Improvement of the Fail-Safe Characteristics in Motion Analysis Using Adaptive Technique

Ayoub Al-Hamadi, Ruediger Mecke, and Bernd Michaelis

Institute for Electronics, Signal Processing and Communications (IESK)
Otto-von-Guericke-University Magdeburg
D-39016 Magdeburg, P.O. Box 4210 Germany
ayoub@iesk.et.uni-magdeburg.de

Abstract. In this paper, we propose an adaptive technique for the automatic extraction and tracking of moving objects in video sequences that works robustly under the influence of image-specific disturbances (e.g. brightness variations, shadow and partial occlusion). For this technique, we apply the colour information, a neural recognition system and a recursive filtering algorithm to the improvement of the matching quality when disturbances occur. This suggested intensity-based technique is adaptive and robust compared to the conventional intensity-based methods.

1 Introduction

The extraction of moving objects and subsequent recognition of their trajectories (tracking) in video sequences is of increasing importance for many applications. Examples are video surveillance, motion estimation and human computer interaction. Generally, the motion or tracking methods can be divided into four groups: A) Three-dimensional-based methods [1]. B) Feature-based methods [2,3]. C) Deformable model-based methods [4] and D) Intensity-based methods [5]. Usually these conventional intensity-based methods of the motion analysis (*e.g. Blockmatching* or *optical flow*) don't operate reliably by the influence of *image-specific* disturbances such as *brightness variations*, *shadow*, *small grey tone gradients* and *partial occlusion*.

Our technique is an intensity-based method which can take advantage of characteristics found in colour scenes. The developed technique here pursues the objective of automatic segmentation of each moving object and furthermore the determination of the motion trajectory of these moving objects. For initial object selection, the motion vector field (full search Blockmatching BM) is used. Thereby moving blocks can be extracted. For the following tracking analysis the blocks with similarity motion parameters (*if they fulfill a homogeneity criterion*) will be combined to object candidates, which are transformed into adaptive colour space. This adaptive colour space is generated as a function of image content. By this adaptive colour space, we'll achieve an optimal channel reduction, separation the luminance and chrominance information on the one hand and on a high dynamic gain the other. The calculation of the colour components is obtained via the **K**arhunen **L**oéve **T**ransformation (KLT) [6,7]. This method

A. Sanfeliu and J. Ruiz-Shulcloper (Eds.): CIARP 2003, LNCS 2905, pp. 104–112, 2003.

provides the optimal subspace which minimizes the mean-square-error between the given set of vectors and their projections on the subspace. The resulting image regions present the object candidates that are used for the solution of the correspondence problem in video sequences.

Next to the successful tracking of the extracted moving regions in the video sequences a **m**ulti **m**atching **m**ethod (**M³**) is applied. This differs from the simple **B**lockmatching (BM) by the fact that in our approach object-adapted regions (*extracted from the* initial *step*) will be used for further tracking in the sequences instead of fixed blocks. Thus problems of fixed block (e.g. aperture problem) are eliminated. This is because the object-adapted regions contain the energetic features (*e.g. edge, corner etc.)* which are used for the improvement of fail safe characteristics by the motion analysis. Another advantage of our approach is the extension of the sample matching on **M³** (*due to colour components*) to obtain reliable results. This is because the colour components of the adaptive colour space are invariant by the modification of the intensity value due to overlay of *shadow* or *brightness fluctuations*. For further improvement regarding the partial *occlusion* the prehistory is evaluated by means of a *recursive filtering algorithm* [8,9]. In general, this paper consists of two main parts: the first section describes the adaptive technique. The second section presents some experimental results of this technique.

2 Adaptive Technique for Motion Analysis

The suggested technique is described by two processing levels (Fig. 1), whereby the first level deals with the segmentation of the moving objects and analysis of the colour information. The second level of this system is specified by the M³, a neural recognition system and a recursive modified algorithm for the estimation of the displacement vectors under the influence of image-specific disturbance (*partial occlusion*).

