

# Práctica 2 - Extracción, descripción y similitud de características locales

Fernández Freire, Luis – Juez Hernández, Rodrigo

## 2 PREGUNTAS TAREAS OPCIONALES

2.1 Aplicando los descriptores de tipo 'hist' de la tarea 2 sobre la imagen camera() de Skimage (paquete skimage.data, <https://bit.ly/2ZsXfN7>), analice como cambia el descriptor si utiliza:

- Tamaño de vecindario con valores 8 y 16
- Número de bins con valores 16 y 32

(0.75 puntos)

[417 224]

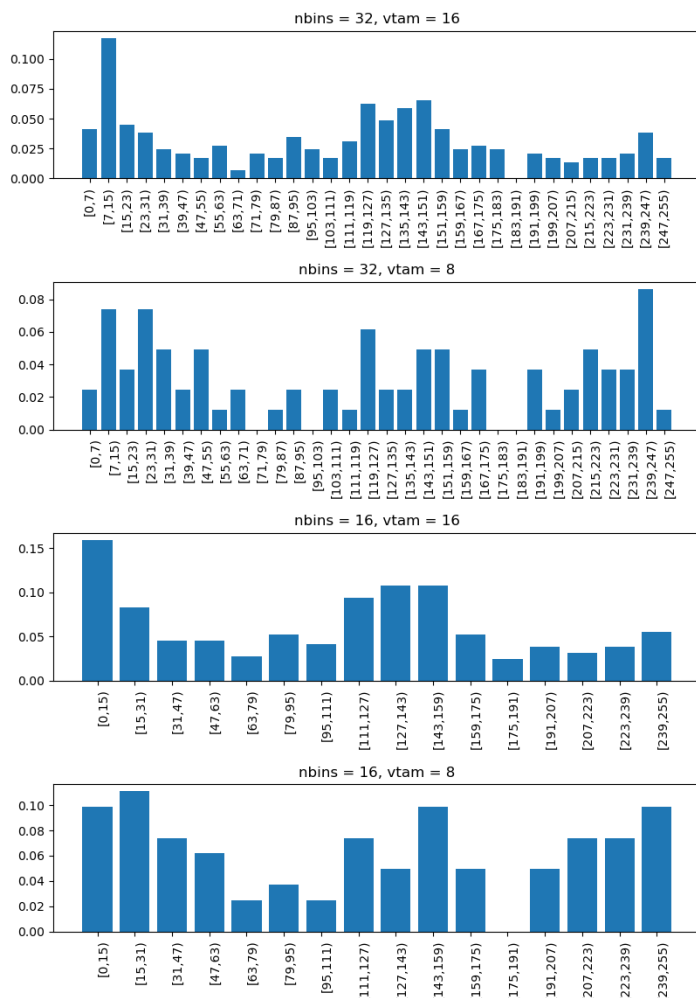


Figura 1

[225 194]

