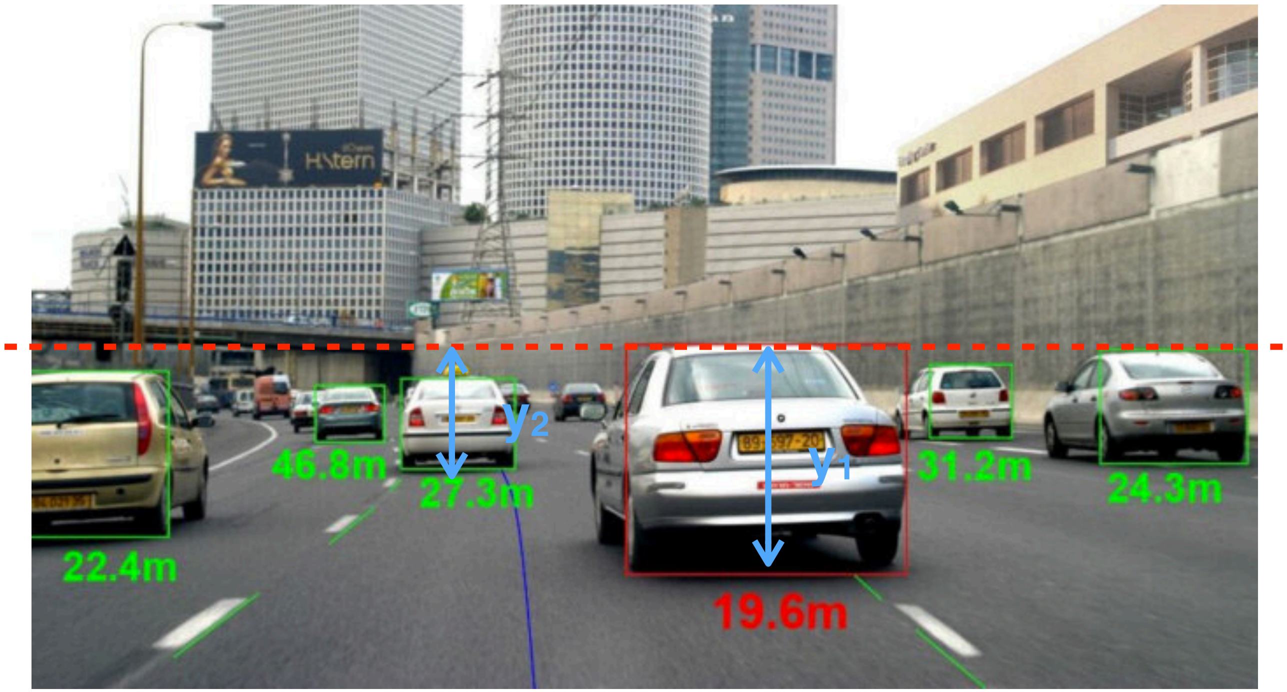
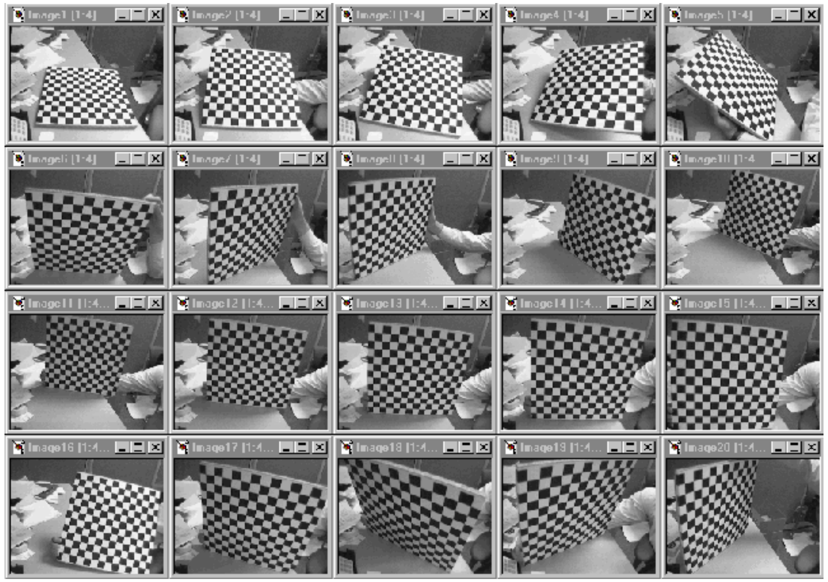
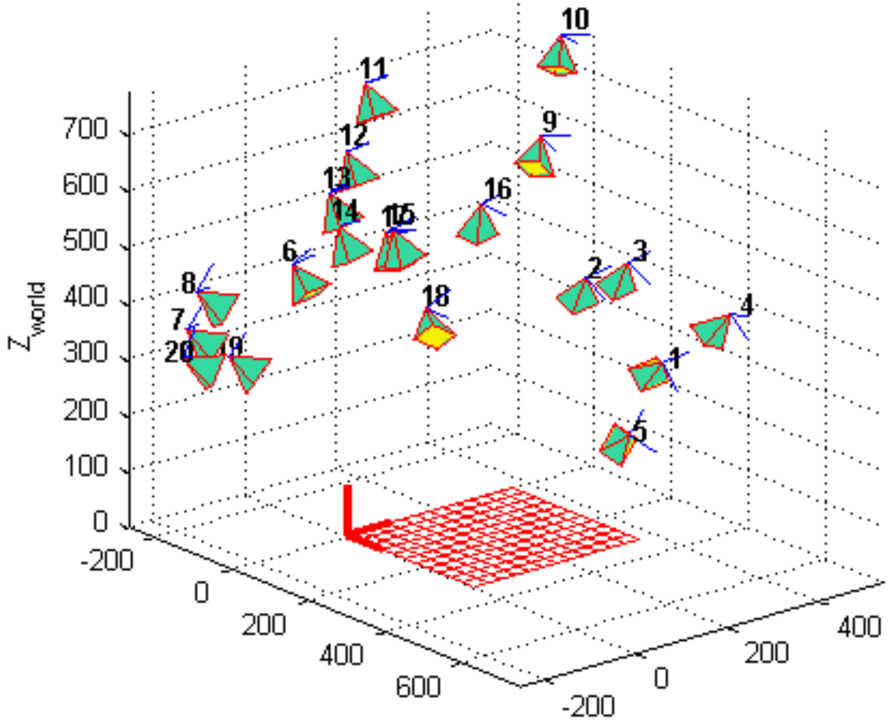
Depth Estimation with Monocular Camera



