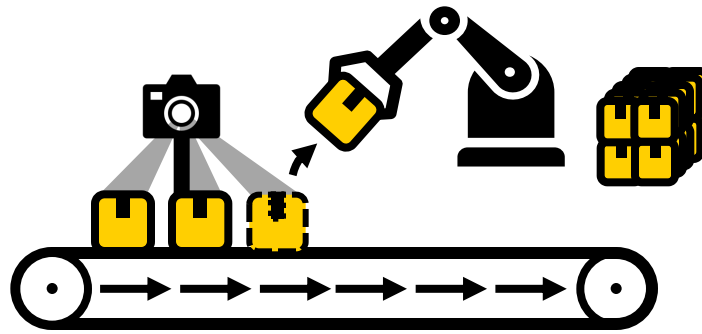# Cross-Spatial Fusion and Dynamic-Range Distribution Based Accelerated Temporal FPGA-GPU Architecture for Ultra-Low Delay Object Pose Estimation

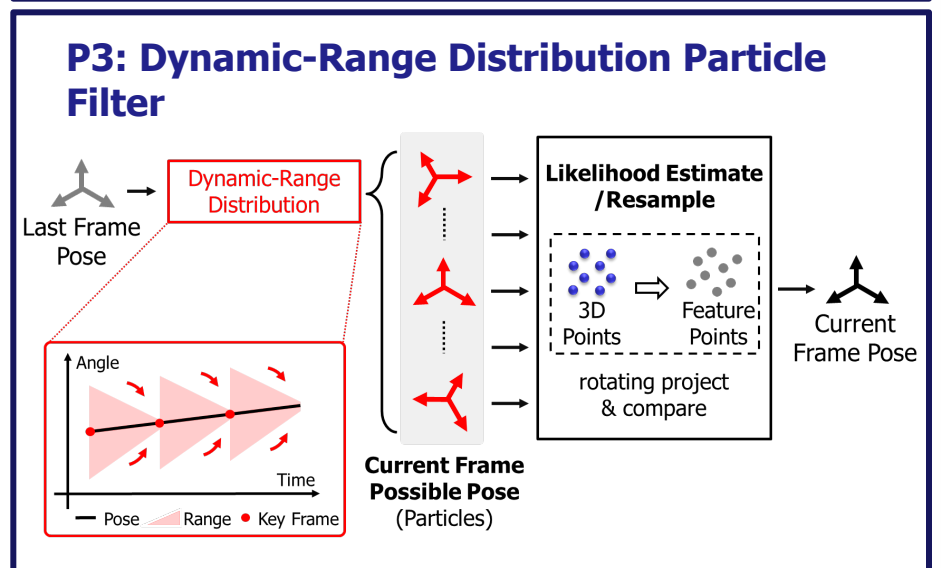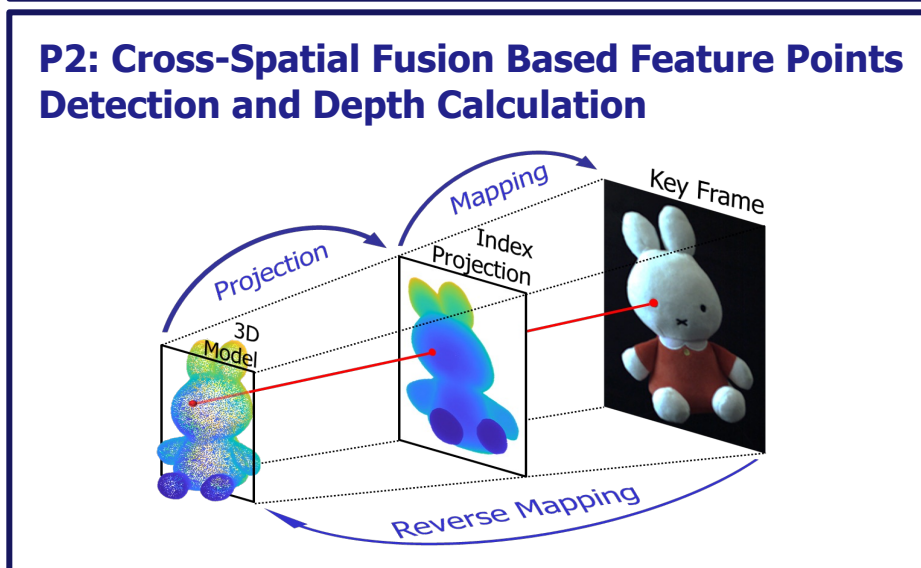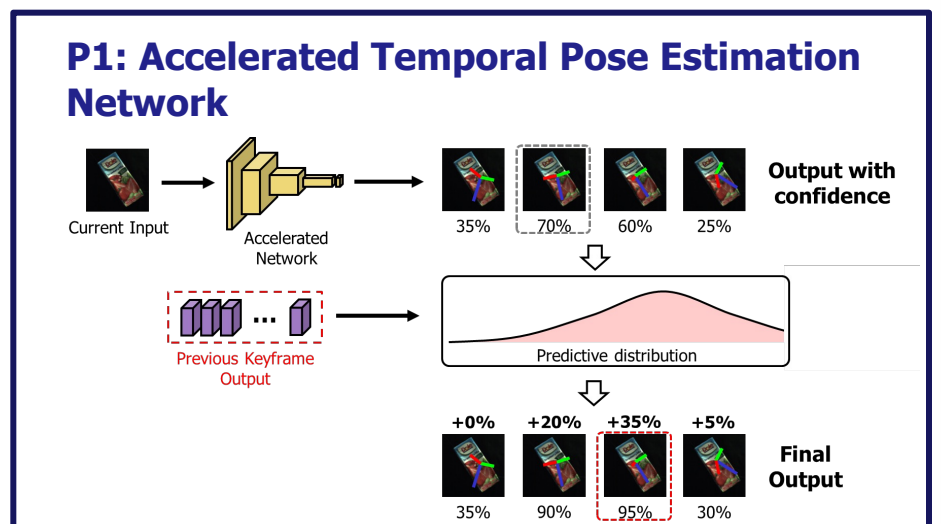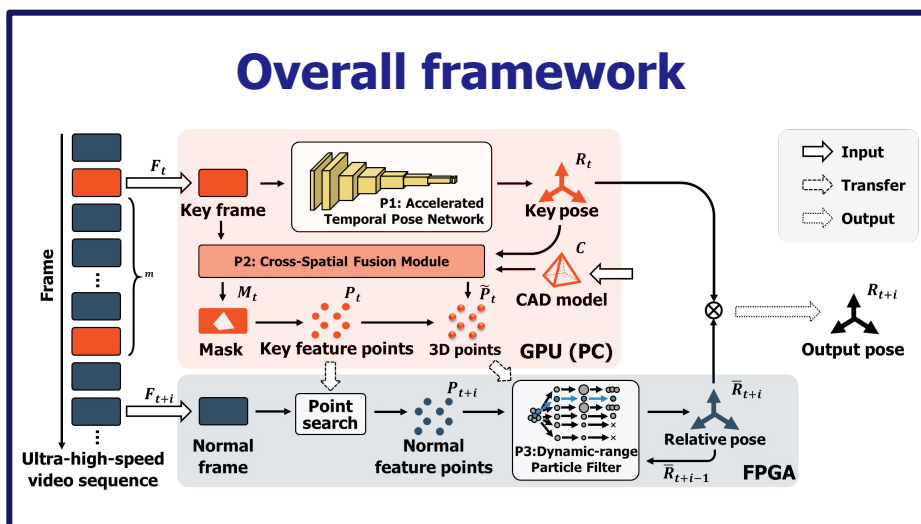王為　池永研究室　修士課程修了

➤ Background

- **Application**
  - Components assembly
  - Cargo handling
  - Quality inspection
  - ...

- **Target**
  - Achieve more than 85% tracking success rate in less than 1ms/frame
- **Challenge**
  - Balance high accuracy and high speed with limited hardware resources
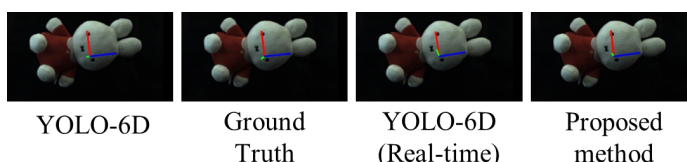
➤ Proposed method

Overall framework

P1: Accelerated Temporal Pose Estimation Network

P2: Cross-Spatial Fusion Based Feature Points Detection and Depth Calculation

P3: Dynamic-Range Distribution Particle Filter

➤ Experiments Result

- **Software performance (Real-time)**

Method	Object									FPS
	bear	bird	box	dog	donut	rabbit	statue	train	Avg.	
YOLO-6D	100.00	43.91	27.73	17.16	19.69	62.77	12.54	9.56	36.67	1000
The proposed method	100.00	100.00	100.00	64.37	97.07	100.00	65.31	58.89	85.71	

- **Visualization Result**

YOLO-6D　　Ground Truth　　YOLO-6D (Real-time)　　Proposed method

- **Hardware performance**

FPGA Logic Utilization	LUT	198589 (86.19%)
	LUTRAM	8004 (7.87%)
	FF	268280 (58.22%)
	DSP	203 (11.75%)
Speed	Frequency	300 MHz
	Process time	0.927 ms/Frame

➤ Conclusion

- The proposed method achieves an accuracy of 85.71% and the running speed on the hardware platform is 0.927 ms/frame, which reaches 1000 FPS.

Graduate School of Information, Production and Systems
Waseda University