

Pràctiques de VC/PSIV

Felipe Lumbreras Ruiz

altra banda veureu que en essència els passos que hem de seguir són els mateixos.

Lab 1: Detecció de vehicles en seqüències de tràfic

1.1 Introducció

Per aquesta pràctica hem escollit un problema senzill on aprendreu a detectar objectes en una seqüència d'imatges, concretament cotxes en una seqüència de tràfic com la que es mostra a la Figura 1.



Figura 1: seqüència de tràfic.

Implementareu un *background subtraction* que consisteix a generar un model de fons on no apareguin cotxes per després treure aquest fons a les imatges de la seqüència, per tal de detectar els objectes que no pertanyen al fons (Figura 2). En essència, el primer pas genera un model de fons estàtic, en el nostre cas una imatge de la carretera sense vehicles, encara que ho refinarem com veurem més endavant. El segon pas extreu els objectes que no pertanyen a aquest fons. Bé, en realitat, cercarem aquells píxels que difereixen en major o menor grau d'aquest fons. Finalment mirarem respecte d'un groundtruth si ens apropem a una bona solució.



Figura 2: segmentació d'un cotxe.

Resoleu el problema amb tots dos entorns perquè veieu les diferències de llenguatge i de funcions a utilitzar, però per

1.2 Objectius

- Treballar amb eines bàsiques per resoldre un problema de visió concret.
- Aprendre a treballar amb imatges, seqüències d'imatges i vídeos. Manipular dades públiques etiquetades d'un problema concret.
- Resoldre un problema de *background subtraction* amb la creació d'un model de fons.
- Adonar-se'n de la diferència entre l'expectativa de la solució i el resultat aconseguit. Quantificar aquestes diferències amb diferents estimadors.

1.3 Puntuació, lliurament i extres

Solucionar el problema en un dels dos entorns puntua fins a 6 punts (inclou bons resultats, codificació polida i respondre a les preguntes en la sessió de lliurament).

Solucionar el problema en l'altre entorn dóna fins a 3 punts més.

Enunciar solucions raonades a la proposta final (bondat dels resultats) puntua fins a 1 punt més.

1.4 Tasques

1.4.1 Tasca 1 - Carregar el dataset i dividir-lo en dues parts iguals, una per modelar el fons i l'altra per detectar [+ 0.5]

El dataset es pot descarregar de la pàgina web: <http://changedetection.net/> → Datasets → 2014 Dataset → Baseline → Highway

Es faran servir les imatges de la carpeta *input* que van de la 1051 (in001051.jpg) a la 1350 (in001350.jpg). Les 150 primeres per entrenar i les 150 restants per segmentar. No obstant això, més endavant farem servir també les de groundtruth per comparar així que manteniu-les a prop.

Funcions de MatLab útils per aquesta tasca:

```
% Llistar imatges d'una carpeta.
files = dir('*.png');
% Llegir una imatge.
im_color = imread('nom_de_la_imatge.png');
% Convertir la imatge en escala de gris.
im_grey = rgb2grey(im_color);
```

Funcions de Python útils per completar aquesta tasca:

```
# Llistar imatges d'una carpeta.
files = os.listdir(path);
# Llegir una imatge en escala de gris.
```

```
skimage.io.imread(fname, as_grey=True)
# Recordeu instal·lar previament skimage i fer
# els imports que pertoqueu.
```

Aquest concepte de tenir diferents subgrups de dades serà una pràctica habitual més endavant quan es vulgui fer aprenentatge. El primer grup de dades (imatges) seria el de **training**, i el segon el de **test**. En cap cas les dades de test han de contribuir a la solució, tant és així que de vegades aquestes dades de tests no són accessibles i queden relegades a una plataforma externa d'avaluació. Finalment, en el cas d'algorismes que tenen paràmetres que s'han d'ajustar moltes vegades fent servir un tercer grup de dades anomenat **validation**.