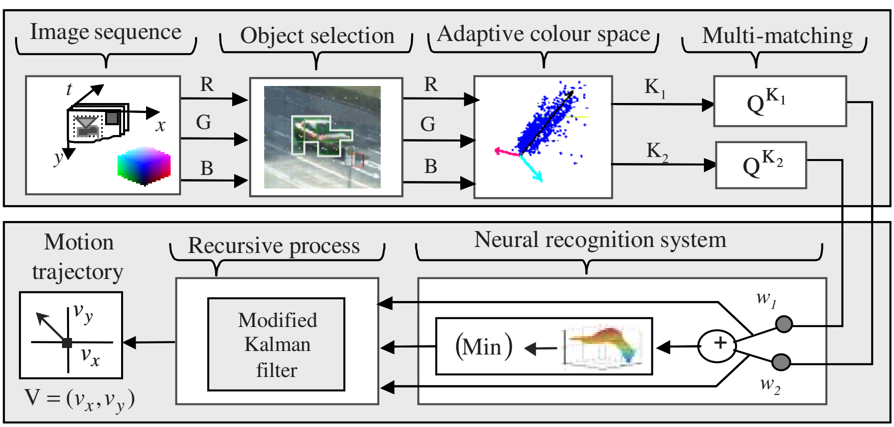

Fig. 1. The simplified suggested adaptive technique for the automatic detection and tracking analysis of objects in colour video sequences

2.1 Segmentation of Objects Using Motion Field

In the motion analysis, it is desirable to apply automatic techniques for the selection of moving objects from image sequences. In this work we use a motion vector field (*full search BM*) for automated initial object selection in the RGB image sequence. One receives a displacement vector, which describes the determined motion of the represented image region as a result of the BM for each reference block. The calculation of the displacement vectors takes place via a Q-criterion (e.g. **m**ean **a**bsolute **d**ifference MAD). Due to the small implementation expenditure the criterion of the MAD was used. More specifically, denoting I(**s**,k) the intensity values of the reference image at pixel **s** and time k, and R the search region, the displacements vector $\mathbf{v} = (v_x, v_y)$ is obtained by *minimizing* MAD over the search region. M and N (Eq. 1) are the dimensions of the reference block.

$$\mathrm{MAD(v)} = (\mathrm{M} \cdot \mathrm{N})^{-1} \cdot \sum_{s \in R} \left| \mathrm{I}(s,k) - \mathrm{I}(s-v, k-\Delta k) \right| \qquad (1)$$

The resulting vector field of the individual blocks of the overall view can be analysed as blocks, in which blocks with identical (*length, direction and neighbourhood*) displacement vectors are combined into an object. These blocks establish the *object candidates* (cluster). The blocks outside of this cluster are analysed as outliers or other objects because the displacement vectors don't belong to this cluster (*different direction)*. Increasing the accuracy of the exact object delimination can only be achieved by extending this segmenting procedure hierarchy. For the following *tracking analysis*, the selected object candidates are transformed into an adaptive colour space, so that a good result is achieved under the influence of *disturbances* (e.g. *brightness variations*, *shadow and small grey tone gradients*) in the sequence.

2.2 The Adaptive $K_1K_2K_3$- Colour Space

Good properties for a colour space with respect to tracking are among others illumination invariance and separability. The colour of the object then remains constant and distinct which makes tracking and detection easier and more reliable. A further demand would be the increase of the dynamic gain. To obtain optimal channel reduction and a high dynamic gain, we use an adaptive colour space after the segmentation of the moving region for the fulfilment of this demand. In the following linear transformation, the components K_i in the adaptive $K_1K_2K_3$- colour space for a pixel in the RGB colour space are obtained as:

$$[\mathrm{K}_1 \ \mathrm{K}_2 \ \mathrm{K}_3]^T = [Q_{ij}] \cdot [\mathrm{R} \ \ \mathrm{G} \ \ \mathrm{B}]^T \qquad (2)$$