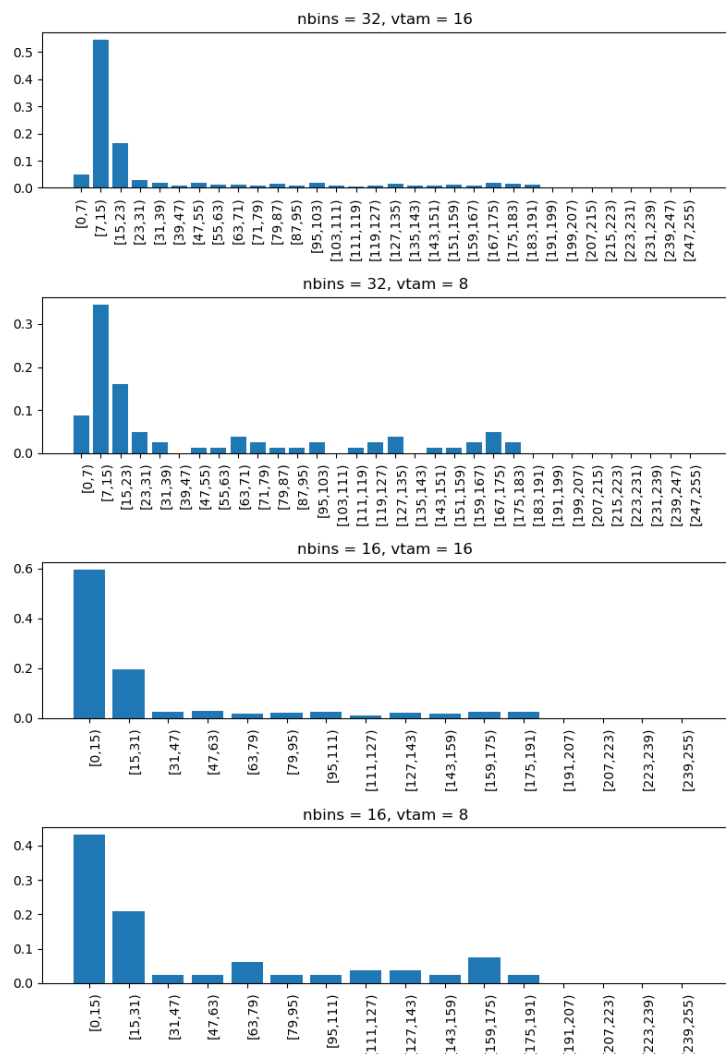


Figura 2

[133 288]

[140 285]

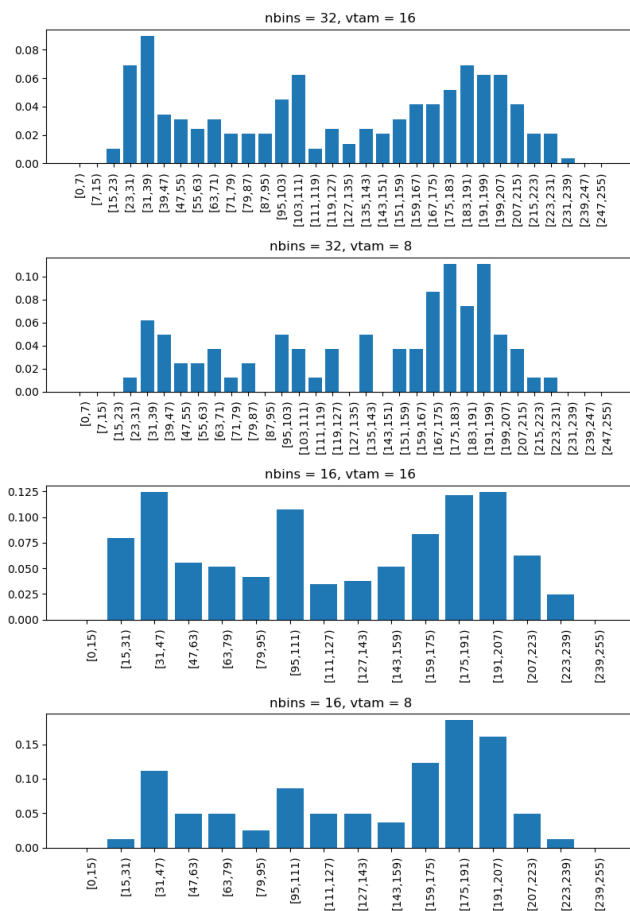


Figura 3

Se no pide evaluar los histogramas obtenidos de una imagen para un punto de interés específico.

En este apartado se nos pide utilizar un número de bins de 16 y 32, lo que significa que los histogramas obtenidos, el rango de valores con 32 bins, tendrá el doble de rangos que el de 16 bins. Si se utiliza bins grandes, los datos obtenidos respecto al punto de interés serán más específicos. Mientras que con bins más pequeños serán más genéricos.

También se nos pide utilizar un tamaño de vecindario de 8 y 16, lo que significa que se tomarán más o menos datos alrededor del punto de interés a la hora de representar el histograma. En este caso, el segundo tomará una mayor cantidad de datos respecto al segundo:

$$vtam=8 \ (8+1)2=81 \text{ datos} ; vtam=16 \ (16+1)2=289 \text{ datos}$$

Para este ejercicio desarrollamos una función que hace multiplots en vertical que muestran todas las posibles combinaciones pedidas para un punto de interés en específico.

