

**Slovak University of Technology in Bratislava
Institute of Information Engineering, Automation, and Mathematics**

PROCEEDINGS

of the 18th International Conference on Process Control

Hotel Titris, Tatranská Lomnica, Slovakia, June 14 – 17, 2011

ISBN 978-80-227-3517-9

<http://www.kirp.chtf.stuba.sk/pc11>

Editors: M. Fikar and M. Kvasnica

Seibold, P.: Comparison of Two Methods for Determining the Optical Flow, Editors: Fikar, M., Kvasnica, M., In *Proceedings of the 18th International Conference on Process Control*, Tatranská Lomnica, Slovakia, 490–494, 2011.

Full paper online: <http://www.kirp.chtf.stuba.sk/pc11/data/abstracts/091.html>

Comparison of two methods for determining the optical flow

Dipl.-Ing. P. Seibold*

* FernUniversität Hagen, D-58084 Hagen
 (Tel: +49 (2331) 987-1100; e-mail: Peter.Seibold@FernUni-Hagen)

Abstract: An unmanned Aerial Vehicle (UAV) outfitted with autonomous control devices shall navigate to predefined positions. By means of cameras and optical flow the position, height above ground, orientation and velocity is determined. Two flow methods, differential technique by Bruce D. Lucas and Takeo Kanade and normalized cross correlation are presented and compared.

1. INTRODUCTION

An unmanned airship shall be regulated in conjunction with other methods by means of image processing in its position, orientation and speed [1][2]. In order to receive 3D information two cameras are mounted under an airship. Every 20ms simultaneously a picture is taken from the left and right camera. The picture of the left camera is compared to the right and to the preceding left picture. The first comparison delivers the height above ground. A measure of the translation and the rotation is obtained by evaluating the relation of the two sequential pictures. With the results the new position and orientation can be determined in respect to the preceding position. With the known starting point the absolute position is cumulated by dead reckoning.

The optical flow [3] [4] determines the relative position of corresponding image dots or regions of two images. The images may have been taken at different times by one camera or simultaneously by slightly shifted positions of two cameras. So far no other light weight system can measure the height above ground in a wider area than cameras. In contradiction to land based vehicles an airship has six degrees of freedom (x,y,z and roll, yaw, pitch). This aggravates the detection of the pose parameters.

2. OPTICAL FLOW TECHNIQUES

In this paper a differential technique is compared with a region-based matching[4].

2.1 SIMPLIFIED EXPLANATION OF THE DIFFERENTIAL METHOD

For the one-dimensional case (e.g., a picture line) a first approximation for the intensity I_1 around a pixel is given by the linear equation:

$$I = x \left. \frac{dI}{dx} \right|_{x=x_1} \quad (1)$$

With the intensity I_2 of a pixel of the second image the displacement Δx can be calculated by:

$$\frac{dI}{dx} \Delta x + \Delta I = 0 \quad (2)$$

more general:
$$\frac{\delta I}{\delta x} dx + \frac{\delta I}{\delta t} dt = 0 \quad (3)$$

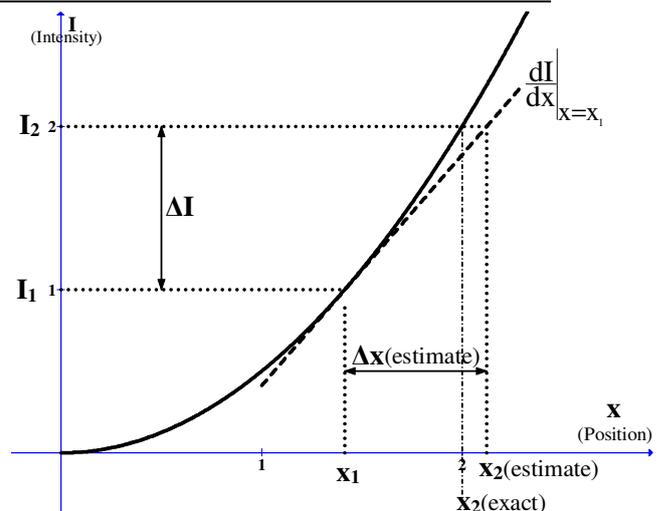


Fig. 1 Differential method for determining optical flow for the one-dimensional case

Fig.1 visualizes eq. (2). Knowing the intensity I_1 at the position x_1 an estimate x_2 for the position of I_2 can be calculated by using the derivate of $I(x)$ at x_1 . The exact position of x_2 can be approximated by iteration [5]. The first iteration step is to calculate $\left. \frac{dI}{dx} \right|_{x=x_2\text{estimate}}$ and to get the x-value of this tangent for I_2 . The remaining error is already after this step less than 1 per cent.

