



(19)
 Bundesrepublik Deutschland
 Deutsches Patent- und Markenamt

(10) **DE 10 2009 026 091 A1 2010.04.01**

(12)

Offenlegungsschrift

(21) Aktenzeichen: **10 2009 026 091.9**

(22) Anmeldetag: **02.07.2009**

(43) Offenlegungstag: **01.04.2010**

(51) Int Cl.⁸: **B25J 19/06 (2006.01)**

(66) Innere Priorität:
10 2008 030 938.9 02.07.2008

(71) Anmelder:
**ELAN Schaltelemente GmbH & Co. KG, 35435
 Wettberg, DE**

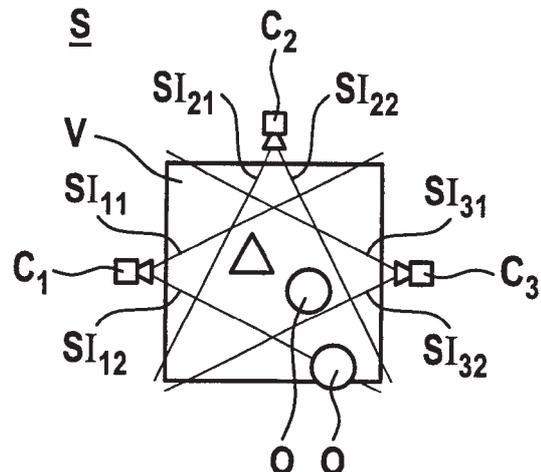
(74) Vertreter:
**Stoffregen, H., Dipl.-Phys. Dr.rer.nat., Pat.-Anw.,
 63450 Hanau**

(72) Erfinder:
Henrich, Dominik, Prof. Dr., 91126 Schwabach, DE

Die folgenden Angaben sind den vom Anmelder eingereichten Unterlagen entnommen

(54) Bezeichnung: **Verfahren und System zur Überwachung eines dreidimensionalen Raumbereichs mit mehreren Kameras**

(57) Zusammenfassung: Die Erfindung bezieht sich auf ein Verfahren zur Überwachung eines dreidimensionalen Raumbereichs (V), insbesondere einer Arbeitszelle einer automatisiert arbeitenden Maschine wie Roboter, bei dem Bilder (B_i) verwendet werden, die in zeitlicher Folge von wenigstens zwei, richtungs-diversitär bzw. richtungs-disimilär angeordneten Kameras (C₁, C₂ ... C) erfasst wurden, deren Sichtstrahlen (SI₁₁, SI₁₂; SI₂₁, SI₂₂; SI₃₁, SI₃₂) sich wenigstens teilweise in einem Anfragevolumen (v) in dem Raumbereich (V) überschneiden und bei dem anhand einer Szenenanalyse der Bilder (B_i) in das Anfragevolumen (v) eindringende Objekte (O_j) und/oder gegebenenfalls Hindernisse (H_k) detektiert werden. Um die Sicherheit zu erhöhen und um nicht einsehbare Bereiche hinter Hindernissen (Verdeckungen) auch bei der Überwachung von beliebigen Anfragevolumina zu minimieren wird vorgeschlagen, dass in jeder oder für jede Kamera (C₁, C₂ ... C_c) eine unabhängige Szenenanalyse der Bilder (B_i) durchgeführt wird, dass Ergebnisse der unabhängigen Szenenanalysen der richtungs-diversitären Kameras (C₁ ... C_c) redundant ausgewertet und dadurch ein oder mehrere fehler-tolerante Überwachungssignale gebildet werden.



Beschreibung

[0001] Die Erfindung bezieht sich auf ein Verfahren zur Überwachung eines dreidimensionalen Raumbereichs, insbesondere einer Arbeitszelle einer automatisiert arbeitenden Maschine wie Roboter, bei dem Bilder verwendet werden, die in zeitlicher Folge von wenigstens zwei, richtungs-diversitär bzw. richtungs-disimilar angeordneten Kameras erfasst wurden, deren Sichtstrahlen sich wenigstens teilweise in einem Anfragevolumen in dem Raumbereich überschneiden und bei dem anhand einer Szenenanalyse der Bilder in das Anfragevolumen eindringende Objekte und/oder gegebenenfalls Hindernisse detektiert werden, sowie auf ein System zur Überwachung eines dreidimensionalen Raumbereichs, insbesondere einer Arbeitszelle einer automatisiert arbeitenden Maschine wie Roboter, mit zumindest zwei dezentral und richtungs-diversitär bzw. richtungs-disimilar angeordneten einzelnen Kameras zum Aufnehmen von zumindest zwei Bildern des Raumes, wobei die Kameras versetzt zueinander angeordnet sind, und mit einer Auswerteeinheit die mittels einer Szenenanalyse der Bilder Objekte und/oder Hindernisse in dem Raum detektiert.

