



(12)

Offenlegungsschrift

(21) Aktenzeichen: **10 2016 208 056.3**

(22) Anmeldetag: **11.05.2016**

(43) Offenlegungstag: **16.11.2017**

(51) Int Cl.: **G01C 11/06** (2006.01)

G01B 11/24 (2006.01)

B60R 1/00 (2006.01)

G06T 7/20 (2017.01)

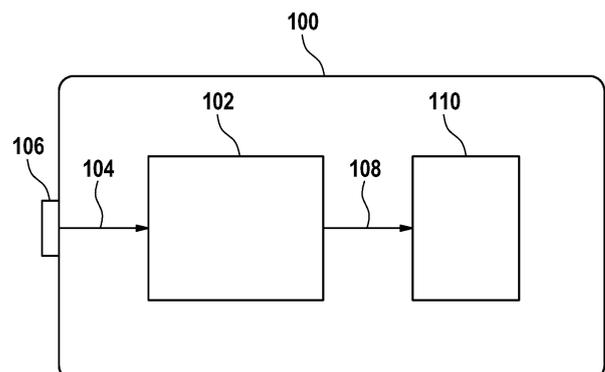
(71) Anmelder:
Robert Bosch GmbH, 70469 Stuttgart, DE

(72) Erfinder:
Simon, Stephan, 31079 Sibbesse, DE

Die folgenden Angaben sind den vom Anmelder eingereichten Unterlagen entnommen.

(54) Bezeichnung: **Verfahren und Vorrichtung zur Verarbeitung von Bilddaten und Fahrerassistenzsystem für ein Fahrzeug**

(57) Zusammenfassung: Die Erfindung betrifft ein Verfahren zur Verarbeitung von Bilddaten (104). Das Verfahren umfasst einen Schritt des Bestimmens von Korrespondenzdaten als Ergebnis einer Korrespondenzbildung zwischen ersten Bilddaten, die auf von einer Kamera (106) zu einem ersten Zeitpunkt erfassten Bilddaten (104) basierende Daten repräsentieren, und vorangegangenen Bilddaten, die auf von der Kamera (106) zu einem vorangegangenen Zeitpunkt erfassten Bilddaten (104) basierende Daten repräsentieren. In einem Schritt des Bestimmens werden Tiefendaten auf Basis einer unter Verwendung der Korrespondenzdaten durchgeführten Eigenbewegungsschätzung bestimmt. In einem Schritt des Ermitteln werden Fusionsdaten unter Verwendung der Tiefendaten und übertragenen Fusionsdaten ermittelt, wobei die Fusionsdaten eine auf den Bilddaten basierende Tiefeninformation umfassen.



Beschreibung

Stand der Technik

[0001] Die Erfindung geht aus von einer Vorrichtung oder einem Verfahren nach Gattung der unabhängigen Ansprüche. Gegenstand der vorliegenden Erfindung ist auch ein Computerprogramm.

[0002] Fahrerassistenzsysteme von Fahrzeugen können von Kameras erfasste Daten verwenden. Dazu können die Daten vor der Verwendung durch die Fahrerassistenzsysteme vorverarbeitet, beispielsweise verdichtet werden.

Offenbarung der Erfindung

[0003] Vor diesem Hintergrund werden mit dem hier vorgestellten Ansatz ein Verfahren zur Verarbeitung von Bilddaten, weiterhin eine Vorrichtung, die das Verfahren verwendet, ein Fahrzeugassistenzsystem sowie schließlich ein entsprechendes Computerprogramm gemäß den Hauptansprüchen vorgestellt. Durch die in den abhängigen Ansprüchen aufgeführten Maßnahmen sind vorteilhafte Weiterbildungen und Verbesserungen der im unabhängigen Anspruch angegebenen Vorrichtung möglich.

[0004] Der beschriebene Ansatz kann beispielsweise im Zusammenhang mit kamerabasierten Fahrerassistenzsystemen oder im Zusammenhang mit anderen Systemen mit Kameras, z. B. mobilen Robotern, eingesetzt werden.

[0005] Kamerabasierte Informationsgewinnung ist typischerweise geprägt von einer zügigen Verdichtung der Information. Das originale Messsignal kann dabei z. B. ein Flussfeld (optical flow map) oder eine Disparitätskarte (stereo disparity map) sein. Ergebnis der Verdichtung sind z. B. Stixel (spaltenweise Pixelgruppen) oder Objektboxen (bounding box). Verfolgung und gegebenenfalls Fusion erfolgt dann auf höherer Ebene.

[0006] Die frühe Verdichtung hat historische und architektonische Gründe. So kann beispielsweise der optische Fluss mit überschaubarem Ressourcenaufwand auf einem FPGA oder ASIC berechnet werden und ein dichtes Ergebnisfeld (hohe Datenmenge pro Bild) produzieren.

[0007] Den Flaschenhals stellt die weitere Auswertung und Verarbeitung dar, insbesondere wenn sie auf einem General Purpose Prozessor durchgeführt wird. Bei der notwendigen Verdichtung geht ein enormer Anteil an Information verloren.

[0008] Diese frühe Informationsverdichtung behindert auch die Entwicklung neuer Funktionen, da die

Architektur für die bestehenden Funktionen optimiert ist.

[0009] Ein „Occupancy Grid“ genanntes Belegungsgitter, das geeignet ist, eine Fusion von mehreren Messungen, die auch von mehreren Sensoren stammen können, zu unterstützen, kann genutzt werden. Dies wird z. B. in der Robotik häufig verwendet, beispielsweise zur Pfadplanung. Das Koordinatensystem entspricht häufig dem einer Landkarte (Vogelperspektive).

[0010] Bei „3D Punktwolken“ (point clouds) stellt ein einzelner solcher 3D Punkt eine 3D Messung eines Weltpunkts dar, für den ggf. auch Messunsicherheiten angegeben werden können. Oberflächen werden häufig durch Interpolation anhand von Punktwolken dargestellt. Auf Embedded-Systemen können solche Punktwolken die reale 3D-Welt aus Aufwandsgründen meist nur sehr dünn (sparse) abbilden, da die Verwaltung und Nutzung von großen Punktwolken hohe Anforderungen an die Rechnerplattform stellt.

[0011] Der hier beschriebene Ansatz kommt ohne Occupancy Grid und 3D Punktwolken aus. Stattdessen wird eine spezielle Fusion auf dem Pixelgitter vorgeschlagen, mit der hohe Datendichten auch auf kostengünstigen Embedded Systemen effizient bewältigt werden können.

[0012] Ein Verfahren zur Verarbeitung von Bilddaten umfasst die folgenden Schritte:
Bestimmen von Korrespondenzdaten als Ergebnis einer Korrespondenzbildung zwischen ersten Bilddaten, die auf von einer Kamera zu einem ersten Zeitpunkt erfassten Bilddaten basierende Daten repräsentieren, und vorangegangenen Bilddaten, die auf von der Kamera zu einem vorangegangenen Zeitpunkt erfassten Bilddaten basierende Daten repräsentieren;
Bestimmen von Tiefendaten auf Basis einer unter Verwendung der Korrespondenzdaten durchgeführten Eigenbewegungsschätzung; und
Ermitteln von Fusionsdaten unter Verwendung der Tiefendaten und übertragenen Fusionsdaten, wobei die Fusionsdaten eine auf den Bilddaten basierende Tiefeninformation umfassen.

[0013] Bei den zu verarbeitenden Bilddaten kann es sich um von zumindest einer Kamera bereitgestellte Bilddaten handeln. Unter einer Kamera kann eine Bilderfassungseinrichtung verstanden werden, die zu aufeinanderfolgenden Zeitpunkten ein Umfeld der Bilderfassungseinrichtung abbildende Bilddaten bereitstellen kann. Bei den ersten Bilddaten sowie den vorangegangenen Bilddaten kann es sich jeweils um Daten handeln, wie sie direkt von einer Kamera bereitgestellt werden oder um bereits aufbereitete Daten handeln. Der vorangegangene Zeitpunkt kann zeitlich vor dem ersten Zeitpunkt liegen.

Bei den vorangegangenen Bilddaten kann es sich somit um gespeicherte Bilddaten handeln. Die Korrespondenzbildung kann eine Korrespondenzanalyse umfassen. Die Korrespondenzdaten können eine Information über zueinander korrespondierende Merkmale, beispielsweise zueinander korrespondierende Bildpunkte, auch Pixel genannt, in den ersten Bilddaten und den vorangegangenen Bilddaten umfassen. Die Korrespondenzdaten können Informationen darüber umfassen, wie sich Positionen korrespondierender Bildpunkte in den ersten Bilddaten und den vorangegangenen Bilddaten zueinander verschoben haben. Unter Durchführung der Eigenbewegungsschätzung kann eine räumliche Struktur eines in den Bilddaten abgebildeten Objekts unter Verwendung korrespondierender Merkmale in den Bilddaten erkannt werden. Die dazu erforderliche Information über die korrespondierenden Merkmale kann dabei in den Korrespondenzdaten enthalten sein. Die Tiefendaten können ein Tiefenbild darstellen. Die Bestimmung der Tiefendaten kann in zwei Schritten erfolgen. Zunächst kann eine Bestimmung der Egomotion, also der Eigenbewegung, durchgeführt werden. Dazu werden nicht alle Korrespondenzen benötigt, sondern es genügt eine kleine Auslese. Als Nächstes wird die Bestimmung der Tiefendaten für alle Pixel durchgeführt, für die Korrespondenzdaten vorliegen. Für den Anwendungsfall einer Mono-Kamera kann zur Bestimmung der Tiefendaten eine Eigenbewegungsschätzung durchgeführt werden. Unter der Eigenbewegungsschätzung kann die Bestimmung der relativen Orientierung der beiden Kamerapositionen (6 Parameter, von denen 5 bestimmbar sind, also alles außer Maßstab) bei einer bewegten monokularen Kamera verstanden werden. Bei einer Stereo-Kamera sind bei einer starren Stereoanordnung die relative Orientierung und der Maßstab normalerweise bekannt. Falls die relative Orientierung trotzdem geschätzt werden soll, kann anstelle einer Eigenbewegungsschätzung eine Online-Kalibrierung durchgeführt werden. Sobald die relative Orientierung und der Kameraabstand bekannt sind, können die gefundenen Korrespondenzen durch Triangulation in Tiefen umgerechnet werden.

[0014] Durch die Fusion mit den übertragenen Fusionsdaten kann eine Qualität der Tiefendaten verbessert werden. Die Fusion kann beispielsweise eine Mittelwertbildung umfassen. Die Tiefeninformation kann somit ein Tiefenbild repräsentieren.

