Statistical Approach to Classification of Flow Patterns for Motion Detection

Joachim Denzler[a] & *Volker Schleß*[b,a] & *Dietrich Paulus*[a,1] & *Heinrich Niemann*[a]

Universität Erlangen–Nürnberg[a]
Lehrstuhl für Mustererkennung (Informatik 5)
Martensstr. 3, D–91058 Erlangen, Germany
email: denzler@informatik.uni-erlangen.de

Daimler–Benz Forschungszentrum Ulm[b]
F3MS, Wilhelm Runge Str. 11, D–89081 Ulm, Germany
email: schless@dbag.ulm.DaimlerBenz.com

ABSTRACT

We present a new approach for egomotion computation and the detection of independent motion in the scene. In contrast to related work we apply statistical methods which are based on the normal optical flow field. We extract features for supervised and unsupervised training from the normal optical flow field in order to train a Gaussian–distribution classifier (GDC) and a Kohonen feature map. Finally, in a test phase the egomotion computation is done by classifying features extracted from the normal optical flow field into the unknown motion direction. For the detection of independent motion, the scene is divided into regions. For each region a decision is made, whether the normal flow in this region is based on the camera motion or an independently moving object. We present results of this approach which show a recognition rate of up to 97% for the egomotion classification and a detection rate of moving objects of up to 87%.

1. MOTIVATION

Applications of state of the art image analysis can be found in autonomous ground vehicles, robotics, industrial production, etc. Since the camera is itself a moving part in many of these systems, estimation of the viewer's motion is as important as the detection of independently moving objects in the scene. [1, 2, 8]

Similar problems have to be solved in active vision systems where — in addition to various changes of the camera parameters — the camera is moved purposively in order to solve the vision tasks more efficiently.

[1] This work was funded partially by the German Research Foundation (DFG) under grant number SFB 182. Only the authors are responsible for the contents.

We have to know the egomotion in order to detect independently moving objects in the scene, and conversely, we also have to know about moving objects in order to estimate the egomotion. A solution to this chicken–egg problem is to compute global features from the motion field [4].

In this contribution we present a new approach, which extends global feature detection by statistical classification in order to detect independently moving objects and to estimate camera motion. Additionally, a Kohonen feature map [6] is trained which allows for an unsupervised clustering of the features into different motion classes.

2. THEORETICAL BACKGROUND

In [4] an approach for the calculation of egomotion has been proposed. It could be shown, that for every egomotion, described by the rotation (α, β, γ) and translation (u, v, w), there exists a certain motion pattern (cf. Fig. 1, left) in the image plane calculated from the normal flow, i.e. the optical flow projected on the image gradient between two images. The unknown motion parameter of rotation (α, β, γ) and the focus of expansion $(x_0, y_0) = (\frac{uf}{w}, \frac{vf}{w})$ are estimated by a search in the motion parameter space $(\alpha, \beta, \gamma, u, v, w)$ to match the significant motion patterns with the observed motion field in the image. More details of this global feature extraction scheme can be found in [4, 5].

This method of motion analysis is based on an computationally expensive exhaustive search for the motion parameters by separating the translation from the rotation, and fitting the so called *coaxis patterns*; coaxis vectors are defined as normal flow vectors perpendicular to conic sections with the image plane; the conic sections are defined by the egomotion of the observer.

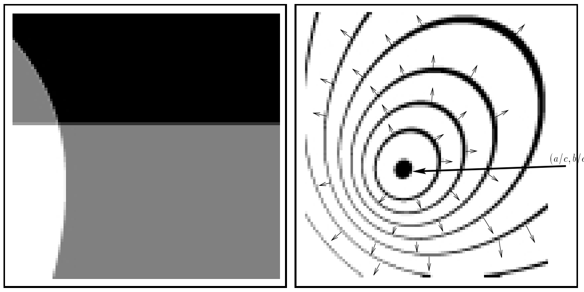

Fig. 1: Left: Motion pattern. Right: example of positive coaxis vectors (from [4])

The comparison of the direction of the normal flow vectors and the vectors perpendicular on the conic sections defines the *sign* of a coaxis vector which will be used in the following. Details on these definitions can be found in [4]. Examples are shown in 1 and Fig. 2.