[0002] Ein Verfahren sowie ein System der eingangs genannten Art ist in der EP 1 635 108 A1 beschrieben. Die EP 1 635 108 A1 beschreibt ein Verfahren und eine Vorrichtung zum Erfassen eines Objektes innerhalb eines Überwachungsbereichs, bei dem zumindest zwei Sensoren jeweils mindestens einen Teil des Überwachungsbereichs überwachen. Jeder Sensor erzeugt ein digitales Einzelbild des gesamten Überwachungsbereichs, wobei die Einzelbilder aller Sensoren in einem einheitlichen Koordinatensystem erzeugt werden. Von jedem Sensor werden diejenigen Koordinatenpunkte der Einzelabbilder als „sicher“ gekennzeichnet, die er als frei von dem Objekt erkennt, wobei alle anderen Koordinatenpunkte der Einzelabbilder als „unsicher“ gekennzeichnet werden. Die derart erzeugten Einzelabbilder werden jeweils zu einer Auswerteeinheit übertragen und die übertragenen Einzelabbilder werden von der Auswerteeinheit entsprechend ihren gemeinsamen Koordinatensystemen zu einem resultierenden Gesamtbild überlagert, in dem diejenigen Koordinatenpunkte des Gesamtbildes als „sicher“ gekennzeichnet werden, die zumindest in einem der von den Sensoren übertragenen Einzelabbildern als „sicher“ gekennzeichnet sind, während alle anderen Koordinatenpunkte des Gesamtbildes als „unsicher“ gekennzeichnet werden. Durch das Verfahren soll eine genauere Erfassung von Objekten möglich sein, insbesondere soll bei der Erfassung eines Objektes auch der „Schattenbereich“ des Objektes als solcher erkannt und als „sicher“ eingestuft werden können, so lange wenigstens ein Sensor diesen Schattenbereich einsehen kann. Eine redundante Auswertung der Sensoren ist in dieser Druckschrift nicht angesprochen.

[0003] In einem Dokument von Ebert D., Henrich D.: "Safe Human-Robot-Cooperation: Imagebased Collision Detection for Industrial Robots", IEEE International Conference on Intelligent Robots and Systems, (IROS'02), Lausanne, 30. September bis 04. Oktober 2002 wird die Überwachung eines dreidimensionalen Arbeitsraumes mittels mehrerer stationärer Videokameras beschrieben. In dem zu überwachenden Arbeitsraum befindet sich ein Roboter, bekannte statische Hindernisse sowie zu detektierende, dynamische, 3-dimensionale Hindernisse, die sich von dem Hintergrund unterscheiden. Ein Lösungsansatz zur bildbasierten Kollisionserkennung zwischen dynamischen Hindernissen und einer beliebigen Testkonfiguration durch Roboter wird beschrieben, ohne dass allerdings sicherheitsgerichtete Aspekte betreffend die Auswertung der Videokameras angesprochen werden.

[0004] Die EP 1 543 270 B1 bezieht sich auf ein Verfahren und eine Vorrichtung zum Absichern eines Gefahrenbereichs, insbesondere des Gefahrenbereichs einer automatisiert arbeitenden Maschine, bei dem mit zumindest einer ersten und einer zweiten Bildaufnahmeeinheit zumindest zwei Bilder des Gefahrenbereichs aufgenommen werden, wobei die Bildaufnahmeeinheiten versetzt zueinander angeordnet sind und bei den anhand einer Sequenzanalyse der Bilder Fremdobjekte in dem Gefahrenbereich detektiert werden. Der dort beschriebenen Lösung liegt der Gedanke zugrunde, eine dreidimensionale Auswertung von mindestens zwei Bildern auszuführen, die mit zu einander versetzten Bildaufnahmeeinheiten von dem Gefahrenbereich aufgenommen werden. Es erfolgt eine Szenenanalyse anhand der mindestens zwei Bilder unter Anwendung mindestens zweier algorithmisch unterschiedlicher Verfahren. Der Gefahrenbereich wird abgesichert, wenn mindestens eines der Analyseverfahren eine Fremdobjektdetektion liefert. Bei diesem Verfahren sind die Kameras räumlich nicht diversitär angeordnet. Eine Auswertung eines Raumvolumens ist ebenfalls nicht angesprochen. Bei dem Verfahren werden Kamerasysteme eingesetzt, welche eine bauliche Einheit mit drei bzw. vier zentral angeordneten Kameras bilden. Die zentrale Anordnung der Kameras ist dabei notwendig, um mittels der Stereoskopie Entfernungen ermitteln zu können. „Blicken“ solche Kamerasysteme jedoch auf ein Hindernis im Überwachungsraum, kann der dahinter liegende, verdeckte Raum nicht mehr überwacht werden. Daraus ergeben sich Verdeckungen („Tot-Zonen“), die entweder den Systemnutzen einschränken, beispielsweise frei Schutzräume bzw. Schutzräume ohne Hindernisse erfordern oder den Einsatz trotz allem nicht angezeigt erscheinen lassen. Eine Kaskadierung ist dagegen unter Kostengesichtspunkten kritisch.