[0015] Gemäß einer Ausführungsform wird durch das Verfahren eine Fusion von Tiefenmessungen im Pixelgitter mittels optischem Fluss durchgeführt.

[0016] Dies ermöglicht eine hohe Reichweite der Bildfassung durch eine nahtlose Einbeziehung beziehungsweise Fusion von zusätzlichen Flussfeldern mit größerem zeitlichen Abstand. Der Abstand kann beispielsweise 2T, 4T oder 8T betragen, wobei T ei-

nen Zeitschritt darstellen kann. Ferner wird eine hohe Genauigkeit in allen Entfernungsbereichen durch Reduktion der Messunsicherheit ermöglicht. Wenn zwei Kameras eingesetzt werden, ist zudem ein nahtloser Übergang des Systemverhaltens von Stereo nach Mono durchführbar, beispielsweise wenn eine der beiden Kameras gestört oder nicht verfügbar ist. Bei videobasierten Fahrerassistenzsystemen ist nur ein geringer Entwicklungsaufwand erforderlich, da verschiedene Kamerasystem-Ausprägungen (z. B. Monokulares System, Stereo-System, Multi-Kamera-System) auf dem hier beschriebenen einheitlichen Konzept zur Basissignalverarbeitung aufbauen können.

[0017] Der beschriebene Ansatz basiert gemäß einer Ausführungsform auf einer Anreicherung von Information auf dem Pixelgitter. Dabei werden Güte und Genauigkeit der Messung erheblich verbessert (im Vergleich zu den Einzelmessungen). Die Anreicherung erfolgt unabhängig von der späteren Verdichtung und kann massiv parallel z. B. auf einer FPGA- oder ASIC-Plattform durchgeführt werden.

[0018] Dazu findet zu vordefinierten Zeiten oder zu variablen Zeiten eine Fusion statt, die ein altes und ein oder mehrere neue Bilder von Messergebnissen gewichtet kombiniert. Beispielsweise kann zu jedem Zeitschritt eine Fusion stattfinden. Es ist jedoch nicht unbedingt notwendig, nach jedem festen Zeitabschnitt T die Fusion durchzuführen. Der Zeitabstand darf durchaus auch unregelmäßig sein, z.B. 2T, wenn zwischendurch ein Flussfeld fehlt (z.B. weil die Recheneinheit mit etwas anderem beschäftigt war), oder sogar 1.2739T und beim nächsten Mal 0.9864T. Gemäß einer Ausführungsform können die Kameras auch nicht mit einer festen Zeitbasis laufen, sondern etwa nach einer festen gefahrenen Strecke getriggert oder ausgelesen werden. Somit kann die Fusion auch zu von einer Zeitbasis unabhängigen Zeitpunkten durchgeführt werden. Bei den nachfolgend beschriebenen Ausführungsbeispielen wird lediglich der Einfachheit halber überwiegend ein fester Zeitabstand T angenommen.

[0019] Unterstützt werden zum einen optische Flüsse zu einem oder mehreren Vorgängerbildern sowie optional Stereo-Disparitäten (ebenfalls vorzugsweise dicht), und zum anderen sowohl monokulare Kamerasysteme als auch Stereo- oder sogar Multi-Kamerasysteme. Die optischen Flüsse können dicht, also möglichst für jedes Pixel vorhanden sein. Vorteilhafterweise können die optischen Flüsse jedoch lediglich regionenweise dicht anstelle von ganz dicht sein. Es ist z.B. nicht notwendig und auch nicht gut möglich, den optischen Fluss auf dem blauen Himmel zu bestimmen. Nur wenn ausreichend Textur vorhanden ist, kann der optische Fluss sinnvoll bestimmt werden. Wo keine ausreichende Textur vorhanden ist, kann auf den Fluss verzichtet werden. Die Stereo-

Disparitäten können ebenfalls ganz oder regionenweise dicht sein.

[0020] Das genannte Verfahren umfasst gemäß einer Ausführungsform einen Schritt des Bestimmens der übertragenen Fusionsdaten als Ergebnis einer Übertragung von in einem vorangegangenen Schritt des Ermitteln ermittelten Fusionsdaten und den Korrespondenzdaten. Durch die Übertragung kann eine seit dem vorangegangenen Schritt erfolgte Bewegung berücksichtigt werden, sodass die übertragenen Fusionsdaten im Schritt des Ermitteln direkt mit den Tiefendaten fusioniert werden können.

[0021] Im Schritt des Bestimmens der Korrespondenzdaten können Flussvektoren als die Korrespondenzdaten bestimmt werden. Die Flussvektoren können dabei Bildpunkten der ersten Bilddaten und zusätzlich oder alternativ Bildpunkten der vorangegangenen Bilddaten zugeordnet sein. Dabei weisen die Korrespondenzen in der Regel Sub-Pixel-Genauigkeit auf, zeigen also nicht exakt von Bildpunktmittle zu Bildpunktmittle, sondern liegen mindestens in einem der beiden Bilder nichtmittig auf dem Raster. Generell sind Ganz-Pixel-Bewegungen ebenso zulässig. Somit kann es auch Korrespondenzen geben, die exakt von Bildpunktmittle auf Bildpunktmittle zeigen. Vorteilhafterweise kann für die Bestimmung von Flussvektoren auf bekannte Verfahren zurückgegriffen werden.

[0022] Im Schritt des Bestimmens von Tiefendaten können Tiefenwerte als die Tiefendaten bestimmt werden. Die Tiefenwerte können Bildpunkten der ersten Bilddaten und zusätzlich oder alternativ Bildpunkten der vorangegangenen Bilddaten zugeordnet sein. Tiefenwerte stellen eine wichtige Information dar, die beispielsweise für Anwendungen im Bereich von Fahrerassistenzsystemen oder mobilen Robotern eingesetzt werden können.

[0023] Beispielsweise können die Fusionsdaten im Schritt des Ermitteln unter Verwendung einer Mittelwertbildung zwischen den Tiefendaten und den übertragenen Fusionsdaten ermittelt werden. Eine solche Mittelwertbildung kann mit einfachen Mitteln durchgeführt werden.

[0024] Für den Fall, dass keine übertragenen Fusionsdaten vorliegen, können die im Schritt des Ermitteln ermittelten Fusionsdaten den Tiefendaten entsprechen. Dies bietet sich beispielsweise bei einer Initialisierung des Systems an.

[0025] Wenn keine Tiefendaten vorliegen, können die Fusionsdaten den übertragenen Fusionsdaten entsprechen. Dies bietet sich beispielsweise dann an, wenn keine neuen Bilddaten mehr bereitgestellt werden.

[0026] Im Schritt des Ermitteln der Fusionsdaten kann eine Konsistenz zwischen den Tiefendaten und den übertragenen Fusionsdaten überprüft werden. Abhängig von einem Ergebnis der Überprüfung der Konsistenz kann die Fusion zwischen den Tiefendaten und den übertragenen Fusionsdaten durchgeführt oder ausgesetzt werden. Auf diese Weise kann eine Verfälschung vermieden werden, wenn eine Inkonsistenz zwischen den zu fusionierenden Daten vorliegt.

[0027] Das Verfahren kann erweitert werden, indem weitere Korrespondenzdaten, beispielsweise unter Verwendung der ersten Bilddaten und zu weiteren vorangegangenen Zeitpunkten erfassten Bilddaten, bestimmt werden und diese weiteren Korrespondenzdaten zur Bestimmung weiterer Tiefendaten verwendet werden, die dann mit den erstgenannten Tiefendaten und den übertragenen Fusionsdaten fusioniert werden können, um die Fusionsdaten zu ermitteln. Auf diese Weise kann die Genauigkeit und Qualität der Fusionsdaten weiter erhöht werden.

[0028] Somit kann das Verfahren zumindest einen Schritt des Bestimmens von weiteren Korrespondenzdaten als Ergebnis einer Korrespondenzbildung zwischen den ersten Bilddaten und weiteren vorangegangenen Bilddaten, die auf von der Kamera zu einem weiteren vorangegangenen Zeitpunkt erfassten Bilddaten basierende Daten repräsentieren umfassen. Entsprechend kann das Verfahren zumindest einen Schritt des Bestimmens von weiteren Tiefendaten auf Basis einer unter Verwendung der weiteren Korrespondenzdaten durchgeführten Eigenbewegungsschätzung umfassen. Im Schritt des Ermitteln können die Fusionsdaten ferner unter Verwendung der weiteren Tiefendaten ermittelt werden.

[0029] Gemäß einer Ausführungsform kann das Verfahren einen Schritt des Bestimmens von Stereo-Korrespondenzdaten als Ergebnis einer Korrespondenzbildung zwischen den ersten Bilddaten und zweiten Bilddaten umfassen. Die zweiten Bilddaten können auf von einer weiteren Kamera zu dem ersten Zeitpunkt erfassten Bilddaten basierende Daten repräsentieren. In einem Schritt des Bestimmens können Stereo-Tiefendaten, die Tiefendaten und/oder einen Maßstab repräsentieren, auf Basis einer unter Verwendung der Stereo-Korrespondenzdaten durchgeführten Triangulation bestimmt werden. Durch den Maßstab kann eine zurückgelegte Strecke abgebildet werden. Im Schritt des Ermitteln können die Fusionsdaten ferner unter Verwendung der Stereo-Tiefendaten und/oder des Maßstabs ermittelt werden. Auf diese Weise können die Bilddaten einer weiteren Kamera, beispielsweise eines Stereo-Kamerasystems in die Ermittlung der Fusionsdaten einfließen. Unter dem Maßstab wird gemäß einer Ausführungsform ein Umrechnungsfaktor verstanden, um die Tiefendaten z.B. von einer Einheit 1/s (TTC) in

einen metrischen Abstand umrechnen zu können. Ein monokulares System kann diesen Maßstab nicht selbst bestimmen. Beispiel: Bei einer aus einer fahrenden Lokomotive aufgenommenen Videosequenz kann anhand der Bewegung im Bild nicht festgestellt werden, ob es sich um eine Modelleisenbahn- oder eine echte Lokomotive handelt. Erst wenn von außen eine Maßstabsinformation bereitgestellt wird, z.B. gefahrener Weg in Metern zwischen zwei Bildaufnahmen oder Höhe der Kamera über den Schienen, wird die Angabe metrischer Tiefendaten möglich. Der Maßstab kann auch von einer Stereokamera stammen, deren Abstand bekannt ist (Basisweite).

[0030] Beispielsweise kann das Verfahren einen Schritt des Erfassens der Bilddaten unter Verwendung zumindest einer Kamera umfassen. Zusätzlich oder alternativ kann das Verfahren einen Schritt des Verwendens der Fusionsdaten in einem Fahrerassistenzsystem für ein Fahrzeug oder in einer Steuereinrichtung für einen mobilen Roboter. Auf diese Weise kann das Verfahren in ein Gesamtsystem eingebunden werden.

