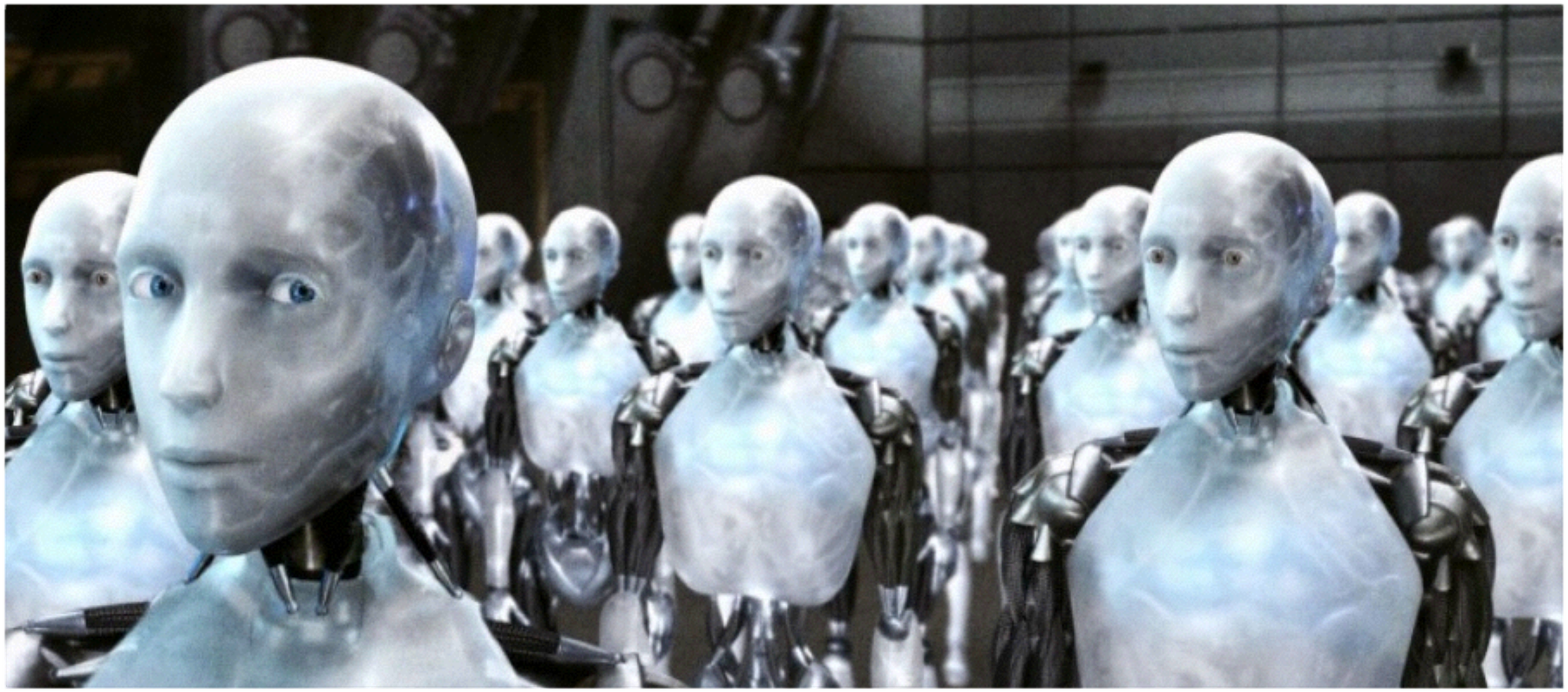


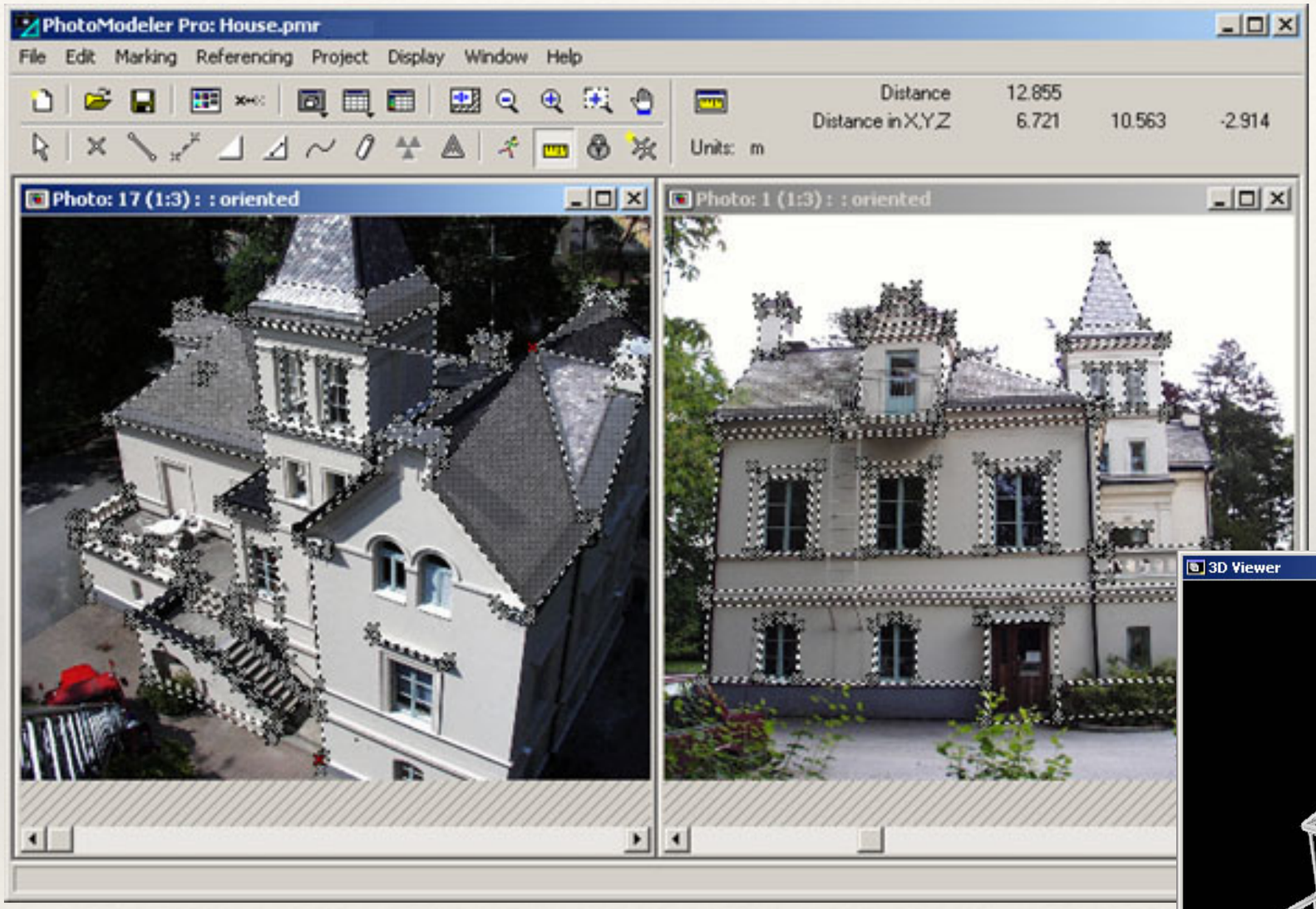
COMP3204/COMP6223: Computer Vision

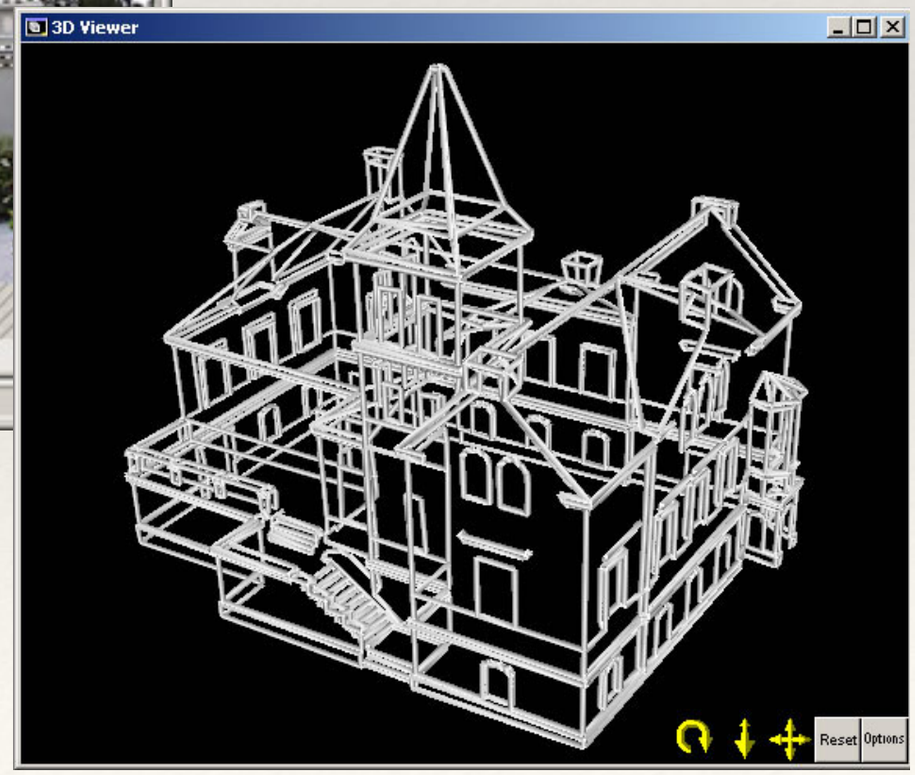
Towards 3D vision

Jonathon Hare
jsh2@ecs.soton.ac.uk

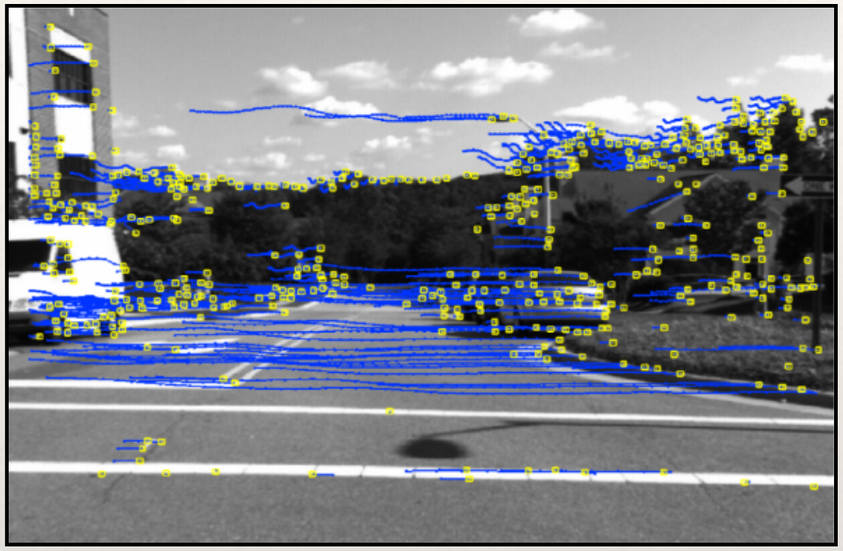
Applications

Architecture



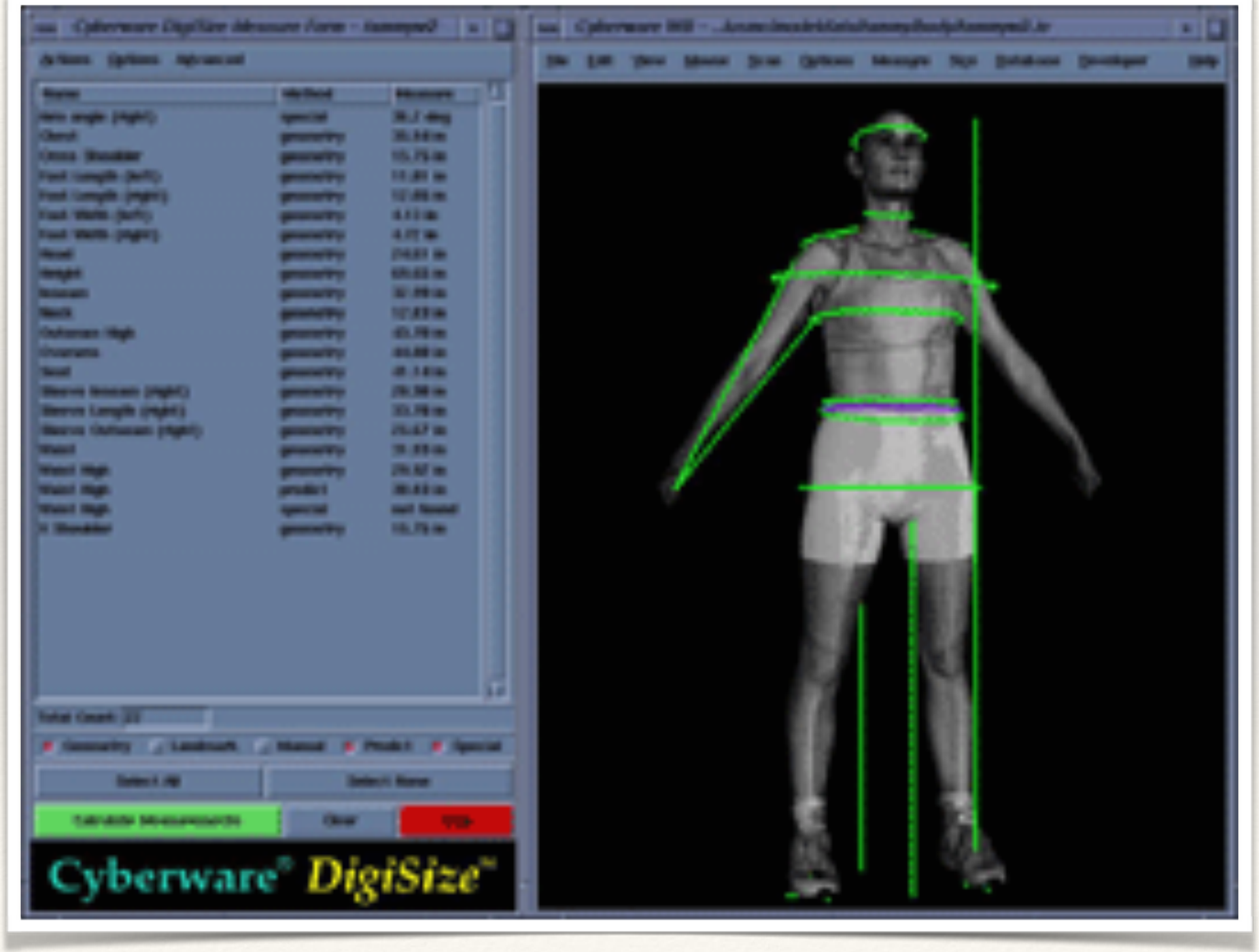


Urban Planning

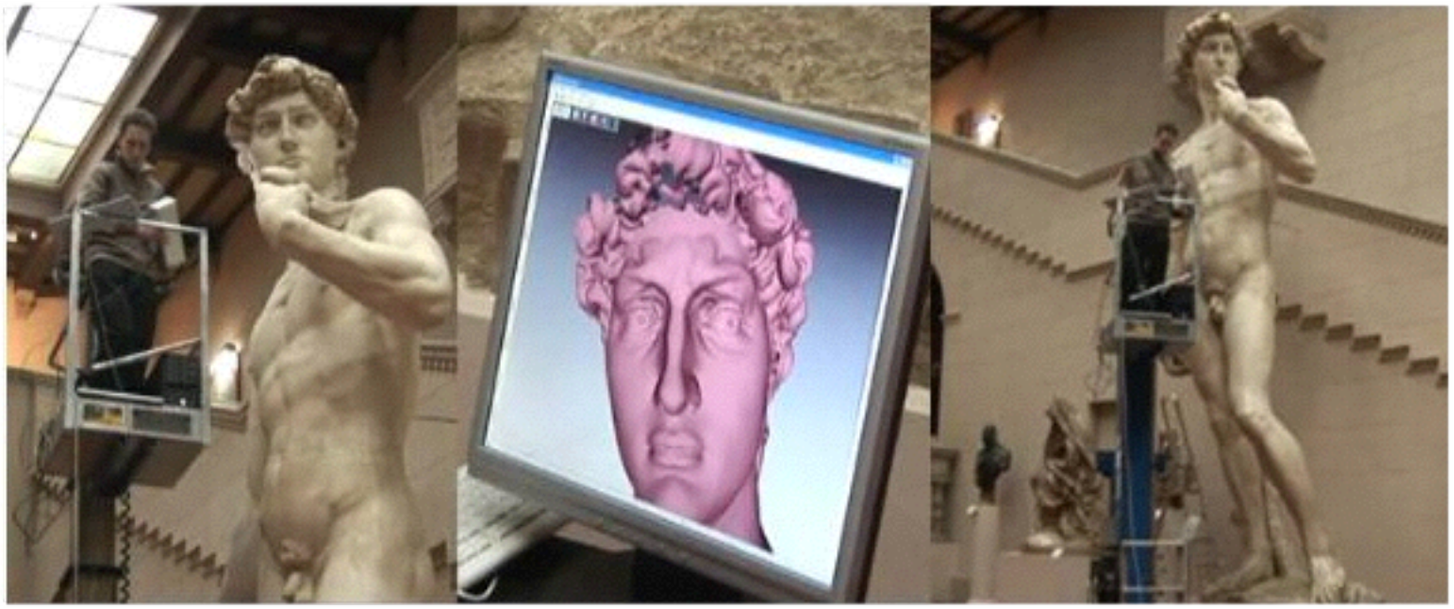


Virtual Tourism

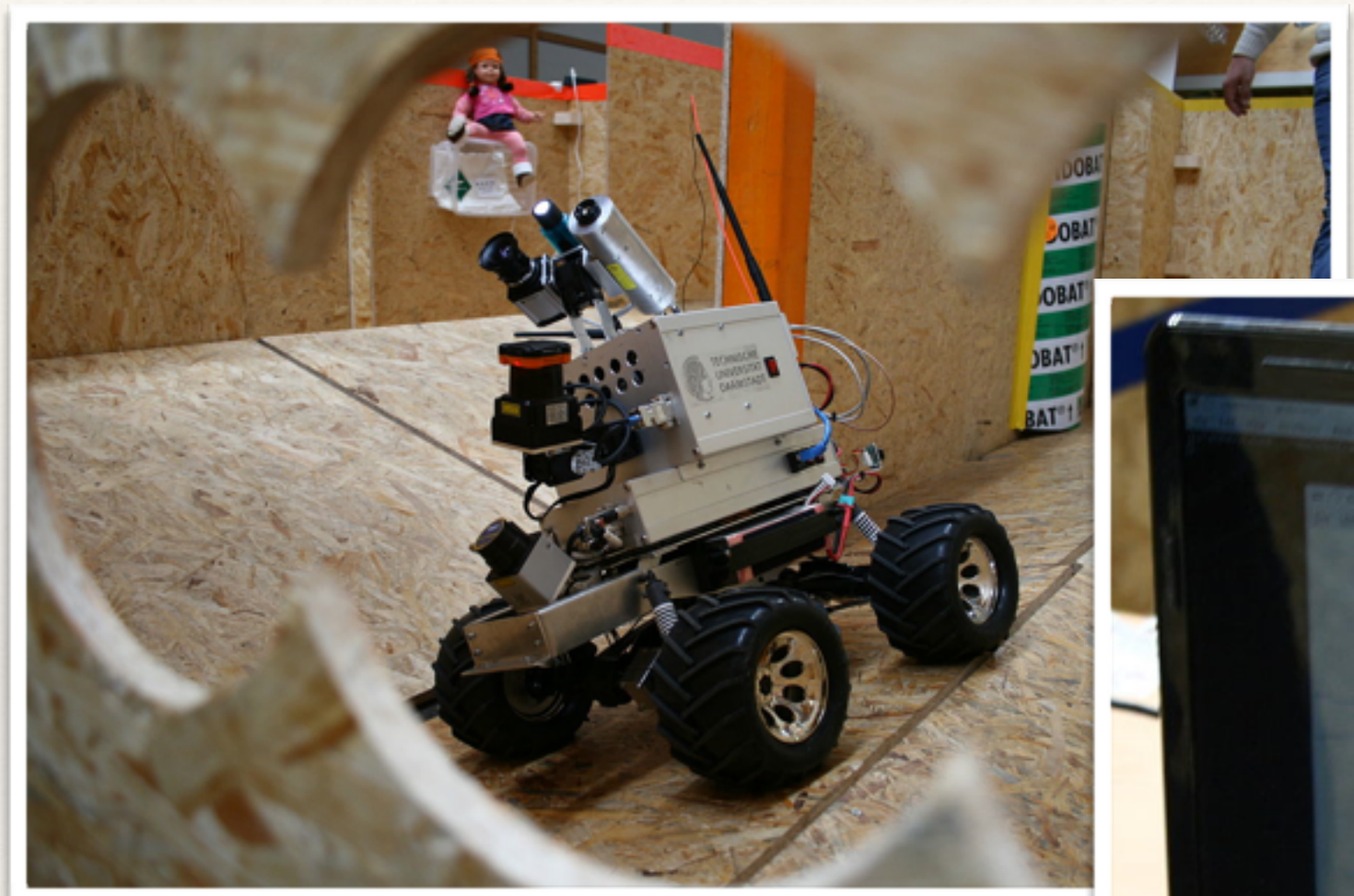
Clothing & body measurement

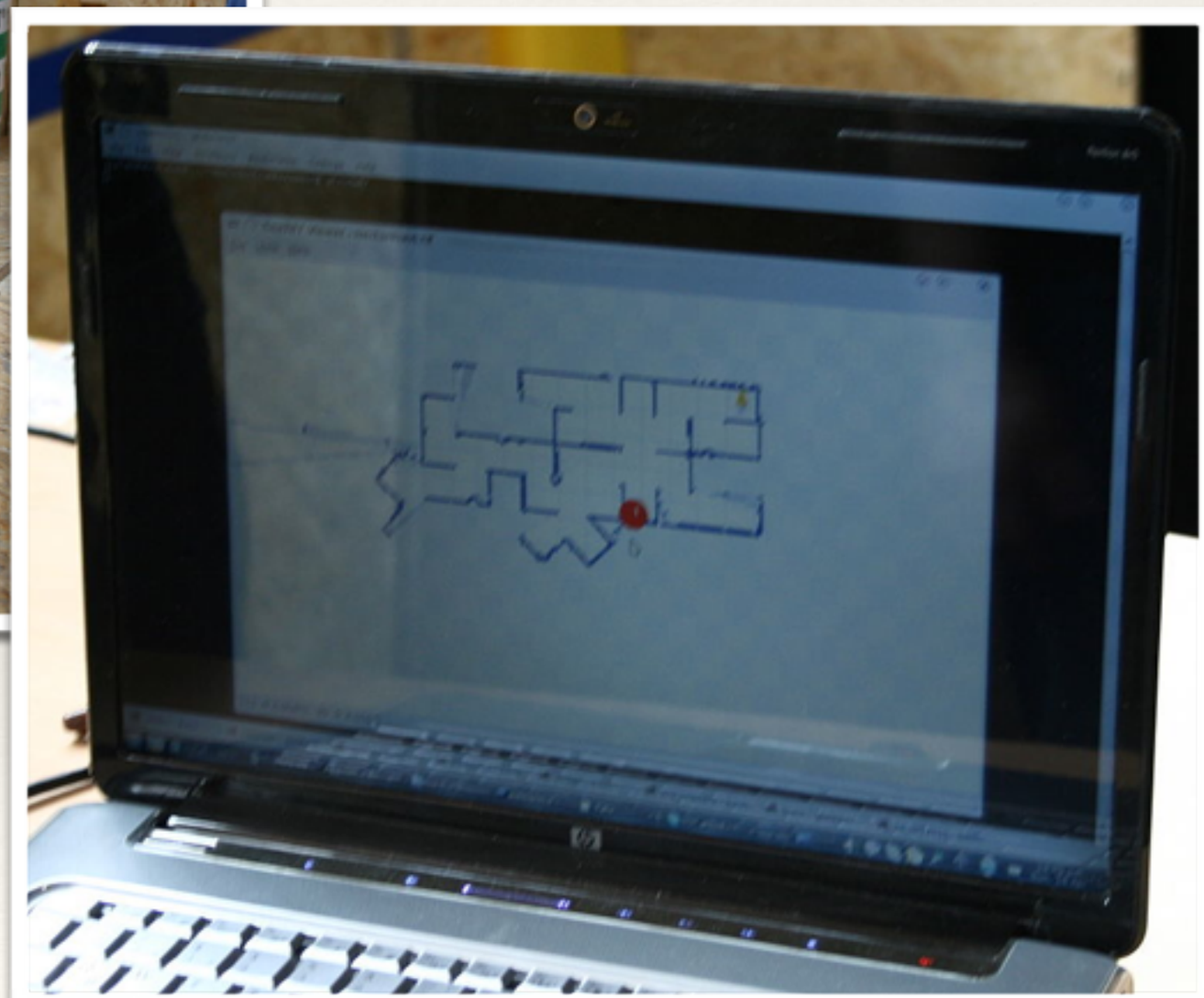


Art

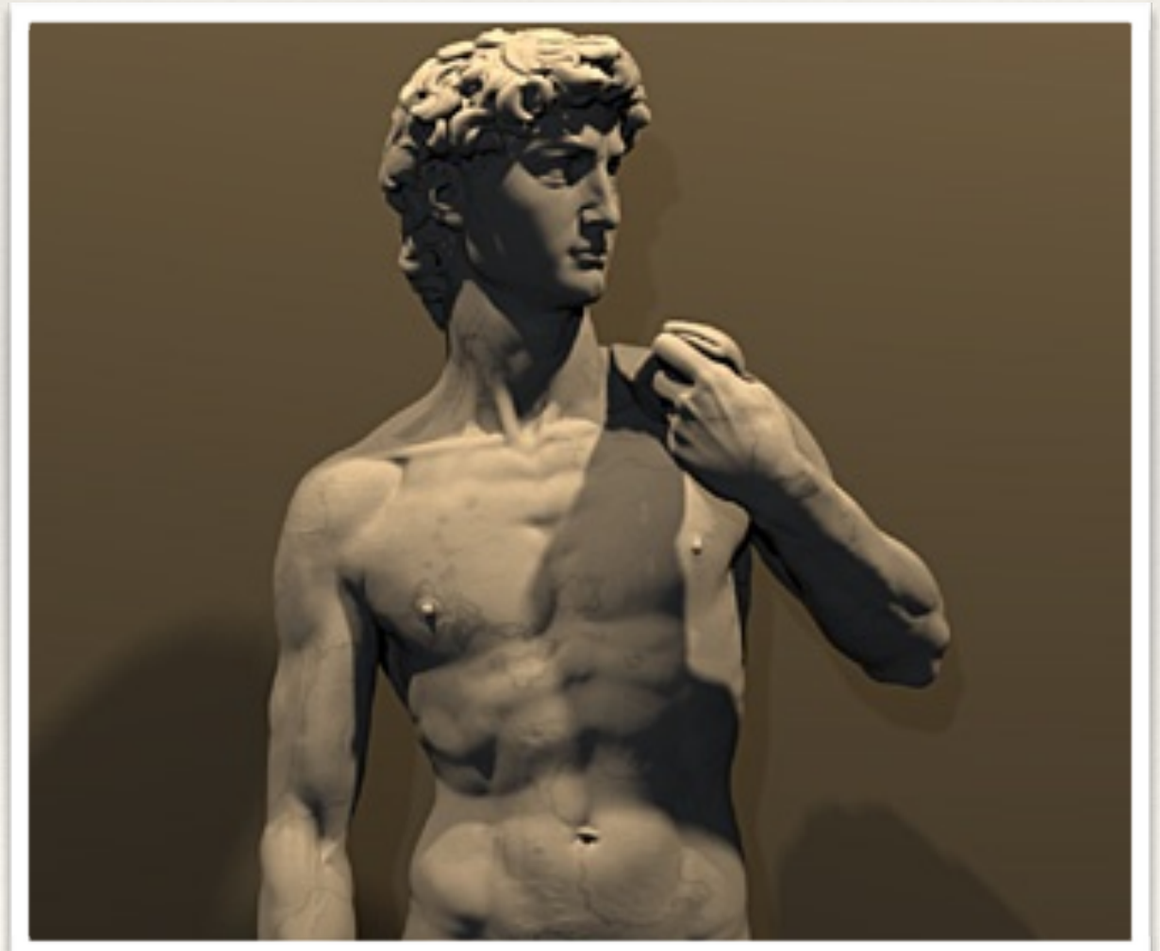


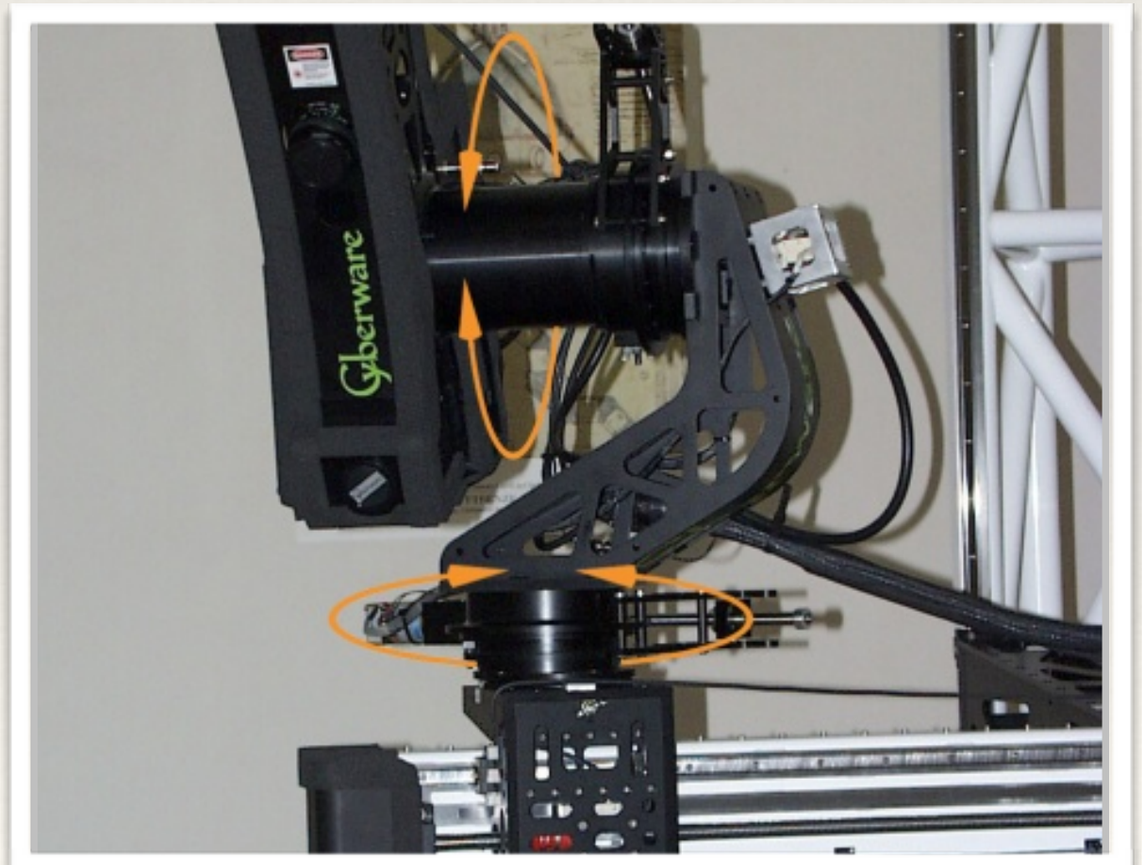
SLAM



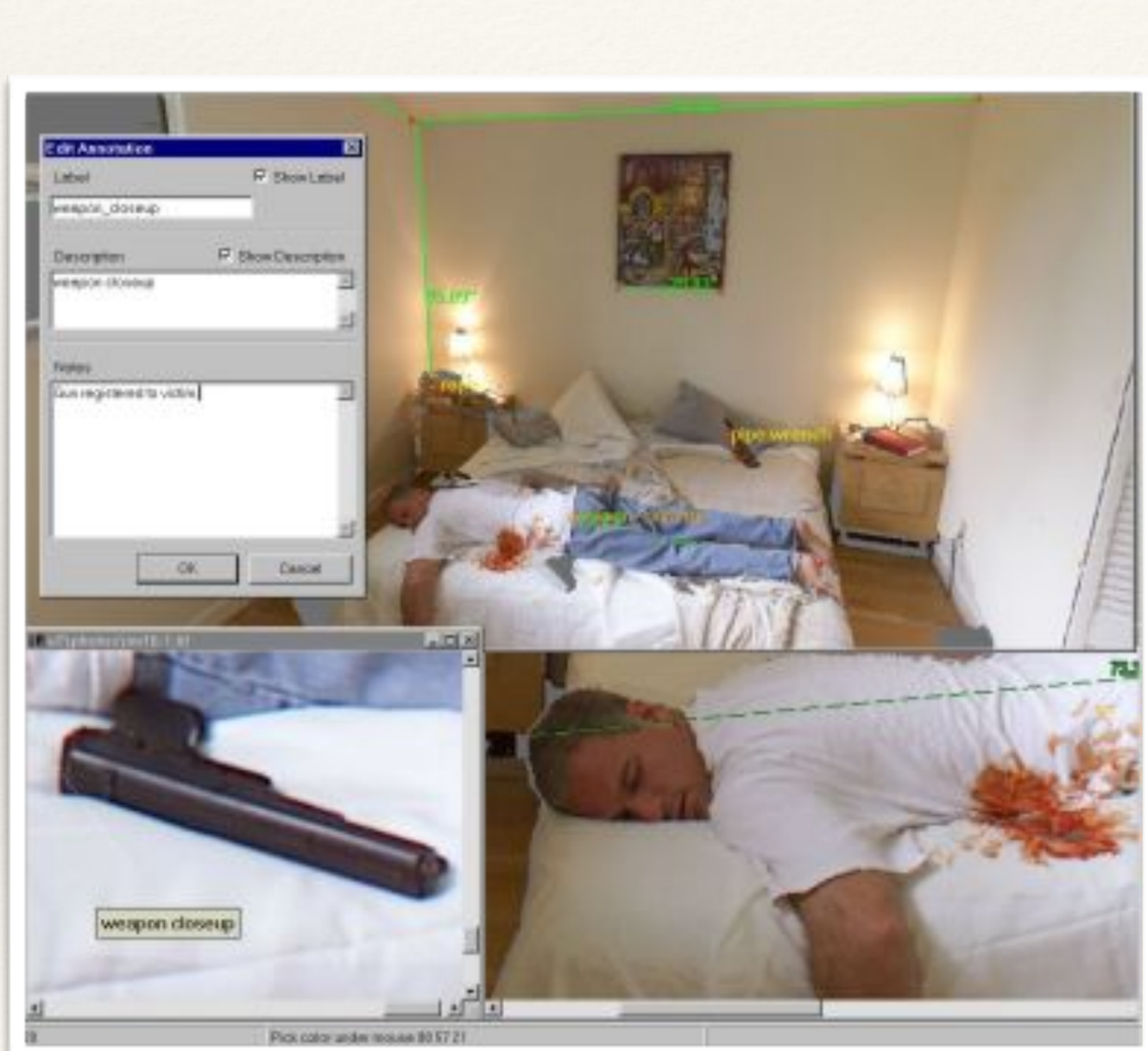


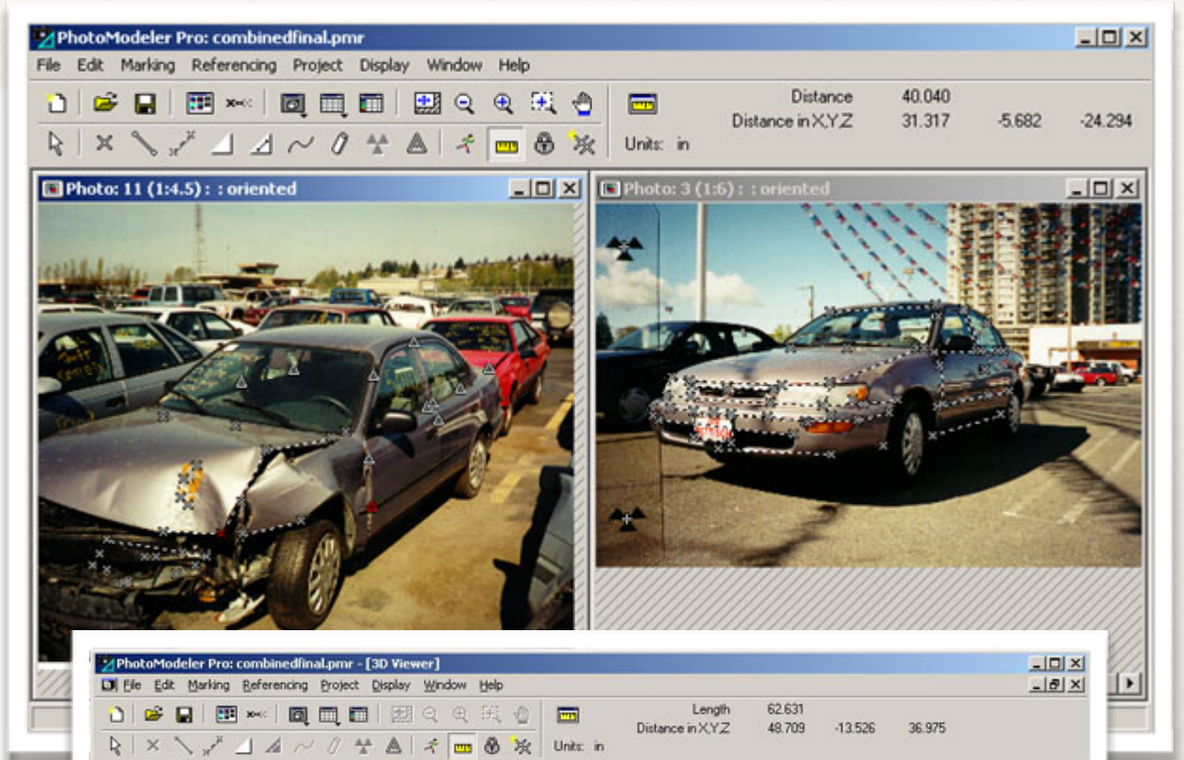
Cultural Heritage

