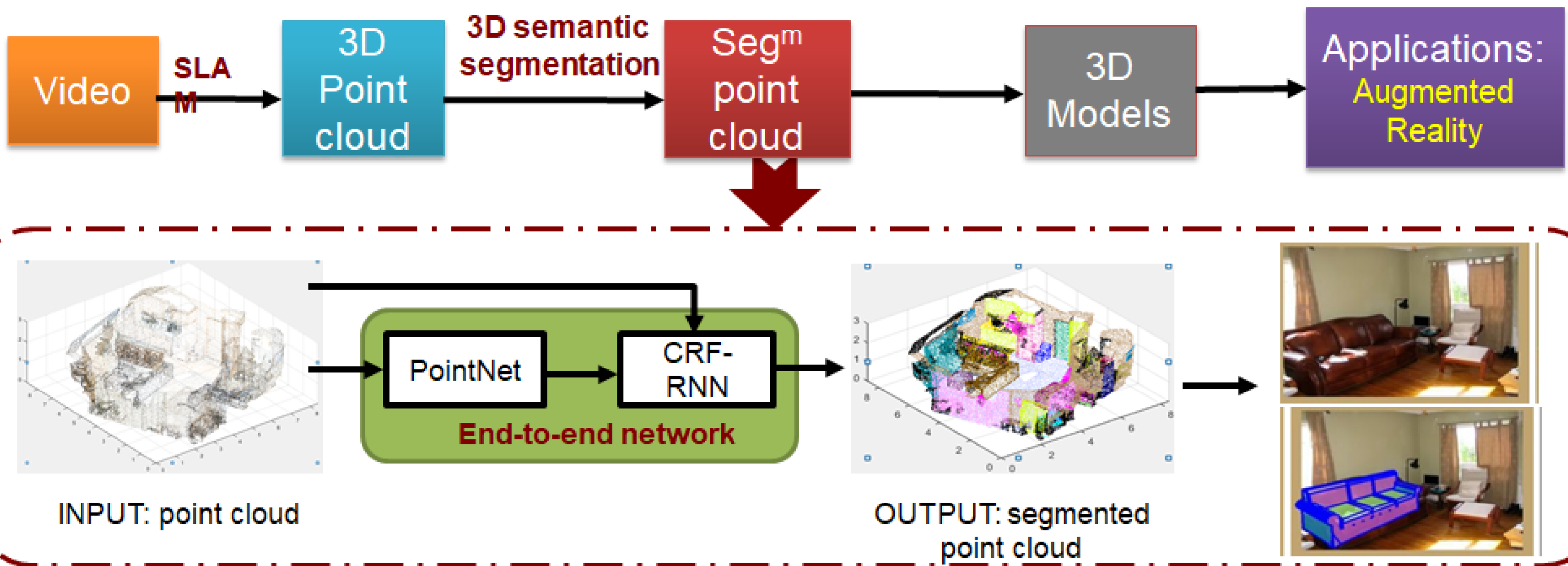


Zaw Lin Kyaw, Tat-Seng Chua

kyawzl@comp.nus.edu.sg , chuats@comp.nus.edu.sg

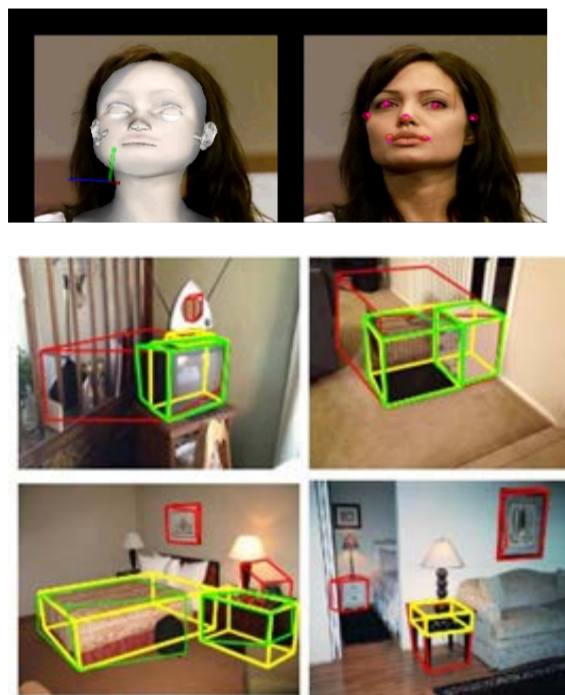
- Our world is inherently three dimensional, but there is a insufficient exploitation of 3d information in video analysis application
- We tackle 3d pose and shape generation in this work. This is a work in progress.

## 3d Extraction Pipeline



## Existing Approaches

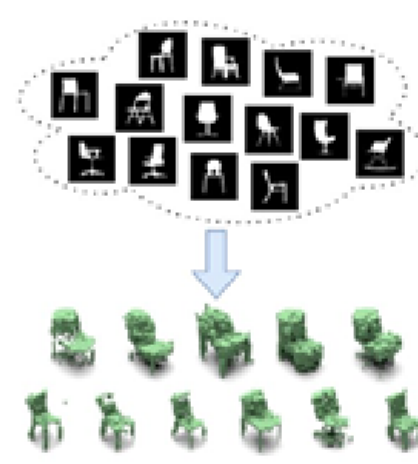
### Pose



After classical POSIT algorithm a number of works has achieved good results on 3d pose estimation using non-deep learning methods. However, they require a known model and model cannot be adapted after fitting.



### 3d Shape Generation



Most successful methods use occupancy map as representation and apply volumetric GAN to achieve shape generation. However they can only generate static poses.

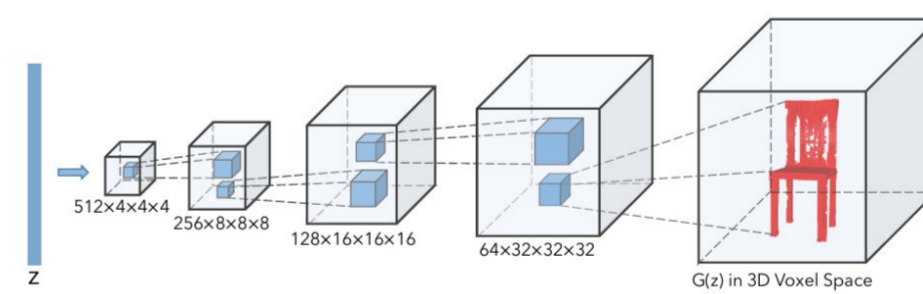
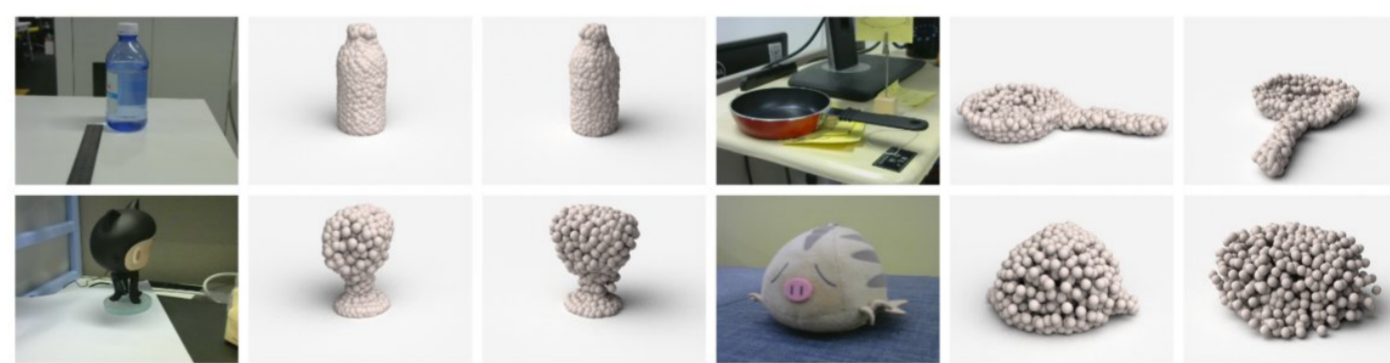


