

Práctica 3. Construcción de un Mosaico

1. Objetivo

El objetivo de la práctica es que el alumno sea capaz de realizar un mosaico o vista panorámica a partir de un conjunto de imágenes obtenidas sin desplazamiento aparente. Implica que el alumno sea capaz de detectar características visuales, filtrar dichas características evitando falsas detecciones, estimar una homografía, y componer una imagen global a partir de unas imágenes parciales.

2. Escenarios

Se trabajará con uno de los tres escenarios mostrados en las figuras 1, 2 y 3 (opcionalmente con varios de ellos). El primer escenario numerado como Escenario3 está formado por dos imágenes, el segundo escenario numerado como Escenario4 está formado por 3 imágenes y el tercer escenario numerado como Escenario8 está formado por 7 imágenes. Las imágenes se denominan EscenarioX_ImagenY. Se suministran las imágenes en formato jpg para su composición y visualización en color, y en formato pgm para procesamiento de SIFT en escala de grises.



Figura 1. Escenario3 (Ría del Nervión)

3. Software utilizado

El desarrollo de la práctica se realizará en Matlab. Para ello se utiliza unos programas desarrollado por David Lowe de la University of British Columbia (<http://www.cs.ubc.ca/~lowe/keypoints/>), a partir de una versión de Álvaro y J.J. Guerrero de la Universidad de Zaragoza. Programas suministrados:

Originales

- match.m: obtiene, entre dos imágenes, los descriptores que se corresponde con la proyección de una misma característica tridimensional (se han modificado algunos parámetros de la versión original). Debe ser ejecutado externamente en el desarrollo de la práctica. Parámetros de entrada:
 - umbral: Umbral para la correspondencia (distRatio). Admite valores entre 0.0 y 1.0. Un valor bajo obliga a que los descriptores sean muy parecidos.
 - representalmagen:

- o 1: Representa la primera imagen con los puntos encontrados
- o 2: Representa la segunda imagen con los puntos encontrados
- o 3: Representa ambas imágenes con una línea que une los puntos casados
- image1: Nombre de la primera imagen en formato .pgm
- image2: Nombre de la segunda imagen en formato .pgm

Resultados devueltos

- puntosMatch: Puntos casados. En cada fila devuelve el valor x y el valor y del punto correspondiente en la primera imagen, y el valor x y el valor y del punto correspondiente en la segunda imagen. Posee tantas filas como puntos casados.
- im1: Matriz imagen1, en formato numérico (empleado para la representación).
- im2: Matriz imagen2, en formato numérico (empleado para la representación).
- sift.m: calcula los descriptores SIFT. Las imágenes tienen que estar en formato pgm. No es necesario ejecutarlo externamente.
- siftWin32.exe: ejecutable para calcular los descriptores SIFT. No es necesario ejecutarlo externamente.
- showkeys.m: muestra los descriptores encontrados sobre la imagen. No es necesario ejecutarlo externamente.
- LICENSE: información de la licencia.