Camera Calibration





Output:

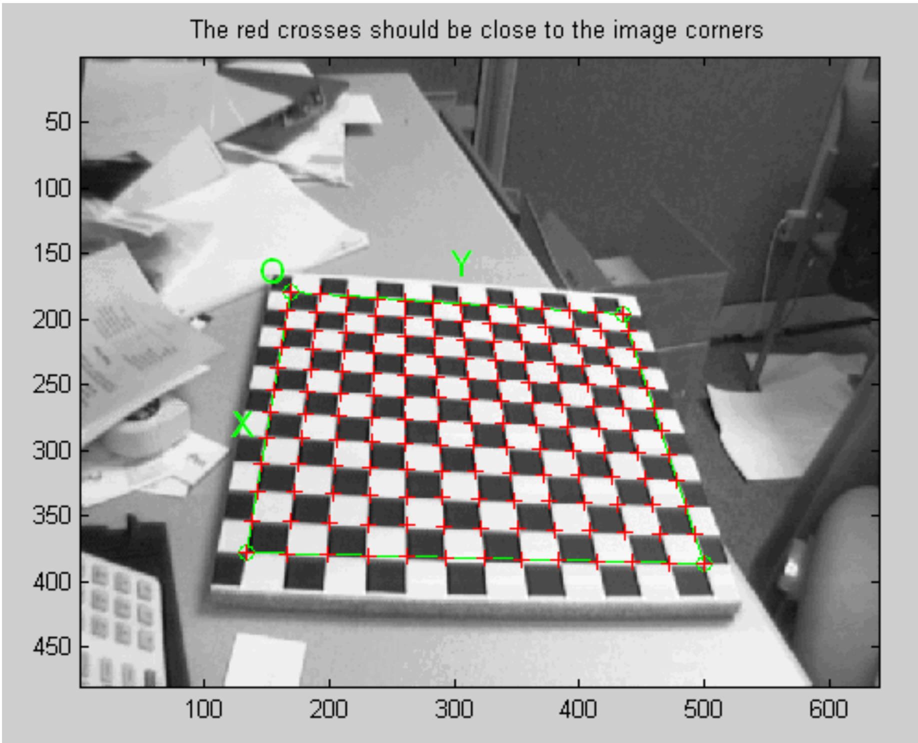
One intrinsic camera parameter (*A* or *K*),

and *n* extrinsic camera parameters (*R* and *T*)