[0005] Weitere Ansätze, die auf einen sicherheitsgerichteten Einsatz von Kameras bzw. Kamerasystemen beziehen, jedoch eine zweidimensionale Arbeitsweise zugrunde legen, existieren beispielsweise auf dem Gebiet der Normung „Vision based protective devices“. Allerdings können derartige Systeme nicht ohne Weiteres einen Raum dreidimensional überwachen bzw. verdeckte Bereiche einsehen.

[0006] Davon ausgehend liegt der Erfindung die Aufgabe zugrunde, ein Verfahren und ein System der eingangs genannten Art derart weiterzubilden, dass die Sicherheit erhöht wird und nicht einsehbare Bereiche hinter Hindernissen (Verdeckungen) minimiert werden. Auch soll eine sichere Raumüberwachung für beliebige Anfragevolumina ermöglicht werden.

[0007] Die Aufgabe wird erfindungsgemäß dadurch gelöst, dass in jeder oder für jede Kamera eine unabhängige Szenenanalyse der Bilder durchgeführt wird, dass Ergebnisse der unabhängigen Szenenanalysen der richtungs-diversitären Kameras redundant ausgewertet und dadurch ein oder mehrere fehlertolerante Überwachungssignale gebildet werden.

[0008] Die Erfindung basiert auf der Idee, zumindest zwei vorzugsweise zumindest drei einzelne Kameras dezentral und dabei richtungs-diversitär (bzw. richtungs-disimilar) zur Überwachung des dreidimensionalen Raumbereichs anzuordnen. Dadurch können Verdeckungen, das heißt nicht einsehbare Bereiche hinter Hindernissen minimiert bzw. besser eingesehen werden. In bevorzugter Ausführungsform kann ein Sichtstrahl einzelner oder mehrerer Kameras aufgeteilt werden.

[0009] Zunächst erfolgt eine Kalibrierung der Kameras, sowohl bezogen auf deren intrinsische Parameter (z. B. Verzerrung, Fokus) als auch bezogen auf extrinsische Parameter (z. B. Ort und Richtung ihrer Installation). Dadurch lässt sich relativ genau bestimmen, wo sich Sichtstrahlen oder aufgeteilte Sichtstrahlen der einzelnen Kameras im dreidimensionalen Raum kreuzen bzw. überlappen. Ergänzend kann noch ermittelt werden, welche Sichtstrahlen und wie viele Sichtstrahlen sich kreuzen bzw. überlappen. Zusätzlich wird der zu überwachende Raum in frei wählbare so genannte Anfragevolumina unterteilt, welche zwar vorzugsweise konvex aber nicht notwendigerweise disjunkt sind oder den Überwachungsraum vollständig abdecken.

[0010] Sofern eine bestimmte Anzahl an Kameras ein freies Anfragevolumen „sieht“, kann angenommen werden, dass der betreffende Raum frei ist. Sofern darin Hindernisse „erlaubt“ bzw. gegeben vorhanden sind, vermindern sich so durch die richtungs-diversitäre Anordnung der Kameras die verdeckten Bereiche in einem Überwachungsraum.

[0011] Des Weiteren zeichnet sich die Erfindung durch einen eigenerfinderischen Gedanken dadurch aus, dass nicht alle Kameras das Anfragevolumen gleichermaßen „frei“ sehen müssen, sondern ein Voting beispielsweise über eine Auswahl-schaltung möglich ist, wenn beispielsweise aus Sicht einer Kamera das Anfragevolumen hinter einem Hindernis liegt und somit verdeckt ist. In diesem Fall müssen nun seitens der anderen Kameras konsistente Informationen in redundanten bzw. sicherheitstechnisch akzeptabler Form vorliegen, zum Beispiel eine Auswahl von drei (zwei out of drei, 2 o 3) um in der bzw. in den nachfolgenden Stufe(n) der Signalverarbeitung der Auswerteelektronik eine Anfrageeinheit dennoch als „frei“ zu bewerten, andernfalls würde ein Signal zur Warnung oder Abschaltung oder Umsteuerung der gefahrbringenden Bewegung gegeben werden.

[0012] Vorzugsweise werden Hindernisse wie statische oder dynamische Hindernisse vor Inbetriebnahme des Systems in dieses „eingelernt“ oder in diesem hinterlegt, so dass der den Sichtbereich einer Kamera beschränkende verdeckte Raum und die somit dahinter liegenden Verdeckungsbereiche dem System bekannt sind. Somit kann fortlaufend, vorzugsweise per Einzelbildverfahren, eine Art von Soll-/Ist-Vergleich erfolgen, der in das System-Voting intelligent eingebunden werden kann. Die Hindernisse, definiert bzw. interpretiert als erlaubte Objekte im Raum, können dabei nahezu beliebige Geometrie haben.

