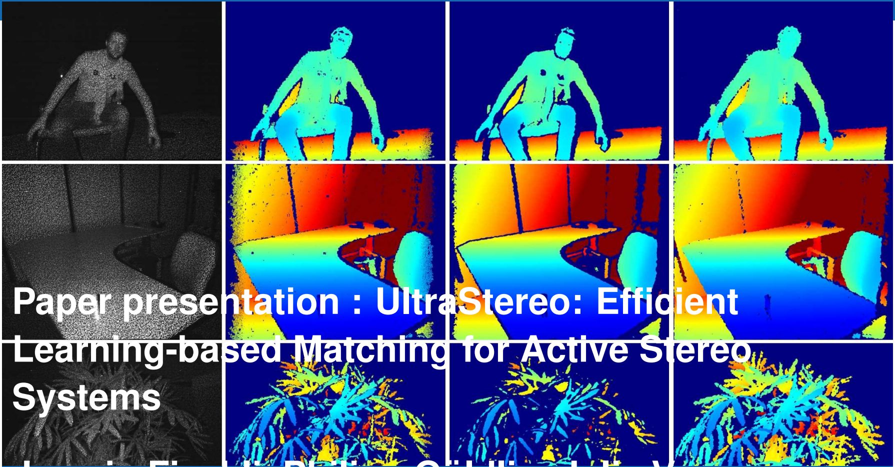


Paper presentation : UltraStereo: Efficient Learning-based Matching for Active Stereo Systems

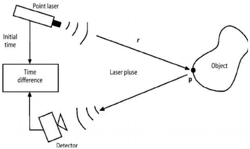
Jasmin Fischli, Philipp Göldlin, Julie Veya, Jonathan Burkhard

Supervisor: Sattler Torsten

Table of contents

Time of flight

Varying active illumination to reconstruct a scene from multiple images

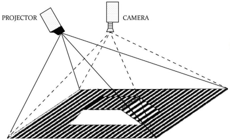


Shortcomings:

- Low mapping rate (30Hz)
- Motion artifacts
- Multipath-interference

Structured Light

Spatial patterns projected on the scene

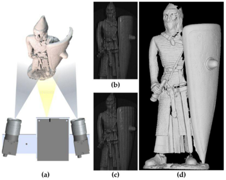


Shortcomings:

- Robustness issues
- Motion artifacts
- Limited depth-range
- Overlapping sensors cause interference

Active Stereo

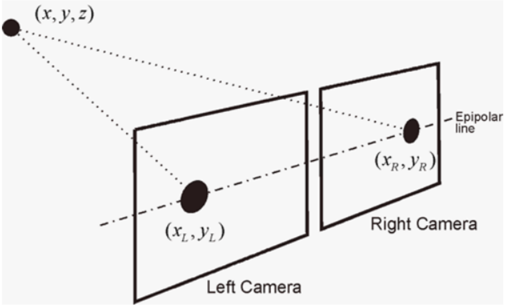
Using 2 calibrated cameras and 1 light source - solves a lot of the mentioned issues



- Mitigates multipath reflections
- Improves robustness
- Avoids interference between multiple systems
- BUT correspondence search has high computational cost!

UltraStereo

- 2 IR cameras: monochrome Ximea cameras, 1280x1024 pixel at 210Hz
- DOE projector



$$disparity = \frac{bf}{x_L - x_R} \quad (1)$$

f = focal length

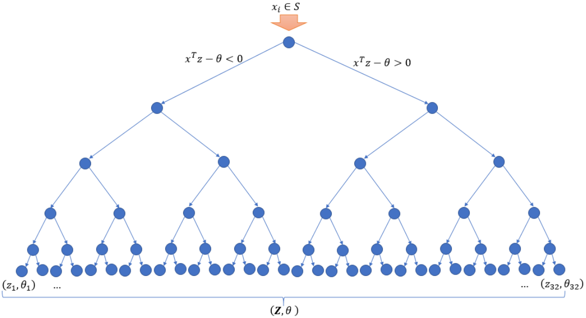
b = baseline

Dimensionality Reduction

$$y = sign(x^T Z - \theta)$$

$x \in R^W$ with W = 11x11
$Z \in R^b$ with b = 32
$\theta \in R^b$ $y \in (0, 1)^b$