For the generation of the transformation matrix Q_{ij} we use the KLT [6], in which the components K_i are aligned in the direction of the largest variances to obtain the largest possible contrast. The first component K_1 contains the brightness information. A larger dynamic gain is obtained by this component, compared with the conventional average value-based brightness H [7]. The other components K_2 and K_3 represent the chrominance information, which one uses for the suppression of the influence by *shadow* and *brightness variations*. This is obtained, since this channel occurs as *difference relation*. This behaviour corresponds to the visual colour perception in such a way that by the variation of the saturation or the brightness of a region the hue remains approx. constant. For the description of the variance of each component (*part variance*), a quality measure (Eq. 3) is defined as control criterion for the data in the respective channel. While the quality measure E_i indicates the preservation of the part variance in the respective channel, the E_3 describes the variance removed by the component K_3. The evaluation of the real image scenes shows that the value of the quality measure of E_3 is approx. above 98%. This is because grey tones and little saturated colours predominantly occur in the real scenes. Therefore the component K_3 will *be removed* by the tracking analysis of objects without significant information loss.

$$E_i = \lambda_i \cdot \left(\sum_{i=1}^{n=3} \lambda_i \right)^{-1} \quad \text{and} \quad E_3 = \frac{\lambda_1 + \lambda_2}{\lambda_1 + \lambda_2 + \lambda_3} \qquad (3)$$

2.3 The Motion Analysis

Next to the successful tracking of the extracted moving regions in the video sequences a **m**ulti **m**atching **m**ethod (**M**^{3}) will be used for the determination of the displacement vectors. In this concept, the improvement of the tracking quality during the correspondence determination in image sequences is clearly reached by the fact that on the one handside *object-adapted* image regions are utilized, and on the other *colour information* is evaluated for the improvement of fail safe characteristics. In order to exploit the characteristic of the individual channels optimally (table 1), it is necessary to summarise the channel-specific Q-criterion (e.g. MAD) for a combined total criterion (MAD_{K1K2}, Eq. 4).

Table 1. The characteristics of MAD-criterion under the influence of disturbances

	Small grey tone gradients	**Shadow and brightness fluctuation**	**Partial occlusion**	***Another disturbance***
MAD_{K1}	*reliable*	*unreliable*	*unreliable*	*reliable*
MAD_{K2}	*reliable*	*reliable*	*unreliable*	*unreliable*
MAD_{K1K2}	*reliable*	*reliable*	*unreliable*	*reliable*

From this the respective displacement vector is calculated. For this purpose a priority of the MAD criterion is suggested according to its reliability.

$$MAD_{K_1K_2}(V)^t = \left(w_1^t + w_2^t\right)^{-1} \cdot \left\{ w_1^t \cdot MAD_{K_1}(V)^t + w_2^t \cdot MAD_{K_2}(V)^t \right\}$$

$$K_1, K_2: \text{ Channels of the adaptive colour space.} \qquad (4)$$

$$w_i^t: \text{weight factor at time } t. \quad 0 < w_i \le 1; \text{ Whereby } w_i = \begin{cases} 0.4 \le w_i \le 1 & \text{reliable} \\ \text{else} \quad 0 & \textit{un}\text{reliable} \end{cases}$$

Inappropriate weighting factors will be determined in this equation by means of an artificial neural recognition system. This motivates the next part of the paper.

2.3.1 Artificial Neural Recognition System

The fundamental idea of generating the total criterion (Eq. 4) is based on the fact that the neural network supplies an output activation that weights the Q- criterion. Using these weighting factors w_i the two channel-specific Q-criteria are summed up. As a result, the influence of *shadow* or *lighting changes* could be reduced considerably by using the total criterion (Eq. 4). This means good results for the determination of displacement vectors can be obtained. For the determination of the weights factors (Eq. 4), a three-layer feed-forward network topology of a **M**ulti-**L**ayer **P**erception (MLP) is applied.

