# VISIÓ PER COMPUTADOR

*Short Project*

**Facultat d'Informàtica de Barcelona**

**Manel Frigola**

**Joan Climent**

**Barcelona, Novembre de 2021**

1. **Objectius del projecte**

L’objectiu del projecte és implementar un sistema automàtic per detectar la mirada mitjançant visió per computador. El sistema ha de ser capaç de detectar on són els ulls i posteriorment mesurar l’angle horitzontal de la mirada en base a la posició del iris. Les imatges amb les que es treballaran seran amb un enquadrament tipus bust d’un usuari sense ulleres en una orientació vertical amb rotacions del cap poc apreciables; tal com es mostra a continuació en la imatge de la figura següent:



Fig 1. Imatge d’exemple d’una detecció d’ulls (imatge extreta de thispersondoesnotexist.com)

1. **Treball a realitzar**
2. Construir un detector de ulls de codi propi seguint les següents passes:
3. Recopilar una base de dades d’imatges d’ulls i no-ulls. Les imatges d’ulls hauran de ser de la parella d’ulls sencera enlloc de imatges contenint un únic ull.
4. Programar un codi que calculi un vector de característiques (anomenat descriptor) d’imatges de nivells de gris. Aquest descriptor serà d’inspiració pròpia del grup i **no estarà basat en característiques bàsiques d’imatges binaritzades** (funció *bwconncomp*).
5. Calcular la taula de descripcions de les imatges de la base de dades de l’apartat i) conjuntament amb l’output (Ulls / No-Ulls). Desar la taula d’observacions en un fitxer format ***csv***.
6. Construir un classificador amb algunes de les funció de classificació que Matlab incorpora (vegeu *Train Ensemble Classifiers Using Classification Learner App*)
7. Realitzeu un estudi estadístic de les capacitats del classificador/s d’ulls: matriu de confusió, precisió, etc., que heu calculat en l’apartat a).
8. Donada una imatge d’un ull, mesurar (utilitzant codi propi) el desplaçament horitzontal del iris en base al llagrimall, l’iris i la comissura contraria al llagrimall.



Fig 2 Exemple idealitzat de la detecció d’iris (rectangle central), llagrimal (rectangle dret), comissura (rectangle esquerra) i desviació de la mirada (en blau).

1. Mostrar la detecció d’ulls i mirada en seqüències de vídeo.
2. Els grups de laboratori que estiguin conformats per tres membres hauran de detectar, a més a més la posició de les celles mesurant l’amplada de l’entrecella.
3. Opcional: a partir de la desviació de l’iris (i la mesura de l’entrecella) aproximar la direcció de la mirada i moure horitzontalment un cursor en pantalla.
4. **Es demana**
5. **Abans de la data del *Checkpoint***

Cal que lliureu un informe comentant de cadascuna de les etapes amb una breu descripció, resultats obtinguts i destacant les parts més rellevants del vostre programa, quins problemes us heu trobat i com els heu resolt. Mostreu també en quins casos el vostre programa no funciona. En especial l’informe ha de contenir:

1. Una explicació elaborada del mètode de classificació implementat. És important detallar: Mida de les imatges i quantitat de les imatges de la BD d’ull/no-ulls. Descripció del vector de característiques. Dimensions del vector de característiques i velocitat de càlcul.
2. Resultats obtinguts, mostrant el rati d'encerts global, falsos positius i falsos negatius.
3. Enumeració de les funcions utilitzades, deixant ben clar quines són les implementades per vosaltres. Tot el software provinent d'altres fonts, ha d'estar correctament referenciat. EN CAS CONTRARI ES CONSIDERARÀ PLAGI.
4. L’última secció de l’informe serà un annex amb tot el codi
5. **Abans de la data d’entrega final**

Cal que lliureu un nou informe que contingui

1. Les novetats respecte de la primera entrega. Una explicació detallada de les modificacions que s’han hagut de fer a l’algorisme de l’etapa 1. Criteris utilitzats per a detectar la mirada. Descripció dels nous mètodes implementats.
2. Enumeració de les noves funcions utilitzades, deixant ben clar quines són les implementades per vosaltres. Tot el software provinent d'altres fonts, ha d'estar correctament referenciat. EN CAS CONTRARI ES CONSIDERARÀ PLAGI.
3. L’última secció de l’informe serà un annex amb tot el codi nou
4. **El dia de la presentació**

Haureu d'executar programa MATLAB que implementi el sistema de reconeixement demanat en aquesta pràctica. Tingueu en compte que el dia de la presentació es provarà el vostre programa de reconeixement amb noves imatges, no incloses en el *dataset*.

