Tracking of Dynamic Objects Based on Optical Flow

Torsten Radtke, Volker Zerbe

Faculty of Informatics and Automation
Ilmenau Technical University
P.O.Box 10 05 65, 98684 Ilmenau
Germany

Abstract

In this paper we describe a system, which incorporates the optical flow to navigate in unknown environments. With the ability to identify and track moving objects it will be able to compute their trajectory and future positions to avoid collisions with them. The system incorporates an algorithm, which eliminates the optical flow induced by the motion of the camera (egomotion).

Real-world and synthetic data is used in all stages of the development of the system, simulation is used as a primary design aid.

Keywords: digital signal processing, optical flow, tracking

1 Introduction

Our goal is to design and implement an integrated navigation and tracking module which incorporates optical flow as well as acceleration sensors and Kalman filter algorithms.

The system will be simulated at mission level, which includes functional model, architectural model and environmental model. The functional model contains the estimation of the optical flow, while the architectural model contains description of actual hardware and corresponding properties. The environmental model creates a test scenario, where the whole system can be simulated in (e.g. a car on a street with buildings and pedestrians).

One of the main components of the system is the optical flow computation which will be described in this paper. It is responsible for the estimation of the flow and the correction of the flow induced by the motion of the camera.

2 Optical Flow

The optical flow is the projection of the environment of a moving camera onto the image plane of the camera. It is used in various applications from recovering the motion of the camera (egomotion) [5, 8] to reconstruction of the surrounding environment as in [9]. For a detailed description of computation methods and comparisons see [1, 2, 4, 6].

2.1 Optical Flow in Biology

Optical flow can be found in biology [3], too: two elements of the facet eye of a fly create a motion detector which measures the movement of the world seen through the facets.

Flies use this visual technique to stabilize their flight: if they turn to the left (by wind) the seen image moves to right. If the fly follows the movement of the image she is able to compensate the disturbance.

2.2 Egomotion and Optical Flow

Camera motion causes three primary forms of optical flow which can occur in any combination:

- Translation of the camera in the image plane or by rotation (small angles) about an axis perpendicular to the optical axis results in parallel flow vectors of equal length.
- Rotation is caused by rotating the camera about the optical axis.

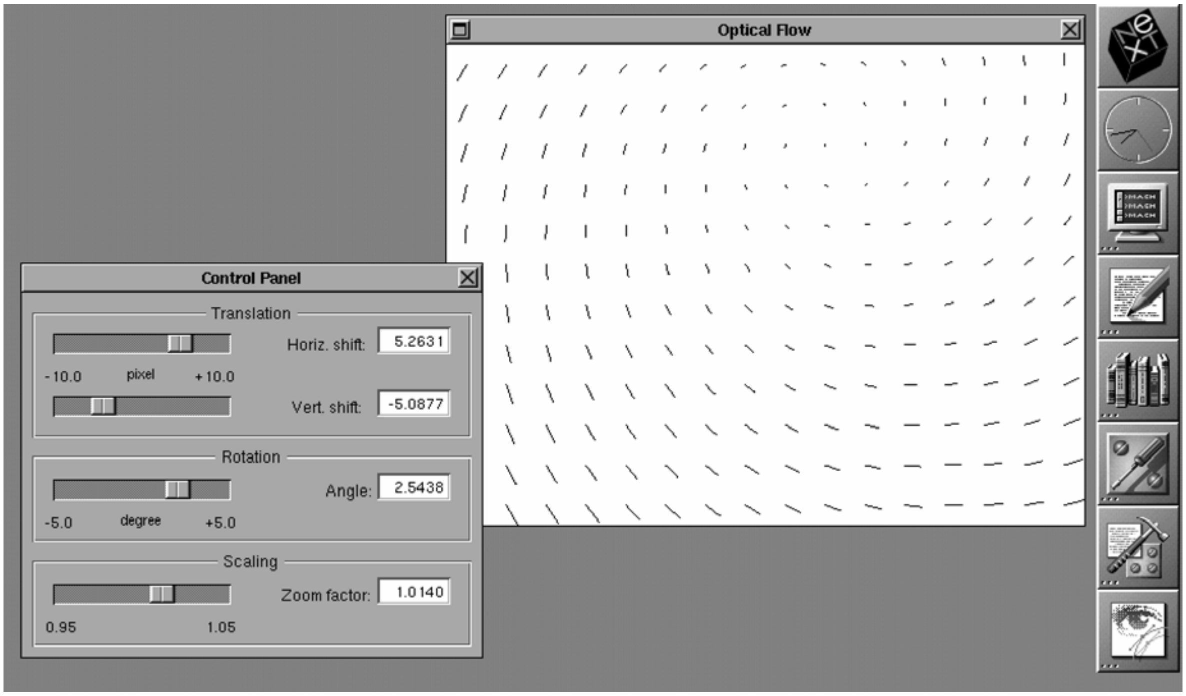

Figure 1: Optical flow visualization tool

- A movement of the camera along the optical axis results in scaling.