The differential method does not accept intensity variations for the corresponding pixels. Also displacements larger than one pixel may cause errors[4]. If larger displacements are to be detected, the images have to be consecutively reduced to smaller sizes and the results must be transposed to enlarged sizes up to the original [6].

For the two dimensional case equation (3) extends to eq. (4) with two unknowns dx and dy.

$$\frac{\partial I}{\partial x} dx + \frac{\partial I}{\partial y} dy + \frac{\partial I}{\partial t} dt = 0 \quad (4)$$

Therefore, other constraints are introduced, such as the assumption of the constancy of the optical flow around a region of the pixel at x, y as proposed in 1981 by Bruce D. Lucas and Takeo Kanade [7].

2.2 EXPLANATION OF A REGION-BASED MATCHING

In contrast to the differential method, which determines the optical flow for each pixel, the region-based matching delivers only the displacement values of a region.

One method is the cross-correlation. By convolution and its maximum, the position of a partial image is detected in the other image. Since the image energy varies in different regions and due to intensity variations the detection may fail. This is avoided by normalizing the values of the overlapping areas. In addition the average intensities are subtracted. We obtain the zero mean normalized cross-correlation (NCC) [8] [9].

$$\gamma_{u,v} = \frac{\sum_{x,y} [f(x,y) - \bar{f}_{u,v}][t(x-u, y-v) - \bar{t}]}{\sqrt{\sum_{x,y} [f(x,y) - \bar{f}_{u,v}]^2} \sqrt{\sum_{x,y} [t(x-u, y-v) - \bar{t}]^2}} \quad (5)$$

$\gamma_{u,v}$ is the correlation coefficient at the location u, v .

Because of the zero mean the range is $-1 \leq \gamma_{u,v} \leq 1$

$f(x,y)$ is the intensity value at x, y of the image to be searched.

$t(x,y)$ is the intensity value at x, y of the image sought (template).

\bar{f} and \bar{t} are respectively the average intensities in the vicinity of u, v . The area has the size of the template

The mean-free template τ is:

$$\tau(x,y) = t(x,y) - \bar{t} \quad (6)$$

The mean values of \bar{t} and $\bar{f}_{u,v}$ are:

$$\bar{t} = \frac{1}{M_x M_y} \sum_{x=1}^{M_x} \sum_{y=1}^{M_y} t(x,y) \quad (7)$$

$$\bar{f}(u,v) = \frac{1}{M_x M_y} \sum_{x=u}^{u+M_x-1} \sum_{y=v}^{v+M_y-1} f(x,y) \quad (8)$$

$M_x M_y$ is the size of the template.

The mean value \bar{t} has to be calculated only once since it is independent of u,v .

Because of the zero mean of τ :

$$\sum_{x,y} \tau(x-u, y-v) = 0 \quad (9)$$

the numerator N of $\gamma_{u,v}$, eq. (5), can be transformed to this:

$$N = \sum_{x,y} [f(x,y)\tau(x-u, y-v)] - \bar{f}_{u,v} \sum_{x,y} \tau(x-u, y-v) \quad (10)$$

$$= \sum_{x,y} f(x,y)\tau(x-u, y-v)$$

Eq. (10) is a convolution of the image with the reversed template $\tau(-x, -y)$ and can be computed by $F^{-1}\{F(f)F^*(\tau)\}$.

F^{-1} is the inverse Fourier transformation. F^* is the complex conjugate.

Each not normalized $\gamma'_{u,v}$, which is the numerator N from eq.(10), has to be divided by the denominator in order to get

the normalized coefficients. The equation for the denominator shall be simplified in order to reduce computation effort.

The first term of the denominator is:

$$\sum_{x,y} [f(x,y) - \bar{f}_{u,v}]^2 = \quad (11)$$

$$\sum_{x,y} [f(x,y)]^2 - \sum_{x,y} 2[f(x,y)\bar{f}_{u,v}] + \sum_{x,y} [\bar{f}_{u,v}]^2$$

$\sum_{x,y} [f(x,y)]^2$ is easily computed with sum tables [11], see example below.

