
Pràctica visió per computador

Tutor de la pràctica : Josep Forest

Gil Gassó Rovira i Marc Sànchez Pifarré, GEINF
(UDG-EPS)

2019-04-14

Contents

Pràctica 2 - Robòtica	2
Descripció general del document	2
Context en el que es realitza la pràctica	3
Hardware	3
Il·luminació	3
tauler d'escacs	3
Calibratge	3
Descripció de l'acció de calibrar	3
Procediment del calibratge	4
Pose Estimation	4
Anàlisi dels paràmetres.	5
T - Vector de translació	7
E - Matriu essencial	8
F - Matriu Fonamental	8
Annexes	9
Fitxers de la pràctica	9
Exemple de Calibratge amb una sola càmera	9
Captures	9
References	12

Pràctica 2 - Robòtica

Sobre què parlar :

- Eixos de referència
- Guanyança de la càmera

Parts de la pràctica :

Descripció general del document

En aquest document s'hi presenta el resum de conclusions extretes de la realització de la pràctica 2 de Visió per computador. En aquesta pràctica se'ns demana realitzar el calibratge d'una càmera i la representació d'objectes 3D en realitat augmentada.

Context en el que es realitza la pràctica

Hardware

El tipus de càmeres que s'utilitzen són les càmeres embbeded dels portàtils o les webcams del laboratori que són càmeres amb òptica pinhole que utilitzen sensors de tipus CMOS. De fet aquest tipus de sensors son els que trobem al 99% dels models de càmeres actuals.

Il·luminació

La il·luminació que s'utilitza per realitzar aquesta pràctica no és una il·luminació directe cap a l'objecte en qüestió (el tauler d'escacs) sinó que s'aprofita la il·luminació ambiental que es disposa al laboratori.

No s'ha aplicat cap tipus d'il·luminació diferent a l'ambiental però un altre tipus d'il·luminació que hauriem pogut utilitzar per a destacar els contrastos de les cantonades del tauler d'escacs pot ser qualsevol tipus d'il·luminació directe que realci el contrast, com per exemple un Ring o un Spot.

tauler d'escacs

Tal com es comenta a l'enunciat el tauler d'escacs és una eina simple que ens permet, mitjançant el contrast dels quadrats blancs i negres, realitzar un calibratge prou precís per a la representació de figures 3D en realitat augmentada.

El **contrast**, representat per les interseccions entre els quadrats blancs i negres del tauler d'escacs, hi juga un paper molt clau a l'hora de realitzar el calibratge per que ens proporciona uns punts fàcilment distingibles situats en un pla de l'espai.

Calibratge

Descripció de l'acció de calibrar

Donat un punt focal que té l'origen al sistema de coordenades de la càmera i un tauler d'escacs situat a un pla del sistema de coordenades del món, la feina del calibratge emana sobre l'acció de representar els punts del pla imatge dins del sistema de coordenades del món i discretitzar-lo de manera que mitjançant la triangulació es pugui determinar propietats de l'objecte representat en el pla imatge, en aquest cas el tauler d'escacs.

Procediment del calibratge

Mitjançant el codi proporcionat i la seva execució s'han enregistrat diferents fotogrames on hi apareix l'objecte que s'utilitza per calibrar la càmera, (tauler d'escacs), representat en el diferents plans dins del sistema de coordenades del món.

Per cada representació esmentada el calibratge realitza els següents passos :

1. Situar l'origen del sistema de coordenades de la càmera en un punt del sistema de coordenades del món.
2. Projectar els punts, en aquest cas les interseccions dels quadrats del tauler d'escacs, al pla imatge dins del sistema de coordenades del món i determinar *entre altres* la distància focal.
3. Cada fotograma incorpora una petita distorsió proporcionada per la len pinhole, en concret una distorsió radial.
4. Es situa el sistema de coordenades de la imatge en 2D sobre el pla imatge.

En general OpenCV2 utilitza factors radials i tangencials per calcular la matriu de distorsió que permetrà calibrar la càmera per que el tipus de lent no afecti al resultat final de la discretització de la imatge, veure la referència 1 - *Camera Calibration with OpenCV*.

