Depth estimation with images captured from multiple views

Prasan Shedligeri
ee16d409@ee.iitm.ac.in
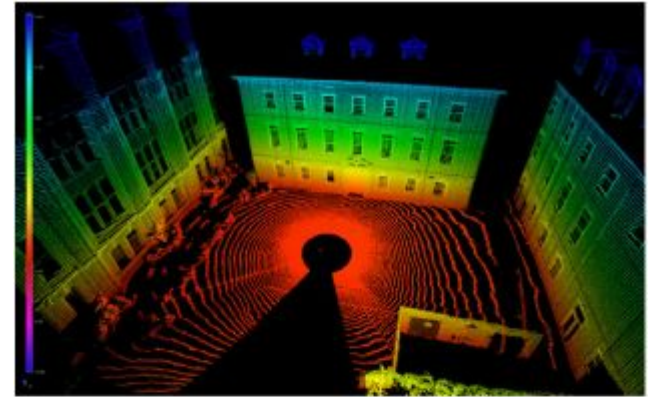
What is depth estimation?

- Determining the distance of each scene point with respect to the camera

Applications

- Autonomous Navigation
  - Indoor (Humanoid)
  - Roads (Cars)
  - Aerial (Drones)
- Virtual/Augmented Reality
  - Virtual Meeting Rooms
  - Gaming
  - Robotic Surgeries
- Scene Understanding
  - Semantic Summarization
  - Large Scale 3D Mapping
  - Architecture/Heritage Capture

How to Estimate Depth ?

- Depth from single image
- **Depth from Multi-view Images / Video**
- Depth from X: Shading, Focus, …
- Depth from Active Sensing : Kinect; LIDAR; Structured Light etc

