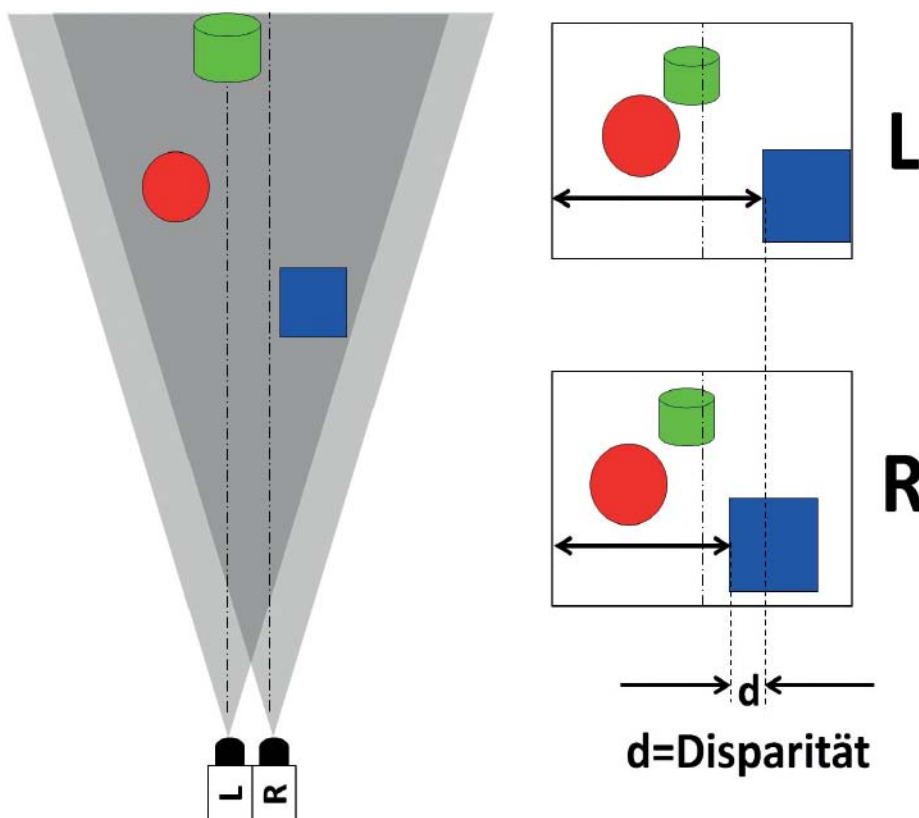


Schnelles Sehen im Raum

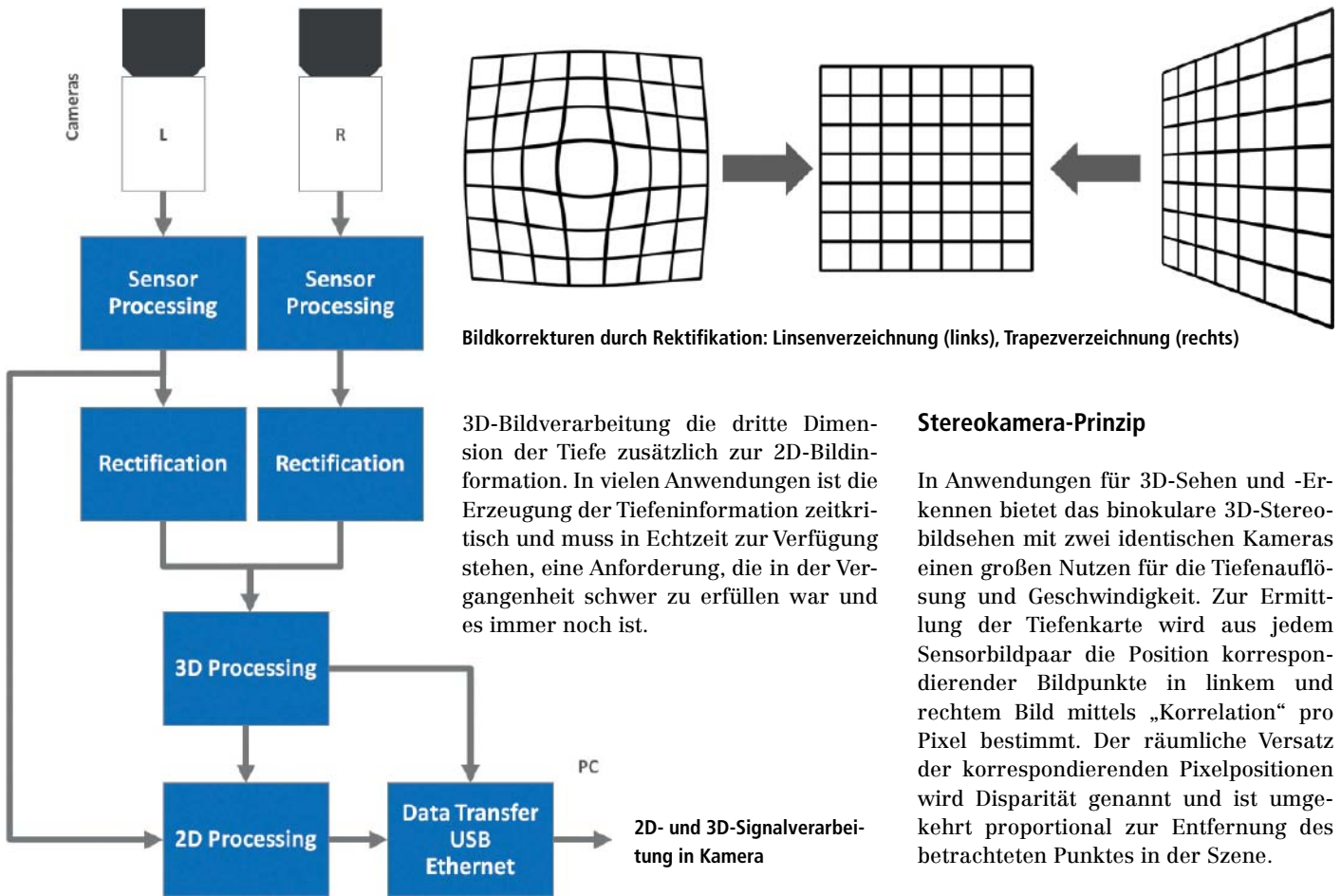
IP-Blocks für die Echtzeit-3D-Bildverarbeitung



Disparitätsberechnung aus linkem und rechtem Kamerabild

Unsere Welt ist 3D. Die Orientierung im Raum lernen wir Menschen als Teil unserer komplexen Entwicklung der Sinne. Spätestens wenn das Auto den Zaunpfiler touchiert, fragt man sich, ob oder wieso diese Entwicklung bereits abgeschlossen ist.

Menschliches dreidimensionales Sehen ist keine genaue Tiefenmessung und die menschliche Einschätzung von Größe und Entfernung des Wahrgenommenen ist ein Ausdruck des persönlichen Sinnesvermögens und Erfahrungsschatzes. Demgegenüber benötigen industrielle Anwendungen zur Lagebestimmung und Roboterführung („Griff in die Kiste“) eine präzise und schnelle Orientierung im Raum. Rauminformationen sind essenziell bei der Erzeugung von dreidimensionalen Schutzzonen für Maschinen oder Roboter und für die Charakterisierung von Objekten in Fahrerassistenzsystemen. Für diese Anwendungen liefert die



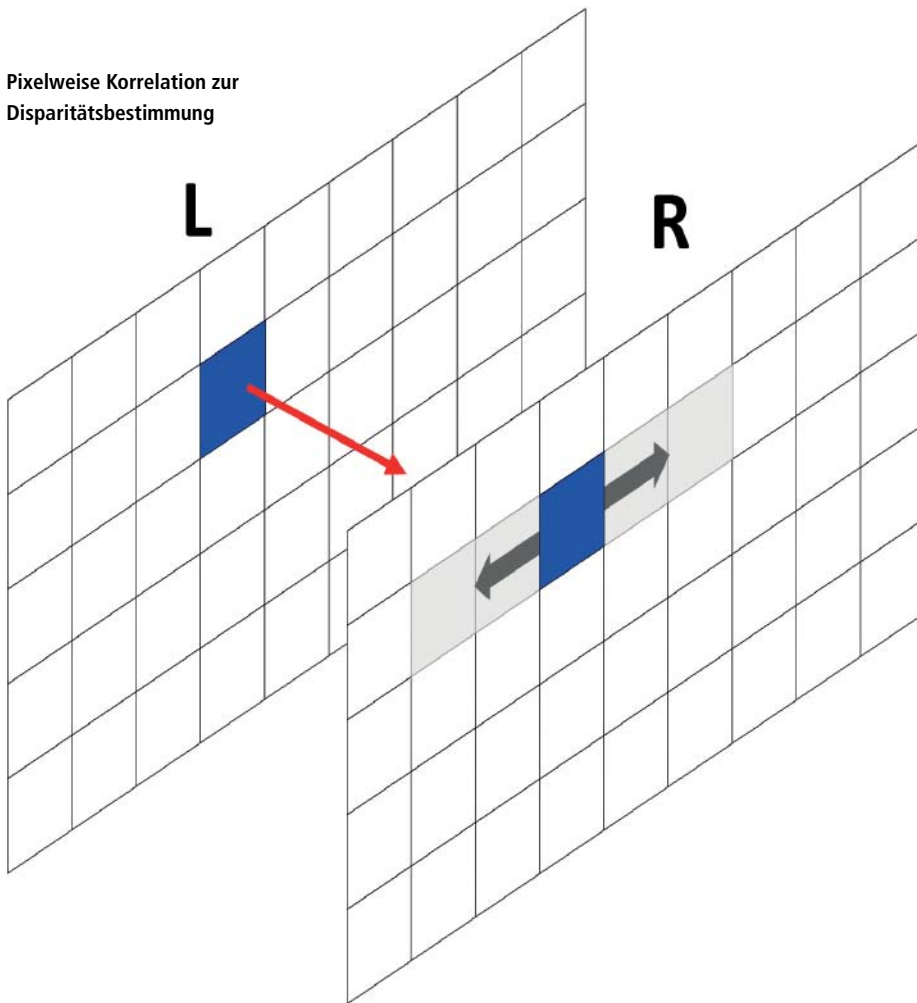
Bildkorrekturen durch Rektifikation: Linsenverzeichnung (links), Trapezverzeichnung (rechts)

3D-Bildverarbeitung die dritte Dimension der Tiefe zusätzlich zur 2D-Bildinformation. In vielen Anwendungen ist die Erzeugung der Tiefeninformation zeitkritisch und muss in Echtzeit zur Verfügung stehen, eine Anforderung, die in der Vergangenheit schwer zu erfüllen war und es immer noch ist.

Stereokamera-Prinzip

In Anwendungen für 3D-Sehen und -Erkennen bietet das binokulare 3D-Stereobildsehen mit zwei identischen Kameras einen großen Nutzen für die Tiefenauflösung und Geschwindigkeit. Zur Ermittlung der Tiefenkarte wird aus jedem Sensorbildpaar die Position korrespondierender Bildpunkte in linkem und rechtem Bild mittels „Korrelation“ pro Pixel bestimmt. Der räumliche Versatz der korrespondierenden Pixelpositionen wird Disparität genannt und ist umgekehrt proportional zur Entfernung des betrachteten Punktes in der Szene.

Pixelweise Korrelation zur Disparitätsbestimmung



Bildverarbeitung in der Kamera

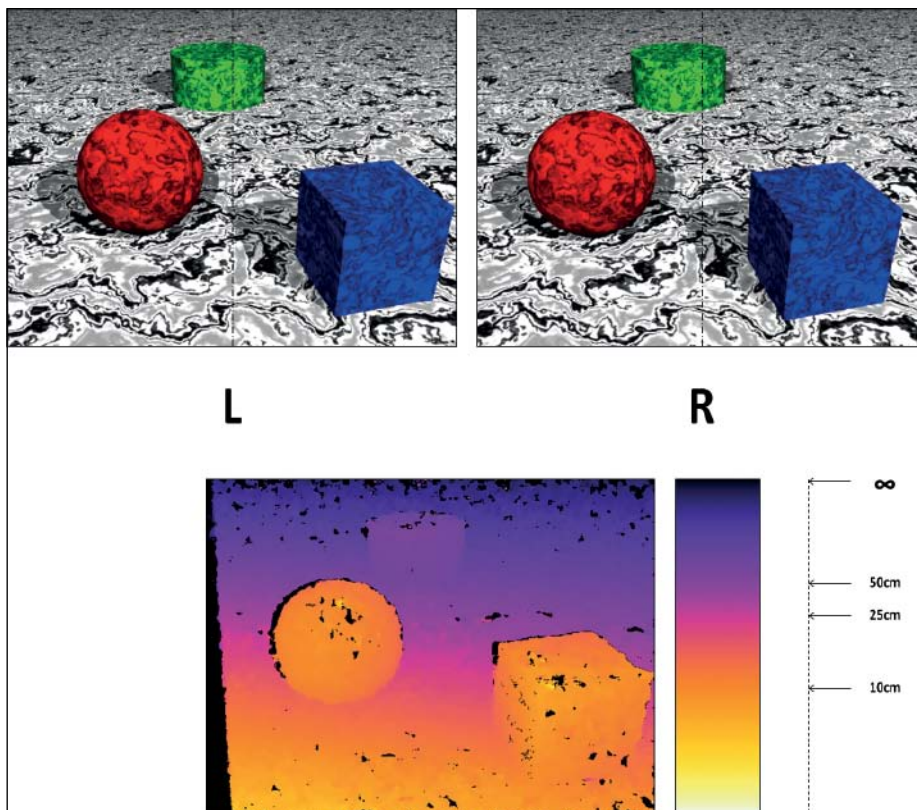
Die Bildverarbeitungs-komponenten einer industriellen Kamera umfassen Datenakquise, Bildsensoransteuerung, Bildvorverarbeitung und -konditionierung. Als weitergehende 2D-Bildverarbeitung sind eine globale bzw. lokale Bildanalyse verschiedener Bildeigenschaften (Farbe, Textur) oder auch einfache Basisoperationen, wie Binarisierung und morphologische Verfahren, in der Kamera möglich.

Zur 3D-Bildverarbeitung und speziell zur Berechnung von Tiefenkarten müssen unerwünschte Effekte, dazu zählen z.B. Linsenverzeichnungen, der Versatz der optischen Zentren der Bildsensoren und Trapezverzerrungen, korrigiert und die Abbildungen des linken und rechten Bildsensors in ein gemeinsames Koordinatensystem überführt werden. Diesen Schritt nennt man Rektifikation. Er ist die Grundlage für eine 3D-Bildverarbeitung. Dazu werden Korrekturmatri-zen für beide Kameras durch Kalibrierung bestimmt, indem Objektpunktfelder, häufig in Form von Schachbrettmustern, von beiden Kameras aufgenommen, analysiert und daraus Korrekturmatri-zen berechnet werden.

Mit dem 3D-Verarbeitungsschritt der Korrelation wird in den von der Rektifikation vorbereiteten Bildern pro Pixel nach Korrespondenzen in dem jeweils anderen Bild gesucht. Durch die Rektifikation wird sichergestellt, dass diese aufwändige Suche nur in einer Dimension, der x-Richtung, erfolgt. Vereinfacht dargestellt, wird ein Pixel von einem Bild der Stereokamera auf einer horizontalen Ebene des Bildes der anderen Stereokamera gesucht. Das Ergebnis der Suche ist der Wert der Disparität, welcher somit als ein Faktor der Pixelgröße des Bildsensors vorliegt.

Nach dieser Suche können die Disparitäten mittels Triangulation in einen echten Tiefenwert überführt werden, der ein Längenmaß senkrecht zur Stereokameraebene darstellt. Dazu sind die genauen Koeffizienten der Rektifikation, der Abstand der Kameras zueinander und die Brennweite der Kamera notwendig.

Die Tiefenkarte kann zusammen mit den Merkmalsextraktionen aus der 2D-Verarbeitung zu einer genaueren Bestimmung der geometrischen Struktur und Flächen im Raum verwendet werden.



Tiefenkarte mit Farbkodierung aus zwei Stereobildern

FPGA in der Bildverarbeitung

Mit der Verbesserung der Leistungsfähigkeit von FPGA haben sich neue Wege

in der Bildverarbeitung ergeben. Die intelligente Kamera mit FPGA kann heute Bildverarbeitungsfunktionen mit extrem hoher Rechenanforderung übernehmen. Die Parallelität der Verarbeitung in FPGA (alle Verarbeitungsschritte laufen immer und gleichzeitig) erlauben Ausführungsgeschwindigkeiten in Echtzeit bei deterministischen Durchlaufzeiten, wie sie z.B. in der Sicherheitstechnik oder bei Fahrerassistenzsystemen gefordert sind.

In der 3D-Bildverarbeitung ist die Berechnung der Tiefenkarte für hochauflösende Bildsensoren ein Beispiel für immense Rechenanforderungen. Die Korrelationsberechnung und Bestimmung der Disparität pro Pixel bedeutet je nach Algorithmus eine Rechenanforderung von mehreren tausend Operationen pro Pixel, was für hochauflösende Bildsensoren und bei einer Pixelrate von 75 Mpix/s schnell die Leistungsfähigkeit modernster CPU im Desktop-PC übersteigt. Ein FPGA kann viele Operationen parallel ausführen und damit z.B. eine Tiefenkarte mit 64 Disparitäten für HD Bildsensoren in weniger als 1 ms ausrechnen.

Bildverarbeitungsfunktionen in Kameras

Die Firma NET bietet FPGA-basierte Kameras mit Bildverarbeitung durch Verwendung von modularen IP-Blöcken an. Die IP-Bibliothek von NET umfasst speziell für die industrielle Bildverarbeitung Funktionen zur Farbverarbeitung, Linsenkorrektur, Rektifizierung und 3D-Korrelation, die als Grundlage für weiterführende Bildverarbeitung im PC oder auch im FPGA dienen. Das FPGA in der Kamera ist sowohl skalierbar als auch konfigurierbar und ermöglicht die weitere Integration für oder vom Kunden entwickelter Bildverarbei-

tungsfunktionen. IP-Blöcke sind kombinierbar und die Ergebnisse der verschiedenen Bildverarbeitungsfunktionen werden in der Kamera zur Übertragung über Standard-PC-Schnittstellen, wie USB oder Ethernet, zusammenfasst.

Moderne FPGA können sehr komplexe Aufgaben der Bildverarbeitung überneh-

men. Für rechenintensive und zeitkritische Aufgaben bietet der Einsatz von Kameras mit FPGA neue Möglichkeiten. Speziell für die 3D-Bildverarbeitung für hochauflösende Bildsensoren erreicht man durch Verwendung des FPGA in der Kamera deutlich schnellere Verarbeitungsgeschwindigkeiten als auf einem Desktop-PC.

► **Autor**
Alf Rieckmann,
Project Manager



► **Kontakt**
NET New Electronic Technology,
Finning
Tel.: 08806/9234-0
Fax: 08806/9234-77
info@net-gmbh.com
www.net-gmbh.com