Bewegungskompensierte Akkumulation verrauschter Bildsequenzen

Oliver Grau

Institut für Theoretische Nachrichtentechnik und Informationsverarbeitung.
Abteilung "Automatische Bildinterpretation", Prof. Dr.–Ing. C.–E. Liedtke
Universität Hannover, Appelstr. 9A, 30167 Hannover,
EMail: grau@tnt.uni–hannover.de

Zusammenfassung. Der Beitrag zeigt ein Verfahren zur bewegungskompensierten Akkumulation von Bildobjekten aus Bildfolgen. Es dient der Reduktion von Artefakten und zur Verbesserung des Signal/Rauschabstandes im Bild. Dazu wird im ersten Schritt die relative dreidimensionale Bewegung des betrachteten Objektes zur Kamera gemessen. Als Form des Objektes wird hier ein Rechteck angenommen, das sich durch die Szene bewegt. Aus den gewonnenen Bewegungsparametern folgen affine Abbildungsparameter, mit denen die Abbildungen des Objektes perspektivisch entzerrt und paßgenau überlagert werden. Wird die Pixelauflösung im akkumulierten Bild gegenüber dem Originalausschnitt vergrößert, erhält das Resultat Superresolution.

Schlüsselwörter. Bildfolgenauswertung, Bewegungskompensation, Rauschunterdrückung, Superresolution

1 Einleitung

An Mustererkennungssysteme in industriellen Umgebungen oder beim Außeneinsatz werden harte Randbedingungen gestellt. Durch wechselnde oder unzureichende Beleuchtung stellen die verwendeten Bildsensoren oft nur Daten mit überlagertem Rauschen oder störenden Reflexionen zur Verfügung, die vor der Weiterverarbeitung aufbereitet werden müssen.

Für die Restaurierung des Bildmaterials bei systematischen Fehlern, wie insbesondere bei Bewegungsunschärfe, sind eine Reihe von Verfahren vorgeschlagen worden [1], die meist im Frequenzbereich arbeiten.

Störungen durch komplexe Artefakte, die beispielsweise wie in Abbildung 2 durch wechselnde Reflexionen auf der betrachteten Oberfläche entstehen, können in der Praxis in Einzelbildern nicht restauriert werden. Dies ist jedoch prinzipiell möglich, wenn die Information aus mehreren Bildern akkumuliert wird.

Der vorliegende Ansatz wertet deshalb eine Bildsequenz aus. Dabei wird durch eine bewegungskompensierte Überlagerung von Einzelbildern einer Bildsequenz eine Verbesserung des Signal/Rauschabstandes erreicht. Artefakte durch lokale Übersteuerung des Sensors, hervorgerufen durch wechselnde Beleuchtung, werden durch die Mittelung über die Bildsequenz ebenfalls unterdrückt.

In bestehenden Ansätzen wie [2] und [3] wird zur Kompensation der Objektbewegung nur translatorische Bewegung und Rotation in der Bildebene modelliert. Infolge der dreidimensionalen Bewegung des Objektes durch die Szene ändert sich neben der Lage der Objekte im Bild jedoch auch deren Größe und Form durch die perspektivische Abbildung. Der vorliegende Ansatz schätzt deshalb die dreidimensionale Bewegung.

Ist die dreidimensionale Bewegung bekannt, können die Einzelbilder in eine Normlage transformiert und aufakkumuliert werden. Wird für die Normlage eine höhere Auflösung als im Originalbildausschnitt gewählt, so bietet das Verfahren eine Rekonstruktion mit Superresolution. Desweiteren ergibt sich aus der Mittelung über die Bildsequenz eine Verminderung der Artefakte.

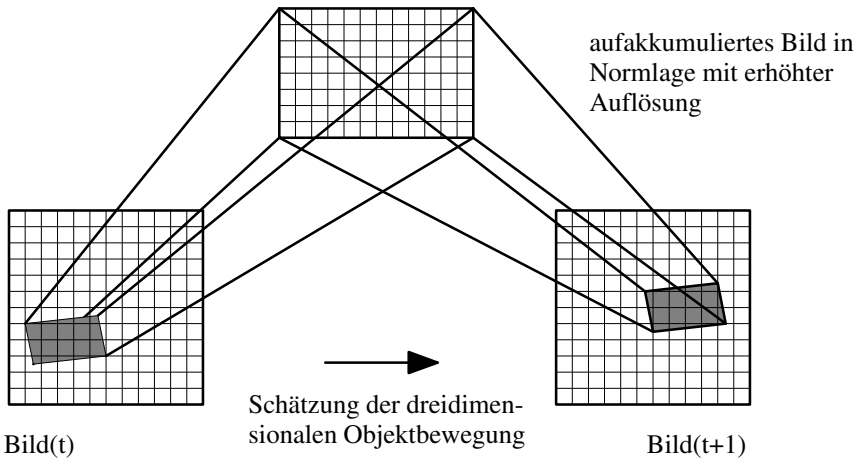

Abb. 1. Die Objekte aus zwei Bildern einer Sequenz werden bewegungskompensiert überlagert. Bei erhöhter Auflösung des akkumulierten Bildes erhält man Superresolution.

