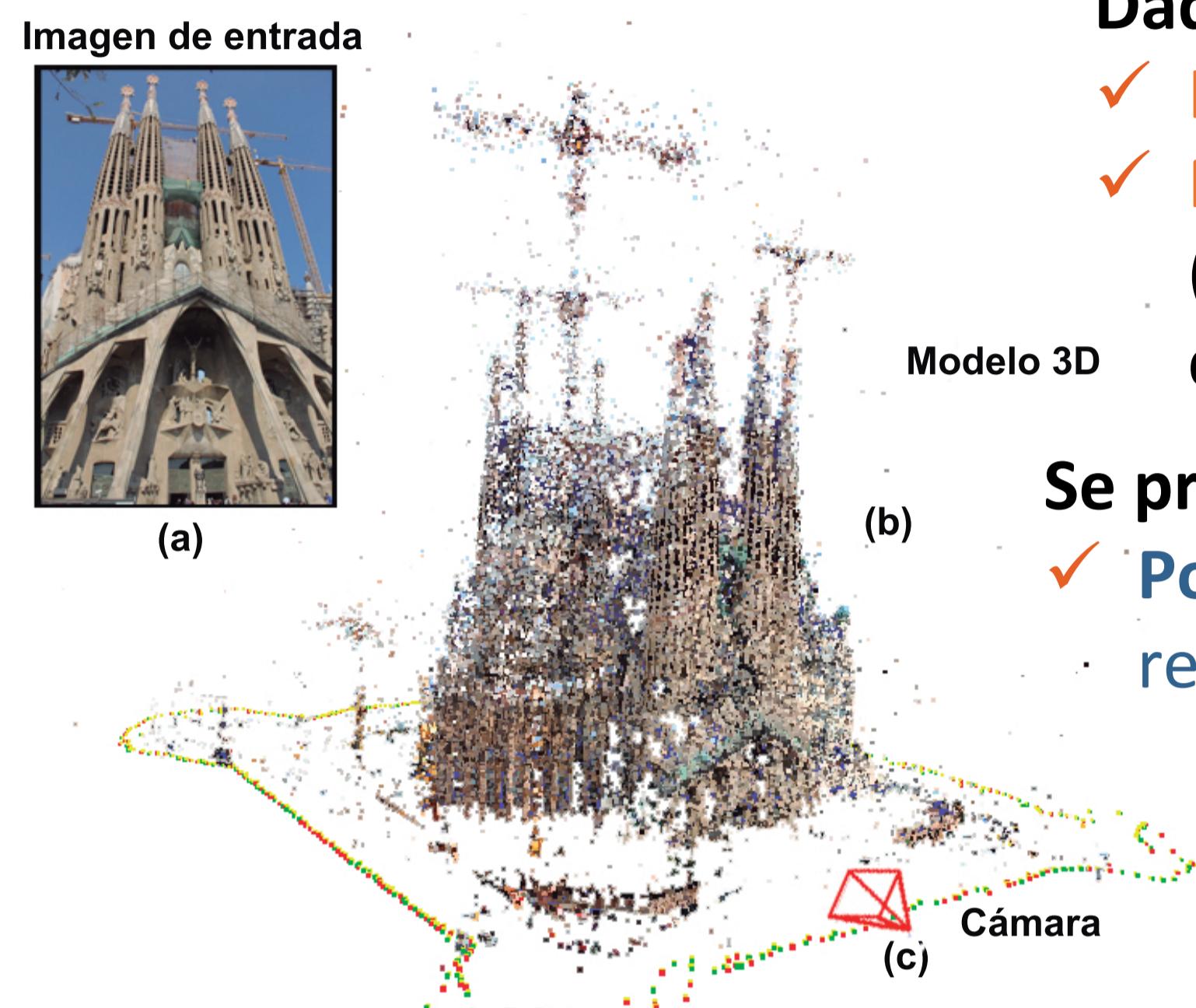


PROBLEMA:

- Dada una imagen de entrada (a) y un modelo 3D del entorno conocido (b), el problema consiste en **encontrar la pose de la cámara** que capturó la imagen con respecto al modelo (c).



