

Verbesserte Bewegungsbestimmung durch adaptive Farbraumtransformation

Ayoub Al-Hamadi, Rüdiger Mecke, Bernd Michaelis
Institut für Prozeßmeßtechnik und Elektronik (IPE)
Otto-von-Guericke-Universität Magdeburg
D-39016 Magdeburg, P.O. Box 4120
e-Mail: Ayoub@ipe.et.uni-magdeburg.de
URL: <http://ipe.et.uni-magdeburg.de/TI/>

Kurzfassung

Fragen der Analyse bewegter Szenen haben stark an Bedeutung gewonnen. Für ungestörte Situationen existieren leistungsfähige Verfahren zur Bestimmung der Bewegungsvektorfelder bzw. für das Tracking von Objekten. Diese erfüllen jedoch oftmals die Forderungen hinsichtlich der Robustheit nur unzureichend bzw. versagen häufig bei nichtkooperativen Meßsituationen wie beispielsweise schlechten Beleuchtungsverhältnissen, Schattenwurf oder geringen Grauwertgradienten im Bild. Durch die Nutzung von Farbbildern werden generell Verbesserungen erwartet. Diese hängen jedoch stark vom verwendeten Farbraum und der konkreten Aggregation der Information der Farbkomponenten ab.

In der vorliegenden Arbeit wird eine Farbraumdefinition in Abhängigkeit von den Bildinhalten verwendet. Dieser *adaptive* Farbraum gewährleistet in Verbindung mit einem Blockmatching-Verfahren die Bestimmung der Bewegungsgrößen eines Objektes auch in ungünstigen Situationen. Es erfolgt gegebenenfalls eine Auswahl der Wichtung der genutzten Farbraumkomponenten in Abhängigkeit vom Szeneninhalte. Die erreichten Verbesserungen werden anhand ausgewählter Meßsituationen demonstriert, die bei Anwendung des konventionellen Blockmatching-Verfahrens zu fehlerhaften Ergebnissen führen.

1 Berücksichtigung der Farbinformation

Mit der zunehmenden Verfügbarkeit preiswerter Farbkameras und der Leistungssteigerung der Bildaufnahme- und Rechentechnik liegt es nahe, die Farbinformation auch für die intensitätsbasierte Bewegungsanalyse zu nutzen. Da mit der Einbeziehung von Farbe zusätzliche Informationen über die beobachteten Objekte zur Verfügung stehen, kann prinzipiell von Verbesserungen hinsichtlich der Zuverlässigkeit und Genauigkeit im Vergleich zu grauwertbasierten Verfahren ausgegangen werden [1].

Betrachtet man die Farben im RGB-Farbraum, so liegen drei farbspezifische Kanäle (Rot-, Grün- und Blaukanal) vor. Dieser Farbraum entspricht unmittelbar dem Sensorsignal der meisten CCD-Farbkameras, d.h. die pro Farbkanal aufgenommene Bildinformation wird nach der Bildaufnahme in der Regel nicht in andere Farbräume transformiert. Für die hier angestrebte Bewegungsanalyse erscheint zunächst der sensornahe RGB-Farbraum als Ausgangsbasis sinnvoll.

In der vorgestellten Arbeit werden translatorische Bewegungsgrößen für quadratische Bildblöcke ermittelt. Verglichen mit anderen Ansätzen, die die Bewegungsanalyse auf bestimmte Merkmale in Intensitätsbildern wie z.B. Liniensegmente [2] oder flexible Konturmodelle [3] beschränken, werden in der vorgestellten Arbeit unmittelbar die Intensitäten des Bildsignals verwendet. Deshalb ist hier keine explizite Merkmalsextraktion erforderlich und es werden keine konkreten Annahmen über bestimmte Objektmerkmale vorausgesetzt. Es wird ein Vergleich von zeitlich aufeinanderfolgenden Bildinhalten (Bildblöcke) durchgeführt mit dem Ziel, aus der besten Korrespondenz die Bewegung zu ermitteln. Dieser Vergleich erfolgt üblicherweise über die Berechnung eines Ähnlichkeitsmaßes (z. B. MAD-Kriterium) gemäß folgender Beziehung:

$$\text{MAD}_{\text{RGB}}(v_x, v_y) = \frac{1}{255 \cdot M^2} \sum_{y=y_0}^{y_0+M} \sum_{x=x_0}^{x_0+M} \left| \frac{1}{3} \mathbf{R}_k(x, y) - \mathbf{R}_{k+1}(x + v_x, y + v_y) \right| \quad (1)$$

$$+ \frac{1}{3} \left| \mathbf{G}_k(x, y) - \mathbf{G}_{k+1}(x + v_x, y + v_y) \right|$$

$$+ \frac{1}{3} \left| \mathbf{B}_k(x, y) - \mathbf{B}_{k+1}(x + v_x, y + v_y) \right|$$

wobei $\mathbf{R}_k, \mathbf{G}_k, \mathbf{B}_k$: Bild im Rot-, Grün-, Blaukanal zum Zeitpunkt k

M : Blockgröße; v_x, v_y : horizontale und vertikale Bewegung

Die Bewegung des Bildblockes (Bewegungsvektor $\mathbf{v} = [v_x, v_y]^T$) wird durch die Lagekoordinaten des MAD-Minimums repräsentiert. Das angegebene herkömmliche MAD-Kriterium wird zu gleichen Anteilen aus den Betragsdifferenzen der einzelnen Farbkanäle gebildet.

In der Abb. 1 sind die Betragsdifferenzen pro Farbkanal (kanalspezifische Ähnlichkeitsmaße $\text{MAD}_{\text{R,G,B}}$) sowie das Gesamtkriterium gemäß der Gleichung (1) für einen exemplarischen Farbbildblock dargestellt. Zum Vergleich ist in der Abb. 1 (unten rechts) das Ähnlichkeitsmaß MAD_{SW} angegeben, das nach Konvertierung des Farbbildes in ein Grauwertbild berechnet werden kann. Es wird deutlich, daß in diesem besonders konstruierten Beispiel keinerlei Gradienten im Grauwertbild vorliegen und folglich keine Bewegung bestimmt werden kann, da das Ähnlichkeitsmaß kein Minimum aufweist. Das Beispiel zeigt exemplarisch, daß die Einbeziehung von Farbinformation wesentliche Vorteile mit sich bringen kann.