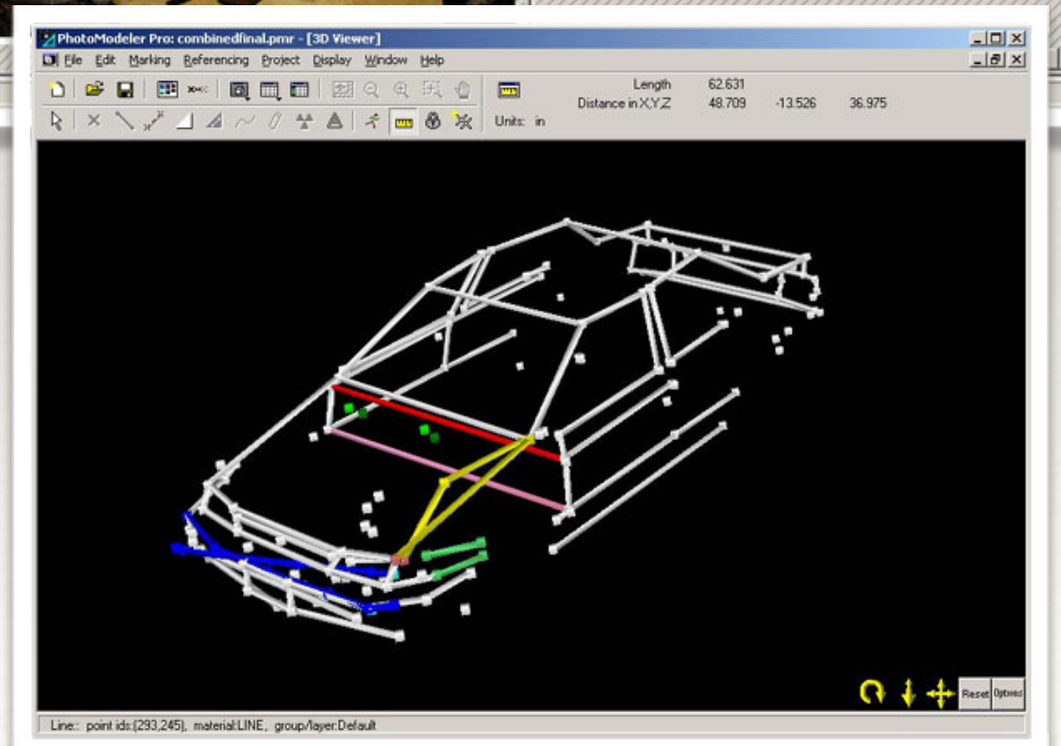




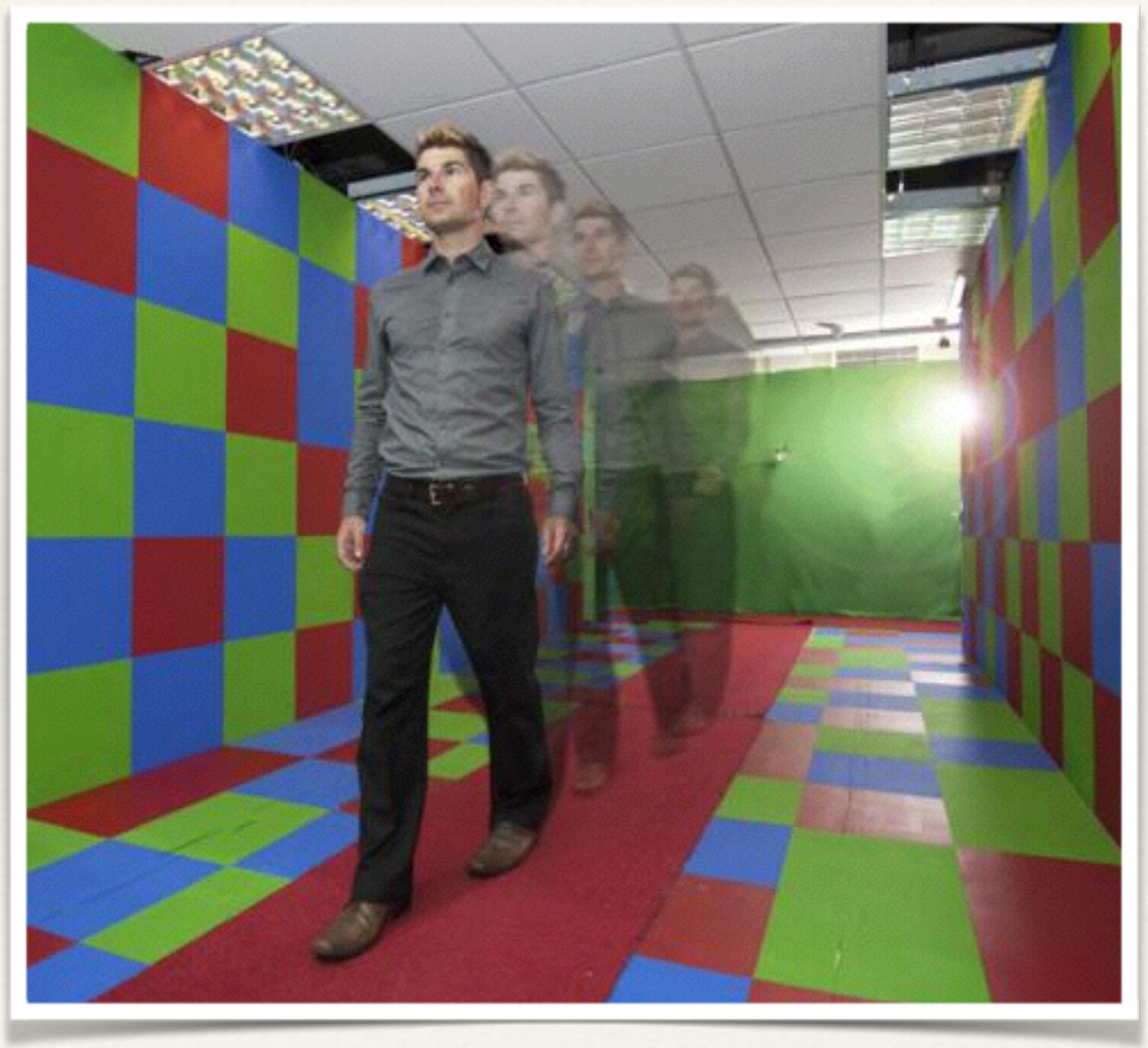
Forensics


Edit Annotation
Label
Show Label
weapon_closeup
Description
Show Description
weapon closeup
Notes
Cancel
weapon closeup

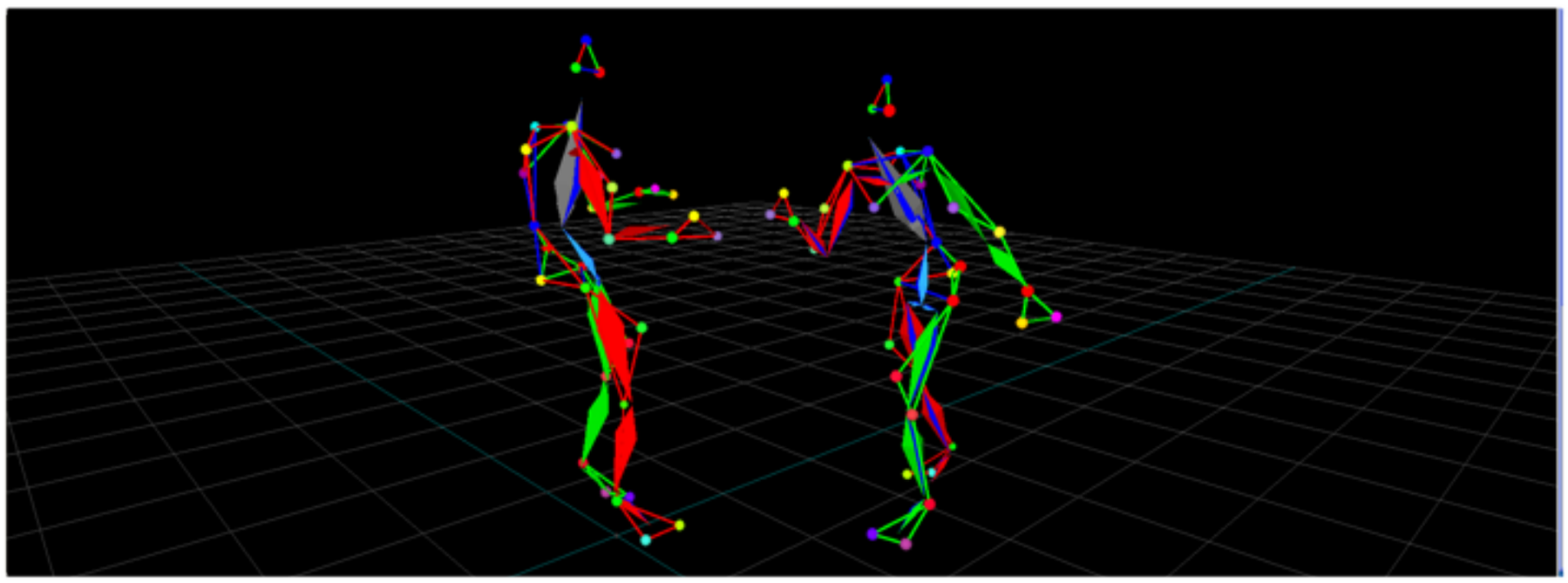

PhotoModeler Pro: combinedfinal.pmr
File Edit Marking Referencing Project Display Window Help
Distance 40.040
Distance in X,Y,Z 31.317 -5.682 -24.294
Units: in
Photo: 11 (1:4.5) : : oriented
Photo: 3 (1:6) : : oriented


PhotoModeler Pro: combinedfinal.pmr - [3D Viewer]
File Edit Marking Referencing Project Display Window Help
Length 62.631
Distance in X,Y,Z 48.709 -13.526 36.975
Units: in
Reset
Line:: point ids:(293,245), material:LINE, group/layer:Default

Surveillance

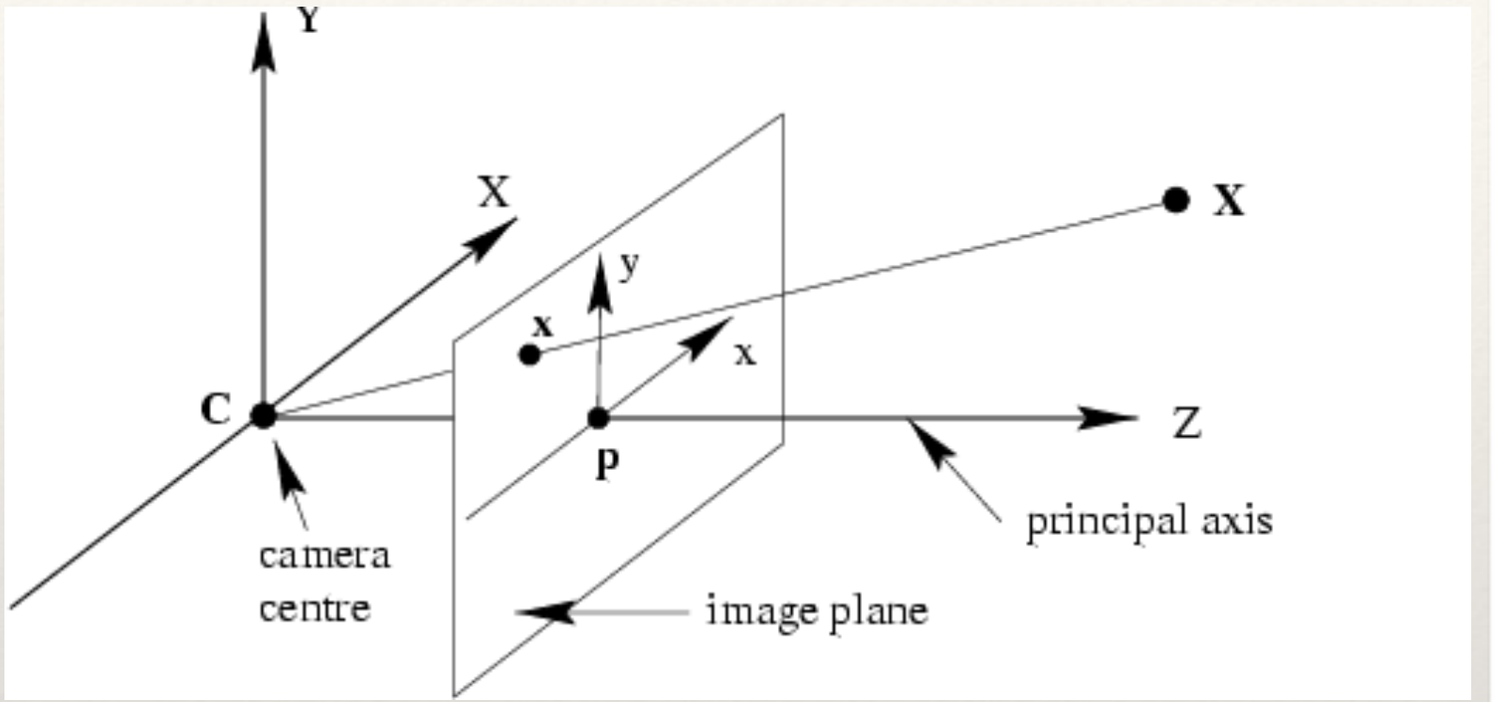