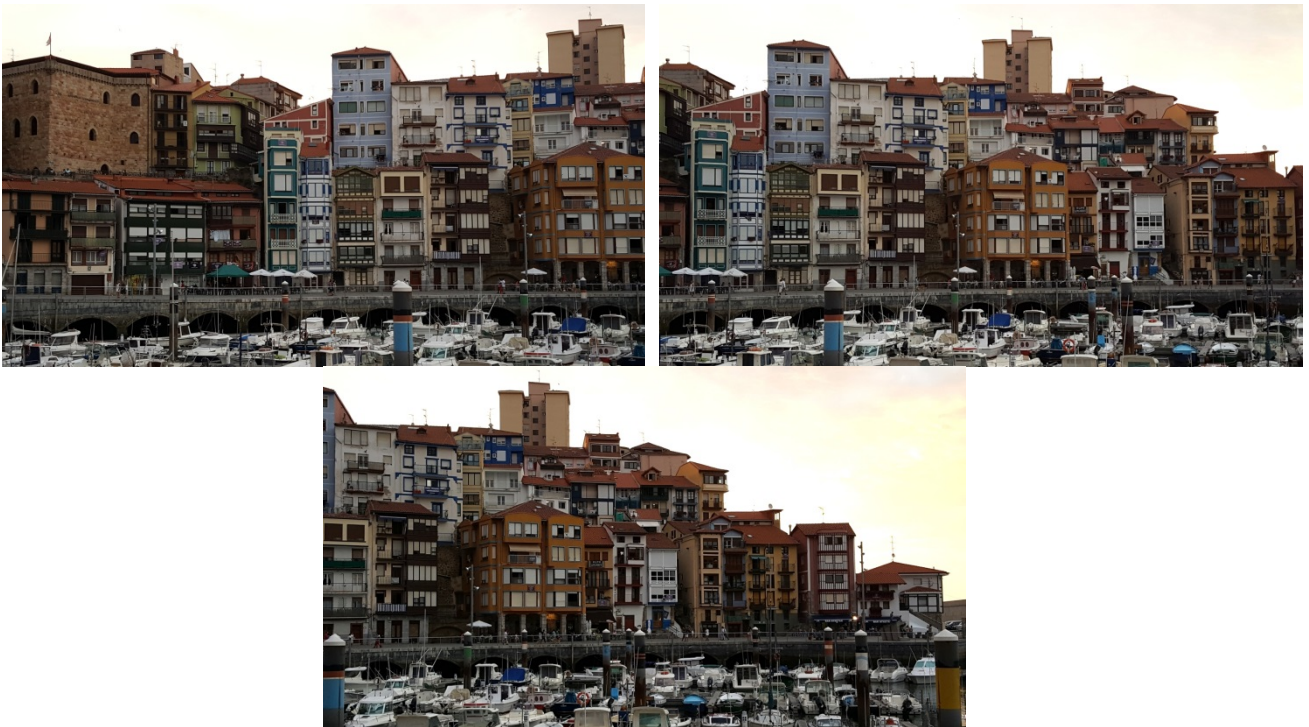


Figura 2. Escenario4 (Bermeo)

Otros

- Prueba3.m: programa de prueba para el Escenario3.
- Prueba4.m: programa de prueba para el Escenario4.
- Prueba8.m: programa de prueba para el Escenario8.