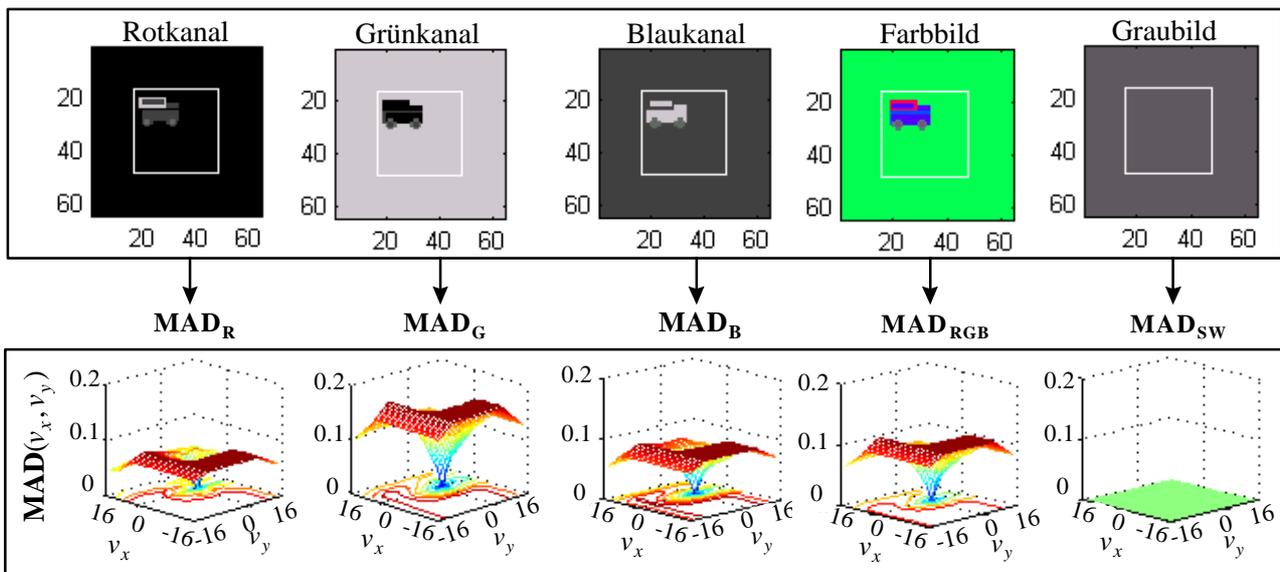


Abb. 1: Kanalspezifische Ähnlichkeitsmaße sowie MAD-Gesamtkriterium zu den darüber dargestellten Bildblöcken (s. auch Abb. 7)

Ausgehend von diesem Beispiel kann weiterhin festgestellt werden, daß je nach Bildinhalt des Farbkanals die einzelnen kanalspezifischen Ähnlichkeitsmaße eine unterschiedliche Ausprägung beispielsweise hinsichtlich des Wertebereiches oder der Minimumform besitzen. Im Sinne einer möglichst genauen und robusten Bewegungsbestimmung sind im allgemeinen steile Minima (s. Abb. 1, Grünkanal) wünschenswert. Da nicht zwangsläufig in allen Farbkanälen ausgeprägte Extremwerte vorliegen und das konventionelle Gesamtkriterium gemäß der Gleichung (1) eine Mittelwertbildung realisiert, werden ausgeprägte kanalspezifische Minima meist 'geglättet'. Besser wäre es, die Berechnung der Ähnlichkeitsmaße von vornherein für solche Bildstrukturen bzw. -kanäle zu realisieren, die eine genaue und zuverlässige Bewegungsbestimmung erwarten lassen. Die Voraussetzung hierfür ist eine Farbraumdefinition in Abhängigkeit von den Bildinhalten.

2 Adaptive Farbraumtransformation und Bewegungsbestimmung

Erfahrungsgemäß kann für Bildbereiche, die ausgeprägte Intensitätsgradienten (z. B. Hell-Dunkel-Übergänge, Ecken, Texturen usw.) enthalten, eine genaue und zuverlässige Bewegungsbestimmung realisiert werden. Hieraus resultiert der Ansatz, vor der Berechnung des Ähnlichkeitsmaßes eine adaptive Farbraumtransformation vorzunehmen. Durch die folgende lineare Transformation erhält man für einen Bildpunkt im RGB-Farbraum die Komponenten K_i :

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

Zur Bestimmung der Koeffizienten q_{ij} eignet sich die Karhunen-Loève-Transformation, bei der die Komponenten K_i im Sinne eines möglichst großen Kontrastes in Richtung größter Varianzen ausgerichtet werden (Hauptachsentransformation). Das entspricht einer Ermittlung der Eigenvektoren der Kovarianzmatrix C , die durch nachfolgende Beziehung aus den Farbwerten $F_i = [R \ G \ B]^T$ der Pixel des betrachteten Bildfensters berechnet wird:

$$C = E \left\{ (F_i - E\{F_i\})(F_i - E\{F_i\})^T \right\} \quad \text{wobei } E\{\cdot\} \dots \text{Erwartungswert} \quad (3)$$

Die orthonormalen Vektoren $\mathbf{q}_i = [q_{i1} \ q_{i2} \ q_{i3}]^T$ als Lösung des Eigenwertproblems bilden die Achsen eines neuen Koordinatensystems. Der zum größten Eigenwert gehörige Eigenvektor $\mathbf{q}_1 = [q_{11} \ q_{12} \ q_{13}]^T$ stellt die Hauptachse dar und weist in Richtung größter Streuung der Farbwerte, wenn jeweils der Inhalt des konkreten Bildfensters für die Transformation zugrunde gelegt wird.