Motion Capture (Films & Games)



Cameras

Camera Geometry



$$\lambda \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & & p_x \\ & f_y & p_y \\ & & 1 \end{bmatrix} \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \end{bmatrix} \begin{bmatrix} \mathtt{R} & \mathtt{t} \\ & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

Camera Geometry

This is a point in the image

$$\lambda \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & & p_x \\ & f_y & p_y \\ & & 1 \end{bmatrix} \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \end{bmatrix} \begin{bmatrix} \mathtt{R} & \mathtt{t} \\ & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

This is a point in the world

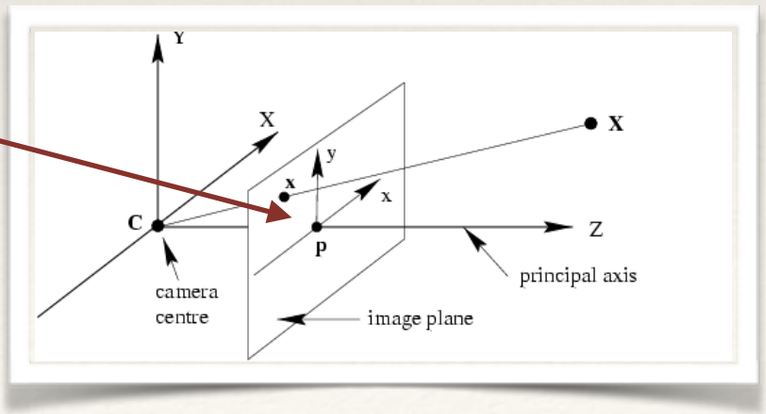
Camera Geometry

These are the "*intrinsic*" parameters

$$\lambda \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & & p_x \\ & f_y & p_y \\ & & 1 \end{bmatrix} \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \end{bmatrix} \begin{bmatrix} \mathtt{R} & \mathtt{t} \\ & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

focal length

position of the *principal point* in the image



Camera Geometry

These are the "*extrinsic*" parameters

$$\lambda \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & & p_x \\ & f_y & p_y \\ & & 1 \end{bmatrix} \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \end{bmatrix} \begin{bmatrix} \mathtt{R} & \mathtt{t} \\ & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

