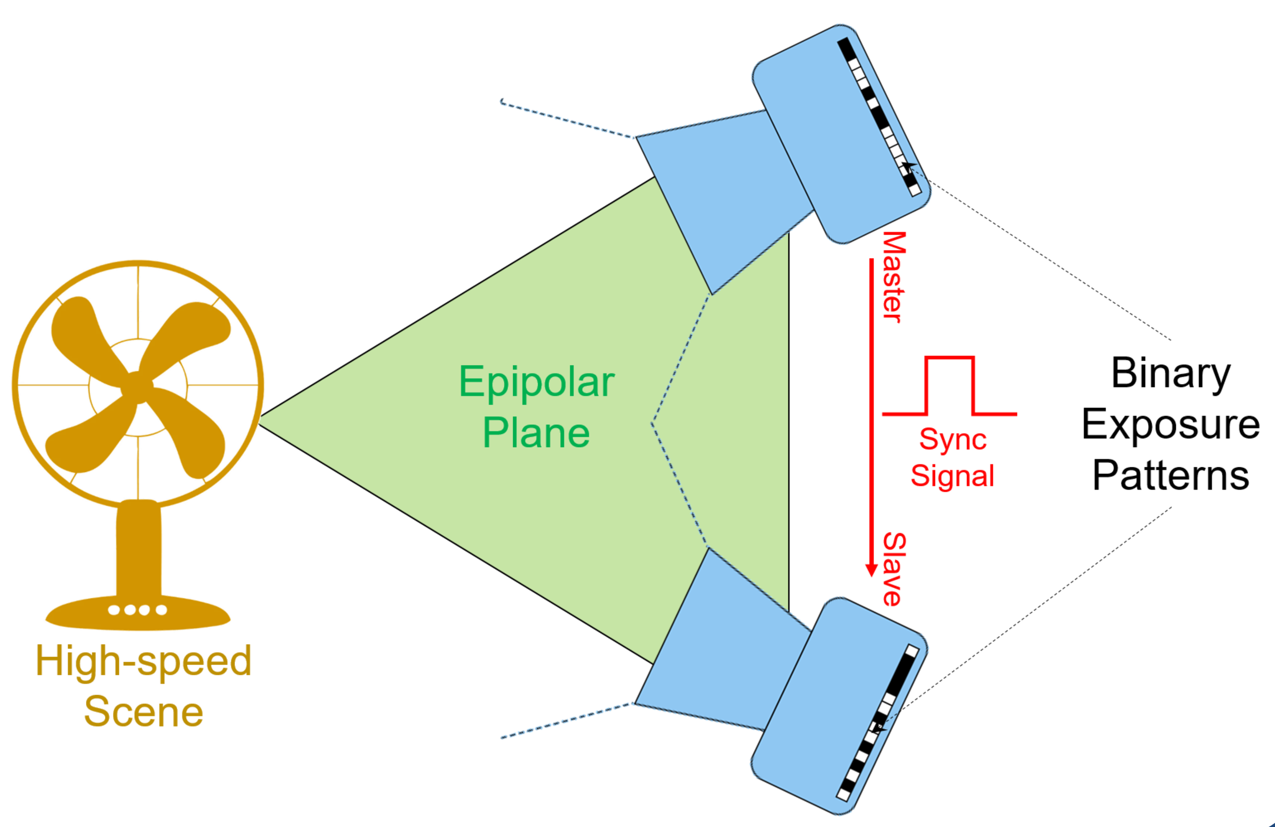
High-Speed Multi-Camera Depth Estimation

Hao Yang, Chu King Kung
Electrical and Computer Engineering (ECE), University of Toronto

Motivation

- The integration of high-speed imaging and depth estimation for real-time applications remains underexplored despite significant advancements in both fields.
- This fusion presents critical challenges:
  - i. **Computational Complexity**: Frame-based stereo methods struggle with agile motion scenarios, such as fast flying or immediate obstacle avoidance, failing to meet real-time performance requirements.
  - ii. **Hardware Constraints**: High-speed cameras often require high-performance computing resources, which are impractical for applications like autonomous flight or mobile robotics due to payload and power limitations.
  - iii. **Synchronization**: Precise synchronization of high-speed cameras at elevated frame rates remains a significant challenge.
- This work explores the use of pixel-wise coded exposure cameras [1] for high-speed imaging, demonstrating their seamless integration with conventional depth-estimation algorithms.
- The proposed system addresses the problem of high-speed depth-estimation, offering a cost-effective and efficient solution.



Related Work

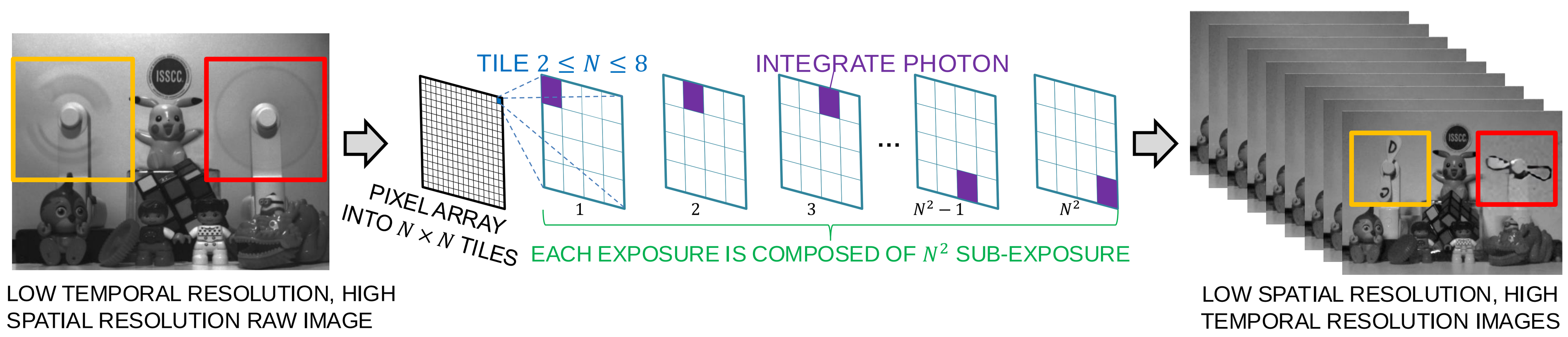
- Time-of-flight cameras and other active sensors offer remarkable results but suffer from limited range.
- Traditional frame-based cameras and algorithms face latency issues and struggle with agility during rapid movements.
- [2], [3] introduced an event-driven solution for stereo depth estimation to achieve real-time efficiency compared to traditional methods.
- However, synchronization is a crucial challenge when using event sensors for high-speed stereo depth estimation, due to event timestamp inaccuracy, event rate differences and motion-induced asynchrony.

References

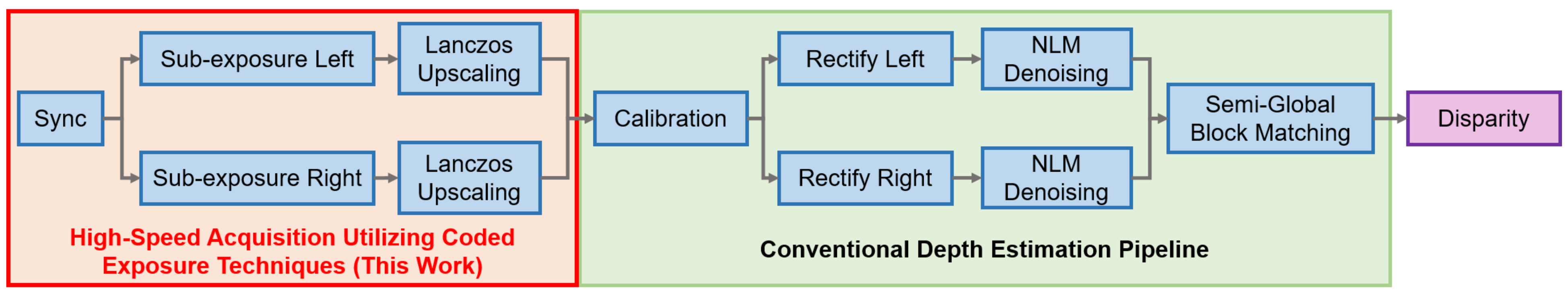
[1] R. Rangel et al., "23,000-Exposures/s 360fps-Readout Software-Defined Image Sensor with Motion-Adaptive Spatially Varying Imaging Speed," VLSI 2024.
[2] J.N. Martel et al., "An active approach to solving the stereo matching problem using event-based sensors," ISCAS 2018.
[3] Kim et al. "Real-time hetero-stereo matching for event and frame camera with aligned events using maximum shift distance." IEEE Robotics and Automation Letters 2022.

