FORESIGHT

CONSISTENT PIXEL MATCHING USING TEMPORAL UPDATES

Omri Danziger

Algorithms developer

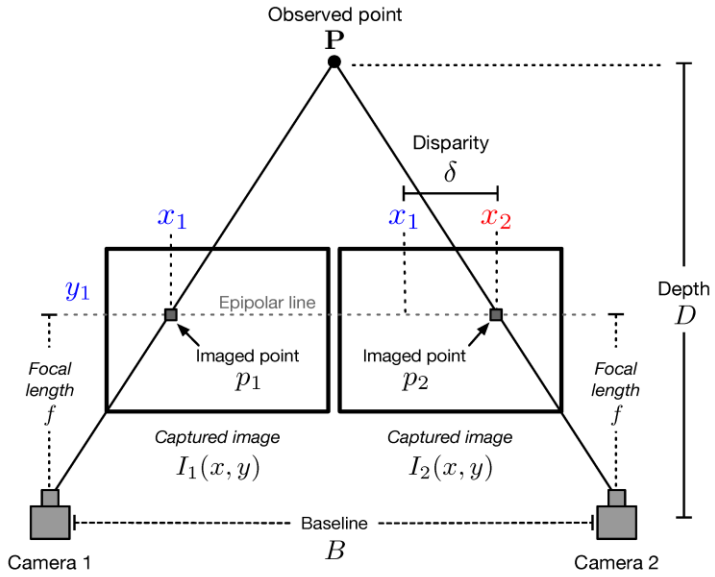
8.4.2024

ForesightAuto Company

- Autonomous vehicle solutions that provide real-time value to customers around the globe
- Our advanced stereoscopic vision technology supports large baselines that enable high accuracy at long ranges and don’t require stringent mechanical requirements.
- **foresightauto.com** for more details

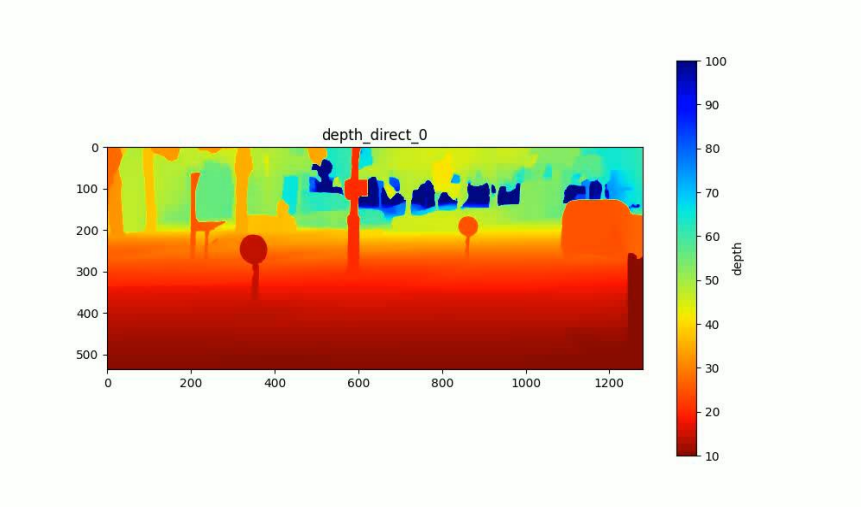
3D Reconstruction using Stereo Matching

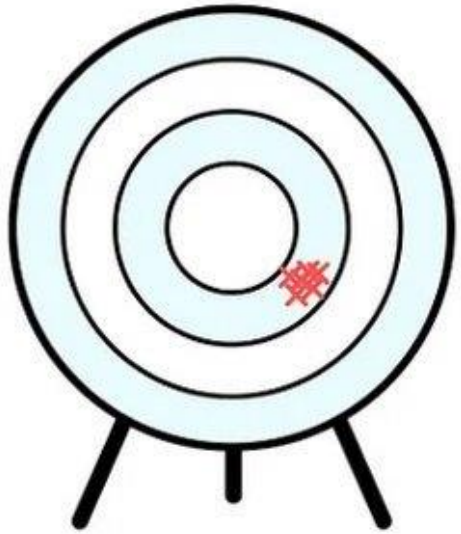
- calibrated cameras with known relative pose
- Matching pixels between images
- Triangulating to form a 3D reconstruction



