ESTIMATION OF OPTICAL FLOW FOR LARGE DISPLACEMENTS

TORSTEN RADTKE
Faculty of Informatics and Automation
Technische Universität Ilmenau
P.O.Box 10 05 65
98684 Ilmenau, Germany
email: torsten.radtke@theoinf.tu-ilmenau.de

HORST SALZWEDEL
Faculty of Informatics and Automation
Technische Universität Ilmenau
P.O.Box 10 05 65
98684 Ilmenau, Germany
email: horst.salzwedel@theoinf.tu-ilmenau.de

ABSTRACT
In this paper we present a new method to estimate optical flow for large displacements. It is based on prediction of global flow field parameters, performs better than multi-resolution estimation methods and has been verified using standard test sequences as well as real-world data. Global flow field parameters can be estimated from optical flow measurements in all flow regions. They can then be used to predict the flow in optical flow regions of the next frame. This technique reduces the complexity in comparison to hierarchy-based methods, while the flow field parameters can also be used to compensate optical flow produced by egomotion.

KEY WORDS
Motion Detection and Estimation, Multi-Resolution and Multi-Spectral Processing, Optical Flow, Vision

1. Introduction

In recent research much attention has been paid to more precise estimation techniques of optical flow, which are computationally expensive. Most of these techniques use more than 2 frames to estimate the flow and cannot achieve real-time performance (see [1] for a possible solution).

Another disadvantage of the published methods is that they only work well for small optical flow. But computational efficiency and the property to estimate large displacements accurately are important preconditions for optical flow techniques in navigation.

2. Review: Optical Flow

Optical flow is based on the constant-brightness assumption. In equation 1, E is the brightness of the image, E_x, E_y, and E_t are derivatives in horizontal, vertical and temporal direction, respectively.

$$E_x u + E_y v + E_t = 0 \quad (1)$$

Optical flow can not be computed locally, since flow velocity has horizontal and vertical components u and v, but only one independent and measurable variable is available at each point of an image.

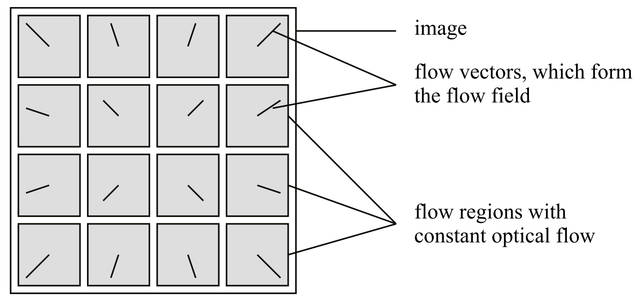

Figure 1. Relationship between image, flow regions, vectors and field

To solve the problem, an additional constraint condition is needed. Horn [2] uses the spatial smoothness of the optical flow as a constraint. However, there is no plausible model to justify the assumption that minimal variation of the optical flow over the image approximates the motion field. We parametrize the optical flow and use the assumption that the flow is constant in small image regions. So one can use multiple pixels to estimate the two unknown velocity components with the method of least squares. In equation 2, u and v denote the horizontal and vertical component of the flow, while E_x, E_y and E_t denote spatial (horizontal and vertical) and temporal derivatives.

$$\begin{aligned} u &= \left(\frac{\sum E_x E_y \sum E_y E_t - \sum E_x E_t \sum E_y^2}{\sum E_x^2 \sum E_y^2 - \sum E_x E_y \sum E_x E_y} \right) \\ v &= \left(\frac{\sum E_x E_y \sum E_x E_t - \sum E_x^2 \sum E_y E_t}{\sum E_x^2 \sum E_y^2 - \sum E_x E_y \sum E_x E_y} \right) \end{aligned} \quad (2)$$

For a better understanding of the dependencies between image, flow vectors, flow regions and flow field, see figure 1. For a detailed review of optical flow estimation, we refer to [3, 4, 5].

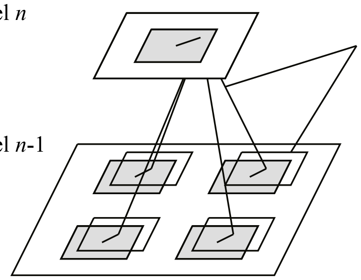

Figure 2. Multi-resolution estimation

3. Treatment of Large Displacements

There are several possibilities to estimate large displacements of single flow regions with sufficient accuracy:

- increased size of flow regions
- computation of flow at multiple resolutions
- prediction of global flow field parameters

3.1 Increased size of flow regions

For this paper we compute optical flow by equation 2 (for further details see [6, 7]). By increased size of the flow region larger displacements can be estimated. For regions of size 12×12 pixels estimates are accurate for displacements of up to 2 pixels in each direction.

Unfortunately, the regions can not be enlarged too much: the assumption that optical flow is constant in small areas does not hold for large regions — averaging effects occur.

