

Projektbeszámoló

Képfeldolgozás haladóknak MSc gyakorlat

Ábrahám Gábor Kiss Marcell Sajtos Gyula

2015. október 1.

1. Feladatspecifikáció

A projekt témájának megnevezése „Logófelismerés”, azonban ez meglehetősen nagy teret ad az értelmezésnek. Ezért fontosnak tartottuk, hogy pontosan definiáljuk a megoldandó problémát.

A feladatot úgy értelmeztük, hogy adott bemeneti képen a megvalósítandó módszer felismeri az azon látható termék feltételezett márkáját annak logója alapján, és kimenetként megadja annak nevét. Amennyiben több lehetséges termék is található a felvételen, akkor azok közül egyet azonosít. Hogy tovább finomítsuk a specifikációt és egyben csökkentsük az erőforrásigényeket, a tervezett alkalmazás csak alkoholos italok, sörök címkéjén szereplő logó felismerésére készítjük fel. Még vita folyik arról, hogy a módszer része legyen-e felhasználói beavatkozás a logó helyének kijelölésében, így segítve az alkalmazást.

2. Fejlesztési folyamat

2.1. Környezet

Az alkalmazható módszerek teszteléséhez és prototipizáláshoz a Matlab csomagot választottuk masszív képfeldolgozási eszköztára miatt. A megbeszélések során felvetődött az OpenCV keretrendszer használata. A csapat közös döntése értelmében ennek alkalmazását csak egy jól teljesítő prototípus elkészítése utánra halasztottuk.

A fejlesztési folyamat során előálló dokumentumok és forráskódok verziókezelésére a webes GitHub szolgáltatást jelöltük ki. A kezdeti kódrészletek

és dokumentáció (beleértve ezen fájlt is) migrálása a GitHubra folyamatban van.

2.2. Módszer

Úgy látjuk, hogy a probléma során felmerülő példányok nagyon változatosak, még hozzá több szempontból is. Egyrészt a bemeneti képen lévő potenciális logó többféle zajjal is terhelt, ez jön a kép készítésének megvilágítási problémáiból, de geometria torzulással is számolunk kell, hiszen a legtöbb alkoholos ital hengeralakú csomagolásban árusított. Másrészt maguk a detektálandó logók sem írhatóak le egyszerűbb matematikai modellel, sajátos grafikai struktúrájuk van.

Ezen tényezőknek nagy hatása van a tervezett módszer komponenseire. Szükségszerűen osztályoznunk kell a bemeneti képet egy tanuló ágens segítségével, továbbá ehhez szükség van annak valamilyen numerikus leírásához. Kezdeti módszernek SVM-et használunk a képfeldolgozás területén bizonyított teljesítménye miatt (ez természetesen változhat). Jelenleg a SIFT [1] és SURF [2] feature leírók használatával kapcsolatos tesztelések folynak, erről pár minta a 3. fejezetben látható. Ezen módszerek illenek a probléma természetéhez, mivel skála- és forgásinvariánsak, sőt bizonyos mértékig ellenállóak az affin transzformációkkal, a különböző nézőpontokból származó torzulásokkal szemben is.

Az aktuális problémát a többcsatornás képek jelentik – jelen pillanatban az $RGB \rightarrow HSV$ konverzió utáni csatornánkénti feature kinyeréssel próbálkozunk.

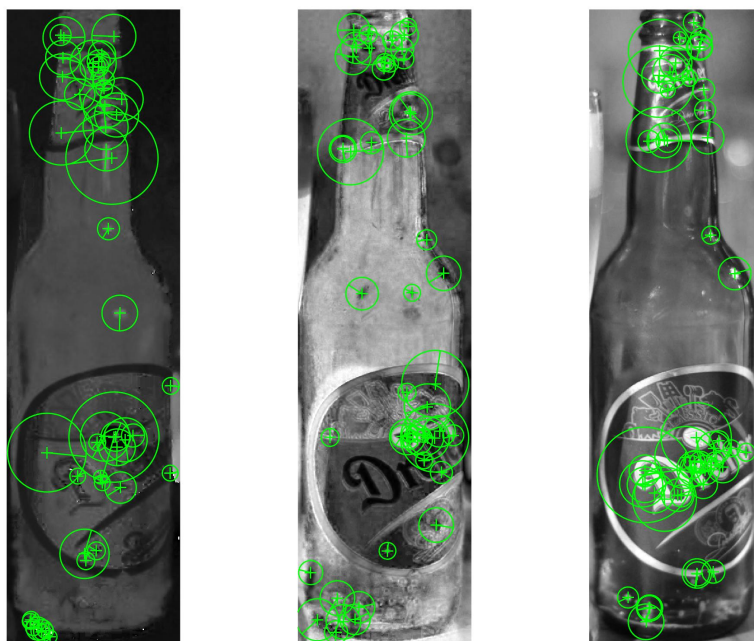
3. Részeredmények

Az alábbi képeken látható a kísérletezés során kapott kimenet egy példára. A SURF detektorral észlelt pontok felső korlátja 50 volt. Jól látszik ezen a példán, hogy sikerül elkapnia a logó emberi szem számára is jellemző pontjait.