How to chose the parameters



- k« W -> sparse hyperplanes
- Information gain: $I(\delta) = H(S) - \sum_{d \in L,R} \frac{S_d(\delta)}{S} H(S_d(\delta))$

δ: set of learned parameters (z, θ)

Matching Framework

- Based on PatchMatch framework
- Iterative and randomized process
- 3 steps: initialization, propagation, post-processing

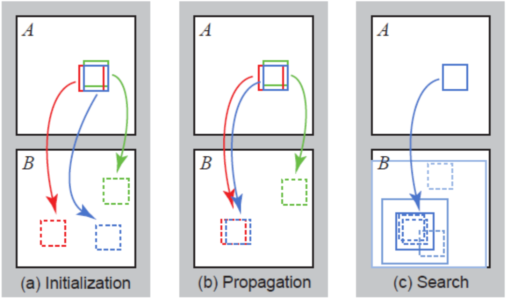


Figure 2: *Phases of the randomized nearest neighbor algorithm: (a) patches initially have random assignments; (b) the blue patch checks above/green and left/red neighbors to see if they will improve the blue mapping, propagating good matches; (c) the patch searches randomly for improvements in concentric neighborhoods.*

Computational Analysis

- Linear with respect to the image size
- O(1) complexity of the binary representation

Evaluation

- 2500 training images and 900 test images (500 articulated hand images and 400 interior images for 5 different environments)

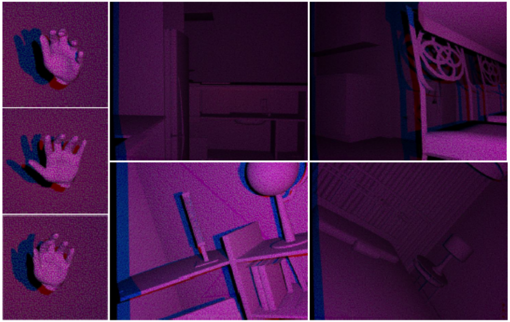


Figure 2. **Synthetic Data.** Representative examples of our synthetic dataset which comprises several indoor environments and a hand sequence.

Bias

- Average absolute depth error present in the whole set

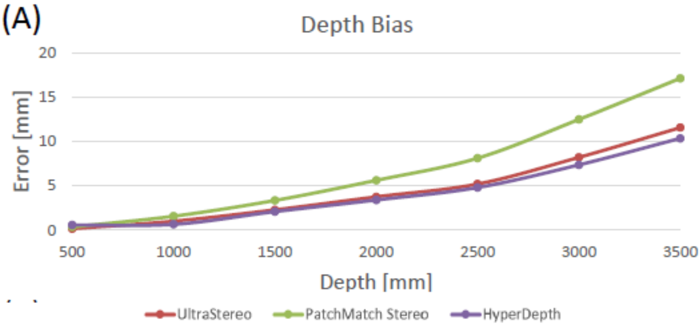


Figure: synthetic data

Bias

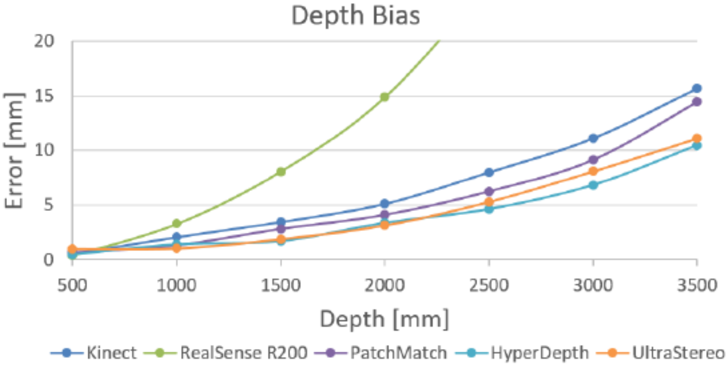


Figure: real data

Invalidation

Issues which the pixels from the final depth image won't contain any estimations :

- Occlusions
- Saturation of infrared sensors
- Low signal noise ratio

How limits this errors ?