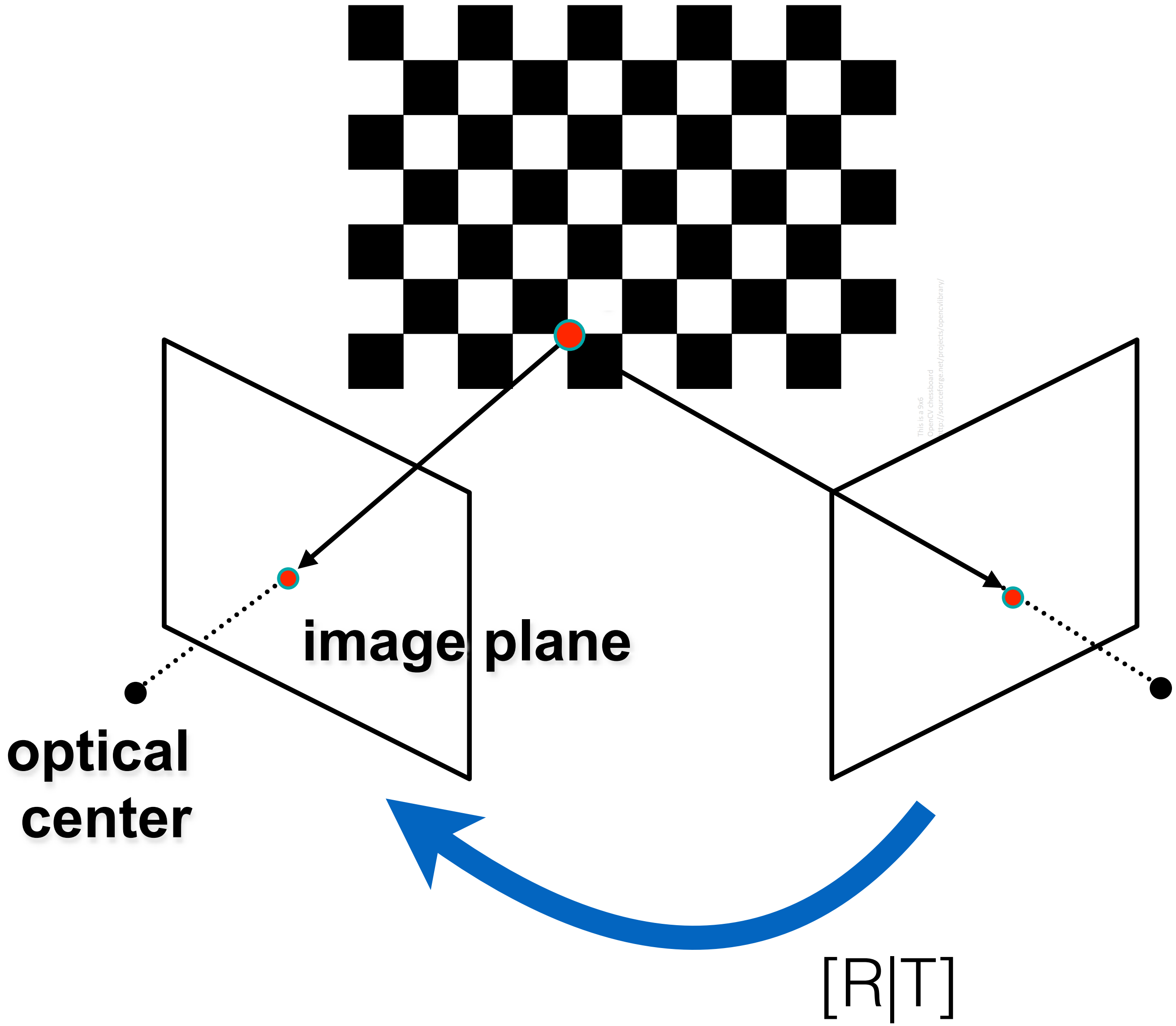
OpenCV: solvePnP

Finds an object pose from 3D-2D point correspondences.

C++: bool **`solvePnP`**(InputArray **objectPoints**, InputArray **imagePoints**, InputArray **cameraMatrix**, InputArray **distCoeffs**, OutputArray **rvec**, OutputArray **tvec**, bool **useExtrinsicGuess**=false, int **flags**=ITERATIVE)



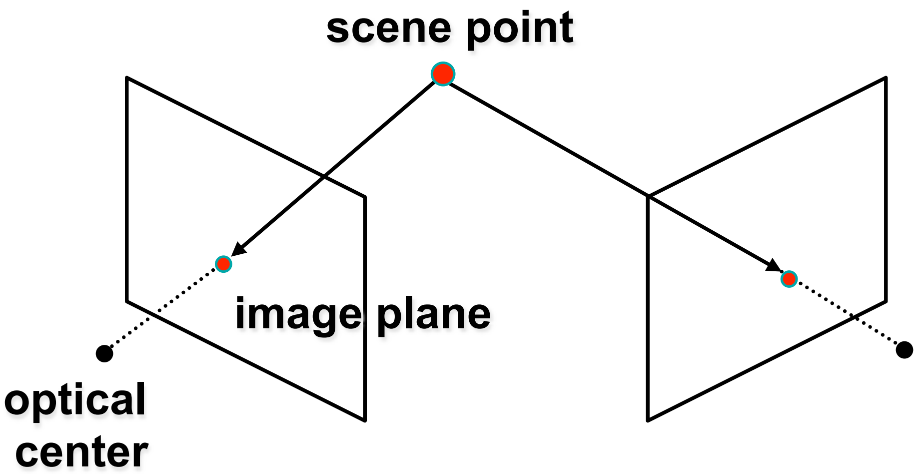
Camera Calibration among Cameras

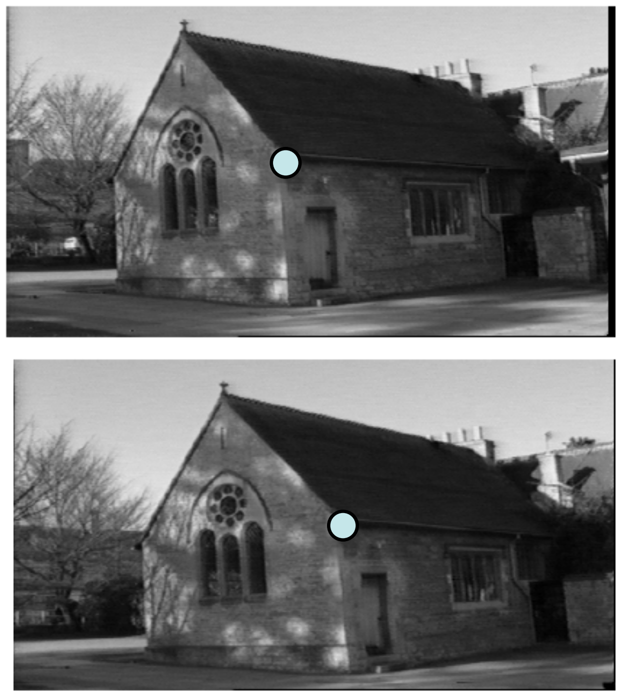


Estimating depth with stereo

- **Stereo**: shape from "motion" between two views

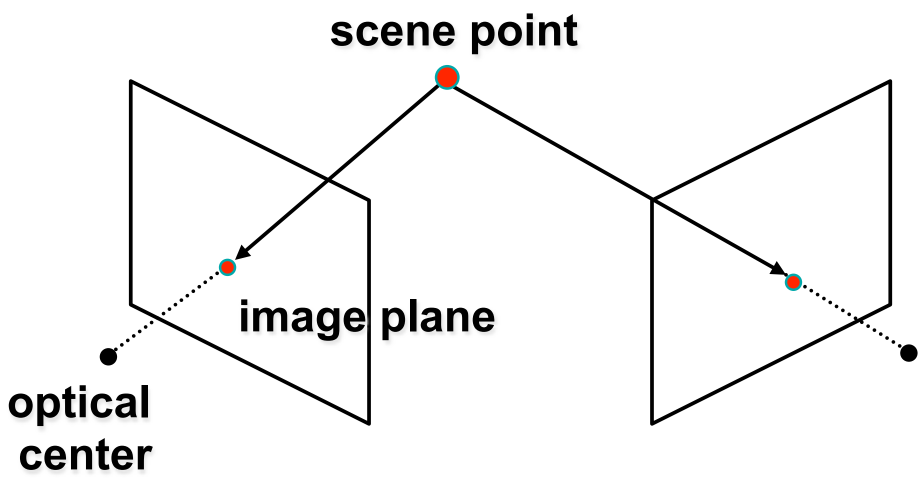
- What we need:
  - Info on camera pose ("calibration")
  - Image point **correspondences**