[0031] Dieses Verfahren kann beispielsweise in Software oder Hardware oder in einer Mischform aus Software und Hardware beispielsweise in einem Steuergerät implementiert sein.

[0032] Der hier vorgestellte Ansatz schafft ferner eine Vorrichtung, die ausgebildet ist, um die Schritte einer Variante eines hier vorgestellten Verfahrens in entsprechenden Einrichtungen durchzuführen, anzusteuern bzw. umzusetzen. Auch durch diese Ausführungsvariante der Erfindung in Form einer Vorrichtung kann die der Erfindung zugrunde liegende Aufgabe schnell und effizient gelöst werden.

[0033] Hierzu kann die Vorrichtung zumindest eine Recheneinheit zum Verarbeiten von Signalen oder Daten, zumindest eine Speichereinheit zum Speichern von Signalen oder Daten, zumindest eine Schnittstelle zu einem Sensor oder einem Aktor zum Einlesen von Sensorsignalen von dem Sensor oder zum Ausgeben von Daten- oder Steuersignalen an den Aktor und/oder zumindest eine Kommunikationsschnittstelle zum Einlesen oder Ausgeben von Daten aufweisen, die in ein Kommunikationsprotokoll eingebettet sind. Die Recheneinheit kann beispielsweise ein Signalprozessor, ein Mikrocontroller oder dergleichen sein, wobei die Speichereinheit ein Flash-Speicher, ein EPROM oder eine magnetische Speichereinheit sein kann. Die Kommunikationsschnittstelle kann ausgebildet sein, um Daten drahtlos und/oder leitungsgebunden einzulesen oder auszugeben, wobei eine Kommunikationsschnittstelle, die leitungsgebundene Daten einlesen oder ausgeben kann, diese Daten beispielsweise elektrisch oder optisch aus einer entsprechenden Datenübertragungsleitung einle-

sen oder in eine entsprechende Datenübertragungsleitung ausgeben kann.

[0034] Eine entsprechende Vorrichtung zur Verarbeitung von Bilddaten weist die folgenden Merkmale auf:

eine Bestimmungseinrichtung, die ausgebildet ist, um Korrespondenzdaten als Ergebnis einer Korrespondenzbildung zwischen ersten Bilddaten, die auf von einer Kamera zu einem ersten Zeitpunkt erfassten Bilddaten basierende Daten repräsentieren, und vorangegangenen Bilddaten, die auf von der Kamera zu einem vorangegangenen Zeitpunkt erfassten Bilddaten basierende Daten repräsentieren, zu bestimmen; eine Bestimmungseinrichtung, die ausgebildet ist, um Tiefendaten auf Basis einer unter Verwendung der Korrespondenzdaten durchgeführten Eigenbewegungsschätzung zu bestimmen; und eine Ermittlungseinrichtung, die ausgebildet ist, um unter Verwendung der Tiefendaten und übertragenen Fusionsdaten zu ermitteln, wobei die Fusionsdaten eine auf den Bilddaten basierende Tiefeninformation umfassen.

[0035] Ein Fahrzeug weist zumindest eine Kamera, eine genannte Vorrichtung zum Verarbeiten der Bilddaten der zumindest einen Kamera und ein Fahrerassistenzsystem auf. Das Fahrerassistenzsystem ist ausgebildet, um unter Verwendung von der Vorrichtung bereitgestellten Fusionsdaten eine Fahrerassistenzfunktion auszuführen.

[0036] Für den Anwendungsfall eines mobilen Roboters kann der mobile Roboter ebenfalls eine Kamera und die genannte Vorrichtung umfassen. Die Fusionsdaten können in diesem Fall beispielsweise von einer Einrichtung zum Steuern einer Funktion oder Bewegung des Roboters verwendet werden.

[0037] Von Vorteil ist auch ein Computerprogrammprodukt oder Computerprogramm mit Programmcode, der auf einem maschinenlesbaren Träger oder Speichermedium wie einem Halbleiterspeicher, einem Festplattenspeicher oder einem optischen Speicher gespeichert sein kann und zur Durchführung, Umsetzung und/oder Ansteuerung der Schritte des Verfahrens nach einer der vorstehend beschriebenen Ausführungsformen verwendet wird, insbesondere wenn das Programmprodukt oder Programm auf einem Computer oder einer Vorrichtung ausgeführt wird.

[0038] Ausführungsbeispiele des hier vorgestellten Ansatzes sind in den Zeichnungen dargestellt und in der nachfolgenden Beschreibung näher erläutert. Es zeigt:

[0039] Fig. 1 eine schematische Darstellung eines Fahrzeugs mit einer Vorrichtung zur Verarbeitung von Bilddaten gemäß einem Ausführungsbeispiel;

[0040] Fig. 2 ein Ablaufdiagramm eines Verfahrens zur Verarbeitung von Bilddaten gemäß einem Ausführungsbeispiel;

[0041] Fig. 3 ein Blockschaltbild einer Vorrichtung zur Verarbeitung von Bilddaten gemäß einem Ausführungsbeispiel;

[0042] Fig. 4 ein Blockschaltbild einer Vorrichtung zur Verarbeitung von Bilddaten gemäß einem Ausführungsbeispiel; und

[0043] Fig. 5 ein Blockschaltbild einer Vorrichtung zur Verarbeitung von Bilddaten gemäß einem Ausführungsbeispiel.

[0044] In der nachfolgenden Beschreibung günstiger Ausführungsbeispiele der vorliegenden Erfindung werden für die in den verschiedenen Figuren dargestellten und ähnlich wirkenden Elemente gleiche oder ähnliche Bezugszeichen verwendet, wobei auf eine wiederholte Beschreibung dieser Elemente verzichtet wird.

[0045] Fig. 1 zeigt ein Fahrzeug **100** mit einer Vorrichtung **102** zur Verarbeitung von Bilddaten **104** einer Kamera **106** des Fahrzeugs **100** gemäß einem Ausführungsbeispiel. Die Vorrichtung **102** ist ausgebildet, um die Bilddaten **104** der Kamera **106** über eine Schnittstelle zu empfangen und unter Verwendung der Bilddaten **104** Fusionsdaten **108** zu erzeugen und bereitzustellen. Die Fusionsdaten **108** repräsentieren gemäß einem Ausführungsbeispiel ein Tiefenbild, das unter Verwendung von zu unterschiedlichen Zeitpunkten von der Kamera **106** erfassten Bilddaten **104** ermittelt wurde.

[0046] Die Vorrichtung **102** ist ausgebildet, um die Fusionsdaten **108** über eine weitere Schnittstelle an eine weitere Einrichtung **110** des Fahrzeugs **100**, beispielsweise ein Fahrerassistenzsystem **110** des Fahrzeugs **100** bereitzustellen.

[0047] Beispielsweise können die Fusionsdaten **108** von dem Fahrerassistenzsystem **110** zur Lösung einer Fahraufgabe des Fahrzeugs **100** verwendet werden.

[0048] Gemäß einem alternativen Ausführungsbeispiel zeigt Fig. 1 einen mobilen Roboter **100**. In diesem Fall können die von der Vorrichtung **102** bereitgestellten Fusionsdaten **108** von einer Steuereinrichtung **110** des Roboters **100** zum Ausführen einer Funktion des mobilen Roboters **100** verwendet werden. Bei dem mobilen Roboter **100** kann es sich auch um ein hoch automatisiertes Fahrzeug handeln.

[0049] Wenn mehr als eine Kamera **106** vorhanden ist, kann die Vorrichtung **102** ausgebildet sein, um die

Bilddaten **104** der mehreren Kameras **106** zu verwenden, um die Fusionsdaten **108** zu ermitteln.

[0050] Fig. 2 zeigt ein Ablaufdiagramm eines Verfahrens zur Verarbeitung von Bilddaten gemäß einem Ausführungsbeispiel. Das Verfahren kann beispielsweise unter Verwendung der anhand von Fig. 1 beschriebenen Vorrichtung zur Verarbeitung von Bilddaten ausgeführt werden.

[0051] In einem Schritt **220** werden Korrespondenzdaten als Ergebnis einer Korrespondenzbildung zwischen ersten Bilddaten und vorangegangenen Bilddaten bestimmt. Die ersten Bilddaten basieren dabei auf von einer Kamera zu einem ersten Zeitpunkt erfassten Bilddaten. Die vorangegangenen Bilddaten basieren auf von der Kamera zu einem vorangegangenen Zeitpunkt erfassten Bilddaten. Gemäß einem Ausführungsbeispiel werden dabei Flussvektoren als die Korrespondenzdaten bestimmt. Die Flussvektoren können dabei den Bildpunkten der ersten Bilddaten zugeordnet sein. Zusätzlich oder alternativ können die Flussvektoren Bildpunkten der vorangegangenen Bilddaten zugeordnet sein.

[0052] In einem Schritt **222** werden basierend auf den Korrespondenzdaten Tiefendaten bestimmt. Dazu wird unter Verwendung der Korrespondenzdaten eine Eigenbewegungsschätzung durchgeführt. Dabei können Tiefenwerte als die Tiefendaten bestimmt werden. Die so bestimmten Tiefenwerte können dabei den Bildpunkten der ersten Bilddaten und/oder den Bildpunkten der vorangegangenen Bilddaten zugeordnet sein.

[0053] In einem Schritt **224** werden Fusionsdaten unter Verwendung der Tiefendaten und übertragenen Fusionsdaten ermittelt. Die Fusionsdaten werden dabei durch eine Fusionierung der Tiefendaten und übertragenen Fusionsdaten ermittelt.

[0054] In einem optionalen Schritt **226** werden die übertragenen Fusionsdaten aus den in einem vorangegangenen Schritt des Ermittelns ermittelten Fusionsdaten und den Korrespondenzdaten bestimmt.

[0055] Die Schritte **220**, **222**, **224**, **226** können fortlaufend wiederholt ausgeführt werden. Auf diese Weise können beispielsweise die ersten Bilddaten bei einem nachfolgenden Durchgang den vorangegangenen Bilddaten entsprechen.

[0056] Gemäß einem Ausführungsbeispiel wird im Schritt **220** für jeden Bildpunkt oder für eine Anzahl der Bildpunkte der ersten Bilddaten und/oder der vorangegangenen Bilddaten ein Flussvektor bestimmt. Es ist dabei nicht notwendig, für jeden Bildpunkt einen Flussvektor angeben zu können. Es dürfen also „Lücken“ vorhanden sein. Somit umfassen die Korrespondenzdaten für jeden oder zumindest für eine An-

zahl von Bildpunkten der ersten Bilddaten und/oder der vorangegangenen Bilddaten je einen Flussvektor.

