

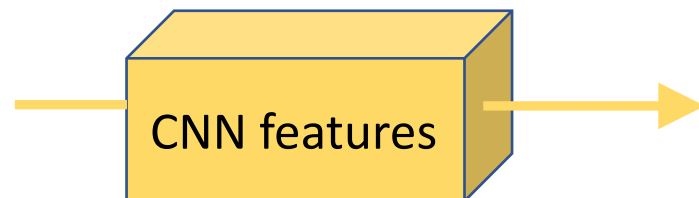
Right eye image



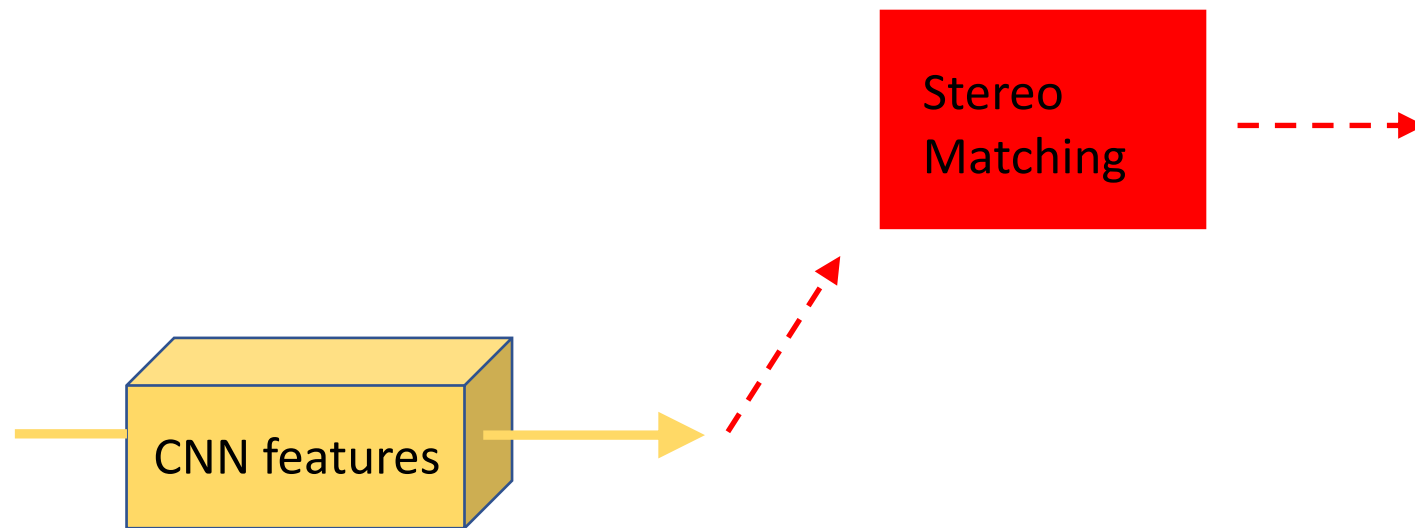
Left eye image



Step 1: CNN feature map  
extraction

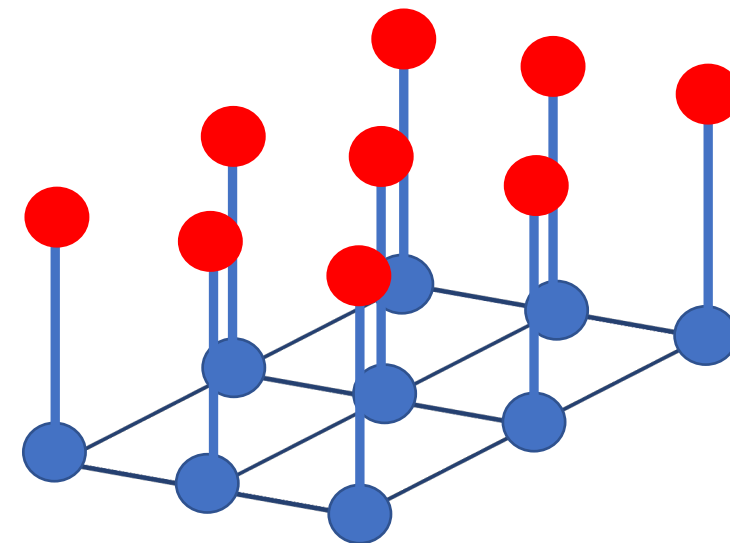


Step 2: Local stereo matching  
in CNN feature space



Step 3: Conditional random field  
smoothing

Observable variables  
e.g. local stereo disparity



Latent variables  
e.g. ground truth depth