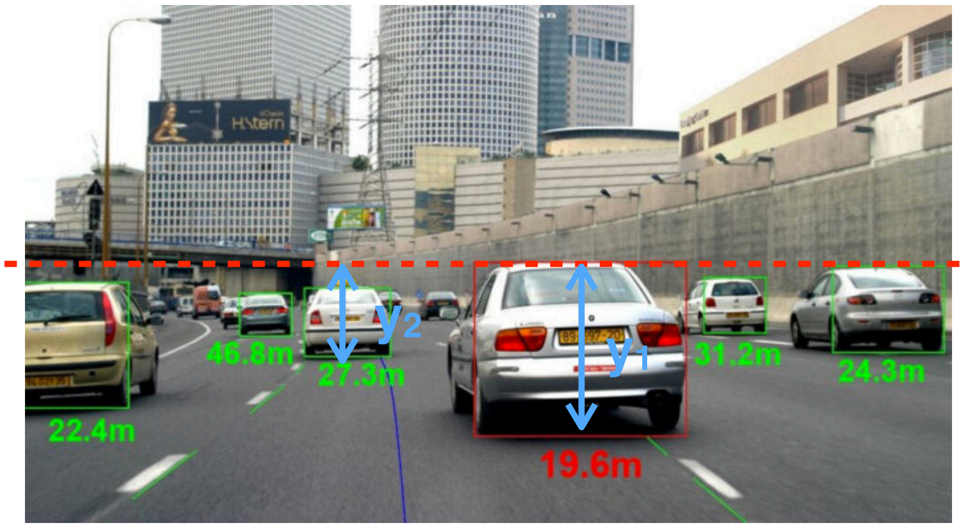
Depth Estimation with Monocular Camera

- To estimate depth, one camera seems not enough



Depth Estimation with Monocular Camera

- To estimate depth, one camera seems not enough
- But, for some applications, it will become possible
  ex. ADAS:
  - Relative pose of cameras to the ground plane is fixed
  - The objects (vehicles or pedestrians) must be on the ground plane



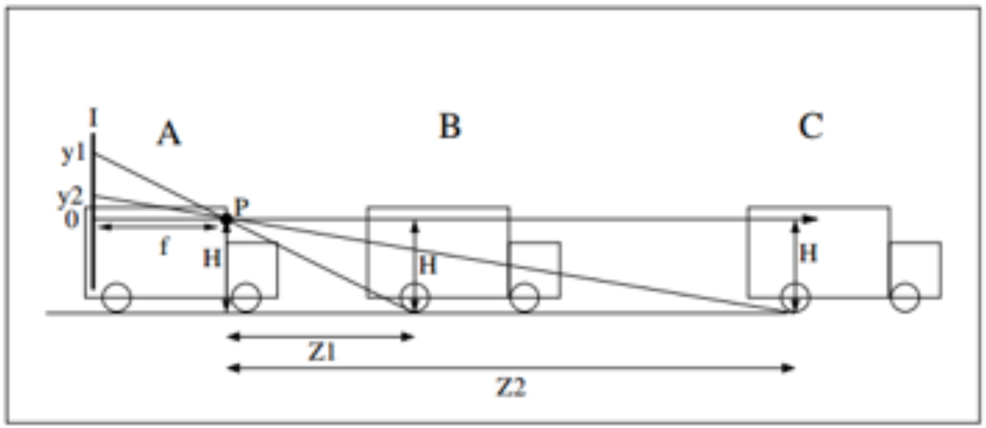
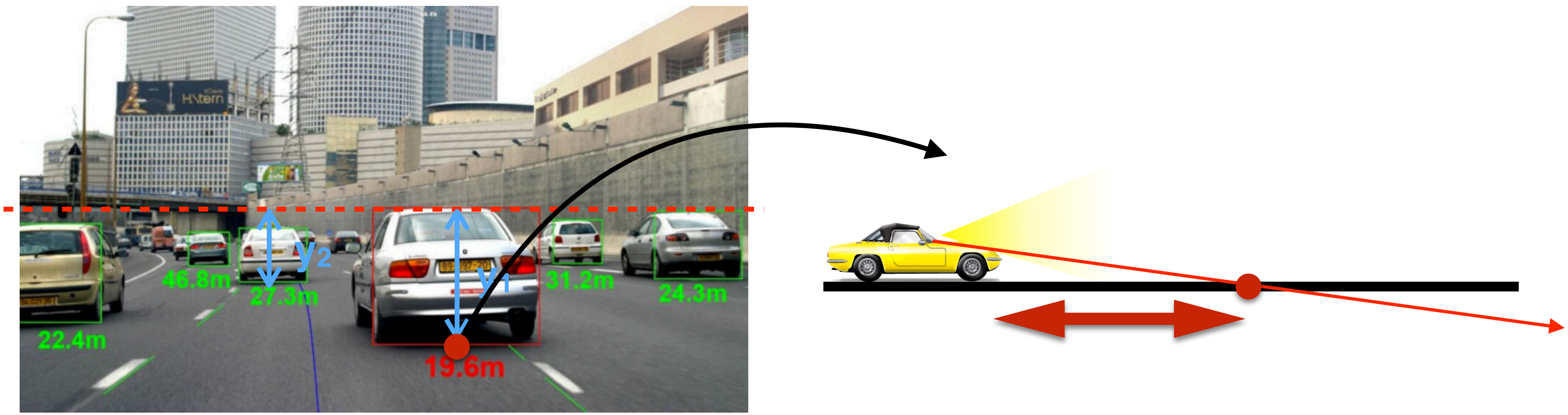