In case of wide spread vectors, i.e. a sparse normal flow field due to lack of structure in the scene, there exists no unique solution for the motion parameters given a coaxis pattern. Also, for separating the rotation from the translation, the camera needs a very large field of view.

In this contribution we apply clustering techniques (Kohonen feature maps) and statistical decisions (Gaussian–distribution classifier, GDC) in order to avoid such global search techniques.

3. CLASSIFICATION OF MOTION FIELDS

By visually inspecting the motion pattern in Fig. 1 left one can observe three homogeneous regions, one with positive and one with negative coaxis vectors and one region, in which both types of coaxis vectors may occur. This depends on the structure and depth in the scene. The idea of our approach is the following:

1. We calculate the normal flow. Since this can only be done on edges in the image, we normally get a sparse normal flow field.
2. Then, we compute the sign of the coaxis vectors (cf. Sect. 2, [5]).
3. Finally, we search for areas in the image which contain positive coaxis vectors, and for those which contain negative vectors.

The area of the positive coaxis vectors will overlap with the area of the negative coaxis vectors. The remaining two areas, one with positive and one with negative coaxis vectors only, will define the egomotion of the observer.

Now statistical methods are applied to estimate the egomotion. One can calculate various features from the areas. One example is the mean and the variance of the position of coaxis vectors. We applied this procedure to the $(1,0,0)$-, $(0,1,0)$- and $(0,0,1)$-coaxis vectors (see [5]) and finally got a 24–dimensional feature vector. This vector was concatenated with a class label for the direction of movement that caused the movement vectors. Sets of these labeled feature vectors are now used for training and testing a GDC.

Kohonen feature maps are well known for their ability to map similar features into neighboring areas in the feature map (topology preserving maps [7]). Thus, a second approach trains Kohonen feature maps to cluster unlabeled feature vectors into different areas. Each area corresponds to a certain motion direction (see Fig. 3). This is motivated by the observation, that similar movements result in similar normal flow patterns (i.e. features).

On the other hand, this approach can be used to detect independently moving objects. Knowing the egomotion of the observer, one can look for outliers in the coaxis vector fields. For this the image is divided into equal sized quadratic areas, which are searched for such outliers. A significant amount of outliers in one area indicates an independently moving object. For example, one can inspect the $(0,1,0)$–coaxis vectors if a counterclockwise rotation around the x–axis is done. Then all the $(0,1,0)$–coaxis vectors in the whole image should be positive. If some areas with negative coaxis vectors can be observed, this is a hint to an independently moving object.

Fig. 2: Examples for positive and negative (0,0,1)–coaxis vectors for a real scene.

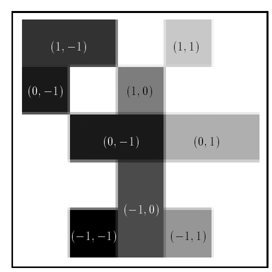

Fig. 3: Best Kohonen feature map: the areas, which correspond to certain motion directions are shown.

4. EXPERIMENTS

We made experiments both for classification of the egomotion and for the detection of independently moving objects. For the classification of the egomotion eight

class (pan,tilt)	correct GDC	correct Kohonen
(-1,-1)	98%	83%
(-1, 0)	91%	92%
(-1, 1)	96%	93%
(0,-1)	90%	63%
(0, 1)	85%	45%
(1,-1)	98%	87%
(1, 0)	93%	98%
(1, 1)	92%	83%

Tab. 1: Evaluation of correct classification according to various motion directions (motion of 2 – 6 pixels/image): results for the Gaussian–distribution classifier (GDC) and the Kohonen feature map (Kohonen).

classes for the eight possible motion directions of a pan–tilt camera unit have been chosen. This allows for a qualitative classification of the motion directions

$$\begin{aligned}(\text{pan,tilt}) \quad = \quad & \{(1,0),(0,1),(-1,0),(0,-1),\\ & (1,1),(-1,-1),(-1,1),(1,-1)\}.\end{aligned}$$

A +1 means a counterclockwise rotation, -1 a clockwise rotation of the axis.

