

### Video Capture

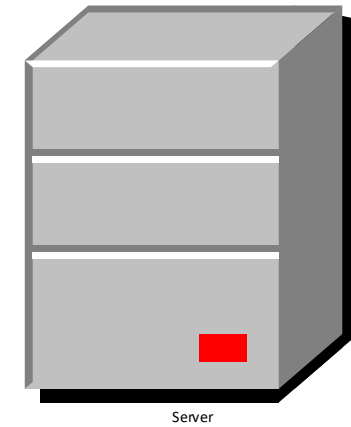
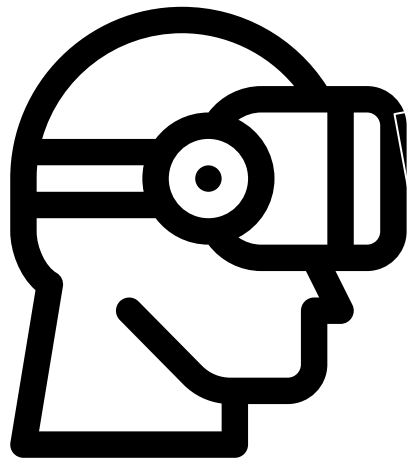
- Videos are captured by the cameras on AR glasses or mobile devices.
- Captured videos are sent to the server for further processing.

### Video Object Segmentation

- Use background subtraction to generate the trimap.
- Use Radial Basis Functions to model color similarity.
- Use superpixels and graph cuts to finalize the binary mask.

### Pose Estimation

- Model-based
- Use silhouettes rather than features to estimate poses
- Real-time



Server

### Synchronization

- The server is never blocked by any client.
- Enable synchronization among multiple clients.
- Track the views of different clients.

### Remote Rendering

- Offload complex rendering tasks to the server.
- Low-fidelity models are stored on the clients, and high-fidelity models are stored on the server.
- Only render key models to save bandwidth and reduce network latency.
- Each client has its own view.
- Use two-pass rendering approach to solve the occlusion problem.

