

PracticaStereo2018-Andres

December 4, 2018

1 Práctica de reconstrucción. Parte II. Visión estéreo

Visión Computacional 2018-19 Practica 2. 29 de octubre de 2018

Este enunciado está en el archivo "PracticaStereo.ipynb" o su versión "pdf" que puedes encontrar en el Aula Virtual.

1.1 Objetivos

Los objetivos de esta práctica son: * reconstruir puntos de una escena a partir de una serie de correspondencias manuales entre dos imágenes calibradas; * determinar la geometría epipolar de un par de cámaras a partir de sus matrices de proyección; * implementar la búsqueda automática de correspondencias que use las restricciones impuestas por la geometría epipolar, aplicando para ello métodos de cortes de grafos; * realizar una reconstrucción densa de la escena.

1.2 Requerimientos

Para esta práctica es necesario disponer del siguiente software: * Python 2.7 ó 3.X * Jupyter <http://jupyter.org/>. * Las librerías científicas de Python: NumPy, SciPy, y Matplotlib. * La librería OpenCV * La librería PyMaxFlow

El material necesario para la práctica se puede descargar del Aula Virtual en la carpeta `MaterialesPractica` del tema de visión estéreo. Esta carpeta contiene: * Una serie de pares estéreo en el directorio `images`; el sufijo del fichero indica si corresponde a la cámara izquierda (`_left`) o a la derecha (`_right`). Bajo el directorio `rectif` se encuentran varios pares estéreo rectificados. * Un conjunto de funciones auxiliares de Python en el módulo `misc.py`. La descripción de las funciones puede consultarse con el comando `help` o leyendo su código fuente. * El archivo `cameras.npz` con las matrices de proyección del par de cámaras con el que se tomaron todas las imágenes con prefijo `minoru`.

1.3 Condiciones

- La fecha límite de entrega será el martes 20 de noviembre a las 23:55.
 - La entrega consiste en dos archivos con el código, resultados y respuestas a los ejercicios:
1. Un "notebook" de Jupyter con los resultados. Las respuestas a los ejercicios debes introducirlas en tantas celdas de código o texto como creas necesarias, insertadas inmediatamente después de un enunciado y antes del siguiente.

2. Un documento "pdf" generado a partir del fuente de Jupyter, por ejemplo usando el comando `jupyter nbconvert --execute --to pdf notebook.ipynb`, o simplemente imprimiendo el "notebook" desde el navegador en la opción del menú "File->Print preview". Asegúrate de que el documento "pdf" contiene todos los resultados correctamente ejecutados.
- Esta práctica puede realizarse en parejas.

1.4 1. Introducción

En los problemas de visión estéreo se supondrá la existencia de un par de cámaras calibradas cuyas matrices de proyección \mathbf{P}_i vienen dadas por

$$\mathbf{P}_1 = \mathbf{K}_1 \cdot [\mathbf{I} \quad \mathbf{0}] \cdot \begin{bmatrix} \mathbf{R}_1 & \mathbf{t}_1 \\ \mathbf{0}^T & 1 \end{bmatrix},$$

$$\mathbf{P}_2 = \mathbf{K}_2 \cdot [\mathbf{I} \quad \mathbf{0}] \cdot \begin{bmatrix} \mathbf{R}_2 & \mathbf{t}_2 \\ \mathbf{0}^T & 1 \end{bmatrix}.$$

En esta práctica se usarán las matrices de proyección de dos cámaras para determinar la posición tridimensional de puntos de una escena. Esto es posible siempre que se conozcan las proyecciones de cada punto en ambas cámaras. Desafortunadamente, esta información no suele estar disponible y para obtenerla es preciso emplear el contenido de las imágenes (sus píxeles) en un proceso de búsqueda conocido como puesta en correspondencia. Conocer las matrices de proyección de las cámaras permite acotar el área de búsqueda gracias a las restricciones que proporciona la geometría epipolar.

```
In [1]: # uncomment to show results in a window
        %matplotlib tk
        import numpy as np
        import scipy.misc as scpm
        import scipy.ndimage as scnd
        import matplotlib.pyplot as ppl
        import numpy.linalg as npla
        import maxflow.fastmin
        import misc
```

```
In [2]: from pprint import pprint as pp
```

```
In [3]: import importlib as il
        # import pixiedust as pd
        # import pdb
```

```
In [4]: il.reload(misc)
```

```
Out[4]: <module 'misc' from '/Users/andreshg/Github/UPM/Computer_Vision/Segunda/Parte 2/misc.p
```

1.5 1. Reconstrucción

Teniendo un conjunto de correspondencias entre dos imágenes, con matrices de calibración P_i conocidas, es posible llevar a cabo una reconstrucción tridimensional de dichos puntos. En el fichero `cameras.npz` se encuentran las matrices de proyección para las dos cámaras. Para cargar este fichero:

```
In [5]: cameras = np.load("cameras.npz")
        P1 = cameras["left"]
        P2 = cameras["right"]
```

Todas las imágenes con el prefijo `minoru` comparten este par de matrices de proyección.