Vergleicht man das aus der Komponente K_1 berechnete Ähnlichkeitsmaß mit dem Gesamtkriterium gemäß der Gleichung (1) (s. Abb. 2), erkennt man einen deutlich steileren Minimumverlauf des MAD_{K_1} -Kriteriums.

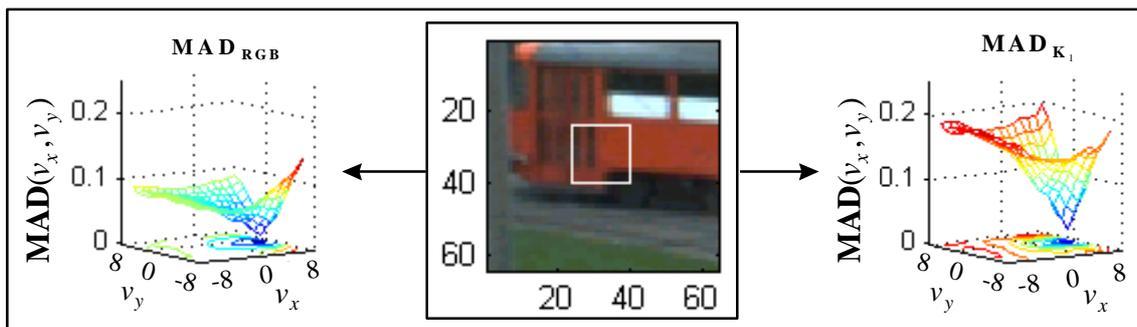


Abb. 2: Gesamtkriterium gemäß der Gleichung (1) (links) und aus der Komponente K_1 berechnetes Ähnlichkeitsmaß (rechts) für einen Bildblock (s. auch Abb. 8)

Da der jeweils resultierende Meßwert für die Bewegung mit Subpixelgenauigkeit durch Interpolation aus den Stützstellen des Minimums ermittelt wird, sind bei Verwendung des MAD_{K_1} -Kriteriums Verbesserungen der Genauigkeit möglich.

Vergleichende Untersuchungen an realen Bildsequenzen zeigten, daß der vorgeschlagene adaptive Farbraum gewisse Ähnlichkeiten zum wahrnehmungsorientierten HUV-Farbraum [4] aufweist. So repräsentiert die Komponente K_1 grob die Helligkeit (Luminanz) der betrachteten Bildregion. Die beiden Komponenten K_2 und K_3 stellen Farbdifferenzen dar. Damit sind diese Kanäle im Gegensatz

zur Komponente K_1 weitestgehend unabhängig von Helligkeitsänderungen und liefern beispielsweise in Bereichen konstanter Luminanz zusätzliche Bildmerkmale für die Analyse.

Das ist insbesondere bei der Bewegungsbestimmung in realen Bildsequenzen von Bedeutung, da hier häufig Beleuchtungsänderungen sowie Bereiche nahezu konstanter Luminanz auftreten und die ausschließliche Verwendung des Kanals K_1 somit zu Fehlern führen kann. In der Abb. 3 ist hierfür ein Beispiel aus einer realen Bildsequenz dargestellt. Darin wird ein bewegtes Objekt (Auto) von einem Schatten überlagert.