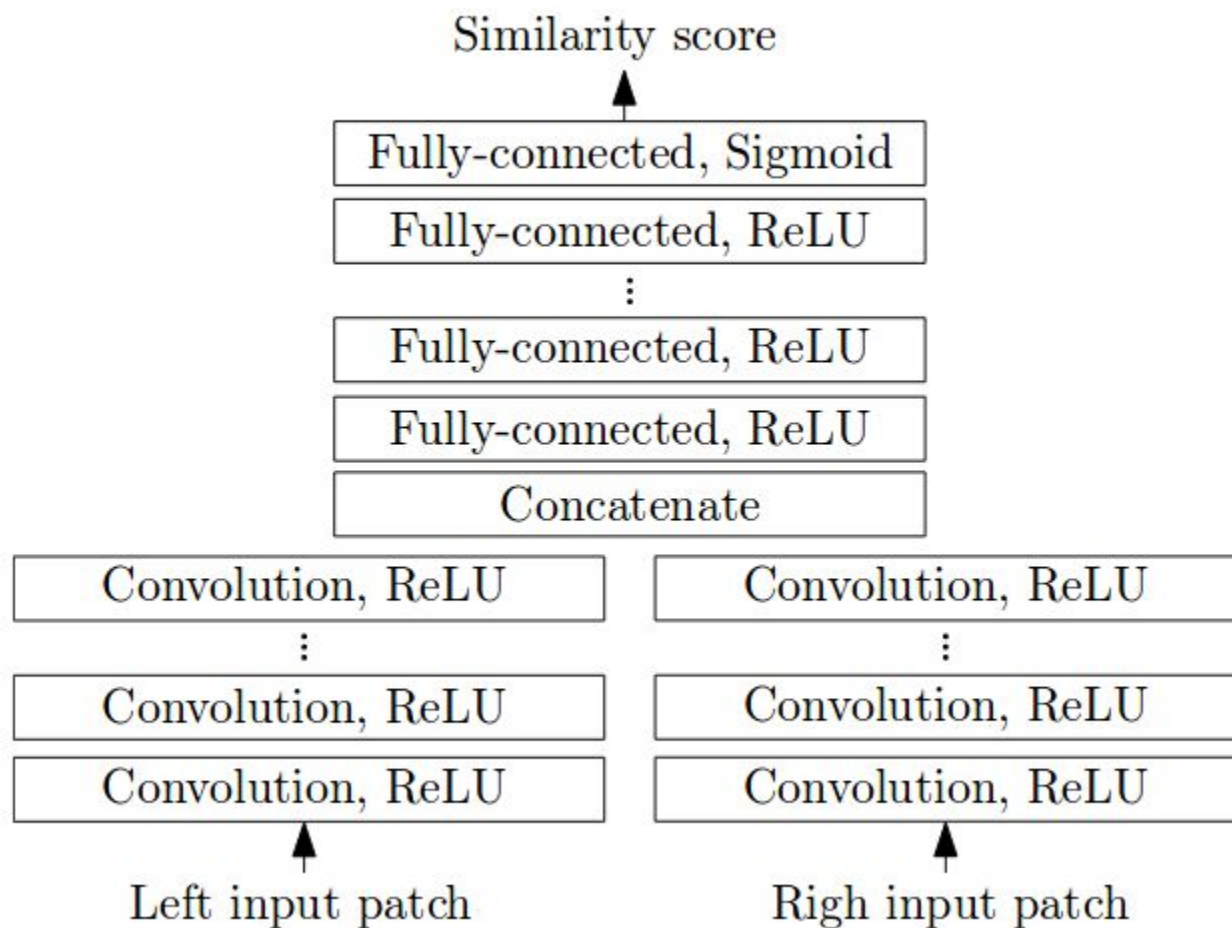# Simple case: Depth from stereo images

Parallax effect

Images taken from two cameras

One of the cameras is horizontally shifted with respect to the other

Stereo Matching by Training a Convolutional Neural Network to Compare Image Patches

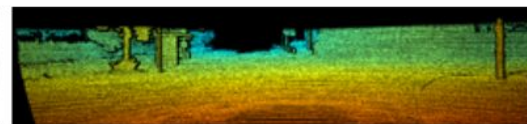Fig courtesy: Zbontar et al. 2015

Results

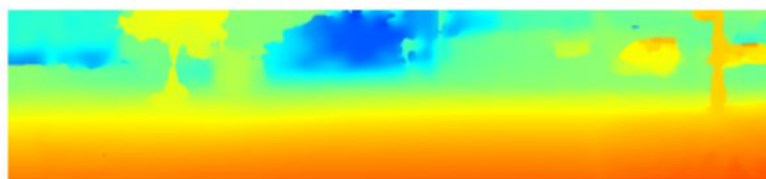Left input image   Right input image   Ground truth
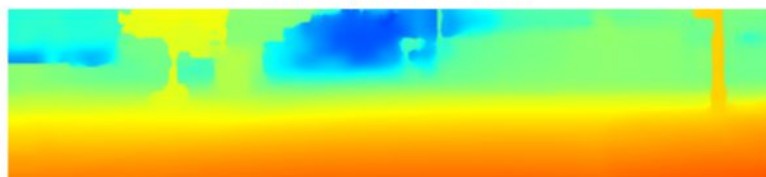
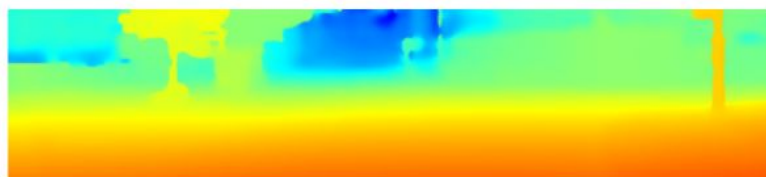Census   Error: 4.63 %

