

Dokumentáció

Készítette: Geiger Boldizsár (RP6OUG)

Előkészületek, bemutatás

Egyetemi beadandó feladatomban egy olyan alkalmazásra esett a választás, amelyben az adott arc érzékelésére és feldolgozására kerül sor. A bemeneti érték lehet egy állókép, de igény szerint lehetőség van camera feedből kinyert információval dolgozni (ennek természetesen feltétele valamilyen USB-n csatlakoztatható vagy beépített rögzítő eszköz, webcam).

A projekt Python 3.6-ban készült, az alábbi könyvtárak felhasználásával:

- OpenCV
- NumPy
- dlib
- os
- glob
- PIL
- matplotlib

A megfelelő futás érdekében ajánlott az általános könyvtárakon kívül telepíteni a külsőket is, amik a listában más színnel kerültek jelölésre. A telepítés legegyszerűbben terminalon keresztül ajánlatos, mindegyik könyvtár elérhető hivatalos forrásból.

Először telepítsük a „PIP” (Pip Installs Packages)-et, egy cross-platform csomagkezelőt ami segít a Python csomagok installálásában:

```
# apt install python3-pip
```

Ezután „pip” által telepítsük a szükséges könyvtárakat, figyelve arra, hogy a megfelelő Python verzió mappájába kerüljön:

```
# pip3 install opencv-python numpy dlib pillow matplotlib
```

Következő lépésként klónozzuk a github repo-t, hogy az állomány futtatható legyen:

```
# git clone https://github.com/gboldi19/Background\_removal
```

Amint ezekkel megvagyunk, a program futásra kész. Működésével kapcsolatos információkkal a dokumentum további részében találkozhat.

(Opcionális): Bár a repository tartalmazza, lehetőség van saját HAAR Cascade dataset letöltésére, ha nem feltétlenül a frontális arcdetektálás a cél.

Megvalósítás módszere

Mintaillesztés az egyik legegyszerűbb módszer, amely objektumok felismerését teszi lehetővé. Egyetlen problémája, hogy főként ideális-, majdnem laborkörnyezetben használható hatalmas pontossággal. Több tényező is ronthat a végeredményen, köztük zaj, távolság, fényviszonyok, egyéb objektum a kép látószögében, ami takar, hogy a detektálni kívánt felület vagy mozgásban van, élessége kérdéses, stb.

A mintaillesztés folyamán 2 képet hasonlítunk össze, az egyik a mintakép ami tartalmazza azt, amit keresünk, a másik pedig az a kép, ami azt tartalmazza amin keresünk.

Arc és arcélek detektálása tanított modellek (dataset) segítségével is lehetséges, minél nagyobb méretű egy ilyen állomány, annál jobban segítheti elő a végeredmény precízségét.

2 ilyen datasetet is alkalmaztam a program vázaként, az egyik Haar-szerű jellegzetességek alapján ismeri fel a kívánt objektumot, a másik HOG (Histogram of Oriented Gradients) jellegzetesség által.

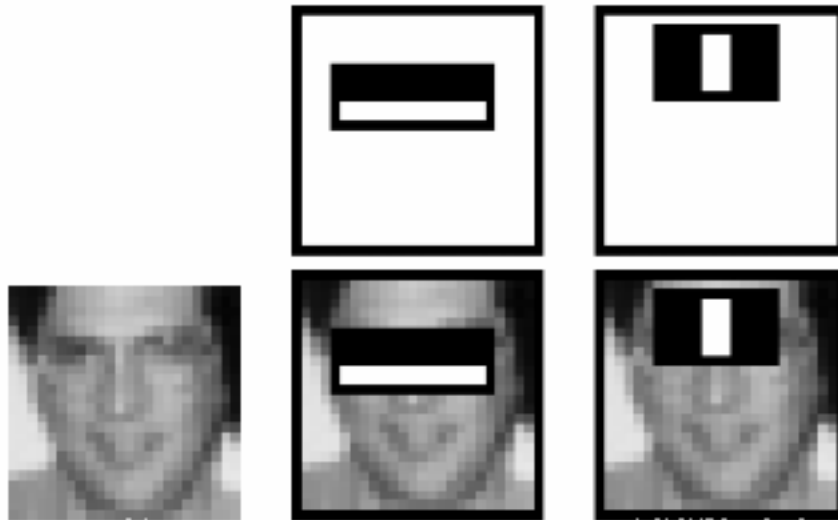
A kaszkád osztályozó alapú objektum felismerésnél felhasznált a Haar-szerű jellegzetességek (Haar-like features) nevüket a Haar wavelet-ekhez való hasonlóságuk kapcsán kapták. A módszer eredeti változatát Paul Viola és Michael Jones publikálták 2001-ben.