New Technique

- The ability to configure exposure patterns at the pixel level enables us to implement a variety of techniques that facilitate high-speed imaging at a relatively low cost compared to traditional high-end cameras.
- One effective technique is multiplexing each exposure temporality, which allows us to trade off spatial resolution for increased speed [1]:



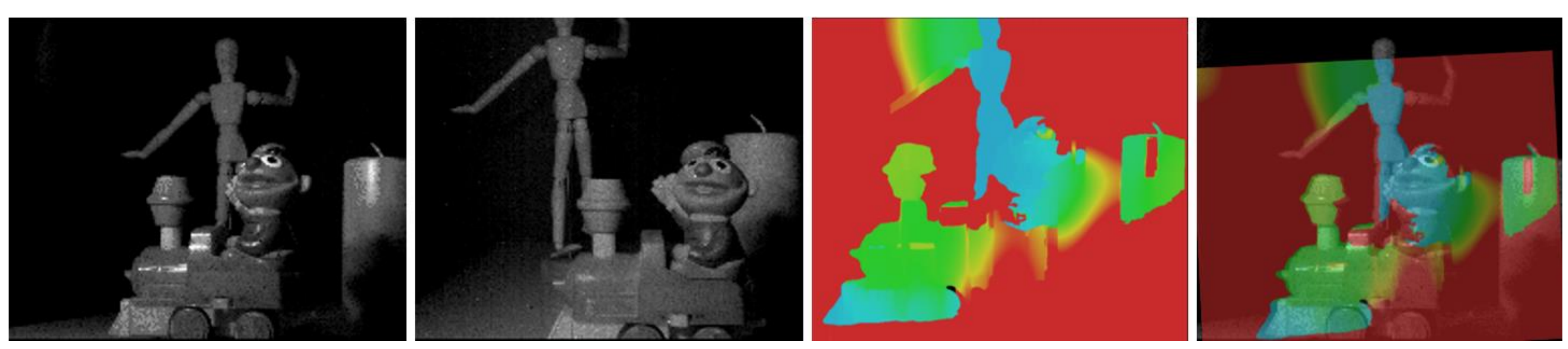
- Subsequently, we can input the high-speed images captured by our cameras into a conventional processing pipeline for depth estimation.
- The block diagram below outlines the pipeline we utilize to generate sub-frame-level disparity maps.



- One of the primary challenges in high-speed depth estimation is achieving synchronization among multiple cameras operating at high speeds.
- To address this issue, we utilize hardware techniques that enable the configuration of multiple cameras in a master-slave arrangement, allowing them to work in synchronization at the sub-frame level.

Experimental Results

- The diagram below illustrates an exemplary depth estimation result.
- From left to right, the components are as follows: the left grayscale image, the right grayscale image, the corresponding depth map, and the depth map overlaid on the right grayscale image.



- The diagrams above represent the results obtained from a static scene.
- In contrast, the subsequent diagrams display the results from three dynamic scenes (with a sub-frame rate of ~350 fps):
  a) A fan rotating at 300 rpm.
  b) A toy car being pushed toward the screen.
  c) A candle being rolled toward the screen.