

Ořezávání účtenky ve fotografii

Specifikace ročníkového projektu
25.2.2019 Jiří Mayer

Cíl projektu

Vytvořit program, který dostane fotografii účtenky a vrátí vrcholy čtyřúhelníku, který co nejlépe ohraničuje účtenku. Oříznutí podle čtyřúhelníku by mělo účtenku připravit na OCR.

Vlastnosti vstupu

Fotografie přímo z chytrého telefonu. Fotografie je ostrá a má dostatečné rozlišení. Je vyfocená na výšku. Před zpracováním se její šířka normalizuje na nějakou danou velikost.

Očekávaná přesnost

Program by měl detekovat fotografie:

- s jednotnou texturou na pozadí (dřevo, kov, plast, látka)
- pokud jsou na fotografii kromě účtenky jiné předměty nebo okraj stolu
- pokud účtenku někdo přidržuje prsty nebo je zatížená nějakým předmětem
- osvětlení by nemělo hrát roli (den, noc, svítlna na mobilu)
- část účtenky je mimo okraj fotografie (malá část, hlavně nahoře nebo dole)

Situace, které nejspíš nebudou dobře detekovány:

- pozadí má podobnou texturu, jako účtenka (noviny, jiné účtenky, bílá plocha)
- účtenka je moc zmačkaná a tak má špatný tvar nebo texturu
- přes účtenku leží předměty tak, že její tvar je moc zvláštní nebo dokonce nesouvislý

Pokud účtenka není nalezena, program informuje uživatele.

Použité technologie

Program bude realizovaný jako konzolová aplikace v Pythonu s použitím knihoven OpenCV a Numpy.

Představa implementace

Začnu podobným postupem, jaký je uvedený v článku Donoser, Bischof, Wilsche https://www.researchgate.net/publication/224057460_Color_blob_segmentation_by_MSER_analysis a zkusím místo explicitního vybírání počáteční oblasti poskytnout nějakou sadu dobrých distribucí pixelů. Nebo můžu kombinovat obojí. Použiju MSER na detekci blobu účtenky, případně jejího konvexního obalu.

Pro detekci čtyřúhelníku z blobu nechci používat bounding rectangle, doufám že najdu lepší algoritmus, který se bude snažit trefit okraje účtenky a ne minimalizovat obsah oblasti.

Jelikož mám problém v dost úzké doméně, tak si můžu dovolit předpokládat víc věcí (a tak například určit některé parametry na základě dat).

Předpokládám, že konkrétní implementace se ještě v průběhu změní nebo upraví, protože teď nedokážu říct, zda bude zvolený postup splňovat všechny požadavky.

Rozšíření

Pokud by program nebyl dostatečně rozsáhlý, můžu přidat buď další krok v OCR pipeline a sice binarizaci fotografie, nebo můžu implementovat jiný detektor a výsledek nějak průměrovat z obou (více) detektorů.