A technikát első sorban az arcfelismerés motiválta, ezzel is demonstrálták, és ezzel a technikával sikerült az első valós idejű arcfelismerők egyikét is megvalósítani, amely annak idején egy Intel Pentium III processzoron 15 képkocka/másodperc sebességgel futott. A publikált objektum felismerési keretrendszer a Haar-szerű jellegzetességek és három főbb ötlet felhasználásával képes elérni a gyors és hatékony objektum felismerést.

A technika nem közvetlenül kép intenzitás értékekkel dolgozik, hanem olyan jellegzetességek készletét használja fel, amelyek a Haar-wavelet-ekre emlékeztetnek, ezek a Haar-szerű jellegzetességek. A felismerési eljárás a képeket egyszerű jellegzetességek értékei alapján osztályozza. A jellegzetességek felhasználásának számos előnye van a hagyományos, a pixelek RGB értékeinek (intenzitásainak) közvetlen felhasználásához képest. Kizárólag pixeleken dolgozni 10 meglehetősen számításigényes folyamat, amelynél a jellegzetesség alapú rendszerek jelentősen gyorsabbak. Ezen túl a jellegzetességek rögzíthetik az olyan alkalmi ismereteket is, melyeket nehéz megtanulni véges számú tanítási adat alapján. A módszerben felhasznált egyszerű jellegzetességek a Haar wavelet-ekhez hasonlítanak. Egy Haar-szerű jellegzetesség szomszédos, azonos méretű és alakú téglalap alakú területeket vizsgál, az ilyen területeken kiszámítja a pixel intenzitások összegeit, és ezeknek az összegeknek a különbségeit vizsgálja (egymáshoz képest).

Ezen különbségek alapján az egyes területeket (egymáshoz képest) sötétnek vagy világosnak nyilvánítja, mely alapján a kép részeit kategorizálni tudja. Az eredeti publikációban három fajta jellegzetesség szerepel: úgynevezett kettő-téglalap, három-téglalap és négy-téglalap jellegzetesség, melyek 2, 3 illetve 4 téglalap alakú területet vizsgálnak. A kettő-téglalap jellegzetesség két horizontálisan vagy vertikálisan szomszédos területen vizsgálja a pixelek összegeinek a különbségét. A három-téglalap jellegzetesség két szélső terület összegének és a

köztük lévő középső terület összegének a különbségét vizsgálja. A négy-téglalap két-két átlós terület összegei közti különbséget vizsgálja.



Ezen jellegzetességek működésének szemléltetésére jó példa, hogy mivel egy emberi arcon a szemek régiója általában sötétebb az az alatti orca régiójánál, egy gyakori Haar-szerű jellegzetesség lesz egy kettő-téglalap, amelynek a felső része egy szem területén helyezkedik el és sötétebb, az alsó része pedig az arc területén van és világosabb. Egy adott osztályozóban felhasznált jellegzetesség jellemzői az alakja (az előbbi készletből), mérete és pozíciója. A detektor alapértelmezett 24×24 pixeles felbontásán elhelyezhető teljes készlete ezeknek a jellegzetességeknek így több, mint százezer.

A téglalap jellegzetességek meglehetősen primitívek sok más alternatívához képest. Ezek a jellegzetességek bár érzékenyek a sávokra, élekre és egyéb egyszerű struktúrákra, viszonylag durvának mondhatóak, önmagában gyenge osztályozók és nincs sok információtartalmuk. Téglalap jellegzetességek megfelelően nagy halmaza azonban már képes az objektum részletes, pontos jellemzésére. A Viola–Jones objektum-felismerési keretrendszerben az osztályozók halmazai kaszkádokba szerveződnek, hogy így együtt egy erős osztályozót alkossanak. A jellegzetességeknek a nagy száma és az integrál képek segítségével is biztosított hatékonysága kompenzálja az egyszerűségüket.