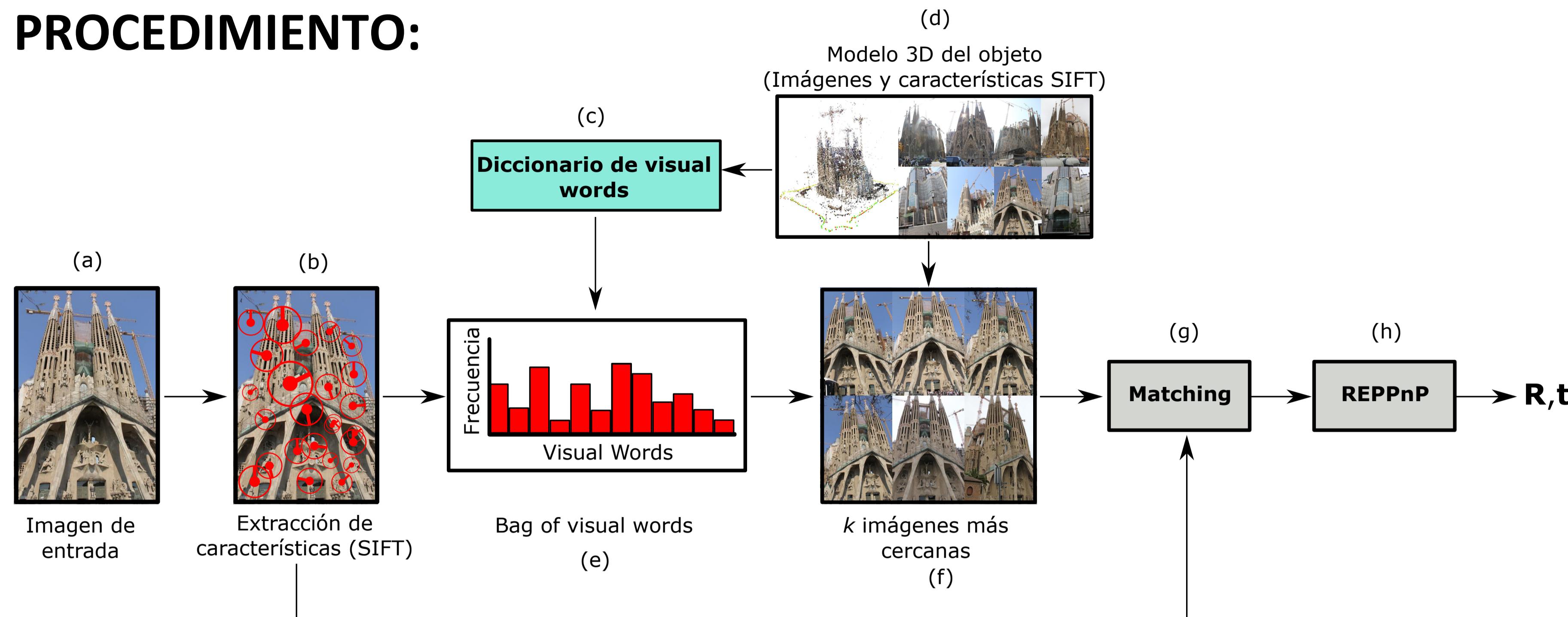
Dados:

- ✓ Imagen de entrada
- ✓ Modelo 3D complejo (en este caso, con más de 100.000 puntos)

Se pretende obtener:

- ✓ Pose de la cámara con respecto al modelo $[R|t]$

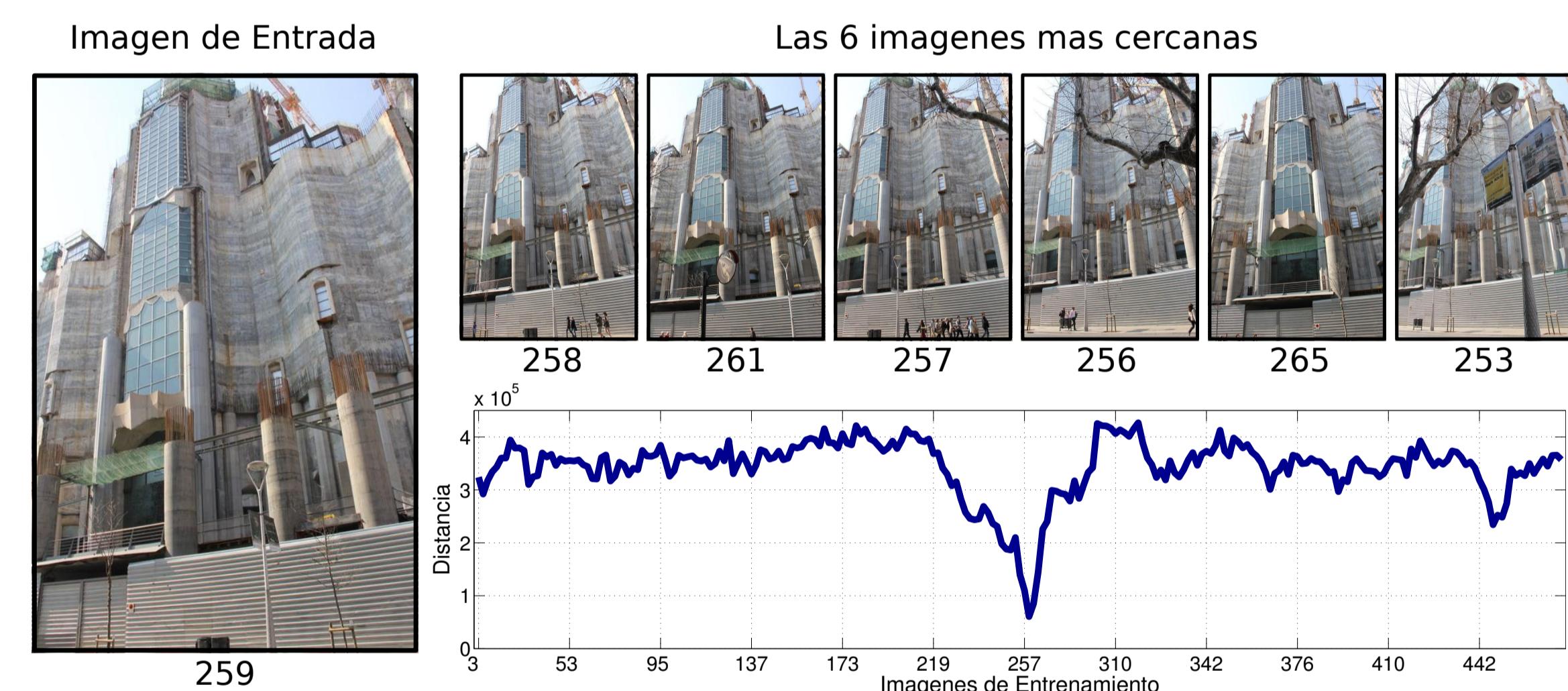
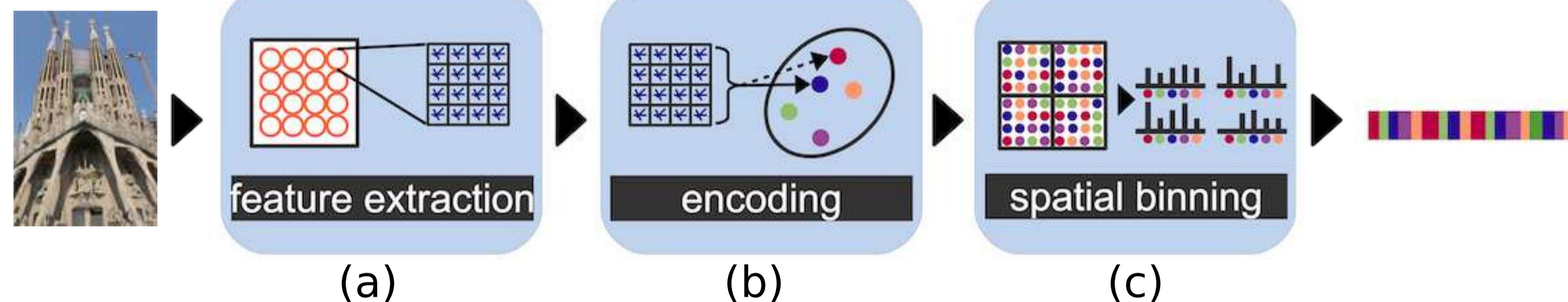
PROCEDIMIENTO:



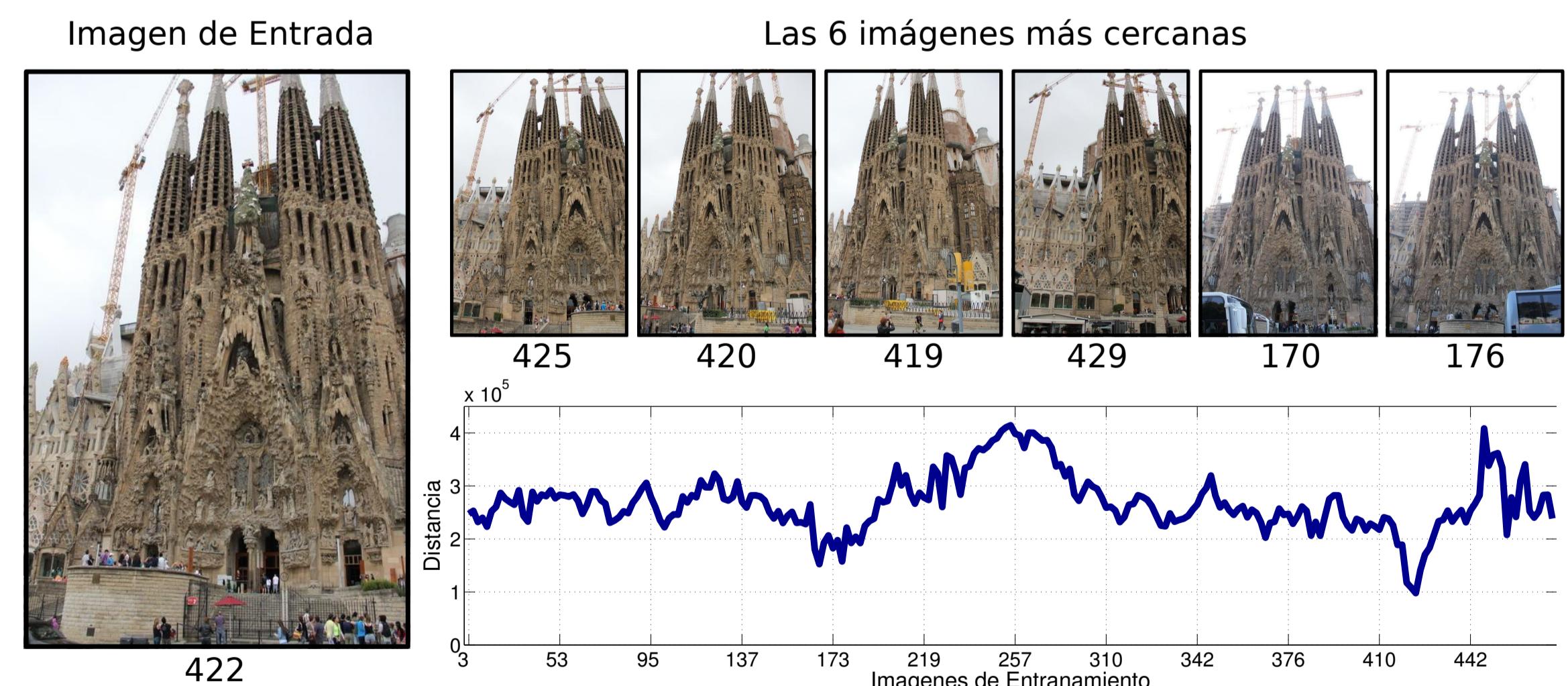
- Dada una **imagen de entrada** (a), se extraen sus **descriptores SIFT** [1] (b). Estos son usados después para codificar la apariencia de la imagen mediante el cómputo del **histograma de BoVW** [2] (e), que se compara con un **diccionario** (c) previamente calculado a partir del **modelo 3D** (d). Se obtienen así las **k imágenes más cercanas** (f), usadas para el proceso de **matching** (g) y, finalmente, para el **REPPnP** [3] (h), obteniendo la pose $[R|t]$ de la cámara.

TÉCNICAS:

①. Bag of Visual Words (BoVW)