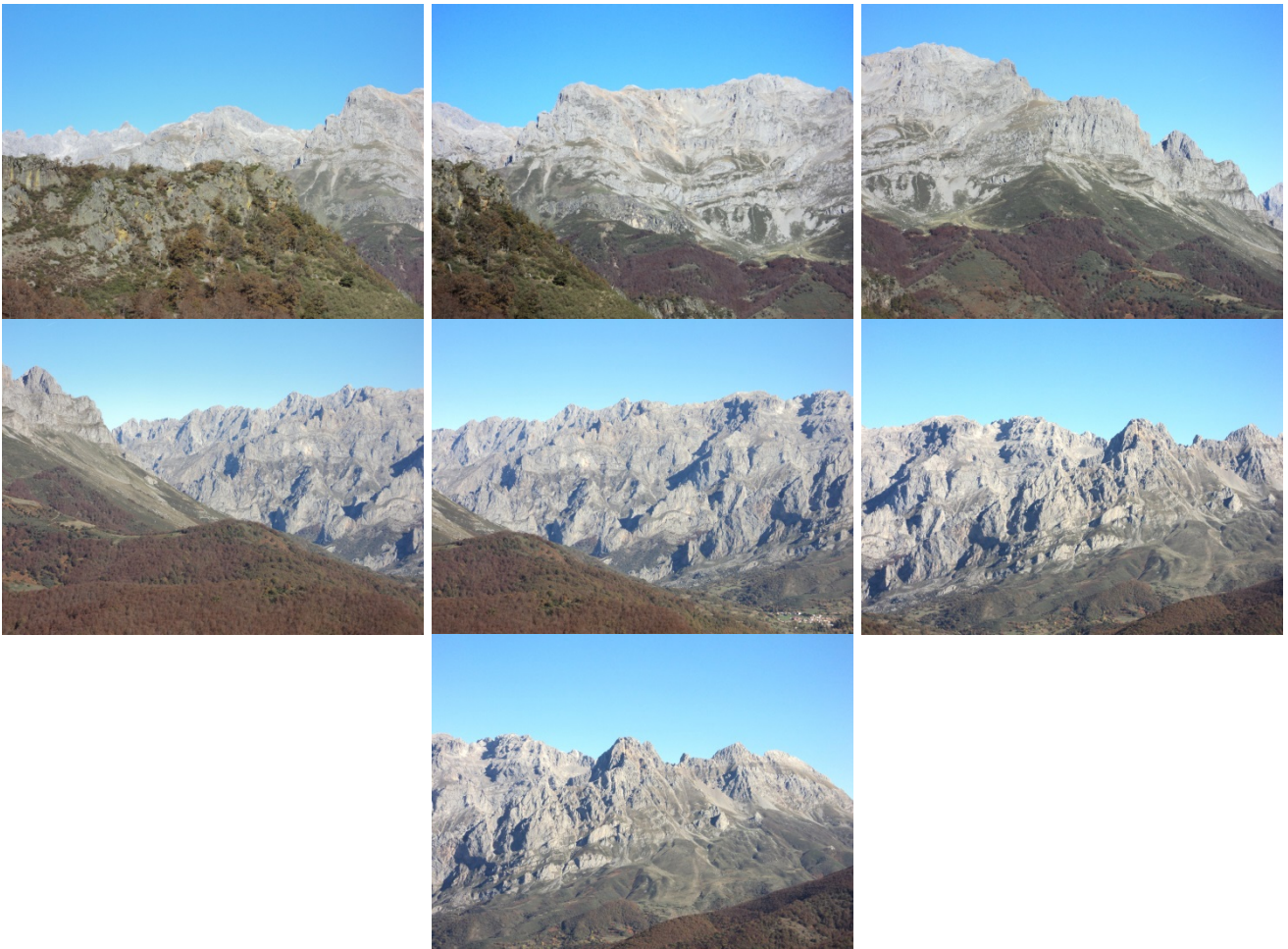


Figura 2. Escenario8 (Picos de Europa)

4. Descripción de la práctica

A continuación se detallan las etapas de las que se compone la ejecución de la práctica. En cada una de ellas se describe una consideración sobre su carácter (obligatoria, opcional o recomendable).

4.1. Detección de características visuales (Obligatorio).

Las características empleadas son las SIFT (Scale Invariant Feature Transform), definidas por D. Lowe en 2004. Son características invariantes ante el cambio de escala, la iluminación, la rotación y ante cualquier transformación afín. En el contexto de la práctica se calculan en la función `sift.m`, a través de la función `match.m`

4.2. Establecimiento de correspondencia entre SIFT (Obligatorio).

El establecimiento de correspondencias (o casamiento de puntos) se realiza mediante la función `match.m`, que a su vez se basa en el cálculo de los descriptores SIFT. Para guiar el número de correspondencia se utiliza el parámetro *umbral* (entre 0.0 y 1.0). Valores muy bajos exigen una alta similitud entre los descriptores, por lo que se detectarán pocos puntos. Valores más altos permiten la detección de más puntos, pero con mayor posibilidad de falsas correspondencias. Los valores utilizados en el parámetro *umbral* tienen que poseer tres cifras decimales, todas ellas distintas del dígito cero. En la figura 3 se muestran las correspondencias detectadas entre la imagen 1 y la imagen 2 del Escenario8, antes de realizar un filtrado. Se han detectado 1414 puntos correspondientes con un umbral de 0.400