Figure 2: *Schematic diagram of the imaging geometry (see text).*

Some Methods

- Assume camera is calibrated. $s\tilde{m} = A[R \mid t]\tilde{M}$
- **1. find the intersection of a ray and the ground plane**
  - transform a 2D point in image to a 3D ray (with intrinsic parameters)
  - find the intersection (a 3D point) of the ray and the ground plane (with extrinsic parameters)

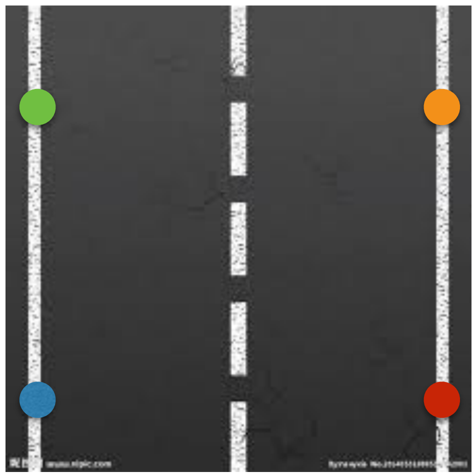


Some Methods

- Assume camera is calibrated.

$$s\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

- **2. using homography between image plane and the ground plane**
  - estimate the homography with at least 4 points
  - transform the coordinate in image to the coordinate on the 3D ground plane