[0057] Gemäß einem Ausführungsbeispiel wird im Schritt **222** für jeden Bildpunkt der ersten Bilddaten und/oder der vorangegangenen Bilddaten ein Tiefenwert bestimmt. Somit umfassen die Tiefendaten für jeden Bildpunkt der ersten Bilddaten und/oder der vorangegangenen Bilddaten je einen Tiefenwert.

[0058] Gemäß einem Ausführungsbeispiel wird im Schritt **224** für jeden Bildpunkt der ersten Bilddaten und/oder der vorangegangenen Bilddaten ein Tiefenwert bestimmt. Somit umfassen die Fusionsdaten für jeden Bildpunkt der ersten Bilddaten und/oder der vorangegangenen Bilddaten je einen Tiefenwert.

[0059] Gemäß einem Ausführungsbeispiel werden im Schritt **226** die Koordinaten für jeden Bildpunkt eingelesener Fusionsdaten auf die Koordinaten nachfolgender Bilddaten, beispielsweise der ersten Bilddaten übertragen.

[0060] Fig. 3 zeigt ein Blockschaltbild einer Vorrichtung **102** zur Verarbeitung von Bilddaten gemäß einem Ausführungsbeispiel. Dabei kann es sich um ein Ausführungsbeispiel der anhand von Fig. 1 beschriebenen Vorrichtung handeln.

[0061] Gemäß diesem Ausführungsbeispiel ist eine erste Kamera **106** zum Bereitstellen von Bilddaten **104** und eine zweite Kamera **306** zum Bereitstellen von Bilddaten **304** vorgesehen. Die Vorrichtung **102** ist gemäß diesem Ausführungsbeispiel ausgebildet, um die Bilddaten **104**, **304** der Kameras **106**, **306** über eine geeignete Schnittstelle zu empfangen und zu verarbeiten.

[0062] Die Vorrichtung **102** weist eine Mehrzahl von Speichereinrichtungen **330**, **332**, **334**, **336**, **338** auf, in denen zu unterschiedlichen Zeitpunkten erfasste Bilddaten **104** der Kamera **106** gespeichert sind. Die Anzahl der Speichereinrichtungen **330**, **332**, **334**, **336**, **338** ist hier nur beispielhaft gewählt. Zwischen zeitlich aufeinanderfolgenden Zeitpunkten liegt gemäß diesem Ausführungsbeispiel beispielhaft die Zeitdauer T . Alternativ kann zwischen zeitlich aufeinanderfolgenden Zeitpunkten eine andere vorgegebene oder variable Zeitdauer liegen.

[0063] In der Speichereinrichtung **330** sind von der ersten Kamera **106** aktuell oder zuletzt bereitgestellte Bilddaten **104** als erste Bilddaten **340** gespeichert. Die ersten Bilddaten **340** werden gemäß diesem Ausführungsbeispiel auch mit $I_L(t)$ bezeichnet.

[0064] In der Speichereinrichtung **332** sind von der Kamera **106** zu einem ersten vorangegangenen Zeitpunkt bereitgestellte Bilddaten **104** als erste vorange-

gangenen Bilddaten **342** gespeichert. Die ersten vorangegangenen Bilddaten **342** werden gemäß diesem Ausführungsbeispiel auch mit $I_L(t - T)$ bezeichnet.

[0065] In der Speichereinrichtung **334** sind von der Kamera **106** zu einem zweiten vorangegangenen Zeitpunkt bereitgestellte Bilddaten **104** als zweite vorangegangene Bilddaten **344** gespeichert. Die zweiten vorangegangenen Bilddaten **344** werden gemäß diesem Ausführungsbeispiel auch mit $I_L(t - 2T)$ bezeichnet.

[0066] In der Speichereinrichtung **336** sind von der Kamera **106** zu einem dritten vorangegangenen Zeitpunkt bereitgestellte Bilddaten **104** als dritte vorangegangene Bilddaten **346** gespeichert. Die dritten vorangegangenen Bilddaten **346** werden gemäß diesem Ausführungsbeispiel auch mit $I_L(t - 3T)$ bezeichnet.

[0067] In der Speichereinrichtung **338** sind von der Kamera **106** zu einem vierten vorangegangenen Zeitpunkt bereitgestellte Bilddaten **104** als vierte vorangegangene Bilddaten gespeichert. Die vierten vorangegangenen Bilddaten werden gemäß diesem Ausführungsbeispiel auch mit $I_L(t - 4T)$ bezeichnet.

[0068] Nach Ablauf eines aktuellen Zeitschritts werden die ersten Bilddaten **340** an die Speichereinrichtung **332**, die ersten vorangegangenen Bilddaten **342** an die Speichereinrichtung **334**, die zweiten vorangegangenen Bilddaten **344** an die Speichereinrichtung **336** und die dritten vorangegangenen Bilddaten **346** an die Speichereinrichtung **338** übergeben.

[0069] Ferner weist die Vorrichtung **102** eine weitere Speichereinrichtung **350** auf, die ausgebildet ist, um von der zweiten Kamera **306** aktuell oder zuletzt bereitgestellte Bilddaten **304** als zweite Bilddaten zu speichern.

[0070] Gemäß einem Ausführungsbeispiel handelt es sich bei den Kameras **106**, **306** um eine linke Kamera **106** und eine rechte Kamera **306**. Gezeigt sind dabei verschieden alte Bilddaten **340**, **342**, **344**, **346** sowie Verbindungen, zwischen denen Berechnungen zur Korrespondenzbildung stattfinden.

[0071] In Fig. 3 ist dabei ein allgemeineres Ausführungsbeispiel gezeigt. Später werden anhand des in Fig. 4 gezeigten einfacheren Beispiels wichtige Details erläutert und zuletzt wird das allgemeinere Beispiel anhand von Fig. 5 weiter detailliert.

[0072] Das in Fig. 3 dargestellte System enthält zwei Kameras **106**, **306**. Hierbei kann es sich beispielsweise um ein Stereo-Fahrerassistenzsystem mit linker Kamera **106** und rechter Kamera **306** handeln, die in Vorwärts-Richtung eines Fahrzeugs blicken. Die Kameras **106**, **306** liefern je ein aktuelles Bild **104**, **304**, die auch mit $I_L(t)$ und $I_R(t)$ bezeichnet werden, wo-

bei t für den aktuellen Zeitpunkt (neuestes Bild) steht. Die Bilder $I_L(t - n \cdot T)$ mit $n = 1, 2, 3, \dots$ stammen von früheren Zeitpunkten und wurden in einem Speicher **330**, **332**, **334**, **336**, **338** aufbewahrt. Statt Originalbilder als Bilddaten **340**, **342**, **344**, **346** aufzubewahren, können auch weiterverarbeitete Bilddaten (Feature-Bilder) als Bilddaten **340**, **342**, **344**, **346** aufbewahrt werden.

[0073] Von der rechten Kamera **306** werden in diesem Ausführungsbeispiel keine alten Bilder aufbewahrt.

[0074] Die Pfeile (mit Pfeilspitzen) im Blockschaltbild stehen für Verbindungen, auf denen Daten **340**, **342**, **344**, **346** transportiert oder im Speicher bereitgehalten werden, z. B. in einem Ringpufferspeicher.

[0075] Des Weiteren sind in **Fig. 3** vier Verbindungen eingezeichnet, die jeweils auf einer Seite mit einem Kringel enden. Auf jeder dieser „Verbindungen“ wird eine Korrespondenzanalyse durchgeführt. Auf den drei Verbindungen zwischen verschiedenen alten Bildern der linken Kamera **106** handelt es sich bei der Korrespondenzanalyse jeweils um einen optischen Fluss Algorithmus (optical flow). Das Ergebnis sind jeweils Flussvektoren pro Pixel.

[0076] Der Kringel zeigt jeweils an, in Bezug zu welchem Bild die Ergebnisse vorliegen sollen. Hier befinden sich alle Kringel am aktuellen linken Bild, das durch die ersten Bilddaten **340** repräsentiert wird. Dies bedeutet, dass für jedes Pixel in diesem Bild ermittelt wird, woher es im jeweiligen Vorgängerbild stammt. Die Suche nach der Korrespondenz erfolgt also in Rückwärtsrichtung (Rückwärtsfluss).

[0077] Die vierte Verbindung mit Kringel besteht zwischen dem aktuellen linken Kamerabild, das durch die ersten Bilddaten **340** repräsentiert wird, und dem rechten Kamerabild, das durch die zweiten Bilddaten **350** repräsentiert wird, wobei auch hier das Ergebnis in Bezug zum aktuellen linken Bild gesucht wird. Auf dieser vierten Verbindung werden sogenannte Stereo-Korrespondenzen zwischen einem Kamerapaar, hier bestehend aus den Kameras **106**, **306**, gebildet. Das Ergebnis ist hier ein Disparitätsvektor pro Pixel. Bei einem kalibrierten Stereosystem ist die zweite Komponente des Disparitätsvektors Null und kann ignoriert werden, es verbleibt also ein eindimensionaler Disparitätswert pro Pixel.

[0078] **Fig. 4** zeigt ein Blockschaltbild einer Vorrichtung **102** zur Verarbeitung von Bilddaten **104** gemäß einem Ausführungsbeispiel. Dabei kann es sich um ein Ausführungsbeispiel der anhand von **Fig. 1** beschriebenen Vorrichtung handeln.

[0079] Wie bereits anhand von **Fig. 3** beschrieben, umfasst die Vorrichtung **102** eine Speichereinrich-

tung **330** zum Speichern von ersten Bilddaten **340**, die von der hier alleinigen Kamera **106** aktuell oder zuletzt bereitgestellte Bilddaten **104** darstellen. Die ersten Bilddaten **340** werden wiederum auch mit $I_L(t)$ bezeichnet. Ferner umfasst die Vorrichtung **102** die bereits anhand von **Fig. 3** beschriebene Speichereinrichtung **332** zum Speichern von vorangegangenen Bilddaten **342**, die von der Kamera **106** zu einem vorangegangenen Zeitpunkt bereitgestellte Bilddaten **104** repräsentieren. Die vorangegangenen Bilddaten **342** werden wiederum auch mit $I_L(t - T)$ bezeichnet.

