An Open Source and Open Hardware Embedded Metric Optical Flow CMOS Camera for Indoor and Outdoor Applications

Dominik Honegger, Lorenz Meier, Petri Tanskanen and Marc Pollefeys
ETH Zürich, Switzerland

Abstract— **Robust velocity and position estimation at high update rates is crucial for mobile robot navigation. In recent years optical flow sensors based on computer mouse hardware chips have been shown to perform well on micro air vehicles. Since they require more light than present in typical indoor and outdoor low-light conditions, their practical use is limited. We present an open source and open hardware design [1] of an optical flow sensor based on a machine vision CMOS image sensor for indoor and outdoor applications with very high light sensitivity. Optical flow is estimated on an ARM Cortex M4 microcontroller in real-time at 250 Hz update rate. Angular rate compensation with a gyroscope and distance scaling using a ultrasonic sensor are performed onboard. The system is designed for further extension and adaption and shown in-flight on a micro air vehicle.**

I. INTRODUCTION

The Successful navigation of mobile robots depends on robust position and velocity information. In particular for undamped systems, such as vertical takeoff and landing of micro aerial vehicles (MAVs), sufficient update rate and low-latency sensors are required to maintain track during operation.

In recent years optical flow sensors based on computer mouse sensors have been successfully used for this purpose [6]. Faced to the ground these sensors can be used for accurate velocity, and with integration, position measurements. However, mouse sensors require strong lighting to provide accurate measurements. The issue can be alleviated with onboard active lighting in the infrared range, such as high-brightness red LEDs. This however conflicts with the limitations on power consumption and ground distance. Automotive CMOS image sensors are substantially more light sensitive and allow operation in indoor environments and during adverse outdoor conditions without artificial lighting. However, to our knowledge there is no CMOS based, lightweight sensor available that could be easily integrated into a robotics research system. The Parrot ARDrone has an onboard camera and computes optical flow in an embedded Linux environment [2] but the hardware design as well as the software implementation is closed source and can only be modified within certain bounds. In this work, we present PX4FLOW, an ARM Cortex M4 based sensor system that performs optical flow processing at 250 frames per second at a subsampled resolution of 64x64 pixels using a CMOS machine vision sensor. An ultrasonic range sensor is used to measure the distance towards the scene and to scale optical flow values to metric velocity values. Angular velocity is

compensated using an onboard gyroscope to estimate correctly the translational velocity. Automatic exposure control allows usage in outdoor and indoor environments. Figure 1 shows the sensor system with mounted lens and ultra sonic distance sensor. The CMOS sensor-microcontroller system is low-power, low-latency and low-cost and therefore suitable for micro aerial vehicle applications.

Fig. 1. The presented PX4FLOW optical flow sensor with mounted lens on the left and ultra sonic distance sensor on the right.

First we provide a summary of the relation between motion field and velocity including all relevant parameters. We present an efficient system setup to perform computer vision tasks with a microcontroller at high frame rates. The system is validated with several experiments comparing the presented flow sensor to a standard mouse sensor and GPS and measurements with ground truth using a VICON motion tracking system.

II. RELATED WORK

Dedicated computer mouse hardware sensors are successfully used for navigation and obstacle avoidance of micro aerial vehicles [6]. Due to the small tracked image area several mouse sensors are used within one vehicle for multiple direction flow detection in [1]. This limitation can be addressed in our design with a wider angle lens and an according distribution of pixel locations where optical flow is calculated. More complex maneuvers as autonomous take-off and landing based on optical flow sensors have been presented in [9]. A mouse sensor based optical flow module for quadrotor control is shown in [3]. All these systems are

not suitable for indoor applications since standard mouse sensors require stronger lighting than present in normal indoor conditions. Systems using a CMOS sensor for optical flow computation have been built using a wireless link for sending the camera images to a computer on the ground, which is processing the images and sending the computed flow value back to the MAV [5] [8].

Dedicated hardware designs implemented in field-programmable gate arrays (FPGA) are used to perform all computations onboard in real-time. The authors of [7] showed optical flow computation in real-time for micro air vehicles. In [4] a stereo camera pair computes optical flow and dense stereo for metric flow in 3D. FPGA systems are, compared to our approach, expensive, large and require further knowledge in hardware description languages.

We combine a low-cost machine vision CMOS sensor and a low-cost, low-power standard microcontroller to compute optical flow in indoor and outdoor environments. The software, written completely in C, is compiled with the GNU GCC compiler and is available under the BSD open source license. The hardware design is available as CC-BY-SA creative-commons licensed open hardware.

III. BACKGROUND

This section summarizes the relation between the pixel-based motion field and metric velocity.

A. Basic Equations of the Motion Field

The motion field is created by projecting the 3D velocity field on the image plane. Let $\mathbf{P} = [X, Y, Z]^\top$ be a point in the three dimensional camera reference frame. Let the optical axis be the Z-axis of this frame, let f denote the focal length and let the center of projection be in the origin. The projected pixel coordinates of $\mathbf{P}$ on the image plane are given by

$$\mathbf{p} = \mathrm{f}\frac{\mathbf{P}}{Z}. \tag{1}$$

Since the focal length f is equal to the distance of the image plane to the origin, the third coordinate of $\mathbf{p}$ is constant $\mathbf{p} = [x, y, \mathrm{f}]^\top$. The relative motion between the camera and $\mathbf{P}$ is given by

$$\mathbf{V} = -\mathbf{T} - \omega \times \mathbf{P}, \tag{2}$$

where ω is the angular velocity and $\mathbf{T}$ the translational component of the motion. Taking the derivative with respect to time of both sides of (1) leads to the relation between the velocity of $\mathbf{P}$ in the camera reference frame and the velocity or the flow of $\mathbf{p}$ in the image plane

$$\frac{\mathbf{flow}}{\mathbf{\Delta}\mathrm{time}} \; \widehat{=} \; \mathbf{v} = \mathrm{f}\frac{Z\mathbf{V} - V_z\mathbf{P}}{Z^2}. \tag{3}$$

Expressed in x and y components and substituting (2) the motion field can be written as

$$v_x = \frac{T_z x - T_x f}{Z} - \omega_y f + \omega_z y + \frac{\omega_x xy - \omega_y x^2}{f} \tag{4}$$

$$v_y = \frac{T_z y - T_y f}{Z} + \omega_x f - \omega_z x + \frac{\omega_x y^2 - \omega_y xy}{f}. \tag{5}$$

The motion field components are equal to pure translational parts plus pure rotational parts. The rotational parts are not dependent from Z and therefore the angular velocity does not carry scene depth information.

The translational components in (4) and (5) are scaled with the focal length and the current distance Z to the scene. If the translational velocity is needed, e.g., if the rotational veclocity is zero or known (measured by a gyroscope) and compensated from the motion field, it is possible to compute the translational velocity in metric scale by

$$\mathbf{v}_{\mathrm{m,trans}} = \mathbf{v}\frac{Z}{\mathrm{f}}. \tag{6}$$

The combination of the translational part of the motion field and distance measurements of the scene leads to a translational velocity in metric scale if it can be assumed that the distance to the scene is approximately constant. This is especially the case if the camera is faced perpendicular to the ground.

IV. SYSTEM SETUP

In the following, we describe an efficient system setup to perform computer vision tasks on a microcontroller. An overview of the setup is shown in Figure 2.
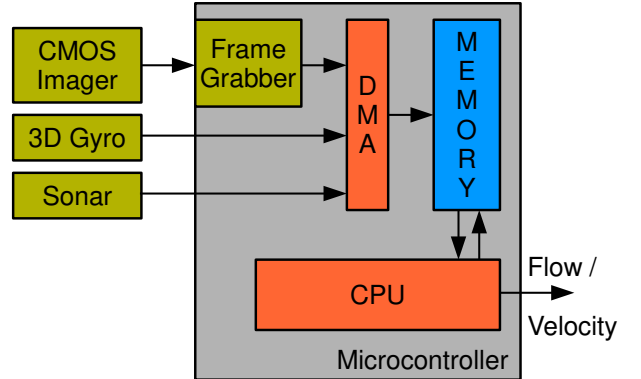

Fig. 2. PX4FLOW system setup, the CMOS imager is directly connected to the microcontroller. Image data from the frame grabber module as well as angular rate and distance measurements are stored in system memory using DMA. The flow values scaled with the corresponding distance are sent out.

The sensor system performs optical flow calculation on images acquired from the CMOS machine vision sensor. It is directly connected to the ARM Cortex M4 microcontroller to a special imager bus peripheral. The micro controller processes the images in real-time. A frame grabber module captures frames from the sensor and stores them in memory. Optical flow values between two succeeding frames are calculated in the flow module.

A refinement to achieve subpixel accuracy and angular rate compensation using the measurements of a gyroscope are performed on the resulting flow values. Finally metric scaling of the flow values as shown in (6) is done using the distance measurements of an ultrasonic sensor.

A. Frame Grabber

Pixel data is streamed in the microcontroller using a parallel interface. The frame grabber module samples pixel values at the corresponding pixel clock of the CMOS sensor. Direct Memory Access (DMA) with double buffering is used to transfer image data to the memory. Only the current and the preceeding frames are stored.

B. Flow Computation

Optical flow is calculated between two successive frames. The sum of absolute differences (SAD) block matching algorithm is used. The SAD value of a reference block of pixels of the current and preceeding frame is compared to SAD values within the search area. The position of the best match in the search area is selected as the resulting flow value. The search range as well as the block size are fully parameterizable to support various applications.

C. Subpixel refinement, Rotation Compensation and Metric Scaling

After the best match is computed a refinement step is made. The flow is computed using bilinear interpolation on subpixel accuracy inside the pixel with the best match. The onboard gyroscope values are stored over DMA into the main memory of the microcontroller. It delivers the angular rotation rates of the camera. These can be used for rotation compensation as the rotational parts of the motion field can be calculated using angular rates and focal length of the lens. The flow introduced by rotations is substracted of the computed optical flow value.

The search window for the SAD block matching can be shifted according to the direction of the rotation. This increases the maximum allowed camera rotation rate, the translational flow can be estimated with a small search range even if the camera is rotating.

If the optional sonar sensor is attached, the refined optical flow value is scaled to its metric value by assuming that the camera looks at a planar surface with the distance measured by the ultra sonic sensor as in (6).

D. Components

The frame is captured through a 16 mm M12 lens (21°FOV) including IR-block coating with an Aptina MT9V034 imager with 6μm pixel size. This device can provide up to 60 FPS at full resolution of 752Hx480V. When pixel binning is enabled, the frame rate increases further. At 4x binning, the resulting resolution of 188Hx120V allows up to 250 Hz frame rate. Pixel binning is performed inside the imager on the digital pixel data. In 4x binning mode one pixel sent out of the imager is the average of 4x4 pixels in normal mode. The images are captured with a STM32F407 32bit microcontroller with Cortex M4F core at 168 MHz. It provides 192 KB RAM and a hardware floating point unit for IEEE 754 single precision floating point operations. The system is completed with a L3GD20 low-noise MEMS gyroscope (16 bit resolution, up to 2000°/s) and an onboard parameter storage EEPROM. Mounting provisions for

Maxbotics ultrasonic sensors allow a compact form factor, the millimeter-resolution model HRLV-EZ4 is recommended in this setup.

V. IMPLEMENTATION

In this section we describe an efficient implementation of the optical flow estimation module and the involved frame maintenance tasks.

A. Frame Grabber

The ARM Cortex M4F microcontroller offers a fully parametrizable camera interface which allows to configure the horizontal as well as the vertical frame dimensions and the pixel color depth. We use 8 bit resolution per pixel to be able to process 4 pixels at a time with special 32 bit instructions. The camera interface stores the incoming pixel data in the embedded main system memory using DMA.

B. Optical Flow

Optical flow estimation is based on SAD block matching. Dedicated integer vector instructions of the Cortex M4 are used. They allow for the computation of the SAD value of four pixels in parallel within a single CPU clock. The SAD value of a 8x8 pixel block is calculated within a search area of ± 4 pixels in both directions. The position of the minimum SAD block value out of the 81 candidates is taken as the flow value at the corresponding sample point. A total of 64 sample points is processed per frame. A subsequent histogram filter takes into account every sample point and choses the histogram bin with the highest value. This results in optical flow values with one pixel resolution. A search range of ± 4 pixels in combination with a 16mm focal length lens corresponds to ± 1.5 meters per second for an object at one meter distance at 250 frames per second.

C. Refinement

After the best flow value is computed on pixel basis a subpixel refinement step is done. The optical flow is estimated with half pixel step size in all directions from the pixel with the best match result. The pixels are bilinearly interpolated again using the integer vector instructions of the Cortex M4. The best match of the directions around the best match including the old result is selected as the final refined optical flow result.

D. Angular Rate Compensation

The equation of the motion field in x component (4) in case of constant distance to the scene leads to

$$v_x = \frac{-T_x f}{Z} - \omega_y f + \omega_z y. \qquad (7)$$

The terms divided by the focal length are neglected since they are more than one order of magnitude smaller compared to the other summands. Under these conditions the effects of angular rates on the motion field can be compensated.

E. Lens Distortion

Using a 16mm M12 lens leads to a maximum radial distortion displacement of 0.15 pixel for the 64x64 pixels patch at full resolution. Since 4x pixel binning is used and the subsampled 64x64 pixels patch is located next to the center of distortion, no correction is required.

The available onboard memory allows for a look-up table based undistortion. That can be implemented for lenses with high radial distortion as typical wide angle lenses.

VI. APPLICATIONS

The fully configurable aspect ratio and the large resolution of the flow sensor system allows multiple use cases, out of which only one is presented here in depth. Equal aspect ratio allows for 2D flow detection on a surface for translational velocity estimation. Wide screen aspect ratio combined with a wide angle lens and detecting flow at both extreme ends of the image supports navigation in corridor environments.
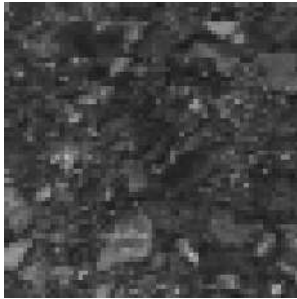
A. Velocity Estimation

Fig. 3. Typical street texture as seen from the flow sensor from 0.8 m altitude through a 16 mm M12 lens.

The two dimensional flow field scaled with the distance to the current scene and angular rate compensation is equal to the translational velocity in metric scale. Included in a mobile robot navigation system the flow sensor supports the position estimator with accurate velocity measurements at high update rates. Figure 3 shows a typical 64x64 pixel frame as captured by the optical flow camera.

B. Direction Estimation

A wide aspect ratio (i.e. 120Hx32V) in combination with a wide angle lens (180°FOV) allows to evade obstacles. The ability to define the image dimensions and to define at which positions optical flow is to be computed allows to use only one sensor, instead of employing multiple mouse sensors. Figure 4 shows a 120x32 pixel frame captured by the PX4FLOW sensor.

Fig. 4. An image from the flow sensor using a wide angle lens with 120x32 pixel resolution resulting in approximately 90° viewing angle showing an office environment with computers on tables.

VII. RESULTS

To evaluate the sensor different experiments were performed. We evaluated the velocity estimation to ground truth during a hovering flight of a quadrotor. An experiment evaluating the minimum illumination needed for reliable flow estimation and comparing it to a typical mouse sensor was done. An indoor trajectory was made integrating the velocity values. Finally, a long outdoor trajectory was created using the attitude estimate from a quadrotor IMU and it is compared to the result of a mouse sensor. Table I shows the specifications of the PX4FLOW sensor.

TABLE I

PX4FLOW SENSOR SPECIFICATION

	PX4FLOW
Size	45mm x 35mm
Power Consumption	115mA @ 5.0V = 0.575 W
Update Rate	250 Hz
Maximum Flow	± 1.5 rad/s = $\pm 86\,°/s$

A. Flight Performance

The flow sensor was attached via UART to a PIXHAWK Cheetah quad rotor, which performed the flight and data logging onboard. The sensor sends ground speed estimates in the MAVLink protocol format to the onboard autopilot, where the individual measurements are rotated with the current heading and integrated to a position estimate in global coordinates.
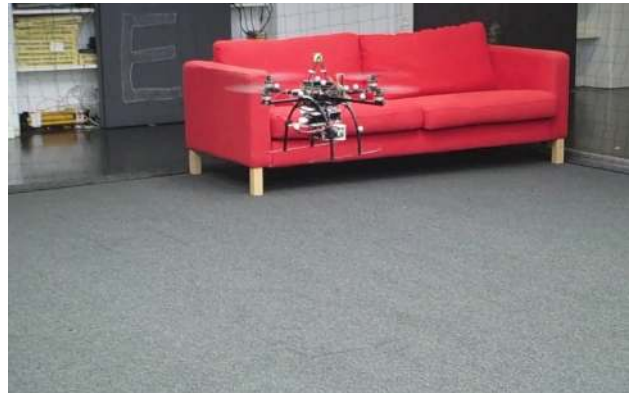

Fig. 5. PIXHAWK Cheetah hovering in an indoor lab environment with PX4FLOW faced downwards

The flow sensor was evaluated with a Vicon motion tracking system as ground truth. Figure 6 shows the raw metric output in x and y direction compared to Kalman filtered flow and ground truth. The plot shows in-flight data captured during on-spot hovering in an indoor lab shown in Figure 5.

B. Illumination Test

A particular benefit of the use of an automotive imager over a normal mouse sensor is the substantially higher light sensitivity, in particular in indoor lighting conditions. Measurements were performed in a static setup with a rotating
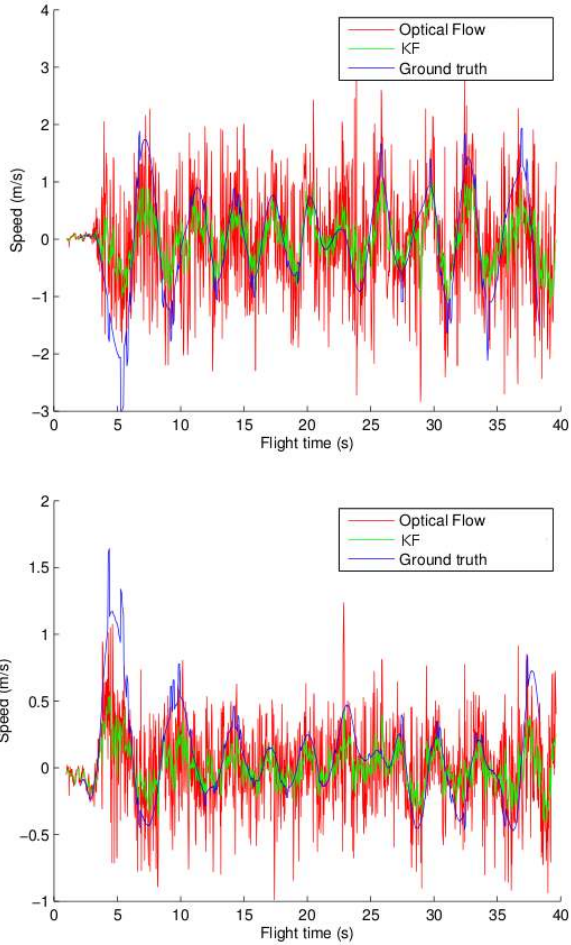
Fig. 6. Results of x and y velocities during a hovering flight estimated by the flow sensor, the Kalman filtered velocities and ground truth.

The typical office lighting is 500-700 Lux, we measured 603 Lux in the experimental lab. These lighting conditions are already quite good and not the average in all locations of a building. The mouse sensor is therefore unsuitable for reliable and robust indoor operation.

C. Indoor Trajectory

The flow sensor was moved along a square shaped indoor trajectory. Measured velocity values and integrated position estimates are shown in Figure 7. No mouse sensor comparison is presented as this sensor is unable to deliver outputs at typical office lighting conditions. Start and stop position are only 0.25 m spaced apart and the overall trajectory has a length of 28.44 meters. No filtering or motion model was used to obtain the integrated position.
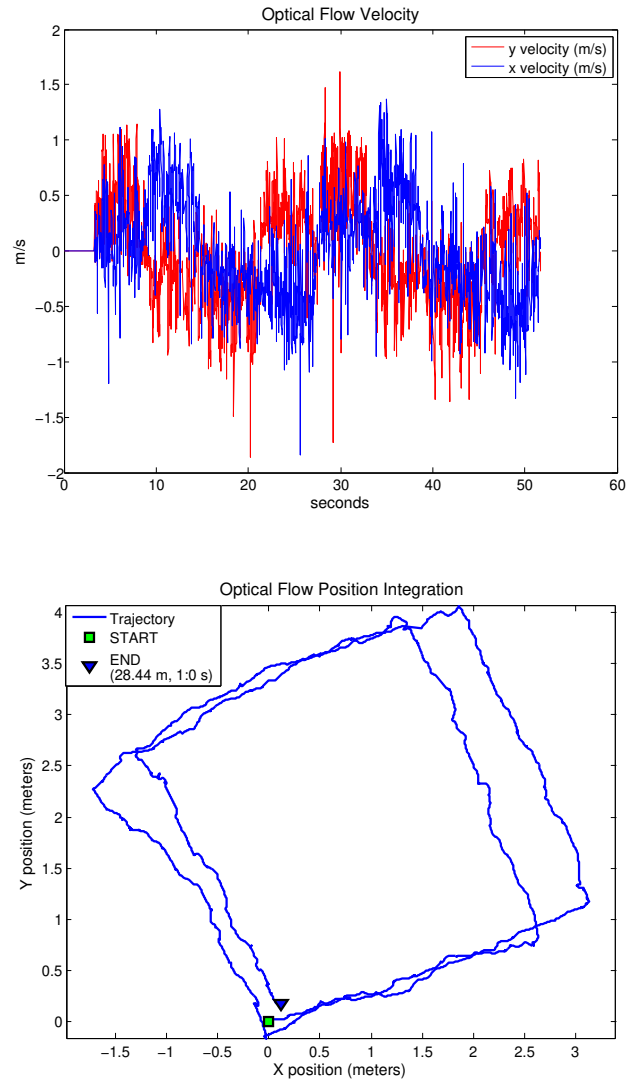

Fig. 7. Square shaped indoor trajectory, x and y velocity values measured during movement along square shaped trajectory (top), integrated x and y position values (bottom).

disc as constant flow source. Table II shows that the mouse sensor requires one order of magnitude more infrared lighting to provide a flow response. Also in the fluorescent lamp spectral band around 2700K it still requires almost one order of magnitude more light. The first value in the table indicates the value measured next to the sensor facing to the floor (Radiance), the second value is the intensity measured on the floor level, orienting the Lux meter towards the illumination source (Irradiance). The sensor to ground distance was 0.6 m and the flow target a rotating disk with coarse salt and pepper noise. The quality criterion for the mouse sensor was according to its (black-box) internal quality indicator. The quality criterion for PX4FLOW was that the mean of the measured speed agreed with the measurement under optimal lighting conditions.

TABLE II

MINIMUM ILLUMINATION FOR EACH SENSOR TYPE.

Sensor (Type)	Min. Illum. (CCFL)	Min. Illum. (Halogen)
ADNS-2080	47/800 Lux	34/680 Lux
PX4FLOW	6/45 Lux	3/25 Lux

D. Outdoor Test

The presented flow sensor was also compared to the ADNS-2080 mouse sensor. Both sensors were mounted on a rigid test rig and carried along an outdoor trajectory on typical road surface. The velocities were transformed to the earth oriented coordinate frame using the estimated attitude of the onboard IMU. The flow values of the mouse sensor were scaled with the height measurements of the PX4FLOW optical flow sensor. Figure 8 shows the integrated metric output of both sensors.

Fig. 9. Integrated metric output of the PX4FLOW sensor with orthophoto overlay for a flight along a promenade.
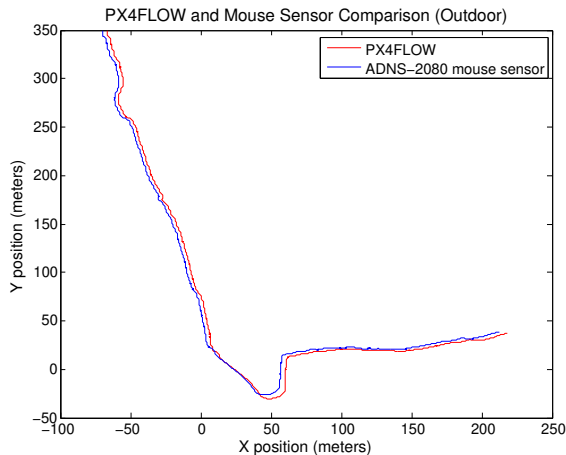

Fig. 8. Comparison of the ADNS-2080 mouse sensor and the PX4FLOW module on a long outdoor trajectory.

Both sensors show similar results on a several hundred meters long trajectory. The main benefit of the PX4FLOW module is the application in indoor and low light conditions. The resulting drift is mainly induced by metal structures and objects on the trajectory that influence the magnetometer and therefore affect the heading estimation.

Figure 9 shows the integrated metric output of the PX4FLOW sensor with orthophoto overlay for a manual flight with the PX4FMU autopilot on a 7"-propeller small quad rotor along a promenade in a park. The plot shows that the overall consistency with the aerial photo is very high and that the trajectories when closing the loop largely overlap. The estimated trajectory is the pure integration of the measured velocity at each time step. In order to purely show the sensor accuracy, we did not include any further filtering or any motion model. The overall trajectory had a length of 192.16 meters.

VIII. CONCLUSION

This paper has shown that a smart camera computing optical flow and compensating for rotations can be implemented with low-cost components. The flow estimation performance matches typical mouse sensors but works also without strong infrared lighting in typical indoor environments.
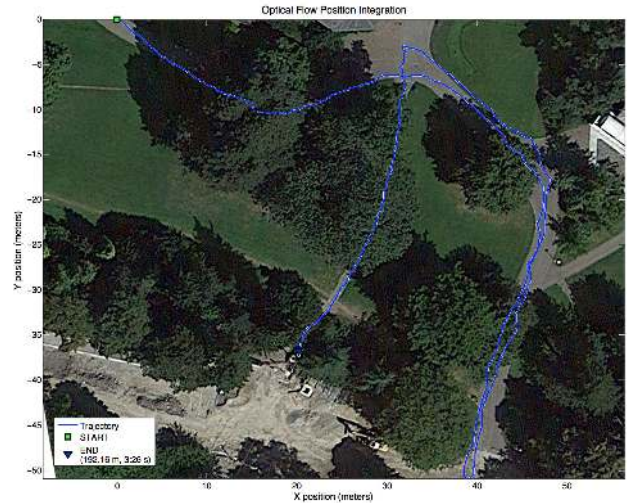
The software implementation and hardware design is open sourced and available to the scientific community. Since the system is light weight, it is suitable for any mobile robot or micro aerial application.

Future work will include the augmentation of the compass with a vision-based heading reference to overcome local magnetic anomalies, which are induced regularly by large metal objects in urban environments.

REFERENCES

[1] A. Beyeler, J.-C. Zufferey, and D. Floreano. Vision-based control of near-obstacle flight. *Autonomous Robots*, 27:201–219, 2009.
[2] J.-P. Bristeau, F. Callou, D. Vissiere, and N. Petit. The navigation and control technology inside the ar.drone micro uav. In *18th IFAC World Congress*, pages 1477–1484, Milano, Italy, 2011.
[3] F. Fraundorfer, L. Heng, D. Honegger, G. Lee, L. Meier, P. Tanskanen, and M. Pollefeys. Vision-Based Autonomous Mapping and Exploration Using a Quadrotor MAV. In *Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on*, oct 2012.
[4] D. Honegger, P. Greisen, L. Meier, P. Tanskanen, and M. Pollefeys. Real-time velocity estimation based on optical flow and disparity matching. In *Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on*, oct 2012.
[5] Farid Kendoul, Isabelle Fantoni, and Kenzo Nonami. Optic flow-based vision system for autonomous 3d localization and control of small aerial vehicles. *Robotics and Autonomous Systems*, 57(67):591 – 602, 2009.
[6] Cedric Pradalier, Samir Bouabdallah, Pascal Gohl, Matthias Egli, Gilles Caprari, and Roland Siegwart. The coax micro-helicopter: A flying platform for education and research. In Ulrich Rckert, Sitte Joaquin, and Werner Felix, editors, *Advances in Autonomous Mini Robots*, pages 89–99. Springer Berlin Heidelberg, 2012.
[7] D. Watman and H. Murayama. Design of a miniature, multi-directional optical flow sensor for micro aerial vehicles. In *Robotics and Automation (ICRA), 2011 IEEE International Conference on*, pages 2986 –2991, may 2011.
[8] S. Zingg, D. Scaramuzza, S. Weiss, and R. Siegwart. MAV navigation through indoor corridors using optical flow. In *Robotics and Automation (ICRA), 2010 IEEE International Conference on*, pages 3361 –3368, may 2010.
[9] J.-C. Zufferey, A. Beyeler, and D. Floreano. Autonomous flight at low altitude using light sensors and little computational power. *International Journal of Micro Air Vehicles*, 2:107–117, 2010.