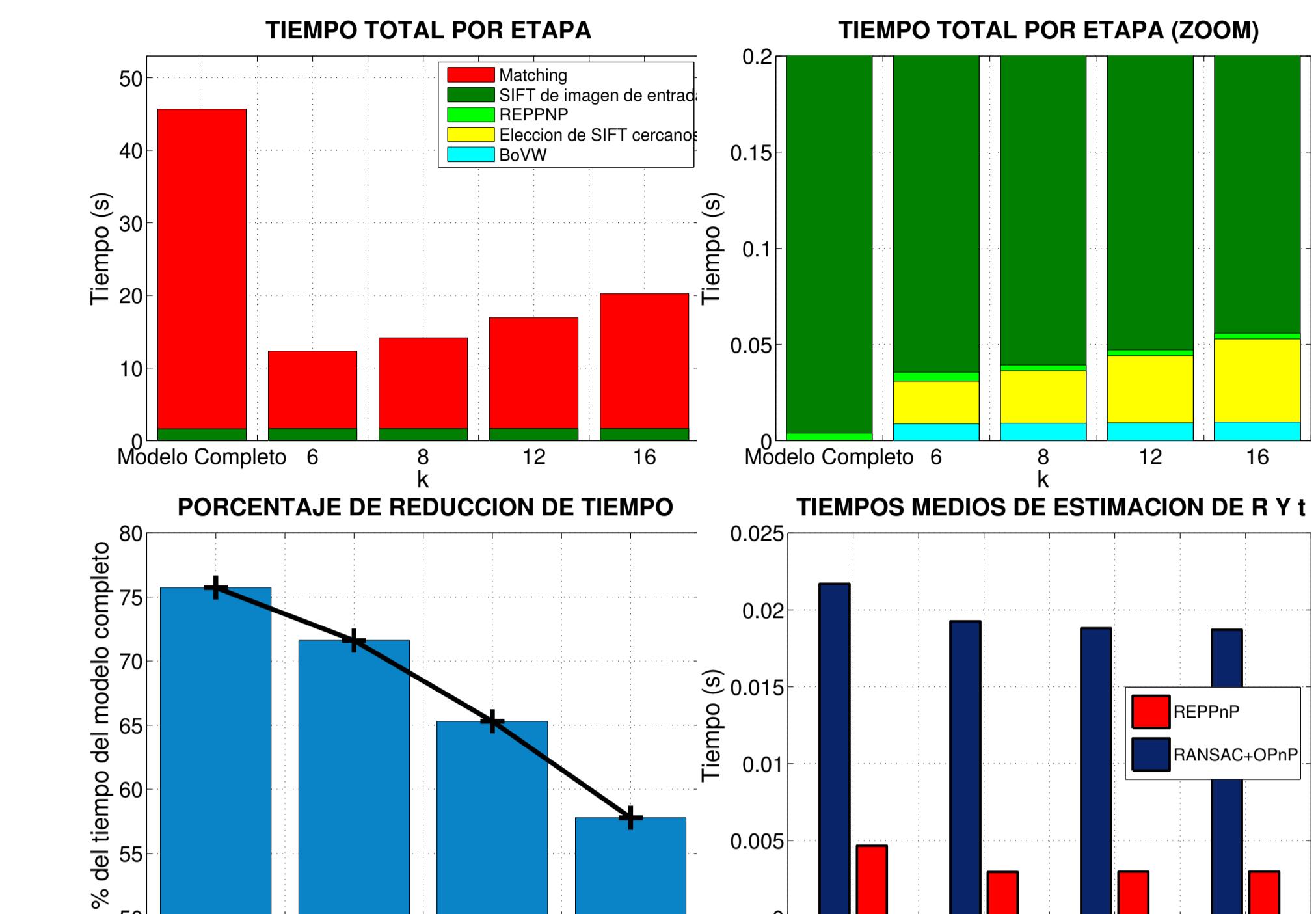
- Creación previa de un **diccionario de visual words** agrupando los descriptores SIFT de las imágenes de entrenamiento mediante el algoritmo *k-means*.
- Al recibir una imagen:
 - ✓ Se extraen los descriptores **SIFT** (a)
 - ✓ Se asignan a los correspondientes **visual words** del diccionario (fase de **encoding**) (b)
 - ✓ Se utilizan **histogramas espaciales** para añadir información espacial (c)
 - ✓ Se extraen las imágenes más cercanas usando el histograma de BoVW.
- Esta etapa **reduce drásticamente el número de imágenes con las que se compara** la de entrada, obteniendo de este modo una reducción considerable en el tiempo empleado en la etapa de **matching**.