Rotation of the camera in world space

Translation of the camera in world space

Camera Calibration

$$\lambda \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & & p_x \\ & f_y & p_y \\ & & 1 \end{bmatrix} \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \end{bmatrix} \begin{bmatrix} \mathtt{R} & \mathtt{t} \\ & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

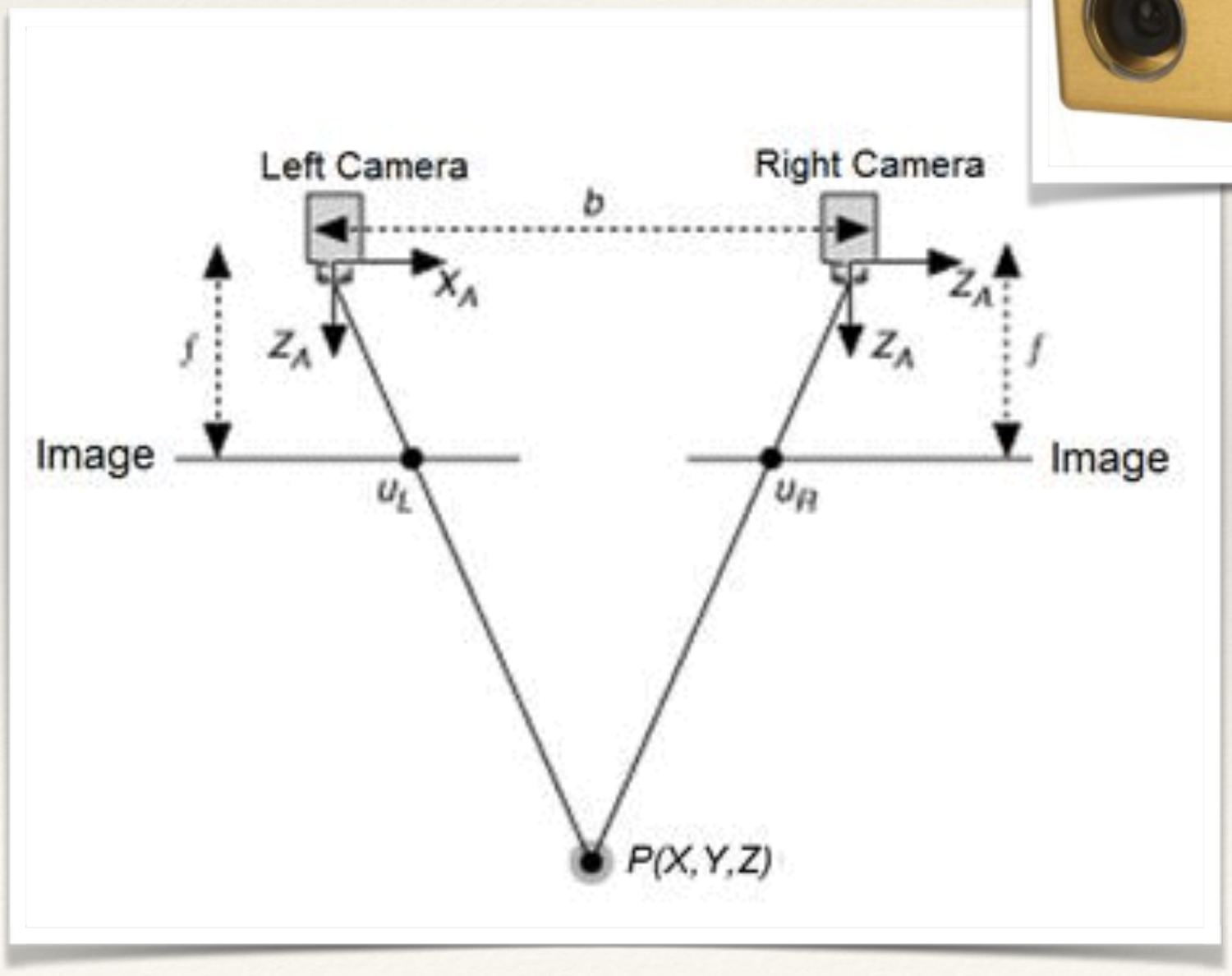
- Camera calibration is the process of estimating the intrinsic parameters of a camera
  - Also deals with learning non-linear radial distortion parameters of real camera lenses
  - Typically determined by solving sets of point correspondences from images of “calibration patterns”

Camera Calibration Demo

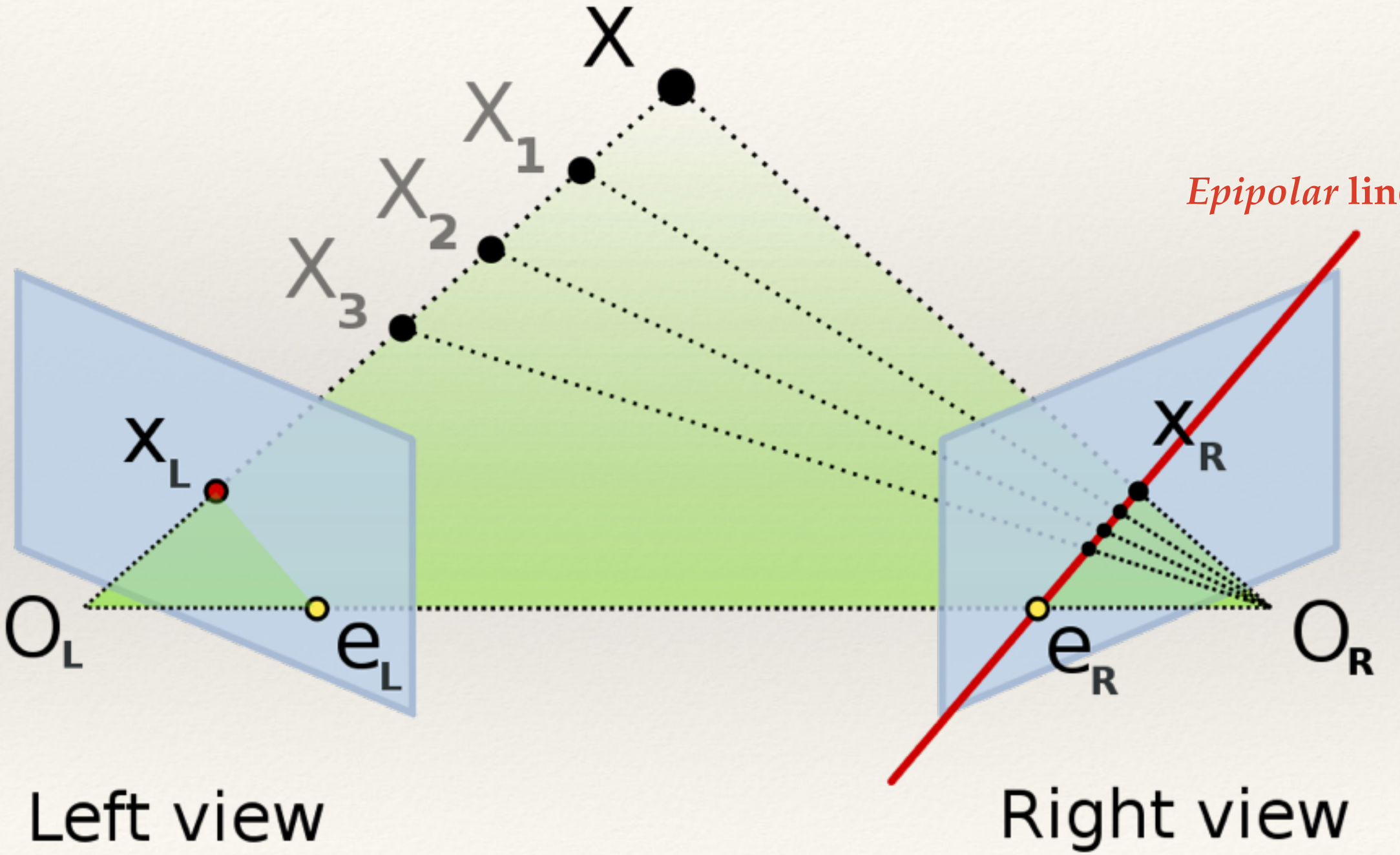
Measuring Depth

Narrow Baseline Stereo

Stereo Camera



Epipolar geometry



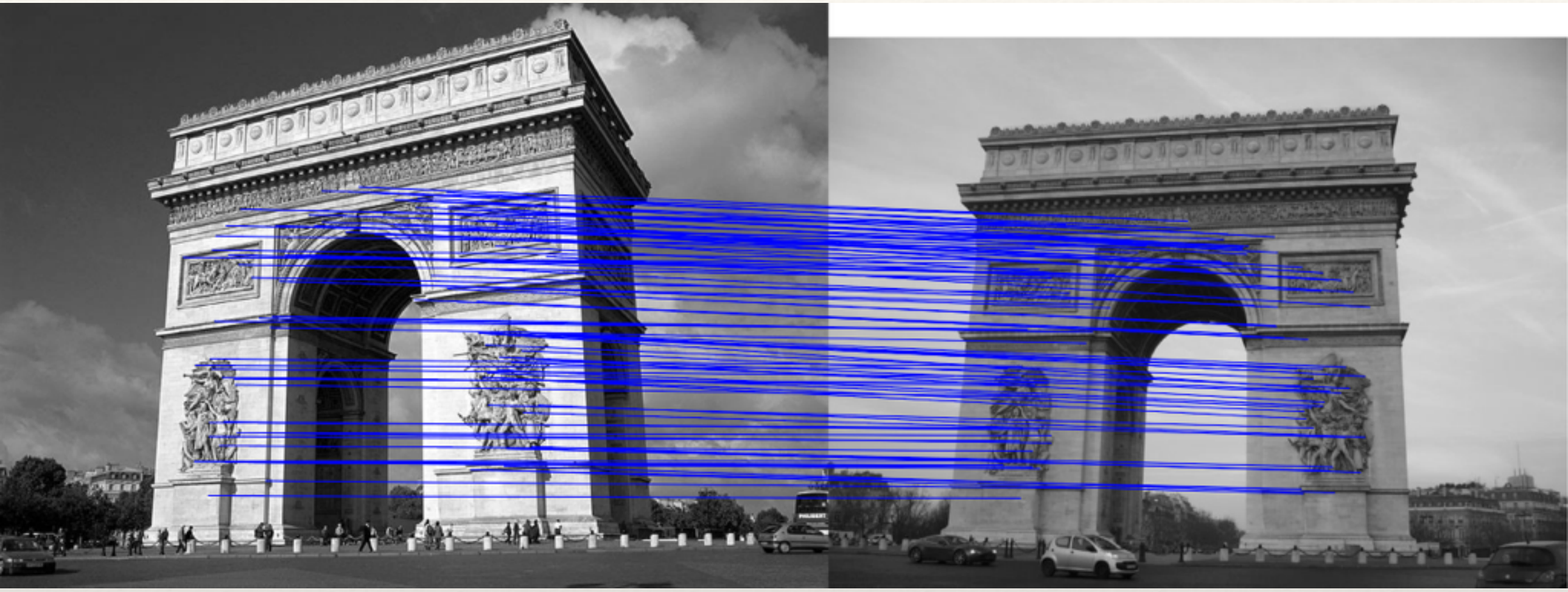
Dense narrow-baseline stereo

Warp images to simplify epipolar geometry

Compute disparity map by matching pixels along the epipolar lines

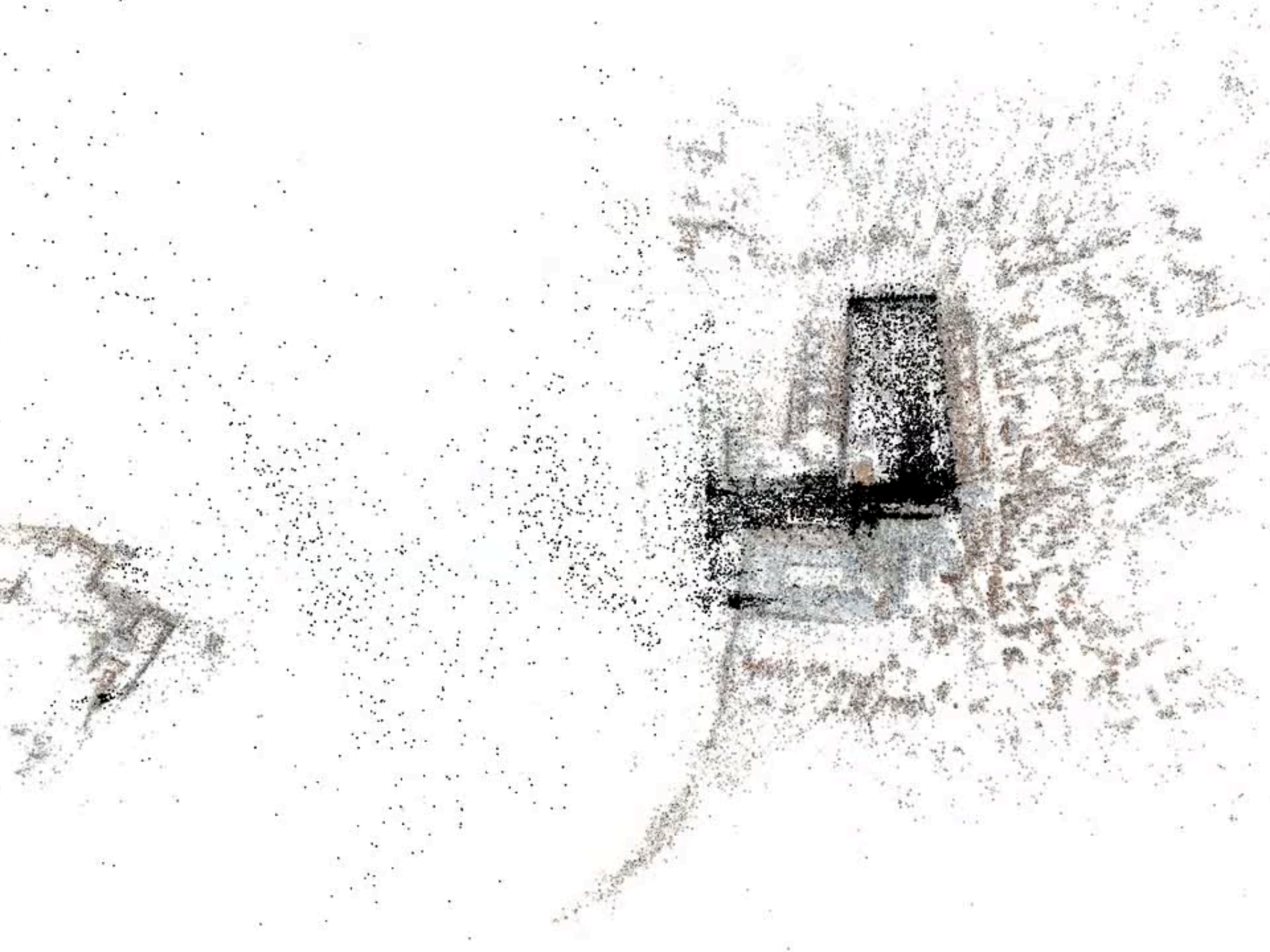
Wide Baseline Stereo

Multiple images can be used to jointly infer 3D structure, and the camera pose and intrinsics of each camera



