

Trabajo Práctico

Reconstrucción 3D Proyectiva y Euclídea

1. Objetivo

El objetivo de la práctica es que el alumno sea capaz de realizar una reconstrucción tridimensional a partir de la información de una escena suministrada por la adquisición de imágenes desde distintos puntos de vista.

2. Escenarios

Se trabajará con uno de los dos escenarios mostrados en las figuras 1 y 2 (opcionalmente con los dos). Cada escenario está formado por tres planos con impresiones con un alto contenido de información. Se adquieren tres imágenes desde distinto punto de vista de cada escenario, denominadas EscenarioX_Imagen1, EscenarioX_Imagen2 y EscenarioX_Imagen3. Las más alejadas son la 1 y la 3. Se suministran las imágenes en formato jpg para su visualización en color, y en formato pgm para procesamiento en escala de grises (en este caso con una resolución de 1500 filas y 2000 columnas).



Figura 1. Escenario1

En el escenario se han dispuesto 12 marcas (círculos negros) con posiciones conocidas. Su utilización (el conocimiento externo de su posición) está limitada a la representación y a la Reconstrucción Euclídea. No se pueden utilizar para el cálculo de la Matriz Fundamental, pues supone un conocimiento del entorno del que normalmente no se dispone.



Figura 2. Escenario2

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:

- **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.
 - representImagen:
 - 1: Representa la primera imagen con los puntos encontrados
 - 2: Representa la segunda imagen con los puntos encontrados
 - 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.
- **DibujaDisparidad.m:** muestra el vector de disparidad sobre los puntos casados en la imagen.
- **DibujaPuntos.m:** muestra los puntos detectados en ambas imágenes (ya incluido en match.m).
- **DibujaPuntosCasados.m:** muestra los puntos detectados con una línea que enlaza los puntos (ya incluido en match.m).
- **DibujaProfundidad.m:** muestra los puntos reconstruidos con la profundidad (Z) codificada con un mapa de color.
- **LICENSE:** información de la licencia.

Otros

- **Prueba1.m:** programa de prueba para el primer escenario.
- **Prueba2.m:** programa de prueba para el segundo escenario.
- **MarcasEscena1.m:** devuelve los valores de las marcas artificiales (círculos negros) tanto en las imágenes 1 y 3 (en píxeles) de la escena 1, como los tridimensionales de la escena (en mm).
- **MarcasEscena2.m:** devuelve los valores de las marcas artificiales (círculos negros) tanto en las imágenes 1 y 3 (en píxeles) de la escena 2, como los tridimensionales de la escena (en mm).

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. Puede ser recomendable la utilización de distintos umbrales en función del apartado: por ejemplo valores más exigentes en el cálculo de la matriz fundamental y valores menos exigentes en la reconstrucción para obtener un mapa más denso. 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 3 del primer escenario, antes de realizar un filtrado. Se han detectado 1205 puntos correspondientes.