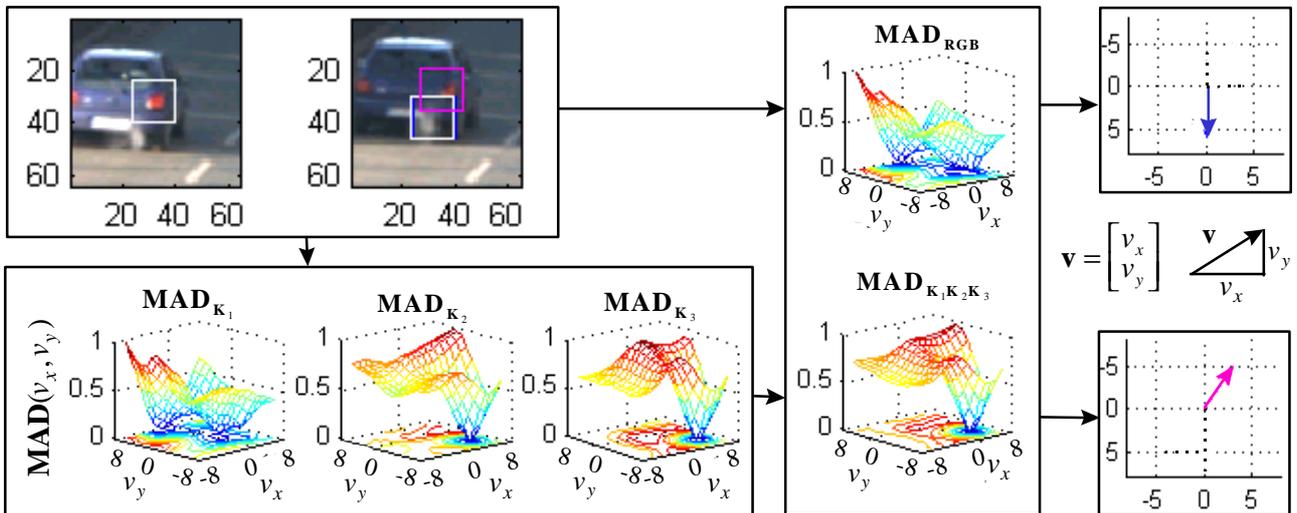


Abb. 3: Ausschnitt aus einer realen Bildsequenz und verschiedene Ähnlichkeitsmaße für einen exemplarischen Bildblock bei Schattenwurf (s. auch Abb. 8)

Im Ergebnis dieser Überlagerung treten im Ähnlichkeitsmaß des Kanals K_1 sowie im Gesamtkriterium gemäß der Gleichung (1) Deformationen und mehrere Minima auf, wodurch dann fehlerhafte Ergebnisse im resultierenden Bewegungsvektor vorliegen (s. Abb. 3, rechts). Das führt bei der Analyse dazu, daß der betrachtete Bildblock scheinbar von dem ursprünglich verfolgten Block wegdriftet. Die Ähnlichkeitsmaße der beiden Komponente K_2 und K_3 ermöglichen aufgrund der guten Minimausprägung hingegen eine genaue Bewegungsbestimmung. Die Ursache hierfür ist die bereits genannte weitgehende Unabhängigkeit der Farbdifferenzen in diesen Kanälen von Helligkeitsänderungen.

Ausgehend von diesem Beispiel erscheint es sinnvoll, aus den kanalspezifischen Ähnlichkeitsmaßen des vorgeschlagenen adaptiven Farbraumes ein kombiniertes Gesamtkriterium abzuleiten (s. Gleichung(2)). Bei diesem Gesamtkriterium wird über die Faktoren w_i eine Wichtung der einzelnen kanalspezifischen Ähnlichkeitsmaße (Betragsdifferenzen) entsprechend ihrer Zuverlässigkeit ermöglicht. Untersuchungen zeigten, daß solche Wichtungsfaktoren mittels eines künstlichen neuronalen Netzes aus abgeleiteten Merkmalen des MAD-Kriteriums ermittelt werden können [5].

$$\text{MAD}_{K_1 K_2 K_3}(v_x, v_y) = \frac{1}{255 \cdot M^2} \sum_{y=y_0}^{y_0+M} \sum_{x=x_0}^{x_0+M} w_1 \left| \mathbf{K}_{1k}(x, y) - \mathbf{K}_{1k+1}(x + v_x, y + v_y) \right| \quad (2)$$

$$+ w_2 \left| \mathbf{K}_{2k}(x, y) - \mathbf{K}_{2k+1}(x + v_x, y + v_y) \right|$$

$$+ w_3 \left| \mathbf{K}_{3k}(x, y) - \mathbf{K}_{3k+1}(x + v_x, y + v_y) \right|$$

wobei $\mathbf{K}_{1k}, \mathbf{K}_{2k}, \mathbf{K}_{3k}$: Kanäle im adaptiven Farbraum zum Zeitpunkt k