[0013] Ferner besteht gemäß einer bevorzugten Verfahrensweise die Möglichkeit, im Rahmen des Votings Bilder von Anfragevolumina bestimmter Kameras als nicht redundant bzw. sicherheitstechnisch akzeptabel auszuklammern, zum Beispiel wenn die Richtungs-Diversität der Kameras nicht gegeben wäre.

[0014] Eine weitere bevorzugte Ausführungsform der Erfindung mit eigenständigem Erfindungsgedanken besteht darin, Zustände einer bestimmten Anzahl von Anfragevolumina zu kombinieren bzw. zusammen zu fügen und aus der Kombination (aus dem Zusammenfügen) der Zustände einer bestimmten Anzahl von Anfragevolumina Plausibilitätskontrollen, das heißt Schutzfeldanalysen vorzunehmen.

[0015] So können beispielsweise einzelne im Raum schwebende Anfragevolumina, die als nicht „frei“ erkannt werden, dennoch toleriert und – bedingt durch eine Störung – als unkritisch uminterpretiert werden, wenn es sich beispielsweise aufgrund der geometrischen Anordnung nicht um einen Menschen handeln kann.

[0016] Ebenso erlaubt diese Verfahrensweise auch eine Mustererkennung bzw. Texturerkennung der Plausibilitätskontrolle, zum Beispiel um im Rahmen einer Muting-Funktion die Schutzfunktion der Schutzeinrichtung zeitlich begrenzt und automatisch zu überbrücken. So können beispielsweise Waren in einem gefahrbringenden Bereich hinein – oder aus ihm heraus – befördert werden, ohne dass die Schutzeinrichtung die Schutzfunktion auslöst.

[0017] Auch besteht die Möglichkeit Menschen zu detektieren, die unerlaubter Weise auf einem Transportband mitfahren. Auch in diesem Fall würden „erlaubte“ und/oder „nicht erlaubte“ Geometrien dem System vorher eingelernt.

[0018] Die besonders bevorzugten Aspekte des Verfahrens können wie folgt zusammengefasst werden:

- Mehrere kalibrierte Kameras (z. B. Tiefen- oder Intensitätskameras) beobachten von unterschiedlichen Richtungen ein Überwachungsvolumen (Anfragevolumen).
- Es liegen ausreichend Informationen über die Verdeckungsgebiete im Überwachungsvolumen vor (z. B. der Ort und die Geometrie der Hindernisse oder die maximale Größe der Verdeckungen).
- Aus den aktuellen Kamerabildern und der Verdeckungsinformation können unbekannte Objekte in beliebiger Anzahl sicher detektiert werden.
- Die Detektion der Objekte wird dadurch sicher, dass eine vorgegebene Anzahl von Fehlern durch redundante Auswertung der Kamerabilder kompensiert werden (z. B. pro Voxel, Strahl, Sichtbereich, Überwachungsraum).

[0019] Das Verfahren kann zur Raumrekonstruktion, Sicht-/Verdeckungsanalyse, Schutzzonenüberwachung, Kollisionserkennung, Kollisionsvermeidung oder als virtueller Schalter angewendet werden.

[0020] Ferner ist das Verfahren auf mehrere Überwachungsräume, welche jeweils von einer Menge an Kameras überwacht werden, erweiterbar. Diese Mengen können diskunkt, teilweise oder vollständig überlappend sein.

[0021] Die Aufgabe wird erfindungsgemäß durch ein System dadurch gelöst, dass die Auswerteeinheit Mittel zur redundanten Auswertung von Ergebnissen von unabhängigen Szenenanalysen der einzelnen Kameras sowie Mittel zur Erzeugung zumindest eines sicherheitsrelevanten Überwachungssignals auf der Grundlage der Auswertung aufweist. Das System besteht dabei aus einer Anzahl von zumindest zwei, vorzugsweise zumindest drei Kameras mit zugehöriger Auswerteelektronik. Sicherheitstechnisch soll dabei eine Hardwarefehlertoleranz von mindestens 1 und alle weiteren Maßnahmen realisiert werden, die zu einem Safety-Integrity-Level von mindestens 2 führen. (Siehe hierzu auch EN IC 61508-1/-7.)

[0022] In dem System können sowohl handelsübliche Kameras als auch speziell ertüchtigte Kameras eingesetzt werden.

[0023] Weitere Einzelheiten, Vorteile und Merkmale der Erfindung ergeben sich nicht nur aus den Ansprüchen, den diesen zu entnehmenden Merkmalen – für sich und/oder in Kombination –, sondern auch aus der nachfolgenden Beschreibung von den Zeichnungen zu entnehmenden bevorzugten Ausführungsformen.