With eq. (8) the second summation is:

$$\sum_{x,y} 2[f(x,y)\bar{f}_{u,v}] = 2\bar{f}_{u,v} \sum_{x,y} f(x,y) \quad (12)$$

$$= \frac{2}{M_x M_y} \left[\sum_{x,y} f(x,y) \right]^2$$

The second term of the denominator is calculated once:

$$\sqrt{\sum_{x,y} [t(x-u, y-v) - \bar{t}]^2} \quad (13)$$

and is valid for all u,v .

This results into:

$$\gamma_{u,v} = \frac{\sum_{x,y} f(x,y)\tau(x-u, y-v)}{\sqrt{\sum_{x,y} [f(x,y)]^2 - \frac{1}{M_x M_y} \left[\sum_{x,y} f(x,y) \right]^2}} \cdot \frac{1}{\sqrt{\sum_{x,y} [t(x-u, y-v) - \bar{t}]^2}} \quad (14)$$

The sums $\sum_{x,y} [f(x,y)]^2$ and $\sum_{x,y} f(x,y)$ are computed with sum tables [11] and they need only an addition of three values for each u,v . The sum tables must be calculated only once.

Example for a sum table:

The image pixel shall have intensity values as shown in table 1. This results to the table 2 as the sum table. Its origin is at the lower left corner.

1	2	3	4
5	6	7	8
9	1	2	3
4	5	6	7

Table 1: Image intensity values

19	33	51	73
18	30	45	63
13	19	27	37
4	9	15	22

Table 2: Sum table values

To get the sum of the center square of the image, table 1, (6+7+1+2=16), take the value of the upper right edge of this square in the sum table and subtract the sums of adjacent areas on the left and bottom side. The 4 is subtracted twice, therefore it has to be added to the subtractions. The result is:

$$45 - 18 - 15 + 4 = 16 \quad (15)$$

3. COMPARISON OF BOTH METHODS WITH 512X512 PIXEL IMAGES

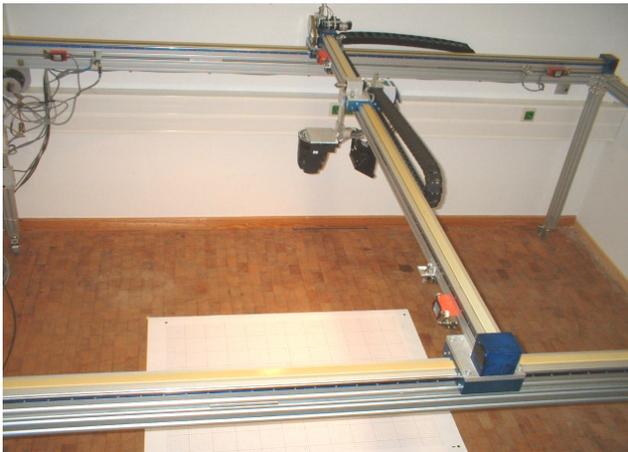


Fig. 2 XY-table with mounted camera

A camera is moved by a XY-table over a scene. The exact position of the camera is given by XY-coordinates which can be set with a precision of better than 0.1mm.

The following tables 3 and 4 show the translation in x and y direction and the angle of the relative rotation of consecutive images as calculated with B. Lucas and T. Kanade method and normalized cross correlation (NCC) in respect to the expected values (Exp.) as determined by the XY-table.

3.1 TRANSLATION

Experiments with real scenes reflect that the differential method of Lucas and Kanade (L&K) detects shifts of up to 30 pixels and rotations of 15°. NCC can identify any shifts as long as the template is inside the image. However, rotations over 5° are detected increasingly uncertain.