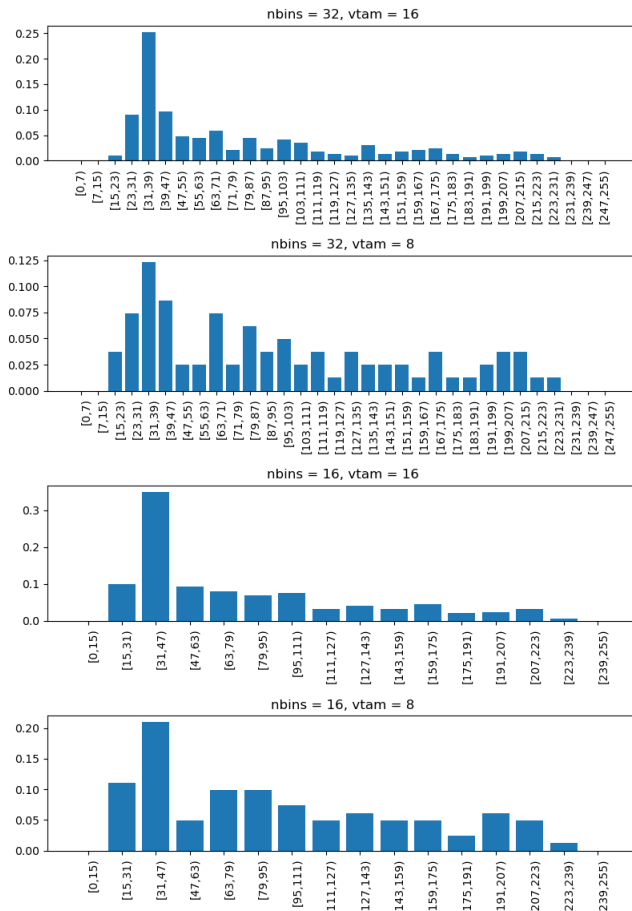


Figura 4

En primer lugar, debemos comentar cómo varía dependiendo del número de vecindarios.

Si nos fijamos en la tendencia general no hay mucha diferencia entre ambos.

Donde más claro se ve es en la Figura 4, las diferencias que hay son debido a que el histograma con 32 bins desgana más el rango, por lo que en el de 16 bins simplemente se juntan más valores por cada rango. Se puede observar claramente en los valores de la izquierda que cada rango de 16 es la suma de dos rangos del de 32 exactamente. Por ejemplo, en la Figura 2 para el vtam 16, si comparamos los bins 32 y 16 el rango [0,15) (0.59) es claramente la suma de los rangos [0,7).

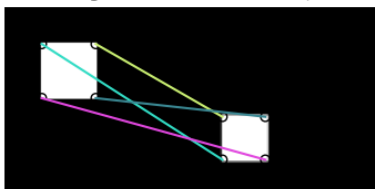
[7,15). Estas diferencias pueden significar que cuando se aplican algoritmos de correspondencia el algoritmo que usa menos bins pueda generalizar más, ya que si tiene

ruido la imagen y hay pequeñas variaciones en un rango de valores muy pequeño los elimina, o también puede ocurrir que tenga más fallo de error, ya que puede que esa información que elimina no sea ruido y sea relevante en la imagen, por ejemplo cambios de tonalidad en la manera que son muy sutiles y con un número reducido de bins no se notarán.

**2.2 Analice las correspondencias obtenidas en la tarea 3 con los descriptores ‘hist’ y ‘mag-ori’ sobre la primera imagen de test<sup>1</sup>. Visualice y razone porque los cambios que observe en los experimentos que realice (0.75 puntos).**

### hist Correspondencia

Imagen test\_basica: original vs. transformada (resultado estudiante)



### mag-ori Correspondencia

Imagen test\_basica: original vs. transformada (resultado estudiante)

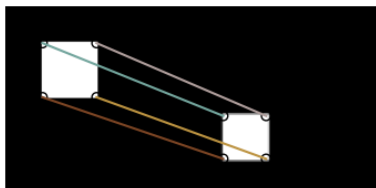
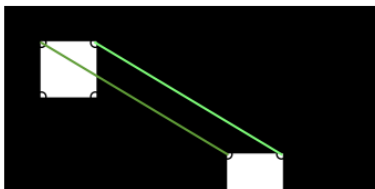


Figura 5

### hist Correspondencia

Imagen test\_basica: original vs. transformada (resultado estudiante)



### mag-ori Correspondencia

Imagen test\_basica: original vs. transformada (resultado estudiante)

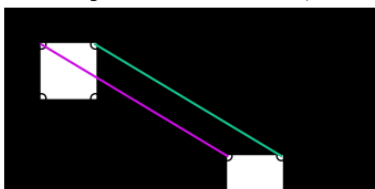
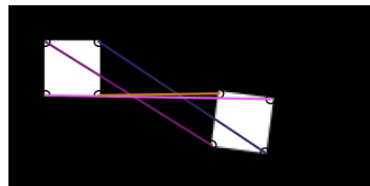