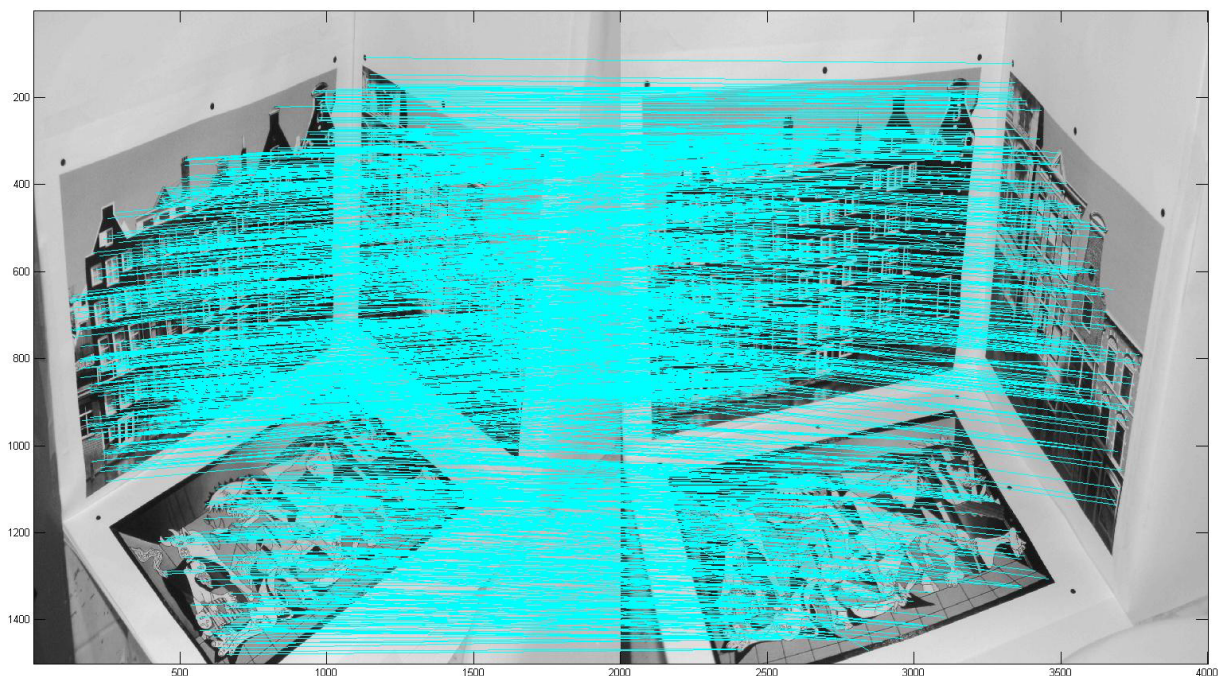


Figura 3. Escenario1. Puntos casados entre la imagen 1 y la imagen 3, sin filtrado

En la figura 4 se representa la disparidad de los puntos casados entre la imagen 1 y la imagen 3 del primer escenario antes de realizar un filtrado (**DibujaDisparidad.m**). El punto en cian representa la posición en la imagen 1 (sobre la que se representa), y el extremo de la línea azul la posición en la imagen 3.

Si $m_1 = [x_1 \ y_1 \ 1]^T$ se corresponde con $m_2 = [x_2 \ y_2 \ 1]^T$ la disparidad será: $d = [x_2 - x_1 \ y_2 - y_1]$

