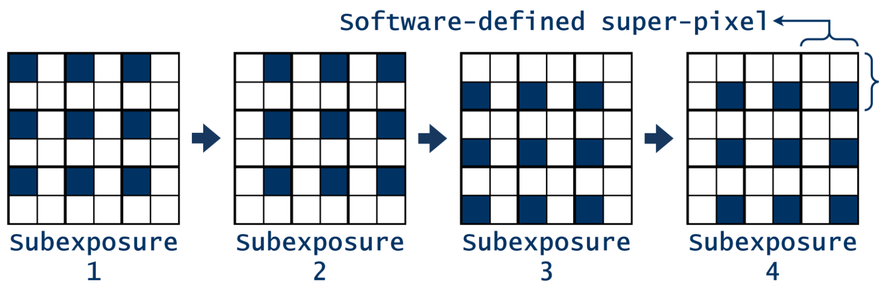
Multi-View Synthesis with Dual-Tap Coded-Exposure Camera Array

M Ahmed Mansoor[1], Jainish Solanki[2], Karanpreet Raja[3]

1 Department of Electrical and Computer Engineering, University of Toronto
2 University of Toronto Institute of Aerospace Studies
3 Department of Electrical Engineering and Computer Science, York University

Motivation

- Most multi-view capture pipelines employ standard frame rate cameras because high-speed camera systems are typically bulky, power inefficient, and difficult to synchronize as arrays, so true high-fps multi-view capture for novel-view synthesis remains underexplored
- We propose the use of **dual-tap coded-exposure cameras** [1] that can **trade spatial resolution for temporal resolution** on demand, enabling synchronized capture at up to 23000 subexposures per second, but without the power, cost, and bulk overhead of traditional high-speed cameras



Related Work

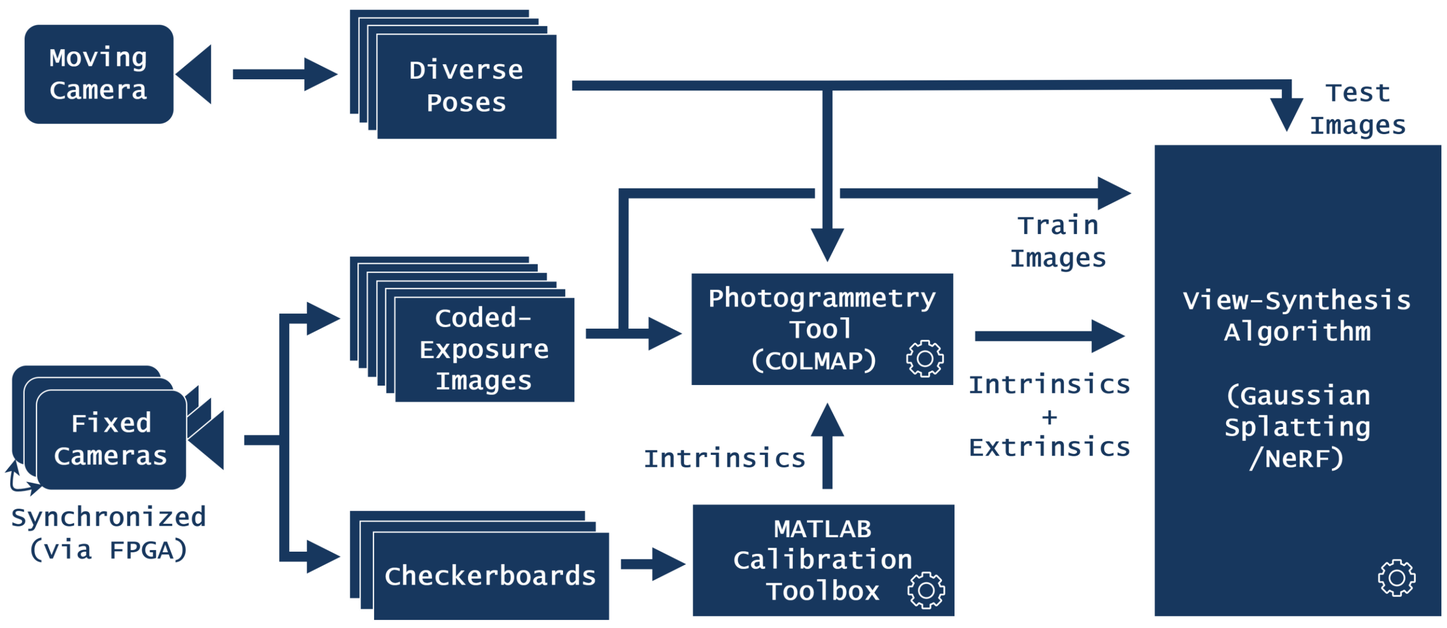
- Time aware NeRF implementations like K-planes [2] and 4D Gaussian Splatting [3] can produce bullet-time and limited slow motion from ordinary frame rate captures. However, when motion outpaces the frame rate, the results degrade.
- Event-based sensing has also been explored [4], but requires specialized pipelines.