Figura 6

### hist Correspondencia

Imagen test\_basica: original vs. transformada (resultado estudiante)



### mag-ori Correspondencia

Imagen test\_basica: original vs. transformada (resultado estudiante)

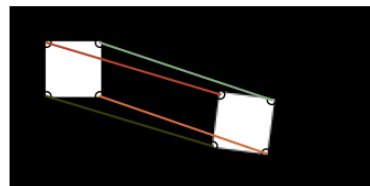


Figura 7

Tal y como se puede observar en el resultado de la correspondencia con el descriptor “hist”, las esquinas que ha vinculado con respecto a la imagen original son todas distintas. Las esquinas vinculadas tienen un error de 90°. Es decir, si cambiáramos la asignación por la esquina siguiente en sentido de las agujas del reloj, la vinculación sería la correcta.

Ahora bien, el resultado obtenido con la correspondencia del descriptor “mag-ori”, identifica correctamente las esquinas correspondientes de la imagen original con respecto a la modificada.

Pasemos a la Figura 6.

En esta ocasión, al eliminar la mitad inferior de la imagen, el descriptor de correspondencia “hist” detecta correctamente las esquinas correspondientes de la imagen original respecto a la modificada. Y como se puede apreciar, la correspondencia con el descriptor “mag-ori” no falla tampoco en su detección.

Finalmente, hablemos de la Figura 7.

Con mirar mínimamente las imágenes de ambos descriptores, la de la correspondencia “hist” falla en su vinculación de puntos de interés con respecto a la de su imagen original. De hecho, los valores vinculados nos indica que los puntos de interés superiores de la imagen original los vincula con los puntos de interés inferiores de la imagen transformada, y viceversa.

Por el otro lado, la correspondencia del descriptor “mag-ori” acierta todos los puntos de interés de la imagen original.

Por tanto, la razón por la que falla la correspondencia “hist” y no falla la correspondencia “mag-ori” se debe a que la segunda obtiene las orientaciones y las magnitudes de la imagen original gracias a la derivada respecto al eje X y respecto al eje Y. Esas orientaciones y magnitudes son resultado de cambios en la imagen y guardan información de “a donde mira el borde” es decir la normal, por ejemplo en el cuadrado movido de lugar (**Figura 5**) el histograma no diferencia bien las esquinas porque los colores son iguales, es decir aunque en la matriz no estén posicionados en los mismos sitios los colores en el vecindario siempre van a sumar la misma cantidad, sin embargo el mag-ori calcula la normal al borde y esta orientación si que varía en las esquinas porque en una mira a la izquierda, en otra a la derecha, etc.

Así es como diferencia las esquinas bien, esa es la razón para que funcione en el cuadrado rotado. Hay que tener cuidado porque este algoritmo no aplica ningún algoritmo de normalización de orientaciones como si hace SIFT, aunque en el caso del cuadrado funciona bien.

**2.3 Aplique las funciones de la parte 1, 2 y 3 sobre los pares de imágenes proporcionados en la carpeta ‘img’. Discuta las diferencias en las esquinas devueltas (si las hay) para cada par de imágenes y qué fenómenos del mundo real o efectos de la cámara pueden haber causado estas diferencias. Proporcione al menos un problema para cada par de imágenes. Incluya ejemplos visuales en la memoria de los experimentos que realice. (1 punto)**