Leemos las imágenes y marcamos al menos seis puntos correspondientes en cada una de ellas.

```
In [6]: img1 = scpm.imread("images/minoru_cube3_left.jpg")
        img2 = scpm.imread("images/minoru_cube3_right.jpg")
```

```
/Users/andreshg/anaconda/lib/python3.6/site-packages/ipykernel_launcher.py:1: DeprecationWarning:
`imread` is deprecated in SciPy 1.0.0, and will be removed in 1.2.0.
Use ``imageio.imread`` instead.
```

```
"""Entry point for launching an IPython kernel.
```

```
/Users/andreshg/anaconda/lib/python3.6/site-packages/ipykernel_launcher.py:2: DeprecationWarning:
`imread` is deprecated in SciPy 1.0.0, and will be removed in 1.2.0.
Use ``imageio.imread`` instead.
```

```
In [7]: pt1, pt2 = misc.askpoints(img1, img2)
```

```
/Users/andreshg/anaconda/lib/python3.6/site-packages/matplotlib/backend_bases.py:2453: MatplotlibDeprecationWarning:
warnings.warn(str, mplDeprecation)
```

Ejercicio 1. Implementa la función `M = reconstruct(points1, points2, P1, P2)` que, dados una serie de N puntos 2D `points1` de la primera imagen y sus N homólogos `points2` de la segunda imagen (ambos en coordenadas homogéneas, $3 \times N$), y el par de matrices de proyección `P1` y `P2` de la primera y la segunda cámara respectivamente, calcule la reconstrucción tridimensional de cada punto. De ese modo, si `points1` y `points2` son $3 \times N$, la matriz resultante `M` debe ser $4 \times N$.

El tipo de reconstrucción debe ser algebraico, no geométrico.

```
In [8]: def reconstruct(points1, points2, P1, P2):
        """Reconstruct a set of points projected on two images."""

        # Transform homog to cartesian co-ordinates
        cp1 = points1[:2]
        cp2 = points2[:2]
```

```

# build coefficient matrix and compute reconstruction by least-squares.
# Useful functions are npla.lstsq() and npla.pinv()
cm = []
m = []
for i in range(len(cp1[0])):
    a0 = P1[0] - P1[2]*cp1[0,i]
    a1 = P1[1] - P1[2]*cp1[1,i]
    ha0 = P2[0] - P2[2]*cp2[0,i]
    ha1 = P2[1] - P2[2]*cp2[1,i]

    Am = np.array([a0,a1,ha0,ha1])

    # La última columna resta para poder calcular los mínimos cuadrados
    cm.append(npla.lstsq(Am[:, :3], -Am[:, 3])[0].tolist())
return cm

```

Reconstruye los puntos marcados y pinta su estructura 3D.

```

In [9]: print(pt1)
# reconstruct
mM=reconstruct(pt1, pt2, P1, P2)

mM_T = np.array(mM).T.tolist()
# convert from homog to cartesian

# plot 3D
misc.plot3D(mM_T[0],mM_T[1],mM_T[2])

[[105.47849462 207.67204301 304.99923195 288.23732719 202.80568356
 115.75192012]
 [ 25.10344086  58.08654378  26.72556068 150.54737327 204.07732719
 146.22172043]
 [ 1.          1.          1.          1.          1.
 1.          ]]

```

/Users/andreshg/anaconda/lib/python3.6/site-packages/ipykernel_launcher.py:21: FutureWarning: To use the future default and silence this warning we advise to pass `rcond=None`, to keep using

Out[9]: <mpl_toolkits.mplot3d.axes3d.Axes3D at 0x1164bde80>

Ejercicio 2. Reproyecta los resultados de la reconstrucción en las dos cámaras y dibuja las proyecciones sobre las imágenes originales. Comprueba que las proyecciones coinciden con los puntos seleccionados en el ejercicio anterior. Para dibujar los puntos puedes usar la función `plthom` de la práctica anterior o la versión que se distribuye con esta práctica (`misc.plthom`).

```

In [10]: # Proyecto los puntos en ambas cámaras
mM_T.append([1] * len(mM_T[0]))

```

```

proy1 = np.dot(P1,mM_T)
proy2 = np.dot(P2,mM_T)

# Pinto con misc.plothom()
ppl.figure()
misc.plothom(proy1,'r.')
ppl.imshow(img1)
ppl.show()

ppl.figure()
misc.plothom(proy2,'r.')
ppl.imshow(img2)
ppl.show()

```

1.6 2. Geometría epipolar

La geometría epipolar deriva de las relaciones que aparecen en las proyecciones de una escena sobre un par de cámaras. La matriz fundamental F , que depende exclusivamente de la configuración de las cámaras y no de la escena que éstas observan, es la representación algebraica de dicha geometría: a partir de ella se pueden calcular los epipolos y las líneas epipolares. La relación entre un par de cámaras P_1 , P_2 y la matriz fundamental es de $n-1$ (salvo factor de escala). Es decir, dadas dos cámaras calibradas, sólo tienen una matriz fundamental (excepto un factor de escala); dada una matriz fundamental existen infinitas configuraciones de cámaras posibles asociadas a ella.