The feature vector (*learning data*) for the MLP is gained by an individual MAD criterion (*absolute features*) and/or from *two* consecutive MAD criteria (*difference features*). Such difference features are e.g. the change of the *minimum value*, the change of the MAD *margin values*, the change of the *average value* and of the *surrounding region values* of the MAD function. Thus, a better description of the image interference is obtained by combination the *difference* and *absolute* features. With the weighting factors (w_i) the reliability of the MAD criteria is not only evaluated under the influence of shadow or lighting changes, a partial *occlusion* can also be detected. This can take place, if the weighting factors w_1 and w_2 of the two components underrun a threshold (*approx. zero*) at the same time. Then a partial occlusion occurs as disturbance in the sequence. In this case the computed motion vector is unreliable. To solve this problem, a *recursive algorithm* is used. This will be clarified in the following section.

2.3.2 A Modification of the Kalman Filter

Own experiments on real-world *greyscales* images have shown that measuring values (motion vectors) calculated by this procedure are corrupted by noise. Especially, if there are applications with high requirements regarding the quality of motion vectors, it is necessary to minimize the influence of noise. A recursive filtering algorithm (Kalman filter (KF) [7,8]) is used for this task. The position of the minimum of the Q-criterion (*motion vector*) is the input of the filter.

Besides the capability of the filter to reduce the influence of noise, there are some other *important* advantages. For instance the inclusion of the internal system model makes it possible to *predict* the motion vector at the next time step. This can be used effectively to reduce the *search area* in the matching algorithm. However, in case of problematic image situations (e.g. *partial occlusion* of the tracked image region), outliers in the calculated motion vectors cause false estimates of the KF.

In order to cope with this problem, a modification of the conventional KF has been proposed by [10]. This modified KF is used in the proposed technique. It estimates the quality of the respective motion vector based on the weighting factors w_1 and w_2 (Fig. 1). That means that the incoming (false) motion vectors are weighted less than before and the KF uses its internal model for the estimation increasingly. As a result, the influence of the *partial occlusions* of the tracked image region on the KF estimates is reduced considerably.

3 Results of the Analysis

In the following sequence we demonstrate the suitability and capability of the proposed technique by the motion analysis. The object of interest is overlaid by *shadow* in $image_{k=5}$ and *lighting modifications* in $image_{k=10}$ in this sequence. These influences lead to clear modifications of the intensity values (*e.g. darkening Fig. 2 A*). In the *first step* the object is selected using the motion vector field. The resulting vector field is evaluated by the combination of blocks with identical displacement vectors to one object candidate. These blocks establish the object candidates, which are transformed afterwards into the adaptive colour space for the improvement of fail-safe characteristics (*suppression of the influence of shadow and lighting modifications*). For the object-adapted image region according to Fig. 2 ($image_{k=4}$), the calculated components K_i are forwarded to the second processing level of the technique. There, the motion analysis of the image region takes place via M^3. This also guarantees the determination of reliable motion vectors in extreme cases, if the real image objects are overlaid by disturbances. Generally, the M^3 operates like the BM and shows good results, since the block dimension is adapted to the object boundaries and *two channels* of the adaptive colour space for the computation of the Q-criterion (MAD) are used. Here the MAD function will be calculated according to Eq. 4 for all discrete displacement vectors v_x and v_y in the channels (K_1 and K_2). *For* this overlay of shadow in $image_{k=5}$, deformations and several minima occur in the Q-criterion. This occurs in the light-intensity-dependent component (K_1) (Fig.2 C) as well as in the total Q-criterion, if the matching is accomplished by using the image contents in the *RGB colour space* or in *Greyscale* image sequences. These incorrect results in the determination of the displacement vectors lead to the fact that the treated image region leaves the originally pursued region (Fig. 2 E) and the motion trajetory doesn't describe the actual position of the object. In contrast to it the Q-*criterion* of the component K_2 allow an error free motion estimation due to good minimum development (Fig. 2 C). For the automatic weighting of the channel-specific MAD (Eq. 4), a neural recognition system is used to suspend the faulty MAD function.