In Tab. 1, second column, classification results for 1000 test patterns of the various motion directions are shown. For most of the classes the recognition rates are over 90% for the GDC, which has been trained with 15000 examples. Analyzing the errors, it is worth noting that all the wrong classifications are into neighboring classes, for example the patterns of class $(1,0)$ are classified as $(1,-1)$.

To verify the suitability of our classifier in new environments, we applied the egomotion classification to a different scene, which has not been used for training. Then, we get an average recognition rate of about 63%.

In Fig. 3 the best feature map (25 neurons) after 15000 training patterns is shown with the areas which correspond to the motion classes. It can be seen, that motion classes, which correspond to opposite motion directions are also located opposite in the feature map (topology preserving maps). In Fig. 4 and 5 two other feature maps (9 and 100 neurons) can be seen. In the case of nine neurons the class centers have a short distance in the feature map, which results in more misclassifications. In the case of 100 neurons, the results are similar to the best feature map with 25 neurons, but the computational effort and the size of the map grows, too. In Tab. 1, third column, the classification results with the Kohonen feature map (25 neurons) are shown. Compared to the GDC, the results are worse in general, but acceptable, except for the two motion classes $(0,1)$ and $(0,-1)$ (tilt movement of the camera). We assume, that this depends on the positional low accuracy of the tilt axis; this has to be verified in our future experiments.

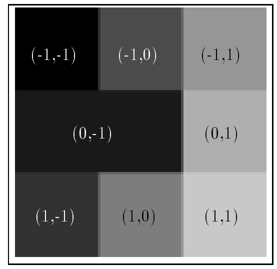

Fig. 4: Kohonen feature map for 9 neurons.

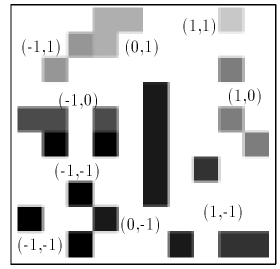

Fig. 5: Kohonen map (100 neurons).

Finally, we have performed experiments for the detection of independent motion in the scene (a moving toy train) under a known camera motion. The 512×512 sized image was divided into 100 areas of the same size. Then, for each area the extracted normal flow directions are compared with the normal flow directions, corresponding to the known egomotion. As soon as more than 66% of wrong coaxis vectors (outliers) occurred, we decided this area as an area containing independently moving objects. In Fig. 6 (left) a correct result of the detection of an independently moving object is shown, based on outlier classification. In Fig. 6 (right) one of those typical errors is illustrated, which mostly occur on strong background edges. The error rates on a larger test set are presented in Tab. 2 and Tab. 3. In Tab. 2 the percentages of images are shown, in which a moving object has been correctly detected. An object has not been detected if areas are marked which does not contain the object or the moving object has not been detected at all. In Tab. 3 we tested whether areas are marked in the case that no moving object is in the scene.