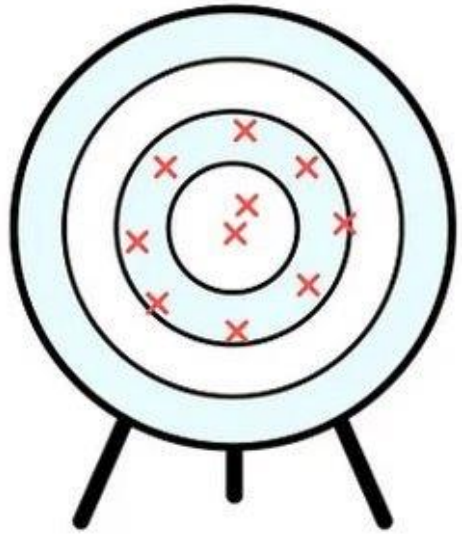
Consistency over time

- May be described as the opposite of "jumpy".
- "Jumpy" models are often not ideal, especially for visualization
- Main causes:
  - Error distribution
  - Multiple "good" decisions
- Balanced with accuracy





Biased, consistent



Unbiased, inconsistent

Popular solutions

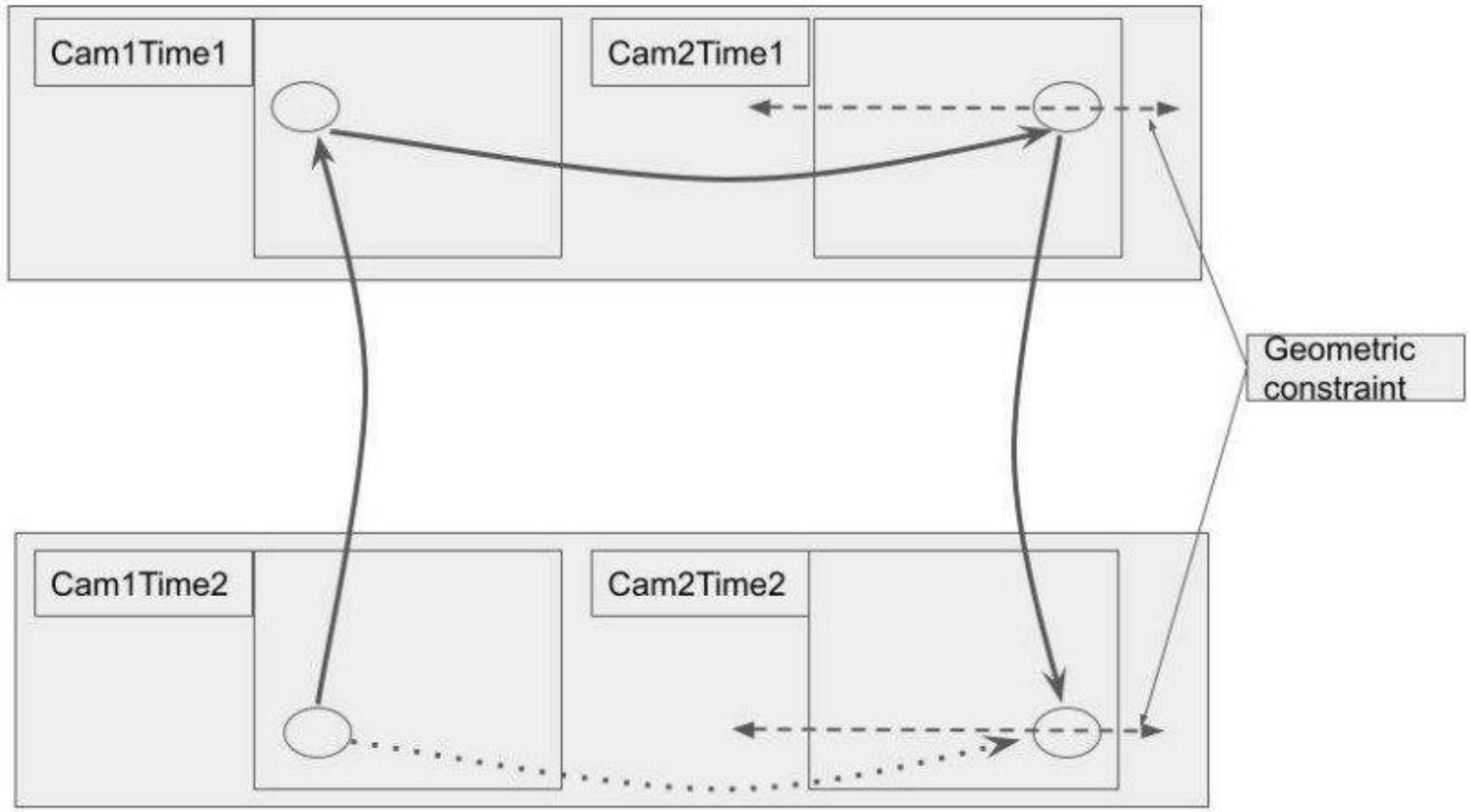
- Filter with confidence
  - Less results
  - Difficulty with multiple good options
- Training Augmentations
  - Difficult to replicate perspectives and occlusions
- Temporal models – SLAM based
  - Dynamic objects difficulty
- Temporal models – Disparity change based
  - Unique blocks