1. **Indicacions**

Les següents indicacions són orientatives, no són d’obligat compliment.

* 1. **Construcció de la base de dades**

Per construir el vostre detector d’ulls haureu de obtenir una base de dades d’imatges d’ulls i no ulls. Normalment s’utilitzen bases de dades molt extenses i variades, tenir accés a aquestes base de dades no està al nostre abast, però existeixen a Internet algunes DB prou decents, podeu veure un compendi d’aquestes base de dades a *https://www.face-rec.org/databases/*. A la carpeta de l’assignatura Short Project QT18 trobareu una serie d’imatges de cares *BioID\_NNNN.png* associades a un fitxer *BioID\_NNNN.eye* que indica la posició (x,y) dels ulls. Si us voleu fer una BD d’imatges d’ulls, a la web *thispersondoesnotexist.com*, sempre es genera una nova imatge d’una persona (que no existeix) amb la posició dels ulls en les mateixes coordenades. La BD de no ulls, hauria de ser molt més extensa que la de ulls amb una relació 1/50 amb altres parts de la cara i possibles fons d’interiors. Per construir la BD de no-ulls podeu agafar aleatòriament elements de la cara que no siguin ulls i també del fons de les imatges anteriors.

* 1. **Mida de les imatges**

Escaleu les mides de les imatges per codi (*imresize*) a valors adequats que el vostre computador pugui manegar en un temps prou raonable. Treballar a una escala reduïda, de per exemple 256x256 píxels per detectar el parell d’ulls, i a una resolució superior (1024x1024 píxels) per detectar l’iris és una bona estratègia que comporta un bon compromís entre velocitat i precisió.

* 1. **Detecció de la cara**

Per detectar la posició aproximada de la cara de l’usuari podeu fer us (no és imprescindible) de la funció *CascadeObjectDetector* de la llibreria *Computer Vision Toolbox*. En el següent codi (Fig. 3) d’exemple es mostra com es pot utilitzar la variable *Detector* per localitzar la posició de la cara (funció *step*).

% Create a cascade detector object.

Detector = vision.CascadeObjectDetector('FrontalFaceLBP');

% Read a video frame and run the face detector.

videoReader = VideoReader('face 1.webm');

while hasFrame(videoReader)

% get the next frame

videoFrame = readFrame(videoReader);

bbox = step(Detector, videoFrame);

% Draw the returned bounding box around the detected face.

videoFrame = insertShape(videoFrame, 'Rectangle', bbox);

imshow(videoFrame);

end

Fig 3. Codi d’exemple per detectar la cara utilitzant la llibreria *Computer Vision Toolbox*.

* 1. **Reducció de l’àrea de cerca**

En base a la detecció de la posició de la cara del usuari es pot reduir l’àrea de cerca dels ulls i d’aquesta manera reduir el temps d’execució del programa. Amb la mateixa idea de reduir el temps d’execució del codi, es pot mesurar el desplaçament de la mirada en base a la posició detectada dels ulls. El codi que mesuri el desplaçament de la mirada pot centrar-se en només un únic ull, i així realitzar només un codi específic per localitzar, p.ex. l’ull dret i no caldria fer-ho pel ull esquerra.

1. **Valoració del projecte**

Tot i que un correcte funcionament del programa serà benvingut, no és l’objectiu principal del projecte. Els factors que més es valoraran del vostre projecte són:

* Un informe del projecte ben estructurat i complet.
* Comparació del resultat de diferents classificadors, diferents descriptors, etc.
* Utilització de noves/diferents estratègies.
* Anàlisi estadístic del rendiment del detector i sintonització dels paràmetres de les funcions utilitzades.
* Avaluar el funcionament del vostre programa amb imatges de test prou variades.