Scale-Invariant Feature Transform (SIFT): SIFT is a feature extraction technique that can be used to determine the size of an object in an image by comparing it to a known reference object.

Structure from Motion (SfM): SfM is a technique that can be used to determine the three-dimensional structure of an object from two-dimensional images captured by a moving camera. The size of the object can be estimated from the estimated camera pose and the known dimensions of reference objects in the scene.

Monocular Depth Estimation: Monocular depth estimation is a technique that uses a single camera to estimate the depth of objects in an image. The size of an object can be estimated from its depth and the known intrinsic parameters of the camera.

Image Segmentation: Image segmentation is a technique that can be used to separate the foreground objects from the background in an image. The size of the object can be estimated by computing its bounding box and using the known intrinsic parameters of the camera.