M : Blockgröße; v_x, v_y : horizontale und vertikale Bewegung

w_i : Wichtungsfaktoren wobei $\sum_{i=1}^3 w_i = 1$

Die Bewegungsinformation unzuverlässiger Kanäle wird somit unterdrückt, wohingegen diejenigen Kanäle, die eine hohe Zuverlässigkeit erwarten lassen, mit einer hohen Wichtung einbezogen werden. Das so ermittelte $\text{MAD}_{K_1 K_2 K_3}$ -Gesamtkriterium für ein Beispiel (s. Abb. 3) zeigt exemplarisch erreichbare Verbesserungen. In der Abb. 4 sind die resultierenden Bewegungsvektoren für 6 Folgebilder einer Testsequenz im Vergleich mit dem herkömmlichen Verfahren (gemäß Gleichung 1) dargestellt.

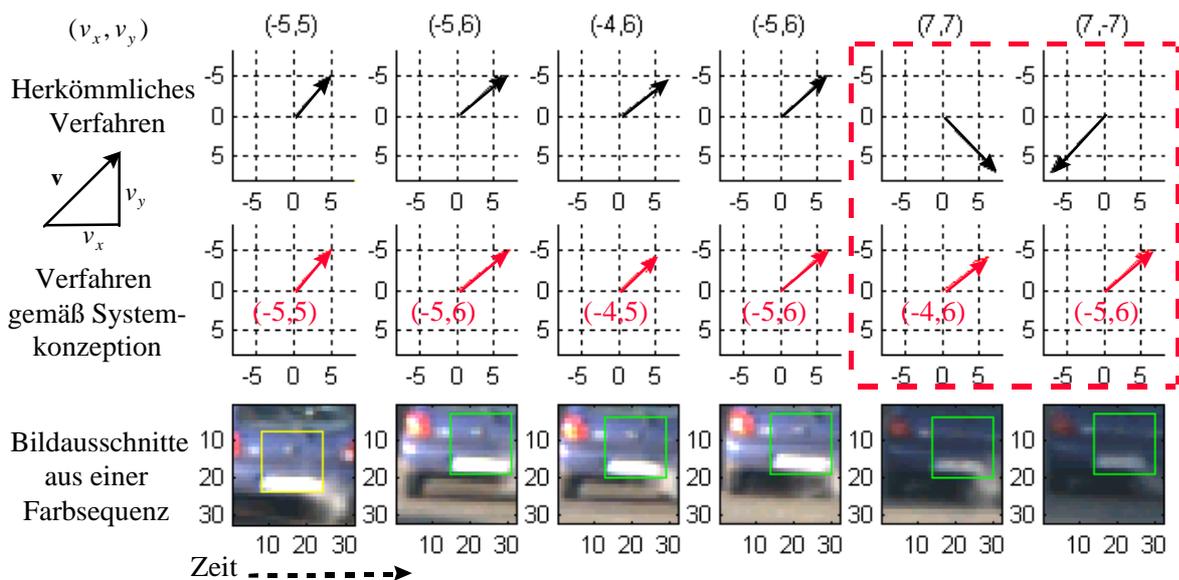


Abb. 4: Ermittelte Bewegungsvektoren für einen Block einer Farbbildsequenz (s. Abb. 9)

Bei Verwendung des herkömmlichen Verfahrens würde die Bewegungstrajektorie beim Eintritt des Objektes in den Schattenbereich nicht mehr die tatsächliche Objektbewegung beschreiben. Dahingegen ermöglicht die vorgeschlagene Systemkonzeption die Verfolgung des betrachteten Bildblockes über die gesamte Bildsequenz (s. Abb. 6).