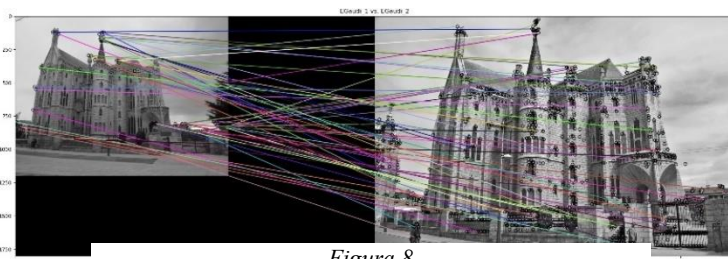


Figura 8



Figura 9

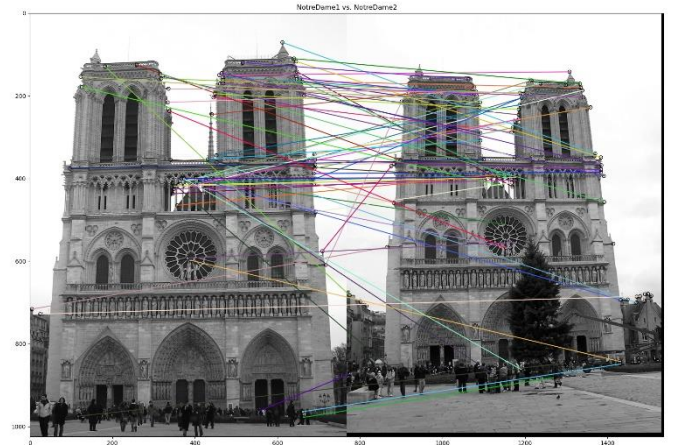


Figura 10



Figura 11

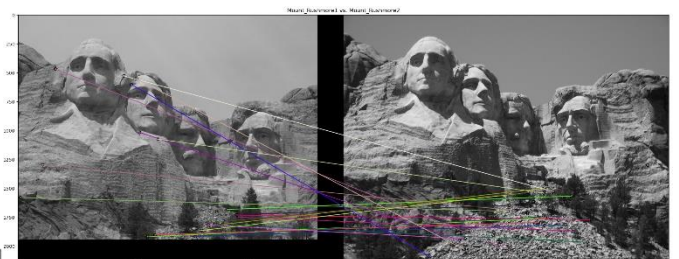


Figura 12

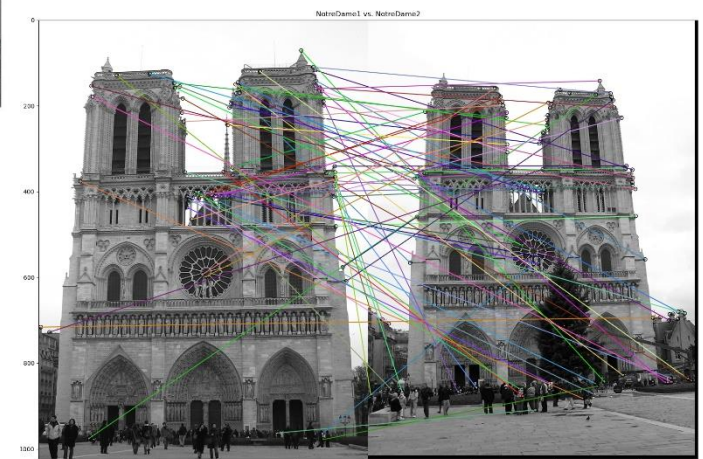


Figura 13



### Diferencias EGaudi:

- **Hist (Figura 8):** Se detectan una gran multitud de esquinas entre ambas esquinas. De los puntos de interés que identifica, muchos de ellos no los vincula correctamente. Esto se puede deber al cambio de iluminación, posición desde donde se realiza la foto y nuevas figuras que aparecen en la imagen respecto a la imagen 1. Por ejemplo, podemos ver como identifica un punto de interés a un transeúnte en la segunda imagen. Y como los puntos de interés de las ventanas se vinculan a los del cerco de la segunda imagen. Esto es porque ambos son bordes de claro oscuro, y los histogramas pueden salir similares, pero los colores están posicionados de manera distinta en el vecindario.
- **mag-ori (Figura 11):** Al igual que con “hist”, esta correspondencia vincula la mayoría de los puntos de interés de forma errónea. En este caso, la razón puede deberse a la perspectiva con la que se realiza la foto, aunque mag-ori encuentra la orientación la manera de describir es haciendo un histograma con la magnitud de los vectores de las orientaciones, por lo que un borde con la misma forma y en la misma orientación puede coincidir, aunque la zona no sea la correcta. Si nos fijamos los errores que tiene ahora son con bordes con la misma forma, por ejemplo las ventanas y el porche ambos tienen un borde que es un semicírculo y los confunde.

### Diferencias Monte Rushmore:

- **hist (Figura 9):** En esta ocasión, se detecta una menor cantidad de puntos de interés respecto a la “EGaudi”. Detecta una mayor cantidad de vínculos en el pequeño montículo con grava. Esto puede deberse a la cantidad de esquinas que hay en dicha zona por las grietas que se generan debido a la erosión de la tierra. También influye la distancia de dicho montículo respecto al monte Rushmore, como una foto está más lejos tiene menos definición las piedras y hay menos bordes y variaciones.
- **mag-ori (Figura 12):** Si observamos el resultado, la aplicación de la correspondencia por orientación tampoco se libra del montículo de piedras, lo que genera esa zona una captación de vínculos de puntos de interés. Sin embargo, como los bordes de las piedras tienen orientaciones aleatorias no se libran de los errores.