3.2 Multi-resolution estimation

The optical flow can be estimated at different levels of an image hierarchy, starting at the top level (coarsest resolution). For each level of the hierarchy, image rows and columns are downsampled by a factor of 2 and the measurable displacements double. Flow vectors at a given level n are projected to the level $n-1$ to reduce the remaining flow there (see figure 2).

This technique has also limits: the number of levels can not be increased infinitely. The top level of the hierarchy must be sufficiently large for at least one flow region.

Figure 3 (left) shows the results of flow estimation over the first 20 frames of the "translating tree" sequence. Each frame is shifted about 2 pixels to the left in comparison to the previous frame. Estimation is done between the first frame and all other frames, so the flow between frame 1 and 11 is 20 pixels to the left. One can see that a 3-level hierarchy (each new level has been downsampled and filtered by a 3×3 separable low-pass filter) with flow regions of size 12×12 can handle flows up to 12 pixels/frame, higher flows lead to unusable results.

3.3 Prediction of global flow field parameters

For a given flow field it is possible to estimate parameters describing the movement of the camera (egomotion). Rotation about an axis perpendicular to the optical axis (yaw, pitch) causes a translation of the projected image in horizontal and vertical direction t_x and t_y, while rotation about the optical axis (roll) causes a rotation of the projected image about the image center (α). If the camera moves forward or backward along the optical axis, a zoom effect z applies to the image. These parameters are computed as shown in equation 3.

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

In the equation above, $m_x(x,y)$ and $m_y(x,y)$ denote the computed flow vector components in a flow region at position (x,y), sums include all flow regions of the image.

Given a set of 6 to 8 previously computed parameter quadruples, it is possible to estimate a linear model for each parameter by the method of least squares. Based on this model we can extrapolate the 4 parameters for the next frame and tune the flow regions accordingly. The flow regions will be pre-adjusted to the predicted flow, hence the remaining flow will be small enough (less than 3 pixels/frame) to be handled by a flow region.

Figure 3 (right) shows the results of flow estimation over the same series of 20 frames as above. Here the flow can be estimated accurately at up to 40 pixels/frame (this result can not be achieved with a more than 3-level multiresolution estimation because another level is not large enough to contain a flow region).

Figure 4 shows frames 1, 10, and 25 of the "diverging tree" sequence (top row) and corresponding optical flow in form of needle diagrams (bottom row, middle and right). The lower left graph shows the performance of the global parameter based method on the whole sequence.

4. Comparison

A comparison of the three methods gives the following results:

1. increased size of flow regions
   - – problems at image borders (large flow regions result in no flow vectors at the borders)
   - – the flow is not constant in large flow regions, this leads to estimation of mean values for the flow
   - + no special adaption of the algorithm is necessary

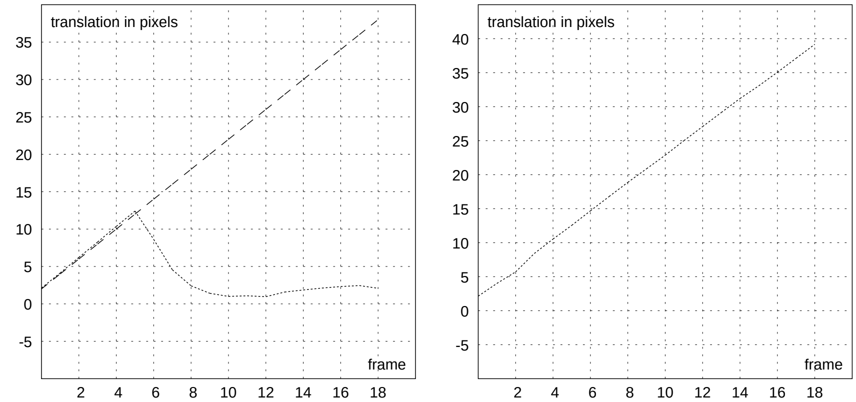

Figure 3. Comparison between multi-resolution and global prediction based estimation (translating tree sequence)

2. computation of the flow at multiple resolutions

   – hierarchy construction and the increased number of flow fields lead to higher computational complexity

   – wrong flow estimation at a high hierarchy level affects all lower levels and leads to large errors

   + good results for flow fields which are hard to estimate with the model used in method 3 (problem example: a car moves from left to right in front of a fixed camera)

3. prediction of global flow field parameters

   – bad results for flow fields which are hard to estimate with the given model (i.e. the camera moves in a static environment)

   ± lower computational complexity than method 2 but estimation of flow field parameters is necessary

   + theoretically the largest range for displacements, it is not required that the top level of the hierarchy contains a flow region

5. Conclusion

A new method to estimate the optical flow for large displacements has been presented. The results are comparable or better than those from multi-resolution estimation methods. For image sequences with small dimensions, hierarchy-based estimation methods do not have enough levels to handle large displacements well. In such cases, the new method performs as good as on other image dimensions.

The estimation of global flow field parameters and model estimation not only allows the prediction of the next quadruple of parameters but also the possibility to compensate the optical flow produced by egomotion. This information can be used to track dynamic objects from a moving camera as described in [8].