Voraussetzung für eine genaue Bewegungsschätzung ist, daß die Form des betrachteten Objektes bekannt ist. Für den hier betrachteten Fall wurde ein Rechteck angenommen.

Vor der Akkumulation der verfolgten Bildausschnitte werden die Abbildungen in eine Normposition transformiert. Dazu dient eine affine Abbildung. Die dafür notwendigen Abbildungsparameter werden aus den Bewegungsparametern gewonnen.

Der folgende Abschnitt beschreibt das Verfahren der bewegungskompensierten Akkumulation. Abschnitt 3 faßt die Ergebnisse einer konkreten Anwendung zusammen. In Abschnitt 4 wird das Verfahren und die erzielten Ergebnisse abschließend diskutiert.

2 Bewegungskompensierte Akkumulation

2.1 Bestimmung der Bewegungsparameter

Zur Bestimmung der dreidimensionalen Bewegungsparameter wird ein differentielles Verfahren angewendet, das in [4] beschrieben wurde.

Ein Punkt $\mathbf{P}_{(k-1)}= (P_x, P_y, P_z)^T$ im Raum, der zum Bild (k–1) der Sequenz gehört, wird unter der Annahme der Zentralprojektion an die Stelle $\mathbf{B}_1 = (B_x, B_y)^T$ der Bildebene projiziert. Liegt der Brennpunkt für die Zentralprojektion im Koordinatenursprung und hat die Bildebene ihren Ursprung an der Stelle $\mathbf{P}_{bild}= (0,0,F)^T$, so gilt:

$$B_x = F\frac{P_x}{P_z} \qquad B_y = F\frac{P_y}{P_z} \qquad \text{mit der Brennweite F} \tag{1}$$

Durch die Objektbewegung errechnet sich die neue Position des Punktes zu:

$$\mathbf{P}_{(k)} = [\mathbf{R}_G] \cdot (\mathbf{P}_{(k-1)} - \mathbf{G}) + \mathbf{G} + \mathbf{T} \tag{2}$$

mit $\mathbf{T} = (T_x, T_y, T_z)^T$ Translationsvektor

$\mathbf{G} = (G_x, G_y, G_z)^T$ Mittelpunkt des Objektrahmens

$[\mathbf{R}_G]$ Rotationsmatrix aus dem Rotationsvektor $\mathbf{R} = (R_x, R_y, R_z)^T$

Der Rotationsvektor **R** beschreibt eine komponentenweise Rotation eines Punktes um die drei Achsen des Koordinatensystems. Damit beschreiben die 6 Parameter (T_x, T_y, T_z, R_x, R_y, R_z) eine allgemeine dreidimensionale Bewegung eines Punktes. Durch Einsetzen in Gleichung (1) kann daraus die Position $\mathbf{B}_2$ im Folgebild bestimmt werden. Aus dem Displacement des Punktes $\mathbf{d} = \mathbf{B}_2 - \mathbf{B}_1$ und dem Grauwertgradienten **g** ergibt sich für ein linearisiertes Bildsignal:

$$\begin{aligned} \Delta I_{p1} &= \mathbf{g} \cdot \mathbf{d} = g_x \cdot d_x + g_y \cdot d_y \\ &= g_x \cdot (B_{2x} - B_{1x}) + g_y \cdot (B_{2y} - B_{1y}) \end{aligned} \tag{3}$$

ΔI_{p1} stellt somit die Grauwertdifferenz zwischen dem Bildpunkt und dem durch die Abbildunggleichungen und die Objektbewegung errechneten neuen Position im Folgebild dar. Durch Einsetzen in Gleichung (1) folgt:

$$\Delta I_{p1} = f \cdot g_x \cdot \left(\frac{P_{(k)x}}{P_{(k)z}} - \frac{P_{(k-1)x}}{P_{(k-1)z}}\right) + f \cdot g_y \cdot \left(\frac{P_{(k)y}}{P_{(k)z}} - \frac{P_{(k-1)y}}{P_{(k-1)z}}\right) \tag{4}$$