1.6.1 2.1 Estimación de la matriz fundamental

Ejercicio 3. Implementa la función $F = \text{projmat2f}(P_1, P_2)$ que, dadas dos matrices de proyección, calcule la matriz fundamental asociada a las mismas. F debe ser tal que, si m_1 de la imagen 1 y m_2 de la imagen 2 están en correspondencia, entonces $m_2^T F m_1 = 0$.

```

In [11]: def projmat2f(P1,P2):
          """ Calcula la matriz fundamental a partir de dos de proyeccion"""
          A,B = P1[:, :3], P2[:, :3]
          a,b = P1[:, 3], P2[:, 3]
          iA,iB = npla.inv(P1[:, :3]), npla.inv(P2[:, :3])
          #Calculamos la matrix antisimétrica
          temp = misc.skew(np.dot(iB,b)-np.dot(iA,a))
          M = np.dot(iB.T,np.dot(temp,iA))
          return M

```

Se utiliza la siguiente fórmula para calcular la matriz fundamental

```

In [12]: # compute Fundamental matrix
          F = projmat2f(P1, P2)

```

Ejercicio 4 ¿Cómo es la matriz fundamental de dos cámaras que comparten el mismo centro? (Por ejemplo, dos cámaras que se diferencian sólo por una rotación.)

$$\bar{m}'^T \underbrace{\mathbf{B}^{-T} [\mathbf{B}^{-1} \bar{d} - \mathbf{A}^{-1} \bar{b}]_x \mathbf{A}^{-1}}_{\substack{\mathbf{F}_{3 \times 3} \\ \text{Matriz Fundamental}}} \bar{m} = 0$$

ecuacion_fundamental.png

1.6.2 Resultado

Dado el vector que une los dos centros es 0 (ya que comparten centro) (representado por la componente marrón en la fórmula de mas arriba), y la función de la matriz fundamental depende de este vector (se multiplican todas las componentes necesarias), la matriz fundamental será una matriz de 0.

1.6.3 2.2 Comprobación de F (OPCIONAL)

En los siguientes dos ejercicios vamos a comprobar que la matriz F estimada a partir de P1 y P2 es correcta.

Ejercicio 5. Comprueba que F es la matriz fundamental asociada a las cámaras P1 y P2. Para ello puedes utilizar el resultado 9.12, que aparece en la página 255 del libro Hartley, Zisserman. "Multiple View Geometry in Computer Vision." (second edition). Cambridge University Press, 2003.

```
In [13]: # P2.T * (F * P1) = A
         FxP1 = np.dot(F,P1)
         A = np.dot(P2.T, FxP1)
         # Esta matriz tiene que ser antisimétrica con respecto a su transpuesta
         print(A)
         print()
         print(A.T)

[[-1.96435472e-15 -7.03485803e-01 -6.00041642e+01  1.09871679e+04]
 [ 7.03485803e-01  1.04217722e-16 -3.08433134e+00  1.98740801e+02]
 [ 6.00041642e+01  3.08433134e+00  1.87111641e-14 -3.12199487e+04]
 [-1.09871679e+04 -1.98740801e+02  3.12199487e+04 -2.83341949e-10]]

[[-1.96435472e-15  7.03485803e-01  6.00041642e+01 -1.09871679e+04]
 [-7.03485803e-01  1.04217722e-16  3.08433134e+00 -1.98740801e+02]
 [-6.00041642e+01 -3.08433134e+00  1.87111641e-14  3.12199487e+04]
 [ 1.09871679e+04  1.98740801e+02 -3.12199487e+04 -2.83341949e-10]]
```

También se puede comprobar geoméricamente la bondad de una matriz F, si las epipolares con ella estimadas pasan por el homólogo de un punto dado en una de las imágenes.

Dada la matriz fundamental \mathbf{F} entre las cámaras 1 y 2, se puede determinar, para un determinado punto m_1 en la imagen de la cámara 1, cuál es la recta epipolar l_2 donde se encontrará su homólogo en la cámara 2:

$$l_2 = \mathbf{F}m_1.$$

Las siguientes dos funciones sirven para comprobar esta propiedad. En primer lugar, se necesita una función que dibuje rectas expresadas en coordenadas homogéneas, es decir, la versión de `plot` para rectas en lugar de puntos.