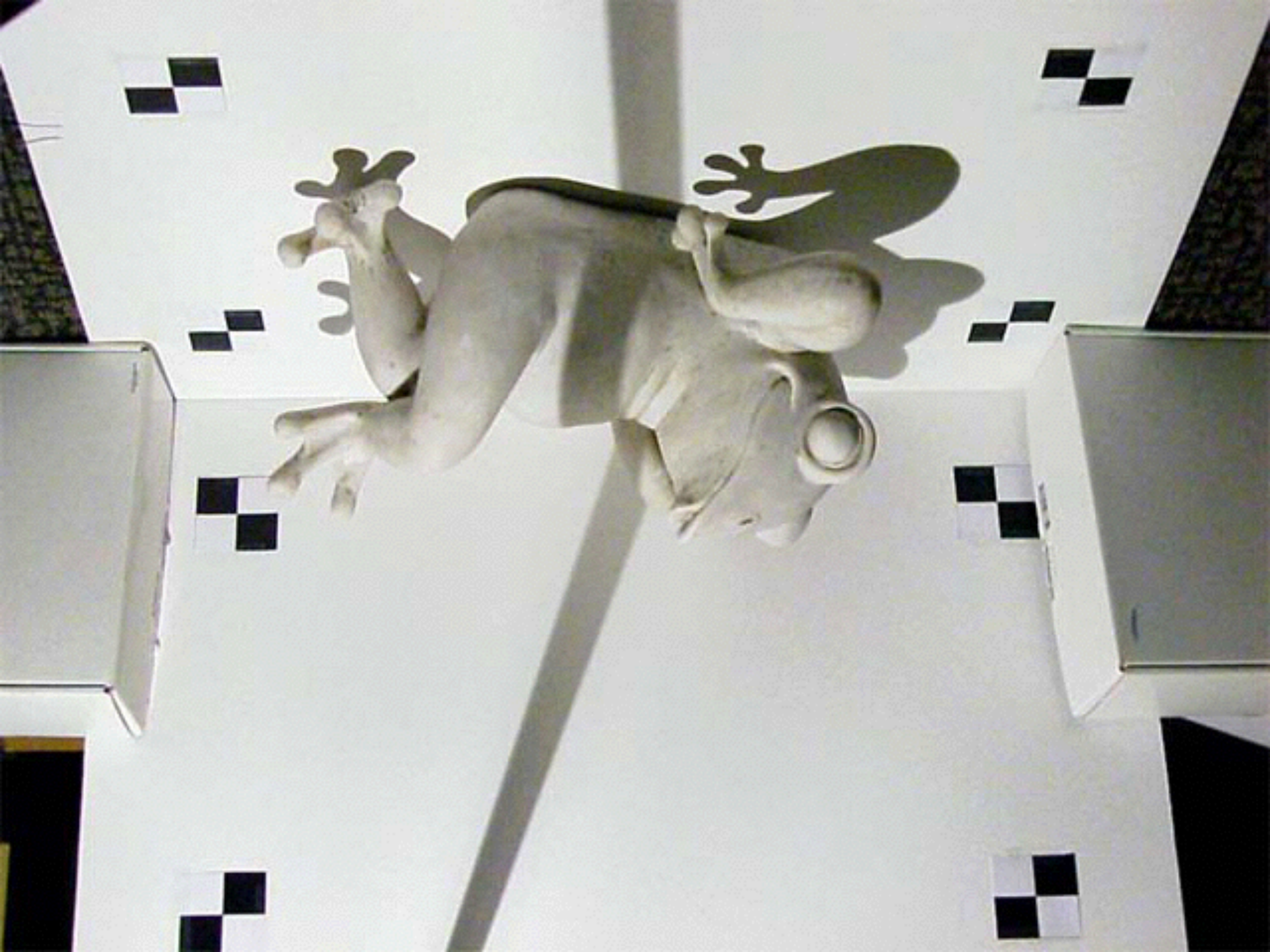
Point matches (i.e. SIFT) are used as the basis for triangulating 3D points from the 2D images