[0080] Die Vorrichtung **102** umfasst gemäß diesem Ausführungsbeispiel ferner zwei Bestimmungseinrichtungen **460**, **462**, **464** und eine Ermittlungseinrichtung **466**. Die Bestimmungseinrichtung **460** ist ausgebildet, um eine Korrespondenz zwischen den ersten Bilddaten **340** und den vorangegangenen Bilddaten **342** zu bilden und durch eine solche Korrespondenzbildung bestimmte Korrespondenzdaten **470** bereitzustellen. Die Bestimmungseinrichtung **462** ist ausgebildet, um unter Verwendung der Korrespondenzdaten **470** eine Eigenbewegungsschätzung durchzuführen und Tiefendaten **472** auf Basis der Eigenbewegungsschätzung bereitzustellen. Die Ermittlungseinrichtung **466** ist ausgebildet ist, um unter Verwendung der Tiefendaten **472** und übertragener Fusionsdaten **474** Fusionsdaten **476** zu ermitteln und bereitzustellen. Gemäß diesem Ausführungsbeispiel werden die übertragenen Fusionsdaten **474** von der Bestimmungseinrichtung **464** bereitgestellt. Die Bestimmungseinrichtung **464** ist ausgebildet, um die übertragenen Fusionsdaten **474** unter Verwendung der von der Ermittlungseinrichtung **466** bereitgestellten Fusionsdaten **474** und den von der Bestimmungseinrichtung **460** bereitgestellten Korrespondenzdaten **470** zu bestimmen.

[0081] In **Fig. 4** ist somit ein vereinfachtes Blockschaltbild der Vorrichtung **102** gezeigt, bei dem nur die bereits anhand von **Fig. 3** beschriebene linke Kamera **106** betrachtet wird. Durch die Korrespondenzdaten **470** repräsentierte Korrespondenzen (Corresp) werden zwischen dem durch die ersten Bilddaten **340** repräsentierte neuen Kamerabild und dem durch die vorangegangenen Bilddaten **342** repräsentierte vorausgegangenem Kamerabild berechnet. Basierend auf diesen Korrespondenzen wird unter Verwendung der Bestimmungseinrichtung **462** die Eigenbewegung (Egomotion) und Tiefe (Depth) berechnet (EM&D). Das als Tiefendaten **472** vorliegende Ergebnis hiervon wird fusioniert mit früheren Ergebnissen, die zunächst ins Pixelgitter des aktuellen Bildes gewarpt werden. Dazu wird die Bestimmungseinrichtung **464** verwendet.

[0082] Zur Erläuterung weiterer Details wird zunächst ein anhand von **Fig. 4** gezeigtes einfacheres Ausführungsbeispiel betrachtet, das auch als Aus-

schnitt aus dem anhand von **Fig. 3** gezeigten Beispiel angesehen werden kann.

[0083] Mit der Bestimmungseinrichtung **460**, auch als „Corresp“ bezeichnet, ist hier die Korrespondenzbildung bezeichnet, in die über den Doppelpfeil die (ggf. vorverarbeiteten) Bilddaten **340** $h(t)$ und **342** $h(t - T)$ eingehen.

[0084] Anschließend wird in der Bestimmungseinrichtung **462**, auch als „EM&D“ (Egomotion & Depth) bezeichnet, eine Eigenbewegungsschätzung durchgeführt und zu jedem Pixel eine Tiefe bestimmt. Dazu wird gemäß einem Ausführungsbeispiel ein sogenanntes „Structure from Motion“ Verfahren verwendet.

[0085] Bei einer monokularen Kamera **106** können von den sechs Parametern der Eigenbewegung (3 Rotationswinkel und ein 3-dimensionaler Translationsvektor) bekanntlich nur fünf Parameter bestimmt werden (3 Rotationswinkel und die Richtung der Translation), der Maßstab (zurückgelegte Strecke) bleibt also zunächst unbekannt oder kann optional aus einer anderen Quelle z. B. von einem GPS-Empfänger oder von Radsensoren (Odometrie), bezogen werden.

[0086] Es ist aber auch möglich und unproblematisch, die weitere Verarbeitung trotz zunächst unbekanntem Maßstab durchzuführen. Dazu werden die als „Tiefe“ bezeichneten Zwischenergebnisse vorzugsweise als TTC pro Pixel oder Kehrwert $1/TTC$ pro Pixel ausgedrückt, wobei TTC (Time To Collision) für eine Zeit steht, beispielsweise für die Zeit, die bei unveränderter Relativbewegung zwischen Kamera und Szene vergeht, bis der beobachtete Welt-punkt die gedachte Ebene durchschreitet, die durch das Projektionszentrum der Kamera **106** verläuft und senkrecht zu ihrer Hauptachse (optische Achse) steht.

[0087] Aus numerischen Gründen wird statt TTC auch gerne der Kehrwert $1/TTC$ verwendet. Die Einheit der TCC kann beispielsweise in Sekunden oder durch die Anzahl der Kamerabilder (Frames) angegeben sein.

[0088] Nachfolgend wird verallgemeinernd von „Tiefe“ gesprochen. Damit kann eine räumliche Distanz oder eine Zeitdifferenz oder eine beliebige nichtlineare Funktion (z. B. Kehrwert) von einer der beiden Größen gemeint sein.

[0089] Bis hierher sind dem Fachmann die verwendeten Begriffe und Verfahren bekannt. Nachfolgend wird ein Kerngedanke des hier beschriebenen Ansatzes näher beschrieben. Dabei geht es im Wesentlichen um die Blöcke in dem in **Fig. 4** gezeigten Block-

schaltbild, die die Bestimmungseinrichtung **464** und die Ermittlungseinrichtung **466** betreffen.

[0090] Zunächst wird dazu eine erste Variante einer Fusion im Pixelgitter beschrieben. Gemäß diesem Ausführungsbeispiel ist die Ermittlungseinrichtung **466** ausgebildet, um einen auch als „Fusion“ bezeichneten Fusionsalgorithmus auszuführen.

[0091] An den Fusionsalgorithmus werden also durch die Tiefendaten **472** repräsentierte „Tiefenbilder“ mit dem beschriebenen Inhalt pro Pixel übergeben.

[0092] Optional wird pro Pixel auch eine Information über die Messunsicherheit pro Pixel mitgeliefert. Diese kann als Bild vorliegen oder als Rechenvorschrift. Die Rechenvorschrift benötigt weniger Speicherplatz und weniger Übertragungsbandbreite. Die Messunsicherheit kann beispielsweise in Form einer Standardabweichung bezüglich des jeweiligen Tiefenwerts angegeben sein.

[0093] Die hier beschriebene erste Variante der in der Ermittlungseinrichtung **466** durchgeführten Fusion im Pixelgitter ist dadurch gekennzeichnet, dass die durch die Tiefendaten **472** repräsentierten Tiefen vorab in der Bestimmungseinrichtung **460** so berechnet werden, dass das durch die Tiefendaten **472** repräsentierte Tiefenbild in den Koordinaten des älteren der beiden Bilder vorliegt, die durch die Bilddaten **340**, **342** repräsentiert sind.

[0094] Das ist so zu verstehen, dass der Tiefenwert eines Pixels dem entsprechenden Pixel an gleicher Position im älteren der beiden Originalbilder, die durch die Bilddaten **340**, **342** repräsentiert sind, zuzuordnen ist.

[0095] In dem vereinfachten Blockschaltbild gemäß **Fig. 4** gehen keine weiteren Ergebnisse ein, sondern es findet nur eine Fusion zwischen dem gerade ermittelten Tiefenbild (ggf. mit Messunsicherheiten) **472** und dem durch die übertragenen Fusionsdaten **474** repräsentierten Inhalt des Fusionspeichers, der durch die Bestimmungseinrichtung **464** dargestellt ist, statt. Bei diesem Inhalt **474** handelt es sich um das Ergebnis der Fusion für die vorausgegangenen Zeitschritte, welches nun zunächst im Block **464** von alten Pixelkoordinaten auf die aktuellen Pixelkoordinaten übertragen (gewarpt) wird, denn die Bildszene hat sich in der Zwischenzeit verändert, und zwar im Wesentlichen pixelweise verschoben, wobei diese pixelweise Verschiebung durch das optische Flussfeld bekannt ist, das durch die Korrespondenzdaten **470** repräsentiert wird. Wie oben bereits angemerkt muss es sich dabei nicht um ganzzahlige pixelweise Verschiebungen handeln, sondern es treten im Allgemeinen auch subpixelweise Verschiebun-

gen auf, sodass dann optional eine Interpolation beim Warming angewendet wird.

[0096] Dazu werden die Korrespondenzdaten **470** als das bereits vorliegende Ergebnis der in der Bestimmungseinrichtung **460** durchgeführten Korrespondenzbildung verwendet.

[0097] Den in der Bestimmungseinrichtung **464** durchgeführten Schritt des Warpings, auch Schritt des Übertragens genannt, kann man sich vorstellen wie ein Bild auf einem Gummituch, das verzerrt wird, indem man jeden Punkt des Gummituchs mit dem zugehörigen Flussvektor verschiebt, der durch die Korrespondenzdaten **470** bereitgestellt wird.

[0098] Vorteilhafterweise wird beim Warming auch eine Interpolation durchgeführt. Diese ist damit begründet, dass die pixelweisen Verschiebungen im Allgemeinen nicht ganzzahlig sind, sondern auch Nachkommaanteile aufweisen. Die Interpolation ist also ratsam, wenn mit der in der Ermittlungseinrichtung **466** durchgeführten Fusion eine möglichst hohe Genauigkeit erreicht werden soll.

[0099] Nach dem in der Bestimmungseinrichtung **464** durchgeführten Warming sind die Bildkoordinaten der übertragenen Fusionsdaten **474** bereits auf den aktuellen Stand gebracht, die Inhalte, also die Tiefenwerte der übertragenen Fusionsdaten **474** jedoch noch nicht. Sie beziehen sich noch auf den vorherigen Zeitpunkt.

[0100] Das ist in dieser ersten Variante auch so gewollt, denn die neuen Tiefenwerte der Tiefendaten **472** werden ebenfalls so berechnet, dass sie damit kompatibel sind, also auch dem vorherigen Zeitpunkt zuzuordnen sind.

[0101] Wenn mit Messunsicherheiten gearbeitet wird, so wird für diese ebenfalls ein Warming durchgeführt. Dies kann analog zum Warming der Tiefenwerte in der Bestimmungseinrichtung **464** erfolgen. Bei diesem Schritt ist vorteilhafterweise zu berücksichtigen, dass das Warming selbst zu einer Erhöhung der Messunsicherheiten führt. Dies lässt sich dadurch veranschaulichen, dass die Messung „altert“! Eine Begründung hierfür sind z. B. die Unsicherheit bei der Bestimmung des zum Warming verwendeten Flussfelds und unter Umständen das „Altern“ der Daten, denn die Szene kann sich verändert haben, seitdem die letzte Messung stattgefunden hat. Die Erhöhung der Messunsicherheiten kann im einfachsten Falle durch eine konstante additive Korrektur aller Messunsicherheiten erfolgen.