[0024] Es zeigen:

[0025] Fig. 1 eine schematische Darstellung eines Systems zur sicherheitsgerichteten Überwachung eines Raumes umfassend Kameras, Objekte sowie Hindernisse,

[0026] Fig. 2a)–e) eine schematische Darstellung verschiedener Anwendungen wie Sicht-/Verdeckungsanalyse, Raumrekonstruktion, Zonenüberwachung, sowie Kollisionsvermeidung,

[0027] Fig. 3 eine schematische Darstellung einer Arbeitszelle mit Industrieroboter, statischen Hindernissen, dynamischen Hindernissen sowie unbekanntem Objekten (z. B. Mensch),

[0028] Fig. 4 eine schematische Darstellung mehrerer Überwachungsräume mit mobilem Industrieroboter,

- [0029] Fig. 5 ein Ablaufdiagramm des erfindungsgemäßen Verfahrens,
- [0030] Fig. 6 Detektion mit einer Kamera: (Rück-)Projektion,
- [0031] Fig. 7 Detektion mit einer Kamera: verallgemeinerte Abstandsbilder,
- [0032] Fig. 8a)–c) Detektion mit einer Kamera: abgedeckte Technologien wie Abstandsbilder, Hybridbilder sowie Differenzbilder,
- [0033] Fig. 9a)–c) Detektion mit einer Kamera: Definition verschiedener Abstände zur Kamera,
- [0034] Fig. 10 Detektion mit einer Kamera: Klassifikation von Pixeln,
- [0035] Fig. 11 Tabelle,
- [0036] Fig. 12 Detektion mit mehreren Kameras: Formen der Redundanz,
- [0037] Fig. 13a), b) Detektion mit mehreren Kameras: Definition von Detektionsdiagrammen: (a) Kontinuierliche Darstellung, (b) Diskrete Darstellung
- [0038] Fig. 13c) Detektion mit mehreren Kameras: Beispiel eines Detektionsdiagramms
- [0039] Fig. 14a), b) Detektion mit mehreren Kameras: Beispiele für Entscheidungsstrategien,
- [0040] Fig. 15 Detektion mit mehreren Kameras: Der Zustand „unbestimmt“,
- [0041] Fig. 16 Umgang mit Verdeckungen: Verdeckung als Objekt,
- [0042] Fig. 17 Umgang mit Verdeckungen: Maskierung,
- [0043] Fig. 18 Umgang mit Verdeckungen: minimaler Durchmesser, und
- [0044] Fig. 19 Umgang mit Verdeckungen: Koordinaten-Abschnitt.
- [0045] Fig. 1 zeigt rein schematisch ein System S zur sicherheitsgerichteten Überwachung eines Volumens V im dreidimensionalen Raum, insbesondere einer Arbeitszelle durch mehrere Kameras C_i mit $i = 1 \dots c$. Im vorliegenden Ausführungsbeispiel sind drei kalibrierte Kameras C_1, C_2, C_3 , beispielsweise Zeilen-, Grauwert-, Farb- oder Tiefenkamera ausgebildet, die an einem bekannten Ort und mit bekannter Orientierung (extrinsische Parameter) sowie Fokus und Verzerrung (intrinsische Parameter) angeordnet sind.
- [0046] Das überwachte Volumen $V \subseteq \mathbb{R}^3$ wird teilweise von mehreren Kameras C_i überwacht. In dem Volumen V befinden sich opake Objekte O_j wie Menschen mit unbekannter Anzahl und Geometrie. Gegebenenfalls befinden sich in dem Volumen V noch vorzugsweise opake, konvexe Hindernisse H_k wie Pfeiler, Industrieroboter mit zu jedem Zeitpunkt in dem Volumen V bekannter Geometrie.
- [0047] Jede Kamera C_i liefert ein aufgenommenes (Teil-)Bild B_i des Volumens V gegebenenfalls mit detektierten Objekten O_j bzw. Hindernissen H_k .
- [0048] Eine für die sichere Überwachung geforderte Redundanz r ist vorgegeben. Aufgrund der zuvor erfolgten Kalibrierung der Kameras C_i lässt sich relativ genau bestimmen, wo sich Sichtstrahlen $SI_{11}, SI_{12}; SI_{21}, SI_{22}; SI_{31}, SI_{32}$ im dreidimensionalen Raum V kreuzen bzw. überlappen.
- [0049] Mittels des Systems S erfolgt für alle konvexen Anfragevolumina $v \subseteq V$ eine sichere Detektion des Zustandes $Z(v)$ aus der Menge der möglichen Zustände {Frei, Objekte, Hindernis} mit
- | | |
|------------|----------------------------------------------|
| Frei: | v enthält weder O_j noch H_k |
| Objekt: | v enthält möglicherweise Teile von O_j |
| Hindernis: | v enthält Teile von H_k aber keine O_j |
- [0050] Die Detektion ist sicher, falls Zustand $Z(v)$ nur dann frei ist, wenn das Anfragevolumina v mit mindes-

tens r-facher Redundanz von den Kameras C_i als „frei“ von Objekten O_i erkannt wird.