El que vam fer va ser reproduir el procés explicat a la referència 1 - *Camera Calibration with OpenCV* però en viu, és a dir sense passar-li imatges i fent que la captura de video cada x temps realitzés un fotograma. Quants més fotogrames li passem a la matriu de calibratge més precís és el resultat d'aquest calibratge i conseqüentment millora l'estimació de la posició de l'objecte representat en el sistema de coordenades del món.

Pose Estimation

En aquest apartat s'explicarà amb detall els components de la matriu de calibratge realitzada mitjançant la captura dels 10 fotogrames amb el sistema d'estereovisió (10 fotogrames cada càmera).

Es pot veure la sortida del procés de calibratge realitzat en l'apartat d'estereovisió al fitxer "sortida2cams.txt". En general el calibratge amb dues perspectives és el mateix que el calibratge amb 1 sola perspectiva però el què canvia és que amb 2 perspectives necessites quadrar els temps de captura i unificar els dos fotogrames simultanis mitjançant la instrucció `cv2.stereoCalibrate()`, veure la referència 2 - *Camera Calibration and 3D Reconstruction*.

A mesura que es van discretitzant els fotogrames es va emplenant el que s'anomena el model de la càmera (`camera_model`) que no deixa de ser el diccionari amb tots els punts que s'han recollit expressats de la següent manera :

```
1 Intrinsic_mtx_1 [[553.05486348  0.          324.3926378 ]
2 [  0.          553.45015689 229.8780157 ]
3 [  0.           0.           1.           ]]
4 dist_1 [[ 3.58530635e-02 -9.48637868e-01 -5.26278465e-03 -3.87420281e
5         -03
6         4.29571670e+00]]
7 Intrinsic_mtx_2 [[715.35935416  0.          321.6634394 ]
8 [  0.          714.60277722 231.54998339]
9 [  0.           0.           1.           ]]
10 dist_2 [[-6.14052418e-02  9.93783844e-01 -4.92704324e-03  2.57797233e
11         -03
12         -6.33308179e+00]]
13 R [[ 0.99996558  0.00601564 -0.00571376]
14 [-0.00625317  0.99907669 -0.04250486]
15 [ 0.00545279  0.04253912  0.99907992]]
16 T [[ 4.23885796]
17 [ 0.96565357]
18 [-0.70508252]]
19 E [[ 8.56509219e-04  7.45509567e-01  9.34795664e-01]
20 [-7.28171856e-01 -1.84558831e-01 -4.23092921e+00]
21 [-9.92126618e-01  4.22913516e+00 -1.74654545e-01]]
22 F [[-5.12417102e-09 -4.45691648e-06 -2.06676613e-03]
23 [ 4.36098933e-06  1.10452530e-06  1.23451951e-02]
24 [ 3.23789006e-03 -1.69087327e-02  1.00000000e+00]]
```

El codi anterior és la sortida referent a la fusió de 2 fotogrames mitjançant la funció `cv2.stereoCalibrate` i rebent com a paràmetres els resultats per cada fotograma de la crida a la funció `cv2.calibrateCamera`.

Anàlisi dels paràmetres.

Com que el que es pretén és acabar presentant un punt 3D d'un sistema de referència del món a un punt 2D d'un sistema de referència de la càmera sobre el pla imatge, en el procés hi juguen els elements escenificats a la següent imatge :

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

Figure 1: Transformació d'un punt 3D a un punt 2D

on :

- (X, Y, Z) : son les coordenades del punt 3D en l'espai de coordenades del món.
- (u, v) : Cooedenades del punt 2D projectat sobre el pla imatge en pixels.
- (cx, cy) : Punt principal sobre el pla imatge.
- (fx, fy) : LLargades focals expressades en pixels.

Llavors el que es cerca és la matriu amb les components fx, cx, fy i cy que ens permetrà realitzar aquest procés.