	Δx			Δy			$\Delta \phi$		
	Exp.	L&K	NCC	Exp.	L&K	NCC	Exp.	L&K	NCC
	0,0	-0,1	0,0	33,8	32,9	34,0	0,0	-1,1	0,0
	0,0	-0,2	0,0	33,8	31,5	35,0	0,0	-9,1	0,0
	0,0	0,3	0,0	33,8	32,4	33,0	0,0	-2,9	0,0
	0,0	0,2	0,0	33,8	35,2	35,0	0,0	-0,6	0,0
	0,0	0,0	0,0	33,8	33,2	33,0	0,0	0,0	0,0
	0,0	-0,1	0,0	33,8	34,5	34,0	0,0	0,6	0,0
	0,0	0,1	0,0	33,8	33,6	33,0	0,0	0,0	0,0
	0,0	0,2	0,0	33,8	34,1	34,0	0,0	0,6	0,0
	0,0	0,3	0,0	33,8	33,1	34,0	0,0	1,1	0,0
Cum. Values:	0,0	0,7	0,0	304,2	300,6	304,9	0,0	-11,4	0,0

Table 3: Translation in y-direction

Legend:

- $\Delta x, \Delta y$: displacement in pixel of a region around the center of two consecutive images.
- $\Delta \phi$: rotation angle of two consecutive images.
- Exp.: expected value as set by the XY-table.
- L&K: differential method of B. Lucas and T. Kanade.
- NCC: Normalized Cross Correlation.
- Cum Values: cumulated values.

Due to the application as pose detection of an airship, in contrast to the often-quoted measure of error (angular error) [10], the deviation of the cumulative value is considered. For translations ($\Delta x, \Delta y$) both methods deliver good values, NCC somewhat better. At the orientation $\Delta \phi$, which should be zero, the differential method deviates greatly at one pair of images from the expected value.

3.2 ROTATION

The camera was continuously rotated by 10 degrees around the central axis. As table 4 shows, the values of the differential method are closer to the expected values and is therefore superior to the correlation method.

The advantage of the NCC, is that a failure can be detected with the relationship between a poor value and a low correlation maximum.

Exp.	$\Delta \phi$	
	L&K	NCC
10,0	7,2	6,8
10,0	8,9	6,8
10,0	9,9	11,3
10,0	10,0	11,3
10,0	10,1	9,1
10,0	9,9	8,5
10,0	9,8	8,0
10,0	9,7	9,1
10,0	9,8	1,1
Cum. Values:	90,0	72,0

Cum. Values:

Table 4: Rotation by 10 degrees

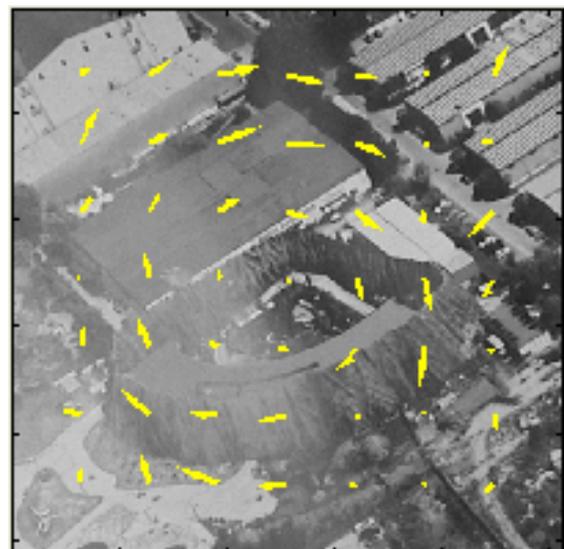


Fig. 3: Flow vectors of a rotated scene. Arrows represent the optical flow of two consecutive images

3.3 BRIGHTNESS CHANGE

Due to brightness change of the scenes the shutter of a camera may alter the exposure. To simulate this effect every 2nd image of a real flight over a landscape has been darkened by a multiplication with 0.95.

As expected, the differential method (Fig. 4) has failed. The composed image calculated using the optical flow is chaotic. The flow vectors as seen in Fig. 5 reflect this irregularity. The composite image by using NCC, Fig. 6, is satisfactory. Without the brightness attenuation the differential method provides the same good result as NCC.



Fig. 4: Composite image, differential method with brightness change

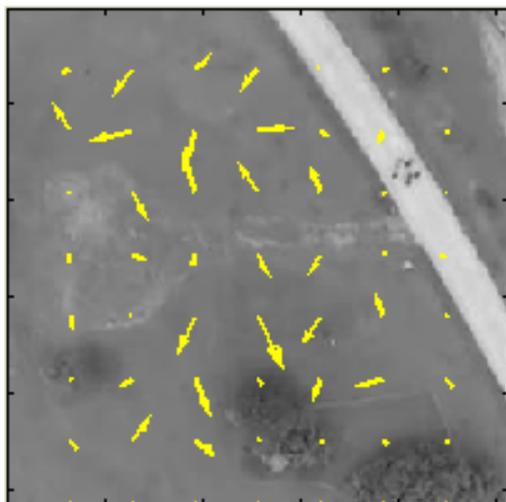


Fig. 5: Chaotic flow vectors of one pair of images with the differential method. Brightness is changed.