The training data for the neural network are formed from a quantity of feature vectors from the MAD function, which are in this case generated from the test sequence which contains sufficiently significant image interferences. Here the unreliable channel-specific MAD_{K1} of the component (K_1) is suppressed. In opposite to MAD_{K1} the reliable similarity criterion (MAD_{K2}) contributes to a high weight in the total criterion (Fig. 2 C). This can be demonstrated clearly via the MLP. When analysing the calculated MAD_{K1} functions, the MLP_{K1} supplies an output activation of zero ($w_1=0$) at the time $_{k=5}$ (Fig. 2 D). The other MLP_{K2} of systems evaluates the MAD_{K2} of the K_2 component as reliable at the same $time_{k=5}$ (w_2= 0.9). The total criterion allows an error free motion estimation (*motion trajectory* Fig. 2 E) due to a good minimum. The total criterion ensures improved results compared to the conventional procedures in the motion analysis (Fig. 2 E). If the moving object is not occluded in one frame and its match is partially occluded in the following frame, then the displacement vectors estimated for this object region may have some error caused by the changed shape of the region due to the partial occlusion. In this case the weighting factors w_1 and w_2 of the two components under-run a threshold (*approx. zero*) at the same time.

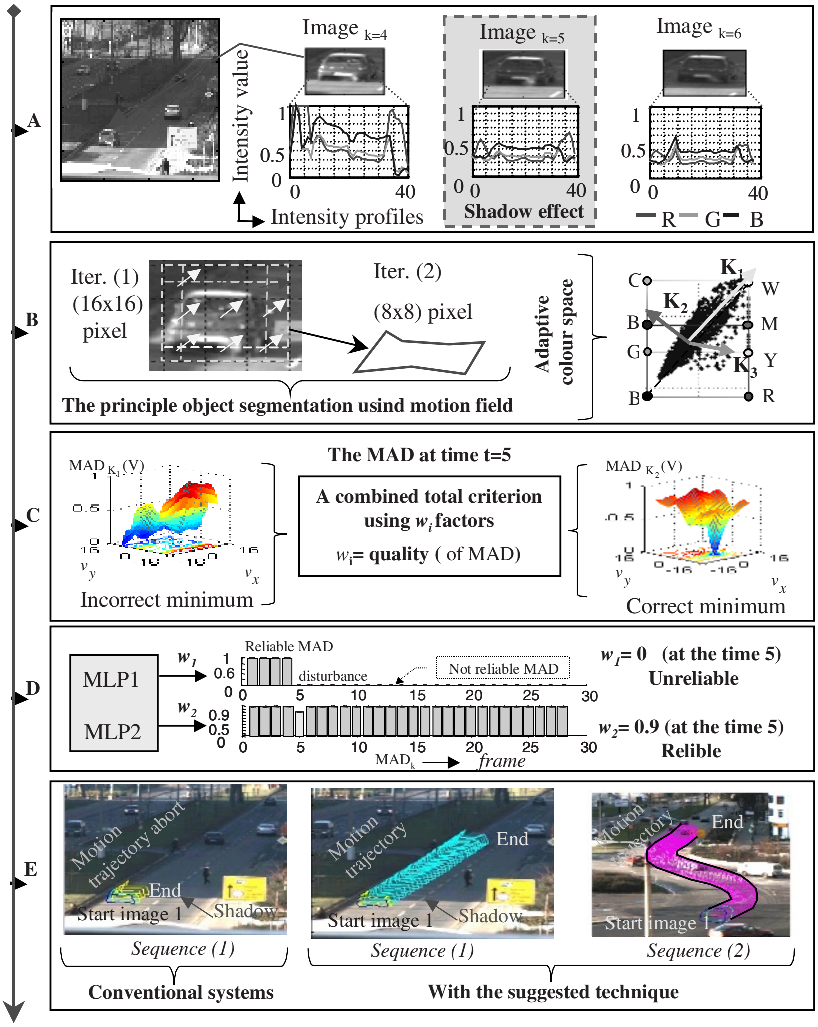

Fig. 2. The analysis of the image sequence under the influence of the shadow effect in the image$_{k=5}$ and lighting modifications in image$_{k=10}$. The calculated MAD functions for the object of interest is shown in C. The weight w_1 shows that the MAD_{K1} criterion starting at the time$_{k=5}$ can't be evaluated anymore (D). The MLP1 supplies an output activation w_1 of zero. In opposite to it MLP2 evaluates MAD_{K2} as reliable. In E) The results of the tracking analysis for an image region in a video sequence with conventional methods and with the suggested system structure (b). The right image shows the analysis for another sequence with the influence of disturbances (shadow).

For the solution of this problem those displacement vectors are used as input data for a recursive process of estimation, which additionally include the temporal *'prehistory'* of the motion in the analysis. This procedure generates a *back-up* trajectory, which is valid within a limited area in case of failure of the measuring information by occlusion. For the solution of this problem those displacement vectors are used as input data for a recursive process of estimation, which additionally include the temporal *'prehistory'* of the motion in the analysis. This procedure generates a *back-up* trajectory, which is valid within a limited area in case of failure of the measuring information by occlusion.