Intrinsic_mtx_1 i Intrinsic_mtx_2

Resultat del cv2.calibrateCamera de la càmera 1 representat en format matriu :

```
1 [
2     [fx, 0,  cx]
3     [0,  fy,  cy]
4     [0,  0,  1 ]
5 ]
```

```
1 Intrinsic_mtx_1 [[553.05486348  0.          324.3926378 ]
2                  [ 0.          553.45015689 229.8780157 ]
3                  [ 0.          0.          1.          ]]
4
5 Intrinsic_mtx_2 [[715.35935416  0.          321.6634394 ]
6                  [ 0.          714.60277722 231.54998339]
7                  [ 0.          0.          1.          ]]
```

dist_1 i dist_2

Vector de distorsions format amb 5 coeficients de distorsió com :

```
1 [k1, k2, p1, p2, k3]
```

on :

- k1, k2 i k3 son coeficients de distorsió radials.
- p1 i p2 son coeficients de distorsió tangencials.

```
1 dist_1 [[ 3.58530635e-02 -9.48637868e-01 -5.26278465e-03 -3.87420281e
          -03  4.29571670e+00]]
2
3 dist_2 [[-6.14052418e-02  9.93783844e-01 -4.92704324e-03  2.57797233e
          -03 -6.33308179e+00]]
```

R - Matriu de rotació

Matriu de rotació que s'aplicaria entre el sistema de coordenades de la primera càmera i el sistema de coordenades de la segona càmera.

```
1 R [[ 0.99996558  0.00601564 -0.00571376]
2     [-0.00625317  0.99907669 -0.04250486]
3     [ 0.00545279  0.04253912  0.99907992]]
```

Podem comparar la matriu anterior amb la matriu identitat i ens n'adonarem que la càmera 2 pràcticament no està rotada en funció de la càmera 1.

```
1 I [[ 1  0  0 ]
2     [ 0  1  0 ]
3     [ 0  0  1 ]]
```

T - Vector de translació

Com que la imatge està presa des de dos posicions diferents tenim 2 punts focals, aquest vector representa la distància a l'espai entre el punt focal de la càmera 1 i el punt focal de la càmera 2.

Si pf1 és el punt focal a l'espai de la càmera 1 i pf2 és el punt focal a l'espai de la càmera 2, aplicant la translació T a pf1 tindriem pf2 sempre que estiguéssin en el mateix sistema de referència.

```
1 T [[ 4.23885796]
2     [ 0.96565357]
3     [-0.70508252]]
```

De fet en la matriu anterior veiem que la posició de la càmera 2 en funció de la càmera 1 és de 4.24cm sobre l'eix x, 0.97 cm sobre l'eix y i -0.7 cm sobre l'eix z, que significa que la càmera 2 està enganxada a la pantalla del portàtil on hi ha la càmera 1 (embedded) i que es situa 4.24cm a la dreta, un pèl més amunt i un fil més enfora.

E - Matriu essencial

Matriu que agafa els punts extrems de les imatges de les dues càmeres i construeix una matriu que representa les posicions relatives entre una càmera i l'altre, acaba relacionant les dues captures per poder extreure el valor conjunt.

```
1 E [[ 8.56509219e-04  7.45509567e-01  9.34795664e-01]
2    [-7.28171856e-01 -1.84558831e-01 -4.23092921e+00]
3    [-9.92126618e-01  4.22913516e+00 -1.74654545e-01]]
```

cv2 ho utilitza aquesta matriu com a transformació entre la vista de la càmera 1 i la càmera 2. La matriu E relativitza la perspectiva entre una càmera i l'altre.

Es fa servir també per al càlcul de la matriu fonamental.