Aquestes dades, per cadascuna de les imatges, també tenen associat un resultat desitjat que anomenarem **groundtruth** i ens permetrà obtenir un resultat objectiu i mesurable, que permetrà comparà diferents propostes de solució.

1.4.2 Tasca 2 - Calcular la mitjana i desviació estàndard [+ 0.5]

Aquests paràmetres ens permetran definir el nostre model gaussià de la variació dels píxels en funció d'on són dins la imatge. Obtenim la mitjana μ punt a punt, que seran els valors que més sovint trobarem a tots els píxels. Aquesta mitjana té forma (mida) d'imatge i per tant serà fàcil de visualitzar. La desviació estàndard σ ens dirà si hi ha molta o poca variació en els píxels, i com el cas anterior també té la mida d'una imatge i la podem representar ajustant la visualització al rang que toca.

Abans de fer les visualitzacions feu una predicció “mental” del que us trobareu. Aquesta manera de testejar els algorismes de visió ens ajuda a depurar mentre estem construint el codi (revisant aquelles parts si el que preveiem no és el que surt), i a descobrir comportaments propis, tendències, singularitats del problema que tenim entre mans.

$$\mu = \frac{1}{N} \sum_{i=1}^N x_i$$

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \mu)^2},$$

Funcions de MatLab útils per completar aquesta tasca:

```
% Mitjana d'un vector.
im_mean = mean(im);
% Buscar una funcio per una paraula clau.
lookfor deviation
% Obtenir informacio sobre com funciona
% una determinada funcio.
help std % informacio a la linia de comandes
doc std % informacio en una pagina aillada
```

Alguns enllaços Python útils per aquesta tasca:
<http://docs.scipy.org/doc/numpy/reference/>

[generated/numpy.mean.html](http://docs.scipy.org/doc/numpy/reference/generated/numpy.mean.html)
<http://docs.scipy.org/doc/numpy/reference/generated/numpy.std.html>

Els resultats que s'obtenen un cop finalitzada aquesta tasca són la mitjana (Figura 3) que ens representa el fons de la imatge sense cotxes i la desviació estàndard, que ens dóna una idea de la variació que hi ha hagut a cada píxel (ho podem interpretar com si un píxel ha canviat molt, han passat un o diversos cotxes, o si no ha canviat gens, ha estat sempre part del background).



Figura 3: Mitjana de la seqüència d'entrenament.

1.4.3 Tasca 3 - Segmentar cotxes restant el model de fons [+ 1.0]

Per cada píxel resteu el fons. Tots els píxels que siguin iguals a 0 és perquè no han variat, per tant pertanyen al fons (valor 0). Tots els altres píxels, independentment del valor han de tenir valor 1 (hi ha hagut moviment). Per no ser excessivament estrictes podeu definir un llindar (*thr*) per decidir a quin conjunt pertany el píxel, els píxels que estiguin a prop de 0 ($0 \pm thr$) seran fons.

Hi ha diverses maneres d'afrontar aquest etiquetatge dels píxels a 0 o 1. Una manera consisteix en “buscar” els que compleixen el criteri i posar-los al valor que volem. Una altra possibilitat consisteix directament a aplicar la condició ($<$ o $>$) sobre la mateixa imatge, que ens retornarà una imatge amb valors booleans (0,1). El que heu d'intentar és no fer servir bucles que recorrin tota la imatge fent aquesta tria píxel a píxel, ja que trigariau molt a trobar la solució.

1.4.4 Tasca 4 - Segmentar cotxes amb un model més elaborat [+ 1.0]

$$P_{i,j} = \begin{cases} 1, & |I_{i,j} - \mu_{i,j}| > \alpha \sigma_{i,j} + \beta \\ 0, & \text{otherwise} \end{cases}$$

on α i β són paràmetres amb els quals podeu jugar.

Els resultats que heu d'obtenir per una imatge com la de la Figura 4 han de ser similars al de la Figura 5



Figura 4: Exemple d'imatge.