[0051] Mit dem oben genannten Ansatz sind verschiedenen Anwendungen denkbar. Dazu gehören die Raumrekonstruktion, die Zonenüberwachung und die Kollisionserkennung. [Fig. 2a\)](#) zeigt ein Beispiel der Raumrekonstruktion, bei welcher mit einer Kamera C_1 sowie einem Hindernis H_1 , der von der Kamera C_1 überwachte Bereich in Sichtbereich SB_1 , SB_2 sowie in einen Verdeckungsgebiet VB unterteilt werden kann. Ein weiteres Beispiel der Raumrekonstruktion ist die Bestimmung der möglichen Aufenthaltsorte von eindringenden Objekten O ([Fig. 2b](#)).

[0052] In [Fig. 2c\)](#) ist ein Beispiel der Zonenüberwachung dargestellt, wobei das Eindringen von (möglichen) Objekten O in eine Warn- bzw. Schutzzone SZ dargestellt ist. Das Eindringen des Objektes O (z. B. eine Hand) in die Schutzzone SZ kann beispielsweise auch als das Betätigen eines virtuellen Schalters interpretiert werden.

[0053] [Fig. 2d\)](#) zeigt ein Beispiel der Kollisionsvermeidung, wobei eine Schutzzone SZ , beispielsweise ein Roboterarm, auf einer Bewegungsbahn BB kollisionsfrei um ein Objekt O zu einem Zielpunkt bewegt wird.

[0054] [Fig. 3](#) zeigt ebenfalls rein schematisch ein dreidimensionales Volumen V in Form einer Arbeitszelle, umfassend ein beispielsweise bewegtes Hindernis H_1 wie Industrieroboter mit einem ebenfalls bewegten Hindernis evtl. als Schutzzone SZ_1 , SZ_2 sowie bekannte, statische Hindernisse H_2 in Form von Säulen, Wänden, Schränken und Tischen. Der Zugang des Objektes O wie Mensch kann mittels der Kameras C_1 , C_2 , C_3 , C_4 erfasst werden.

[0055] In [Fig. 4](#) ist ebenfalls rein schematisch ein System zur sicherheitsgerichteten Überwachung mehrerer Überwachungsräume V_1 , V_2 dargestellt, wobei zum Beispiel ein Hindernis H_1 wie mobiler Industrieroboter über eine Transportvorrichtung T zwischen den Überwachungsräumen V_1 , V_2 transportierbar ist. Der Überwachungsraum V_1 wird durch die Kameras C_1 bis C_4 und der Überwachungsraum V_2 durch die Kameras C_5 bis C_8 überwacht.

[0056] [Fig. 5](#) zeigt ebenfalls rein schematisch ein Ablaufdiagramm des erfindungsgemäßen Verfahrens. Wie bereits zuvor erwähnt erfolgt zunächst eine Kalibrierung der Kameras C_i sowohl bezogen auf deren intrinsischen und extrinsischen Parameter. Weiterhin werden vorab Ort und Richtung, Hindernisse H_j sowie das Anfragevolumina v definiert. Jede Kamera C_i liefert ein Bild bzw. Teilbild B_i des Volumens V mit detektierten Objekten O bzw. Hindernissen H . In einem ersten Verfahrensschritt erfolgt eine Klassifikation der Pixel p . Die Klassifikation erfolgt für das Anfragevolumen jedes Bild $B_1 \dots B_c$ separat. Mit dem Ergebnis der Klassifikationen $K_i(p, v) \in \{\text{Frei, Objekt, Hindernis}\}$ wird die dominante Klasse $D_i(v) \in \{\text{Frei, Objekt, Hindernis}\}$ für das Anfragevolumen v bestimmt. Unter Berücksichtigung der Redundanz r erfolgt sodann eine kamerabasierte Entscheidung mittels eines Detektionsdiagramms, d. h. die Detektion des Zustands $Z^+(v) \in \{\text{Frei, Objekt, Hindernis, Unbestimmt}\}$.

[0057] Unter Berücksichtigung von weiteren Parameter wie Verdeckungsanzahl θ , Abstände d_{\min} , z_{\max} usw. erfolgt dann eine anwendungsbasierte Auflösung unbestimmter Zustände. Als Ergebnis wird der Zustand $Z(v) \in \{\text{Frei, Objekt, Hindernis}\}$ zur Verfügung gestellt.