Reconstructing Venice

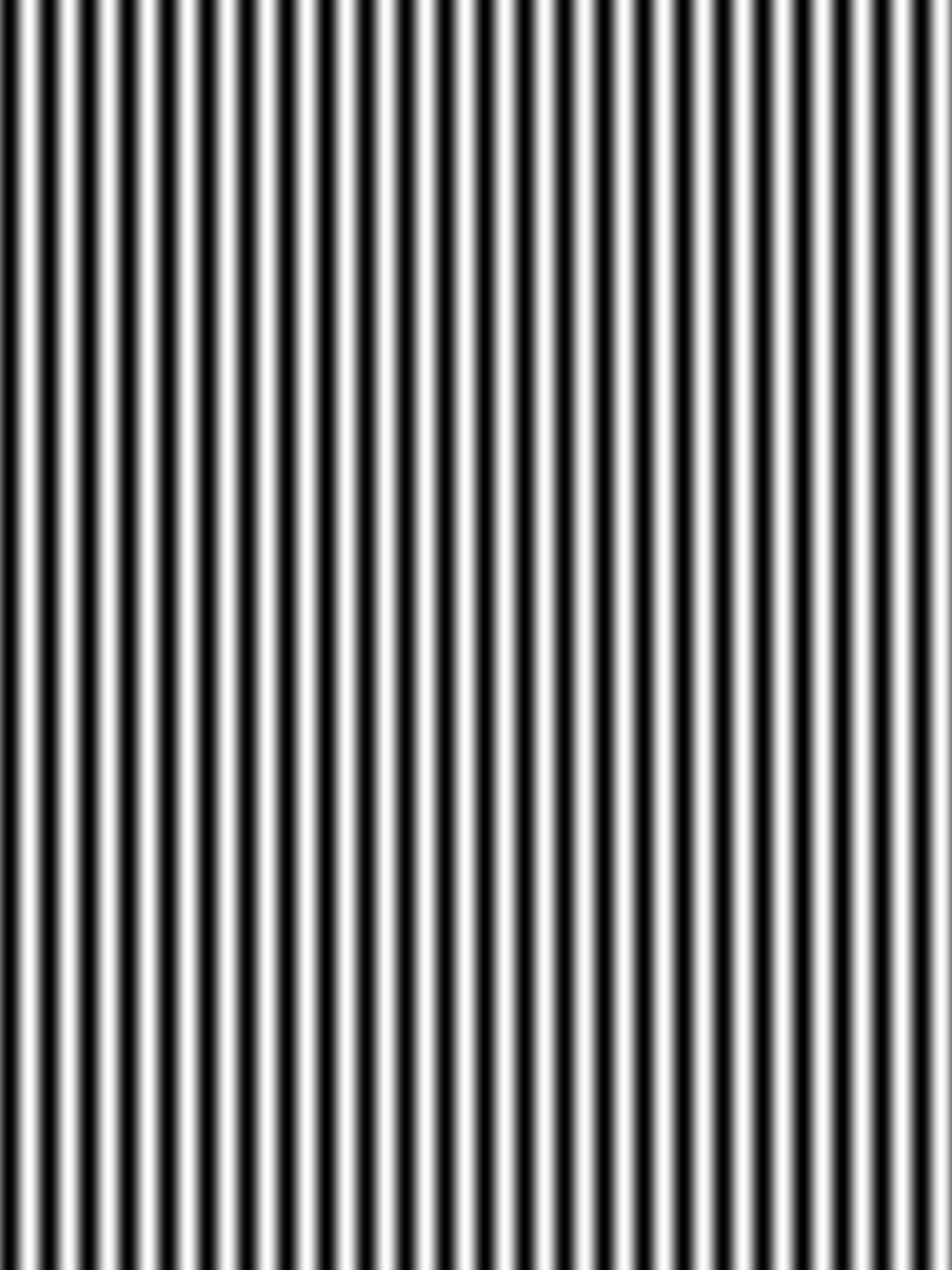


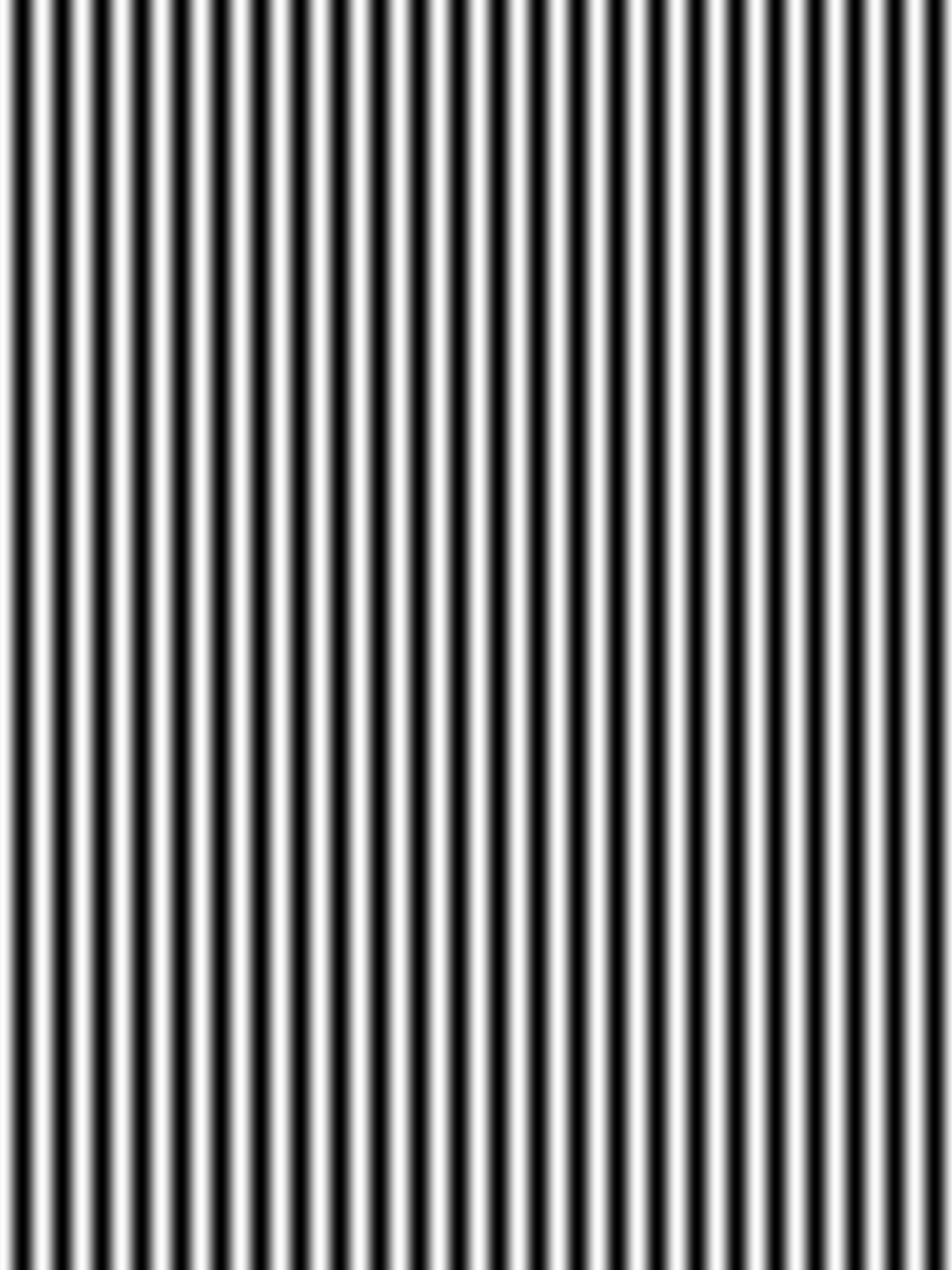
Monocular Vision

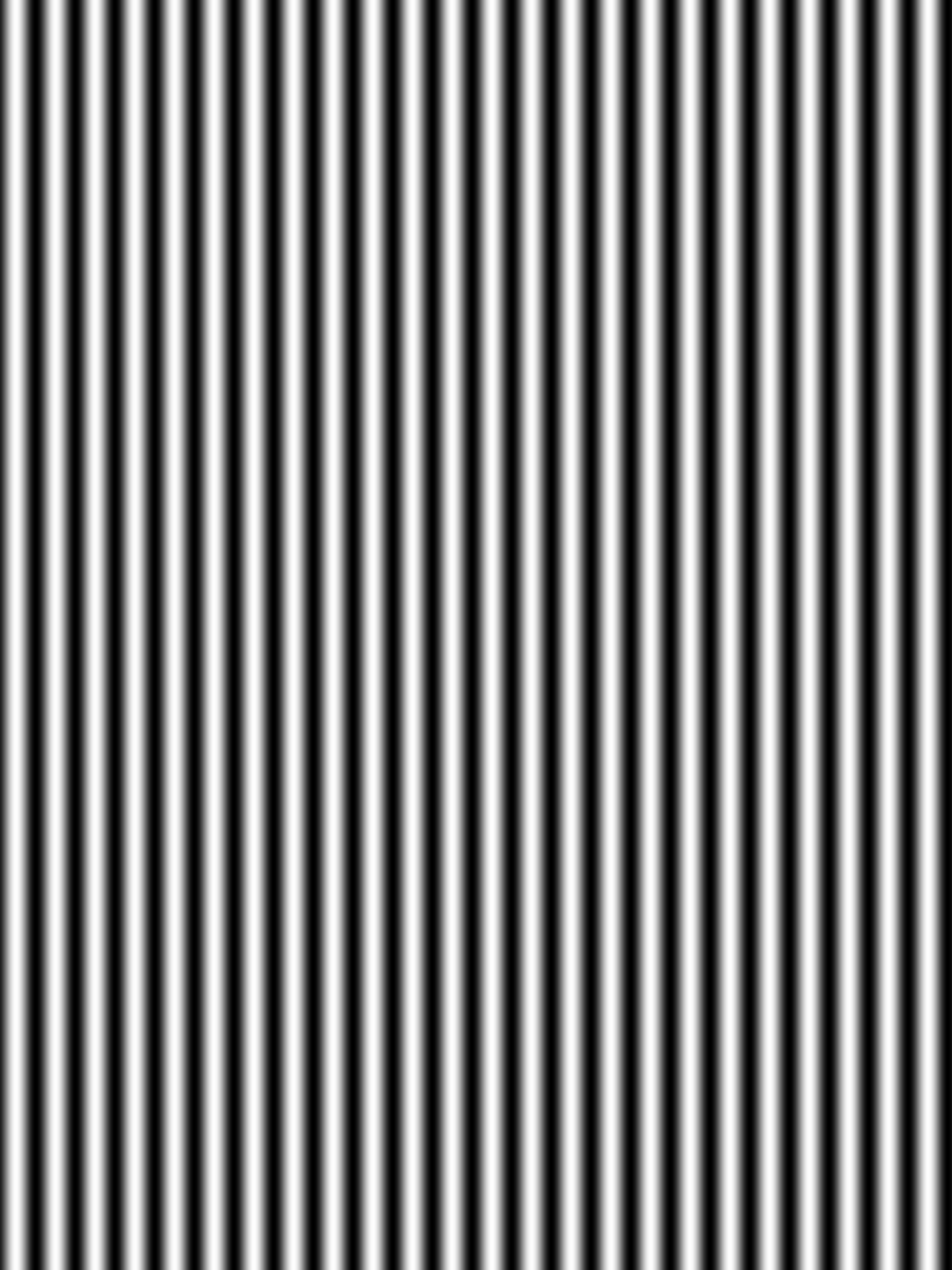
Shadow Scanner

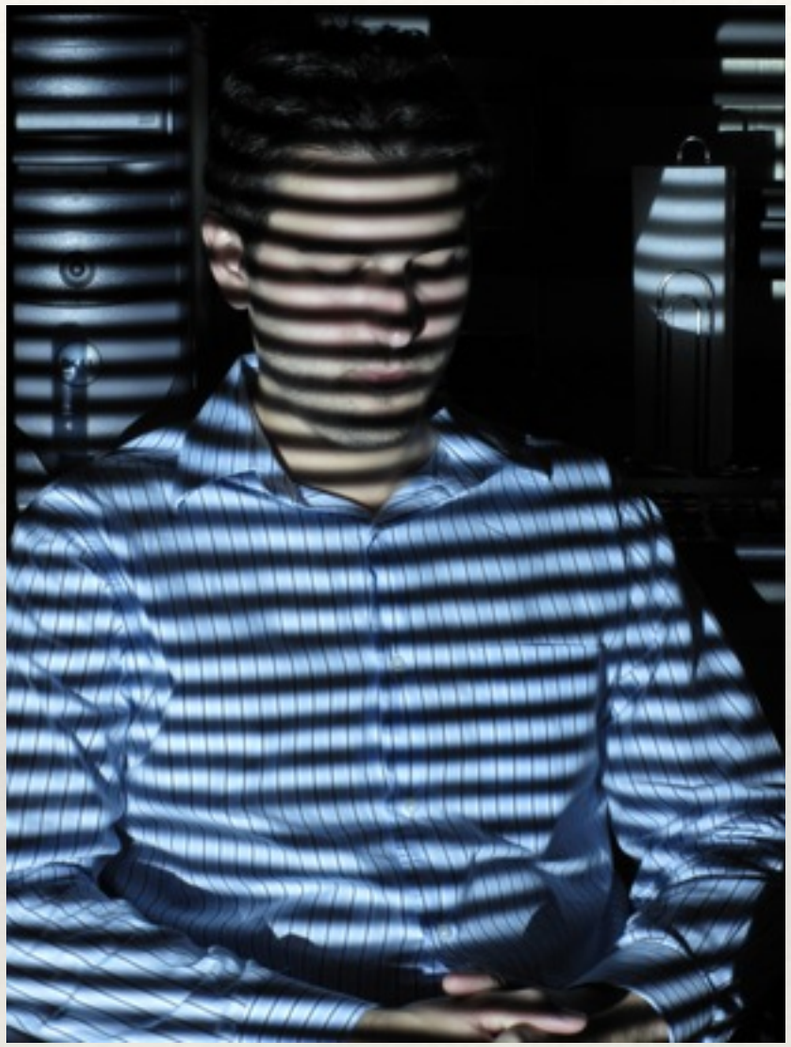
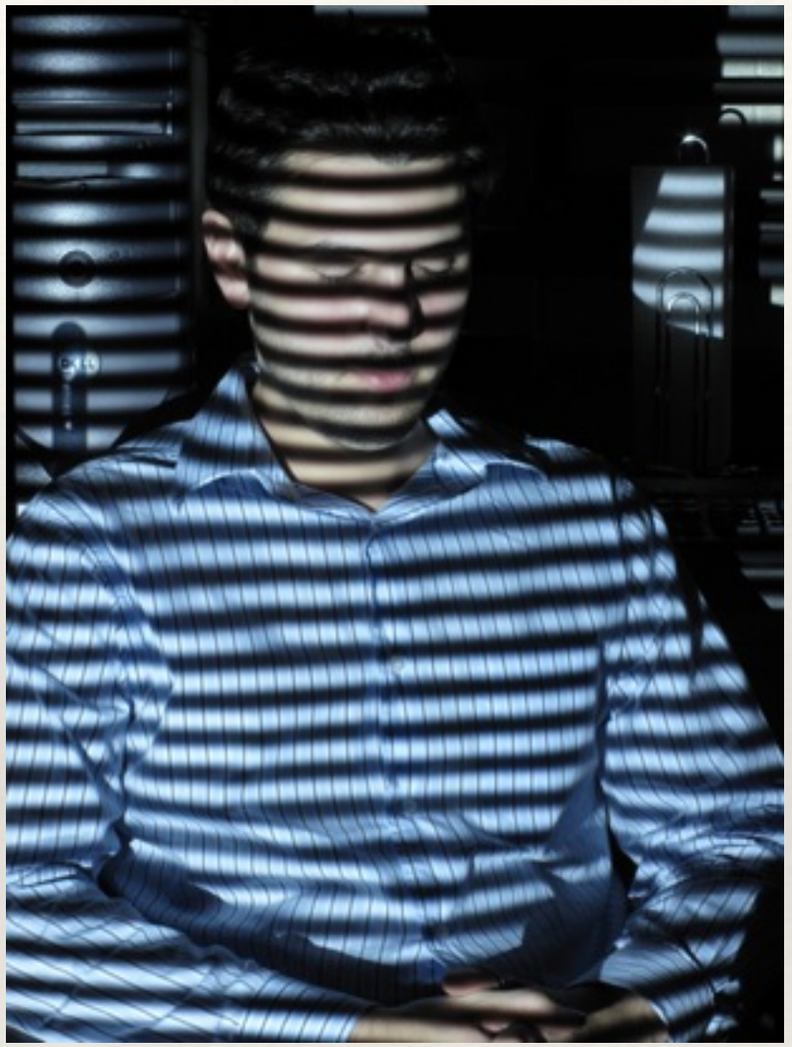