Temporal model – Using Individual Temporal Adjustments

- The ball moves across the left camera images
- The ball moves differently across the right camera images
- Tracking it in both cameras will allow distance estimation

Indirect Stereo Matching

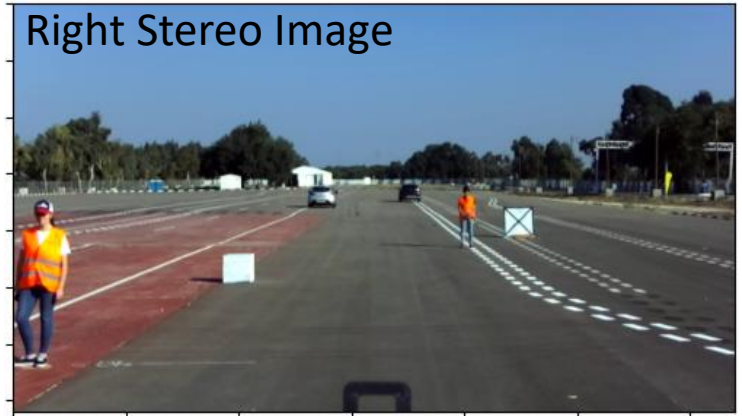
- Create an indirect stereo matching in Time2 using a dense path composition
- Any decision or error from the stereo matching in Time1 will be passed into Time2
- Can only pass objects seen throughout all the images in the composition

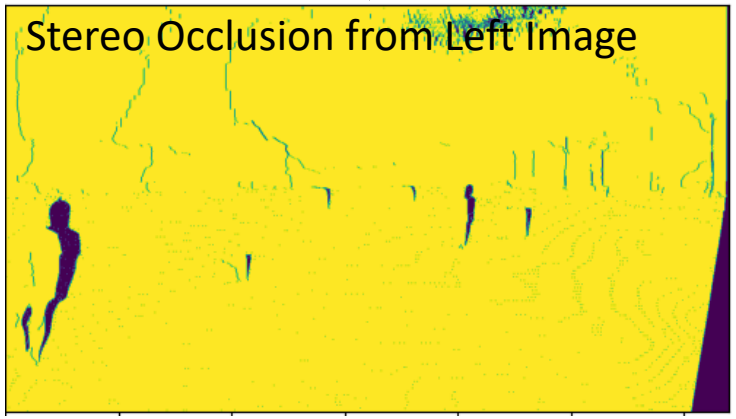


$$L2 \rightarrow R2 = (R1 \rightarrow R2) \circ (L1 \rightarrow R1) \circ (L2 \rightarrow L1)$$

Indirect Stereo Matching - mask

- Stereo occlusion areas are often estimated accurately in Stereo networks, but cannot be us in the optical flow
- Calculated similary to the flow composition