[0058] Zur Erläuterung ist die Detektion eines Volumens mit einer Kamera in [Fig. 6](#) mittels (Rück-)Projektion erläutert. Für ein Volumen v enthält die Projektion $P_i(v)$ in der Kamera C_i alle Pixel $p \in B_i$ auf welche das Volumen v durch eine Optik der Kamera C_i abgebildet wird. Für Pixel $p \in B_i$ der Kamera C_i ist die Rückprojektion $P_i^{-1}(p)$ die Vereinigung aller Teilvolumina v' mit $p = P_i(v')$. $P_i^{-1}(p)$ wird Sichtkegel SK von p genannt. Es gilt immer $v \subset P_i^{-1}(P_i(v))$ und häufig auch noch $v \neq P_i^{-1}(P_i(v))$.

[0059] Ferner können die Bilder der Kamera C_i auf allgemeine Abstandsbilder generalisiert werden, wie dies in [Fig. 7](#) dargestellt ist. In einem verallgemeinerten Abstandsbild enthält ein Kamerabild B_i pro Pixel $p \in P_i$ den Abstand d_p von Kamera C_i zu dem nächsten detektierbaren Objekt O bzw. Hindernis H innerhalb des Sichtkegels SK von Pixel p . Daraus folgt, dass ein Teilvolumen TV im Sichtkegel SK von Kamera C_i bis d_p frei ist, während ein Teilvolumen TV_2 im Sichtkegel SK ab d_p verdeckt ist und gegebenenfalls weitere Objekte O bzw. Hindernisse H enthalten kann.

[0060] Die [Fig. 8a\)](#) bis c) zeigen Beispiele von Technologien, die mit dem verallgemeinerten Abstand d_p pro Pixel $p \in P_i$ erfasst werden können: Abstandsbilder aus Tiefenkameras ([Fig. 8a\)](#)) mit Abstand $0 \leq d_p < \infty$ zu Hindernissen H bzw. eventuell Objekten

O und $d_p = \infty$ ansonsten.

[0061] Hybridbilder aus Intensitätskamas (Fig. 8b) mit Abstand $d_p = 0$ für Objekte (Vordergrund) bzw. $0 < d_p < \infty$ für Hindernisse und $d_p = \infty$ ansonsten (Abstand). Bei bewegten Hindernissen ist hierbei allerdings ein so genanntes „Background Modelling“ nötig. Bei dem Background Modelling wird der Hintergrund der Szene so modelliert (z. B. über Referenzbilder oder Geometriemodelle), dass bekannte Hindernisse H_i nicht als Vordergrund detektiert werden und somit unbekannte Objekte O_j vor diesen Hindernissen als Objekte erkannt werden können.

[0062] Fig. 8c zeigt Differenzbilder aus Intensitätskamas mit Abstand $d_p = 0$ für Objekte bzw. Hindernisse (Vordergrund) und $d_p = \infty$ sonst (Hintergrund).

[0063] Die Fig. 9a) bis Fig. 9c) zeigen weitere wichtige Abstandsinformationen. Dazu sei Pixel $p \in P_i(v)$ aus der Projektion des Anfangsvolumens $v \subseteq V$ in Kamera C_i mit Sichtstrahl $s = P_i^{-1}(p)$ gegeben.

[0064] Aus der Kamerakalibrierung und der gegebenen Hindernis-Geometrie sind dann folgende Abstände $d_p, d_{v,\min}, d_{v,\max}, d_{H,\min}$ sowie $d_{H,\max}$ messbar bzw. berechenbar.

[0065] Fig. 9a) zeigt die Abstände $d_{v,\min}$ bzw. $d_{v,\max}$ als minimalen bzw. maximalen Abstand des Anfragevolumens v zur Kamera C_i entlang des Sichtstrahls s .

[0066] Fig. 9b) zeigt die Abstände $d_{H,\min}$ bzw. $d_{H,\max}$ als minimalen bzw. maximalen Abstand eines Hindernisses H zur Kamera C_i entlang des Sichtstrahls s .

[0067] Fig. 9c) zeigt den gemessenen Abstand d_p zu dem nächsten Objekt, welcher auch mit dem minimalen Hindernisabstand $d_{H,\min}$ zusammenfallen kann. Wegen der Konvexität der Hindernisse befindet sich zwischen $d_{H,\min}$ und $d_{H,\max}$ nur Hindernisvolumen.