An example of the optical flow is shown in Figure 1 together with our visualization tool. The flow shows the result of shifting the image by 5.3 pixel to the right and 5.1 pixel to the bottom, a rotation about 2.54 degree counter-clockwise and scaling by a factor of 1.014. For comparison, the images depicted in figure 4 at the end of this article differ by these movements.

3 Correction of Self-caused Flow

As stated in the introduction we want to eliminate that part of the flow which is caused by the movement of the camera. The correction algorithm starts with the computation of the optical flow in 16 x 12 = 192 evenly distributed areas with the method used in [4] and [9].

Because the computations in each flow region are independent from other regions this is a good position to divide the problem in smaller and independent problems and to do it in parallel.

Equations 1 and 2 show the method of least squares applied to the sum of the three possible motions. A is the set of flow areas, the effect of translation is described by t, rotation by r and scaling by s, while m denotes the measured values. The indices x and y stand for the horizontal and vertical direction, respectively.

$$\sum_{a \in A} (t_x + r_x + s_x - m_x)^2 + \sum_{a \in A} (t_y + r_y + s_y - m_y)^2 = 0 \qquad (1)$$

$$\sum_{a \in A} (t_x + \alpha y + zx - m_x)^2 + \sum_{a \in A} (t_y + \alpha x + zy - m_y)^2 = 0 \qquad (2)$$

Derivation of equation 2 shows that it is possible to compute translation, rotation and scaling independently of the others. Based on the flow in the 192 areas, the horizontal and vertical translation, the rotation about the optical axis and the scaling are estimated using equation 3.

To simplify the computation of the rotation and to allow the simple application of the least squares method, we use a Taylor expansion of the $\sin(x)$ and $\cos(x)$ function. Therefore $\sin(x)$ simplifies x to while $\cos(x)$ simplifies to 1.

$$
\begin{aligned}
t_x &= \frac{\sum_{a \in A} m_x}{|A|} \\
t_y &= \frac{\sum_{a \in A} m_y}{|A|} \\
z &= \frac{\sum_{a \in A} x m_x + \sum_{a \in A} y m_y}{\sum_{a \in A} \left(x^2 + y^2\right)} \\
\alpha &= \frac{\sum_{a \in A} x m_y + \sum_{a \in A} y m_x}{\sum_{a \in A} \left(x^2 + y^2\right)}
\end{aligned}
\tag{3}
$$

Based on the global parameters horizontal and vertical translation, rotation about the center of the image and scaling, it is possible to compute a correction term for each flow area. By this means the flow caused by the egomotion of the camera will be eliminated.

The remaining flow is caused by dynamic image contents like moving or changing persons and objects, or uncertainties and errors of the computation of the optical flow. A threshold will be applied to separate significant and insignificant flow vectors. If significant flow vectors were encountered, they have to be grouped together to track objects instead of flow vectors. One important thing here is the merging and separation of objects (e.g. a person gets in a car or separation of a person from a group).

Figure 2 shows a frame of a test sequence while the corresponding optical flow is shown in Figure 3 in form of a "needle" plot. One can see the large magnitudes of the flow vectors in the bottom right of the image, as the camera "overtakes" the car. Please note that the printed image has been enhanced in brightness and contrast.

3.1 Limits

Computation of the optical flow is restricted by the quality of the image because numerical derivations have to be computed. So the image has to be "smooth" which can be achieved by application of low-pass filters like the 3 x 3 separable filter with coefficients [0.25 0.5 0.25] we used, or lossy wavelet compression algorithms [7].

The correction algorithm which eliminates the self-caused flow has limits caused by the size of the flow region and the height of the image pyramid. All the limits stated below are valid for an image size of 512 x 256 pixel and a pyramid height of 5 with 128 flow regions of size 12 x 12.

Figure 2: Sample frame of the test sequence

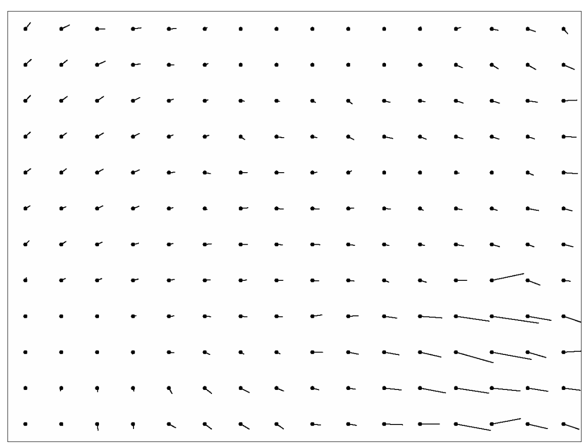

Figure 3: Needle plot of optical flow

The translational movement can be estimated accurately at speeds up to 48 pixel/frame in any direction. Estimation of the rotational movement is limited to 5 degree/frame, here the rotation of the single flow region prevents further rotation range enhancements. For the estimation of zoom a similar reason prevents the accurate computation of larger zoom speeds: the size of flow regions change. We found a limit of about 5 percent/frame.