In der Abb. 5 ist zusammenfassend die vorgeschlagene Systemkonzeption dargestellt, die pro Bildblock die Bestimmung des Bewegungsvektors \mathbf{v} realisiert.

Neben den translatorischen Bewegungsgrößen (Bewegungsvektor) können auf die beschriebene Weise prinzipiell auch rotatorische Bewegungen ermittelt werden. Dafür ist dann die Berechnung eines eindimensionalen Ähnlichkeitsmaßes erforderlich, dessen Minimum die Drehung des Bildblockes zwischen zwei Folgebildern repräsentiert. Die in

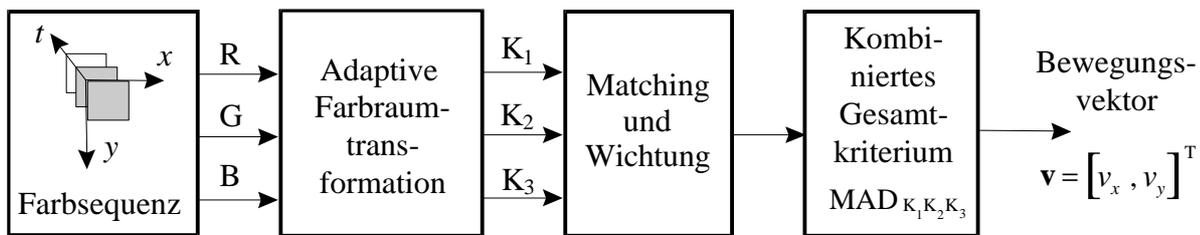


Abb. 5: Vorgeschlagene Systemkonzeption

einem Vektor zusammengefaßten Bewegungsgrößen bilden pro Folgebild einen Meßvektor. Diese Meßvektoren können als Eingangsdaten für ein rekursives Schätzverfahren eingesetzt werden, bei dem zusätzlich die zeitliche 'Vorgeschichte' der Bewegung in die Analyse einbezogen wird. Dadurch sind weitere Verbesserungen der Genauigkeit und Robustheit möglich. So kann beispielsweise der Einfluß statistischer Störeinflüsse erheblich reduziert werden. Insbesondere auch nichtkooperative Meßsituationen (partielle Objektverdeckungen, Beleuchtungsänderungen usw.) lassen sich durch die Einbeziehung der zeitlichen 'Vorgeschichte' bei der Analyse behandeln [5].

3 Zusammenfassung und Ausblick

In diesem Beitrag wurde ein System zur Bestimmung von Bewegungsgrößen in Farbbildsequenzen beschrieben. Es konnte gezeigt werden, daß durch eine adaptive Farbraumtransformation die Ähnlichkeitsmaßberechnung von vornherein für Kanäle mit ausgeprägten Intensitätsgradienten realisiert werden kann. Diese kanalspezifischen Ähnlichkeitsmaße werden entsprechend ihrer Zuverlässigkeit zu einem kombinierten Gesamtkriterium zusammengefaßt, das insbesondere in problematischen Meßsituationen durch adaptive Wichtung der Anteile genauere und robustere Ergebnisse, verglichen mit herkömmlichen Verfahren, liefert. Die Wichtung kann durch Auswertung von Merkmalen der Ähnlichkeitsmaße erfolgen. Ausgehend von den optimierten Bewegungsgrößen lassen sich mittels eines rekursiven Schätzverfahrens weitere Verbesserungen hinsichtlich der Genauigkeit und Robustheit erreichen.

Danksagung

Diese Arbeit wird durch das Land Sachsen-Anhalt gefördert (FKZ: 1441A/8386H).

