# Fulfilment of the Demand for Robustness and Flexibility in the Motion Analysis Systems by Adaptive Colour Space Transformation

*Ayoub Al-Hamadi and Bernd Michaelis*
*Institute for Electronics, Signal Processing and Communictions (IESK)*
*Otto-von-Guericke-University Magdeburg*
*D-39016 Magdeburg, P.O. Box 4210 Germany*
*E-mail: ayoub@iesk.et.uni-magdeburg.de*

## Abstract

In this contribution a system structure for the motion analysis of objects in video sequences will be presented. This system structure is adaptive and works robust in comparison with the conventional procedures of image interference's durably. The colour information builds the core of this system conception in connection with the iconic multi channel Matching (IMCM).

## 1. Introduction

Usually the conventional intensity-based methods of the motion analysis do not operate reliably at non-cooperative measuring situations as for example brightness variations, shadow throw or small grey tone gradients. In order to fulfill the demand for robustness and flexibility, an image analysis system is presented here for the estimation of motion trajectory of an object, which is based on an optimal use of the colour information. Thus general improvements are achieved using motion trajectory (displacement vectors). These improvements depend, however, strongly on the used colour space and the concrete aggregation of the colour component. In the suggested system structure, a colour space definition is used as a function (dependently) of the image contents. One arrives here by the colour information in connection with a multi-channel matching method to a total criterion. This total criterion leads to the clear improvement of the match quality and thus the determination of motion trajectory of the object, also in unfavourable situations, is ensured.

## 2. The System Structure

The suggested system structure is described by two processing levels (Fig. 1), whereby the first level deals with the analysis of the colour information. This contains the colour image recording with a CCD camera, the object selection and the adaptive colour space transformation for the interesting objects. In the second processing level of the system, the motion analysis of the selected objects is executed using the first processing level. The second processing level of the system structure is described by the multi-channel matching, which describes analysis of the combined total criterion with neural image-specific system recognition and the calculation of the displacement vectors.

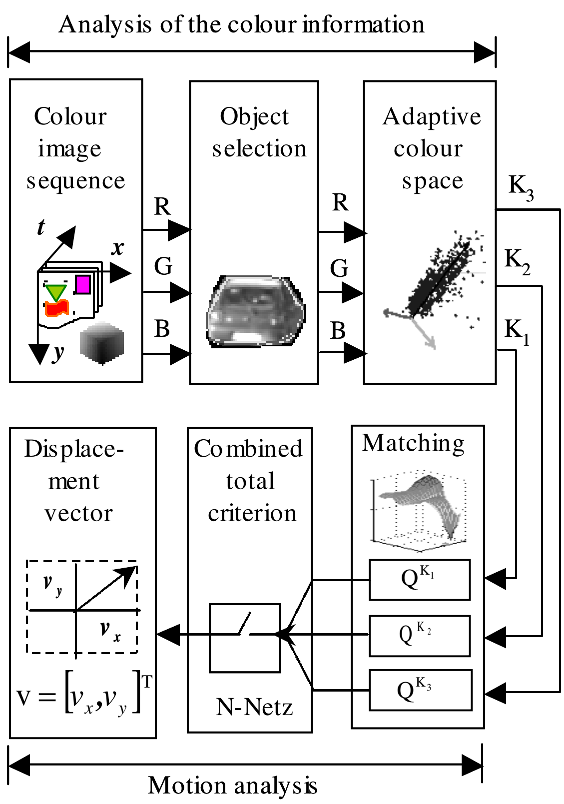


*Figure 1. The suggested system structure to the motion analysis.*

### 2.1 The Analysis of the Colour Information

#### *2.1.1 The object selection*

The target of the object selection (first processing level in the system structure) is the delimitation of the interesting image region, which is transformed afterwards into a new adaptive colour space. One applies here block matching method (full search) for automated initials object selection in the RGB image sequence.[2] One receives a displacement vector, which describes the determined movement of the represented picture or object window as a the result of the block matching for each reference block. The block-by-block calculation of the displacement vectors takes place via a Q- similarity criterion (as an example MAD, and nKKF).[2] This criterion supplies an extreme value, whose position represents the movement. Due to the small implementation expenditure the criterion of the middle absolute difference (MAD) was used:

$$\mathrm{MAD}(v_x, v_y) = \frac{1}{255 \cdot \mathrm{M} \cdot \mathrm{N}} \cdot \sum_{y=1}^{\mathrm{N}} \sum_{x=1}^{\mathrm{M}} \left| \mathrm{I_k}(x,y) - \mathrm{I_{k+1}}(x+v_x, y+v_y) \right| \qquad (1)$$

M and N are the dimensions of the reference block and $v_x$ as well as $v_y$ the displacements along the image axes. Here $I_k$ represents the intensity values of the reference area. The determination of the position of the reference area in the next image takes place through the localisation of the extreme value (minimum) of the MAD function. The resulting displacement vector fields of the individual blocks of the overall view can be analysed as blocks, in which blocks with similar displacement vectors are combined into an object (Fig. 2 a).

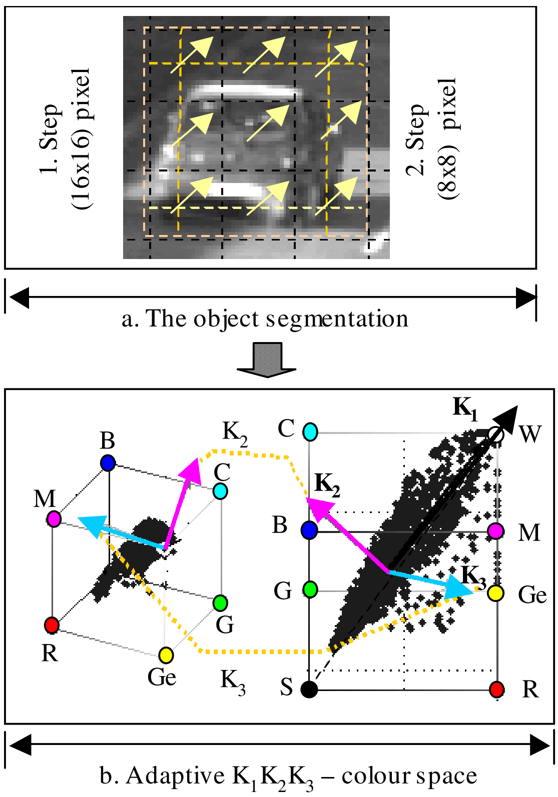


***Figure 2:*** *The principle of the object segmentation and the transformation into the adaptive colour space*

These blocks build the object candidates for the following tracking. Since however the object candidates (RGB of screen windows) in relation to non-cooperative measuring situations are sensitive, the use of adaptive colour space is suggested for the pursuit of the following object, which represents the core of the system.

### *2.1.2 The Adaptive $K_1K_2K_3$- Colour Space*

This adaptive $K_1K_2K_3$- colour space is weighted as a function of the image content, the used colour space components ensure the determination of the displacement vector of an object also in unfavourable situations. 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:

$$\begin{bmatrix} K_1 \\ K_2 \\ K_3 \end{bmatrix} = [A] \cdot \begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} q_{11} & q_{12} & q_{13} \\ q_{21} & q_{22} & q_{23} \\ q_{31} & q_{32} & q_{33} \end{bmatrix} \cdot \begin{bmatrix} R \\ G \\ B \end{bmatrix} \qquad (2)$$