Ejercicio 6. Implementa la función `plotline(line)` que, dada una línea expresada en coordenadas homogéneas, la dibuje.

```
In [14]: def plotline(line, axes = None):
         """Plot a line given its homogeneous coordinates.

         Parameters
         -----
         line : array_like
             Homogeneous coordinates of the line.
         axes : AxesSubplot
             Axes where the line should be plotted. If not given,
             line will be plotted in the active axis.
         """
         if axes == None:
             axes = plt.gca()

         [x0, x1, y0, y1] = axes.axis()

         #      (x0, y0) .------. (x1, y0)
         #      |
         #      |
         #      |
         #      |
         #      |
         #      |
         #      |
         #      (x0, y1) .------. (x1, y1)

         # TODO: Compute the intersection of the line with the image
         # borders.

         # Con dos puntos podemos hacer una recta
         point1 = 0
         point2 = 0
         # Si y es mayor que x...
         if abs(line[0]) > abs(line[1]):
             # "cross product": un vector perpendicular a la linea homogenea y a [0, -1, y]
             point1 = np.cross(line, [0, -1, y0])
             point2 = np.cross(line, [0, -1, y1])
         # Si y es menor que x...
         else:
             # "cross product": un vector perpendicular a la linea homogenea y a [0, -1, x]
             point1 = np.cross(line, [-1, 0, x0])
             point2 = np.cross(line, [-1, 0, x1])
```

```

# Le quitamos Z y lo dejamos en coordenadas de la imagen
point1_Final = point1 / point1[2]
point2_Final = point2 / point2[2]

# TODO: Plot the line with axes.plot.
#axes.plot(...)
# pintamos una linea que cruce (x1,x2) y (y1,y2)
plotline = axes.plot([point1_Final[0], point2_Final[0]], [point1_Final[1], point2_Final[1]])
axes.axis([x0, x1, y0, y1])

"""
# Debug
print("line")
print(line)
print("punto 1 antes de quitarle Z")
print(point1)
print("punto 2 antes de quitarle Z")
print(point2)
print("punto 1 final")
print(point1_Final)
print("punto 2 final")
print(point2_Final)
"""

return plotline

```

Ejercicio 7. Completa la función `plot_epipolar_lines(image1, image2, F)` que, dadas dos imágenes y la matriz fundamental que las relaciona, pide al usuario puntos en la imagen 1 y dibuje sus correspondientes epipolares en la imagen 2 usando `plotline`.

```

In [15]: def plot_epipolar_lines(image1, image2, F):
        """Ask for points in one image and draw the epipolar lines for those points.

        Parameters
        -----
        image1 : array_like
            First image.
        image2 : array_like
            Second image.
        F : array_like
            3x3 fundamental matrix from image1 to image2.
        """
        # Prepare the two images.
        fig = plt.gcf()
        fig.clf()
        ax1 = fig.add_subplot(1, 2, 1)
        ax1.imshow(image1)
        ax1.axis('image')

```



```

ax2 = fig.add_subplot(1, 2, 2)
ax2.imshow(image2)
ax2.axis('image')
ppl.draw()

ax1.set_xlabel("Choose points in left image (or right click to end)")
point = ppl.ginput(1, timeout=-1, show_clicks=False, mouse_pop=2, mouse_stop=3)
while len(point) != 0:
    # point has the coordinates of the selected point in the first image.
    point = np.hstack([np.array(point[0]), 1])
    ax1.plot(point[0], point[1], '.r')

    # TODO: Determine the epipolar line.
    # Lo que hacemos es multiplicar nuestra matriz fundamental por nuestro punto
    # coordenadas homogéneas de la línea que tenemos que dibujar
    line = np.dot(F, point)
    #print(point)
    #print(line)

    # Plot the epipolar line with plotline (the parameter 'axes' should be ax2).
    plotline(line, axes=ax2)

    ppl.draw()
    # Ask for a new point.
    point = ppl.ginput(1, timeout=-1, show_clicks=False, mouse_pop=2, mouse_stop=3)

ax1.set_xlabel('')
ppl.draw()

```

Utiliza esta función con un par de imágenes llamándola de dos formas diferentes: seleccionando puntos en la imagen izquierda y dibujando las epipolares en la imagen derecha y viceversa. Comprueba en ambos casos que las epipolares siempre pasan por el punto de la segunda imagen correspondiente al seleccionado en la primera. Esto confirmará la corrección de la matriz F .

Añade dos figuras una que muestre la selección de puntos en la imagen izquierda y las rectas correspondientes en la imagen derecha, y otra que lo haga al revés. Indica para ambos casos qué matriz fundamental has usado al llamar a `plot_epipolar_lines`.

```

In [17]: # Si ahora cogemos la imagen uno y pintamos epipolares sobre la imagen dos usando la
plot_epipolar_lines(img1,img2,F)

In [19]: # ... entonces si cogemos la imagen dos y pintamos epipolares sobre la imagen uno usando
# la matriz transpuesta tendremos el mismo resultado.
plot_epipolar_lines(img2,img1,F.T)

```

1.7 3. Rectificación de imágenes

Es recomendable trabajar a partir de ahora con imágenes en blanco y negro y con valores reales entre 0 y 1 para cada uno de sus píxeles. Eso se puede conseguir con

```
In [20]: img1 = misc.rgb2gray(img1/255.0)
         img2 = misc.rgb2gray(img2/255.0)
```

La mayoría de algoritmos de puesta en correspondencia, incluyendo el que se va a implementar en esta práctica, requieren que las imágenes de entrada estén rectificadas.