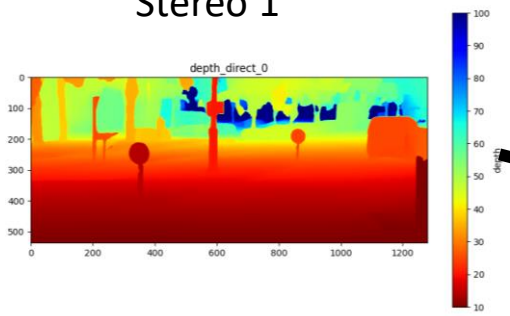




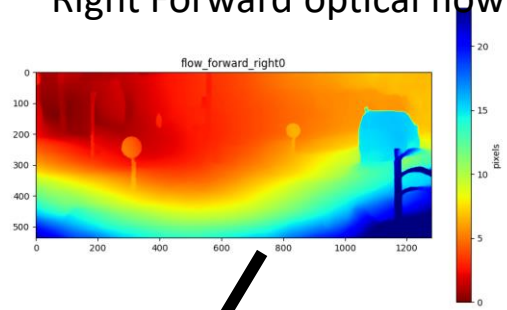
$$\text{mask}_{L_1} = \text{mask}_{L_1 \to R_1} \wedge \text{mask}^{L_1}_{R_1 \to R_2}$$
$$\text{mask}_{L_2 \to R_2} = \text{mask}_{L_2 \to L_1} \wedge \text{mask}^{L_2}_{L_1}$$

