The Top 10 Perception and Control Open Source Software and Datasets for Reproducible Vision-based Drone Research

Davide Scaramuzza

- My lab homepage: [http://rpg.ifi.uzh.ch/](http://rpg.ifi.uzh.ch/)
- Publications: [http://rpg.ifi.uzh.ch/publications.html](http://rpg.ifi.uzh.ch/publications.html)
- Software & Datasets: [http://rpg.ifi.uzh.ch/software_datasets.html](http://rpg.ifi.uzh.ch/software_datasets.html)
- YouTube: [https://www.youtube.com/user/ailabRPG/videos](https://www.youtube.com/user/ailabRPG/videos)
Motivation: Flying Robots to the Rescue!
My Dream Robot: Fast, Lightweight, Autonomous!

Video: https://youtu.be/JDvcBuRSDUU

LEXUS commercial, 2013 – Created by Kmel, now Qualcomm

WARNING!
There are 50 drones in this video but 40 are CGI and 10 are controlled via a Motion Capture System 😊
Research Overview

Real-time, Onboard Computer Vision and Control for Autonomous, Agile Drone Flight

Visual-Inertial State Estimation (~SLAM)

http://rpg.ifi.uzh.ch/research_vo.html

5x

Learning-aided Autonomous Navigation

http://rpg.ifi.uzh.ch/research_learning.html

Autonomous Flight

http://rpg.ifi.uzh.ch/research_mav.html

Event-based Vision for Low-latency Control

http://rpg.ifi.uzh.ch/research_dvs.html
State of the Art on Autonomous Drone Navigation

... but these robots are completely “blind”

Video https://youtu.be/kzLWxn5Z85Q

Source: R. D’Andrea’s research
State of the Art on Autonomous Drone Navigation

Video: https://youtu.be/fXy4P3nvxHQ

...while this robot can “see”
DARPA FLA Program (2015-2018)

- Flight speed: 20 m/s

Video: [https://www.youtube.com/watch?v=LaXc-jmN89U](https://www.youtube.com/watch?v=LaXc-jmN89U)
Robustness to “Stronger Disturbances”!

Video: https://www.youtube.com/watch?v=pGU1s6Y55JI

Faessler, Fontana, Forster, Scaramuzza, Automatic Re-Initialization and Failure Recovery for Aggressive Flight with a Monocular Vision-Based Quadrotor, ICRA’15. **Featured in IEEE Spectrum.**
Autonomous Flight through Narrow Gaps [ICRA’17]

Video:  
https://youtu.be/meSItatXQ7M

SVO: Semi-Direct Visual Odometry for Monocular and Multi-Camera Systems

Christian Forster, Zichao Zhang, Michael Gassner, Manuel Werlberger, Davide Scaramuzza

Abstract—Direct methods for Visual Odometry (VO) have gained popularity due to their capability to exploit information from all intensity gradients in the image. However, low computational speed as well as missing guarantees for optimality and consistency are limiting factors of direct methods, where established feature-based methods instead succeed at. Based on these considerations, we propose a Semi-direct VO (SVO) that uses direct methods to track and triangulate pixels that are characterized by high image gradients but relies on proven feature-based methods for joint optimization of structure and motion. Together with a robust probabilistic depth estimation algorithm, this enables us to efficiently track pixels lying on weak corners and edges in environments with little or high-frequency texture. We further demonstrate that the algorithm can easily be extended to multiple cameras, to track edges, to include motion priors, and to enable the use of very large field of view cameras, such as fisheye and catadioptric ones. Experimental evaluation on benchmark datasets shows that the algorithm is significantly faster than the state of the art while achieving highly competitive accuracy.