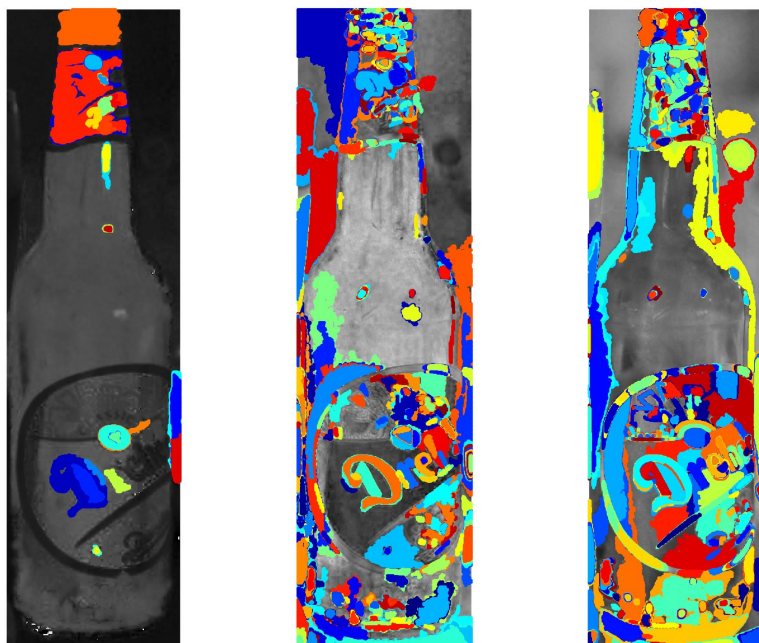
A SURF számunkra nagy előnyét a 4. számú ábra érzékelteti. A baloldalon egy nagyjából steril kép, a jobboldalon pedig egy valós helyzetbeli példa. Ezen kép tanuló algoritmus nélkül, pusztán a két kép domináns feature pontjai összepárosításával készült. Ebben az esetben 100 volt a pontok maximális száma.



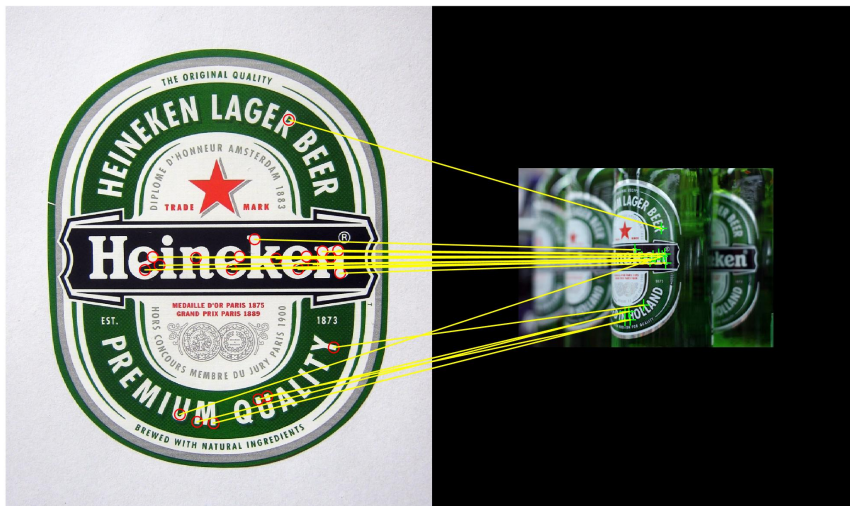
1. ábra. Szürkeárnyaltos képen detektált SURF pontok



2. ábra. Hue, saturation és value csatornán detektált SURF pontok



3. ábra. Hue, saturation és value csatornán detektált régiók MSER-rel



4. ábra. SURF pontok párosítása

Hivatkozások

- [1] Lowe, David G. "Object recognition from local scale-invariant features." *Computer vision, 1999. The proceedings of the seventh IEEE international conference on*. Vol. 2. Ieee, 1999.
- [2] Bay, Herbert, Tinne Tuytelaars, and Luc Van Gool. "Surf: Speeded up robust features." *Computer vision–ECCV 2006*. Springer Berlin Heidelberg, 2006. 404-417.