Fast architecture   Error: 1.01 %

Accurate architecture   Error: 0.91 %

Results
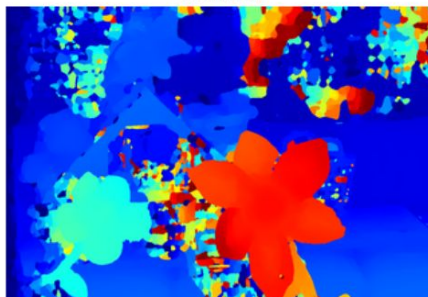

Left input image   Right input image   Ground truth

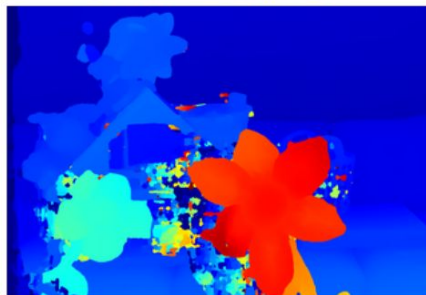Census   Error: 34.65 %
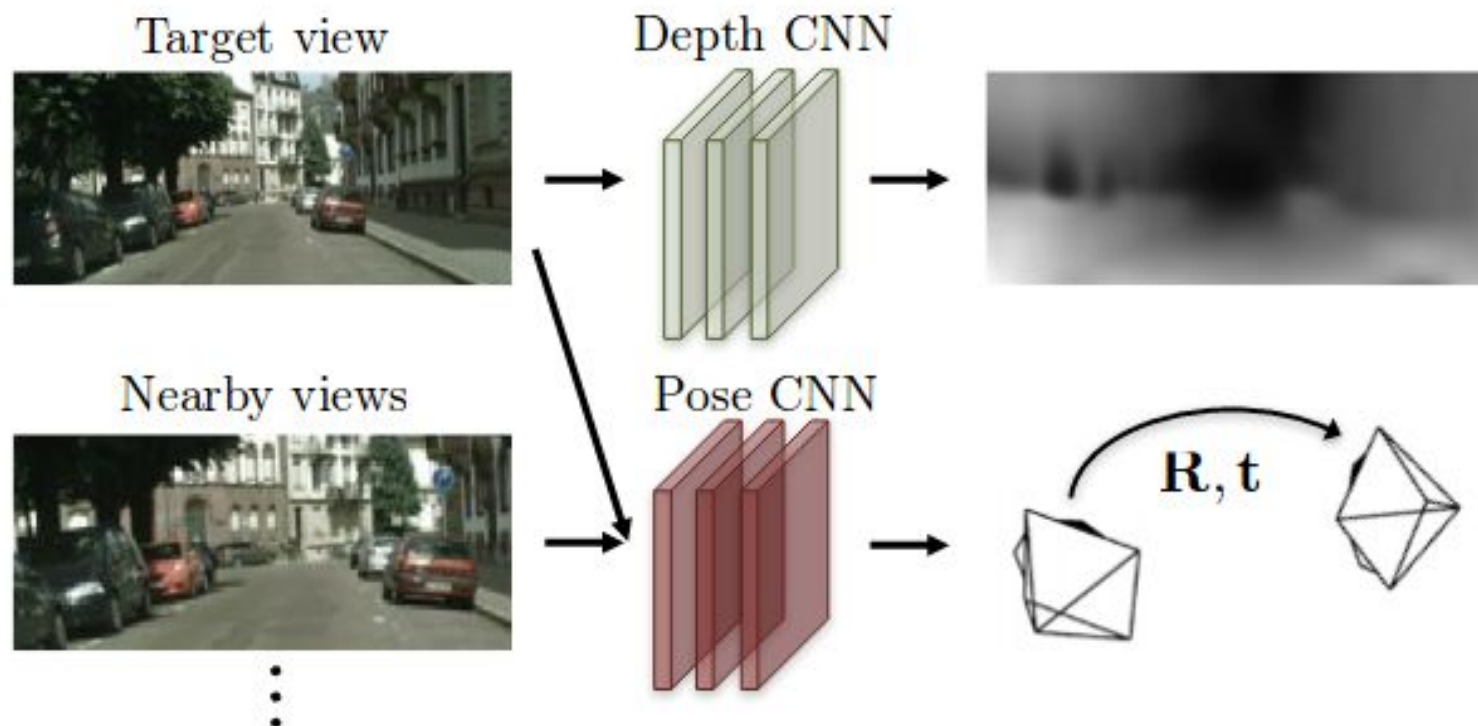
Accurate architecture   Error: 7.25 %