Dos imágenes están rectificadas si sus correspondientes epipolares están alineadas horizontalmente. La rectificación de imágenes facilita enormemente los algoritmos de puesta en correspondencia, que pasan de ser problemas de búsqueda bidimensional a problemas de búsqueda unidimensional sobre filas de píxeles de las imágenes. En el material de la práctica se han incluido dos funciones que rectifican (mediante un método lineal) dos imágenes. La función `H1, H2 = misc.projmat2rectify(P1, P2, imsize)` devuelve, dadas las dos matrices de proyección y el tamaño de las imágenes en formato (filas,columnas), las homografías que rectifican, respectivamente, la imagen 1 y la imagen 2. La función `projmat2rectify` hace uso de `projmat2f`, por lo que es necesario que esta función esté disponible.

Ejercicio 8. Se tienen dos imágenes no rectificadas `im1` e `im2`, y su matriz fundamental asociada `F`. Con el procedimiento explicado, se encuentran un par de homografías H_1 y H_2 que dan lugar a las imágenes rectificadas `O1` y `O2`. ¿Cuál es la matriz fundamental `F` asociada a estas dos imágenes? ¿Por qué?

Nota: `F` depende exclusivamente de `F`, `H1` y `H2`.

1.7.1 Resultado

La matriz fundamental se calcula multiplicando el vector que une el primer centro con el punto real, el vector que une el segundo centro con el punto real, y el vector que une los dos centros. Dado que dichos vectores no cambian al rectificar las imágenes, el la matriz `F` resultante será igual a `F`. Esto se ve reflejada en la fórmula del ejercicio 3.

Ejercicio 9. Rectifica el par de imágenes estéreo `img1` e `img2` y calcula la matriz fundamental asociada a estas imágenes.

```
In [21]: H1, H2 = misc.projmat2rectify(P1, P2, projmat2f, img1.shape)
         O1, O2 = misc.rectify_images(img1, img2, H1, H2)
```

```
In [22]: ppl.figure()
         ppl.imshow(O1)
         ppl.figure()
         ppl.imshow(O2)
```

```
Out[22]: <matplotlib.image.AxesImage at 0x1173636a0>
```

Ejercicio 10. (opcional) Calcula y muestra la matriz fundamental de las imágenes rectificadas. Justifica el resultado obtenido (mira la sección 9.3.1 del libro de Hartley y Zisserman, pág. 248 y 249).

```
In [23]: # La inversa de H2 transpuesta por la matriz fundamental original
         invH2 = np.dot(npla.inv(H2).T, F)
         # La inversa de H1
         invH1 = npla.inv(H1)
         #Matriz fundamental rectificada
         Fr = np.dot(invH2, invH1)
         Fr /= Fr[2,1]
         print(Fr)
```

```
[[ 2.23248531e-22  7.57802232e-17 -2.02988756e-14]
 [-6.15143783e-17 -1.93148584e-17 -1.00000000e+00]
 [ 1.98620144e-14  1.00000000e+00 -1.86082381e-12]]
```

Ejercicio 11. (Opcional) Usa `plot_epipolar_lines` para dibujar varias líneas epiplares de las imágenes rectificadas. Muestra los resultados.

In [24]: `"""`

Vamos a editar el código de la función `plotline` para que imprima los dos puntos que hacen una línea, esto es importante para la justificación.

Es el mismo código que antes sin comentarios pero con dos prints al final.
`"""`

```
def plotline(line, axes = None):
    if axes == None:
        axes = plt.gca()
    [x0, x1, y0, y1] = axes.axis()
    point1 = 0
    point2 = 0
    if abs(line[0]) > abs(line[1]):
        point1 = np.cross(line, [0, -1, y0])
        point2 = np.cross(line, [0, -1, y1])
    else:
        point1 = np.cross(line, [-1, 0, x0])
        point2 = np.cross(line, [-1, 0, x1])
    point1_Final = point1 / point1[2]
    point2_Final = point2 / point2[2]

    plotline = axes.plot([point1_Final[0], point2_Final[0]], [point1_Final[1], point2_Final[1]])
    axes.axis([x0, x1, y0, y1])

    print("punto 1:")
    print(point1_Final)
    print("punto 2:")
    print(point2_Final)
    print("")

    return plotline
```

In [25]: `# Cada vez que pulsemos en un punto, aquí se imprimirán las coordenadas de la línea`
`plot_epipolar_lines(O1, O2, Fr)`

```
punto 1:
[ -0.5  164.15   1. ]
punto 2:
[494.5  164.15   1. ]
```

```

punto 1:
[ -0.5      302.47056452    1.      ]
punto 2:
[494.5      302.47056452    1.      ]

punto 1:
[ -0.5      440.79112903    1.      ]
punto 2:
[494.5      440.79112903    1.      ]

punto 1:
[ -0.5      324.42620968    1.      ]
punto 2:
[494.5      324.42620968    1.      ]

punto 1:
[ -0.5      251.97258065    1.      ]
punto 2:
[494.5      251.97258065    1.      ]

```

1.7.2 Justificación ejercicios 10 y 11

Si ejecutamos los fragmentos de código de los ejercicios 10 y 11 tenemos los siguientes resultados:

Matriz fundamental rectificada:

```

[[ 1.03162575e-22 -1.94177057e-16 -2.03242309e-14] [ 1.83672211e-16 1.24468708e-17 -
1.00000000e+00] [ 2.07441490e-14 1.00000000e+00 -2.02384848e-12]]

```

Puntos: punto 1: [-0.5 469.33346774 1.]

punto 2: [494.5 469.33346774 1.]

o, en otro punto: punto 1: [-0.5 337.59959677 1.]

punto 2: [494.5 337.59959677 1.]