[0102] Nun kann die Fusion der durch die Tiefendaten **472** und die übertragenen Fusionsdaten **474** repräsentierten Tiefenwerte stattfinden, bei der die neue Messung **472** und das alte (gewarpte) Ergebnis

474 zusammengebracht werden. Diese in der Ermittlungseinrichtung **466** durchgeführte Fusion erfolgt Pixel für Pixel und ohne weitere Abhängigkeiten von anderen Daten. Pixel für Pixel durchgeführten Fusionen können also parallel und unabhängig voneinander durchgeführt werden und sind dabei sehr einfach.

[0103] Es sind vier mögliche Fälle zu betrachten. Die ersten drei Fälle sind dabei mehr oder weniger trivial, während der vierte Fall der interessante und wichtige ist.

[0104] Für den ersten Fall ist am jeweiligen Pixel weder ein altes noch ein neues Ergebnis vorhanden. Also gibt es auch kein fusioniertes Ergebnis. Mit anderen Worten enthalten weder die Tiefendaten **472** noch die übertragenen Fusionsdaten **474** für das entsprechende Pixel einen zugeordneten Tiefenwert. Dies bedeutet, dass die neu ermittelten Fusionsdaten **476** für das entsprechende Pixel keinen Tiefenwert enthalten.

[0105] Für den zweiten Fall ist ein neues Ergebnis, jedoch kein altes Ergebnis vorhanden. Das neue Ergebnis wird als fusioniertes Ergebnis übernommen, seine Messunsicherheit ebenfalls. Mit anderen Worten enthalten die Tiefendaten **472** für das entsprechende Pixel einen zugeordneten Tiefenwert, die übertragenen Fusionsdaten **474** jedoch nicht. Dies bedeutet, dass die neu ermittelten Fusionsdaten **476** für das entsprechende Pixel den durch die Tiefendaten **472** für das entsprechende Pixel bereitgestellten Tiefenwert enthalten.

[0106] Für den dritten Fall ist ein altes Ergebnis, jedoch kein neues Ergebnis vorhanden. Das alte Ergebnis wird als fusioniertes Ergebnis übernommen, seine (inzwischen wegen Alterung angestiegene) Messunsicherheit ebenfalls. Mit anderen Worten enthalten die Tiefendaten **472** für das entsprechende Pixel keinen zugeordneten Tiefenwert, die übertragenen Fusionsdaten **474** dagegen schon. Dies bedeutet, dass die neu ermittelten Fusionsdaten **476** für das entsprechende Pixel den durch die übertragenen Fusionsdaten **474** für das entsprechende Pixel bereitgestellten Tiefenwert enthalten.

[0107] Für den vierten Fall sind sowohl altes als auch neues Ergebnis vorhanden. Zunächst wird geprüft, ob diese konsistent sind oder sich widersprechen. Im Falle des Widerspruchs wird keines der beiden Ergebnisse übernommen oder alternativ das neuere oder alternativ dasjenige mit der kleineren Messunsicherheit. Ein Widerspruch liegt vor, wenn beide Ergebnisse weiter voneinander entfernt liegen, als dies mit der kleineren der beiden Messunsicherheiten zu erklären ist. Im gegenteiligen Fall (konsistente Ergebnisse) findet die Fusion statt. Dabei wird vorzugsweise eine varianzgewichtete Mittelung durchgeführt, die dem Fachmann bekannt ist. Dabei werden auch

die Messunsicherheiten fusioniert. Mit anderen Worten enthalten bei diesem vierten Fall sowohl die Tiefendaten **472** als auch die übertragenen Fusionsdaten **474** für das entsprechende Pixel einen zugeordneten Tiefenwert. Dies bedeutet, dass die neu ermittelten Fusionsdaten **476**, sofern die Tiefenwerte zueinander konsistent sind, für das entsprechende Pixel einen Tiefenwert enthalten, der sich aus einer Fusion, beispielsweise einer Mittelung, der Tiefenwerte ergibt, die dem entsprechenden Pixel in den Tiefendaten **472** und den übertragenen Fusionsdaten **474** zugeordnet sind.

[0108] Zuletzt werden die in den Fusionsdaten **474** enthaltenen fusionierten Tiefenwerte beispielsweise unter Verwendung der Bestimmungseinrichtung **464** vom vorherigen Zeitpunkt auf den aktuellen Zeitpunkt übertragen. Maßgeblich hierfür ist die Eigenbewegung zwischen diesen beiden Zeitschritten, die beispielsweise unter Verwendung der Bestimmungseinrichtung **460** bereits geschätzt wurde und damit also vorliegt und angewendet werden kann.

[0109] Diese Fortschreibung der Tiefenwerte ist besonders einfach, wenn die Eigenbewegung aus einer reinen Translation (ohne Rotation) besteht. Dann ist lediglich eine additive Korrektur anzuwenden, die für alle Tiefenwerte identisch ist. Der Begriff „additive Korrektur“ schließt dabei auch das negative Vorzeichen ein. Bei Vorwärtsfahrt und nach vorne gerichteter Kamera (wichtigster Fall) ist das negative Vorzeichen zutreffend.

[0110] Falls die Eigenbewegung zusätzlich zur Translation auch eine Rotation beinhaltet, so ist zur Fortschreibung der Tiefenwerte eine Rotation im dreidimensionalen Raum erforderlich. Hierfür gibt es mehrere Möglichkeiten und Vereinfachungen, die dem Fachmann bekannt sind. Die größtmögliche Vereinfachung ergibt sich aus dem Verzicht auf diese Korrektur. Diese kann gerechtfertigt sein, wenn die Rotationswinkel klein sind, was bei einem schnell vorwärts fahrenden Fahrzeug der Fall ist.

[0111] Nun wird eine zweite Variante der Fusion im Pixelgitter beschrieben. Die zweite Variante der Fusion im Pixelgitter ist dadurch gekennzeichnet, dass die Tiefen vorab so berechnet werden, dass das Tiefenbild, das durch die Tiefendaten **472** repräsentiert wird, in den Koordinaten des neueren der beiden Bilder, die durch die Bilddaten **340**, **342** repräsentiert werden, vorliegt. Das ist so zu verstehen, dass der Tiefenwert eines Pixels dem entsprechenden Pixel an gleicher Position im neueren der beiden Originalbilder, als dem durch die ersten Bilddaten **340** repräsentierten Originalbild, zuzuordnen ist.

[0112] Entsprechend wird also auch die in der Ermittlungseinrichtung **466** durchgeführte Fusion in den neuen Koordinaten durchgeführt.

[0113] Dazu ist es notwendig, den zu fusionierenden Inhalt, in Form der Fusionsdaten **476**, des Fusionspeichers, der in **Fig. 4** durch die Bestimmungseinrichtung **464** repräsentiert wird, in zwei Schritten darauf vorzubereiten. Der eine Schritt ist das oben beschriebene Warping. Der zweite Schritt ist das ebenfalls oben beschriebene Übertragen der Tiefenwerte der Fusionsdaten **476** vom vorherigen Zeitpunkt auf den aktuellen Zeitpunkt, wobei hier die Tiefenwerte des Fusionspeichers zu übertragen sind. Die Reihenfolge der beiden Schritte kann auch umgekehrt werden.

[0114] Nach dieser Vorbereitung kann die Fusion zwischen den übertragenen Fusionsdaten **474** und den Tiefendaten **472** erfolgen.

[0115] Beide Varianten sind vom Aufwand und vom Ergebnis vergleichbar.

[0116] **Fig. 5** zeigt ein Blockschaltbild einer Vorrichtung **102** zur Verarbeitung von Bilddaten **104** gemäß einem Ausführungsbeispiel. Dabei kann es sich um ein Ausführungsbeispiel der anhand von **Fig. 1** beschriebenen Vorrichtung handeln.

[0117] Wie bereits anhand der **Fig. 3** und **Fig. 4** beschrieben, umfasst die Vorrichtung **102** Speichereinrichtungen **330**, **332**, **334**, **336**, **338** sowie die Bestimmungseinrichtungen **460**, **462**, **464** und die Ermittlungseinrichtung **466**.

[0118] Lediglich beispielhaft umfasst die Vorrichtung **102** zwei weitere Bestimmungseinrichtungen **560**, **561**, die entsprechend der Bestimmungseinrichtung **460** ausgeführt sind und zwei weitere Bestimmungseinrichtungen **562**, **563**, die entsprechend der Bestimmungseinrichtung **462** ausgeführt sind.

[0119] Die Bestimmungseinrichtung **560** ist ausgebildet, um wie anhand von **Fig. 4** beschrieben, eine Korrespondenz zwischen den ersten Bilddaten **340** und den vierten vorangegangene Bilddaten **548** zu bilden und durch eine solche Korrespondenzbildung bestimmte Korrespondenzdaten **570** bereitzustellen.

[0120] Die Bestimmungseinrichtung **561** ist ausgebildet, um wie anhand von **Fig. 4** beschrieben, eine Korrespondenz zwischen den ersten Bilddaten **340** und den zweiten vorangegangene Bilddaten **344** zu bilden und durch eine solche Korrespondenzbildung bestimmte Korrespondenzdaten **571** bereitzustellen.

[0121] Die Bestimmungseinrichtung **562** ist ausgebildet, um wie anhand von **Fig. 4** beschrieben, unter Verwendung der Korrespondenzdaten **570** eine Eigenbewegungsschätzung durchzuführen und Tiefendaten **572** auf Basis der Eigenbewegungsschätzung bereitzustellen.

[0122] Die Bestimmungseinrichtung **563** ist ausgebildet, um wie anhand von **Fig. 4** beschrieben, unter Verwendung der Korrespondenzdaten **571** eine Eigenbewegungsschätzung durchzuführen und Tiefendaten **573** auf Basis der Eigenbewegungsschätzung bereitzustellen.

[0123] Die Ermittlungseinrichtung **466** ist gemäß diesem Ausführungsbeispiel ausgebildet ist, um unter Verwendung der Tiefendaten **472**, **572**, **573** und der übertragenen Fusionsdaten **474** die Fusionsdaten **476** zu ermitteln und bereitzustellen.

[0124] Gemäß diesem Ausführungsbeispiel werden die übertragenen Fusionsdaten **474**, wie anhand von **Fig. 4** beschrieben, von der Bestimmungseinrichtung **464** bereitgestellt.

[0125] Gemäß einem Ausführungsbeispiel umfasst die Vorrichtung **102** ferner zwei weitere Bestimmungseinrichtungen **580**, **582**.

[0126] Die Bestimmungseinrichtung **580** ist ausgebildet, um eine Korrespondenz zwischen den ersten Bilddaten **340**, die auf von der ersten Kamera **106** bereitgestellten Daten basieren, und zweiten Bilddaten **584**, die auf von der zweiten Kamera **106** bereitgestellten Daten basieren, zu bilden, und durch eine solche Korrespondenzbildung **580** bestimmte Stereo-Korrespondenzdaten **586** bereitzustellen. Dabei repräsentieren die ersten Bilddaten **340** und die zweiten Bilddaten **584** zu einander entsprechenden Zeitpunkten erfasste Daten.