Nach Substitution von $\mathbf{P_{(k)}}$ als Funktion von $\mathbf{T}$ und $\mathbf{R}$ folgt:

$$\begin{aligned}\Delta I_{B1} \cdot P_z^2 \;=\; & f \cdot g_x \cdot P_z \cdot \mathbf{T_x} + f \cdot g_y \cdot P_z \cdot \mathbf{T_y} - [\, \Delta I_{B1} \cdot P_z + f \cdot P_x g_x + f \cdot P_y g_y \,] \cdot \mathbf{T_z} \\ & - [\, \Delta I_{B1} \cdot P_z \cdot (P_y - G_y) + f \cdot P_x \cdot g_x \cdot (P_y - G_y) + f \cdot P_y \cdot g_y \cdot (P_y - G_y) \\ & \quad + f \cdot P_z \cdot g_y \cdot (P_z - G_z) \,] \cdot \mathbf{R_x} \\ & + [\, \Delta I_{B1} \cdot P_z \cdot (P_x - G_x) + f \cdot P_x \cdot g_x \cdot (P_x - G_x) + f \cdot P_y \cdot g_y \cdot (P_x - G_x) \\ & \quad + f \cdot P_z \cdot g_x \cdot (P_z - G_z) \,] \cdot \mathbf{R_y} \\ & + [\, f \cdot P_x \cdot g_y \cdot (P_x - G_x) - f \cdot P_z \cdot g_x \cdot (P_y - G_y) \,] \cdot \mathbf{R_z}\end{aligned}$$

mit $(P_x, P_y, P_z)^T = \mathbf{P}_{(k-1)}$. (5)

ΔI_{B1} wird für eine Reihe von Meßpunkten bestimmt. Durch Einsetzen der Meßpunkte in Gleichung (5) folgt ein überbestimmtes lineares Gleichungssystem mit sechs Unbekannten (T_x, T_y, T_z ,R_x, R_y, R_z), das mit der Methode der kleinsten Fehlerquadrate gelöst werden kann. Als Meßpunkte dienen Bildpunkte, die einen hohen Gradienten im Grauwertverlauf aufweisen. Die notwendige Anzahl an Meßpunkten muß an das Bildmaterial angepaßt werden. In der Praxis haben sich für die meisten Fälle 10–20% der Punkte im Objekt mit dem höchsten Gradienten als ausreichend erwiesen.

Ist die Orientierung des Objektes in der Szene nicht genau bekannt, so kann für P_z ein zweckmäßiger Anfangswert angenommen werden, der dann lediglich einen Skalierungswert darstellt. In der weiteren Rechnung geht nur die relative Änderung der Z–Komponente durch die Objektbewegung ein.

Da die Annahme eines linearen Bildsignalverlaufs im realen Bildmaterial praktisch immer verletzt ist, wird das oben skizzierte Verfahren iterativ angewendet. Die Iteration wird abgebrochen, wenn sich die quadratische Summe der Grauwertdifferenzen zwischen den Iterationen nicht mehr signifikant verkleinert.

Eine Konvergenz des Verfahrens tritt nur auf, wenn die Objektbewegung zwischen den Einzelbildern nicht zu groß ist. Der Fangbereich des Verfahrens läßt sich jedoch für praktische Anwendungen dadurch erhöhen, daß eine initiale Prädiktion der Bewegung vorgegeben oder eine Schätzung der zweidimensionalen Bewegung in der Bildebene durch einen einfachen Blockmatcher vorgeschaltet wird.

2.2 Affine Abbildung

Um den viereckigen Ausschnitt eines Bildes in die Normlage zu bringen, wird eine affine Abbildung verwendet:

$$B'_X = a_1 B_X + a_2 B_Y + a_3 \quad (6)$$

$$B'_Y = a_4 B_X + a_5 B_Y + a_6 \quad (7)$$

Die Bestimmung der sechs Abbildungsparameter erfolgt aus drei Eckpunkten des zugehörigen Objektrahmens. Dazu werden die geschätzten dreidimensionalen Koordinaten der drei Eckpunkte mit Hilfe von Gleichung (1) in die Bildebene projiziert und in Gleichung (6) und (7) eingesetzt. Es ergeben sich insgesamt sechs Gleichungen aus denen die sechs Abbildungsparameter folgen:

$$
\begin{aligned}
B_{1x}' &= a_1 B_{1x} + a_2 B_{1y} + a_3 & \qquad B_{1y}' &= a_4 B_{1x} + a_5 B_{1y} + a_6 \\
B_{2x}' &= a_1 B_{2x} + a_2 B_{2y} + a_3 & \qquad B_{2y}' &= a_4 B_{2x} + a_5 B_{2y} + a_6 \\
B_{3x}' &= a_1 B_{3x} + a_2 B_{3y} + a_3 & \qquad B_{3y}' &= a_4 B_{3x} + a_5 B_{3y} + a_6
\end{aligned}
\tag{8}
$$

mit drei Eckpunkten $\mathbf{B_1}$, $\mathbf{B_2}$, $\mathbf{B_3}$ des Referenzobjektes und $\mathbf{B'_1}$, $\mathbf{B'_2}$, $\mathbf{B'_3}$ den zugehörenden Eckpunkten des Rahmens in der Normlage. Die Vorgabe der freien Parameter $\mathbf{B'_1}$, $\mathbf{B'_2}$, $\mathbf{B'_3}$ legt die Auflösung des akkumulierten Bildes fest.

Die transformierten Grauwerte werden durch bilineare Interpolation aus den Originalbildern gewonnen.

3 Ergebnisse

Das Verfahren zur bewegungskompensierten Akkumulation wurde unter anderem an dem in Abbildung 2 gezeigten Werkstück getestet. Das Bildmaterial ist stark verrauscht.

Ein manuell vorgegebener Rahmen wurde über insgesamt fünf Bilder verfolgt. Das Werkstück bewegt sich dabei durch den gesamten sichtbaren Bildausschnitt. Die Bewegungsschätzung wurde mit einer groben Prädiktion (zweidimensionale Displacments mit ca. +/– 3 Pel Genauigkeit) initialisiert. Das Verfahren konvergiert dabei nach 2–3 Iterationsschritten.

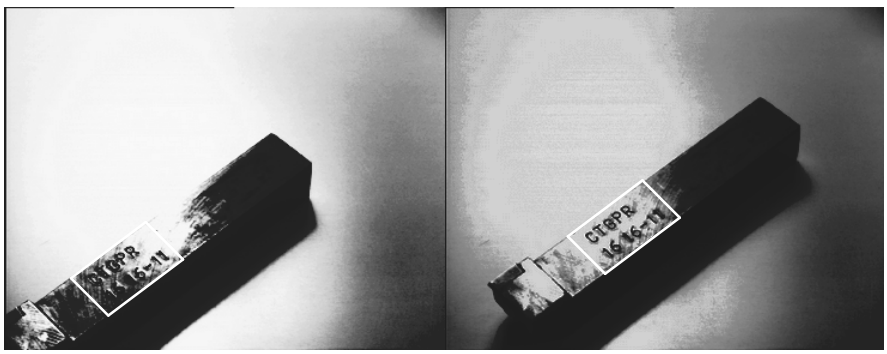

Abb. 2. Bilder eines bewegten Werkstückes mit eingezeichnetem Objektrahmen

Abbildung 3 zeigt die in Normlage transformierten Ausschnitte und Abbildung 4 das aufakkumulierte Resultat. Die Bilder in Normlage haben eine Auflösung von 120 x 200 Pixel, was etwa einer Verdreifachung der Originalauflösung entspricht.

Abb. 3. Sechs Bildausschnitte in Normposition

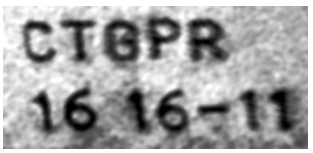

Abb. 4. Ergebnis der Akkumulation

Bei ca. 1000 Meßpunkten, die zum Aufstellen der Bewegungsgleichung dienten, benötigt das Verfahren ca. 0.3 s CPU–Rechenzeit auf einer Sun SPARC–Station 10/20. Das Signal/Rauschleistungsverhältnis hat sich deutlich vergrößert und die ungleiche Beleuchtung wurde ausgeglichen.

4 Diskussion

Das vorgestellte Verfahren eignet sich insbesondere für die Bildrestaurierung verrauschter oder mit störenden Reflexen auf den Objektoberflächen behafteter Bildsequenzen. Durch die Schätzung der dreidimensionalen Bewegung können auch starke Größenänderungen der Objekte in der Abbildung ausgeglichen werden. Weist das Bildmaterial zusätzlich starke Bewegungsunschärfe auf, so kann die Bewegungsschätzung die notwendigen Parameter für ein auf das Einzelbild angewendetes Restaurationsverfahren liefern.

Das Verfahren kann sowohl für Superresolution–Anwendungen als auch als Vorverarbeitungsstufe zur Reduktion von Rauschen und Artefakten eingesetzt werden. Der Aufwand an Rechenzeit ist relativ gering und kann bei Anpassung an das Bild-

material durch Minimierung der notwendigen Meßpunkte für die Bewegungsschätzung optimiert werden.

Literatur

1. Rosenfeld, A.; Kak, A. C.: Digital Picture Processing
   Second Edition, Volume 1, Academic Press, 1982
2. Huang, T. S.; Tsai, R. Y.: Multi–frame image restoration and registration
   in Advances in Computer Vision and Image Processing (T. S.Huang, Ed.), Vol. 1, pp. 317–339, JAI Press, Greenwich, CT, 1984
3. Irani, M.; Peleg S.: Improving Resolution by Image Registration
   CVGIP: Graphical Models and Image Processing, Vol. 53, No. 3, May 1991
4. Kappei, F.: Modellierung und Rekonstruktion bewegter dreidimensionaler Objekte in einer Fernsehbildfolge
   Dissertation, Universität Hannover 1989