SUPPLEMENTARY MATERIAL
Video of the experiments: https://youtu.be/hR8uq1RTUfA
SVO: Semi-direct Visual Odometry [ICRA’14, TRO’17]

- Jointly tracks features and 6DoF motion under mostly-rigid world assumption
- Minimizes both photometric and geometric error

Achieves lowest latency & CPU load
- 2.5ms (400 fps) on i7 laptops
- 10ms (100 fps) on smartphones

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>St.D.</th>
<th>CPU@20 fps</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVO Mono</td>
<td>2.53</td>
<td>0.42</td>
<td>55 ±10%</td>
</tr>
<tr>
<td>ORB Mono SLAM (No loop closure)</td>
<td>29.81</td>
<td>5.67</td>
<td>187 ±32%</td>
</tr>
<tr>
<td>LSD Mono SLAM (No loop closure)</td>
<td>23.23</td>
<td>5.87</td>
<td>236 ±37%</td>
</tr>
</tbody>
</table>

Video: [http://rpg.ifi.uzh.ch/svo2.html](http://rpg.ifi.uzh.ch/svo2.html)

Estimation of fast motions with a single camera

Download from: [http://rpg.ifi.uzh.ch/svo2.html](http://rpg.ifi.uzh.ch/svo2.html)
On-Manifold Preintegration for Real-Time Visual-Inertial Odometry

Christian Forster, Luca Carlone, Frank Dellaert, Davide Scaramuzza

Abstract—Current approaches for visual-inertial odometry (VIO) are able to attain highly accurate state estimation via nonlinear optimization. However, real-time optimization quickly becomes infeasible as the trajectory grows over time; this problem is further emphasized by the fact that inertial measurements come at high rate, hence leading to fast growth of the number of variables in the optimization. In this paper, we address this issue by preintegrating inertial measurements between selected keyframes into single relative motion constraints. Our first contribution is a preintegration theory that properly addresses the manifold structure of the rotation group. We formally discuss the generative measurement model as well as the nature of the rotation noise and derive the expression for the maximum a posteriori state estimator. Our theoretical development enables the computation of all necessary Jacobians for the optimization and a-posteriori bias correction in analytic form. The second contribution is to show that the preintegrated IMU model can be seamlessly integrated into a visual-inertial pipeline under the unifying framework of factor graphs. This enables the application of incremental-smoothing algorithms and the use of a structureless model for visual measurements, which avoids optimizing over the 3D points, further accelerating the computation. We perform an extensive evaluation of our monocular VIO pipeline on real and simulated datasets. The results confirm that our modelling effort leads to accurate state estimation in real-time, outperforming state-of-the-art approaches.

SUPPLEMENTARY MATERIAL

- Video of the experiments: https://youtu.be/CsIkci5lfco

of monocular vision and gravity observable [1] and provides robust and accurate inter-frame motion estimates. Applications of VIO range from autonomous navigation in GPS-denied environments, to 3D reconstruction, and augmented reality.

The existing literature on VIO imposes a trade-off between accuracy and computational efficiency (a detailed review is given in Section II). On the one hand, filtering approaches enable fast inference, but their accuracy is deteriorated by the accumulation of linearization errors. On the other hand, full smoothing approaches, based on nonlinear optimization, are accurate, but computationally demanding. Fixed-lag smoothing offers a compromise between accuracy for efficiency; however, it is not clear how to set the length of the estimation window so to guarantee a given level of performance.

In this work we show that it is possible to overcome this trade-off. We design a VIO system that enables fast incremental smoothing and computes the optimal maximum a posteriori (MAP) estimate in real time. An overview of our approach is given in Section IV.

The first step towards this goal is the development of a novel preintegration theory. The use of preintegrated IMU measurements was first proposed in [2] and consists of combining many inertial measurements between two keyframes into a single relative motion constraint. We build upon this work and present a preintegration theory that properly addresses the manifold structure of the rotation group SO(3).
Full smoothing methods **estimate the entire history of the states** (camera trajectory and 3D landmarks), by solving a large nonlinear optimization problem.

- **Superior accuracy over filtering** methods, which only update the last state.
- Solved using **phactor graphs** (iSAM): only update variables affected by a new measurement.