Veure la referència 3 - *findEssentialMat* per a una explicació més acurada del càlcul.

$$F = cameraMatrix2^{-T} E cameraMatrix1^{-1}$$

Figure 2: Càlcul de la matriu fonamental

F - Matriu Fonamental

Ve a ser al forma bilineal de fer correspondre 2 punts de 2 sistemes de coordenades diferents. Ens permet relacionar el punt entre les dues càmeres.

```
1 F [[-5.12417102e-09 -4.45691648e-06 -2.06676613e-03]
2    [ 4.36098933e-06  1.10452530e-06  1.23451951e-02]
3    [ 3.23789006e-03 -1.69087327e-02  1.00000000e+00]]
```


Annexes

Fitxers de la pràctica

- **calibratge.py** : Captura 10 cops 10 imatges amb una diferència de 0.5 segons cada imatge i calibra la càmera i extreu un fitxer .yaml.
- **edges.py** : Obre el fitxer .yaml resultant de fer la calibració i pinta els eixos d'un objecte 3D sobre els fitxers que hi ha al fitxer imatges. (les variables x i y son les que estipulen el nombre de punts en l'eix x i l'eix y)
- **photomaton.py** : simplement realitza una imatge i la desa al directori imatges. S'ha utilitzat per depurar.
- **stereoCalibration.py** : realitza una calibració dual amb 10 captures per a cada càmera sincronitzades.
- **calibration.yaml** : Exemple de model de calibració resultat de un calibratge.py.
- **calibration.z.npz** : Calibratge resultant en format z.npz.
- **README.md** : Font d'aquest document.
- **sortida2cams.txt** : El resultat de l'execució del calibratge de dues càmeres (el que es va pintar per pantalla portat cap a un fitxer).

Exemple de Calibratge amb una sola càmera

Es veu representada la matriu de coeficients i el vector de distorsions.

```
1 camera_matrix:
2 - [992.2030465290455, 0.0, 672.9783890677685]
3 - [0.0, 988.9628211272332, 322.06622116670366]
4 - [0.0, 0.0, 1.0]
5 dist_coeff:
6 - [-0.09502689555607022, 0.10223723893932801, -0.0036757148328838926,
7   0.0022448212909430294,
8   0.058308034924238285]
```

Captures

A continuació annexem alguna captura per escenificar la feina feta.