3.2 Visualization

We use visualization of the flow in two cases: the flow estimated from a sequence of images is visualized and the effect of translation, rotation and scaling can be visualized in real-time with our tool (see figure 1). Both are used to verify numerical results.

If the computed flow vectors are changed by a random error, it is even possible to test the algorithms for recognition and tracking of moving objects without using real-world images. Because of the usage of the NEXTSTEP tools InterfaceBuilder and ProjectBuilder and the Objective-C programming language, it is possible to concentrate on the primary algorithms rather than on the graphical user interface.

4 Future Work

4.1 Accelerometers

Another way to compute egomotion is to filter and integrate measured acceleration of the camera/system by an IMU (inertial measurement unit). In conjunction with the optical flow we have two possibilities to keep track of the new information:

- a more accurate estimation of the egomotion, or
- the possibility to map the surrounding terrain based on unreliable position data in opposition to [9] who uses high precision position data

4.2 Hardware/Software Realization

For an integrated system we need a stand-alone hard- and software realization to compute the optical flow in real-time. We therefore compare different hardware approaches like digital signal processors and SIMD enhanced microprocessors (PowerPC G4, Pentium III/4, AMD Athlon).

A DSP implementation of the computationally expensive optical flow computation has been finished. For the usage of SIMD enhanced MPUs an important criterion is the support of high level language extensions like vector data types. To our knowledge this feature is available only for the PowerPC G4.

5 Conclusion

In this paper we presented the optical flow estimation part of an integrated navigation and tracking module. We use the optical flow to track dynamic objects, which cause optical flow while they move or change. An important property is the ability to remove the optical flow induced by the motion of the camera.

Two primary application fields of our system are optical tracking and counting of moving objects incorporating a moving camera and collision warning systems for cars.

Bibliography

[1] J.L. Barron, D.J. Fleet, and S.S. Beauchemin, “Performance of optical flow techniques” International Journal on Computer Vision, 12, pp. 43-77, 1994.

[2] S.S. Beauchemin and J.L. Barron, “The computation of optical flow” Computing Surveys, 27, pp. 433-467, 1995.

[3] H. Cruse, J. Dean, and H. Ritter, “Die Entdeckung der Intelligenz oder Können Ameisen denken?” C.H. Beck, München, 1998.

[4] A. Dev, “Visual navigation on optical flow” PhD thesis, University of Amsterdam, 1998.

[5] A.D. Jepson and D.J. Heeger, “Spatial Vision in Humans and Robots”, chapter “Linear subspace Methods for Recovering Translational Direction”, pp. 39-62. Cambridge University Press, New York, 1993.

[6] B.D. Lucas and T. Kanade, “An iterative image registration technique with an application to stereo vision” Proc. 7th International Joint Conference on Artificial Intelligence, pp. 674-679, Vancouver, British Columbia, August 1981.

[7] T. Radtke and V. Zerbe, “Image compression using a modified SPIHT algorithm and linear prediction” Advances in Intelligent Computation

and Multimedia Systems, pp. 49-52, Baden-Baden, Germany, August 1999.

[8] T.Y. Tian, C. Tomasi, and D.J. Heeger, "Comparison of approaches to egomotion computation" Proc. Conference on Computer Vision and Pattern Recognition, pp. 315-320, 1996.

[9] J.M. Walton, "Terrain mapping from a mobile platform through optical flow techniques" Master's thesis, Massachusetts Institute of Technology, May 1995.

Torsten Radtke received the Dipl.-Inf. degree in Computer Science from the Ilmenau Technical University, Germany, in 1999. Since then he has been a Research Assistant at Ilmenau Technical University and is studying towards his Dr.-Ing. degree. His research interests include image and video compression, optical flow and description languages.

Volker Zerbe received the Dipl.-Ing. in Electronics from the Technical University Sofia, Bulgaria in 1986. He attended the Ilmenau Technical University, Germany, receiving the Dr.-Ing. degree in 1991 for his dissertation entitled "Spezifikation und HLL-Implementierung von Multitaskbetriebssystemen". Dr. Zerbe is currently employed as a Senior Scientist. His research interests include Mission Level Design Automation and Design of Parallel Systems.

Figure 4: 2 images which differ by the movements shown in figure 1