Independently moving object		
total images	correct detected	incorrect detected
547	476	71
100%	87.0%	13.0%

Tab. 2: Test of the quality of the detection of an independently moving object

Feature extraction (computation of normal flow vectors, coaxis vectors, and computation of one feature

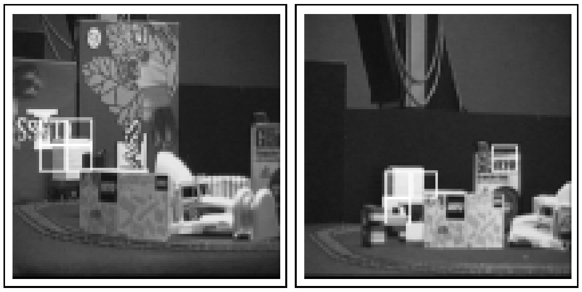

Fig. 6: Left: results for detection of independently moving objects. Four rectangular areas marking the moving toy train. Right: typical error. Additional areas are marked, which do not contain moving objects.

No independently moving object (false alarm)		
total images	correct result (no moving object)	incorrect result (moving object)
453	402	51
100%	88.7%	11.3%

Tab. 3: Test of the quality of the detection of an independently moving object

vector) takes between 1.39 and 2.42 sec on a HP 735 depending on the number of flow vectors caused by the movement. The classification time of one feature vector is about 0.01 sec. The system is implemented in C++ and integrated in an object–oriented environment [9].

Additional experiments concerning results for different resolutions of the images, different scenes, and different pan–tilt devices (Canon VCC1, TRC stereo head) as well as the error rate during the learning process can be found in [10]. Presently, the integration into the system described in [3] is being done.

5. CONCLUSION

In this contribution we presented a statistical approach to classification of normal flow patterns for motion detection. Two methods have been presented: a GDC and a Kohonen feature map. The task has been to classify unknown camera motion qualitatively into nine classes of different pan–tilt movements. For the GDC the average recognition rate is over 90%. The Kohonen feature map is worse compared to the GDC (best result with a 25 neurons feature map: 80%).

In addition we have proposed a method for the detection of independently moving objects during a known camera motion. This method is based on a outlier classification in the directions of the normal optical flow (coaxis patterns). With this approach we detected motion correctly in over 87% of the images.

The limitation of our approach for egomotion detection is, that so far only qualitative results about the motion can be obtained, whereas the approach of [4] yields qualitative data. The advantage of our approach is, that after a training stage the classification is much faster, because no exhaustive search is needed.

6. REFERENCES

[1] W. Burger and B. Bhanu. Estimating 3-d egomotion from perspective images sequences. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(11):1040–1057, 1990.

[2] A. Blake D. Sinclair and D. Murray. Robust estimation of egomotion from normal flow. *International Journal of Computer Vision*, 13(1):57–69, 1994.

[3] J. Denzler and H. Niemann. Combination of simple vision modules for robust real–time motion tracking. *European Transactions on Telecommunications*, 5(3):275–286, 1995.

[4] C. Fermueller. *Basic Visual Capabilities.* PhD thesis, University of Maryland, 1993.

[5] C. Fermueller and Y. Aloimonos. The role of fixation in visual motion analysis. *International Journal of Computer Vision*, 11(2):165–186, 1993.

[6] T. Kohonen. *Self-organization and associative memory.* Springer Verlag, London, Berlin, Heidelberg, New York, Paris, Tokyo, Hong Kong, Budapest, 1984.

[7] M. Kraaijveld, J. Mao, and A. Jain. A nonlinear projection method based an kohonen's topology preserving maps. *IEEE Transactions on Neural Networks*, 6(3):548–559, 1995.

[8] R.C. Nelson. Qualitative detection of motion by a moving observer. *International Journal of Computer Vision*, 7(1):33–46, 1991.

[9] D.W.R. Paulus and J. Hornegger. *Pattern Recognition and Image Processing in C++.* Advanced Studies in Computer Science. Vieweg, Braunschweig, 1995.

[10] Schleß, V. Klassifikation von Bewegungsmustern in Bildfolgen zur Bewegungsberechnung *"Classification of motion pattern in image sequences for motion computation"*. Diploma thesis, Lehrstuhl für Mustererkennung (Informatik 5), Universität Erlangen–Nürnberg, Erlangen, 1996.