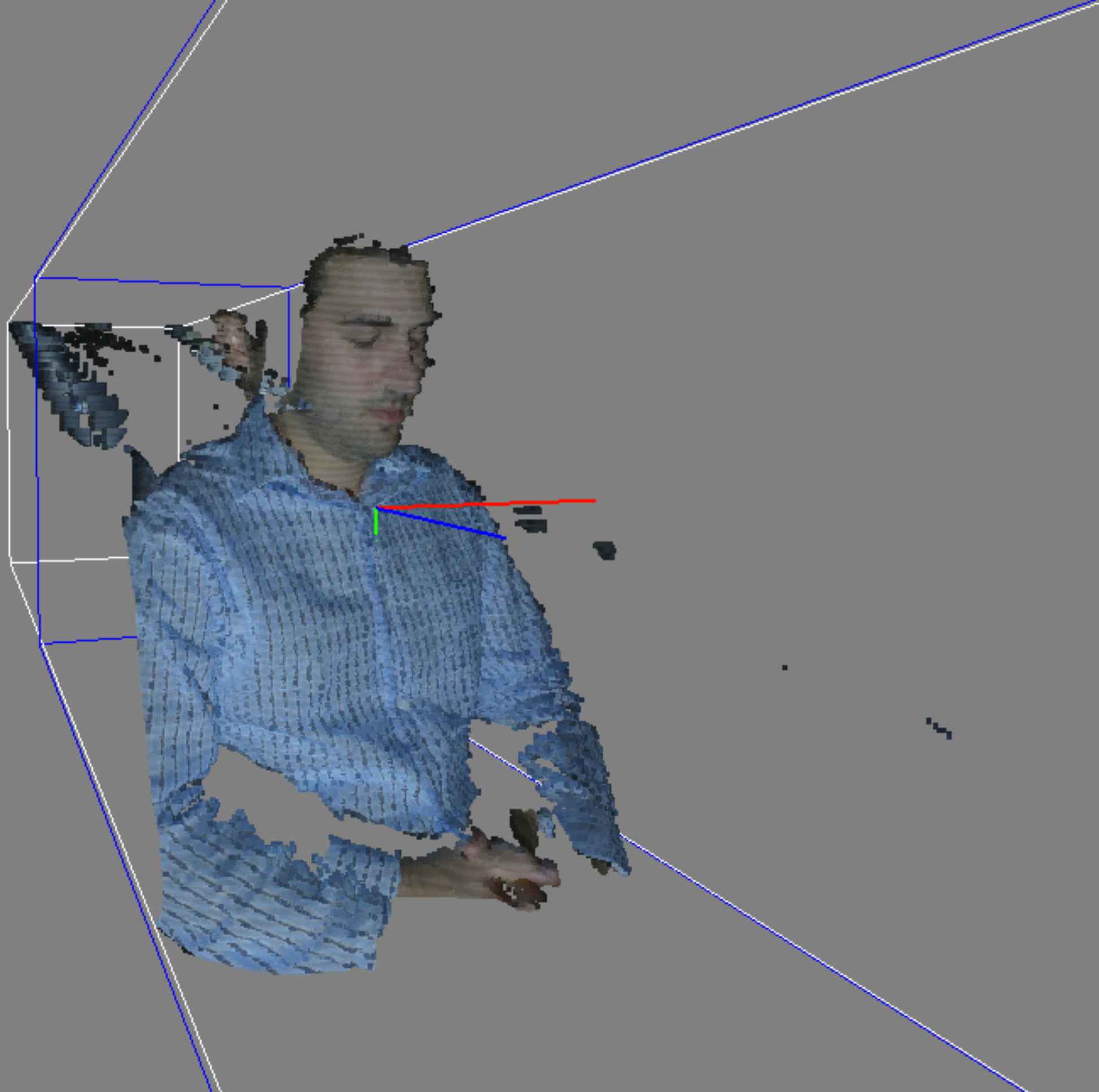
Structured Light Imaging

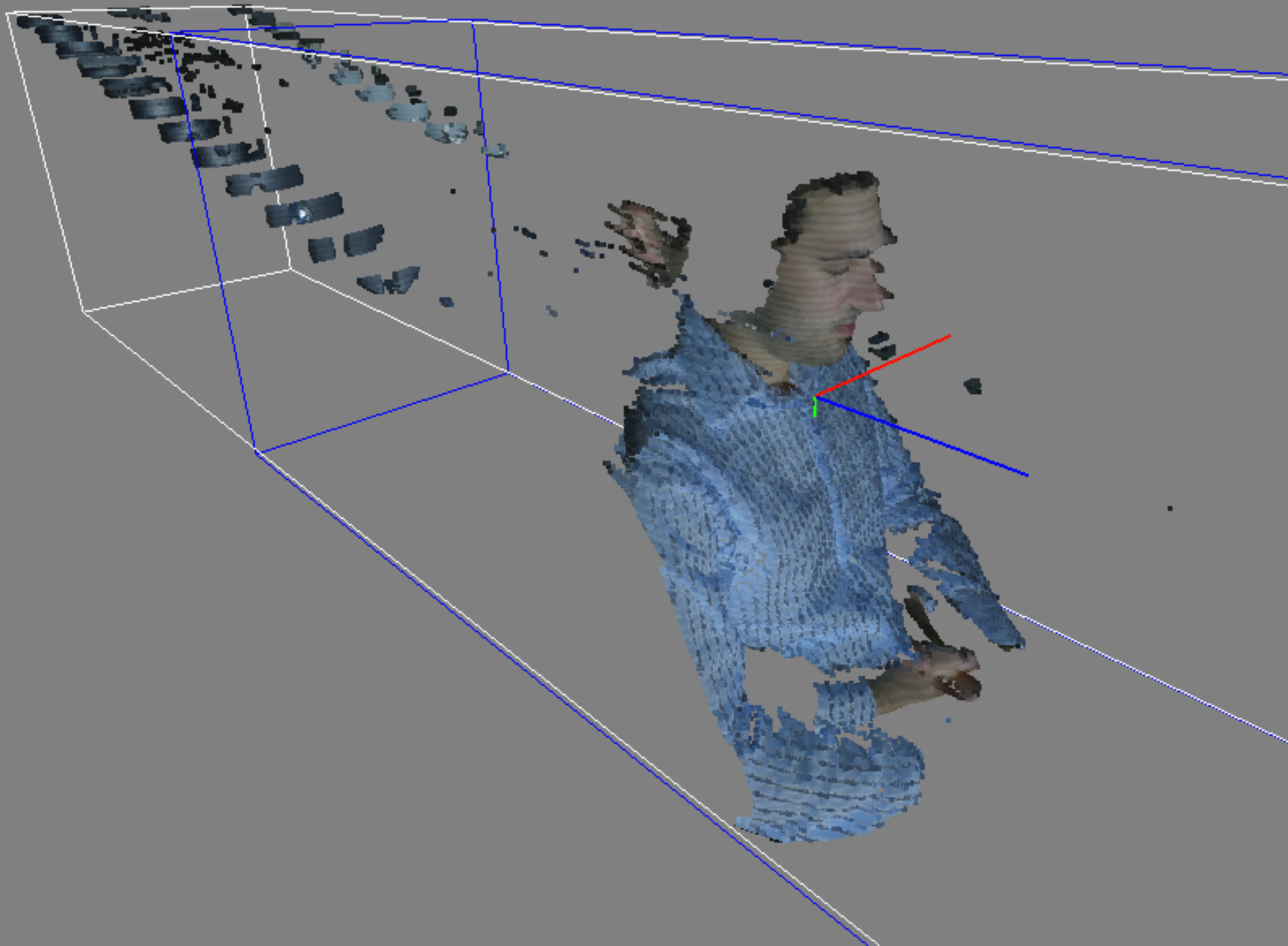




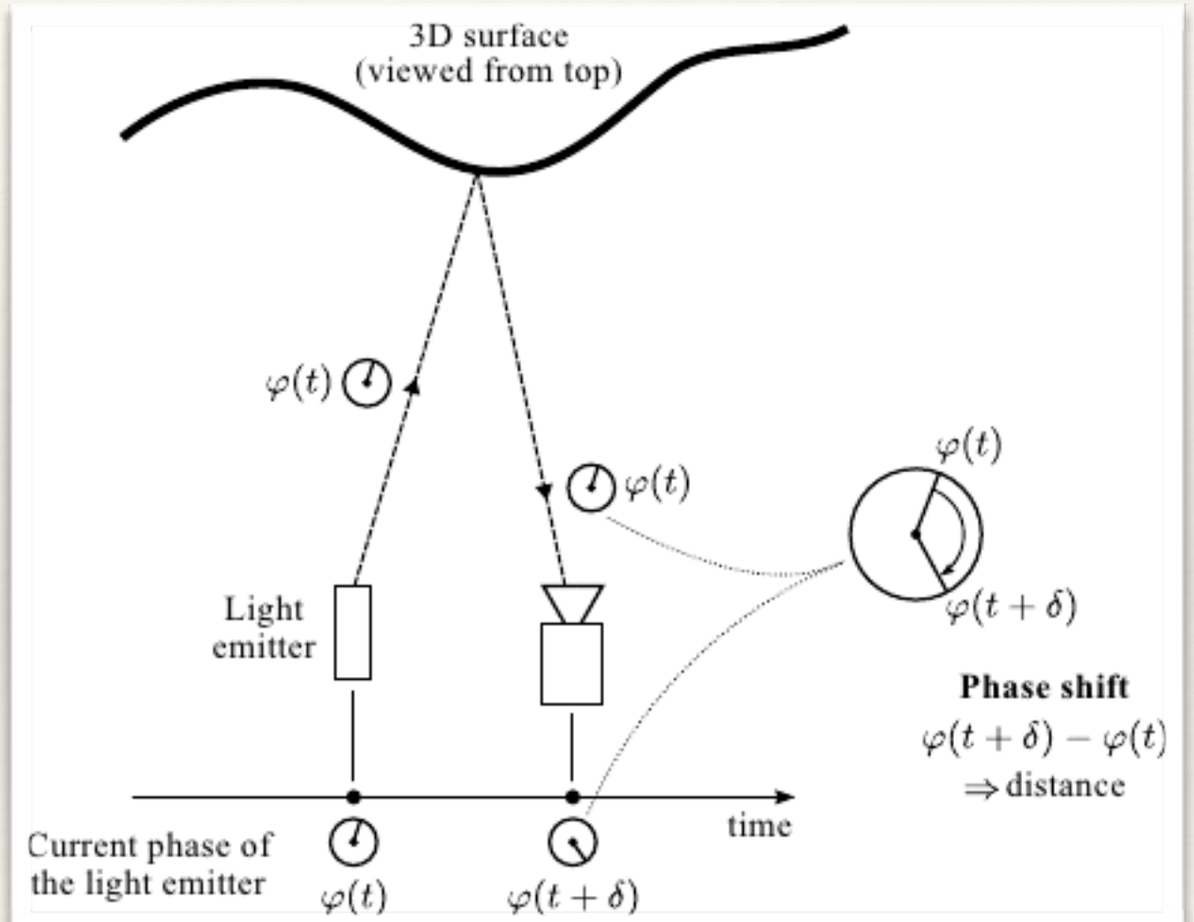


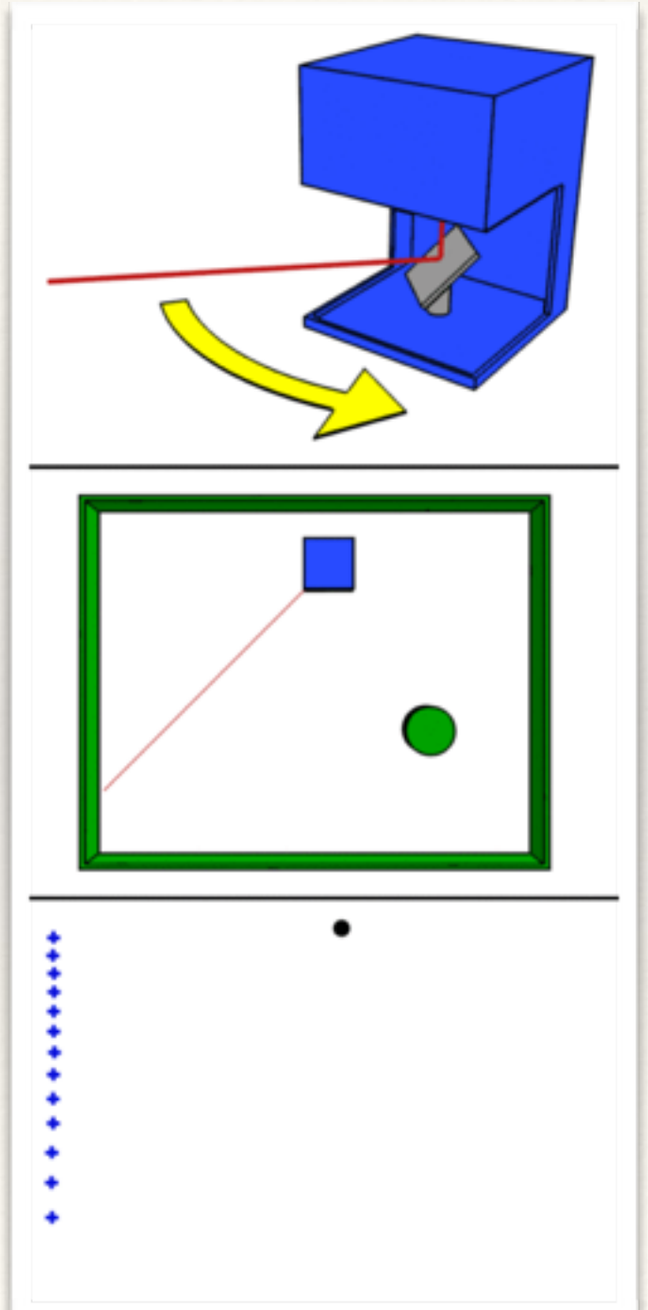


Time of Flight Imaging



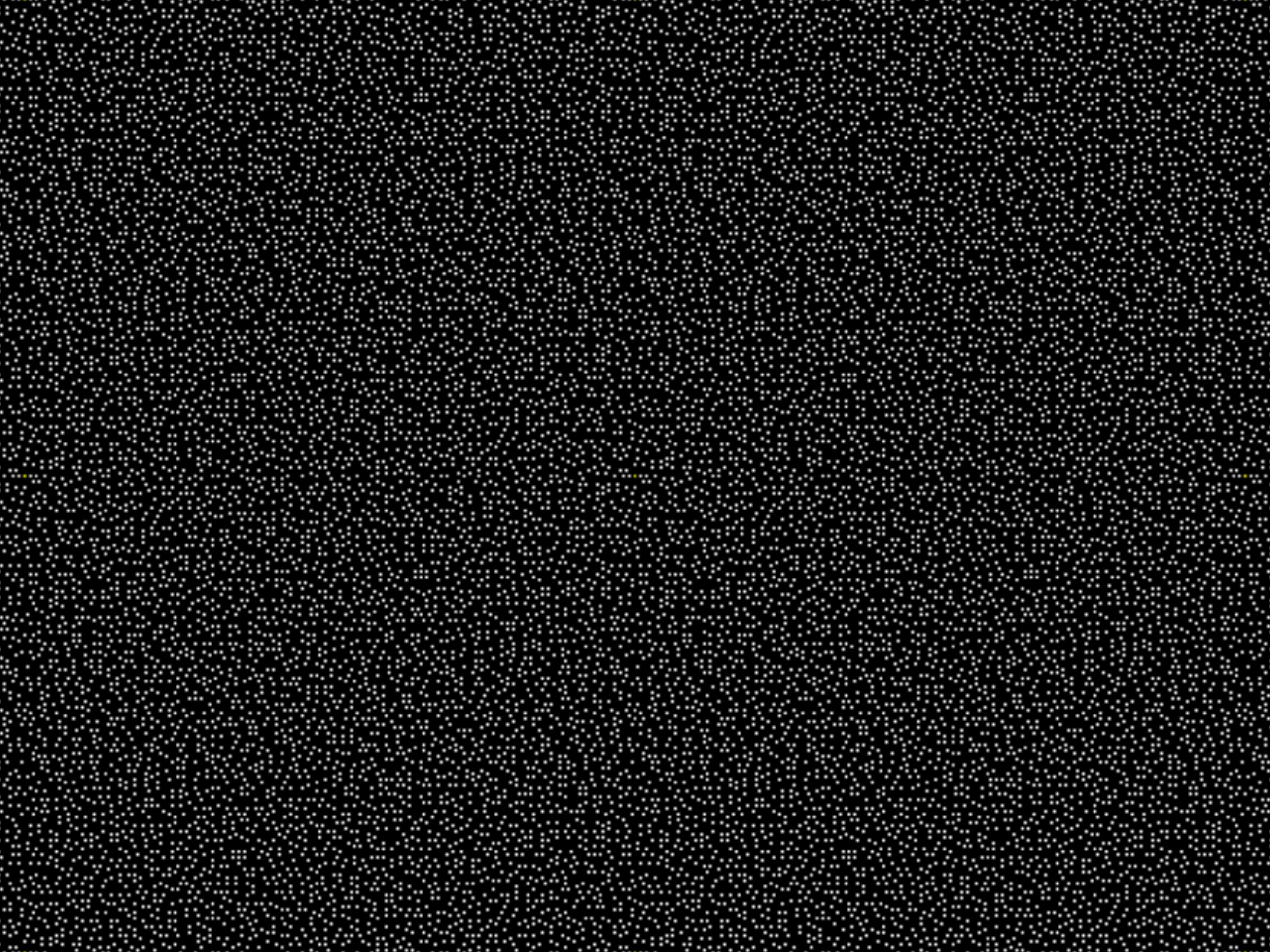
Non-visible techniques

LIDAR



PrimeSense (Kinect)

- Uses *coded* structured IR light
  - IR Laser projects a stationary, random pattern of dots
    - Basically shining light through a opaque stencil with holes in it
  - IR camera records those dots
  - Template matching is used to compare the actual location of the dots from the IR sensor to the known location if the dots were projected on a plane perpendicular to the optical axis at a known distance



Kinect Demo

Summary

- 3D computer vision has lots of practical applications
- Camera models give a mathematical description of how a pixel in a 2D image is related to a point in a 3D scene
  - Camera calibration can be used to find the parameters of a camera
- Multiple views of a scene can be used to infer depth
- There are lots of other techniques for capturing depth that only require a single sensor