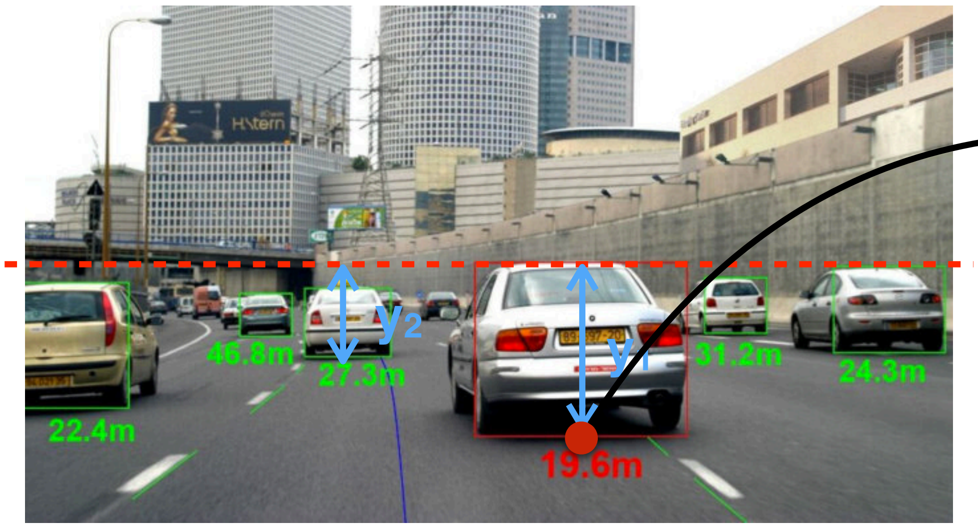


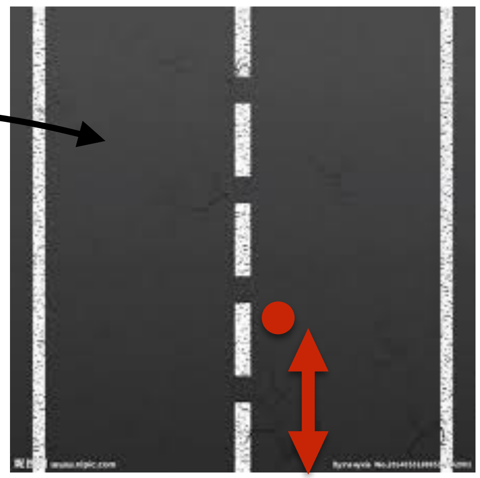
Some Methods

- Assume camera is calibrated.

$$s\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

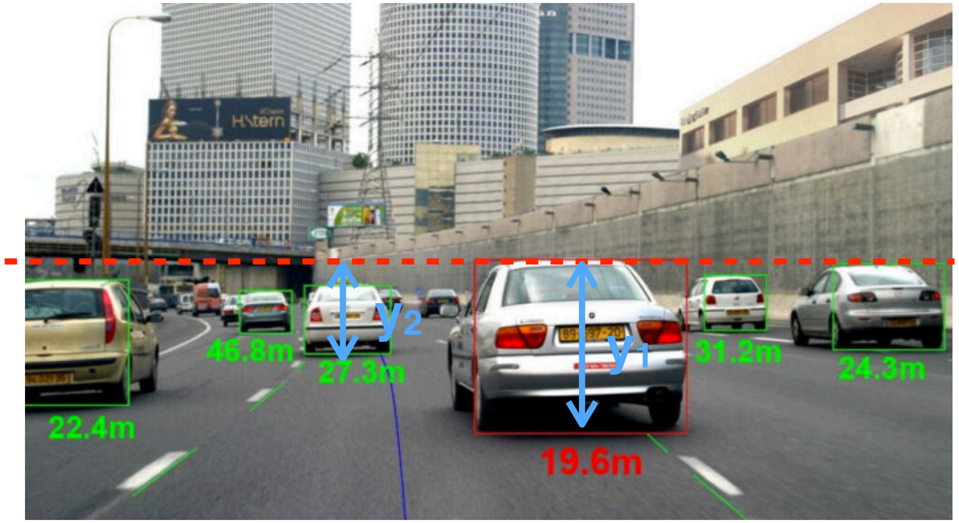
- **2. using homography between image plane and the ground plane**
  - estimate the homography with at least 4 points
  - transform the coordinate in image to the coordinate on the 3D ground plane