For the determination of the linear transformation matrix $q_{ij}$ the Karhunen Loéve transformation[1,6] is suitable, in which the components $K_i$ in the sense of a large possible contrast in the direction of largest variances are aligned. The first main component $K_1$ contains the brightness information. A larger dynamic gain is obtained by the transformation, compared with conventional average value-based brightness component. The components $K_2$ and $K_3$ contain now the chrominance information. These components are orthonormal to each other. The first component in this adaptive colour space describes here the projection straight line (main component). Because this component $K_1$ indicates the largest eigenvalue, the weights $q_{1i}$ are positive in the projection channel. These are to attribute to the larger variance proportion, which this channel possesses. The total variance $\psi$ of the image information of the adaptive colour space contained in the transformed images is defined by the eigenvalues $\lambda_i$ of the covariance matrix (S), which is calculated from the colour values $F_i = [\mathrm{R\,G\,B}]_i^T$ of the regarded image region, in accordance with the following equations:

$$S = E\left\{ (F_i - E\{F_i\}) (F_i - E\{F_i\})^T \right\} \qquad (3)$$

$$\psi = \lambda_1 + \lambda_2 + \lambda_3 = 100\% \qquad (4)$$

The achromatic projection component ($K_1$) includes most of image information. For the description of the section variance or the packing density, a quality measure is inserted here as control criterion for the data in the respective channel. During the quality measure $E_i$ indicates the preservation of the section variance in the respective channel, the $E_3$ describes the variance removed by the component $K_3$. If the value of the quality measure of $E_3$ is situated close to 100%, then most information is collected in the two components $K_1$ and $K_2$.

$$E_i = \frac{\lambda_i}{\sum_{i=1}^{n=3} \lambda_i} \tag{5a}$$

$$E_3 = \frac{\lambda_1 + \lambda_2}{\lambda_1 + \lambda_2 + \lambda_3} \tag{5b}$$

The examination of the real image scenes shows that the value of the quality measure of $E_3$ is approximately above 98%. The $K_1$ component describes more than 90% of the total variance of the examined image scenes. This is to trace back to the fact that in the real image scenes predominantly grey tones and little saturated colours occur and thus percentage is strongly dependant on the respective viewed region and on the colour saturation.[2]

The components $K_2$ and $K_3$ describe colour differences, which can be used for overcoming the problems in particular with shadow throw and lighting modifications. The cause for the elimination of such influences consists of the fact that the weights of these components approximately approaches the average value zero (ε). If one adds a constant (Δ*h*) to the RGB-intensity values, then the (other) two components $K_2$ or $K_3$ remain almost constant.[2,3]

$$K_2(R, G; B) \approx K_2(R + \Delta h, G + \Delta h, B + \Delta h) \tag{6}$$

$$K_3(R, G; B) \approx K_3(R + \Delta h, G + \Delta h, B + \Delta h) \tag{7}$$

Calculated colour components for the image in Fig. 2:

$$K_2 = -0.741 \cdot R + 0.099 \cdot G + 0.663 \cdot B \approx$$
$$-0.741 \cdot (R + \Delta h) + 0.1 \cdot (G + \Delta h) + 0.663 \cdot (B + \Delta h)$$

$$K_2(R + \Delta h, G + \Delta h, B + \Delta h) \approx$$
$$K_2(R, G, B) + \Delta h \cdot \underbrace{(-0.741 + 0.1 + 0.663)}_{\varepsilon = 0.022} \tag{6a}$$

$$K_3 = 0.3237 R - 0.8134 G + 0.4833 B \approx$$
$$0.324 (R + \Delta h) - 0.8134 (G + \Delta h) + 0.483 (B + \Delta h)$$

$$K_3(R + \Delta h, G + \Delta h, B + \Delta h) \approx$$
$$K_3(R, G, B) + \Delta h \cdot \underbrace{(0.324 - 0.8134 + 0.483)}_{\varepsilon = -0.0061} \tag{7b}$$

This behaviour is similar to the characteristics of the U or V component (U and V describe colour differences in opponent colour space) of the perception-oriented HUV colour space and it corresponds to the visual colour perception that with variation of the saturation or the brightness of a colour keeps the colour hue approximately constant.[2,7,9]

### 2.2. The Motion Analysis

Furthermore, the well-known iconic matching[2,4] for the determination of the displacement vectors or the tracking of objects is applied to three channels (iMCM) (second processing level). Due to the different behaviour of the three components $K_1K_2K_3$ during the motion regulation in non-cooperative measuring situations, it is necessary to summarise the channel-specific MAD similarity mass for a combined total criterion ($MAD_{K1K2K3}$ Eq. 8) from which the respective displacement vector is calculated. For this purpose a priority of the MAD-similarity mass is suggested according to its reliability. If only small disturbances are present, frequently the achromatic component $K_1$ is sufficient to determine around the motion trajectories, since the position of the extremum is hardly influenced by smaller disturbances. This reason lies in the fact that the development of the primitive ones in the picture is strengthened by these achromatic components.

$$\begin{aligned} MAD_{K_1K_2K_3}(v_x, v_y) = \{ & w_1 \cdot MAD_{K_1}(v_x, v_y) + \\ & w_2 \cdot MAD_{K_2}(v_x, v_y) + \\ & w_3 \cdot MAD_{K_3}(v_x, v_y) \ \} \end{aligned}$$

Whereby $K_1, K_2, K_3$ : (8)

Channels of the adaptive color space.

$w_i$ : weight factor with $\sum_{i=1}^{3} w_i = 1$

Appropriate priority factors will be determined by means of an artificial neural network (feed forward network) from derived MAD features.[2,8] On this base good results in the sense of a reliable determination could be obtained by displacement vectors. The fundamental idea with the generation of the total criterion consists of the fact that the neural network supplies an output activation, which weight the channels. Using these weight factors $w_i$ the three channels can be summed up. That results in improving the similarity measure regarding the localisation of the minimum. This concept ensures the reduction of the influence of unreliable components.

## 3. Results of the Analysis

For the regarded, object-adapted image region according to Fig. 2, the component $K_1$ contains about 90% of the total variance of the image region. During the enlargement of the image region, the adaptive colour space turns into the HUV colour space. The $K_1$ component resembles the H- channel (brightness) of the HUV of colour space (conversion of straight line between black point and white point) more strongly. Regarding small or object-adapted image regions, $K_1$ strengthens the primitive ones (brightness transitions), while they are weakened with the H-channel. The $K_2$ and $K_3$- components in the adaptive colour space are analysed with the target of the minimising the influence of brightness fluctuations and shadow throw. After the transformation of the regarded object region in accordance with Eq. 2 into the adaptive colour space, the components $K_i$ are passed on into the second processing level of the system (motion analysis).

In the second processing level of the system structure, the motion of the image region (displacement vector) is determined over the position coordinates of the MAD minimum. By the use of object-adapted image regions from the first processing level of the system, there will be turn into to the iconic matching methode. This operates similarly as the block matching procedure. The iconic matching method brings very good results (*problems with the absence of a unique minimum and the screen problem eliminated*) compared to the block matching procedure, since the block dimension is adapted to the object boundaries using the iconic matching. The object boundaries are determined in this work by the use of displacement vector fields by means of full search block matching.[2] If one calculates now the MAD function for all discrete displacement vectors $v_x$ and $v_y$ in the channels, the result will be a matrix of the dimension:

$$(2 \cdot \Delta v_{x\,max} + 1) \cdot (2 \cdot \Delta v_{y\,max} + 1) \cdot 3 \qquad (9)$$

$\Delta v$ describe here maximum of expected shifts in horizontal and vertical direction. They correspond at well-known time to the difference between the next image acquisition and the translation speed of the object. The overlay of the object in this real image sequence by shadow throw leads to clear modifications of the intensity values (e.g. darkening with the shadow throw Fig. 3).

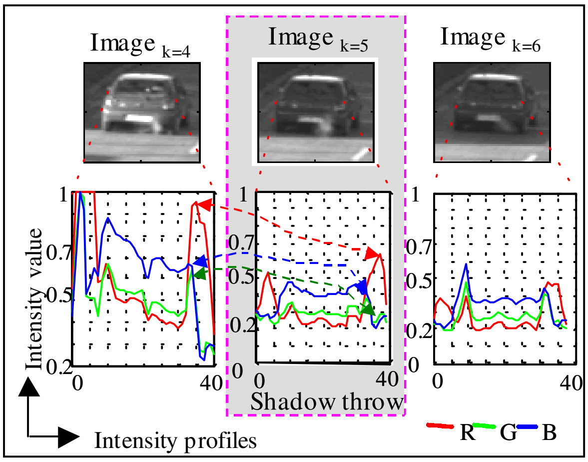


*Figure 3: Image sequence with the influence of the shadow throw in the image k=5.*

In the result of this overlay, deformations and several minima occur in the similarity measure. This occurs both in the light-intensity-dependent component as well as in the total criterion ($MAD_{RGB}$), if the matching is executed using the picture contents in the RGB colour space.[5] These incorrect results in the determination of the displacement vectors lead to the fact that the regarded image region of the originally pursued region jumps away (Fig. 5 a). The similarity *criteria* of the two components of $K_2$ and $K_3$ enable an error free movement regulation due to good minimum development. For the automatic weighting of the channel-specific MAD function (total criterion), a neural network is used here to suspend with the target the faulty MAD function.[2]

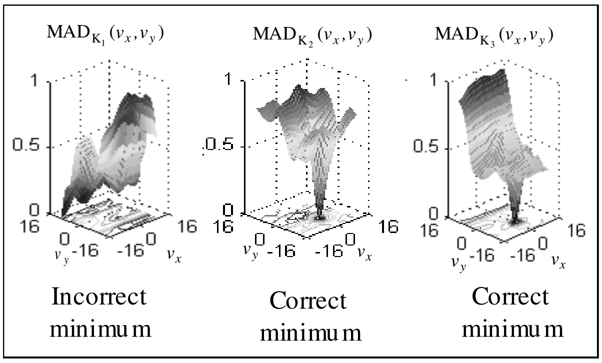


*Figure 4. The calculated MAD functions for the regarded object (during the overlay of a shadow throw at the time k=5 in the image sequence) in the respective channels of the adaptive colour space.*

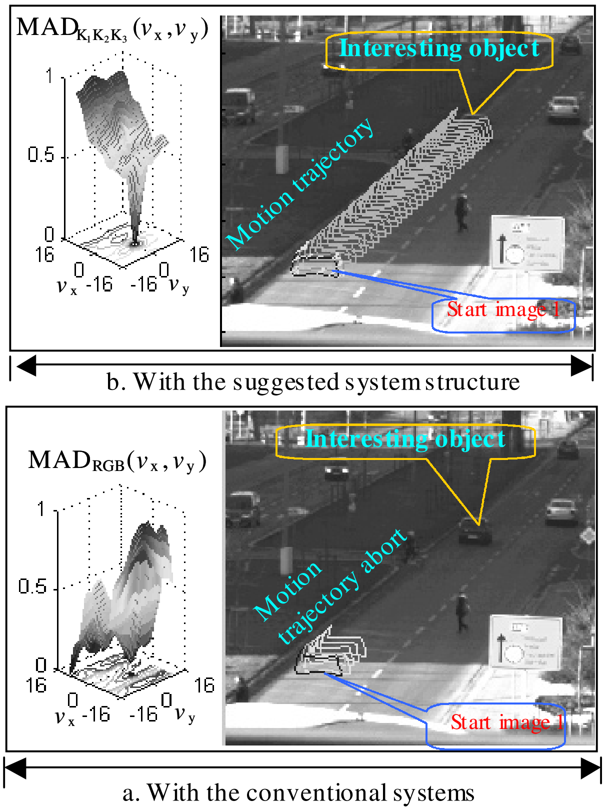


*Figure 5. Results of the motion analysis for a picture region in a picture sequence with conventional methods and with the suggested system structure.*

The training data for the neural network are formed from a quantity of feature vectors from the MAD function, which are generated here from the test sequence, which contains sufficiently significant image interference's. The unreliable channel-specific similarity criteria of the components of the adaptive colour space thereby are suppressed. In the opposite the reliable similarity criteria contribute to a high weight into the total criterion. The total criterion determined in such a way shows attainable improvements compared to the conventional procedures in the motion analysis (as shown in Fig. 5).

## 4. Summary and Conclusion

In this contribution, a system structure was suggested for the movement regulation of objects in colour image sequences. For an automatic object selection, block matching method (full search) was used. Because the block matching method (full search- block specified) for the displacement calculation in non-cooperative measuring situations is relatively sensitive, the use of an adaptive colour space during the tracking the objects was suggested.

It could be shown that by an adaptive colour space transformation the calculation of the similarity criterion can be implemented from the beginning for channels with pronounced intensity gradients. The channel-specific similarity criteria are combined according to their reliability into a combined total criterion, which supplies more exact and durable results, compared with conventional procedures, in particular to problematic measuring situations by adaptive priority of the proportions. By the suggested system structure, reliable results are achieved despite the influence of non-cooperative measuring situations.

The determined motion trajectories can be used as input data for a recursive process of estimation, with which additionally the temporal 'prehistory' of the motion is included in the analysis. Thus further improvements are possible for example with partial object covering.[5,8]

## Acknowledgements

This work was supported be LSA/ Germany grant (FKZ: 1441A/8386H)

## 5. Reference

1. Al-Hamadi A., Mecke, R.,Michaelis, B.: Verbesserte Bewegungsbestimmung bei der adaptiven Farbraumtransformation, FarbBV'99, 7-8.10.99, TU-Ilmenau, S. 75-81.
2. Al-Hamadi A.:Verbesserte Störsicherheit bei der Bewegungsanalyse in monokularen Farbbildsequenzen durch adaptive Farbraumtransformation, Dissertation, UNI-Magdeburg, 2001.
3. Al-Hamadi A.; Hofmeister, H.; Michaelis, B.; Brückner B.; Mecke, R.: Automatisierte Objektselektion und –verfolgung in einem adaptiven Farbraum; 6.FarbV2000; 05-06. Oktober 2000; GFaI Berlin; S. 27-34
4. Rehrmann, V.: Korrespondenzanalyse in Farbbildern, Dynamische Perzeption' 98, S. 23-30.
5. Mecke, R.; Al-Hamadi , A.; Michaelis, B.: A Robust Method for Block-Based Motion Estimation in RGB-Image Sequences. Proceedings of the 14th International Conference on Pattern Recognition (ICPR'98), Brisbane, Australien, 17.-20. August 19998, Volume I, S. 663-667.
6. Wahl F. M.: Digitale Bildsignalverarbeitung; Band 13; Springer-Verlag; 1989.
7. Schuster, R.: Adaptive Modeling in color image sequence. In: Mustererkennung 1994, ed. Kropatsch, W., pp.161-169, Springer Verlag, 1994.
8. Mecke, R.: Grauwertbasierte Bewegungsschatzung in monokularen Bildsequenzen unter besonderer Berücksichtigung bildspezifischer Störungen, Dissertation, Otto-von-Guericke-Universität Magdeburg, Fakultät Elektrotechnik, 1999.
9. Pomierski, T.; Neurophysiologisch motivierte Architektur zur Erzeugung stabiler Farb- und Texturrepräsentationen. PhD Thesis, Technische Universität Ilmenau, 1996.

## 6. Biography

Ayoub Al-Hamadi received his diploma engineer (Dipl.-Ing.) degree in Electro-technology and information technology of the Otto von Guericke University Magdeburg, Germany in 1997 and a PhD thesis in technical computer science (image processing) from the Magdeburg University in 2001. Since 1997 he has worked in the institute for Measurement technology and electronics University Magdeburg. His research work concentrates in the field of the image processing and tracking analysis.