**Open Source**

https://bitbucket.org/gtborg/gtsam

---


2. Delmerico, Scaramuzza, A Benchmark Comparison of Monocular Visual-Inertial Odometry Algorithms, ICRA’18, [PDF](#), [Video](#)
RPG Quadrotor Control Framework
Thrust Mixing, Saturation, and Body-Rate Control for Accurate Aggressive Quadrotor Flight

Matthias Faessler, Davide Falanga, and Davide Scaramuzza

ICRA’17, PDF, Video, Code

Differential Flatness of Quadrotor Dynamics Subject to Rotor Drag for Accurate Tracking of High-Speed Trajectories

Matthias Faessler\textsuperscript{1}, Antonio Franchi\textsuperscript{2}, and Davide Scaramuzza\textsuperscript{1}

Abstract—In this paper, we prove that the dynamical model of a quadrotor subject to linear rotor drag effects is differentially flat in its position and heading. We use this property to compute feed-forward control terms directly from a reference trajectory to be tracked. The obtained feed-forward terms are then used in a cascaded, nonlinear feedback control law that enables accurate agile flight with quadrotors. Compared to state-of-the-art control methods, which treat the rotor drag as an unknown disturbance, our method reduces the trajectory tracking error significantly. Finally, we present a method based on a gradient-free optimization to identify the rotor drag coefficients, which are required to compute the feed-forward control terms. The new theoretical results are thoroughly validated through extensive comparative experiments.

SUPPLEMENTARY MATERIAL

Video: [https://youtu.be/VIQIIwL6M5PA](https://youtu.be/VIQIIwL6M5PA)
Code: [https://github.com/uzh-rpg/rpg_quadrotor_control](https://github.com/uzh-rpg/rpg_quadrotor_control)

Fig. 1: First-person-view racing inspired quadrotor platform used for the presented experiments.

I. INTRODUCTION

...to compute feed-forward control terms directly from the reference trajectory to be tracked. The obtained feed-forward terms are then used in a cascaded, nonlinear feedback control law that enables accurate agile flight with quadrotors on a priori unknown trajectories. Finally, we present a method...
High-Performance Test Platform
Control-System Architecture

High-Level Part

Position Controller

Body Rates
Collective Thrust

Low-Level Part

Body Rate Controller

Body Torques
Collective Thrust

Mixer

Commands

Motors
Modeling and Compensating Aerodynamic Effects

- Aerodynamic effects (e.g., rotor drag) become dominant at high speeds
- We model and compensate them achieving 50% more accurate trajectory tracking at high speeds than without considering aerodynamic effects
- Proof that quadrotor dynamics subject to linear rotor drag is differentially flat in position and heading

Faessler, Franchi, Scaramuzza, Differential Flatness of Quadrotor Dynamics subject to Rotor Drag for Accurate Tracking of High Speed Trajectories, RAL’18, PDF, YouTube, ICRA18 Video Teaser, Open-Source Code
Open-Source Code

RPG Quadrotor Control Framework

• Used in my lab for the past 6 years
• Used in over 500 public live demonstrations from my lab
• has been the basis of over 50 publications from my lab

https://github.com/uzh-rpg/rpg_quadrotor_control
Quadrotor Simulation

- Used for **prototyping and development**
- We use the **same high-level control pipeline** as for our **real quadrotors**
- **Gazebo simulator**: physics, mobile robots kinematics, sensors
- **RotorS Simulator**: Gazebo plugin, simulates quadrotor aerodynamic effects plus sensors (very good for IMUs)
- **Continuous integration test** of control pipeline based on quadrotor simulation
- **Simulation** experiments **must succeed** to be able to merge code into the real **quadrotor** control pipeline

Development in Simulation

- IROS 2017 Autonomous Drone Racing Challenge simulated in Gazebo and visualized with RVIZ
Development in Simulation

- IROS 2017 Autonomous Drone Racing Challenge simulated in Gazebo and visualized with RVIZ

Video: https://youtu.be/ImQ8izGQzuk
PAMPC: Perception-Aware Model Predictive Control for Quadrotors

Davide Falanga*, Philipp Foehn*, Peng Lu, and Davide Scaramuzza

Abstract—We present a perception-aware model predictive control framework for quadrotors that unifies control and planning with respect to action and perception objectives. Our framework leverages numerical optimization to compute trajectories that satisfy the system dynamics and require control inputs within the limits of the platform. Simultaneously, it optimizes perception objectives for robust and reliable sensing by maximizing the visibility of a point of interest and minimizing its velocity in the image plane. Considering both perception and action objectives for motion planning and control is challenging due to the possible conflicts arising from their respective requirements. For example, for a quadrotor to track a reference trajectory, it needs to rotate to align its thrust with the direction of the desired acceleration. However, the perception objective might require to minimize such rotation to maximize the visibility of a point of interest. A model-based optimization framework, able to consider both perception and action objectives and couple them through the system dynamics, is therefore necessary. Our perception-aware model predictive control framework works in a receding-horizon fashion by iteratively solving a non-linear optimization problem. It is capable of running in real-time, fully onboard our lightweight, small-scale quadrotor using a low-power ARM computer, together with a visual-inertial odometry pipeline. We validate our approach in experiments demonstrating (I) the contradiction between perception and action objectives, and (II) improved behavior in extremely challenging lighting conditions.

Supplementary material

This paper is accompanied by a video showcasing the conducted experiments: http://rpg.ifi.uzh.ch/pampc

Fig. 1: An example application of our PAMPC, where a quadrotor is asked to fly at 3 m/s around a region of interest while keeping it visible in the field of view of its camera.

systems with onboard vision, since the motion of a camera can negatively affect the quality of the estimation, posing hard bounds on the agility of the robot. On the other hand, perception can benefit from the robot motion if it is planned considering the necessities and the limitations of onboard vision. For example, to pass through a narrow gap while localizing with respect to it using an onboard camera, it is necessary to guarantee that the gap is visible at all times. Similarly, to navigate through an unknown environment, it is necessary to guarantee that the camera always points towards texture-rich regions.

To fully leverage the agility of autonomous quadrotors, it
Manipulates trajectories to increase visibility of points of interest
- Example hover-to-hover: increased thrust → lower pitch + more altitude → better visibility
- Controls heading to look towards point of interest
- Robustly tracks reference jumps or trajectories
- Execution time: 3.5ms on a single core of a low-power ARM processor (Qualcomm Snapdragon or Odroid). Runs online with 100Hz

Falanga, Foehn, Peng, Scaramuzza, PAMPC: Perception-Aware Model Predictive Control, PDF (arXiv), YouTube, Code
The Zurich Urban Micro Aerial Vehicle Dataset

András L. Majdik¹ Charles Till² and Davide Scaramuzza²

Abstract
This paper presents a dataset recorded on-board a camera-equipped Micro Aerial Vehicle (MAV) flying within the urban streets of Zurich, Switzerland, at low altitudes (i.e., 5-15 meters above the ground). The 2 km dataset consists of time synchronized aerial high-resolution images, GPS and IMU sensor data, ground-level street view images, and ground truth data. The dataset is ideal to evaluate and benchmark appearance-based localization, monocular visual odometry, simultaneous localization and mapping (SLAM), and online 3D reconstruction algorithms for MAVs in urban environments.

Keywords
visual localization, air-ground matching, aerial robotics

Supplementary material
The dataset is available at:
http://rpg.ifi.uzh.ch/zurichmavdataset.html

Introduction
New applications of Micro Aerial Vehicles (MAVs) are envisioned by several companies, e.g., good delivery (e.g., Amazon Prime Air, DHL, Alibaba, Matternet, Swiss Post), inspection and monitoring (e.g., SenseFly, NOVA, DJI, Airbus, Impulse Robotics), and others.
The Zurich Urban Micro Aerial Vehicle Dataset [IJRR’17]

- **2km dataset** recorded with Fotokite drone in Zurich streets at 5-15m altitude
- Fotokite is a tethered drone (first and only drone authorized to fly over people’s heads in USA (FAA approved), France, and Switzerland)
- Ideal to evaluate and benchmark VO /VSLAM and 3D reconstruction for drones
- Data includes **time synchronized**:
  - Aerial images at 30Hz from GoPro H4, 1920×1080 pxl
  - GPS, IMU, Google Street View images

Video: [https://youtu.be/7hTvWbxxmY0](https://youtu.be/7hTvWbxxmY0)

Majdik, Till, Scaramuzza, The Zurich Urban Micro Aerial Vehicle Dataset, IJRR’ 17, [PDF](http://rpg.ifi.uzh.ch/zurichmavdataset.html), [YouTube](http://rpg.ifi.uzh.ch/zurichmavdataset.html)
Data-Efficient Decentralized Visual SLAM

Titus Cieslewski¹, Siddharth Choudhary² and Davide Scaramuzza¹

Abstract—Decentralized visual simultaneous localization and mapping (SLAM) is a powerful tool for multi-robot applications in environments where absolute positioning is not available. Being visual, it relies on cheap, lightweight and versatile cameras, and, being decentralized, it does not rely on communication to a central entity. In this work, we integrate state-of-the-art decentralized SLAM components into a new, complete decentralized visual SLAM system. To allow for data association and optimization, existing decentralized visual SLAM systems exchange the full map data among all robots, incurring large data transfers at a complexity that scales quadratically with the robot count. In contrast, our method performs efficient data association in two stages: first, a compact full-image descriptor is deterministically sent to only one robot. Then, only if the first stage succeeded, the data required for relative pose estimation is sent, again to only one robot. Thus, data association scales linearly with the robot count and uses highly compact place representations. For optimization, a state-of-the-art decentralized pose-graph optimization method is used. It exchanges a minimum amount of data which is linear with trajectory overlap. We characterize the resulting system and identify bottlenecks in its components. The system is evaluated on publicly available datasets and we provide open access to the code.

Supplementary Material
Data and code are at: https://github.com/uzh-rpg/dslam_open

I. INTRODUCTION
Using several robots instead of one can accelerate many tasks. One task is the construction of a spatial map.
Consider $n$ robots:

Goals:
- Have a consistent global map
- Decentralized: No central entity

T. Cieslewski, S. Choudhary, D. Scaramuzza: *Data-Efficient Decentralized Visual SLAM*, ICRA 2018
Challenge

- Place recognition and optimization among all robots
- Exchanging all data is expensive: $O(n^2)$!

Solution

- **Cluster the NetVLAD space**, assign clusters to robots
- Send given query only to **robot assigned to corresponding cluster**
- **Place recognition constant**, rather than linear in robot count!
- Clustering **trained on Oxford Robotcar** Dataset [Maddern 2016]

**Code is open source:**

[https://github.com/uzh-rpg/dslam_open](https://github.com/uzh-rpg/dslam_open)

NetVLAD for Python and Tensorflow
[https://github.com/uzh-rpg/netvlad_tf_open](https://github.com/uzh-rpg/netvlad_tf_open)

T. Cieslewski, S. Choudhary, D. Scaramuzza: *Data-Efficient Decentralized Visual SLAM*, ICRA’18, [PDF], ICRA18 Video Pitch, PPT, Code and Data
Deep-Learning based Navigation
Towards Domain Independence for Learning-based Monocular Depth Estimation

Michele Mancini\textsuperscript{1}, Gabriele Costante\textsuperscript{\textdagger}, Paolo Valigi\textsuperscript{1} and Thomas A. Ciarfuglia\textsuperscript{1}

Jeffrey Delmerico\textsuperscript{2} and Davide Scaramuzza\textsuperscript{2}

Abstract—Modern autonomous mobile robots require a strong understanding of their surroundings in order to safely operate in cluttered and dynamic environments. Monocular depth estimation offers a geometry-independent paradigm to detect free, navigable space with minimum space and power consumption. These represent highly desirable features, especially for micro aerial vehicles. In order to guarantee robust operation in real world scenarios, the estimator is required to generalize well in diverse environments. Most of the existent depth estimators do not consider generalization, and only benchmark their performance on publicly available datasets after specific fine-tuning. Generalization can be achieved by training on several heterogeneous datasets, but their collection and labeling is costly. In this work, we propose a Deep Neural Network for scene depth estimation that is trained on synthetic datasets, which allow inexpensive generation of ground truth data. We show how this approach is able to generalize well across different scenarios. In addition, we show how the addition of Long Short Term Memory (LSTM) layers in the network helps to alleviate, in sequential image streams, some of the intrinsic limitations of monocular vision, such as global scale estimation, with low computational overhead. We demonstrate that the network is able to generalize well with respect to different real world environments without any fine-tuning, achieving comparable performance to state-of-the-art methods on the KITTI dataset.

SUPPLEMENTARY MATERIAL

A video showing the results of our monocular depth estimation approach is available at https://youtu.be/UfoAyLb-5I.

The datasets we collected and the trained models to reproduce our results are available at http://www.sira.diel.unipg.it/supplementary/ral2016/extra.html.

1. INTRODUCTION

or mobile obstacles is crucial for field operations of the vast majority of ground and low altitude flight vehicles. Depth information can be used to estimate proximity to obstacles and, in general, to obtain an understanding of the surrounding 3D space. This perception of the 3D environment can then...
Learning-Based Monocular Depth Estimation

- Goal: Per-pixel depth estimation from a single image
- Training data from simulation only (Microsoft AirSim) & test on real data without any fine-tuning
- Etherogeneous synthetic scenes (urban, forest) to favor domain independence

[Mancini et al., Towards Domain Independence for Learning-Based Monocular Depth Estimation, RAL’17, PDF, PPT, YouTube, Dataset and Unreal-Engine 3D models](http://www.sira.diei.unipg.it-supplementary/ral2016extra.html)
DroNet: Learning to Fly by Driving

Antonio Loquercio*, Ana I. Maqueda †, Carlos R. del-Blanco †, and Davide Scaramuzza*

Abstract—Civilian drones are soon expected to be used in a wide variety of tasks, such as aerial surveillance, delivery, or monitoring of existing architectures. Nevertheless, their deployment in urban environments has so far been limited. Indeed, in unstructured and highly dynamic scenarios, drones face numerous challenges to navigate autonomously in a feasible and safe way. In contrast to traditional “map-localize-plan” methods, this paper explores a data-driven approach to cope with the above challenges. To accomplish this, we propose DroNet: a convolutional neural network that can safely drive a drone through the streets of a city. Designed as a fast 8-layers residual network, DroNet produces two outputs for each single input image: a steering angle to keep the drone navigating while avoiding obstacles, and a collision probability to let the UAV recognize dangerous situations and promptly react to them. The challenge is however to collect enough data in an unstructured outdoor environment such as a city. Clearly, having an expert pilot providing training trajectories is not an option given the large amount of data required and, above all, the risk that it involves for other vehicles or pedestrians moving in the streets. Therefore, we propose to train a UAV from data collected by cars and bicycles, which, already integrated into the urban environment, would not endanger other vehicles and pedestrians. Although trained on city streets from the viewpoint of urban vehicles, the navigation policy learned by DroNet is highly generalizable. Indeed, it allows a UAV to successfully fly at relative high altitudes and even in indoor environments, such as parking lots and corridors. To share our findings with the robotics community, we publicly release all our datasets, code, and trained networks.