Matching Example

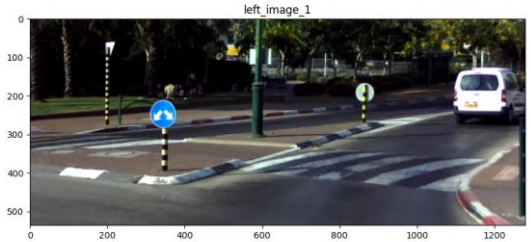
Stereo 1



Right Forward optical flow

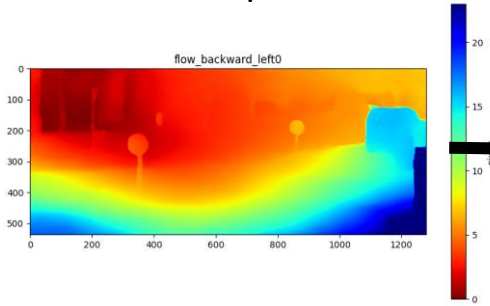


Left Image 1

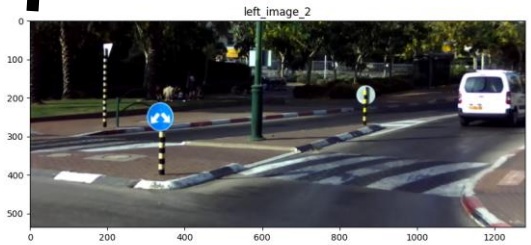


Right Image 1

Left backward optical flow

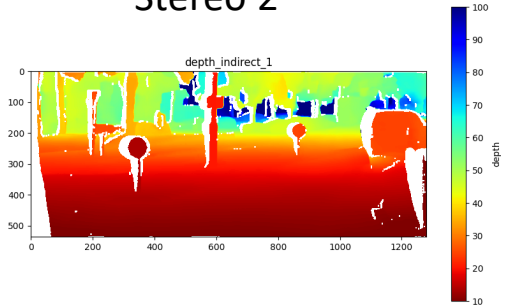