Como se puede observar la matriz fundamental de las imágenes rectificadas tiene una configuración (sobretudo los característicos "unos" en posiciones claves y valores muy próximos a cero) que, al calcular las líneas epipolares, hace que estas líneas sean paralelas unas de otras.

Todas las líneas epipolares tienen la misma pendiente (en este caso 0, todas las líneas son constantes en el eje X) y son paralelas al eje X y perpendiculares al eje Y.

1.8 4. Búsqueda de correspondencias

La búsqueda de correspondencias consigue establecer automáticamente las correspondencias de puntos entre dos imágenes (lo que se ha hecho manualmente en el ejercicio 2) haciendo uso de las restricciones que proporciona la geometría epipolar.

1.8.1 4.1 Cálculo de las medidas de similitud

Una vez rectificadas las dos imágenes de un par estéreo, se pueden buscar las correspondencias. Una matriz de disparidades S indica, para cada píxel de la imagen 1 rectificada, a cuántos píxeles de diferencia está su correspondencia en la imagen 2 rectificada. En nuestra práctica, para simplificar el problema, vamos a considerar que los elementos de S son enteros. Para el píxel en la posición (x, y) en la imagen 1, su correspondiente está en $(x + S[y, x], y)$ en la imagen 2. Si $S[y, x] < 0$, la correspondencia está hacia la izquierda; si $S[y, x] > 0$, la correspondencia está hacia la derecha; si $S[y, x] = 0$, las coordenadas de los dos puntos coinciden en ambas imágenes.

La búsqueda de correspondencias requiere ser capaz de determinar el parecido visual entre píxeles de dos imágenes. Si los píxeles m_1 y m_2 son visualmente parecidos, tienen más probabilidad de estar en correspondencia que otros que sean visualmente diferentes. Como la apariencia (el nivel de gris) de un único píxel es propensa al ruido y poco discriminativa, el elemento de puesta en correspondencia será una ventana centrada en el píxel. Dado un píxel m de una imagen, llamaremos vecindad del píxel de radio K al conjunto de píxeles de la imagen que se encuentren dentro de una ventana de tamaño $(2K + 1)(2K + 1)$ píxeles centrada en m . El número de píxeles de una vecindad de radio K es $N = (2K + 1)^2$. Dadas dos vecindades w_1 y w_2 de dos píxeles, el parecido visual entre ellas puede calcularse con la suma de *diferencias al cuadrado* (SSD) de cada una de sus componentes

$$d_{SSD}(\mathbf{v}, \mathbf{w}) = \sum_{i=1}^N (\mathbf{v}_i - \mathbf{w}_i)^2.$$

La distancia d_{SSD} es siempre positiva, es pequeña cuando dos ventanas son visualmente parecidas y grande en caso contrario.

Ejercicio 12. Implementa la función `C = localssd(im1, im2, K)` que calcua la suma de diferencias al cuadrado entre las ventanas de radio K de la imagen 1 y la imagen 2. El resultado debe ser una matriz del mismo tamaño que las imágenes de entrada que contenga en cada punto el valor de d_{SSD} para la ventana de la imagen 1 y la ventana de la imagen 2 centradas en él. Es decir, $C[i, j]$ debe ser el resultado de d_{SSD} para las ventanas centradas en $im1[i, j]$ e $im2[i, j]$.

Para este ejercicio puede resultar útil la función `scipy.ndimage.convolve`.

```
In [26]: def localssd(im1, im2, K):
         """
         The local sum of squared differences between windows of two images.

         The size of each window is (2K+1)x(2K+1).
         """
         dmatrix = (im1-im2)**2
         masc = np.ones([2*K+1, 2*K+1])
         cmatrix = scnd.convolve(dmatrix, masc, mode="constant", cval = 0)
         return cmatrix
```

Ejercicio 13. Implementa la función `D = ssd_volume(im1, im2, disps, K)` que calcula la suma de diferencias al cuadrado entre las ventanas de la imagen `im1` y la imagen `im2` desplazada horizontalmente. El parámetro `disps` debe ser una lista de valores indicando las disparidades que se usarán para desplazar la imagen `im2`. Por ejemplo, si `disps` es `np.arange(-3, 2)`, se llamará 5 veces a `localssd` para la imagen 1 y la imagen 2 desplazada 3, 2, 1, 0 y 1 píxeles en sentido horizontal. K es el parámetro que indica el radio de las ventanas usado por `localssd`.