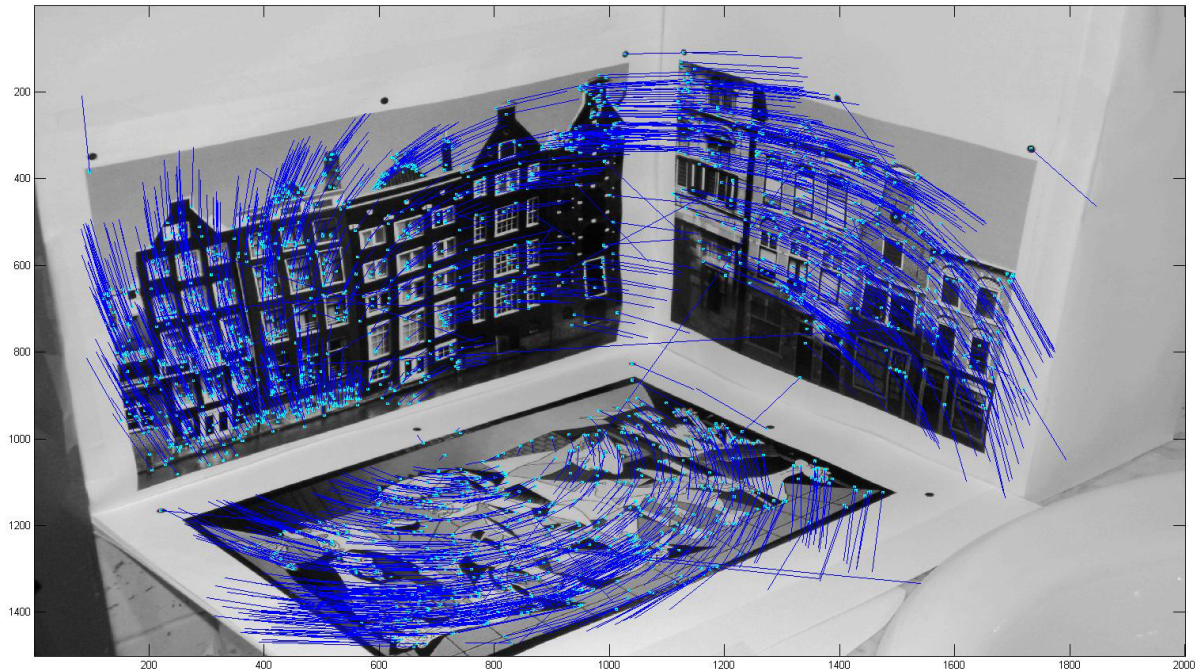


Figura 4. Escenario1. Disparidad entre la imagen 1 y la imagen 3, sin filtrado

En la figura 4 se observan algunas líneas que no siguen la orientación y la longitud de la zona en la que se ubican. Si se supone que la figura posee una cierta continuidad figural, un alto gradiente de disparidad indica la existencia de falsas correspondencias que deben ser eliminadas en una etapa de filtrado.

4.3. Cálculo de la Matriz Fundamental (Obligatorio).

La Matriz Fundamental recoge la información geométrica de la disposición binocular, salvo un escalado en la traslación.

Si el punto de la primera imagen $m_1 = [x_1 \ y_1 \ 1]^T$ se corresponde con el punto de la segunda imagen $m_2 = [x_2 \ y_2 \ 1]^T$ la Matriz Fundamental cumplirá: $m_2^T F m_1 = 0$ o desarrollado:

$$\begin{bmatrix} x_2 & y_2 & 1 \end{bmatrix} \begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix} = 0 \quad \text{donde los puntos son conocidos y las } f_{ij} \text{ son incógnitas}$$

Será necesario calcular también los epipolos de cada imagen: $F e_1 = 0$; $F^T e_2 = 0$

Se pueden utilizar distintos métodos para calcular la matriz Fundamental. El más sencillo es el método lineal, en el que se utilizan más de ocho puntos correctamente casados, se forma un sistema con más ecuaciones que incógnitas y se resuelve por mínimos cuadrados. No se tiene en cuenta que el determinante de la matriz fundamental tiene que ser nulo. En el desarrollo de la práctica se puede emplear cualquier método de los descritos en el temario de la asignatura.

El cálculo preciso de la matriz Fundamental es imprescindible para lograr una buena precisión en la reconstrucción proyectiva. Por tal motivo es recomendable eliminar las falsas correspondencias como se detalla en el siguiente apartado. También es importante emplear puntos casados que procedan de tres planos distintos de la escena. No está permitida utilizar la información suministrada en los ficheros MarcasEscenax.m sobre la posición de las marcas (círculos negros) de la escena.

El cálculo de la Matriz Fundamental se debe realizar entre las vistas más alejadas (Imagen1-Imagen3).

4.4. Filtrado de falsas correspondencias (Obligatorio).

La existencia de falsas correspondencias puede ocasionar un elevado error en los cálculos posteriores (Matriz Fundamental, Reconstrucción Proyectiva). 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 Matriz Fundamental. Tras un primer cálculo de la Matriz Fundamental 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 los valores absolutos (o el error cuadrático) de todos los puntos. 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 Matriz Fundamental con los puntos que permanecen activos (el valor absoluto es menor), y así conseguir eliminar los puntos que pueden ser falsas correspondencias. Se puede repetir el proceso varias veces, para obtener una Matriz Fundamental más precisa. En la figura 5 se representan los puntos filtrados de la figura 4, y como quedaría su disparidad. Se ha calculado la Matriz Fundamental tres veces, eliminando los puntos que posee un error en valor absoluto mayor que tres veces el valor absoluto medio.