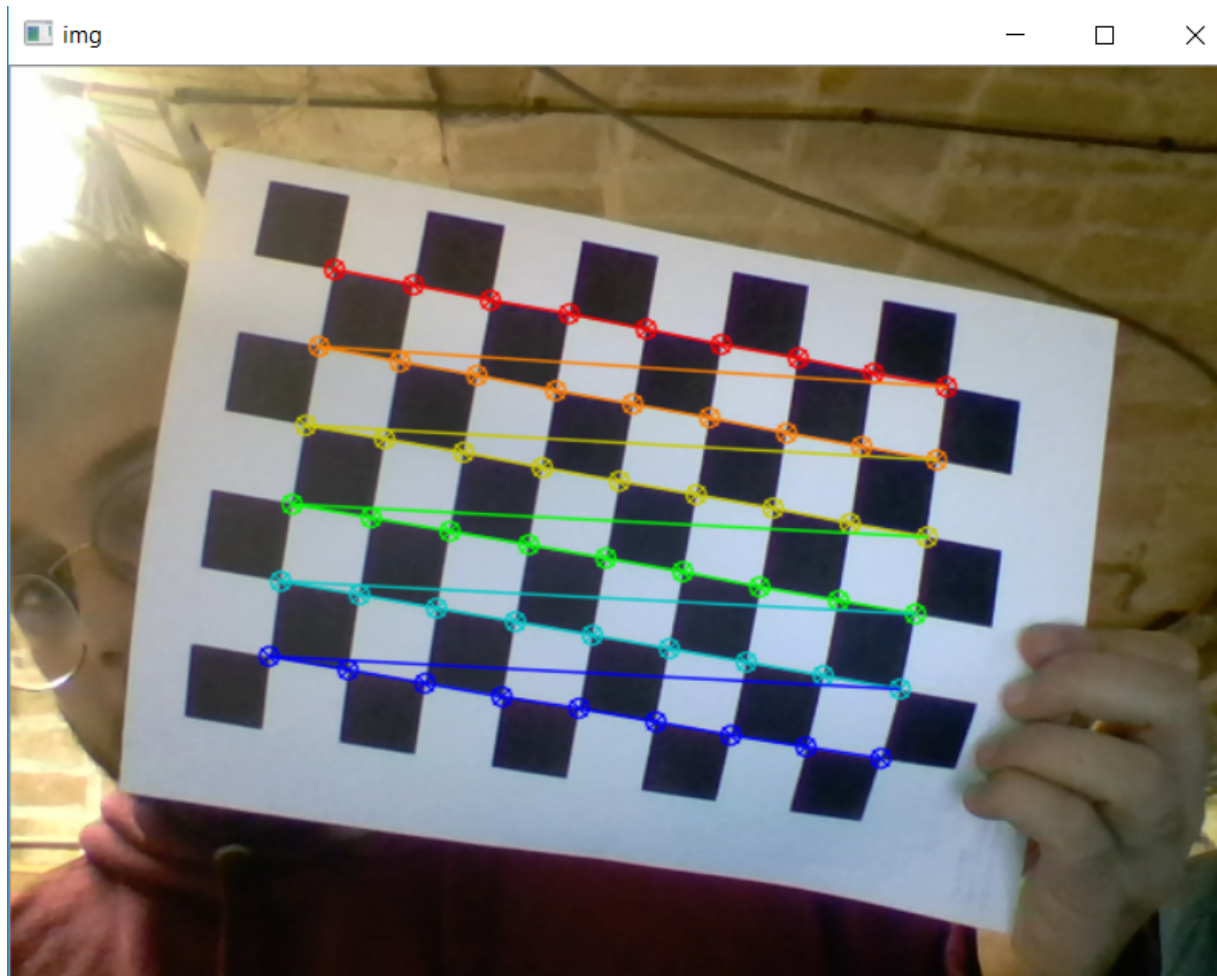


Figure 3: Calibratge

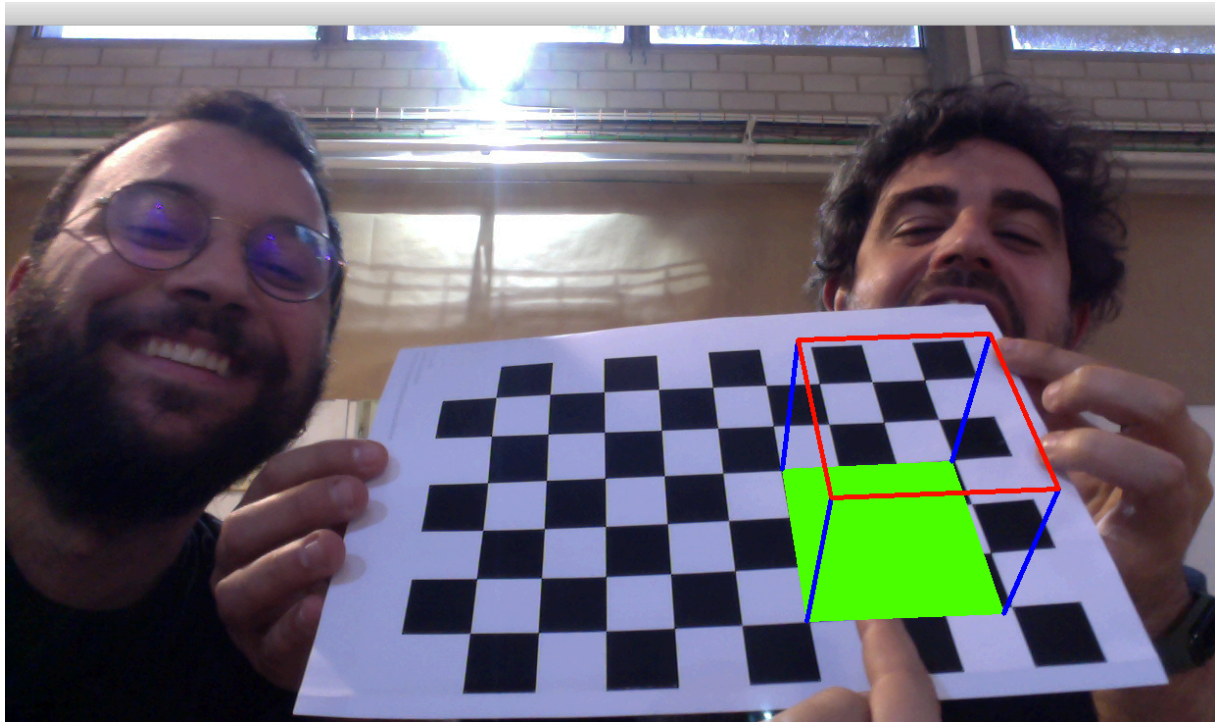


Figure 4: Representació objecte 3D

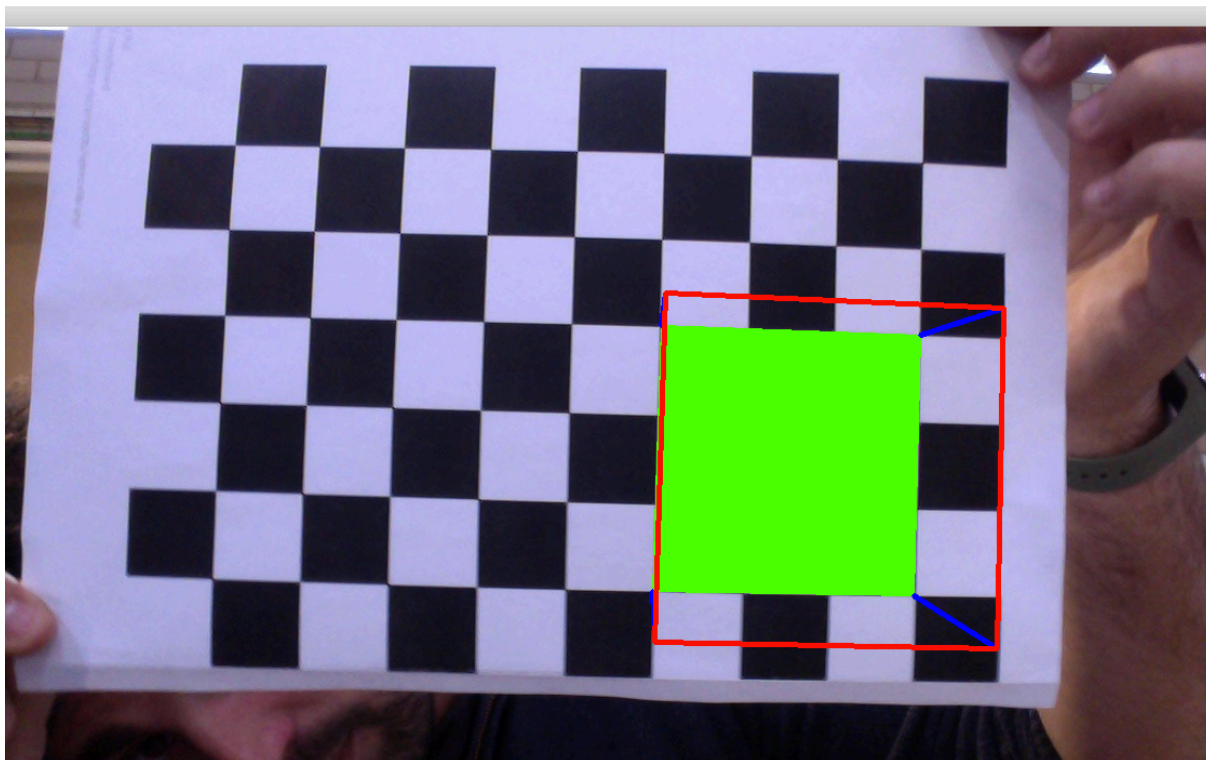


Figure 5: Representació objecte 3D 2

References

- 1 - Camera Calibration with OpenCV
- 2 - Camera Calibration and 3D Reconstruction
- 3 - findFundamentalMat i findEssentialMat