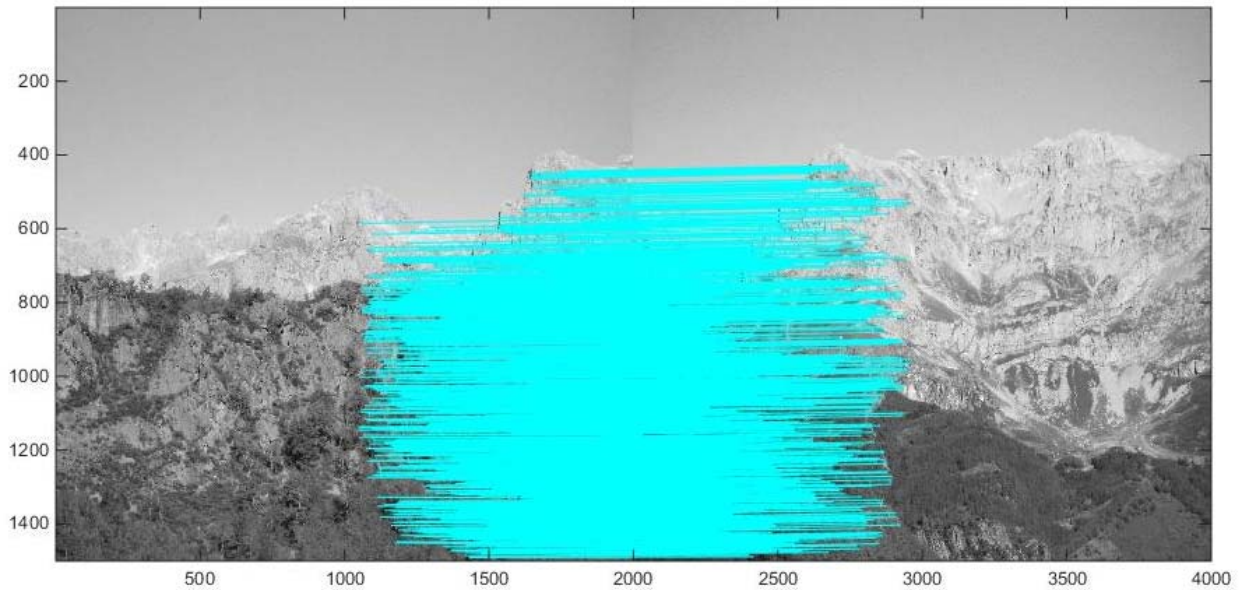


Figura 4. Escenario8. Puntos casados entre la imagen 1 y la imagen 2, sin filtrado

4.3. Cálculo de la Homografía (Obligatorio).

Cuando no hay traslación apreciable entre los centros ópticos de dos adquisiciones de una escena es posible definir una homografía entre los puntos correspondientes de ambas imágenes.

Si el punto de la primera imagen $\tilde{m}_1 = [x_1 \ y_1 \ 1]^T$ se corresponde con el punto de la segunda imagen $\tilde{m}'_1 = [x'_1 \ y'_1 \ 1]^T$ se cumplirá $\tilde{m}'_1 = H \tilde{m}_1$ o desarrollando :

$$\lambda \begin{bmatrix} x'_1 \\ y'_1 \\ 1 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix} \quad \text{donde los puntos son conocidos y las } h_{ij} \text{ son incógnitas.}$$

Se pueden utilizar distintos métodos para calcular la homografía. El más sencillo es el método lineal, en el que se supone que un elemento de la homografía es 1.0, se forma un sistema con más ecuaciones que incógnitas y se resuelve por mínimos cuadrados. Se necesitan 4 o más puntos correspondientes.

Una vez calculada la homografía se puede calcular los puntos en la segunda imagen obtenidos al aplicar la homografía sobre los puntos de la primera imagen:

$$\tilde{p}'_1 = H \tilde{m}_1 \quad \text{donde el error } e'_1 = \|\tilde{p}'_1 - \tilde{m}'_1\| \text{ indicará la bondad de la estimación}$$

4.4. Filtrado de falsas correspondencias (Recomendable).

La existencia de falsas correspondencias puede ocasionar un elevado error en los cálculos posteriores. Para eliminarlas se proponen diversos métodos de filtrados que podrán ser usados por el alumno:

Eliminación de los puntos con elevado error residual en el cálculo de la Homografía. Tras un primer cálculo de la Homografía se puede analizar el error que produce cada punto en la fórmula (que debería ser cero, en un caso ideal), y hallar la media de la distancia entre los puntos en la imagen y los obtenidos al aplicar la homografía. Se propone el rechazo de aquellos puntos cuyo valor absoluto superan en un cierto número de

veces dicho valor medio (por ejemplo tres veces). Se puede volver a calcular la Homografía con los puntos que permanecen activos (distancia menor), y así conseguir eliminar los puntos que pueden ser falsas correspondencias. Se puede repetir el proceso varias veces, para obtener una Homografía más precisa.

Empleo del método RANSAC para eliminar falsas correspondencias. El método RANSAC (Random Sample Consensus) permite eliminar valores atípicos en una muestra.

En la figura 5 se representan las correspondencias de la Escena3, entre la imagen 1 y la imagen2.