[0127] Die Bestimmungseinrichtung **582** ist ausgebildet, unter Verwendung der Stereo-Korrespondenzdaten **586** eine Triangulation durchzuführen und auf Basis der Triangulation Stereo-Tiefendaten **588** bereitzustellen, die Tiefendaten und/oder einen Maßstab umfassen.

[0128] Die Ermittlungseinrichtung **466** ist gemäß diesem Ausführungsbeispiel ausgebildet, um die Fusionsdaten ferner unter Verwendung der Stereo-Tiefendaten und/oder den Maßstab zu ermitteln, die über die Stereo-Tiefendaten **588** bereitgestellt werden.

[0129] Somit werden gemäß diesem Ausführungsbeispiel vier Tiefenmessungen in Form der Tiefendaten **472**, **572**, **573** sowie der Stereo-Tiefendaten **588** und ein fortgeschriebenes früheres Ergebnis in Form der übertragenen Fusionsdaten **474** miteinander fusioniert. Die Anzahl der verwendeten Tiefenmessungen sowie die zur Bestimmung der Tiefenmessungen verwendeten Bilddaten **340**, **342**, **344**, **548**, **584** ist dabei nur beispielhaft gewählt und kann entsprechend den gewünschten Einsatzgebieten der Vorrichtung **102** angepasst werden.

[0130] Für das in **Fig. 5** gezeigte komplexere Ausführungsbeispiel ist die anhand von **Fig. 4** beschriebene zweite Variante zu empfehlen, da es darin keinen gemeinsamen älteren Zeitpunkt gibt, wohl aber einen gemeinsamen neueren Zeitpunkt, auf den sich alle vier neuen Tiefenbildmessungen beziehen.

[0131] In dem in **Fig. 5** gezeigten komplexeren Ausführungsbeispiel ist das bekannte Blockschaltbild aus **Fig. 3** enthalten. Das Blockschaltbild aus **Fig. 4** ist ebenfalls enthalten.

[0132] Das Blockschaltbild in **Fig. 5** enthält insgesamt vier Blöcke „Corresp“, dargestellt durch die Bestimmungseinrichtungen **460**, **560**, **561**, **580**, in denen jeweils die Korrespondenzbildung stattfindet. Dabei wird auf der linken Seite dreimal der optische Fluss bestimmt, also eine Korrespondenzbildung in zeitlicher Richtung mit unterschiedlichen zeitlichen Abständen T , $2T$, $4T$, und auf der rechten Seite einmal eine Korrespondenzbildung zeitsynchron in örtlicher Richtung zwischen linker Kamera **106** und rechter Kamera **306**, für den Stereo-Fall, durchgeführt.

[0133] Nach den „Corresp“-Blöcken schließt sich jeweils wieder eine Egomotion- und Tiefenbestimmung an, dargestellt durch die Bestimmungseinrichtungen **462**, **562**, **563**. Die Bestimmungseinrichtung **582** stellt dabei einen „D&S“-Block (Depth & Scale-Block) dar, der ein Tiefenbild und gemäß einem Ausführungsbeispiel zusätzlich auch einen Maßstab liefert. Eine Egomotion-Schätzung ist im Allgemeinen im Block **582** nicht nötig, da das Stereo-Kamerapaar normalerweise mechanisch fest zueinander steht. Der Maßstab ergibt sich aus dem in der Regel bekannten festen Abstand zwischen den beiden Kameras **106**, **306** (Basisweite). Dieser Maßstab kann innerhalb der in der Ermittlungseinrichtung **466** durchgeführten Fusion unmittelbar weiterverwendet werden oder mit einem Maßstab aus einer anderen Quelle, wie oben beschrieben, fusioniert werden.

[0134] Der durch die Ermittlungseinrichtung **466** dargestellte Fusionsblock hat im Unterschied zu dem anhand von **Fig. 4** beschriebenen Ausführungsbeispiel nun vier Eingangsbilder, statt einem Eingangsbild. Die vier Eingangsbilder werden durch die Bilddaten **340**, **342**, **344**, **548** repräsentiert. Das Prinzip der Fusion bleibt aber das gleiche wie oben beschrieben. Stehen für einen Pixel mehrere Messungen zur Verfügung, so sind diese zu fusionieren. Dazu wird bevorzugt auch hier wieder die varianzgewichtete Mittelung durchgeführt, die dem Fachmann bekannt ist.

[0135] Optional kann diese Fusion auch Schritt für Schritt nacheinander erfolgen, also z. B. Fusion von erster und zweiter Messung, Fusion des Zwischenergebnisses mit dritter Messung, Fusion des weiteren Zwischenergebnisses mit vierter Messung.

[0136] Der in der Bestimmungseinrichtung **464** durchgeführte Warming-Schritt verwendet vorzugsweise auch hier das durch die Korrespondenzdaten **470** repräsentierte Ergebnis der Korrespondenzbildung zwischen den Bildern $I_L(t)$ und $I_L(t - T)$.

[0137] Die beschriebene Fusion im Pixelgitter stellt eine zentrale Schnittstelle im System dar, an die die beteiligten Sensoren, hier die Kameras **106, 306**, ihre Daten **104, 304** einspeisen und an der sich die weiteren Algorithmen bedienen.

[0138] Die bereitgestellten Fusionsdaten **476** können beispielsweise von Fahrerassistenzsystemen oder Informationssystemen genutzt werden, die eine gesteigerte Anforderung hinsichtlich Reichweite und Genauigkeit haben. Das beschriebene einheitliche Konzept der Basissignalverarbeitung kann für alle Kamerasysteme (Mono, Stereo, Duales Mono, Multi-Kamera) eingesetzt werden.

[0139] Umfasst ein Ausführungsbeispiel eine „und/oder“-Verknüpfung zwischen einem ersten Merkmal und einem zweiten Merkmal, so ist dies so zu lesen, dass das Ausführungsbeispiel gemäß einer Ausführungsform sowohl das erste Merkmal als auch das zweite Merkmal und gemäß einer weiteren Ausführungsform entweder nur das erste Merkmal oder nur das zweite Merkmal aufweist.

Patentansprüche

1. Verfahren zur Verarbeitung von Bilddaten (**104; 304**), wobei das Verfahren die folgenden Schritte umfasst:

Bestimmen (**220**) von Korrespondenzdaten (**470**) als Ergebnis einer Korrespondenzbildung zwischen ersten Bilddaten (**340**), die auf von einer Kamera (**106**) zu einem ersten Zeitpunkt erfassten Bilddaten (**104**) basierende Daten repräsentieren, und vorangegangenen Bilddaten (**342**), die auf von der Kamera (**106**) zu einem vorangegangenen Zeitpunkt erfassten Bilddaten (**104**) basierende Daten repräsentieren, wobei die ersten Bilddaten (**340**) Daten repräsentieren, die auf von einer Kamera (**106**) zu einem ersten Zeitpunkt erfassten Bilddaten (**104**) basieren, und wobei die vorangegangenen Bilddaten (**342**) Daten repräsentieren, die auf von der Kamera (**106**) zu einem vorangegangenen Zeitpunkt erfassten Bilddaten (**104**) basieren;

Bestimmen (**222**) von Tiefendaten (**472**) auf Basis einer unter Verwendung der Korrespondenzdaten (**470**) durchgeführten Eigenbewegungsschätzung; und

Ermitteln (**224**) von Fusionsdaten (**476**) unter Verwendung der Tiefendaten (**472**) und übertragenen Fusionsdaten (**474**), wobei die Fusionsdaten (**476**) eine auf den Bilddaten (**340, 342**) basierende Tiefeninformation umfassen.

2. Verfahren gemäß Anspruch 1, mit einem Schritt des Bestimmens (**226**) der übertragenen Fusionsdaten (**474**) als Ergebnis einer Übertragung von in einem vorangegangenen Schritt des Ermittelns ermittelten Fusionsdaten (**476**) und den Korrespondenzdaten (**470**).

3. Verfahren gemäß einem der vorangegangenen Ansprüche, bei dem im Schritt des Bestimmens (**220**) der Korrespondenzdaten (**470**) Flussvektoren als die Korrespondenzdaten (**470**) bestimmt werden, die Bildpunkten der ersten Bilddaten (**340**) und/oder Bildpunkten der vorangegangenen Bilddaten (**342**) zugeordnet sind.

4. Verfahren gemäß einem der vorangegangenen Ansprüche, bei dem im Schritt des Bestimmens (**222**) von Tiefendaten (**472**) Tiefenwerte als die Tiefendaten (**472**) bestimmt werden, die Bildpunkten der ersten Bilddaten (**340**) und/oder Bildpunkten der vorangegangenen Bilddaten (**342**) zugeordnet sind.

5. Verfahren gemäß einem der vorangegangenen Ansprüche, bei dem die Fusionsdaten (**476**) im Schritt des Ermittelns (**224**) unter Verwendung einer Mittelwertbildung zwischen den Tiefendaten (**472**) und den übertragenen Fusionsdaten (**474**) ermittelt werden.

6. Verfahren gemäß einem der vorangegangenen Ansprüche, bei dem die Fusionsdaten (**476**) im Schritt des Ermittelns (**224**) den Tiefendaten (**472**) entsprechen, wenn keine übertragenen Fusionsdaten (**474**) vorliegen, und die Fusionsdaten (**476**) den übertragenen Fusionsdaten (**474**) entsprechen, wenn keine Tiefendaten (**472**) vorliegen.

7. Verfahren gemäß einem der vorangegangenen Ansprüche, bei dem im Schritt des Ermittelns (**224**) der Fusionsdaten (**476**), eine Konsistenz zwischen den Tiefendaten (**472**) und den übertragenen Fusionsdaten (**474**) überprüft wird.

8. Verfahren gemäß einem der vorangegangenen Ansprüche, mit zumindest einem Schritt des Bestimmens (**220**) von weiteren Korrespondenzdaten (**570, 571**) als Ergebnis einer Korrespondenzbildung zwischen den ersten Bilddaten (**340**) und weiteren vorangegangenen Bilddaten (**344; 548**), wobei die vorangegangenen Bilddaten (**344; 548**) Daten repräsentieren, die auf von der Kamera (**106**) zu einem weiteren vorangegangenen Zeitpunkt erfassten Bilddaten (**104**) basieren, und mit zumindest einem Schritt des Bestimmens (**222**) von weiteren Tiefendaten (**572, 573**) auf Basis einer unter Verwendung der weiteren Korrespondenzdaten (**570, 571**) durchgeführten Eigenbewegungsschätzung, und wobei im Schritt des Ermittelns (**224**) die Fusionsdaten (**476**) ferner unter Verwendung der weiteren Tiefendaten (**572, 573**) ermittelt werden.