Left Image 2



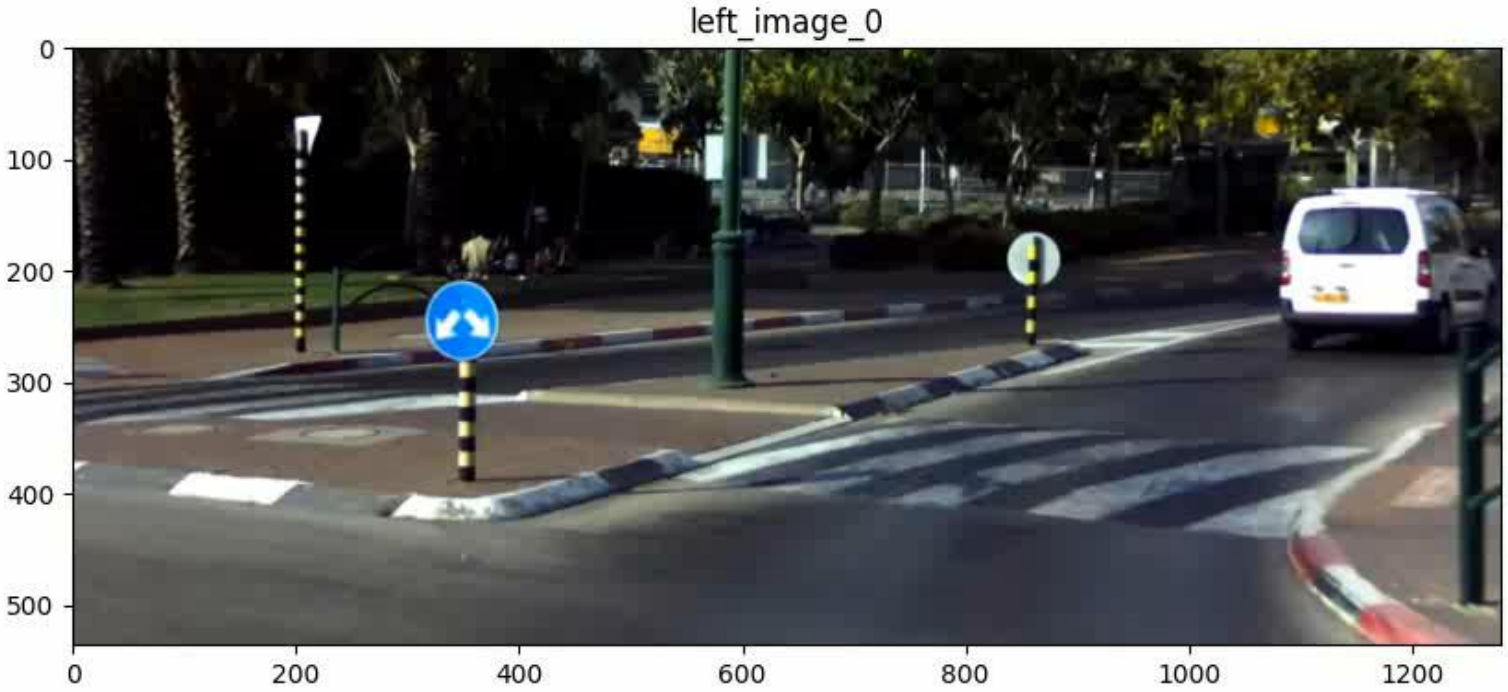
Right Image 2

Stereo 2

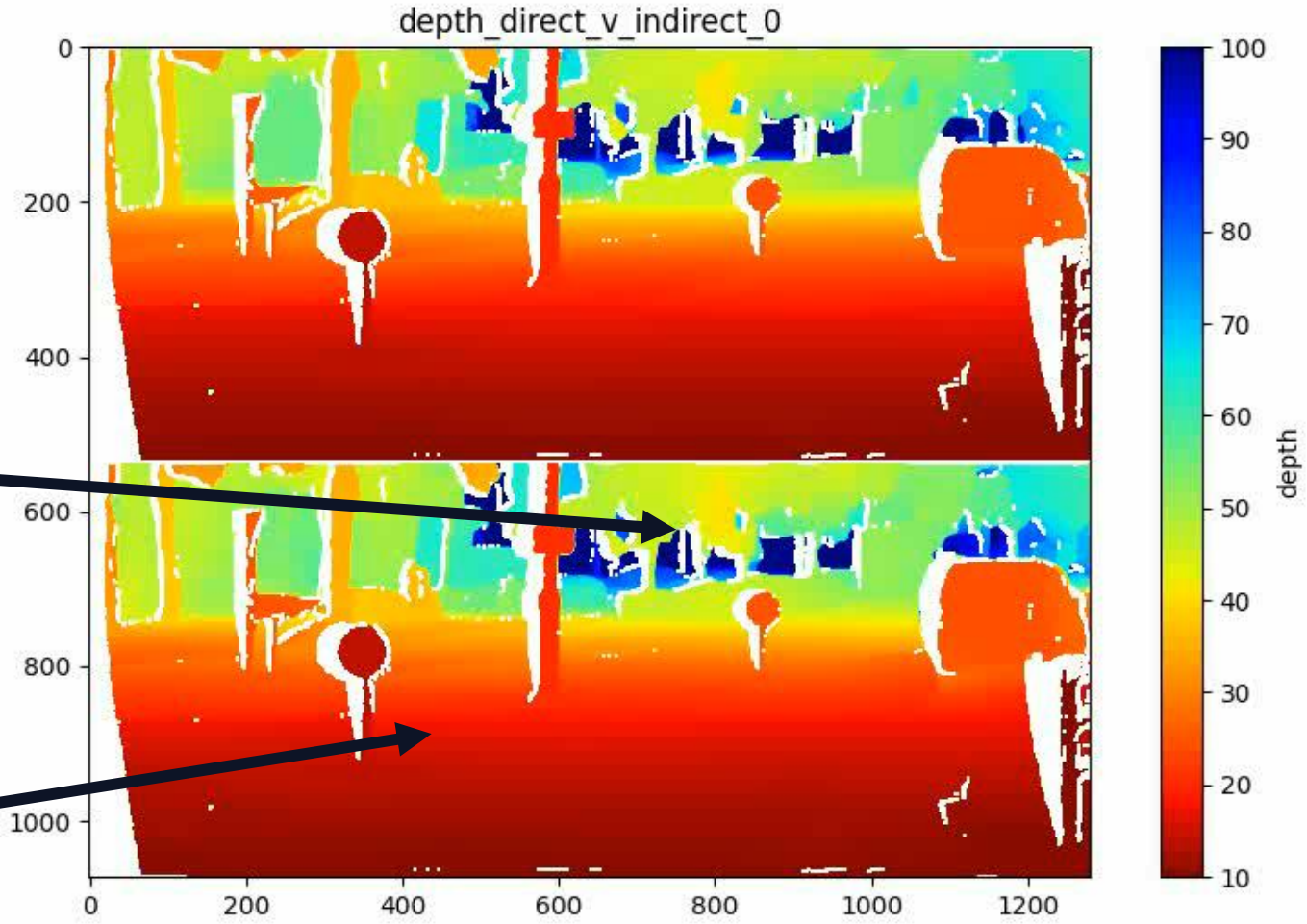


Sequence examples

- Forward motion
- Dynamic objects
- Trees as difficult objects to match



Sequence comparison