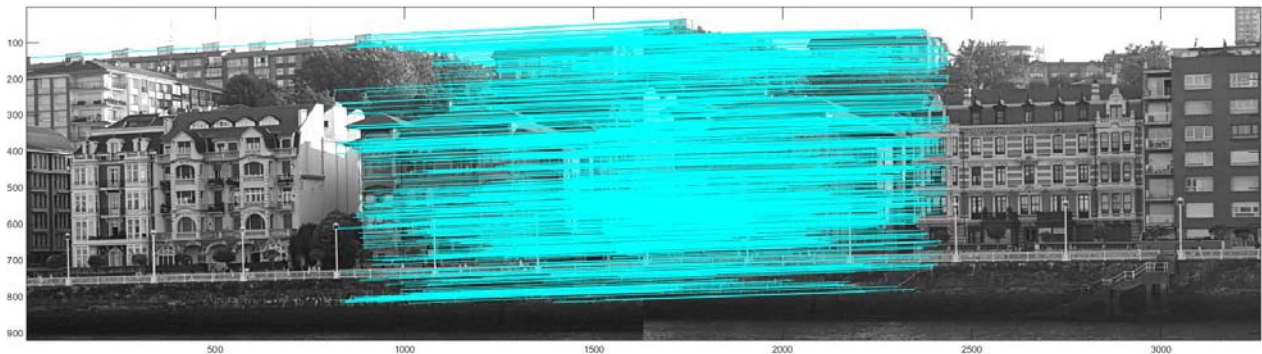


Figura 5. Escenario3. Puntos casados entre la imagen 1 y la imagen 2, sin filtrado

En la figura 6 se estima la homografía, y se calculan los errores sobre la primera imagen. Se representan los puntos casados en cyan (con anchura 3x3) y en rojo una línea que une con el punto obtenido al aplicar la homografía, que solo se visualiza cuando el error supera los 2.0 píxeles. La media del error es de 0.71 píxeles, y el máximo error 81.0 píxeles. La línea roja indica una falsa correspondencia.

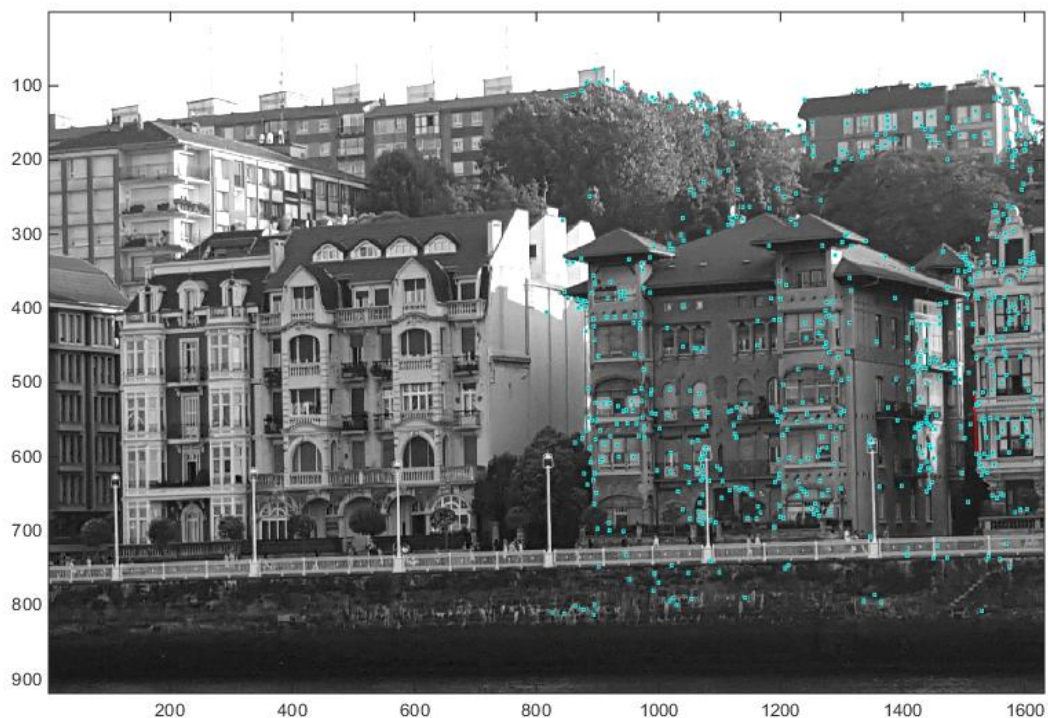


Figura 6. Correspondencias y error al aplicar la homografía (en rojo cuando es mayor de 2 píxeles), sin filtrado

En la figura 7 se representa el resultado cuando se eliminan los puntos que superan tres veces la media. Ahora la media del error es de 0.49 píxeles, y el máximo error 1.99 píxeles.