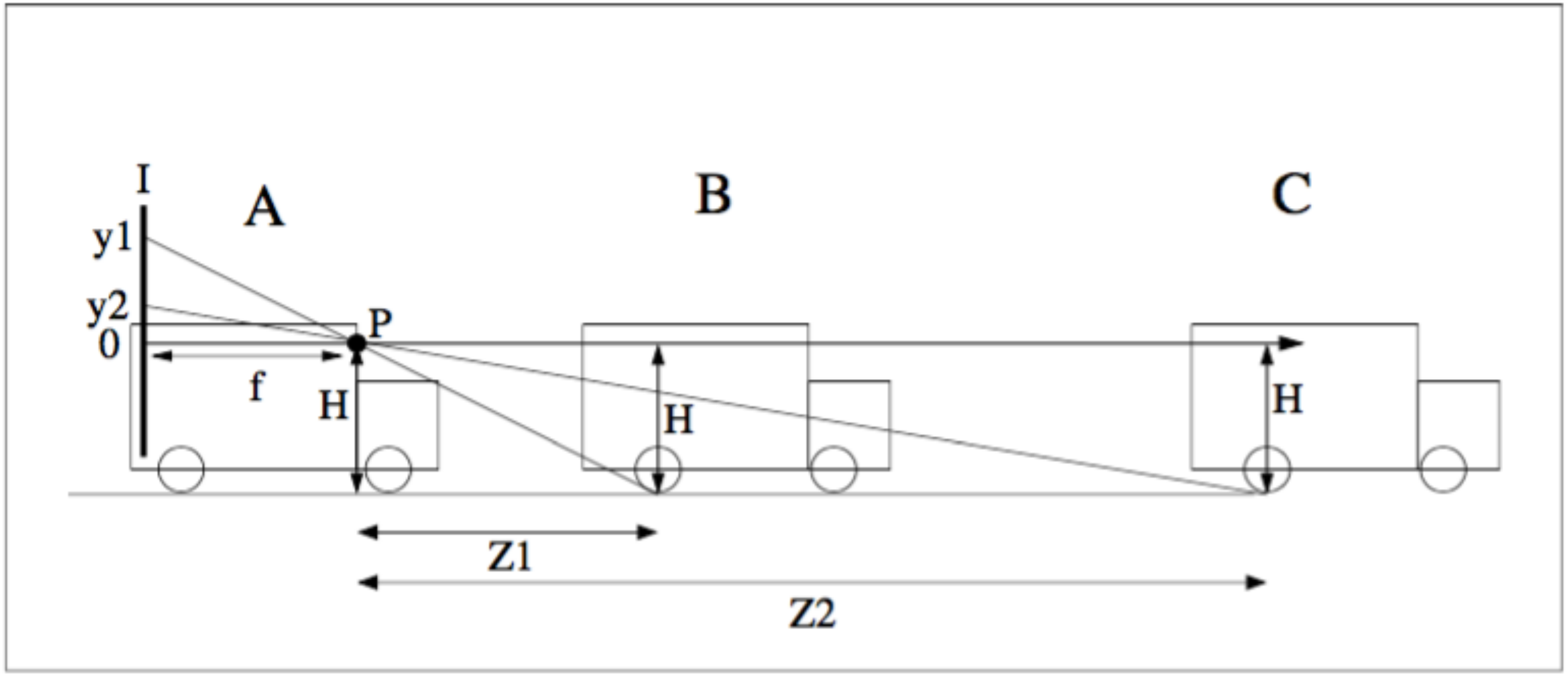




Some Methods

- Assume camera is calibrated.
- **3. using image geometry for dash cam (Mobileye)**



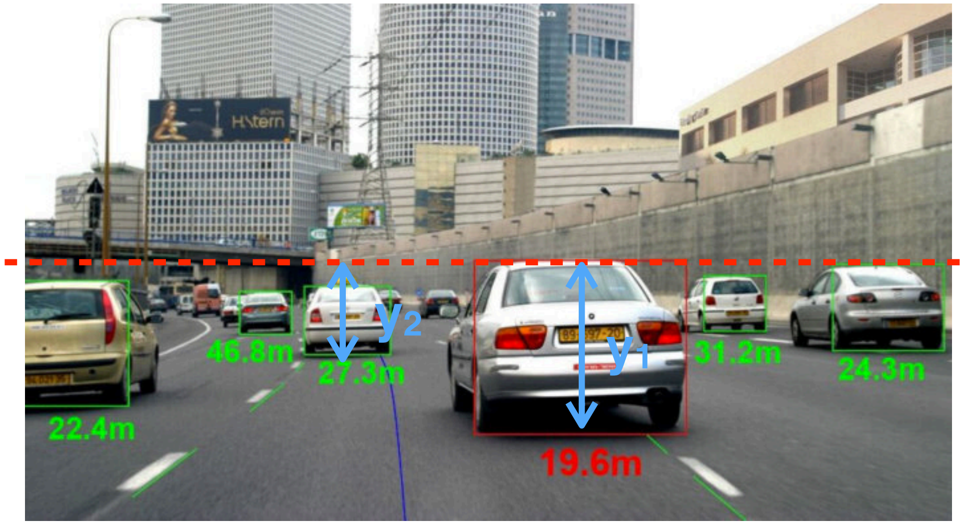


$$Z = \frac{fH}{y}$$

G.P. Stein, O. Mano, A. Shashua, "Vision-based ACC with a single camera: bounds on range and range rate accuracy," *Intelligent Vehicle Symposium*, 2003.

Error Analysis

- **3. using image geometry for dash cam (Mobileye)**



$$Z = \frac{fH}{y}$$

The error in range Z_{err} due to an error of n pixels in location of the contact point is:

$$Z_{err} = Z_n - Z = \frac{fH}{y+n} - Z = \frac{fH}{\frac{fH}{Z}+n} = \frac{nZ^2}{fH+nZ} \quad (3)$$

Typically $n \approx 1$ and $fH >> nZ$ so we get:

$$Z_{err} \approx \frac{nZ^2}{fH} \quad (4)$$

G.P. Stein, O. Mano, A. Shashua, "Vision-based ACC with a single camera: bounds on range and range rate accuracy," *Intelligent Vehicle Symposium*, 2003.