②. REPPnP

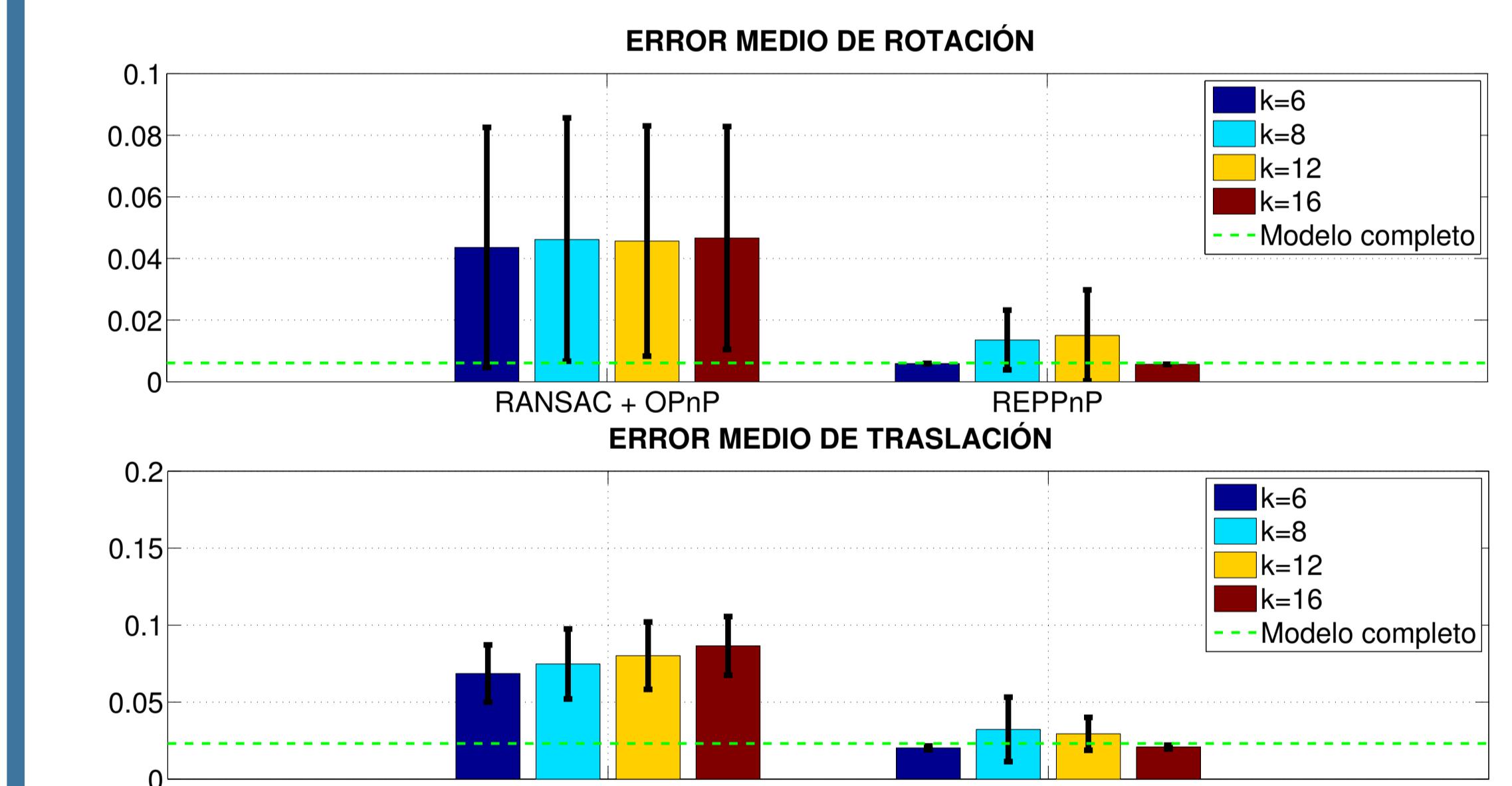
- Método iterativo de **estimación de la pose** que incluye eliminación de **outliers**.
- Hace innecesario por lo tanto el uso de métodos como **RANSAC** combinado con algoritmos P3P.
- Considera el error algebraico del sistema lineal derivado de la formulación del PnP.
- Demuestra **convergencia hasta con un 50% de outliers**.
- Reduce **hasta 100 veces** el tiempo empleado por las estrategias clásicas P3P+RANSAC +PnP.