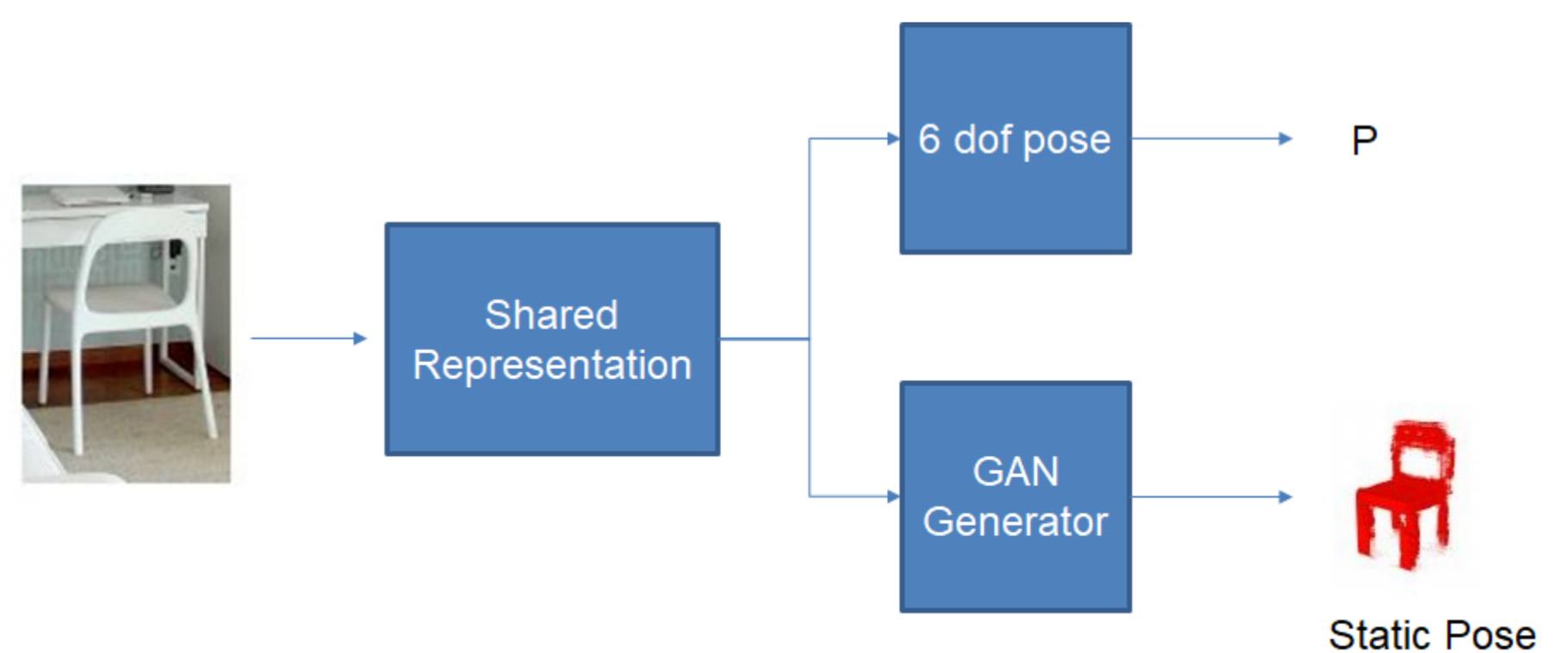
Figure 1: The generator in 3D-GAN. The discriminator mostly mirrors the generator.



## Proposed Approach

- The two inter-related problems are treated independently in literature. We aim to model them jointly.

- Our hypothesis is that shape information and pose information complement each other and results in better generated models which fit a given image or video more precisely.



- Apply  $P$  on generated pose to obtain final pose fitted onto the image