Index Terms—Learning from Demonstration, Deep Learning in Robotics and Automation, Aerial Systems: Perception and Autonomy

Supplementary Material.

For supplementary video see https://youtu.be/ow7aw9H4BcA
The project’s code, datasets and trained models are available at
http://fg.in.uzh.ch/dronet.html

Fig. 1: DroNet is a convolutional neural network, whose purpose is to reliably drive an autonomous drone through the streets of a city. Trained with data collected by cars and bicycles, our system learns from them to follow basic traffic rules, e.g., do not go off the road, and to safely avoid other pedestrians or obstacles. Surprisingly, the policy learned by DroNet is highly generalizable, and even allows to fly a drone in indoor corridors and parking lots.

robotic system facing the above tasks should simultaneously solve many challenges in perception, control, and localization. These become particularly difficult when working in urban areas, as the one illustrated in Fig. 1. In those cases, the autonomous agent is not only expected to navigate while avoiding collisions, but also to safely interact with other agents present in the environment, such as pedestrians or cars.

The traditional approach to tackle this problem is a two-step interleaved process consisting of (i) automatic localization in a given map (using GPS, visual and/or range sensors), and (ii) computation of control commands to allow the agent to
DroNet: Learning to Fly by Driving

- DroNet learns to follow streets autonomously, without interaction
- The hardest problem in Machine Learning is data collection
- Our idea: learn to fly autonomously by mimicking cars and bicycles!

Video: YouTube
https://youtu.be/ow7aw9H4BcA


Low-latency, Event-based Vision
What is an event camera?

- A novel sensor that only transmits brightness **changes**
- Output is a stream of **events**

- **Low-latency** (~ 1 μs)
- **No motion blur**
- **High dynamic range** (140 dB instead of 60 dB)

**Problem**: Traditional computer vision algorithms for standard cameras cannot be applied!

[Mini DVS sensor from IniVation.com](http://youtu.be/LauQ6LWTkxM)
Check out their booth in the exhibition hall
Event SLAM: Low latency, HDR, high speed

[ICRA’14-17, RAL’16-17, BMVC’ 16-17, PAMI’17]


Video: https://youtu.be/DN6PaV_kht0  Video: https://youtu.be/6aGx-zBSzRA
Fast Event-based Corner Detection

Elias Mueggler\textsuperscript{1}
mueggler@ifi.uzh.ch
Chiara Bartolozzi\textsuperscript{2}
chiara.bartolozzi@iit.it
Davide Scaramuzza\textsuperscript{1}
sdavide@ifi.uzh.ch

\textsuperscript{1} Robotics and Perception Group
University of Zurich
Zurich, Switzerland
\textsuperscript{2} iCub Facility
Istituto Italiano di Tecnologia
Genova, Italy

Abstract

Event cameras offer many advantages over standard frame-based cameras, such as low latency, high temporal resolution, and a high dynamic range. They respond to pixel-level brightness changes and, therefore, provide a sparse output. However, in textured scenes with rapid motion, millions of events are generated per second. Therefore, state-of-the-art event-based algorithms either require massive parallel computation (e.g., a GPU) or depart from the event-based processing paradigm. Inspired by frame-based pre-processing techniques that reduce an image to a set of features, which are typically the input to higher-level algorithms, we propose a method to reduce an event stream to a corner event stream. Our goal is twofold: extract relevant tracking information (corners do not suffer from the aperture problem) and decrease the event rate for later processing stages. Our event-based corner detector is very efficient due to its design principle, which consists of working on the Surface of Active Events (a map with the timestamp of the latest event at each pixel) using only comparison operations. Our method asynchronously processes event by event with very low latency. Our implementation is capable of processing millions of events per second on a single core (less than a micro-second per event).
FAST-like Event-based Corner Detection [BMVC’17]

- Operates on Surface of Active Events

- The event is considered a corner if
  - 3-6 contiguous pixels on red ring are newer than all other pixels on the ring, and
  - 4-6 such contiguous pixels on the blue ring.