- An algorithms does an invalidation pass durint the post-processing step

Invalidation

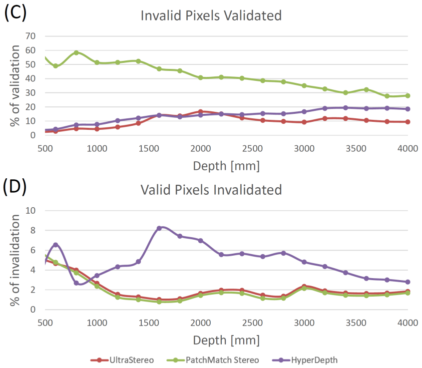


Figure: Quantitative results on syntatic data

Example of depth-map produced with UltraStereo

Look at the thin structures like plants

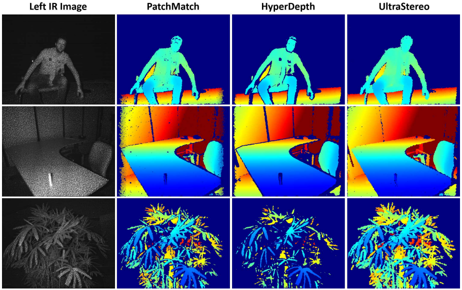


Figure: Qualitative Evaluation

Edge fattening

Other issiue is the edge fattening. To measure it they :

- Used a hand to test their algorithms
- Put a hand at 1 m from the sensors
- Defined key hand pose for each frame

Edge fattening

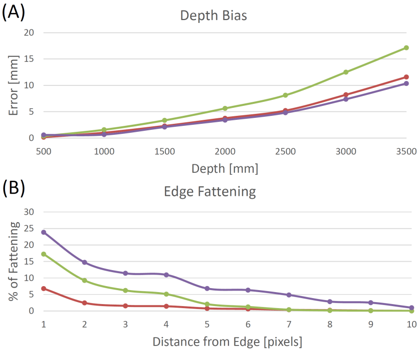


Figure: Quantitative results on syntatic data

Edge fattening

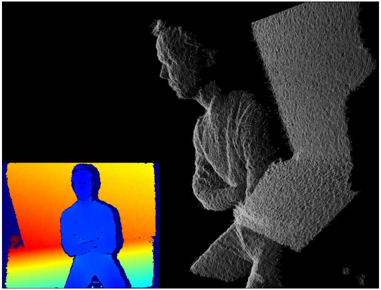


Figure: Example of pointcloud produced with our algorithm. Notice the absence of quantization and flying pixels

Binary representation

- Compare UltraStereo with Census and Locality Sensitive Hashing (LSH)
- Collect 1000 images with the Kinnect

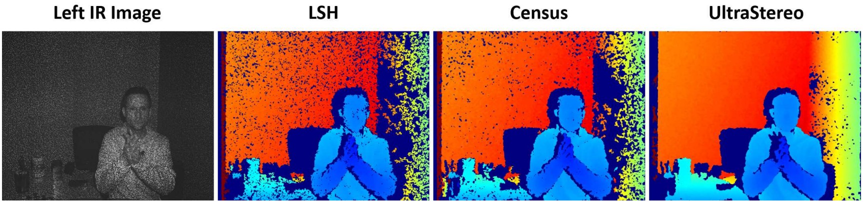


Figure: Census use 121 bits, LSH and UltraStereo use only 32 bits

Interference and Generalization

- Interference caused by multiple sensors
- Camera calibrations

Interference and Generalization

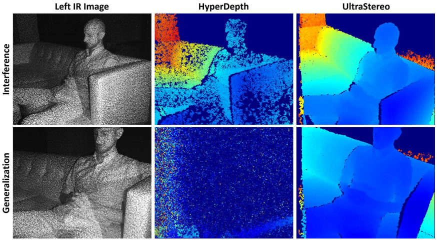


Figure: Examples of depth-maps produced with UltraStereo and state of the art competitors.

Conclusion

- Breakthrough in the field of active stereo depth estimation
- Does not depend of the windows size or the size of the disparity space
- Use machine learning algorithms
- Run on GPU
- Does not suffer from camera calibrations nor interference problems

Questions

?

Thank you for your attention !