1. Introduction

Autonomous visual navigation, i.e. determination of position, attitude and velocity (ego motion) by processing of the images from onboard camera(s), is essential for mobile robots control even in the presence of GPS networks, as the accuracy of GPS data and/or the available map of surroundings can be insufficient. Besides, GPS signals reception can be unstable in many locations (inside buildings, tunnels, in narrow streets, canyons, under trees, etc).

Up to now most of the practical visual navigation solutions have been developed for ground robots moving in cooperative and/or well determined environment. However, future generations of mobile robots should be also capable of operating in complex and non-cooperative 3D environments. Visual navigation in such conditions is much more challenging, especially for flying robots, where full 6DOF pose/motion should be determined. Generally 3D environment perception is required in this case, i.e., determination of a local depth map for the visible scene.

3D scene information can be obtained by stereo imaging; however this solution has certain limitations. It requires at least two cameras, precisely mounted with a certain stereo base (can be critical for small vehicles). Due to fixed stereo base the range of the depth determination with stereo imaging is limited. A more universal solution with less hardware requirements can be achieved with optical flow processing of sequential images from a single onboard camera.

The ego motion of a camera rigidly mounted on a vehicle is mapped into the motion of image pixels in the camera focal plane. This image motion is commonly understood as image flow or optical flow (OF) (Horn & Schunck, 1981). This vector field of 2D image motion can be used efficiently for 3D environment perception (mapping) and vehicle pose/motion determination as well as for obstacle avoidance or visual servoing. The big challenge for using the optical flow in real applications is its computability in terms of its density (sparse vs. dense optical flow), accuracy, robustness to dark and noisy images and its real-time determination. The general problem of optical flow determination can be formulated as the extraction of the two-dimensional projection of the 3D relative motion into the image plane in form of a field of correspondences (motion vectors) between points in consecutive image frames.
This article addresses a real-time solution for high precision optical flow computation based on 2D correlation of image fragments on the basis of an optical correlator. It exploits the principle of Joint Transform Correlation (JTC) in an optoelectronic setup using the Optical Fourier Transform (Goodman, 1968). Based on the experience of the authors with different successful optical processor developments (Tchernykh et al., 2004, Janschek et al., 2004a, Tchernykh et al., 2000, Janschek et al., 2005a) a new optical processor design is presented, which makes use of advanced optoelectronic technology. The proposed optoelectronic optical flow processor (OE-OFP) shows to be very compact with low mass and low power consumption and provides the necessary high performance needed for navigation applications in the field of robotics (ground, aerial, marine) and space flight (satellites, landing vehicles). The paper recalls briefly the underlying principles of optical flow computation and optical correlation, it shows the system layout and the conceptual design for the optoelectronic optical flow processor and it gives preliminary performance results based on a high fidelity simulation of the complete optical processing chain.

2. Requirements to Optical Flow Processor

Considering a flying platform moving in a complex non-cooperative 3D environment (indoor or outdoor) as a target mission, the following requirements to the Optical Flow Processor (OFP) can be formulated.

1. **Image quality tolerance.** As various illumination conditions can be expected, the OFP should be able to process dark and noisy images with low texture contrast.

2. **Optical flow density.** The required resolution of the optical flow fields depends on scene complexity. On the authors’ experience, the depth information should be obtained for at least 32x32 (better 64x64) locations for adequate perception of complex 3D environment (required for navigation). This means, that at least 32x32 (better 64x64) optical flow vectors should be determined for each frame.

3. **Frame rate.** Considering relatively high motion dynamics of the flying robot, processing of up to 10 frames per second is required for navigation purposes.

4. **Accuracy.** Considering a maximal acceptable error of 10 percent for local depth determination to get a reasonable 3D environment perception, the error of the OF vectors determination should be also within 10 percent. This means that OF vectors with magnitude of a few pixels should be determined with sub pixel accuracy (errors should be within a few tenths of a pixel).

5. **Size/mass/power consumption.** To allow installation onboard a flying platform these parameters should be minimized. Roughly the volume of the OFP should be within a few tens of cubic centimetres, mass – within a few tens of grams and power consumption – within a few watts.

3. Existing solutions (optical flow determination background)

The problem of the optical flow computation is being investigated for more than two decades. Many methods for the OF determination have been developed (Beauchemin & Barron, 1995, Bruhn et al., 2003, McCane et al, 1998, Liu et al., 1998). All these methods have in common, that rather dense and accurate OF needs low noise images and requires high computational power, which is hardly realizable with embedded processors (Liu et al., 1998). Existing pure digital high performance solutions based on conventional PC or FPGA
Optical Correlator based Optical Flow Processor for Real Time Visual Navigation

technology (Bruhn et al., 2003, Bruhn et al., 2005, Diaz et al., 2006) additionally consume a lot of power, mass and volume which does not fit the requirements of mobile robotics, especially if application onboard a flying platform is considered. The recently developed and currently very popular SIFT approach (Lowe, 1999, Se et al., 2001) allows a computationally efficient determination of more or less sparse OF fields in well structured environments. Some specialized high speed OF sensors on hybrid analogue-digital technology (Barrows & Neely, 2000, Zufferey, 2005) provide even super real-time performances but are suffering from the required accuracy of the OF vectors for navigation purposes. A most robust approach is the area correlation, applied originally for image registration (Pratt, 1974). Area correlation uses the fundamental property of the cross-correlation function of two images, which gives the location of the correlation peak directly proportional to the displacement vector of the original image shift.

For each pair of sequential images the OF field is determined by subdividing both images into small fragments and 2D correlation of corresponding fragments. As a result of each correlation the local shift vector at the specific location is determined; a whole set of local shift vectors forms an optical flow matrix (Figure 1).

![Figure 1. Principle of correlation based optical flow determination](image)

The optical flow determination method, based on 2D correlation of the image fragments, offers a number of advantages:

- high sub pixel accuracy;
- low dependency on image texture properties (no specific texture features required);
- high robustness to image noise (suitable for short exposures and/or poor illumination conditions);
- direct determination of multi-pixel shifts (suitable for fast image motion).

Simultaneously this method requires a very large amount of computations which prevents practically this method from a real time realization with conventional digital processors onboard a flying robot.

Generally, none of the existing OF determination techniques satisfies all of the requirements to the Optical Flow Processor, suitable for installation onboard a flying platform (listed in section 2). To reach the real time performance and to satisfy the strict size/mass/power limitations while keeping the accuracy and robustness of the 2D correlation based approach, we propose to perform the correlations with an onboard optical correlator.
4. Optical correlator technology

A very efficient method for 2D correlation requiring only double Fourier transform without phase information is given by the Joint Transform Correlation (JTC) principle (Jutamulia, 1992).

The two images \( f_1(x, y) \) and \( f_2(x, y) \) to be correlated are being combined to a joint image \( I(x, y) \) (Figure 2). A first Fourier transform results in the joint power spectrum \( S(u, v) = F[I(x, y)] \). Its magnitude contains the spectrum \( F(u, v) \) of the common image contents augmented by some periodic components which are originating from the spatial shift \( \vec{G} \) of \( f_1 \) and \( f_2 \) in the joint image \( I \). A second Fourier transform of the squared joint spectrum \( J(u, v) = S(u, v)^2 \) results in four correlation functions. The centred correlation function \( C_g(x, y) \) represents the auto-correlation function of each input image, whereas the two spatially shifted correlation functions \( C_{xG}(x, y) \) represent the cross-correlation functions of the input images. The shift vector \( \vec{G} \) contains both the technological shift according to the construction of the joint image \( I(x, y) \) and the shift of the image contents according to the image motion. If the two input images \( f_1 \) and \( f_2 \) contain identical (but shifted) image contents, the cross-correlation peaks will be present and their mutual spatial shift \( \Delta = \vec{G} - (-\vec{G}) \) allows determining the original image shift in a straightforward way.

This principle can be realized in hardware by a specific optoelectronic setup, named Joint Transform Optical Correlator (JTOC). The required 2D Fourier transforms are performed by means of diffraction-based phenomena, incorporating a so called optical Fourier processor (Goodman, 1968). A laser diode (Figure 3) generates a diverging beam of coherent light which passes a single collimating lens focusing the light to infinity. The result is a beam of parallel light with plane wave fronts. The amplitude of the plane wave front is modulated by a transmissive or reflective spatial light modulator (SLM). The SLM actually works as a diffraction grid and the resulting diffraction pattern can be made visible in the focal plane of a second lens (Fourier lens). Under certain geometric conditions the energy distribution of this pattern is equal to the squared Fourier transform (power spectrum) of the modulated wave front. The power spectrum can be read by a CCD or CMOS image sensor located in the focal plane of the Fourier lens of the optical Fourier processor. The position of the correlation peaks in the second power spectrum (correlation image) and the associated shift value can be measured with sub-pixel accuracy using e.g. standard algorithms for centre of mass calculation.

Optical processing thus allows unique real time processing performances of high frame rate video streams.

This advanced technology and its applications have been studied during last years at the Institute of Automation of the Technische Universität Dresden (Tchernykh et al., 2004, Janschek et al., 2004a, Janschek et al., 2007). Different hardware models have been manufactured, e.g. under European Space Agency (ESA) contracts (Figure 4). Due to special design solutions owned by TU Dresden, the devices are robust to mechanical loads and deformations. (Tchernykh et al., 2000, Janschek et al., 2005a). One of the models has been
successfully tested in an airborne test campaign, where very promising performances have been shown (Tchernykh et al., 2004).

Figure 2. Principle of the joint transform correlation

\[ f(x,y) = \left( e^{-\frac{G_1}{2}} \right) + \left( e^{-\frac{G_2}{2}} \right) + \left( e^{-\frac{G_3}{2}} \right) + \left( e^{-\frac{G_4}{2}} \right) \]

\[ F(u,v) = |F(x,y)|^2 \cos(2\pi u G_1 + 2\pi v G_2) \]

\[ C(x,y) = 2C_1(x,y) + C_2(x+G_1,y+G_2) + C_3(x-G_3,y-G_4) \]

Figure 3. Principle of the Joint Transform Optical Correlator (JTOC)

Figure 4. Hardware model of an optical correlator
5. Determination of the optimal size of correlated fragments

5.1 Simulation experiment description
The dimensions of correlated fragments determine both the accuracy and reliability of the 2D correlation operation. Larger window size improves the reliability of the optical flow determination in poorly textured image areas and reduces the errors of the obtained OF vectors. At the same time, increasing the correlated fragments size smooths the obtained OF field, it suppresses small details and produces additional errors in areas with large variations of local depth.

The goal of the simulation experiment was to estimate the optimal size of the correlated fragment, making the best compromise between the accuracy/reliability of correlation and preservation of small details of the underlying 3D scene.

The experiment has been performed with a high fidelity software model of the proposed optical flow processor. The model includes the detailed model of the complete processing chain of the optical correlator. Image processing algorithms simulate all relevant operations of the optoelectronic hardware realization of the optical correlator (optical diffraction effects, dynamic range limitation and squaring of the output images by image sensor, scaling of the output images according to focal length value, etc.).

The experiment has been performed using simulated images from synthetic 3D scenes of a planetary surface generated during an ESA (European Space Agency) study on the visual navigation of a planetary landing vehicle (Janschek et al., 2005b). The images contain parts of rich texture as well as flat low texture regions and dark shadows. The image sequence of an inclined landing trajectory (example image see Figure 5) has been generated on base of a 3D model of the landing site using standard ray tracing software considering the Modulation Transfer Function (MTF) of the camera as well as photonic and readout noise and pixel response non-uniformity.

![Figure 5. Synthetic 3D scene for testing and performance evaluation](image-url)
5.2 Simulation experiment results

The OF fields have been determined with a correlated fragments size varying in the range from 8x8 to 64x64 pixels and they have been compared with a reference (ideal) OF field to determine the OF errors. The reference OF field has been produced directly from the reference trajectory data and the known 3D model of the landing site. Figure 6 shows the results of the optical flow accuracy sensitivity for different correlation window sizes. Images in the left column represent the 2D patterns of the OF vectors magnitudes (brighter pixels represent larger OF vectors), the middle row contains the error patterns, determined as the difference between the reference (ideal) and test OF fields. RMS error values are shown in the diagram at the bottom right corner. According to this sensitivity analysis, minimal OF errors are expected for a window size of 24x24 pixels.

For the selected window size (24x24 pixels) the sensitivity to additive and multiplicative image noise on the OF error has been investigated. It has been found, that random noise with standard deviation within 8% of average image brightness (signal-to–noise ratio above 12 dB) has little influence on the OF field accuracy. Starting from $\sigma = 8\%$, however, the effect of image noise rapidly increases. According to these results, the limit of acceptable image noise for optical flow determination with fragments size 24x24 can be set to $\sigma = 8\%$ of average image brightness.

6. Optical flow processor concept

Based on the result of previous theoretical studies and experimental works (software simulations and hardware models testing) a detailed concept of a compact OptoElectronic Optical Flow Processor (OE-OFP), suitable for installation onboard a flying robot has been developed.
The main purpose of the OE-OFP is the real-time determination of the optical flow field for the visible surrounding environment. Figure 7 shows the general data flow chart for the optical flow computation according to the Joint Transform Correlation (JTC) principle.

Figure 7. Main operations for optical flow determination according to JTC principle

The operations of 2D Fourier transform are most time and resource consuming for a digital realization of the OF processor. With optical realization however, Fourier transform is practically performed instantly (with “speed of light”) and requires power only for SLM illumination. In this case other operations in the data processing chain (images readout/loading, fragments cut and correlation images processing) are practically determining the limits of performance improvement and size/mass/power minimization. Therefore, optimization of these operations is particularly essential for optimal design of the optoelectronic OF processor.

The concept of the OE-OFP has been developed assuming a realization of the input/output and digital processing operations directly on the image sensors and SLM chips. This solution eliminates the need for a dedicated digital processing electronics and reduces dramatically the power consumption.

The unpackaged chips can be mounted close to each other on a single substrate (Chip-on-Board – COB mounting). A small distance between the dies (Figure 8) is offering direct chip-to-chip connections. This avoids the need for powerful buffers inside the processor and in consequence reduces further the OFP power consumption. As the processor outputs only the OF vectors coordinates, the output data rate and therefore the power consumption of the output buffers are also limited.

Figure 8. Realization concept of the OE-OFP

The optical system of the OF processor has been designed using a reflective SLM, which modulates the phase of the reflected wave front. To reduce the overall processor size and to increase the mechanical stability, a folded optical system design on the base of a small block of glass or optical plastic is currently considered. The small dimensions of the optical system allow a realization of the whole OF processor including an interface board and the lens in a compact housing, suitable for installation on a flying platform (Figure 9).
The operation of the presented optoelectronic processor is briefly explained in the following. The lens forms an image of the surrounding environment on the input image sensor. After exposure, the image data are recorded in an on-chip memory within the image sensor. The fragments for correlation are cut from two sequential frames according to the pre-programmed pattern – this operation is also performed within the input image sensor. The fragments prepared for correlation are sent to the SLM. Coherent light, emitted by a laser diode, reflects from the aluminized side of a glass block and illuminates the SLM surface via the embedded lens (can be formed as a spherical bulb on the surface of the block). The phase of the wave front reflected from the SLM, is modulated by the input image. It is focused by the same lens and forms (after intermediate reflection) the amplitude image of the Fourier spectrum of the input image on the surface of the Spectrum/Correlation Image Sensor (SCIS). After a second optical Fourier transform, the correlation image is obtained. The optical flow vector (equal to the shift between the correlated fragments) is calculated from the correlation peaks positions within the correlation image. This operation is performed directly inside the SCIS chip. The coordinates of the OF vectors are sent to the output buffers, installed on a small printed board.

The expected performances of the OE-OFP (Table 1) have been estimated on the base of the conceptual design of the processor and the results of simulation experiments, taking into account also the test results of the existing hardware models of the optical correlator developed within previous projects (Tchernykh et al., 2004, Janschek et al., 2004a).

<table>
<thead>
<tr>
<th>Input</th>
<th>3D scene</th>
</tr>
</thead>
<tbody>
<tr>
<td>Output</td>
<td>optical-flow fields</td>
</tr>
<tr>
<td>Optical-flow resolution (max)</td>
<td>64x64=4096 vectors/field</td>
</tr>
<tr>
<td>Optical-flow resolution (min)</td>
<td>8x8=64 vectors/field</td>
</tr>
<tr>
<td>OF fields rate @ 4096 vectors/field</td>
<td>10 fields/s</td>
</tr>
<tr>
<td>OF fields rate @ 64 vectors/field</td>
<td>500 fields/s</td>
</tr>
<tr>
<td>Processing delay</td>
<td>One frame (0.002 … 0.1 s)</td>
</tr>
<tr>
<td>Inner correlations rate</td>
<td>50000 correlations/s</td>
</tr>
<tr>
<td>OF vectors determination errors</td>
<td>σ = 0.1 … 0.25 pixels</td>
</tr>
<tr>
<td>OF processor dimensions</td>
<td>50x20x8 mm (w/o lens)</td>
</tr>
<tr>
<td>OF processor mass</td>
<td>within 20g (w/o lens)</td>
</tr>
<tr>
<td>Power consumption</td>
<td>within 2 W</td>
</tr>
</tbody>
</table>

Table 1. Expected performances of the Optoelectronic Optical Flow Processor
Comparison of Table 1 with the requirements listed in section 2 shows that the proposed optoelectronic Optical Flow processor is expected to satisfy the requirements, listed in section 2. To compare the proposed processor with other currently available solutions for real time optical flow determination, it is however important to evaluate a performance measure related to mobility, which takes into account also the processor power consumption and volume related to the computing performance in terms of flow vectors per second and accuracy.

Figure 10 shows these performance-to-mobility measures taking into account also the power consumption and the volume of the optical-flow processor module. It follows that the proposed optoelectronic optical flow processor design (OE-OFP) shows unique performances in comparison with the fastest digital optical-flow computation solution currently available (Bruhn et al., 2005, Diaz et al., 2006).

![Figure 10. Performance-to-mobility comparison of optical flow processors](image)

7. Application areas

The proposed optical flow processor is intended to be used mainly in the field of visual navigation of mobile robots (ground, aerial, marine) and space flight (satellites, landing vehicles). The small size, mass and power consumption makes the proposed OE-OFP particularly suitable for application onboard micro air vehicles (MAVs).

From the obtained optical flow, 3D information can be extracted and a 3D model of the visible environment can be produced. The considerable high resolution (up to 64x64 OF vectors) and very high accuracy (errors $\sigma \leq 0.25$ pixels) of the determined optical flow makes such 3D environment models detailed and accurate. These 3D environment models can be used for 3D navigation in complex environment (Janschek et al., 2004b) and also for 3D mapping, making the proposed OF processor ideally suited for 3D visual SLAM. The applicability of the optical flow data derived with the proposed principles (joint transform correlation) and technology (optical correlator) to real world navigation solutions even under unfavourable constraints (inclined trajectories with considerable large perspective distortions) has been proved by the authors in recent work (Janschek et al., 2005b, Tchernykh et al., 2006), some simulation results are also given in the next section.

The anticipated real time performance of the processor (up to 500 frames/s with reduced OF field resolution) provides a wide range of opportunities for using the obtained optical flow for many additional tasks beyond localization and mapping, e.g. vehicle stabilization, collision avoidance, visual odometry, landing and take-off control of MAVs.
8. Application example: visual navigation of the outdoor UAV

The concept of visual navigation for a flying robot, based on 3D environment models matching has been proposed by the authors (Janschek et al., 2005b, Tchernykh et al., 2006) as one of the most promising applications of high resolution real time optical flow. 3D models of the visible surface in the camera-fixed coordinate frame will be reconstructed from the OF fields. These models will be matched with the reference 3D model with known position/attitude (pose) in a surface-fixed coordinate frame. As a result of the matching, the reconstructed model pose in the surface-fixed frame will be determined. With position and attitude of the reconstructed model known in both camera-fixed and surface-fixed frames, the position and attitude of the camera can be calculated in the surface-fixed frame.

Matching of 3D models instead of 2D images is not sensitive to perspective distortions and is therefore especially suitable for low altitude trajectories. The method does not require any specific features/objects/landmarks on the terrain surface and it is not affected by illumination variations. The high redundancy of matching of the whole surface instead of individual reference points ensures a high matching reliability and a high accuracy of the obtained navigation data. Generally, the errors of vehicle position determination are expected to be a few times smaller than the resolution of the reference model.

To prove the feasibility of the proposed visual navigation concept and to estimate the expected navigation performances, a software model of the proposed visual navigation system has been developed and an open-loop simulation of navigation data determination has been performed.

A simulation environment has been produced using the landscape generation software (Vue 5 Infinity from e-on software) on the base of 3D relief, obtained by filtering of a random 2D pattern. Natural soil textures and vegetation have been simulated (with 2D patterns and 3D models of trees and grass), as well as natural illumination and atmospheric effects (Figure 11). A simulation reference mission scenario has been set up, which includes the flight along a predetermined trajectory (loop with the length of 38 m at a height about 10 m over the simulation terrain).

Figure 11. Simulation environment with UAV trajectory (side and top views)

Simulated navigation camera images (Figure 12) have been rendered for a single nadir-looking camera with a wide angle (fisheye) lens (field of view 220°), considering the simulated UAV trajectory.
A reference 3D model of the terrain has been produced in a form of Digital Elevation Model (DEM) by stereo processing of two high altitude images (simulating the standard aerial mapping). Such model can be represented by a 2D pseudo image with the brightness of each pixel corresponding to the local height over the base plane.

The optical flow determination has been performed with a detailed simulation model of the optical correlator. The correlator model produces the optical flow fields for each pair of simulated navigation camera images, simulating the operation of the real optical hardware. Figure 13 shows an example of the optical flow field. The 3D surface models have been first reconstructed as local distance maps in a camera-fixed coordinate frame (Figure 13), then converted into DEMs in a surface-fixed frame using the estimated position and attitude of the vehicle. Figure 14 shows an example of both the reconstructed and reference DEMs.

Navigation data (position, attitude and velocity of the robot) have been extracted from the results of the matching of the reconstructed and reference models and compared with the reference trajectory data to estimate the navigation errors. As a result of the test, the RMS position error for the translation part of the trajectory was 0.20 m and the RMS attitude error was 0.45 degrees. These have been obtained by instantaneous processing of the optical flow.
data, i.e. without any time filtering, and without any additional navigation aids (except the DEM reference map). The navigation accuracy can be further improved by some filtering, and by using data from inertial measurement unit.

Figure 14. Reference and reconstructed DEMs

9. Summary and conclusions

The conceptual design of an advanced embedded optical flow processor has been presented. Preliminary performance evaluation based on a detailed simulation model of the complete optical processing chain shows unique performances in particular applicable for visual navigation tasks of mobile robots. The detailed optoelectronic design work is currently started.

10. References


Computer Vision is the most important key in developing autonomous navigation systems for interaction with the environment. It also leads us to marvel at the functioning of our own vision system. In this book we have collected the latest applications of vision research from around the world. It contains both the conventional research areas like mobile robot navigation and map building, and more recent applications such as, micro vision, etc. The first seven chapters contain the newer applications of vision like micro vision, grasping using vision, behavior based perception, inspection of railways and humanitarian demining. The later chapters deal with applications of vision in mobile robot navigation, camera calibration, object detection in vision search, map building, etc.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following:
