Grzegorz BIESZCZAD, Sławomir GOGLER, Michał KRUPIŃSKI, Andrzej LIGIENZA, Krzysztof SAWICKI
INSTITUTE OF OPTOELECTRONICS, MILITARY UNIVERSITY OF TECHNOLOGY
2 gen. S. Kaliskiego Str., Warsaw, Poland

The concept of thermovision sensor supporting the navigation of unmanned aerial platforms

Abstract

The topic of this paper is a concept of a new type of a sensor intended for unmanned aerial vehicles (UAVs). Its operation is based on processing images acquired from a thermal camera operating in the long-wave infrared band (LWIR) placed underneath a vehicle's chassis. The vehicle's spatial displacement is determined by analyzing movement of characteristic thermal radiation points (ground, forest, buildings, etc.) in pictures acquired by the thermal camera. Magnitude and direction of displacement is obtained by processing stream of consecutive pictures with optical-flow based algorithm in real time. Radiation distribution analysis allows to calculate camera's self-translation vector. Advantages of measuring translation based on thermal image analysis is lack of drift effect, resistance to magnetic field variations, low susceptibility to electromagnetic interference and change in weather conditions as compared to traditional inertial navigation sensors. As opposed to visible light situational awareness sensors, it offers operation in complete darkness (harsh weather nights and indoors).

Keywords: Thermal-vision, UAV, navigation, situational awareness.

1. Introduction

Currently, GPS sensors, barometric, accelerometer and gyroscopic sensors are used as navigation sensors for unmanned vehicles. Sensors that derive position information from relative motion, acceleration or angular momentum are commonly referred to as inertial sensors. Image sensors can also be a rich source of relative motion information. Situational awareness sensors that are intended for obstacle detection and crash avoidance are becoming increasingly common. Such systems are often crafted using computer vision algorithms, processing signal form stereoscopic cameras. Vision sensors provide additional information to navigation system, thus increasing precision of UAVs position estimation. Such systems often employ cameras working in visible part of the spectrum. One of the first successful implementations of such a sensor is PX4FLOW [1]. Thermal inertial sensor that is under development in Military University of Technology is a new type of positioning sensor, that is intended not only to increase the accuracy and reliability of the navigation system, but additionally to enable acquisition of infrared imaging as complementary data. Thermal imaging enables navigation in complete darkness, harsh weather conditions or indoors. Additional visual information can increase reliability and robustness of obstacle detection and navigation system. The conceptual image of developed sensor installed in UAV is shown in Fig. 1.

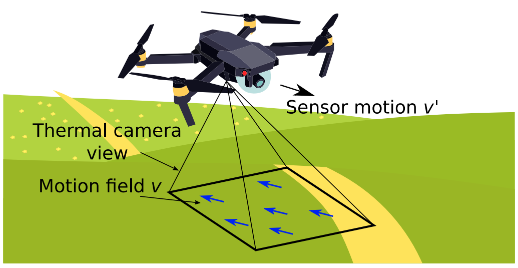

Fig. 1. Concept image of motion flow camera in the drone [www.freepik.com]

2. Equations of the Motion Field

Regardless of the image source used, it is possible to extract motion of points in the picture by optical flow algorithms. Vector field containing information about motion is called motion field. Motion field is created by projecting the 3D velocity field on the image plane. Let $p = [x, y, z]$ be a point in the object space of the camera, the optical axis be the z-axis and f denote the focal length of the camera objective. The center of projection is given by the intersection point of the optical axis with the principal image plane (or the rear nodal point). Projection $p' = [x', y', z']$ of the point p on the focal is given by:

$$\frac{p'}{f} = \frac{p}{z} \rightarrow p' = \frac{fp}{z} \quad (1)$$

Since the focal length f is the distance from the principal image plane to the image plane, the third coordinate of p' is constant: $p' = [x', y', f]$. Assuming classical perspective model of the camera the transformation can be modeled with matrix equation:

$$p' = K \cdot T \cdot M \cdot p \quad (2)$$

where: K – intrinsic camera parameters, T– perspective transformation matrix of the camera, M – geometry transformation matrix.

The perspective transformation equation results in three-dimensional result, hence to obtain position on the image plane one must reduce z component and calculate x, y components with equations:

$$x' \rightarrow \frac{p'_x}{p'_z}, \quad y' \rightarrow \frac{p'_y}{p'_z} \quad (3)$$

Geometry transformation matrix M can be factored to transformation matrices of distinct types of motion, that can be used to model motions of the camera like tracking and booming by translation parallel to image plane, dolling by translation in normal to image plane, rolling by rotating around image axis and finally panning and tilting by rotating in two other axes. Not all motions have to be taken into account in redundant navigation sensor, as it was described by Meier [2] where only translation in plane parallel to the ground and rotation around cameras' optical axis was considered. In such a case, relative motion of the camera and point p is reduced to:

$$v = -v_T - \omega p \quad (4)$$

where $\omega = [\omega_x, \omega_y, \omega_z]$, is the angular velocity and $v_T = [v_{Tx}, v_{Ty}, v_{Tz}]$ the translational component of the motion. Taking the derivative with respect to time of both sides of equation (1) leads to the relation between the velocity of p in the camera reference frame and the flow velocity of p' in the image plane:

$$v' = \frac{\mathrm{d}p'}{dt} = f\frac{z\frac{dp}{dt} - \frac{dz}{dt}p}{z^2} = f\frac{zv - v_z p}{z^2} = [v'_x, v'_y, v'_z] \quad (5)$$

The x and y components of the motion field after substituting (4) and considering image plane geometry (3) can be written as:

$$v'_x = \frac{v_{Tz}x - v_{Tx}f}{z} - \omega_y f + \omega_z y + \frac{\omega_x xy - \omega_y x^2}{f} \quad (6)$$

$$v'_y = \frac{v_{Tz}y - v_{Ty}f}{z} - \omega_x f + \omega_z x + \frac{\omega_x y^2 - \omega_y xy}{f} \quad (7)$$

The rotational parts are not dependent on the Z coordinate and therefore the angular velocity does not carry scene depth information.

The translational components in (6) and (7) are scaled with the focal length and the current distance Z to the scene what results from equation (3). If the Z distance is known (ex. from barometer or rangefinder) and rotational velocity is known and compensated one can compute translational velocity in real scale by:

$$V_m = v' \frac{Z}{f} \quad (8)$$

Above calculations are valid if the distance to the scene is constant in the field of view. This is especially the case if the camera is faced perpendicular to the ground. Inference about camera position is possible only with given motion of observed scene in image plane. Velocity of objects on the image plane can be determined using optical flow methods where it is derived from image to image displacement in time. Basic velocity estimation approach takes into the consideration a two consecutive frame analysis, where past frame is threated as reference frame. In constant frame rate camera like microbolometric thermal camera, the interval between frames is constant.

The motion field in real life cases is affected by other types of motions that can occur beside the translation and the rotation. It was simulated using frame reference approach, how motion in other degrees of freedom affects the motion field and what kind metrics can be used to quantify magnitude of such impact. Exemplary generated motion fields for different degrees of freedom are shown in Figure 2.

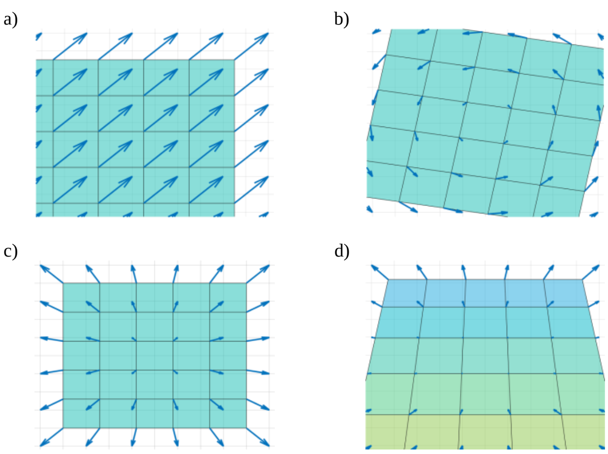

Fig. 2. Motion fields for a) tracking, b) rolling, c) dolling, d) tilting

For idealized case of motion field estimated from image, i.e. with equal reliability for every point in the picture, further inference of camera motion is possible by calculation of motion field metrics, where most of them are straight-forward adaptation of basic vector field theory: Camera tracking and booming is assumed to be a mean vector of all moves in the motion field (9), camera rolling can be calculated from rotation of motion field (10). Relative camera dolling can be calculated from divergence of the motion field (11).

$$\overline{V}_{xy} = \frac{1}{X \cdot Y} \iint_{x,y} V_{xy} \, \mathrm{d}x\mathrm{d}y \quad (9)$$

$$\mathrm{rot}(V_{xy}) = \frac{1}{2\sqrt{X \cdot Y}} \iint_{x,y} \nabla \times V_{xy} \, \mathrm{d}x\mathrm{d}y \quad (10)$$

$$\mathrm{div}(V_{xy}) = \frac{1}{X \cdot Y} \iint_{x,y} \nabla \cdot V_{xy} \, \mathrm{d}x\mathrm{d}y \quad (11)$$

Camera tilting and panning are more complicated cases that occurs in UAVs, where perspective transformation of scene takes part. To extract pan/tilt motion a perspective model described earlier with equation (2) and (3) has been assumed with extension to calculate the distance between original and transformed position:

$$\Delta p' = p' - K \cdot T \cdot \bar{A} \cdot M_{\varphi} \cdot A \cdot p' \quad (12)$$

where: M_{φ} – rotation by φ angle in x axis matrix, $\bar{A}$, A – complementary translation helper matrices.

Information about tilt angle in perspective transformation is carried in both Δp_x and Δp_y shifts on the image and both vectors. This information can even be extracted back. Solving the Δp_x and Δp_y for φ gives solution for x (13) and y (14) shifts in the image however for tilting around the x axis, only the x component is algebraically solvable.

$$\varphi_x = -\sin^{-1}\left(\frac{1[m] \cdot \Delta p_x}{\Delta p_x x - xy}\right) \quad (13)$$

$$\sin \varphi_y = -\frac{1[m] \cdot \cos \varphi_y y - y + \Delta p_y}{\Delta p_y y - y^2} \quad (14)$$

For y component the solution leads to implicit function (14). To solve for φ_y one can use approximation of cosine function for small angles assuming $\cos(\varphi) = 1$. This makes possible to solve for φ_y with regard to perspective transformation:

$$\varphi_y = -\sin^{-1}\left(\frac{\Delta p_y}{\Delta p_y y - y^2}\right) \quad (15)$$

For panning around the y axis, the situation is reversed. Linearization for small angles (smaller than about 5°) is possible in this case, because frame to frame tilt in practical situations is limited. With 60 Hz array readout linearization error smaller than 1° the maximum tilt angle velocity is 300°/s which is acceptable for nonracing drones. The resultant tilt angle is a mean of extracted angles from the pixels in the whole motion field.

Described earlier metrics were computed for artificially generated motion fields and results of developed metrics are summarized in Table 1.

Tab. 1. Developed motion estimation metrics results for distinct types of motion

	$\overline{V}_{xy}$	$\mathrm{rot}(V_{xy})$	$\mathrm{div}(V_{xy})$	$\overline{\varphi_x}$	$\overline{\varphi_y}$
Tracking 10%	0.1	0	0	16-9i	16-9i
Tracking 5%	0.05	0	0	13-8i	13-8i
Rolling 10°	0	0.191	0	0	0
Rolling 5°	0	0.0915	0	0	0
Dolling 10%	0	0	-0.04	0	0
Dolling 5%	0	0	-0.02	0	0
Tilting 10°	0.035	0	-0.0064	10.08	0
Tilting 1°	0,007	0	10^{-5}	1.0001	0

Analyzed motions are: tracking by chosen percentage of image size, rolling with chosen angle, dolling by percentage of distance and tilting by a chosen angle. Results leads to conclusions that:

- Mean motion vector correctly indicates movement of the camera but can also be non-zero on camera tilt because of nonsymmetrical nature of perspective transformation.
- Rotation metric is robust to other types of motion and can be reliably used to estimate angular velocity with respect to cameras' optical axis.

- Divergence of the field properly indicates the motion on *Z* axis but can manifest itself at the occasion of large tilt angles, where perspective distortion becomes significantly asymmetrical.
- Proposed tilt angle estimation can correctly compute motion based on reverse perspective transformation equations but gives incorrect results (ex. complex angles that have no physical meaning) in presence of global, translational camera motion. Future development must focus on computation of complex motions and managing them independently.

Practical measurements suffer from non-equal reliability, because of noise on the image and plateau parts of the image where motion is impossible to extract. Measured motion field extracted from infrared image sequence using Farneback method [3] is shown in Figure 2.

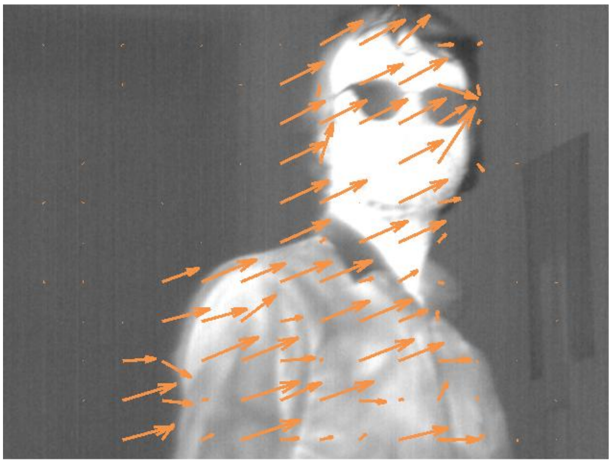

Fig. 2. Optical-Flow estimated from infrared image using Farneback method

In such a real-life case, an additional measure like weight function discriminating reliable data points will be necessary to use developed motion metrics.

3. Device architecture

Navigation sensor contains image sensor and demands computationally intensive image processing in real-time. Architecture of the device must ensure sufficient computational power and flexibility with low size and power consumption at the same time. General architecture of the device includes three main functional elements: sensor module, data processing module and interface module. The sensor module has the task of powering and providing an interface to the infrared sensor array. The data processing module contains programmable FPGA and microprocessor for processing data acquired from the sensor. The interface module brings the signal from the sensor to the flying system's executive systems (eg. on-board autopilot) or telemetry system transmitting data to the remote flight control system.

The detector in the sensor module is an array of uncooled microbolometer detectors. These types of detector arrays are increasingly used in thermal imaging cameras due to relatively low cost and good operational parameters. The microbolometric matrices achieve the temperature resolution NETD (Noise Equivalent Temperature Difference) of 40 mK for an array temperature of 300 K, the optical system F/# = 1.0 and frame reading frequency of 60 Hz. The arrays planned to be evaluated in the project are the Ulis detector types Atto320, Atto640 which are 12 µm detector types and Pico 386 gen2 which is a mature reliable 17 µm type detector. Detector array outputs analog signal that is proportional to the incident infrared radiation which is then converted to digital signal using AD converter. Every infrared detector array has its own unique undesirable characteristic, so-called nonuniformity [4], which has to be digitally compensated for by the internal computing system of the camera. Properly prepared infrared image is used to extract motion field.

Digital image processing system is a developed version of previously proved architecture with use of SoC-FPGA device [5]. The purpose of SoC-FPGA based design for image processing, is to provide hardware and system capabilities for highly specialized thermal image processing. Extensive programmable resources enable customization of hardware peripherals that are needed for interfacing with infrared detector module. Additionally, it provides configurable digital logic for image processing acceleration, especially for computationally intensive optical-flow algorithms. The preprocessed image is then transferred to Hard Processors' System (HPS) operating memory by means of a dedicated frame grabbing DMA. The data flow in developed architecture is shown in Figure 3.

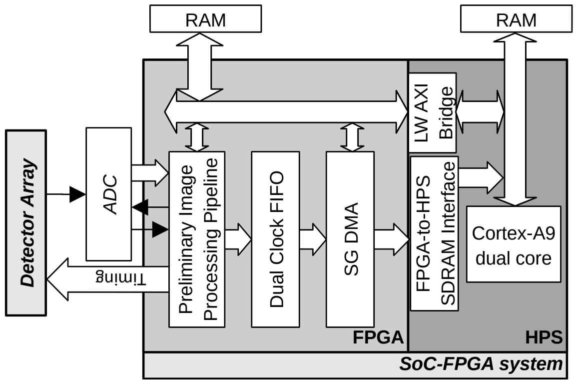

Fig. 3. Architecture of SoC realizing thermal image processing in navigation sensor

Developed navigation sensor is designed to use SoC-FPGA device type 5CSEMA5F31C6N Cyclone V SE which is designed in 28-nm technology and integrates dual-core ARM Cortex-A9 processor, peripherals and FPGA fabric comprising 85k Logic Elements, 288 IOs and 4.45Mb of internal memory.

Dual core processor system is intended to be used in image processing that is hard to be paralleled or accelerated in digital logic and will also serve as the communication and control manager. Embedded Linux operating system installed on the processor part of the SoC is planned to provide hardware handling, network stack and image processing application management.

Device is being developed using TERASIC DE10-Standard Evaluation Kit and specially developed proximity board capable of reading from Pico 386 gen2 or Atto320 Ulis detectors. Specially developed optical system with F# = 1/1.22 and focal length of 35 mm provides optical interface to the detector array. Image is being preprocessed by the FPGA and HPS hardware and software modules. Image can be exported by means of Ethernet network or JTAG interface and processed offline on external PC. Data integrity, preliminary optical-flow implementations and motion flow analysis algorithms described in the paper has been tested in MATLAB software. The image of digital evaluation kit with proximity board installed is shown in Figure 4.

Fig. 4. Test platform to evaluate device architecture and algorithms

4. Conclusions

Developed concept of device architecture for navigation sensor based on infrared image optical-flow analysis has been described. Fundamental functions of developed sensor like image acquisition, infrared image optical flow and basic motion filed analysis has been successfully evaluated on devices' architecture mockup. Device is still under development.

Scientific work is financed by The National Centre for Research and Development in project no DOB-2P/02/09/2018.

5. References

[1] Honegger Dominik, Meier Lorenz, Tanskanen Petri, Pollefeys Marc: An Open Source and Open Hardware Embedded Metric Optical Flow CMOS Camera for Indoor and Outdoor Applications. International Conference on Robotics and Automation (ICRA 2013) in Karlsruhe, Germany, 2013.

[2] Meier Lorenz: Dynamic Robot Architecture for Robust Realtime Computer Vision. ETH Zürich, Doctoral Thesis https://doi.org/10.3929/ethz-a-010874068, 2017

[3] Farneback G: Two-Frame Motion Estimation Based on Polynomial Expansion. In Proceedings of the 13th Scandinavian Conference on Image Analysis, 363 - 370. Halmstad, Sweden: SCIA, 2003.

[4] Krupiński Michał, Bieszczad Grzegorz, Gogler Sławomir, Madura Henryk: Non-uniformity correction with temperature influence compensation in microbolometer detector. Proc. SPIE 9481, 948113, doi: 10.1117/12.2184645, 2015

[5] Bieszczad Grzegorz: SoC-FPGA embedded system for real-time thermal image processing. Proceedings of the 23rd International Conference Mixed Design of Integrated Circuits and Systems, MIXDES 2016, Article number 7529788, pp. 469-473, doi: 10.1109/MIXDES.2016.7529788, 2016

Received: 12.10.2018 *Paper reviewed* *Accepted: 03.12.2018*

Grzegorz BIESZCZAD, PhD, eng.

Grzegorz Bieszczad graduated in 2008 in the Military Technical Academy (2008) at the faculty of Electronics with a specialty of ICT. PhD degree received in 2012. Currently works on issues related to the design of digital systems, programming of microprocessors and FPGAs related to digital image processing, including thermal images.

e-mail: grzegorz.bieszczad@wat.edu.pl

Sławomir GOGLER, MSc, eng.

Military University of Technology Warsaw, graduated in 2011 from Warsaw University of Technoolgy, faculty of Mechatronics. He works on optical and optomechanical design and FEM simulations involving heat flow, magnetic fields and radiative transfer.

e-mail: slawomir.gogler@wat.edu.pl

Michał KRUPIŃSKI, MSc, eng.

He received his MSc degree from the Military University of Technology, in 2010. He works at the same university. His current interests are in the design and programming of digital systems, digital processing and thermographic imaging.

e-mail: michal.krupinski@wat.edu.pl

Andrzej LIGIENZA, MSc, eng.

Andrzej Ligienza graduated BSc in the Military University of Technology (2014) at the faculty of Mechatronics with a specialty of ISM. MSc graduated in 2016 at the faculty of Electronics. Currently PhD student at faculty of Optoelectronics.

e-mail: andrzej.ligienza@wat.edu.pl

Krzysztof SAWICKI, MSc, eng.

Krzysztof Sawicki graduated at Faculty of Electronics of Military University of Technology in 2009 with specialty of ICT networks. Currently works on network steganography especially in wireless networks.

e-mail: krzysztof.sawicki@wat.edu.pl