Direct Stereo Depth

Trees jitter less in Indirect Stereo

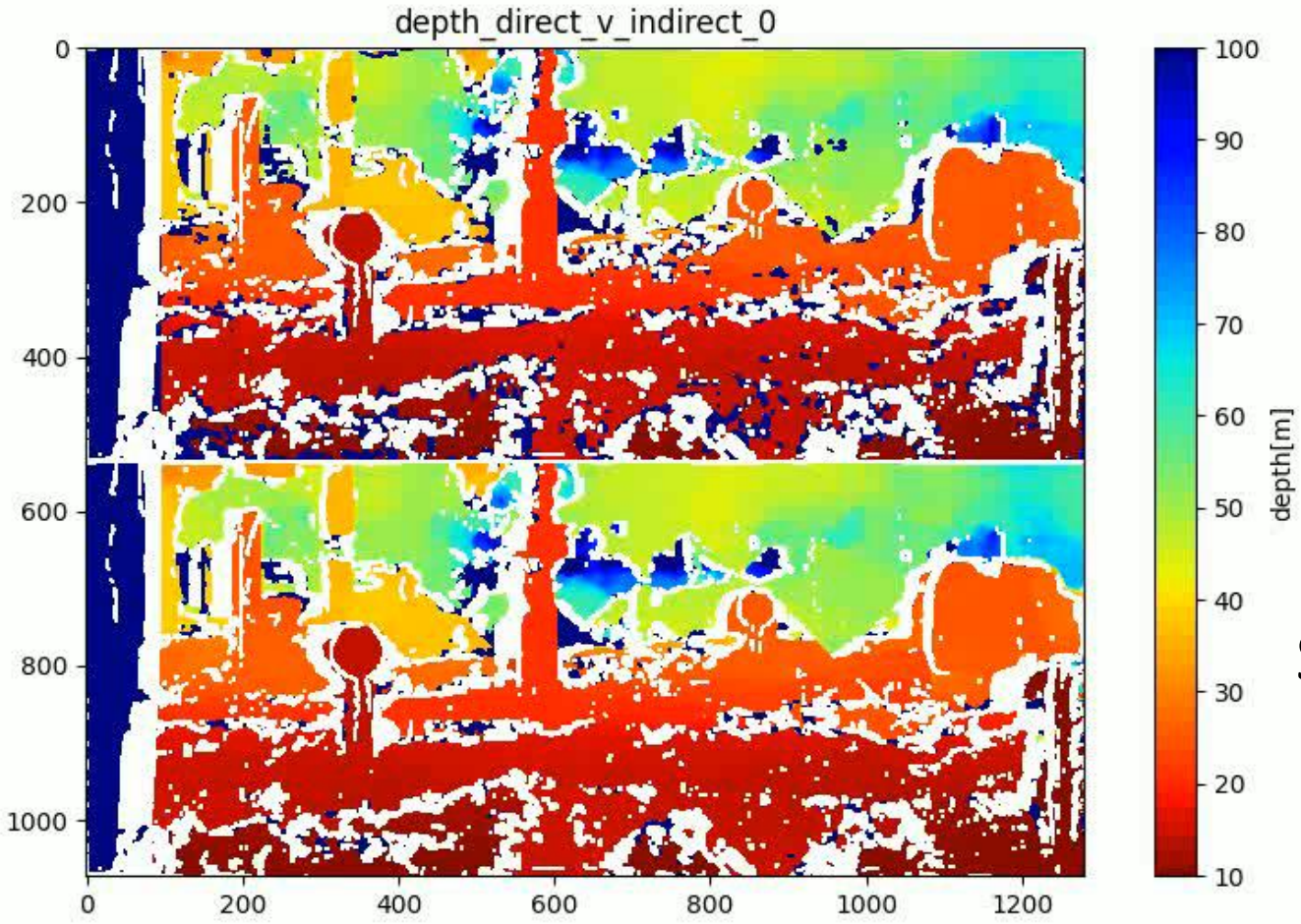
Indirect Stereo Depth

Occlusion

Tracked, expands

Example – Disparity by OpenCV's SGBM

- Method easily switched disparity method
- SGBM filter jitters in Direct estimations

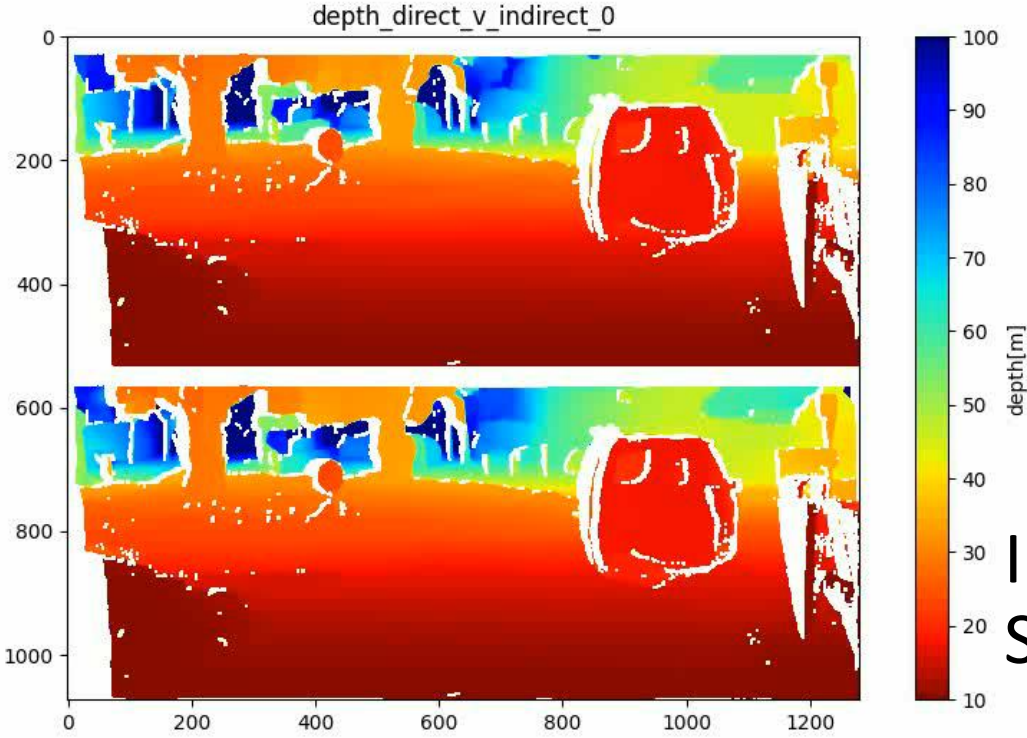


Direct Stereo Depth

Indirect Stereo Depth

Example - Bumper

- Sky is filtered because it is not in the first frames
- More Tree jitters

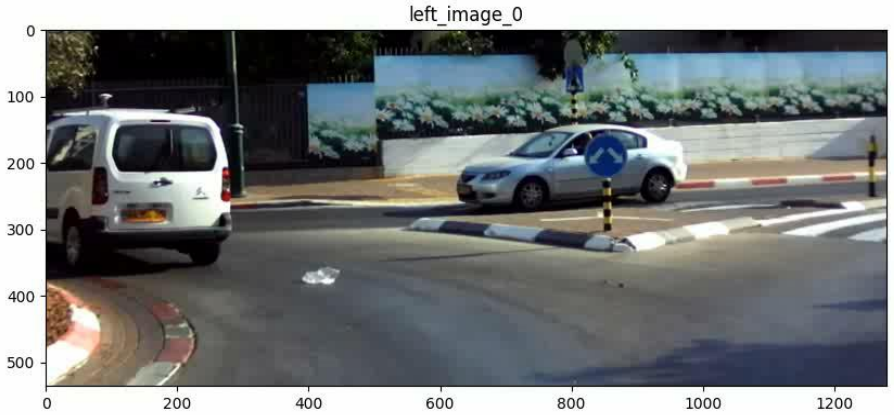


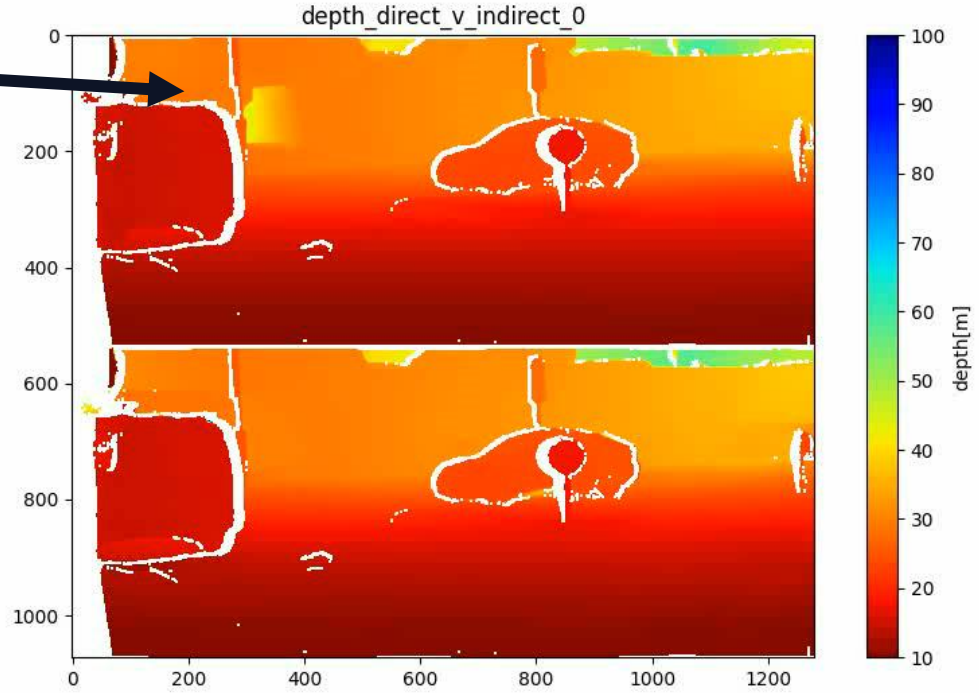
Direct Stereo Depth

Indirect Stereo Depth

Example - Bumper

- Decision on Fence is jittering only on direct Stereo





Direct Stereo Depth

Indirect Stereo Depth

Summary

- Consistency oriented models may be used to avoid jittering
- Tracking objects in both images allows maintaining previous disparity decisions
- Individual temporal updates is a consistency-oriented method that can be built from any optical flow methods
- Accurate for systems with good optical flow capabilities