Eliminación de los puntos que no poseen correspondencia en la imagen intermedia. Si se han utilizado las imágenes 1 y 3 para el establecimiento de correspondencia, se pueden eliminar aquellas correspondencias que no se produzcan también en la imagen intermedia, la imagen 2.

Eliminación de puntos con un alto gradiente de disparidad. Se propone la construcción de un mapa de disparidad, similar al reflejado en la figura 4, y eliminar aquellos puntos cuya disparidad sea significativamente distinta a la de su zona.

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

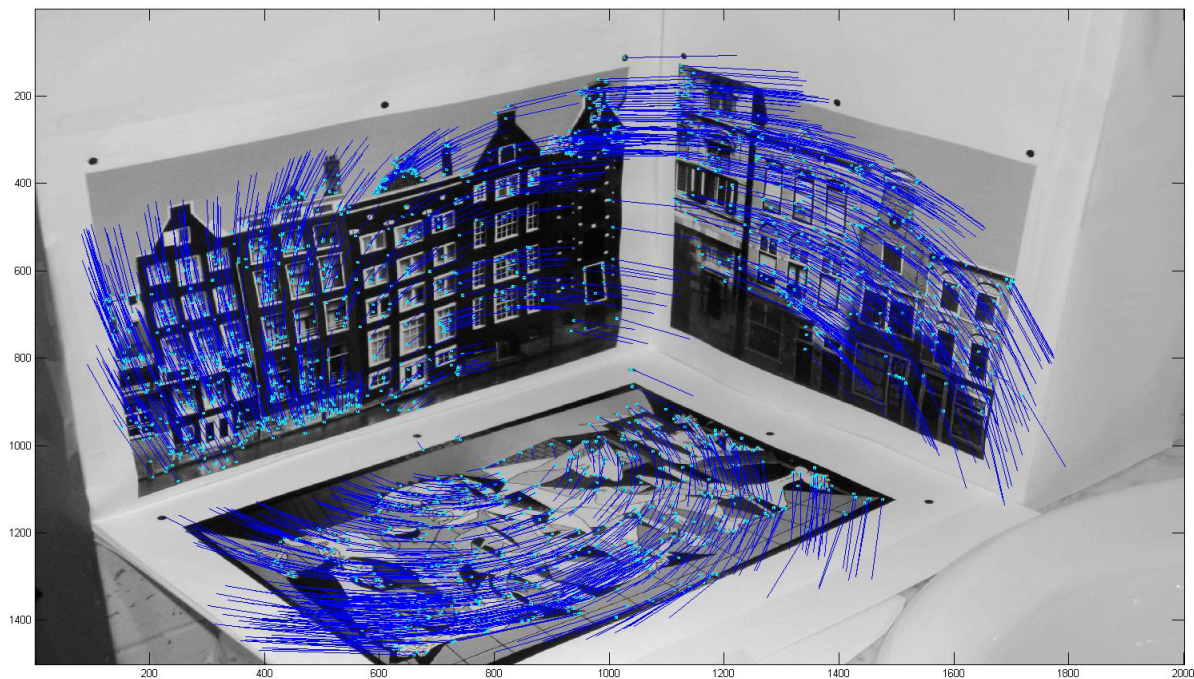


Figura 5. Escenario1. Disparidad entre la imagen 1 y la imagen 3, filtrada

4.5. Reconstrucción Proyectiva 3D (Obligatorio).

La Reconstrucción Proyectiva 3D de la escena permite incorporar información de profundidad en la escena detectando los puntos más cercanos y los puntos más lejanos; suministra pues información relativa. No suministra información métrica de la escena. Para aumentar la precisión es recomendable utilizar imágenes de posición extrema (por ejemplo la 1 y la 3). Se puede emplear un número de puntos distintos del utilizado en el cálculo de la Matriz Fundamental. La Reconstrucción Proyectiva 3D de la escena propuesta en la práctica (existen diversos métodos para obtener una reconstrucción proyectiva) se realiza a partir del establecimiento de correspondencias, y del epipolo de la segunda imagen.