Mueggler, et al., Fast Event-based Corner Detection, BMVC’17, BMVC’17, PDF, Poster, YouTube, Open-Source Code

Code: https://github.com/uzh-rpg/rpg_corner_events
The Event-Camera Dataset and Simulator:
Event-based Data for Pose Estimation,
Visual Odometry, and SLAM

Elias Mueggler¹, Henri Rebecq¹, Guillermo Gallego¹, Tobi Delbruck² and Davide Scaramuzza¹

Abstract
New vision sensors, such as the Dynamic and Active-pixel Vision sensor (DAVIS), incorporate a conventional global-shutter camera and an event-based sensor in the same pixel array. These sensors have great potential for high-speed robotics and computer vision because they allow us to combine the benefits of conventional cameras with those of event-based sensors: low latency, high temporal resolution, and very high dynamic range. However, new algorithms are required to exploit the sensor characteristics and cope with its unconventional output, which consists of a stream of asynchronous brightness changes (called “events”) and synchronous grayscale frames. For this purpose, we present and release a collection of datasets captured with a DAVIS in a variety of synthetic and real environments, which we hope will motivate research on new algorithms for high-speed and high-dynamic-range robotics and computer-vision applications. In addition to global-shutter intensity images and asynchronous events, we provide inertial measurements and ground-truth camera poses from a motion-capture system. The latter allows comparing the pose accuracy of ego-motion estimation algorithms quantitatively. All the data are released both as standard text files and binary files (i.e., rosbag). This paper provides an overview of the available data and describes a simulator that we release open-source to create synthetic event-camera data.

Keywords
Event-based cameras, visual odometry, SLAM, simulation

Dataset Website
All datasets and the simulator can be found on the web:
http://rpg.ifi.uzh.ch/davis_data.html
A video containing visualizations of the datasets.
Event Camera Dataset and Simulator [IJRR’17]

- Publicly available: [http://rpg.ifi.uzh.ch/davis_data.html](http://rpg.ifi.uzh.ch/davis_data.html)
- First event camera dataset specifically made for VO and SLAM
- Many diverse scenes: HDR, Indoors, Outdoors, High-speed
- Blender simulator of event cameras
- Includes
  - IMU
  - Frames
  - Events
  - Ground truth from a motion capture system

Code, papers, videos, companies on event cameras:
- [https://github.com/uzh-rpg/event-based_vision_resources](https://github.com/uzh-rpg/event-based_vision_resources)

Mueggler, Rebecq, Gallego, Delbruck, Scaramuzza,
Summary

Control
1. RPG Quadrotor Control Framework:
   https://github.com/uzh-rpg/rpg_quadrotor_control
2. Perception Aware Model Predictive Control:
   https://github.com/uzh-rpg/rpg_quadrotor_mpc

Visual Inertial State Estimation and Decentralized SLAM
3. SVO 2.0:
   http://rpg.ifi.uzh.ch/svo2.html
4. GTSAM with Pre-Integrated IMU Factors:
   https://bitbucket.org/gtborg/gtsam
5. MultiFOV Dataset
   http://rpg.ifi.uzh.ch/fov.html
6. Data-Efficient Decentralized Visual SLAM:
   https://github.com/uzh-rpg/dslam_open
7. The Zurich Urban MAV Dataset
   http://rpg.ifi.uzh.ch/zurichmavdataset.html

Deep Learning
8. Monocular Depth Estimation for Drones:
   http://www.sira.diei.unipg.it/supplementary/ral2016/extra.html
9. DroNet: Learning to Fly by Driving
   http://rpg.ifi.uzh.ch/dronet.html

Event Cameras
10. Event camera Dataset and Simulator
    http://rpg.ifi.uzh.ch/davis_data.html
11. FAST Corner Detector for Event Cameras
    https://github.com/uzh-rpg/rpg_corner_events