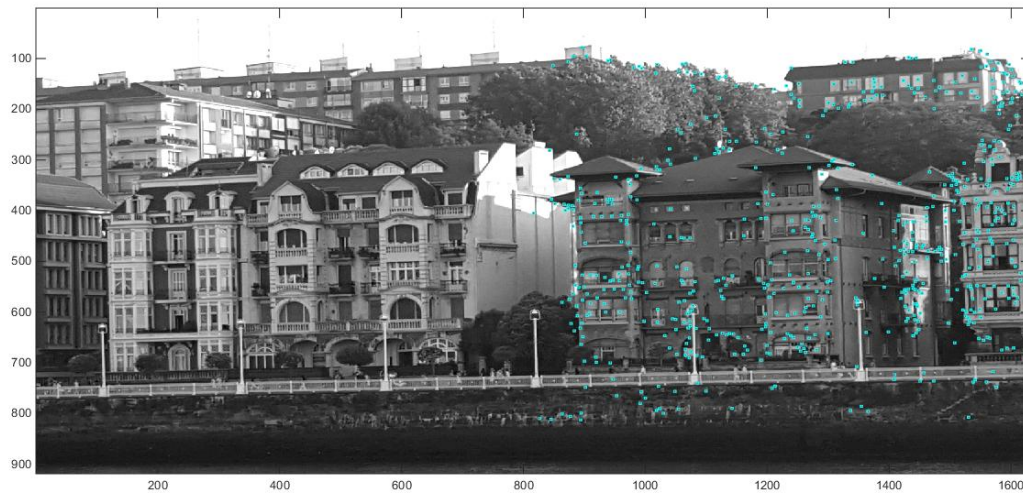


Figura 7. Correspondencias y error al aplicar la homografía (en rojo cuando es mayor de 2 píxeles), con filtrado

En las figuras 8, 9 y 10 se repite el proceso, pero sustituyendo la imagen 2 por otra imagen obtenida cambiando el punto de vista (con desplazamiento). En tal caso se pueden establecer correspondencias, pero no se puede ajustar un modelo con una homografía.

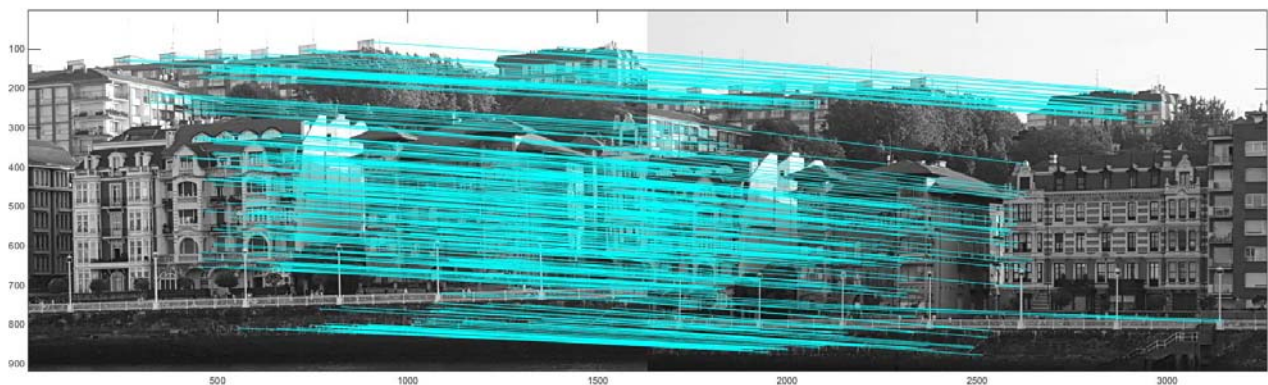


Figura 8. Escenario3. Puntos casados entre la imagen 1 y otra imagen (con desplazamiento), sin filtrado

En este caso (figuras 9 y 10), las líneas rojas no representan necesariamente falsas correspondencias. Son puntos que no cumplen el modelo establecido a partir de una homografía. En la figura 9, sin filtrar, el error medio es de 201.9 píxeles y el máximo de 2951.5 píxeles. En la figura 10, filtrando eliminando los puntos que superan en tres veces la media, el error medio es de 84.0 píxeles y el máximo de 520.4 píxeles. Que disminuya no significa que mejore el modelo; simplemente que se eliminan los puntos más extremos.

