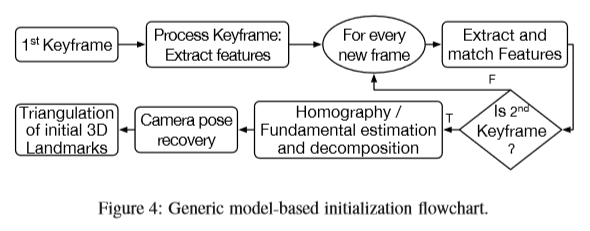
creates the initial 3D map on startup.



the ﬁrst frame processed by the KSLAM system is typically set as the ﬁrst keyframe. Subsequent frames are processed by establishing 2D-2D data associations, which are monitored to decide whether the new frame is the second keyframe or not. The decision criteria is based on the 2D distances between the found matches in both images. The matches are then used to estimate a Homography (degenerate for nonplanar scenes) or a Fundamental matrix (degenerate for planar scenes) using a robust model ﬁtting method (RANSAC or MLESAC). The estimated Homography or the Fundamental matrix are then decomposed as described in [35 (R. Hartley, A. Zisserman, Multiple View Geometry in Computer Vision, Cambridge University Press, 2003.)] into an initial scene structure and initial camera poses. To mitigate degenerate cases, random depth initialization (shown in Fig.3F), as its name suggests, initializes a KSLAM by randomly assigning depth values with large variance to a single initializing keyframe. The random depth is then iteratively updated over subsequent frames until the depth variance converges.

[38] suggested in **LSD SLAM**, and later in DSO, a randomly initialized scene’s depth from the ﬁrst viewpoint, Both systems use an initialization method that does not require two view geometry; it takes place on a single frame: pixels of interest (i.e., image locations that have high intensity gradients) in the ﬁrst keyframe are given a random depth value with an associated

large variance. This results in an initially erroneous 3D map. The pose estimation methods are then invoked to estimate the pose of newly incomingframes using the erroneousmap, which in return results in erroneous pose estimates. However, as the system process more frames of the same scene, the originally erroneous depth map convergesto a stable solution. The initialization is considered complete when the depth variance of the initial scene converges to a minimum.