Fig. 6: NCC
 Composite image with brightness change

3.4 COMPUTING EFFORT

Relevant to the applicability in real time is the calculation time. As a first estimate the calculation time on a PC is used, even when operating in a DSP or FPGA, memory access times are more important than multiplications and additions. A PC reduces memory access times by a large "pre-cache". The differential method requires on a PC 460ms for the calculation of an image pair, while NCC with 170ms is almost three times faster. In order to control the airship autonomously the target process time for two consecutive images shall be less than 20 ms. This task will be accomplished by using DSP for fast sequential arithmetic operation and FPGA for parallel processing.

The amount of multiplications in respect to the processing time are not any more that important since DSPs accomplishes them in one cycle. Other operations like divisions, square roots take many more cycles and should be avoided.

Memory usage in respect to image size M_x (columns), M_y (rows) and a required precision P_i :

The precisions may be set as used for the calculations of tables 3 and 4 to:

- P_1 : one byte
- P_2 : 2 bytes (half precision)
- P_4 : 4 bytes (single precision)

Differential method:

For first, 2nd image and pyramid: $2.6 M_x M_y P_1$

Flow vectors x and y: $2 M_x M_y P_2$

Sum for 512x512 image: 1.2MByte

NCC:

For first and 2nd image: $2 M_x M_y P_1$

FFT (complex) for both images and inverse FFT (real): $5 M_x M_y P_4$

Sum tables: $2 M_x M_y P_4$

Sum for 512x512 image: 7.9MByte

The real amount in bytes is depending on the required accuracy. Even when some memory may be reused, NCC needs by far more memory resources.

4. SUMMARY

NCC is superior to the differential method in the detection reliability, accuracy and speed. Also the allowed displacement range is much wider. While NCC needs only the corresponding region to be within both consecutive images, the differential method starts to fail at more than 30 pixel distance. However the differential method is superior at rotations and for the density of the flow vectors. Since for the application in an airship the maximal rotation angle of two consecutive images is expected to be less than 5 degrees and a density of the optical flow down to one pixel is not necessary, NCC is the preferred method.

Further test flights will be done in order to check the maximum dynamics of movements and the requirement for the optical flow. Also ideas to detect roll and pitch will be developed.

REFERENCES

- [1] Chris Mccarthy and Nick Barnes, Performance of Optical Flow Techniques for Indoor Navigation with a Mobile Robot, Proceedings of the IEEE International Conference on Robotics & Automation, 2004
- [2] Franck Ruffier and Nicolas Franceschini, Optic flow regulation: the key to aircraft automatic guidance, Robotics and Autonomous Systems, 2005
- [3] B. K. P. Horn and B. G. Schunck, Determining optical flow, Artificial Intelligence, 1981
- [4] J. L. Barron et al., Performance of optical flow techniques, International Journal of Computer Vision, 1994
- [5] Andrés et al., Lucas/Kanade meets Horn/Schunck: Combining local and global optic flow methods, 2005
- [6] Peter J. Burt and Edward H. Adelson, The Laplacian Pyramid as a Compact Image Code, IEEE Transactions on Communications, 1983
- [7] Bruce D. Lucas and Takeo Kanade, An Iterative Image Registration Technique with an Application to Stereo Vision, Proceedings of the 7th International Joint Conference on Artificial Intelligence, 1981
- [8] J. P. Lewis, Fast Normalized Cross-Correlation, Industrial Light & Magic, 1995
- [9] A. J. H. Hiiet et al., Fast normalized cross correlation for motion tracking using basis functions, Comput. Methods Prog. Biomed., 2006
- [10] Fleet, D.J. and Jepson, A.D., Computation of Component Image Velocity from Local Phase Information, IJCV, 1990
- [11] Crow, Franklin C., Summed-area tables for texture mapping, SIGGRAPH Comput. Graph., 1984