4 Summary and Conclusion

In this paper, a technique was suggested for the motion estimation of objects in video sequences. For an automatic object selection, a motion vector field was used. Because the BM for the displacement calculation under the influence of disturbance situations is quite sensitive, the use of an adaptive colour space during the tracking of the objects was suggested. For a successful tracking of the extracted moving regions in the video sequences, a $\mathbf{M}^3$ was used. The channel-specific criteria are combined to a total criterion according to their reliability, which show more *exact* and durable results compared to conventional procedures, in particular to problematic measuring situations by adaptive priority of the proportions. In the suggested technique, a modified recursive filtering algorithm was applied for the reduction of the influence of *partial occlusions* on the tracked image region. By the suggested technique, reliable results are achieved despite the influence of disturbance situations.

Acknowledgements. This work was supported by LSA/Germany grant (FKZ: 3133A/0089B).

References

1. Frank T., Haag M., Kollnig H. and Nagel H.: Tracking of occluded vehicles in traffic scenes. 7th European Conference on Computer Vision, pages 485-494, Cambridge, April 1996, Springer-Verlag.
2. Deriche R. and Faugeras O.: Tracking line segments. IVC; 8(4):261–270,1990
3. Rehrmann V.: Object-oriented Motion Estimation in colour image sequences; in Proc. Of the 5th European conference on computer vision; Springer-Verlag, June-1998. Freiburg.
4. Blake A., Curwen R. and Zisserman A: A framework for spatio-temporal control in the tracking of visual contours; Inter. Journal of Computer Vision; 11(2):127–145; 1993.
5. Badenas J.; Sanchitz J.M. and Pla F.: Using Temporal Integration for Tracking Regions in Traffic Monitoring Sequences; 15th International Conference on Pattern Recognition; Barcelona 2000; Sept; III; pp.1137–1140.
6. Devaux J.C.; Gouton P.; Truchetet F.: A erial colour image segmentation by Karhunen-Loeve Transform ; ICPR 2000 ; Barcelona 2000 ; Volume 1 ; pp. 309–312.

7. Al-Hamadi A.; Michaelis B.: Fulfilment of the demand for robustness and flexibility in the motion analysis systems by adaptive colour space transformation; CGIV'2002; April, 2002; Poitiers; France; pp.540–545.
8. Sorenson, H. W.: Kalman Filtering: Theory and Application. IEEE Press, ISBN 0-87942-191–6, 1985.
9. Maybeck, P S.: Stochastic Models, Estimation and Control. Acad. Press Inc., 1–3, 1982.
10. Mecke, R.; Al-Hamadi , A.; Michaelis, B.: A Robust Method for Block-Based Motion Estimation in RGB-Image Sequences. 14th ICPR 1998, Brisbane, Australien, pp. 663–667.