Si el punto de la primera imagen $m_1 = [x_1 \ y_1 \ 1]^T$ se corresponde con el punto de la segunda imagen $m_2 = [x_2 \ y_2 \ 1]^T$, y el epipolo de la segunda imagen se representa por $e_2 = [e_{x2} \ e_{y2} \ 1]^T$ la reconstrucción proyectiva ligada al centro óptico de la primera cámara se obtiene con las expresiones:

$$X_p = x_1 \ ; \ Y_p = y_1 \ ; \ Z_p = \sqrt{(e_{x2} + d_x)^2 + (e_{y2} + d_y)^2} \quad \text{con} \quad d = [d_x \ d_y] = [x_2 - x_1 \ y_2 - y_1]$$

En las figuras 6 y 7 se representan en falso color la reconstrucción proyectiva para los escenarios 1 y 2 respectivamente (**DibujaProfundidad.m**). Los puntos más cercanos (al centro óptico en la posición de la imagen 1) se representan en rojo, luego en naranja, posteriormente en amarillo, verde, cyan y luego azul oscuro. La escala en Matlab es la generada por `colormap('jet')`. ATENCIÓN: puntos reconstruidos con un gran error (por falsas correspondencias), pueden causar un inadecuado uso de la escala de colores.

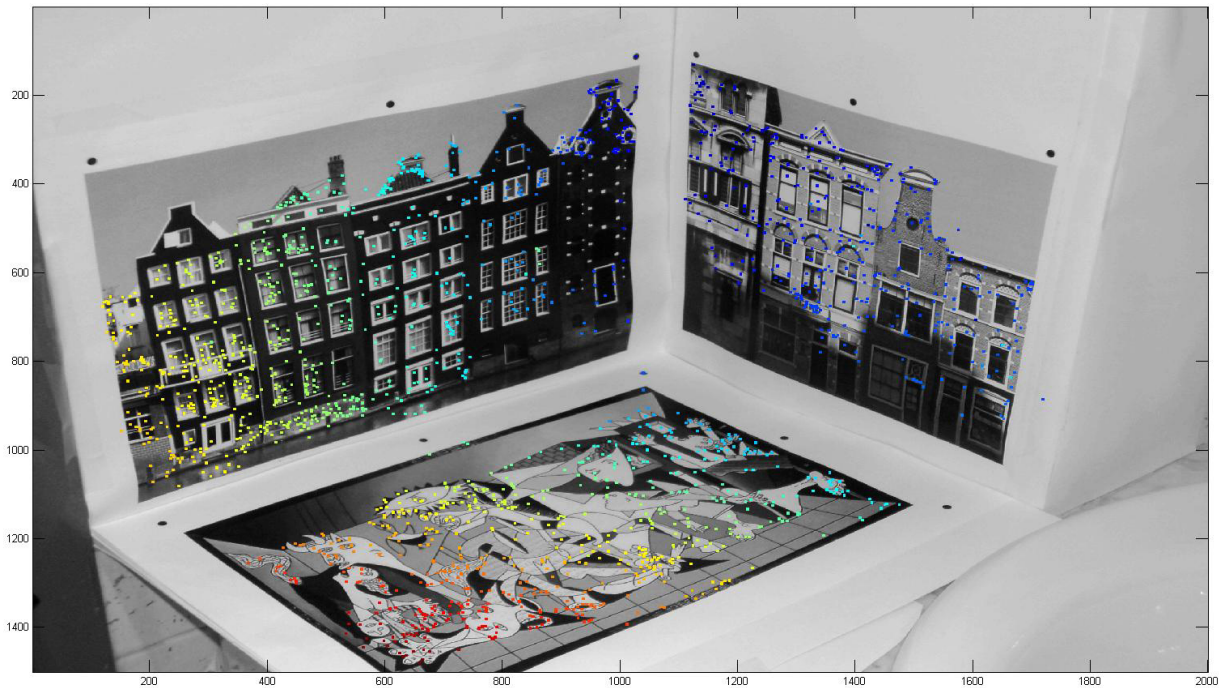


Figura 6. Escenario1. Reconstrucción Proyectiva entre la imagen 1 y la imagen 3, con falso color



Figura 7. Escenario2. Reconstrucción Projectiva entre la imagen 1 y la imagen 3, con falso color

En la figura 8 se representa la Reconstrucción Projectiva obtenida en el escenario 1. La escala de profundidad va de 8000 a 9000 aunque es un dato muy dependiente del cálculo del epipolo. Se han invertido los ejes X e Y para obtener una representación acorde con la imagen captada (el origen de la cámara está en la esquina superior izquierda). Los ejes X e Y se expresan en milímetros. Los puntos obtenidos a partir del descriptor SIFT se representan en azul, mientras que las marcas circulares en asterisco rojo.

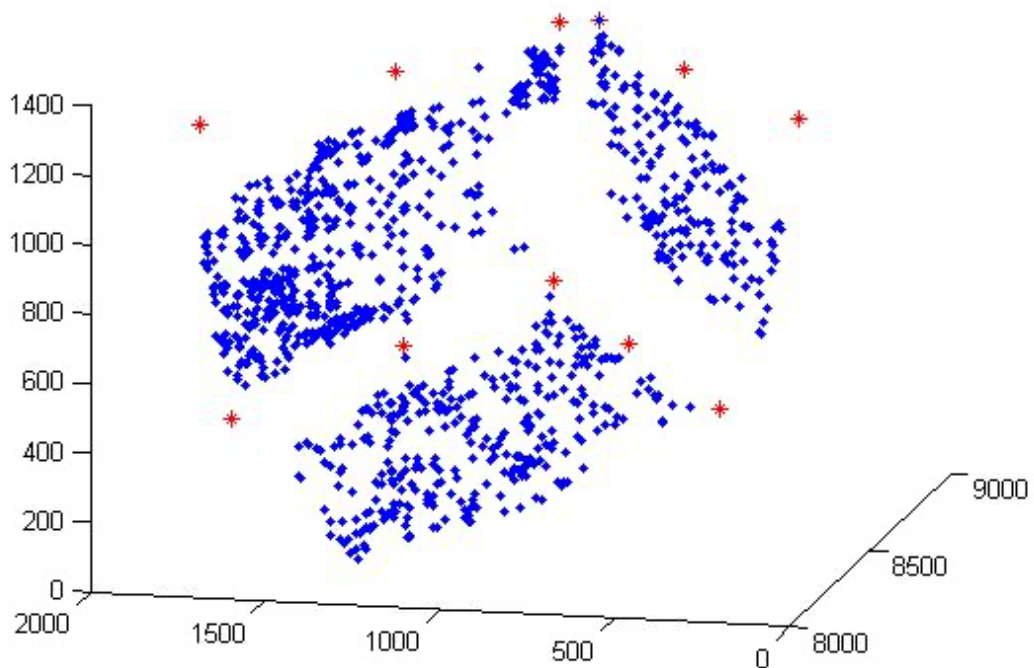


Figura 8. Escenario1. Reconstrucción Projectiva entre la imagen 1 y la imagen 3

4.6. Reconstrucción Euclídea 3D (Recomendable).

Con el conocimiento de algunas marcas de la escena se puede una transformación (una homografía) que relaciona la Reconstrucción Proyectiva con la Reconstrucción Euclídea. Conocida dicha transformación se puede aplicar al resto de los puntos pertenecientes a la Reconstrucción Proyectiva. Para los puntos conocidos se cumplirá (la información se suministra en MarcasEscenaX.m):

$$\lambda \begin{bmatrix} X_E \\ Y_E \\ Z_E \\ 1 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} & h_{14} \\ h_{21} & h_{22} & h_{23} & h_{24} \\ h_{31} & h_{32} & h_{33} & h_{34} \\ h_{41} & h_{42} & h_{43} & h_{44} \end{bmatrix} \begin{bmatrix} X_P \\ Y_P \\ Z_P \\ 1 \end{bmatrix} \quad \text{con los puntos conocidos y las } h_{ij} \text{ incógnitas}$$

La homografía se puede estimar con al menos 5 puntos, formando un sistema lineal que se resuelve por mínimos cuadrados si hay más ecuaciones que incógnitas.