Literatur

- [1] Koschan, A.; Rodehorst, V.; Spiller, K.: Color Stereo Vision Using Block Matching and Active Color Illumination, Proc. of 13th International Conference on Pattern Recognition, IEEE Computer Society Press, 1996, Bd. 1, S. 835-839.
- [2] Zhang, Z.: Token Tracking in a Cluttered Scene. Research Report N°2072, INRIA, [ISSN 0249-6399], Oktober 1993.
- [3] Blake, A.; et. al.: Affine-Invariant Contour Tracking With Automatic Control of Spatiotemporal Scale. In: Proc. of 4. Int. Conf. on ComputerVision, Berlin, 11.-14. Mai, 1993, IEEE Computer Society Press, [ISBN 0-8186-3870-2].
- [4] Schuster, R.: Objektverfolgung in Farbbildfolgen, Dissertation, TU München, 1995.
- [5] Mecke, R.: Grauwertbasierte Bewegungsschätzung in monokularen Bildsequenzen unter besonderer Berücksichtigung bildspezifischer Störungen, eingereichte Dissertation, Otto-von-Guericke-Universität Magdeburg, Fakultät Elektrotechnik, April 1999.



Abb. 6: Erste (links) und letzte (rechts) Aufnahme einer Farbbildsequenz. Der betrachtete Bildblock und die ermittelten Bewegungsvektoren sind gelb dargestellt .

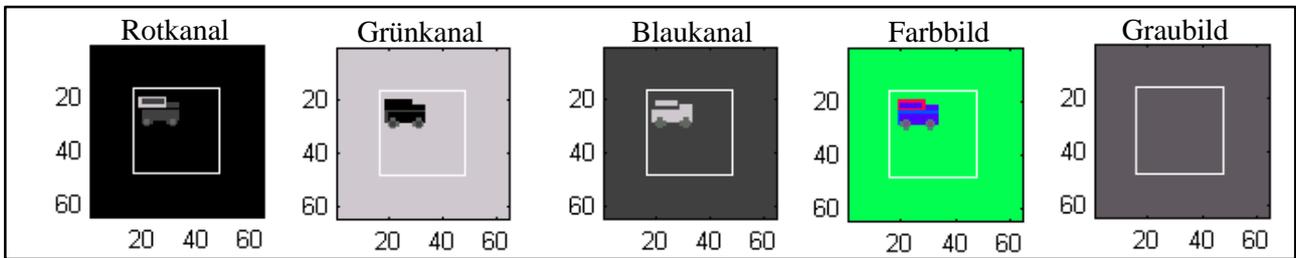


Abb. 7: Farbbilddarstellung zur Abb. 1 .

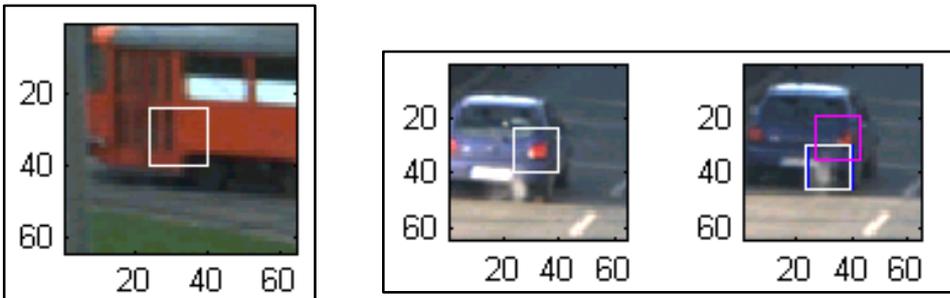


Abb. 8: Farbbilddarstellung zur Abb. 2 und Abb. 3 .

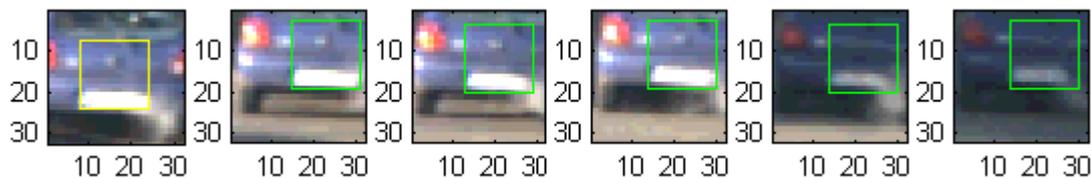


Abb. 9: Farbbilddarstellung zur Abb. 4 .