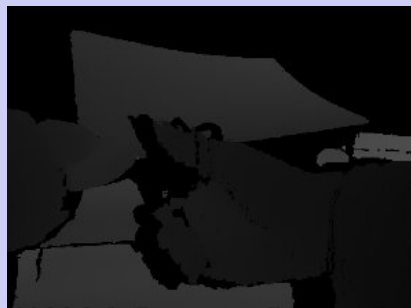


Input at Frame t



RGB@t



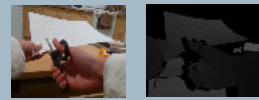
Depth@t



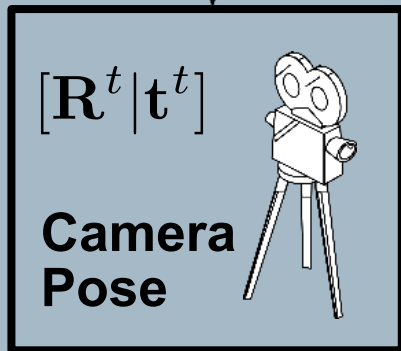
Thermal@t

Localization

sparse map
@t-1



ORB-SLAM
(localization)



Semantic Segmentation

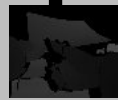


Hand
Detector



Energy
Minimization

$[R^t | t^t]$



Map Update

ORB-SLAM
(mapping)

Dense
Mapping