El valor devuelto D será un array de tamaño MNL , donde L es el número de disparidades indicadas por `disps`, $L = \text{len}(\text{disps})$ (es decir, el número de veces que se ha llamado a `localssd`); M y N son, respectivamente, el número de filas y de columnas de las imágenes de entrada. El elemento $D[y, x, l]$ debe ser la SSD entre la ventana centrada en $\text{im1}[y, x]$ y la ventana centrada en $\text{im2}[y, x + \text{disps}[l]]$.

$D[y, x, l]$ debe ser muy grande para aquellos valores en los que $\text{im2}[y, x + \text{disps}[l]]$ no esté definido, es decir, el índice $(y, x + \text{disps}[l])$ se sale de la imagen 2.

```
In [27]: def ssd_volume(im1, im2, disps, K):
        """
        Calcula el volumen de disparidades SSD
        """
        L = len(disps)
        # matriz_nml = np.zeros(im1.shape, L)
        ssd_vol = np.zeros(im1.shape + (L,))
        for i, val in enumerate(disps):
            #ponemos infinitos para evitar problemas al calcular la disparidad
            array_inf = np.ones((im1.shape[0], abs(val))) * np.inf

            if val > 0:
                imx = im2[:, val:]
                imx = np.hstack((imx, array_inf))
            else:
                imx = im2[:, :val]
                imx = np.hstack((array_inf, imx))
            ssd_vol[:, :, i] = localssd(im1, imx, K)
        return ssd_vol
```

Ejercicio 14. El conjunto de disparidades `disps` debe ser lo más pequeño posible, para mejorar el rendimiento de la optimización. Determina un procedimiento para estimar manualmente el conjunto de disparidades posibles y aplícalo a las imágenes O1 y O2.

```
In [28]: disps = np.arange(110, 210)
```

Aplica la función `ssd_volume` al par de imágenes O1 y O2 con las disparidades estimadas en el ejercicio anterior.

```
In [29]: D = ssd_volume(O2, O1, disps, 5)

        # to speed-up the optimization ahead, discard the par of the image showing only backg
        D = D[:, :400, :]

        D.shape
```

```
Out [29]: (656, 400, 100)
```

1.8.2 4.1 Estimación de la disparidad sin regularizar

La matriz D calculada en el ejercicio anterior proporciona los costes unitarios D_i de una función de energía sin regularización de la forma

$$E(x) = \sum_i D_i(x_i),$$

donde $D_i(l)$ viene dado por $D[l, y, x]$, suponiendo que el píxel i tenga coordenadas (x, y) . Las variables $x = (x_1, \dots, x_{NM})$ indican las etiquetas de cada uno de los píxeles. En este caso, las etiquetas son los índices del array `disps`, que a su vez son las disparidades horizontales. Por eso, a partir de aquí se hablará indistintamente de etiquetas y disparidades. Sólo es necesario recordar que la etiqueta l está asociada a la disparidad `disps[l]`.

Minimizando la energía $x = \arg \min_x E(x)$, se obtiene un vector de etiquetas óptimo x^* que indica, para cada píxel, cuál es su disparidad horizontal entre las dos imágenes.

Ejercicio 15. Minimiza $E(x)$ y muestra las disparidades resultantes.

```
In [30]: res = D.argmin(2) # Minimizar el volumen en 2
         ppl.imshow(res, cmap='gray')
         ppl.colorbar()
```

```
Out[30]: <matplotlib.colorbar.Colorbar at 0x117231eb8>
```

```
In [31]: print(res)
```

```
[[75  3  3 ...  0  0  0]
 [ 3  3  3 ...  0  0  0]
 [ 3  3  3 ...  0  0  0]
 ...
 [99 99 99 ...  0  0  0]
 [99 99 99 ...  0  0  0]
 [99 99 99 ...  0  0  0]]
```

1.8.3 4.2 Estimación de la disparidad regularizada

El etiquetado usando exclusivamente términos unitarios es muy sensible al ruido y propenso a que aparezcan zonas de píxeles cercanos con mucha variación en las etiquetas. Esto es especialmente notable en zonas planas (es decir, sin textura) de las imágenes originales, donde no hay suficiente información para establecer una correspondencia basándose exclusivamente en la apariencia visual de ventanas pequeñas. Por eso es necesario incluir un término de suavizado o regularización en la función de energía. Los tipos de saltos de etiquetas que aparecerán en el resultado final dependerán de cómo sea ese término de suavizado.

La función de energía que utilizaremos para calcular las disparidades en la práctica será el resultado de añadir a la expresión (6) un término que penalice los cambios de disparidad en los píxeles vecinos:

$$E_r(x) = \sum_i D_i(x_i) + \sum_{ij} \lambda |x_i - x_j|.$$

La solución al problema de la correspondencia vendrá dado por el conjunto de etiquetas (disparidades) de los píxeles de la imagen que minimicen $E_r(x)$.