RESULTADOS: ①. Eficiencia:



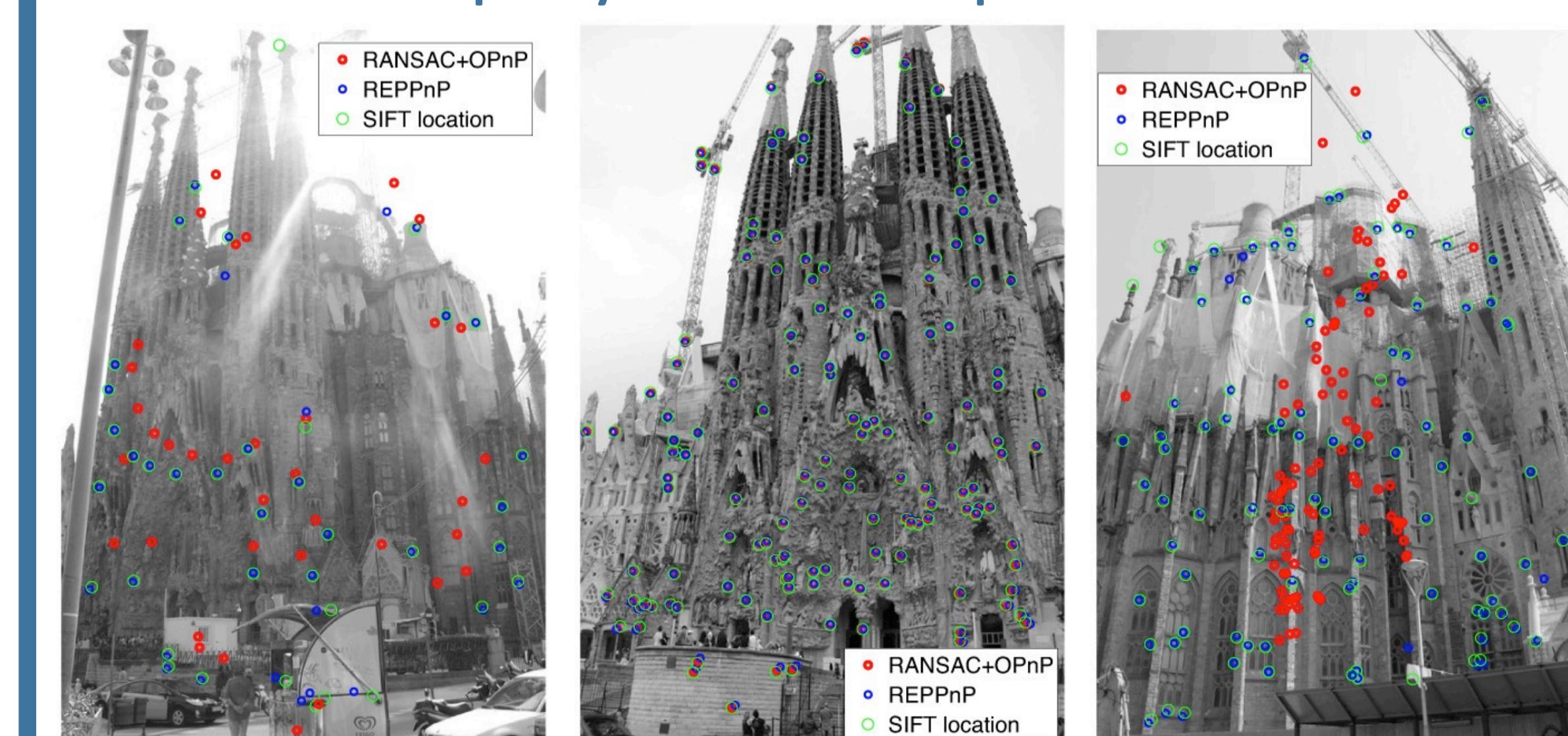
El método propuesto reduce considerablemente el coste de la etapa de *matching* al usar únicamente puntos de las *k* imágenes cercanas seleccionadas por el BoVW en lugar de usar el modelo completo (~100.000 puntos).

②. Precisión:



El método propuesto no afecta a la precisión de la estimación. (Errores relativos. Rotación: norma L_2 entre cuaterniones, traslación: dist. euclídea).

③. Reproyección de puntos SIFT:



REFERENCIAS:

- David Lowe. Distinctive image features from scale-invariant keypoints. In IJCV, 2004.
- Ken Chatfield, Victor Lempitsky, Andrea Vedaldi, and Andrew Zisserman. The devil is in the details: an evaluation of recent feature encoding methods. In BMVC, 2011.
- L Ferraz, X Binefa, F Moreno-Noguer. Very Fast Solution to the PnP Problem with Algebraic Outlier Rejection. In CVPR, 2014.