References

- [1] R. Rangel, et. al., “23,000-Exposures/s 360fps-Readout Software-Defined Image Sensor with Motion Adaptive Spatially Varying Imaging Speed,” in Proc. IEEE Symp. on VLSI Technology and Circuits, 2024.
- [2] G. Fridovich-Keil et. al., “K-Planes: Explicit Radiance Fields in Space, Time, and Appearance,” in CVPR, 2023
- [3] G. Wu et. al., “4D Gaussian
- Splatting for Real-Time Dynamic Scene Rendering,” in CVPR, 2024.
- [4] V. Rudnev et. Al., “Neural Radiance Fields from
- a Single Colour Event Camera,” in ECCV, 2022

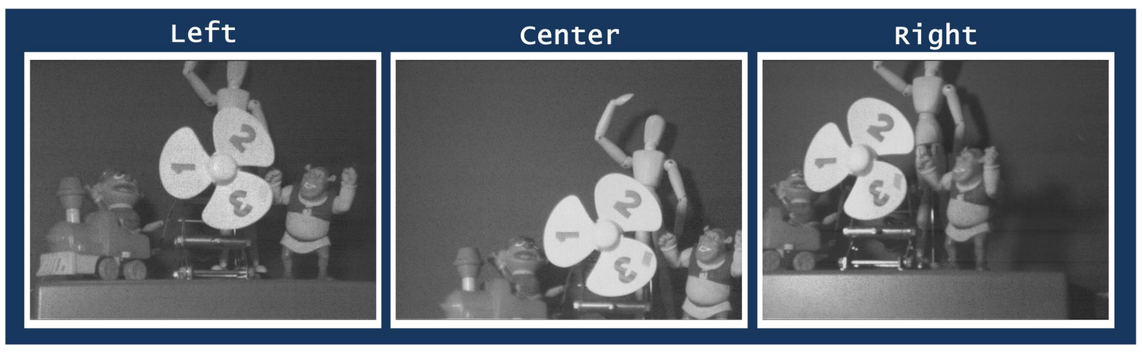
Methodology

- We synchronize an array of 3 coded-exposure cameras by connecting the on-board FPGAs in a single-leader / multi-follower architecture.
- A fourth unsynchronized and moving camera is used to capture more diverse poses for bundle adjustment and as test images
- Camera calibration is performed through traditional means and refined through modern structure-from-motion photogrammetry tools
- Static novel-view synthesis (NeRF / 3DGS) is used as a sanity check before moving on to dynamic scenes (DNeRF / 4DGS)



Results

- Synchronized coded-exposure images from a 3-camera array were successfully captured



- Static novel-view synthesis was performed through both a NeRF and 3D Gaussian Splatting, with the 3DGS producing better results. Dynamic novel-view synthesis is still in progress.
- Results were compared for MATLAB+COLMAP against COLMAP alone, with the later producing better results. Only these are shown below: