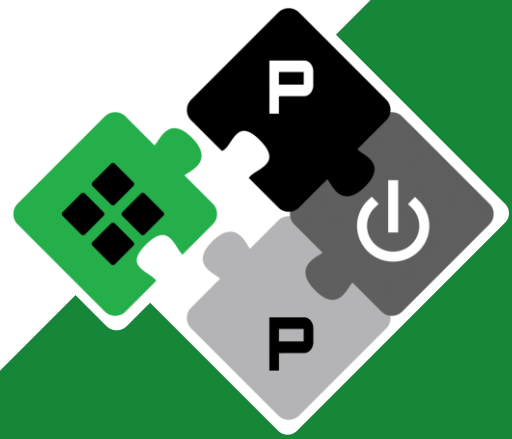
A Fast and Accurate Optical Flow Camera for Resource-Constrained Edge Applications

Integrated Systems Laboratory (ETH Zürich)
Center for Project-Based Learning (ETH Zürich)

Jonas Kühne kuehnej@ethz.ch
Michele Magno
Luca Benini



@pulp_platform
pulp-platform.org
youtube.com/pulp_platform

PULP Platform
Open Source Hardware, the way it should be!

Motivation: Accelerate Computer Vision Tasks

- **Optical Flow is the movement of features from on frame to another**
  - Can be used to determine the ego motion of an agent
  - Can be used to track objects or keypoints

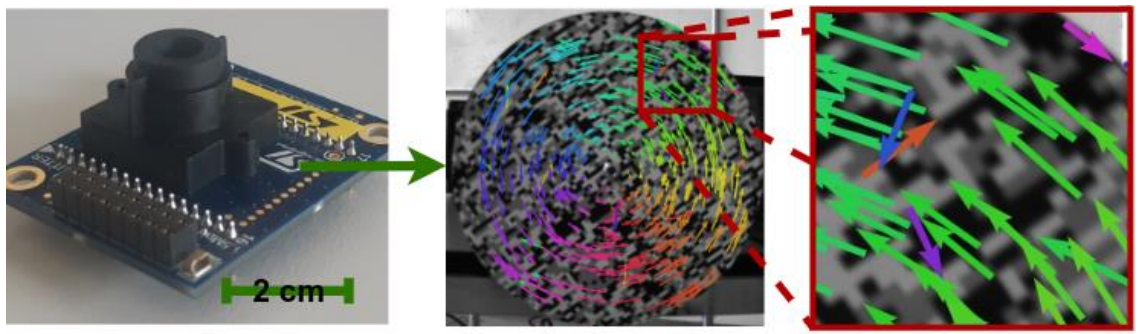
- **Optical Flow calculation is computationally demanding**
  - Often executed on powerful CPUs or GPUs
  - Implementation on small mobile robots or AR/VR glasses is challenging

Vision

- **Hardware acceleration of computer vision tasks like Optical Flow calculation**
  - Enable low latency Optical Flow prediction
  - Enable the use on power- and compute constrained robots (e.g., Nano-UAVs)

Our Contribution

- **Characterization of the VD56G3 sensor by STMicroelectronics**
  - Global shutter grayscale camera 1364 by 1124 pixels
  - Integrated ASIC for the prediction of sparse optical flow
  - Up to 300 frames per second (at reduced resolution)



- **Sensor Characterization**
  - Evaluation of different parameter sets
  - Quantitative and qualitative analysis of tracking accuracy for simple movements
  - Analysis of consistent redetection of distinctive features
  - Analysis of power requirements

Use Case: Localization

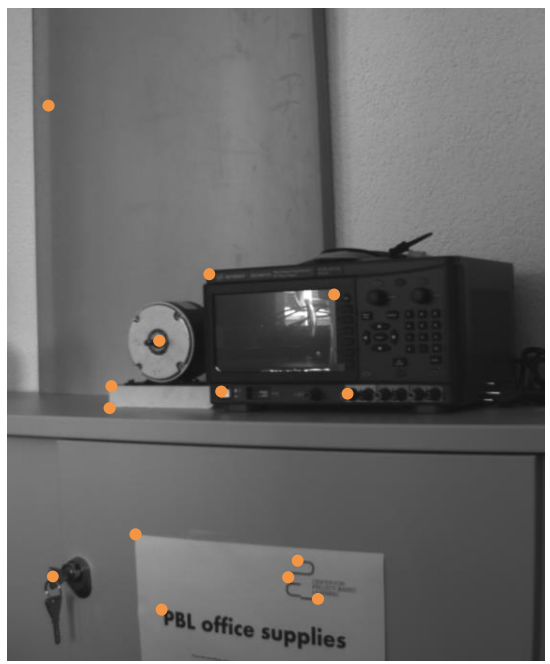
- **Broad adoption of autonomous mobile robots**
  - Service and delivery robots
  - Search and Rescue
  - Surveillance and monitoring

- **Autonomous mobile robots need to self-localize**
  - Assume absence of external systems (e.g., GPS, UWB beacons)
  - → **Use visual sensors**

- **Optimization of existing algorithms**
  - Enable lower latency with constant compute requirements
  - Support power- and compute constrained robots (e.g., Nano-UAVs)

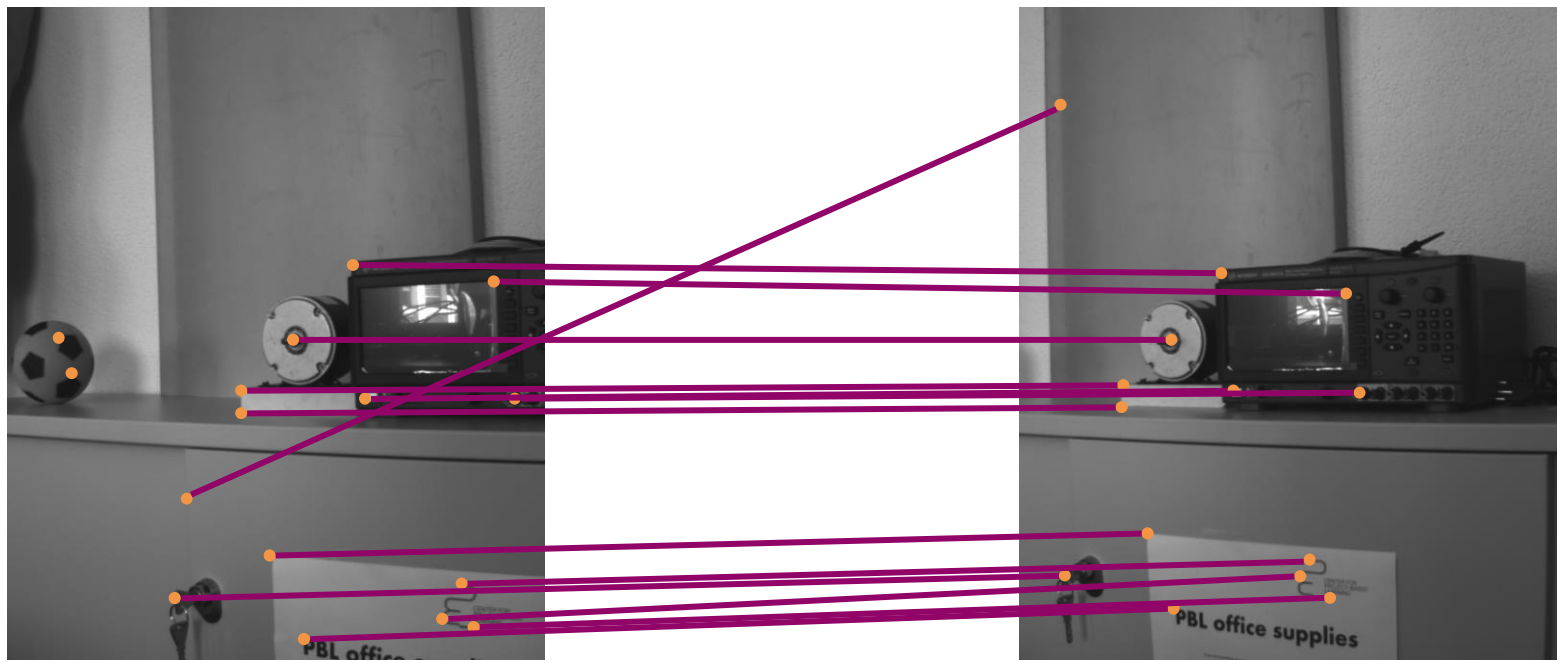
Use Case: Visual Inertial Odometry

- **Reconstruct movement trajectory from visual landmarks**
  - Detect features



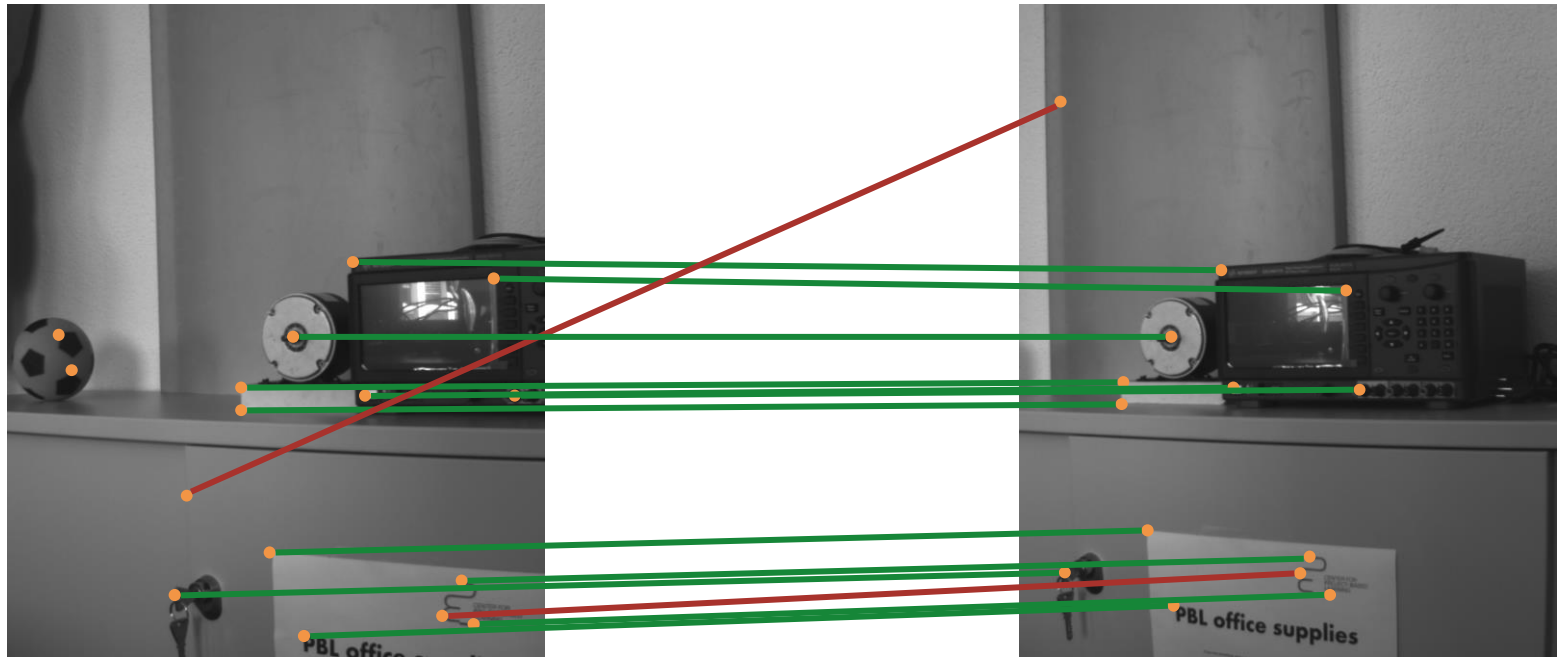
Use Case: Visual Inertial Odometry

- **Reconstruct movement trajectory from visual landmarks**
  - Detect features
  - Match features



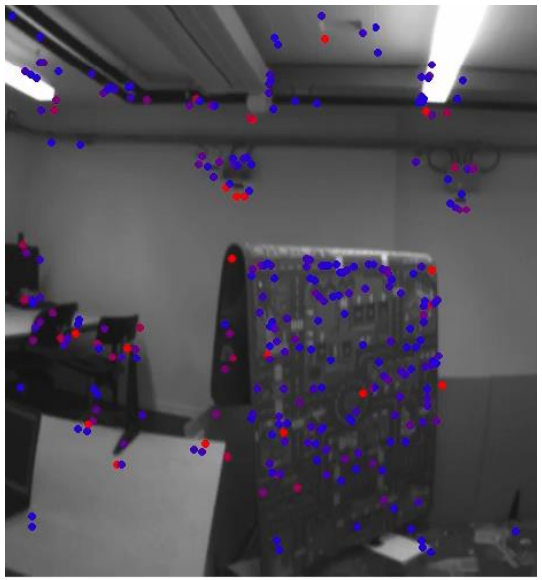
Use Case: Visual Inertial Odometry

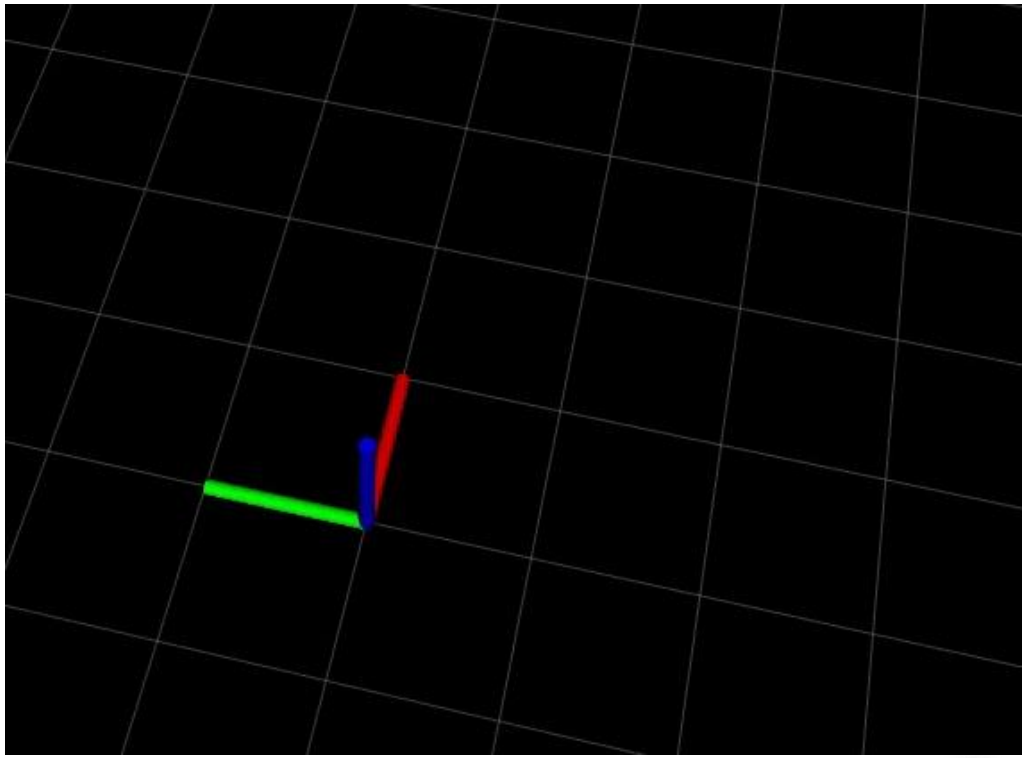
- **Reconstruct movement trajectory from visual landmarks**
  - Detect features
  - Match features
  - Estimate viewpoint difference as rotation and translation
  - Refine estimate



Use Case: Visual Inertial Odometry

- **Value of a sensor like VD56G3**
  - Solves the detection and matching of features
  - The remaining steps would still run on a CPU



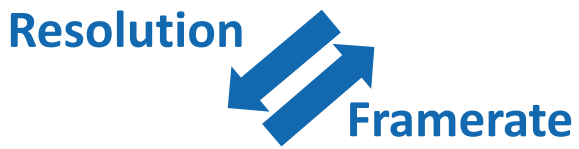


VD56G3 Sensor in Depth

- **Product by STMicroelectronics**
  - MIPI CSI-2 interface
  - Variable resolution (up to 1364 by 1124)
  - Variable frame rate (up to 300 FPS)
  - Supports cropping and subsampling
  - Optical Flow prediction can be disabled
  - Desired number of Optical Flow vectors can be set
    → Technically the number of descriptors

TABLE I
ACHIEVABLE FRAME RATES FOR A GIVEN FRAME HEIGHT AND A GIVEN NUMBER OF OPTICAL FLOW VECTORS WHEN USING THE 10-BIT ADC MODE.

Format	Frame Height [pixel]	# OF Vectors	Frame Rate [1/s]
QVGA	240	1024	338
QVGA	240	2048	288
VGA	480	0	229
VGA	480	1024	205
VGA	480	2048	186
FULL	1364	0	88
FULL	1364	1024	84
FULL	1364	2048	80

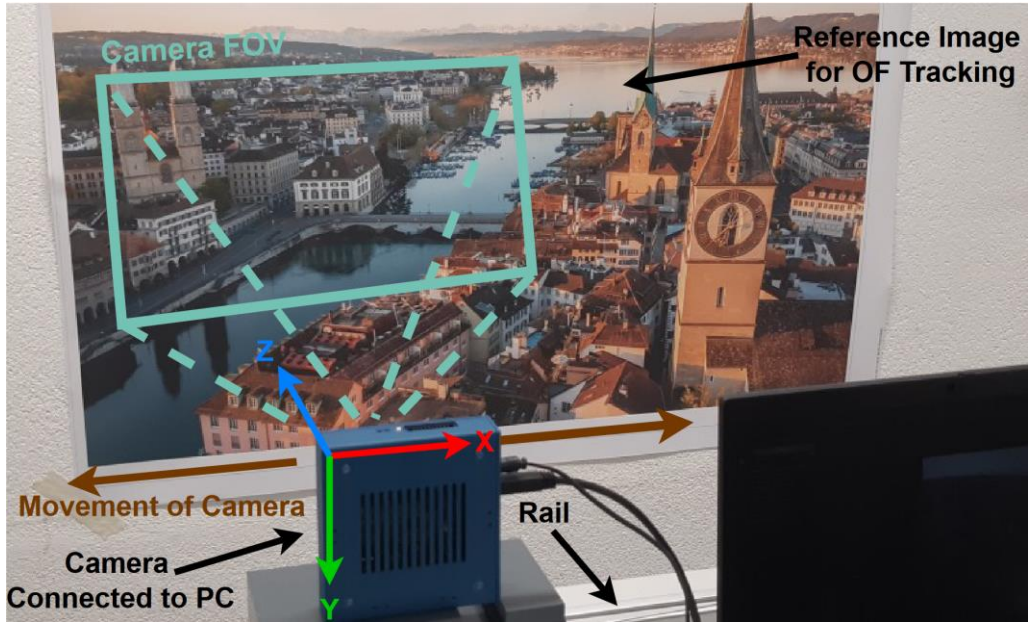


VD56G3 Evaluation Setup

TABLE II
PARAMETER SETS OF CAMERA SETTINGS THAT WERE APPLIED DURING THE RECORDING OF THE DATASET.

Parameter Set	Camera Resolution X	Camera Resolution Y	Crop[a]/Sub-sample	Frame Rate	BRIEF Target	BRIEF Max	OF Spatial Point
1	1124 pixel	1364 pixel	None	60 FPS	1536	2048	2
2	1120 pixel	1344 pixel	Crop (0,0)	60 FPS	1536	2048	2
3[b]	640 pixel	480 pixel	Crop (240,432)	140 FPS	768	1024	4
4	560 pixel	672 pixel	Crop (280,336)	140 FPS	768	1024	4
5	560 pixel	672 pixel	2x sub-sample	140 FPS	768	1024	4
6	272 pixel	336 pixel	Crop (420,504)	240 FPS	384	512	8
7	280 pixel	336 pixel	4x sub-sample	240 FPS	384	512	8

[a]If cropping is applied the numbers in brackets indicate the top left corner of the cropped image relative to the full 1124 by 1364 image.
[b]Operate the optical flow unit at its maximum resolution.



VD56G3 Evaluation Setup

TABLE II
PARAMETER SETS OF CAMERA SETTINGS THAT WERE APPLIED DURING THE RECORDING OF THE DATASET.

Parameter Set	Camera Resolution X	Camera Resolution Y	Crop[a]/Sub-sample	Frame Rate	BRIEF Target	BRIEF Max	OF Spatial Point
1	1124 pixel	1364 pixel	None	60 FPS	1536	2048	2
2	1120 pixel	1344 pixel	Crop (0,0)	60 FPS	1536	2048	2
3[b]	640 pixel	480 pixel	Crop (240,432)	140 FPS	768	1024	4
4	560 pixel	672 pixel	Crop (280,336)	140 FPS	768	1024	4
5	560 pixel	672 pixel	2x sub-sample	140 FPS	768	1024	4
6	272 pixel	336 pixel	Crop (420,504)	240 FPS	384	512	8
7	280 pixel	336 pixel	4x sub-sample	240 FPS	384	512	8

[a]If cropping is applied the numbers in brackets indicate the top left corner of the cropped image relative to the full 1124 by 1364 image.
[b]Operate the optical flow unit at its maximum resolution.

- **Localization**

  - Utilize full field of view
  - Trade off frame rate and sub sampling factor (accuracy) depending on movement speed

VD56G3 Evaluation Setup

TABLE II
PARAMETER SETS OF CAMERA SETTINGS THAT WERE APPLIED DURING THE RECORDING OF THE DATASET.

Parameter Set	Camera Resolution X	Camera Resolution Y	Crop[a]/Sub-sample	Frame Rate	BRIEF Target	BRIEF Max	OF Spatial Point
1	1124 pixel	1364 pixel	None	60 FPS	1536	2048	2
2	1120 pixel	1344 pixel	Crop (0,0)	60 FPS	1536	2048	2
3[b]	640 pixel	480 pixel	Crop (240,432)	140 FPS	768	1024	4
4	560 pixel	672 pixel	Crop (280,336)	140 FPS	768	1024	4
5	560 pixel	672 pixel	2x sub-sample	140 FPS	768	1024	4
6	272 pixel	336 pixel	Crop (420,504)	240 FPS	384	512	8
7	280 pixel	336 pixel	4x sub-sample	240 FPS	384	512	8

[a]If cropping is applied the numbers in brackets indicate the top left corner of the cropped image relative to the full 1124 by 1364 image.
[b]Operate the optical flow unit at its maximum resolution.

- **Eye Tracking** 
  - Narrow region of interest
  - Fast frame rates required

VD56G3 Evaluation Setup

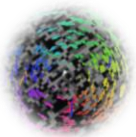
TABLE II
PARAMETER SETS OF CAMERA SETTINGS THAT WERE APPLIED DURING THE RECORDING OF THE DATASET.

Parameter Set	Camera Resolution X	Camera Resolution Y	Crop[a]/Sub-sample	Frame Rate	BRIEF Target	BRIEF Max	OF Spatial Point
1	1124 pixel	1364 pixel	None	60 FPS	1536	2048	2
2	1120 pixel	1344 pixel	Crop (0,0)	60 FPS	1536	2048	2
3[b]	640 pixel	480 pixel	Crop (240,432)	140 FPS	768	1024	4
4	560 pixel	672 pixel	Crop (280,336)	140 FPS	768	1024	4
5	560 pixel	672 pixel	2x sub-sample	140 FPS	768	1024	4
6	272 pixel	336 pixel	Crop (420,504)	240 FPS	384	512	8
7	280 pixel	336 pixel	4x sub-sample	240 FPS	384	512	8

[a]If cropping is applied the numbers in brackets indicate the top left corner of the cropped image relative to the full 1124 by 1364 image.
[b]Operate the optical flow unit at its maximum resolution.

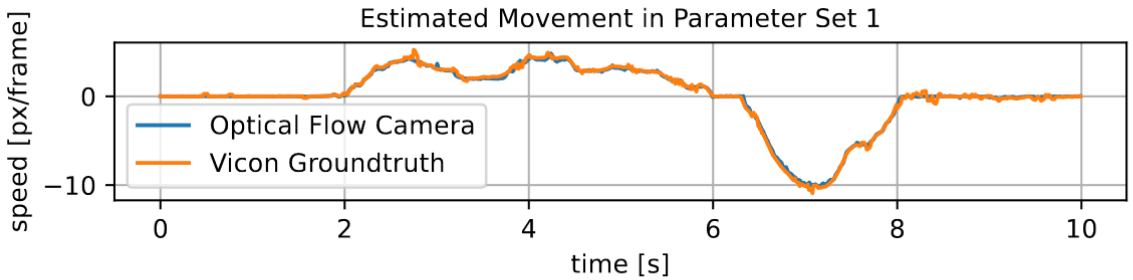
- **Facial Expression / Lip Movement**

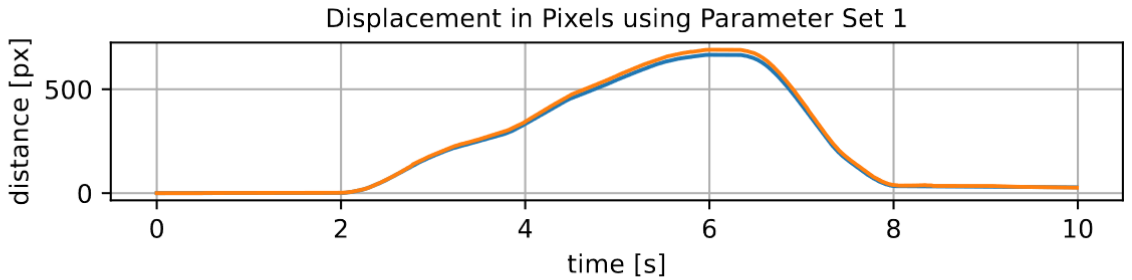
  - Focused region of interest
  - Moderate requirements on frame rate

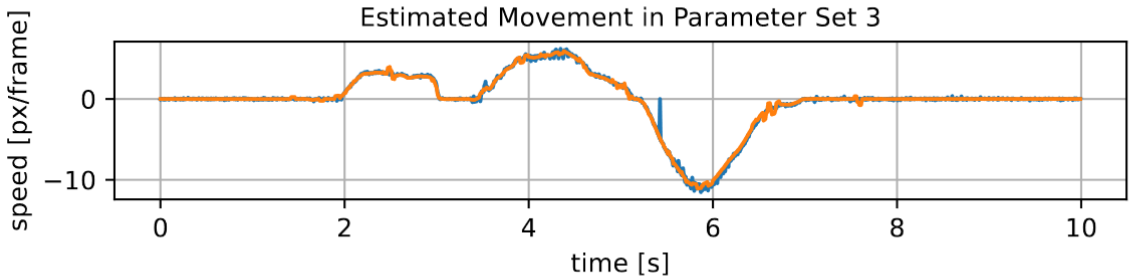


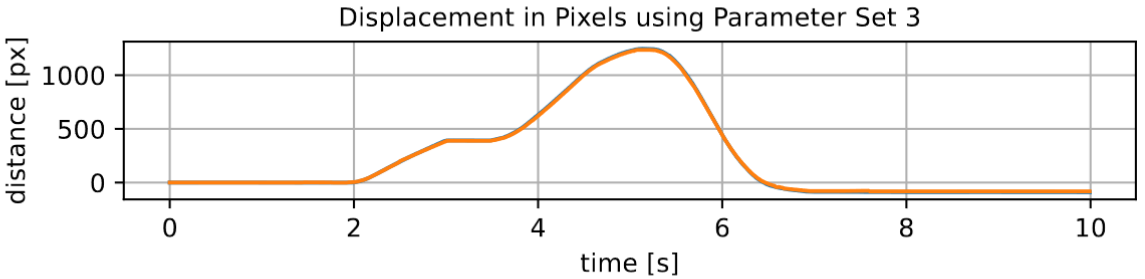
VD56G3 Characterization

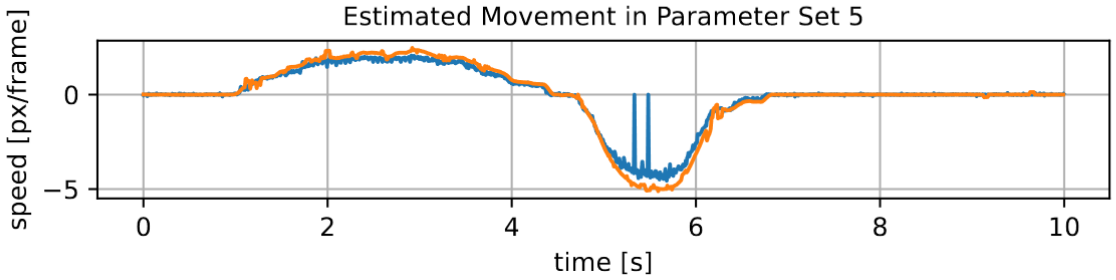
- **Linear Movements – Quantitative Analysis**

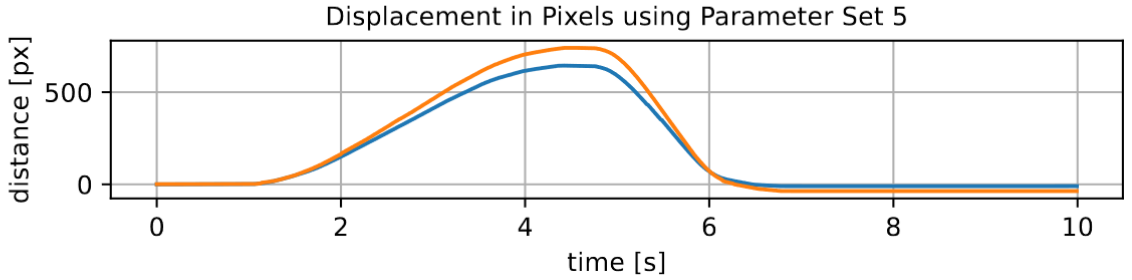






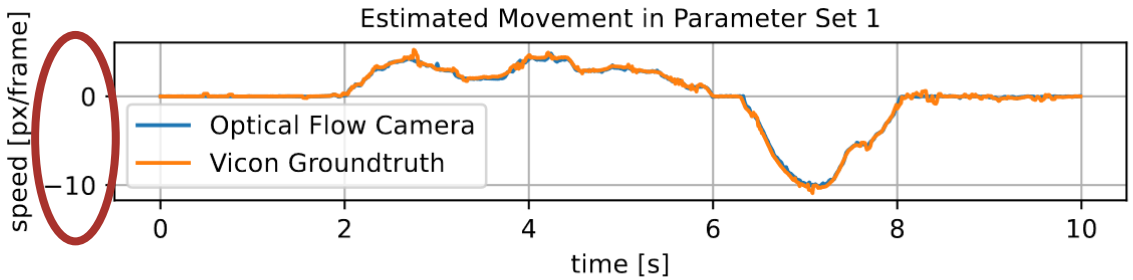


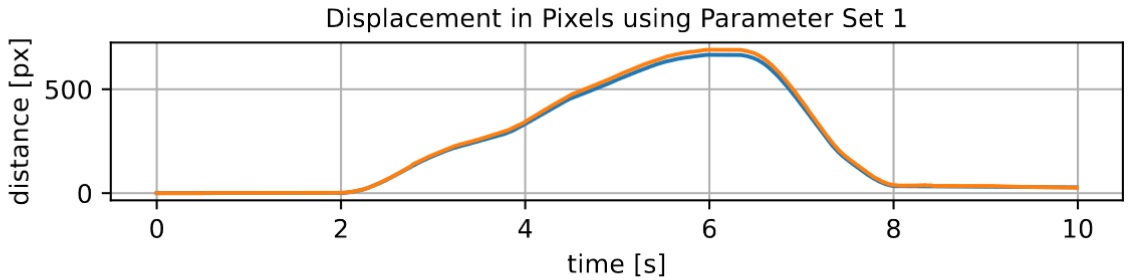


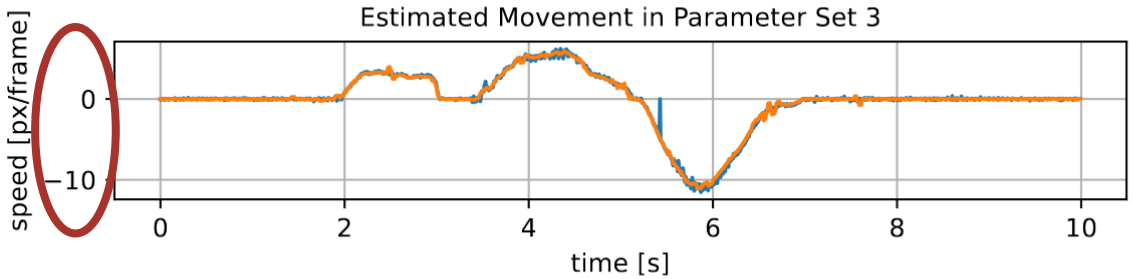


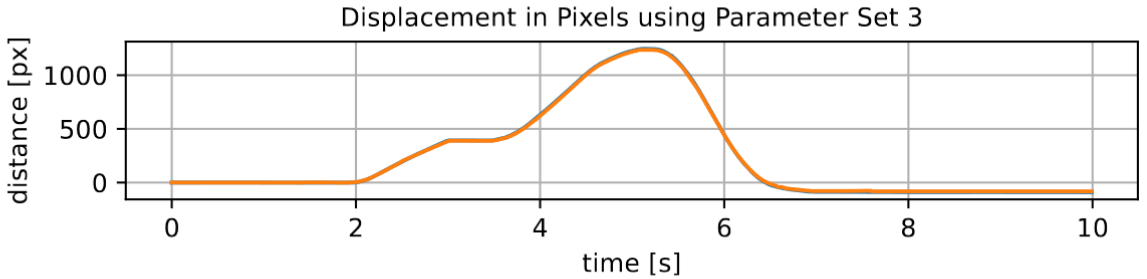
VD56G3 Characterization

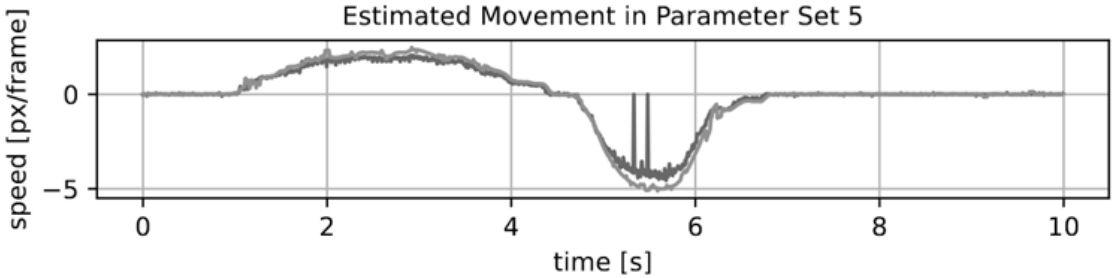
- **Linear Movements – Quantitative Analysis**

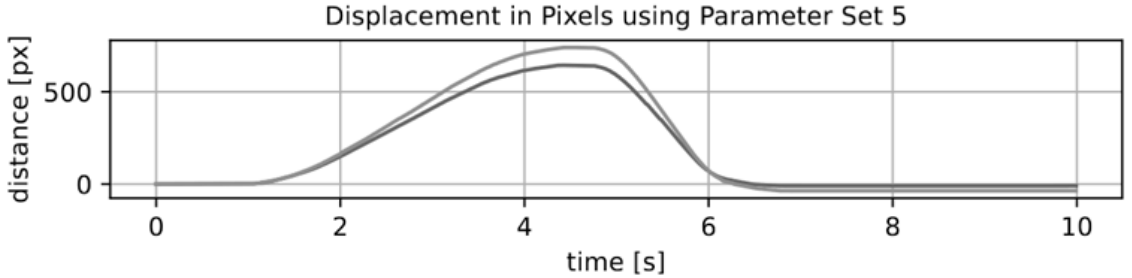






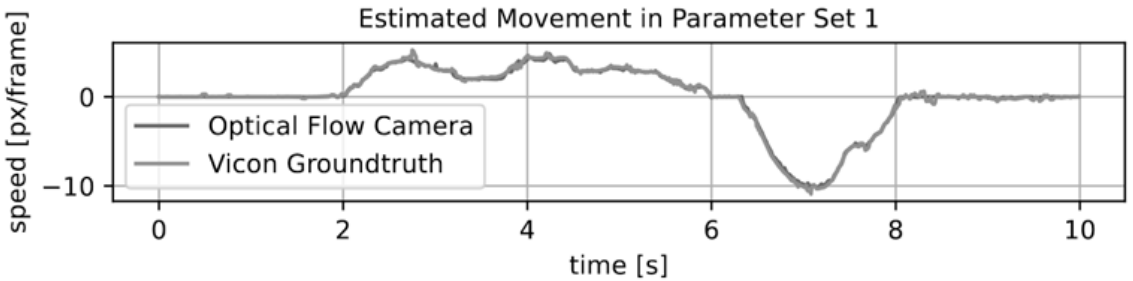


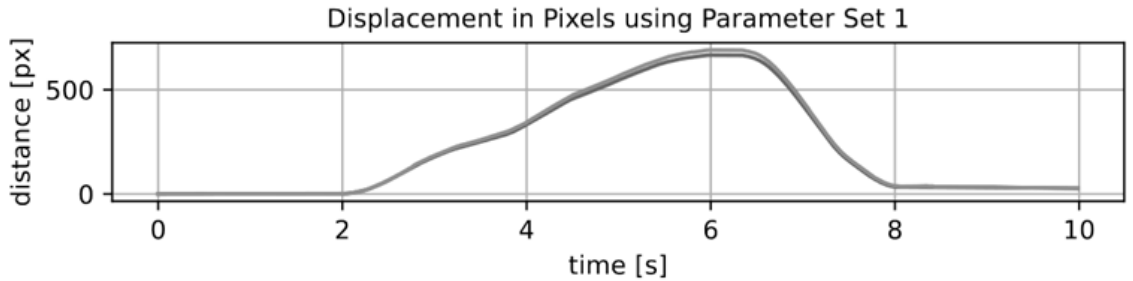


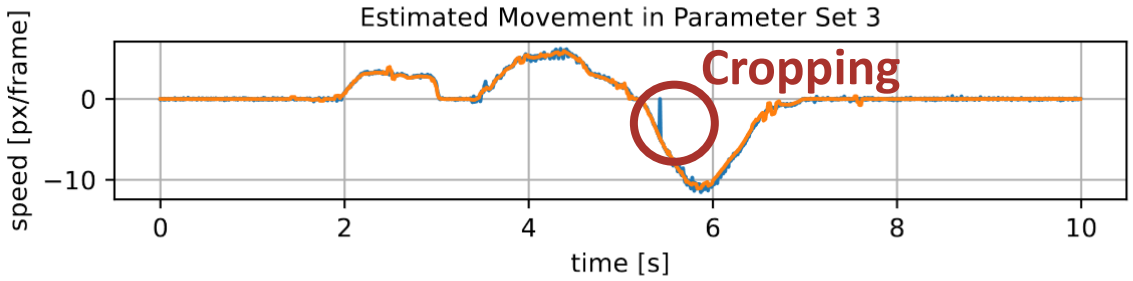


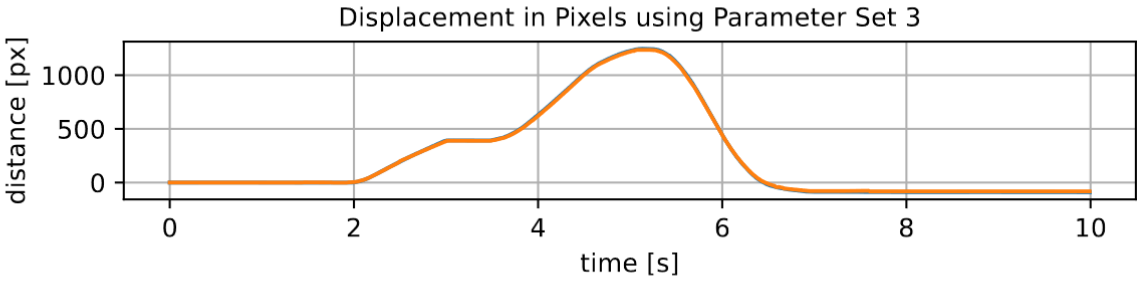
VD56G3 Characterization

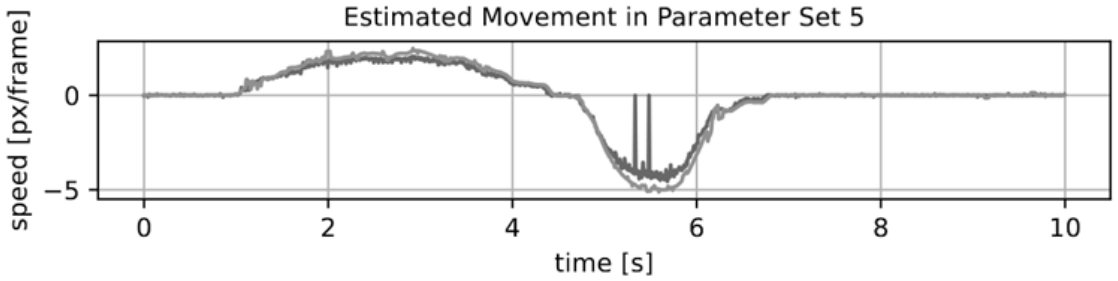
- **Linear Movements – Quantitative Analysis**

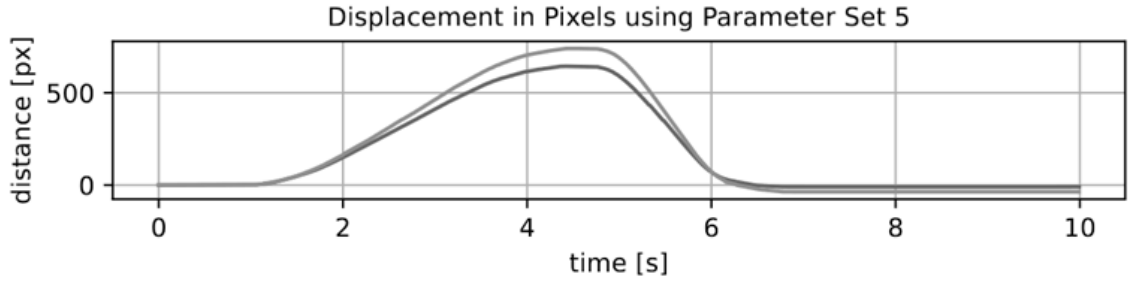






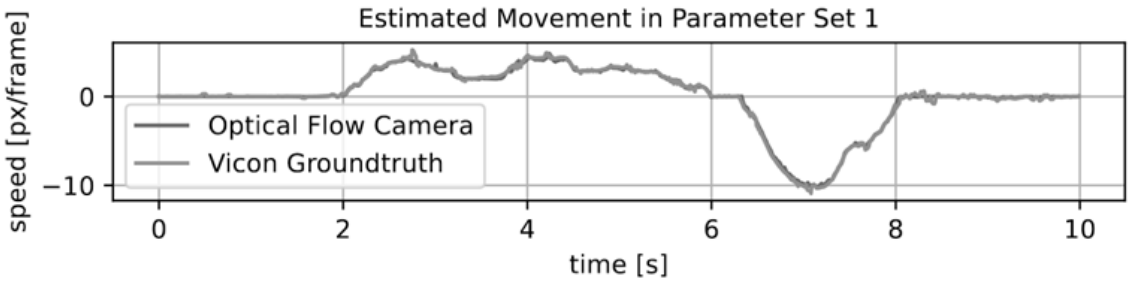


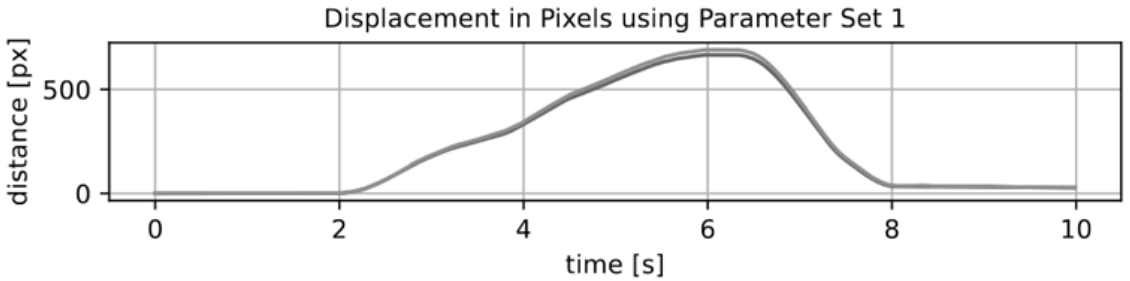


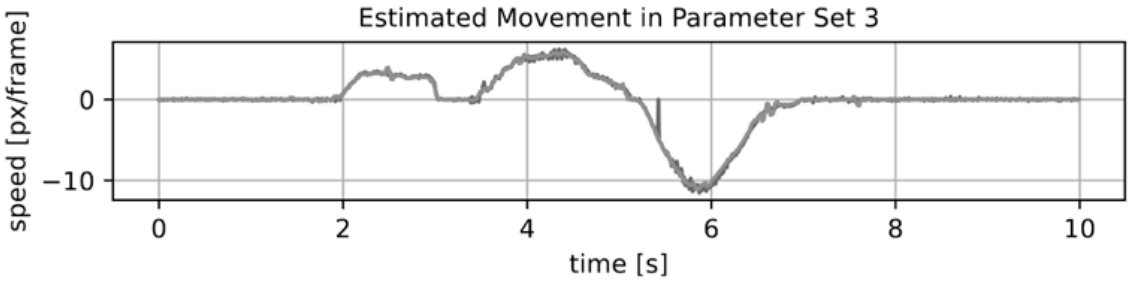


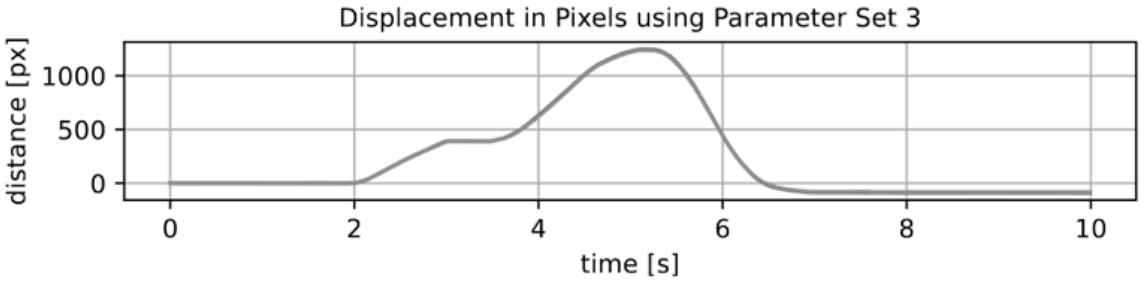
VD56G3 Characterization

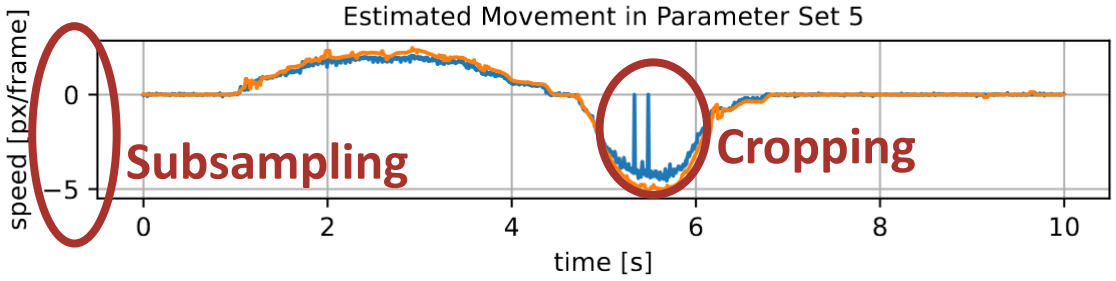
- **Linear Movements – Quantitative Analysis**

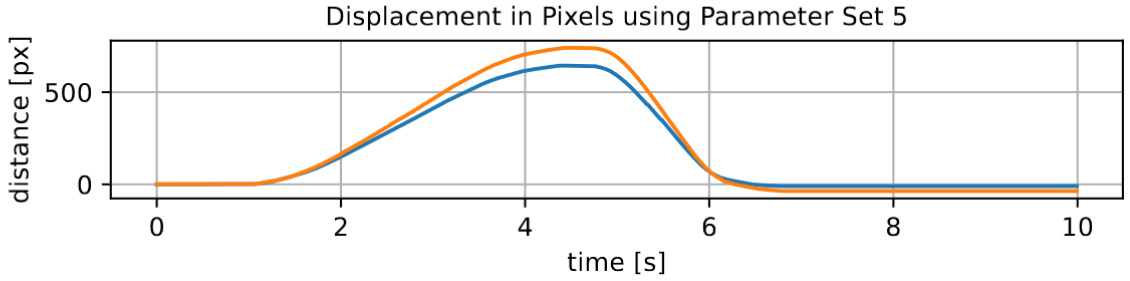






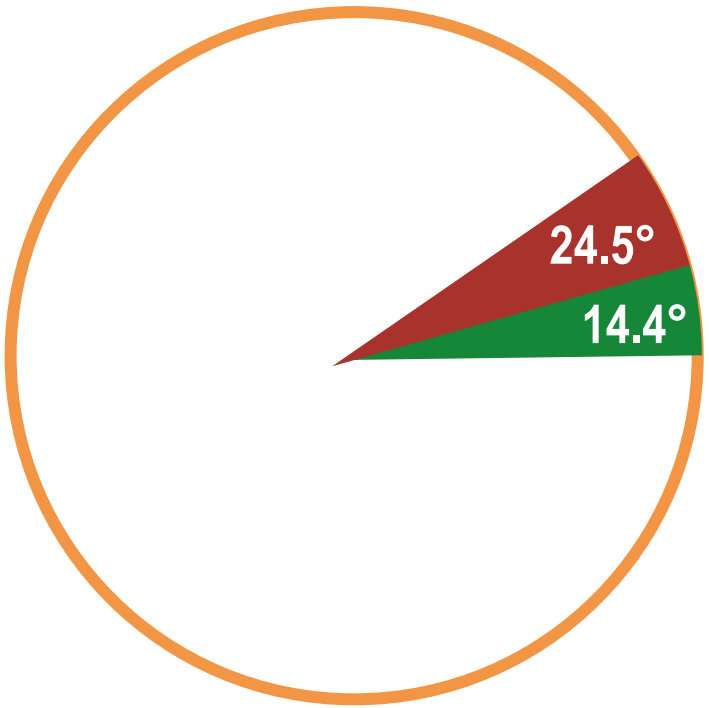


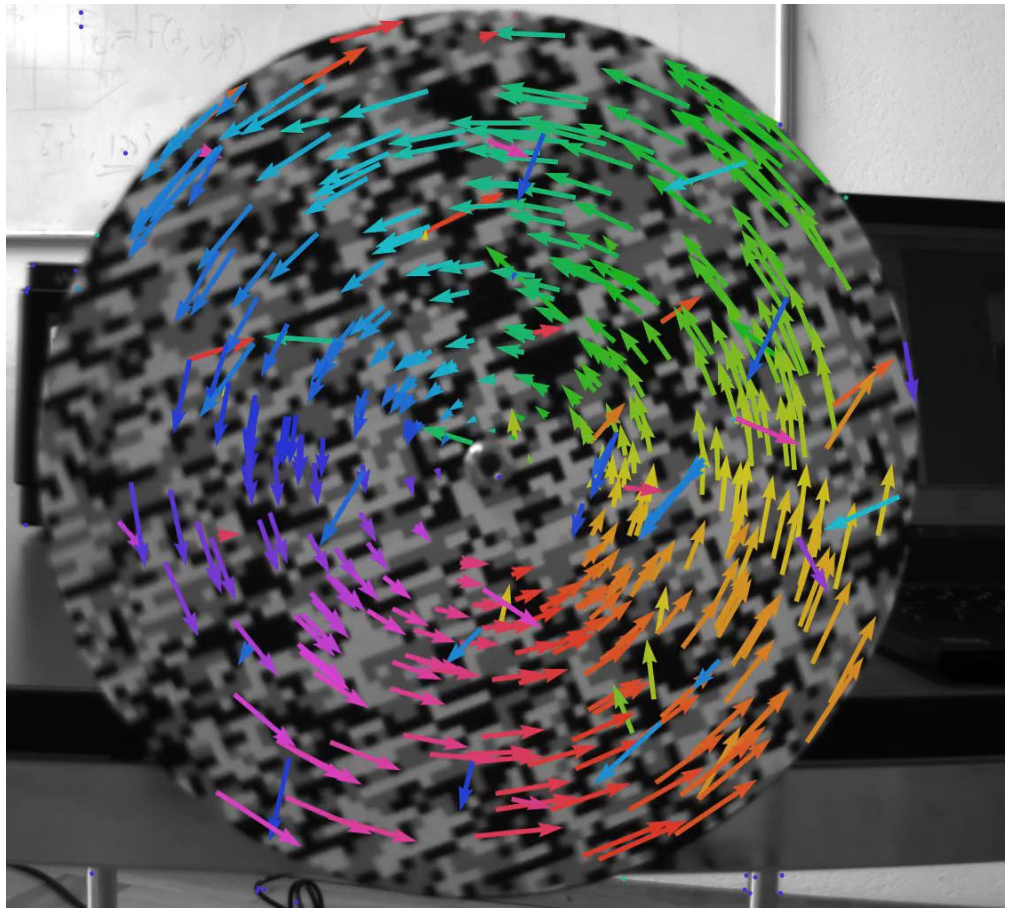




VD56G3 Characterization

- **Rotational Movements – Qualitative Analysis**
  - Change from 1 rotations/s to 1.7 rotations/s
  - Camera Operating at 25 frames per second
  - → Change from 14.4°/frame to 24.5°/frame





VD56G3 Characterization

- **Power Draw**
  - Different framerates and resolutions
  - Tests with disabled/enabled Optical Flow unit

FPS	Resolution	Total Power Draw	
		Optical Flow Off	Optical Flow On
60	high	176.81 mW	219.86 mW
60	medium	153.30 mW	188.15 mW
60	low	142.48 mW	169.68 mW
140	medium	185.07 mW	235.77 mW
240	low	170.98 mW	206.46 mW

high: 1124 pixels by 1364 pixels
medium: 560 pixels by 672 pixels
low: 272 pixels by 336 pixels

Discussion

- **The framerate of the VD56G3 sensor needs to matched with the translational or rotational speed of the object to be tracked to obtain best results.**
- **The detector sensitivity is only updated once per frame (not iteratively within one frame) which can lead to an absence of predictions**
- **The added power consumption of the optical flow unit is comparable to MCU implementations that produce a single Optical Flow prediction at similar frame rates [1]**
  - PX4FLOW: One Optical Flow prediction on MCU at ~250 frames per second
  - VD56G3: Up to 2048 Optical Flow predictions on ASIC at ~250 frames per second

[1] Honegger, D., Meier, L., Tanskanen, P., & Pollefeys, M. – ICRA 2013

Conclusion

- **The VD56G3 sensor is suitable for low-power edge applications that require low-latency optical flow predictions at high frame rates.**
- **We presented a set of configurations and give guidelines on the ideal configuration for a given problem.**

Consistent Redetections