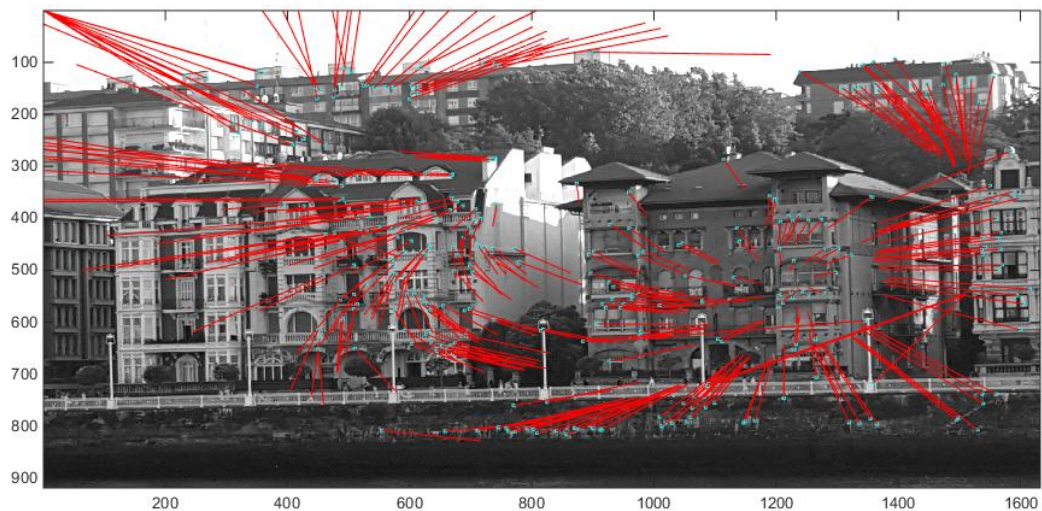


Figura 9. Correspondencias y error al aplicar la homografía (en rojo cuando es mayor de 2 píxeles), sin filtrado

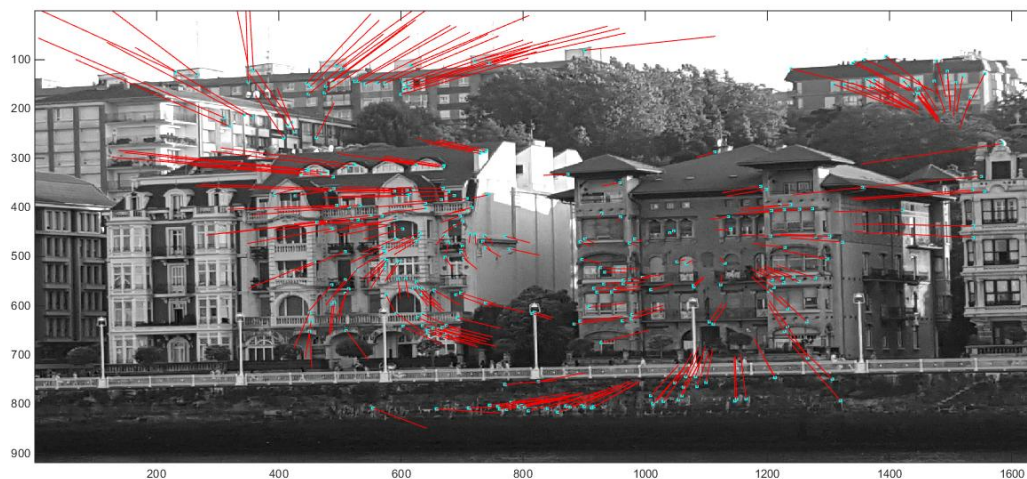
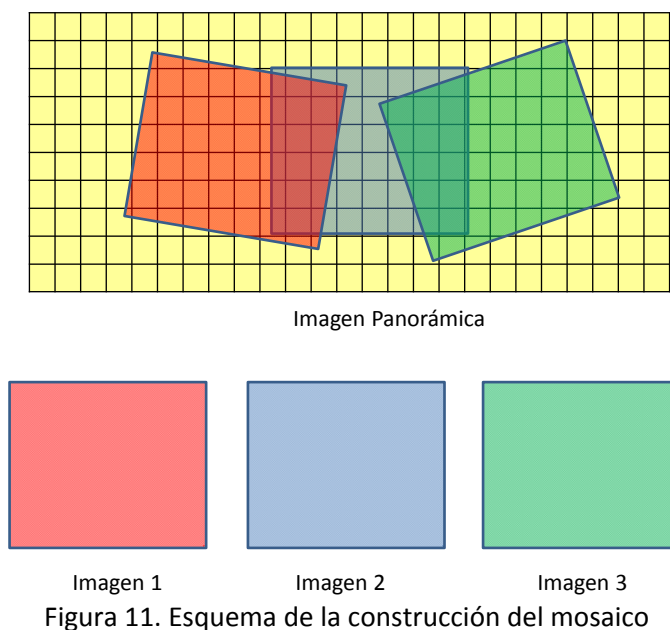


