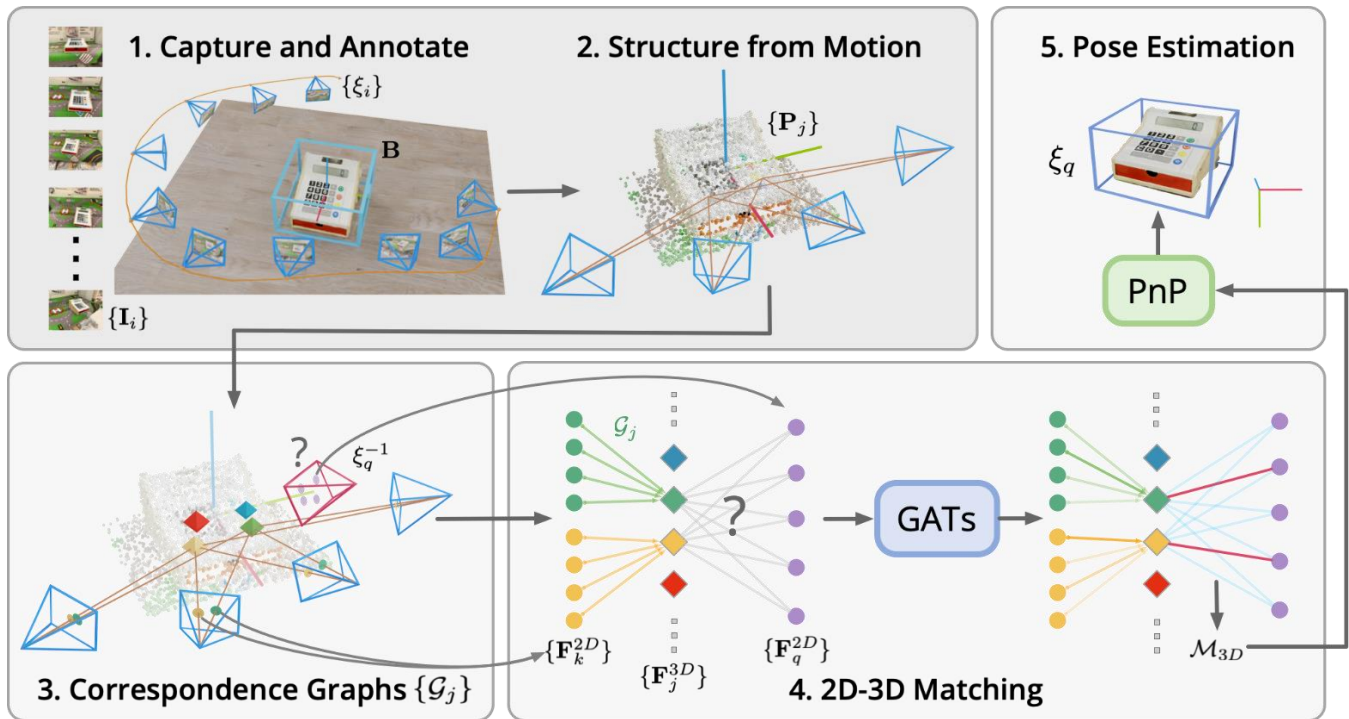


# Reference paper: **OnePose**: One-Shot Object Pose Estimation without CAD Models

Jiaming Sun, Zihao Wang, Siyu Zhang, Xingyi He, Hongcheng Zhao, Guofeng Zhang, Xiaowei Zhou, CVPR, 2022



**1.** For each object, a video scan with RGB frames  $\{I_i\}$  and camera poses  $\{\xi_i\}$  are collected together with the annotated 3D object bounding box  $B$ . **2.** Structure from Motion (SfM) reconstructs a sparse point cloud  $\{P_j\}$  of the object. **3.** The correspondence graphs  $\{G_j\}$  are built during SfM, which represent the 2D-3D correspondences in the SfM map. **4.** 2D descriptors  $\{F_k^{2D}\}$  are aggregated to 3D descriptors  $\{F_j^{3D}\}$  with the aggregation-attention layer.  $\{F_j^{3D}\}$  are later matched with 2D descriptors from the query image  $\{F_q^{2D}\}$  to generate 2D-3D match predictions  $\mathcal{M}_{3D}$ . **5.** Finally, the object pose  $\xi_q$  is computed by solving the PnP problem with  $\mathcal{M}_{3D}$ .

## Uses:

- Robotics
- Augmented Reality (AR)
- Gaming and Virtual Reality (VR)
- Medical Imaging
- Aerospace and Aviation
- Manufacturing and Quality Control
- Automotive
- Architecture and Construction
- Geospatial Applications
- Industrial Automation

Simultaneously performs multiple object detection, accurately determines the depth of specific objects within the scene, and provide robust object classification capabilities.