Una vez obtenida la homografía se aplica la misma transformación para los puntos obtenidos en la Reconstrucción Proyectiva, obteniendo así los puntos de la Reconstrucción Euclídea. En la figura 9 se representa la Reconstrucción Euclídea obtenida en el escenario 1. Los ejes se representan en milímetros. El origen está en la intersección de los tres planos de la escena. Los puntos obtenidos a partir del descriptor SIFT se representan en azul, mientras que las marcas circulares en asterisco rojo.

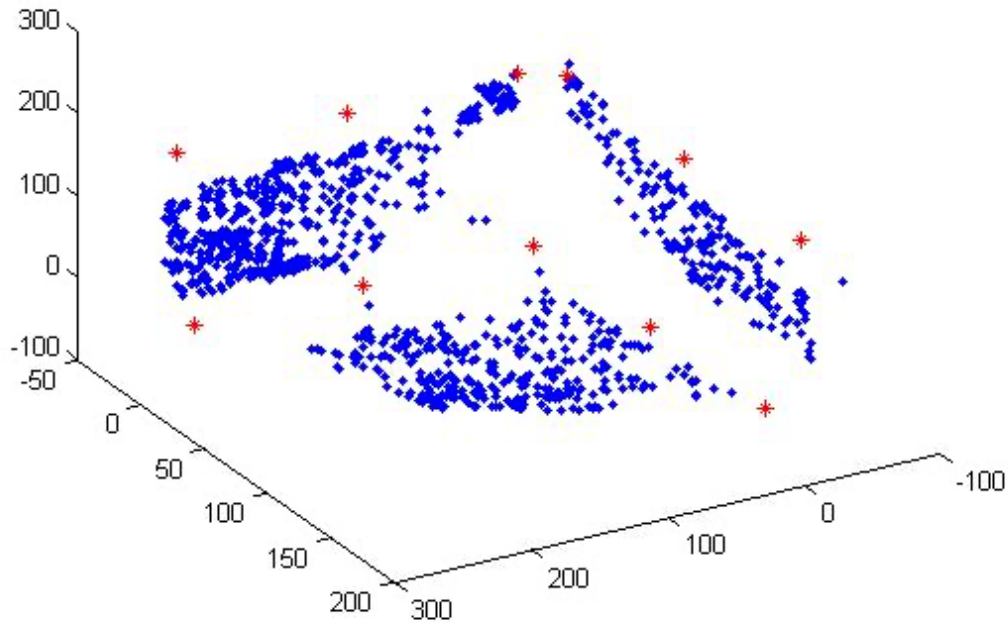


Figura 9. Escenario1. Reconstrucción Euclídea entre la imagen 1 y la imagen 3

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 matriz fundamental, de una homografía o de filtrado utilizando el método RANSAC.

La entrega se realizará en el enlace de tareas de la asignatura del acceso identificado de la UMH. 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 Matriz Fundamental (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. Si se filtran los punto, número de puntos filtrados. Matriz Fundamental y Epipolos obtenidos.

5.4. Filtrado de falsas correspondencias (Obligatorio).

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

5.5. Reconstrucción Proyectiva 3D (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, iniciales y filtrados (si procede). Programa que consiga una representación gráfica de la reconstrucción, similar a alguno de los utilizados para generar las figuras 6, 7 u 8, o cualquier otra que sea representativa.

5.6. Reconstrucción Euclídea 3D (Recomendable).

Homografía obtenida para calcular la Reconstrucción Euclídea a partir de la Reconstrucción Proyectiva. Programa que consiga una representación gráfica de la reconstrucción, similar a alguno de los utilizados para generar la figura 9, o cualquier otra que sea representativa.