Both methods have their benefits and drawbacks, a combination of them could overcome their limitations. So the next step is to combine multi-resolution and global parameter based estimation of optical flow.

References

[1] T. Camus, *Real-Time Optical Flow*, (PhD thesis, Brown University, Providence, 1994).

[2] B.K.P. Horn and B.G. Schunck, Determining optical flow, A.I. Memo No. 572, Massachusetts Institute of Technology, 1980.

[3] 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*, Vancouver, Canada, 1981, 674-679.

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

[5] S.S. Beauchemin and J.L. Barron, The computation of optical flow, *Computing Surveys*, 27, 1995, 433-467.

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

[7] A. Dev, *Visual Navigation on Optical Flow*, PhD diss., University of Amsterdam, Amsterdam, 1998.

[8] T. Radtke and V. Zerbe, Tracking of dynamic objects based on optical flow, in M.R. Syed and O.R. Baiocchi (Ed.) *Intelligent Multimedia, Computing and Communication – Technologies and Applications of the Future*, (New York: John Wiley & Sons, 2001) 20-24.

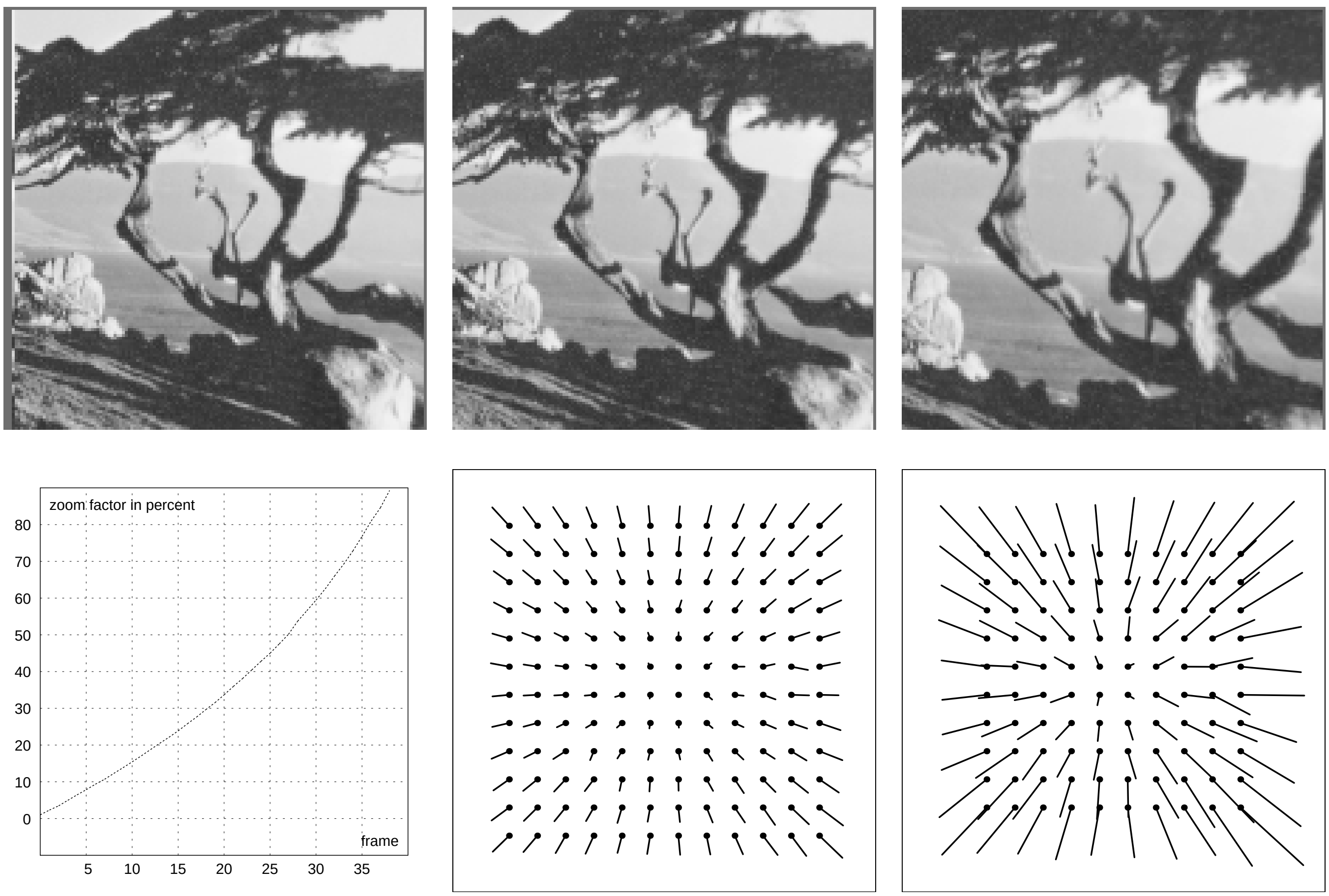

Figure 4. Examples for optical flow in the diverging tree sequence