En [Yuri Boykov, Olga Veksler, and Ramin Zabih. "Fast approximate energy minimization via graph cuts". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23:1222–1239, 2001.] se presentan métodos para resolver algunos problemas de optimización con varias etiquetas empleando algoritmos de cortes de grafos. Es recomendable repasar las secciones 5 y 8.

Ejercicio 16. Escribe la función `find_corresp_aexpansion(D, initLabels, lmb, maxV)`, que recie un volumen `ssd`, `D`, un conjunto inicial de etiquetas, `initLabels`, que puede ser el obtenido en el ejercicio 15, el valor de la constante `lmb` y el valor máximo de la función de coste $|x_i x_j|$, que tendrás que establecer empíricamente. El resultado de esta función serán las etiquetas que minimizan $E_r(x)$. Para ello debes utilizar la función `maxflow.fastmin.aexpansion_grid` del paquete *PyMaxFlow*, que resuelve el problema anterior mediante un algoritmo de cortes de grafos empleando una `a-expansión`.

```
In [32]: def find_corresp_aexpansion(D, initialLabels, lmb, maxV):
          # El segundo parámetro es es la matriz de los costes por pares
          V = np.asarray([[min(lmb*abs(w - c), maxV) for c in range(D.shape[2])] for w in range(D.shape[2])])
          return maxflow.fastmin.aexpansion_grid(D, V, max_cycles=5, labels=initialLabels)
```

Llama a esta función y muestra una figura con las etiquetas que resulten de la minimización de la energía para el volumen `ssd` `D` (este proceso puede durar varios minutos).

```
In [33]: labels = find_corresp_aexpansion(D, res, 5, 70)
```

```
In [34]: ppl.imshow(labels)
          print(labels)
```

```
[[30 30 30 ... 0 0 0]
 [30 30 30 ... 0 0 0]
 [30 30 30 ... 0 0 0]
 ...
 [99 99 99 ... 0 0 0]
 [99 99 99 ... 0 0 0]
 [99 99 99 ... 0 0 0]]
```

La matriz de etiquetas óptimas `X` obtenida de la minimización de la función de energía puede transformarse en la matriz de disparidades `S` indexando en cada una de sus celdas el array de disparidades `disps` `S = disps[X]`. Ahora, el píxel de coordenadas `(x, y)` de la primera imagen rectificada tendrá su correspondencia en el píxel de coordenadas `(x + S[y, x], y)` de la segunda imagen rectificada.

El siguiente ejercicio usa la matriz de disparidades para establecer automáticamente las correspondencias entre un par de imágenes sin rectificar.

Ejercicio 17. Implementa la función `plot_correspondences(image1, image2, S, H1,H2)` que, dado un par de imágenes sin rectificar, la matriz de disparidades entre las imágenes rectificadas y las homografías que llevan de las imágenes sin rectificar a las imágenes rectificadas, pida al usuario puntos en la primera imagen y dibuje sus correspondencias en la segunda.

```
In [35]: def plot_correspondences(image1, image2, S, H1, H2):
          """
```


Ask for points in the first image and plot their correspondences in the second image.

Parameters

```
image1, image2 : array_like
    The images (before rectification)
S : array_like
    The matrix of disparities.
H1, H2 : array_like
    The homographies which rectify both images.
"""
# Prepare the two images.
fig = plt.gcf()
fig.clf()
ax1 = fig.add_subplot(1, 2, 1)
ax1.imshow(image2)
ax1.axis('image')
ax2 = fig.add_subplot(1, 2, 2)
ax2.imshow(image1)
ax2.axis('image')
plt.draw()

ax1.set_xlabel("Choose points in left image (or right click to end)")
point = plt.ginput(1, timeout=-1, show_clicks=False, mouse_pop=2, mouse_stop=3)
while len(point) != 0:
    # point has the coordinates of the selected point in the first image.
    point = np.c_[np.array(point), 1].T
    ax1.plot(point[0], point[1], '.r')

    # Calculamos el punto rectificado
    pointh = np.dot(H1, point)
    pointh /= pointh[2] # Pasamos a cartesianas
    # TODO: Determine the correspondence of 'point' in the second image.
    # perhaps you have to swap the image co-ordinates.
    disp_pointh = S[int(pointh[1]), int(pointh[0])]
    pointh[0] += disp_pointh

    # Se pone la inversa porque estamos llevando de la imagen rectificada a la im
    point_der = np.dot(npla.inv(H2), pointh)
    point_der /= point_der[2]
    # TODO: Plot the correspondence with ax2.plot.

    ax2.plot(point_der[0], point_der[1], '.r.')

    plt.draw()
    # Ask for a new point.
```

```

        try:
            point = ppl.ginput(1, timeout=-1, show_clicks=False, mouse_pop=2, mouse_s
        except Exception as e:
            break

    ax1.set_xlabel('')
    ppl.draw()

In [36]: # Construimos S como la matriz de disparidad real entre dos píxeles, transformamos l
# en la disparidad real (de 110 a 210)
S = np.asarray([[disps[labels[i][j]] for j in range(D.shape[1])] for i in range(D.shap
plot_correspondences(img1, img2, S, H1, H2)

In [ ]:

```