Unsupervised Learning of Depth and Ego-Motion from Video

Arbitrary motion of the camera

Results

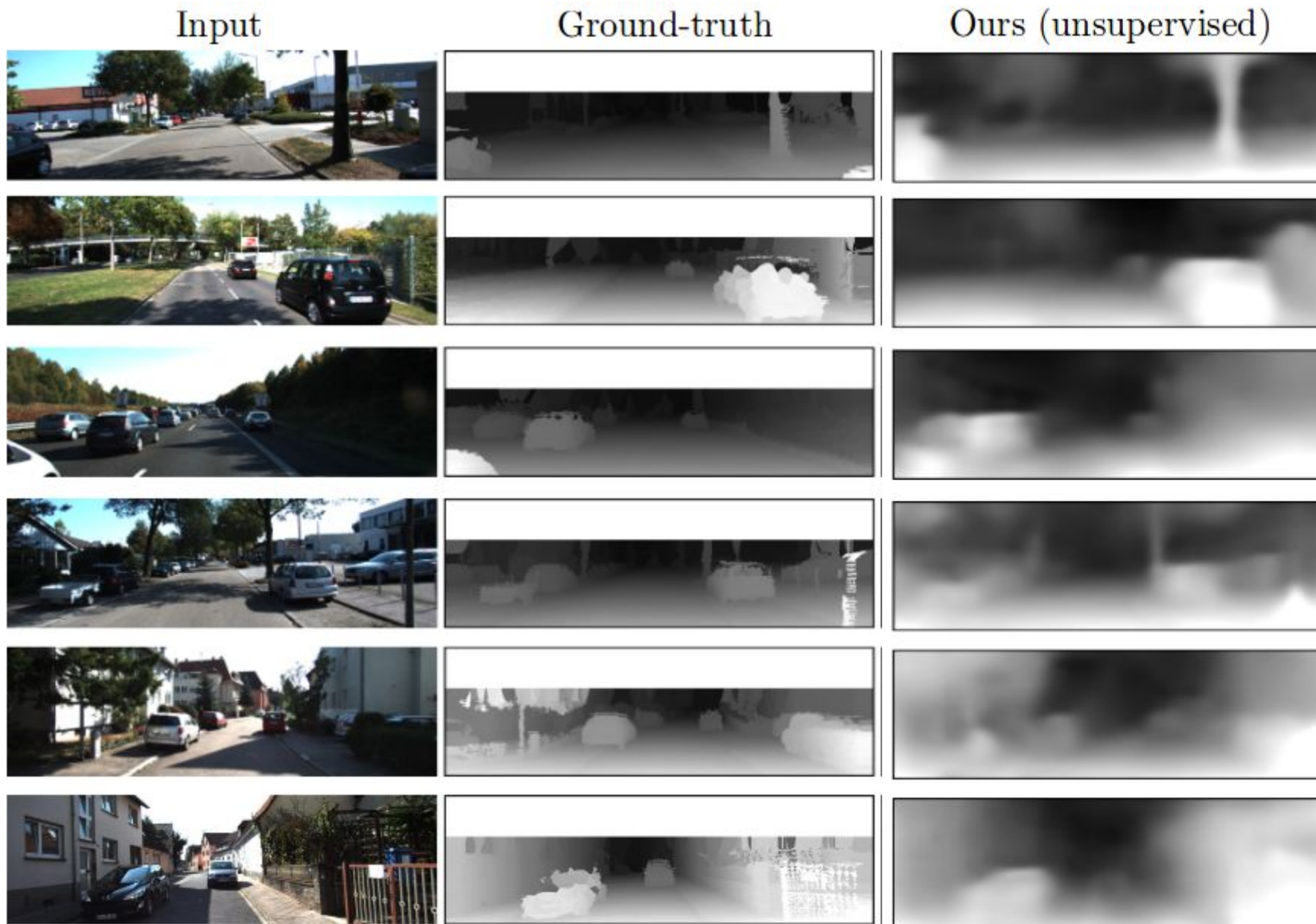| Input	Ground-truth	Ours (unsupervised)

Figure courtesy: Zhou et al. '17

References

1) Ummenhofer, Benjamin, et al. "Demon: Depth and motion network for learning monocular stereo." *IEEE Conference on computer vision and pattern recognition (CVPR)*. Vol. 5. 2017.
2) Zhou, Tinghui, et al. "Unsupervised learning of depth and ego-motion from video." *CVPR*. Vol. 2. No. 6. 2017.
3) Wang, Chaoyang, et al. "Learning Depth from Monocular Videos using Direct Methods." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2018.
4) Mahjourian, Reza, Martin Wicke, and Anelia Angelova. "Unsupervised Learning of Depth and Ego-Motion from Monocular Video Using 3D Geometric Constraints." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2018.
5) Zbontar, Jure, and Yann LeCun. "Computing the stereo matching cost with a convolutional neural network." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015.