9. Verfahren gemäß einem der vorangegangenen Ansprüche, mit einem Schritt des Bestimmens von Stereo-Korrespondenzdaten (586) als Ergebnis einer Korrespondenzbildung zwischen den ersten Bilddaten (340) und zweiten Bilddaten (584), wobei die zweiten Bilddaten (584) Daten repräsentieren, die auf von einer weiteren Kamera (306) zu dem ersten Zeitpunkt erfassten Bilddaten (304) basieren, und mit einem Schritt des Bestimmens von Stereo-Tiefendaten (588), die Tiefendaten und/oder einen Maßstab repräsentieren, auf Basis einer unter Verwendung der Stereo-Korrespondenzdaten (586) durchgeführten Triangulation und wobei im Schritt des Ermittelns (224) die Fusionsdaten (476) ferner unter Verwendung der Stereo-Tiefendaten (588) ermittelt werden.

10. Verfahren gemäß einem der vorangegangenen Ansprüche, mit einem Schritt des Erfassens der Bilddaten (104; 304) unter Verwendung zumindest einer Kamera (106) und/oder mit einem Schritt des Verwendens der Fusionsdaten (476) in einem Fahrerassistenzsystem (110) für ein Fahrzeug (100) oder in einer Steuereinrichtung für einen mobilen Roboter.

11. Vorrichtung zur Verarbeitung von Bilddaten (104; 304), wobei die Vorrichtung die folgenden Merkmale aufweist:

eine Bestimmungseinrichtung, die ausgebildet ist, um Korrespondenzdaten (470) als Ergebnis einer Korrespondenzbildung zwischen ersten Bilddaten (340), die auf von einer Kamera (106) zu einem ersten Zeitpunkt erfassten Bilddaten (104) basierende Daten repräsentieren, und vorangegangenen Bilddaten (342), die auf von der Kamera (106) zu einem vorangegangenen Zeitpunkt erfassten Bilddaten (104) basierende Daten repräsentieren, zu bestimmen;

eine Bestimmungseinrichtung, die ausgebildet ist, um Tiefendaten (472) auf Basis einer unter Verwendung der Korrespondenzdaten (470) durchgeführten Eigenbewegungsschätzung zu bestimmen; und

eine Ermittlungseinrichtung, die ausgebildet ist, um Fusionsdaten (476) unter Verwendung der Tiefendaten (472) und übertragenen Fusionsdaten (474) zu ermitteln, wobei die Fusionsdaten (476) eine auf den Bilddaten (340, 342) basierende Tiefeninformation umfassen.

12. Fahrzeug mit zumindest einer Kamera (106), einer Vorrichtung gemäß Anspruch 11 zum Verarbeiten der Bilddaten (104; 304) der zumindest einen Kamera (106) und einem Fahrerassistenzsystem (110), das ausgebildet ist, um unter Verwendung von der Vorrichtung bereitgestellten Fusionsdaten (476) eine Fahrerassistenzfunktion auszuführen.

13. Computerprogramm, das dazu eingerichtet ist, das Verfahren gemäß einem der vorangegangenen Ansprüche auszuführen.

14. Maschinenlesbares Speichermedium, auf dem das Computerprogramm nach Anspruch 13 gespeichert ist.

Es folgen 4 Seiten Zeichnungen

Anhängende Zeichnungen

FIG. 1

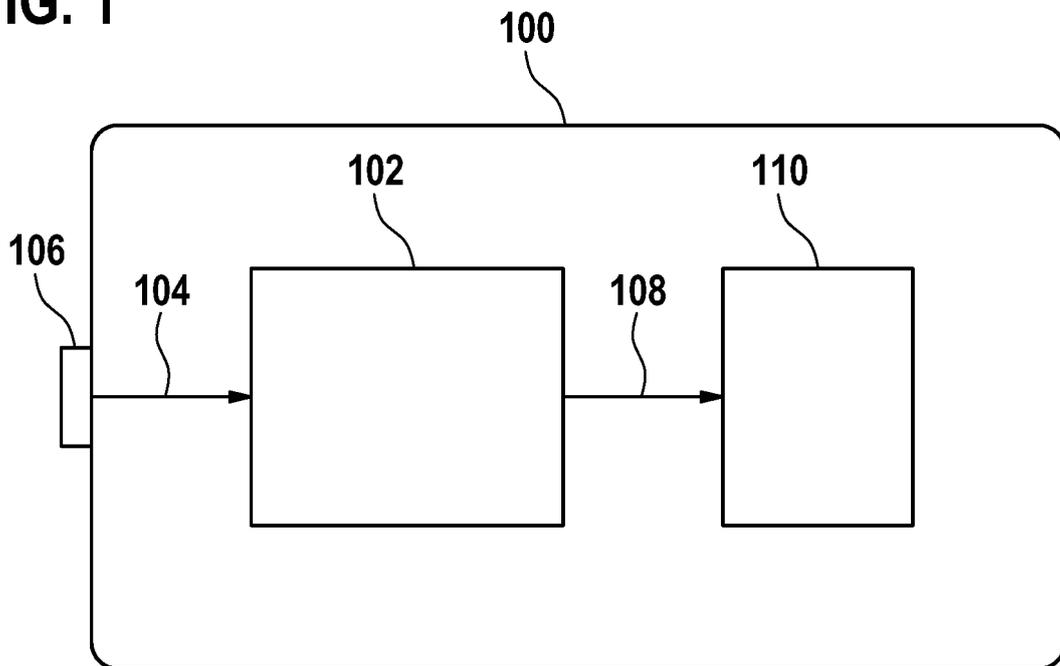


FIG. 2

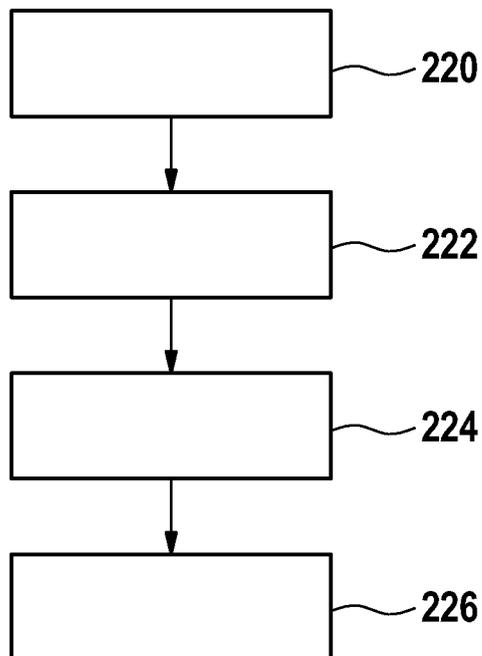


FIG. 3

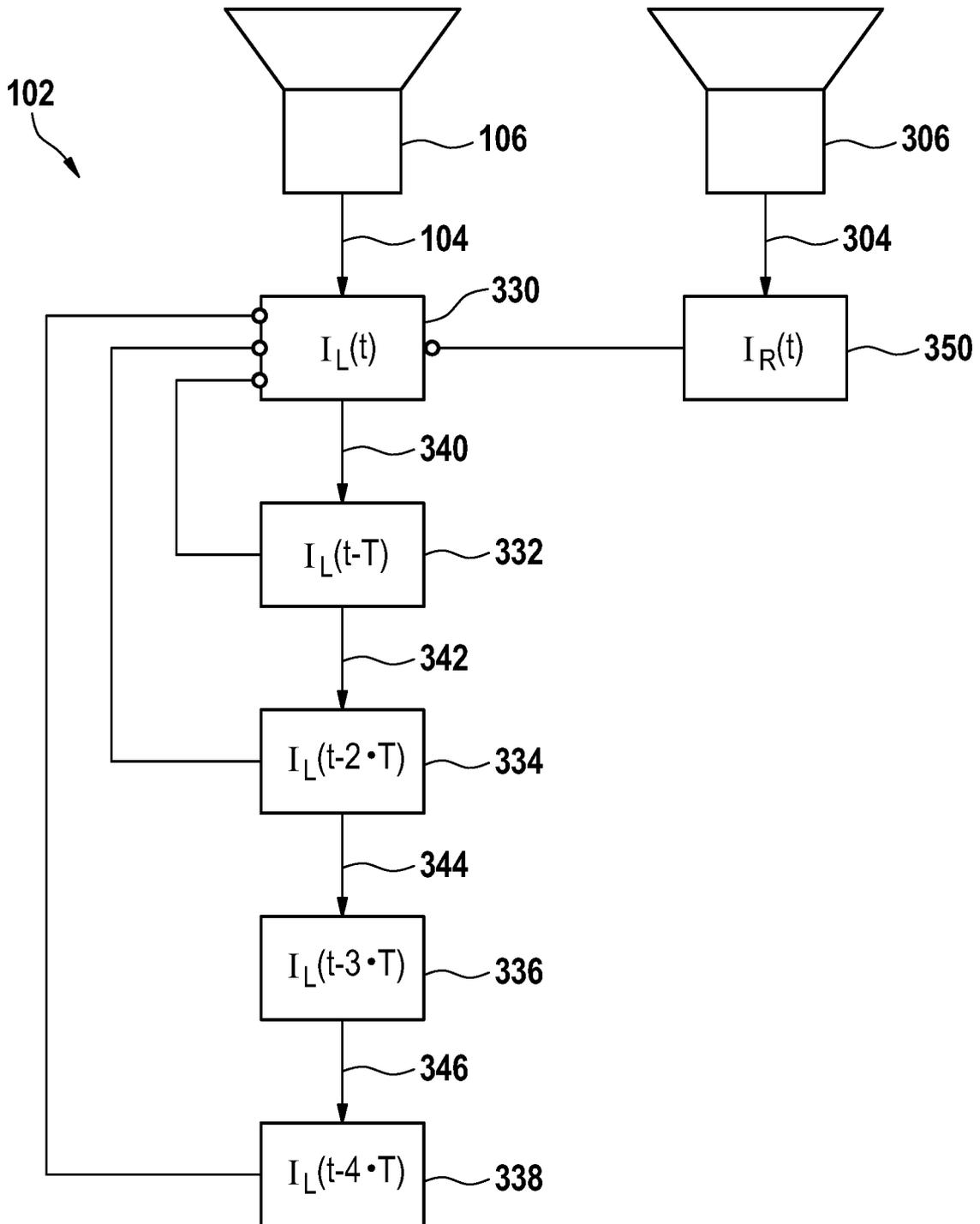


FIG. 4