Figura 5: Exemple de detecció.

1.4.5 Tasca 5 - Gravar un vídeo amb els resultats [+ 2.0]

Graveu un vídeo amb els resultats obtinguts. Aquest vídeo pot tenir el resultat de la subtracció del fons (seqüència binària) o podeu manipular el resultat per donar més informació. Per exemple, amb morfologia (erode, dilate, opening, closing) treure soroll (punts i forats petits aïllats) per tal de compactar el que són les taques dels cotxes.

Alguns enllaços interessants per aquesta tasca:

<https://es.mathworks.com/help/images/dilate-an-image.html>

<https://es.mathworks.com/help/images/erode-an-image.html>

1.4.6 Tasca 6 - Avalua la bondat dels teus resultats [+ 1.0]

Hem agafat 150 imatges de training per modelar el nostre background i hem fet un vídeo amb les altres 150 imatges de test. En aquesta tasca analitzarem com de bé els nostres resultats s'apropen a una solució bona. Per tal de fer això necessitem una "veritat" que anomenem *groundtruth* on s'ha fet una tasca manual d'anotació per part d'experts que decideixen quin és el resultat correcte. Si mires amb deteniment les dades de *grountruth* veuràs que hi ha una anotació de fons – vehicles – ombres – fronteres.

Si per cadascuna de les 150 imatges en les quals podem calcular el resultat tenim com ha de ser el resultat esperat, el que podem fer és avaluar si ens apropem a aquesta solució perfecta. En el codi de suport us donem un conjunt d'indicadors que se solen utilitzar per fer avaluacions

(*accuracy*, *precision*, *recall*, *f1score*). Més endavant en el curs veurem que són, de moment centreu-vos en l'*accuracy* que serà 1 si tot és perfecte, i 0 si és un desastre. Podeu analitzar que passa (si augmenta o disminueix l'*accuracy*) quan es fa algun tipus de filtratge, o si fem servir un model senzill, o si fem servir una determinada α o β .

Treu resultats en mitjana de l'*accuracy* per tres casos concrets.

1.4.7 Tasca 7 - Repetir la pràctica en un altre entorn [+ 3.0]

La part extra per aquesta pràctica consisteix a trobar un altre entorn de processament d'imatges que tingui funcionalitats similars a les que fareu servir i tornar a resoldre el problema per aquest nou entorn. El nou entorn (llenguatge, llibreries, ...) és a la vostra elecció, analitzeu el que hi ha (basats en java, basat en C++, Python, MatLab, ...) i escolliu el que més us agradi.

1.4.8 Tasca 8 - Com obtenir velocitats? [+ 1.0]

Reflexiona com faries un sistema que detectés la velocitat dels cotxes. No l'has de programar, només detectar quins són els punts més conflictius d'aquesta tasca i proposar solucions. No hi ha una solució única. Es valora tenir en compte la casuística d'aquestes imatges i que s'aportin solucions imaginatives. Si no saps com es fa alguna cosa només has de ficar-li nom i descriure el que fa.

1.5 Entrega

El lliurament és presencial a la següent classe de pràctiques. Heu de mostrar els resultats (sense problemes, penjades, parant a cada resultat avaluable), heu de mostrar el codi també i explicar la solució a la tasca 7. Els codis s'han de lliurar fins aquell mateix dia i no hi haurà nota de la pràctica si no es puja el codi. Tot el material (codis sense imatges de cap mena in/out i document associat a la tasca 7 (txt)) es lliura en un fitxer comprimit que es pujarà a Caronte. Un fitxer per grup. Els codis i documents han de portar la informació dels integrants del grup: noms i NIUs. El fitxer ha de tenir el nom "Lab1_NIU1_NIU2.zip".

Agraïments

<http://changedetection.net/>

Referències

- ◊ Apunts de classe
- ◊ <https://docs.scipy.org/doc/numpy/user/numpy-for-matlab-users.html>