Figura 10. Correspondencias y error al aplicar la homografía (en rojo cuando es mayor de 2 píxeles), sin filtrado

4.5. Construcción del mosaico o imagen panorámica (Obligatorio).

En la construcción de la imagen panorámica, es necesario fijar las dimensiones de la imagen final. Se recomienda emplear una imagen más grande (a fin de evitar escritura fuera de la imagen) y posteriormente eliminar los márgenes negros (por ejemplo repitiendo el procesamiento utilizando los límites obtenidos previamente).

Aunque no es el único método, se recomienda partir de la imagen central y adaptar según la homografía (o el producto de las homografías entre imágenes) la posición que tendría un píxel de cualquier imagen sobre la imagen de referencia (que se propone que sea la central). Al aplicar la homografía es normal que sobre un mismo píxel de la imagen final se acumulen más de un punto original. Se propone realizar la media entre ellos.



También puede ocurrir que sobre un mismo píxel de la imagen final no se acumule ningún punto. Esto ocurre en los márgenes de la imagen (no se puede evitar), pero también en píxeles de discontinuidad de la aplicación de la homografía (la homografía es continua, pero la aplicamos sobre los píxeles discretos). Para evitar este efecto se recomienda aplicar también sobre subpíxel (por ejemplo en el extremo y en el punto intermedio del píxel). Este aspecto puede incrementar el tiempo de procesamiento. En la figura 12 se aprecia el efecto que se produce al considerar solo los valores enteros de los píxeles.



Figura 12. Mosaico obtenido considerando solo los valores enteros de los píxeles

En la adquisición de imágenes para una homografía, se recomienda emplear exposición manual y bloquear el balance de blancos. Esto no siempre es posible, y no asegura un perfecto encuadre de las intensidades de la imagen, ya que la óptica no es igualmente sensible en todos los puntos (efecto vignetting). Existen opciones para minimizar estos efectos, pero no se consideran necesarias para la realización de la presente práctica.

En las figuras 12, 14 y 15 se representan las imágenes panorámicas obtenidas en los tres escenarios.



Figura 13. Panorámica del Escenario 3



Figura 14. Panorámica del Escenario4



Figura 15. Panorámica del Escenario8

5. Documentos generados y entregados

La realización de la práctica debe realizarse de forma individual. El plazo de entrega es el 8 de enero. Es necesario entregar **un documento en el que se describa el trabajo realizado** y que incluya los datos relacionados a continuación, **así como los programas empleados en su obtención**. Se valorará positivamente la utilización de programas propios (salvo la detección de puntos SIFT). Por ejemplo para el

cálculo de la homografía o de filtrado utilizando el método RANSAC. La dirección de entrega será jose.sebastian@upm.es. Se entregará un zip, en el que se indique en nombre del alumno y los escenarios utilizados. Por favor, se ruega no mandar las imágenes de entrada, para que no aumente el tamaño del correo. Por apartados cabe destacar los siguientes:

5.1. Detección de características visuales (Obligatorio).

Ningún dato específico.

5.2. Establecimiento de correspondencia entre SIFT (Obligatorio).

Umbral empleado en la correspondencia (con tres cifras decimales, todos los dígitos distintos de cero).
Número de puntos en lo que se establece la correspondencia.

5.3. Calculo de la Homografía (Obligatorio).

La Homografía.

5.4. Filtrado de falsas correspondencias (Recomendable).

Método empleado en el filtrado de falsas correspondencias. Número de puntos iniciales y número de puntos una vez realizado el filtrado.

5.5. Construcción del mosaico o imagen panorámica (Obligatorio).

Programa propio que consiga construir la imagen panorámica, utilizando los descriptores SIFT, de alguno (o de varios) de los escenarios.