dvou faktorů

- (a) Co výskyt v trénovacích datech s klíčovými slovy přenesených v kroku 2 a
- (b) místní frekvence (tj. jak často se vyskytují jako klíčová slova u obrázků I_2 až I_K). Vyber nejvyšší rankink $n - |I_1|$ klíčových slov převedených do \tilde{I} .

Tento algoritmus pro přenos klíčových slov je poněkud odlišný od algoritmů které se běžně používají. Jeden z běžně užívaných funguje na principu, že klíčová slova jsou vybrána od všech sousedů (se všemi sousedy je zacházeno stejně bez ohledu na to jak jsou danému obrázku podobní), jiný užívaný algoritmus k sousedům přistupuje váženě (každý soused má jinou váhu) a to na základě jejich vzdálenosti od testovaného obrázku. Při testování se ovšem ukázalo, že tyto přímé přístupy přináší horší výsledky v porovnání s námi použitým dvoufaktorovým algoritmem pro přenos klíčových slov.

V souhrnu námi použitá metoda je složenina ze svou složeniny obrázkové vzdálenosti měřítka (JEC nebo Lasso) pro nejbližší ranking, kombinuje se s výše popsáním algoritmem na přenášení klíčových slov.

2.4 Vyhodnocení

3 Testovací databáze

Pro natrénování a následné testování byly data z databází ČTK,ESP a IA-PRC.

4 Návrh systému

5 Použité programové prostředky / Požadavky na software

Program byl navržen na operační systému Linux. Jako programovací jazyk byl zvolen Python a to z důvodu jeho jednoduchého použití, což je na prototyp, jako je tento velice výhodné na časovou náročnost. Program využívá knihovnu OpenCV 3.1.

5.1 OpenCV

OpenCV (Open source computer vision) je knihovna vydávána pod licencí BSD a je volně k dispozici jak pro akademické účely, tak pro komerční použití. Je vhodná pro použití v C++, C, Python a Javě. Podporuje operační systémy Windows, Linux, Mac OS, iOS a Android.

Knihovna byla navržena pro výpočetní efektivitu v oblasti počítačového vidění a zpracování obrazu se zaměřením na zpracování obrazu v reálném čase. Z důvodu optimalizace byla napsána v C/C++.

Knihovnu OpenCV je možné stáhnout na adrese: <http://opencv.org/>

6 Výsledek