Ismail Akram
8834
CSc 471 Computer Vision
Prof. Zhu

Question 1

Taking the partial derivative of D w.r.t to disparity d:

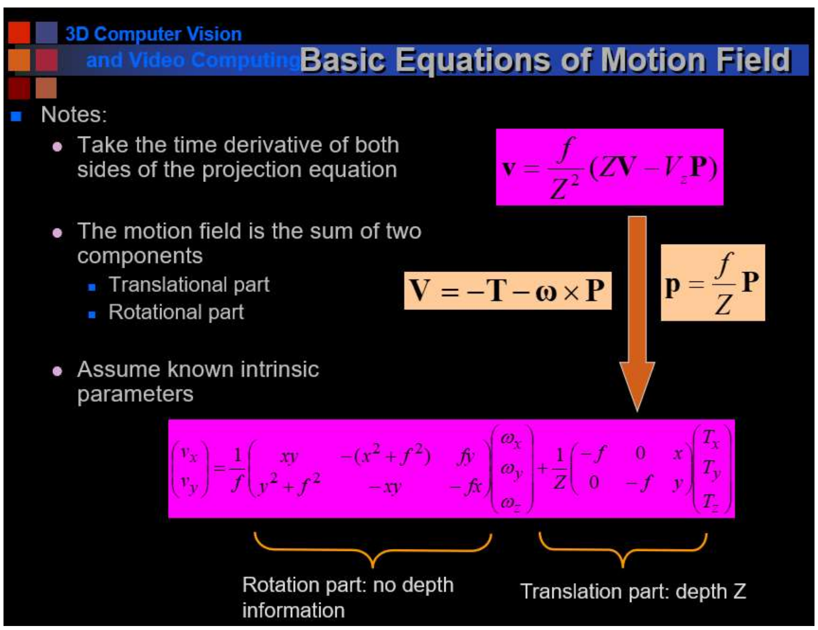
$$D = \frac{fB}{d}$$

$$\frac{\partial D}{\partial d} = \frac{\partial \frac{fB}{d}}{\partial d} = \frac{(fB)\partial\left(\frac{1}{d}\right)}{\partial d} = \frac{\left(-\frac{fB}{d^2}\right)\partial d}{\partial d} = -\frac{fB}{d^2}$$

$$|\partial D| = \left|-\frac{fB}{d^2}\right|\partial d = \left|-(fB)\left(\frac{fB}{D}\right)^2\right|\partial d = \frac{D^2}{fB}\partial d$$

Conclusion:
The dependence of the error in-depth estimation of a 3D point as a function of:
- Baseline length increase means a decrease in the error (inverse relation)
- Focal length increase means decrease in error (inverse relation)
- Stereo matching error increase means increase in error (proportional relationship)
- 3D point depth increase means increase in error (proportional relationship)

Question 2



~~Yes, by rotating a camera around its optical center and viewing multiple frames, one can triangulate 3D information provided they know the rate of rotation and the rotational matrix R and transpose T between the two frames.~~

No, simply rotating about the optical center yields no depth information. With no depth information, we don't have an 3D information.

Ismail Akram
8834
CSc 471 Computer Vision
Prof. Zhu

No, we are simply translating along the optical axis rather than zooming. This means we only get a local reference for the depth and not the true depth from the initial camera position O. However, if the camera were to move along the baseline rather than the optical axis, then we can extract 3D information.

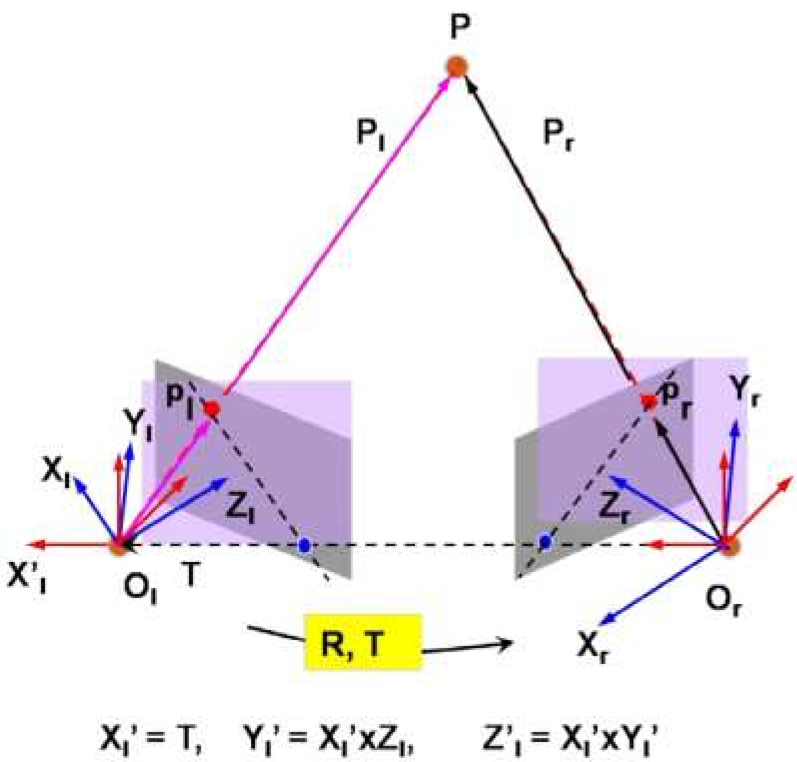


Figure 10. Stereo rectification

We need both rotation and translation (R+T) to net 3D information among the two image planes in reference to the point P.

Ismail Akram
8834
CSc 471 Computer Vision
Prof. Zhu

Question 3

Human:

1. Parallax and stereo motion in our eyes. For example, if we observe the same object in two different scenarios, our brains know to calibrate a weighted combination of information. The object in question would appear “large” In a small room and “small” a large room.
2. We can then relate object size to the size of nearby reference objects, like a brick wall, ignoring perceived sizes SL and SR when making their match.
3. In a static environment, performance depends on the distance of comparison cubes alone.
4. We can infer a lot of information of a blurry image sequence based on motion alone.
5.

Machine:

1. Video Coding and Compression & Video View interpolation
2. Robot navigation (robot vision)
3. Laser Scanning models
4. Recovering 3D models (Computer Vision) & Image-based Rendering
5. Anaglyph from Mars Rover