### Diferencias NotreDame:

- **hist (Figura 10):** Como era de esperar, se detectan una gran cantidad de esquinas, pero los vínculos

realizados entre las 2 imágenes de sus respectivos puntos de interés son erróneos.

- **mag-ori (Figura 13):** Al igual que antes, la identificación de puntos de interés y vinculación sigue siendo un desastre. En este caso, el cálculo de la orientación se hace respecto a las derivadas de la imagen original. En este caso las orientaciones funcionan mucho peor con mag-ori que con hist, esto es debido a que se ha cambiado de orientación la posición de la cámara, por lo que las orientaciones de las normales han variado y ya no coinciden.

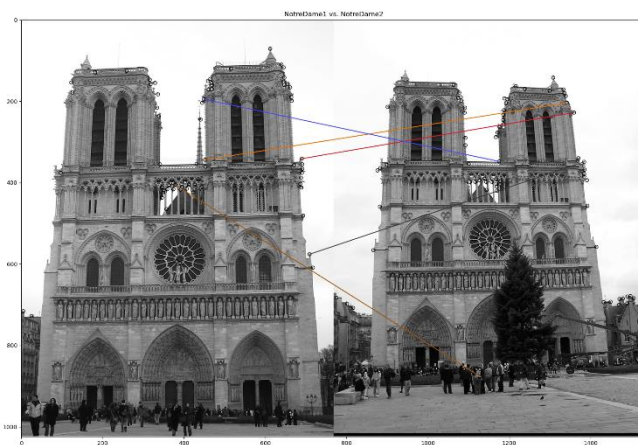


Figura 14

### hist con distancia NDDR (Figura 14):

Finalmente, en esta última imagen de Notre Dame, hemos querido probar la correspondencia “hist” con el cálculo de la distancia por NDDR. Con echar un vistazo se puede ver que el resultado es un desastre. Realiza 5 vinculaciones entre las 2 imágenes y ni siquiera están bien los vínculos. Esto demuestra que utilizar la distancia NDDR para el cálculo de distancias mínimas es inútil.

### Explicación extra:

Tras la implementación del apartado 2b, no hemos obtenido un 100% de aciertos, y hemos obtenido desde nuestros ordenadores personales un 93,27% de aciertos en uno y un 96,15% de aciertos en otro.

Hemos determinado que esto se puede deber a la capacidad de cómputo del procesador de cada ordenador. Ya que otros estudiantes han reportado casos similares en el foro.

### Evidencias:

```
* Finalizado en 6.515 secs
* RESULTADO FINAL: 194/208 DESCRIPTORES CORRECTOS ( 93.27% )
Tests completados = True
```

```
Testeando descriptores tipo MAG-ORI, vecindario 17x17 y 32 bins con imagen #3 Cofee 400x600...
Esquinas...#0 #1 #2 #3 #4 #5 #6 #7 #8 #9 correctos 10/10 descriptores
Testeando descriptores tipo MAG-ORI, vecindario 33x33 y 32 bins con imagen #3 Cofee 400x600...
Esquinas...#0 #1 #2 #3 #4 #5 #6 #7 #8 #9 correctos 8/10 descriptores

* Finalizado en 6.103 secs
* RESULTADO FINAL: 200/208 DESCRIPTORES CORRECTOS ( 96.15% )
Tests completados = True
```

## 4 CARGA DE TRABAJO

Indique brevemente la carga (en horas) de cada tarea de esta práctica. Puede utilizar como ejemplo la Tabla 1.

Tarea	Horas dedicadas (nombre apellido)	Horas dedicadas (nombre apellido)
Tarea 1	0.5	1
Tarea 2a	0	1
Tarea 2b	0.5	1
Tarea 3	0	1
Memo- ria	6	6

## REFERENCIAS

- [1] Juan Carlos San Miguel Avedillo, Tema 2: Extracción de características Puntos de Interés: Harris
- [2] Juan Carlos San Miguel Avedillo, Tema 2: Extracción de características Puntos de Interés: descripción
- [3] Juan Carlos San Miguel Avedillo, Tema 2: Extracción de características Puntos de Interés: similitud