[0068] Fig. 10 zeigt den Verfahrensschritt der Klassifikation von einzelnen Pixeln p einer Kamera C_i . Sei ein Pixel $p \in P_i(v)$ aus der Projektion des Anfragevolumens $v \subseteq V$ in der Kamera C_i mit Sichtstrahl $s = p_i^{-1}(p)$ gegeben. Dann liefert die sensorbasierte Klassifikation $K_i(p, v)$ von Pixel p einen Zustand aus {Frei, Objekt, Hindernisse} bezüglich des Teilvolumens $v \cap s$ mit

$$K_i(p, v) = \begin{cases} \text{Frei,} & \text{falls } d_{v,\max} < d_p \\ \text{Hindernis,} & \text{falls } d_{v,\max} \geq d_p \\ & \text{und } \exists k : (d_p = d_{H_k,\min} \text{ und } d_{v,\max} < d_{H_k,\max}) \\ \text{Objekt,} & \text{falls } d_{v,\max} \geq d_p \\ & \text{und } (\exists k : (d_p = d_{H_k,\min} \text{ und } d_{H_k,\max} \leq d_{v,\max}) \\ & \text{oder } \forall k : (d_p \neq d_{H_k,\min})) \end{cases}$$

[0069] $K_i(P, V)$ ist undefiniert für $p \notin P_i(v)$, da dann $v \cap s = \emptyset$ gilt.

[0070] Die in dem Kamerabild B_i projizierten Pixel $p \in P_i(v)$ können pro Anfragevolumen v Teilmenge von V unterschiedlich klassifiziert sein. Gemischte Klassifikationen $K_i(p, v)$ sind pro Kamera C_i auf eine dominante Klasse $D_i(v)$ zurückführbar wie dies in der Tabelle gemäß Fig. 11 dargestellt ist.

[0071] Zusammen mit den einheitlichen Klassifikationen erfolgt die Bestimmung der dominanten Klasse $D_i(v)$ von Volumen V mit

$$D_i(v) = \begin{cases} \text{Frei,} & \text{falls } \forall p \in P_i(v) : K_i(p, v) = \text{Frei} \\ \text{Objekt,} & \text{falls } \exists p \in P_i(v) : K_i(p, v) = \text{Objekt} \\ \text{Hindernis,} & \text{falls } \exists p \in P_i(v) : K_i(p, v) = \text{Hindernis} \\ & \text{und } \neg(\exists p \in P_i(v) : K_i(p, v) = \text{Objekt}) \end{cases}$$

[0072] Um einen oder mehrere Fehler im System zu tolerieren, wird die Information über mehrere Kameras hinweg redundant ausgewertet. In Fig. 12 sind verschiedene Formen der Redundanz r dargestellt. Sie unterscheiden sich in der konkreten Ausprägung des Anfragevolumens v . Dieses kann beispielsweise ein vorzugs-

weise würfelförmiges Volumenelement (Voxel) (a), ein Sichtstrahl (b), ein Sichtbereich, welcher von allen Kameras identisch klassifiziert wird (c) oder das gesamte, von allen Kameras prinzipiell überdeckte Überwachungsvolumen (d) sein. Dieses Anfragevolumen v wird von mehreren Kameras gleichzeitig überwacht, so dass mehrere Detektionsaussagen über v vorliegen.

[0073] Diese redundanten Detektionsaussagen sind nicht unproblematisch. Liefert beispielsweise eine Kamera C_i für ein beliebiges Teilvolumen v die (pixelbasierte) Detektion $D_i(v) \in \{\text{Frei, Objekt, Hindernis}\}$ kann dies zu widersprüchlichen Aussagen führen, wie beispielsweise $D_i(v) = \text{Frei}$ oder $D_i(v) = \text{Objekt}$.

[0074] Diese Problematik kann mit den in [Fig. 13a](#), b) eingeführten Detektionsdiagrammen behoben werden. Es sei (f, o, h) der Detektionsvektor für Volumen $v \subseteq V$ mit

$$f := \text{card}\{i | D_i(v) = \text{Frei}\}$$

$$o := \text{card}\{i | D_i(v) = \text{Objekt}\}$$

$$h := \text{card}\{i | D_i(v) = \text{Hindernis}\}$$

[0075] Für eine Anzahl von c Kameras gilt immer $f + o + h = c$.

[0076] Das Detektionsdiagramm entspricht dabei einer Dreiecksfläche im Koordinatensystem von (f, o, h) wie dies in [Fig. 13a](#)) kontinuierlich dargestellt ist. Für eine diskrete Anzahl von Kameras entsteht das diskrete Detektionsdiagramm in [Fig. 13b](#)). Beispiele für drei unterschiedliche Detektionsvektoren sind in [Fig. 13c](#)) gegeben. Oben in [Fig. 13c](#)) ist der Detektionsvektor $(f, o, h) = (2, 0, 0)$, in der Mitte $(f, o, h) = (1, 1, 0)$ und unten $(f, o, h) = (0, 2, 0)$ gegeben.

[0077] In dem Detektionsdiagramm bestimmt nun die Entscheidungsstrategie aus der Information mehrerer Kameras C_i den Zustand $Z(v)$ des Anfragevolumens v . Kamerabasierte Entscheidungsstrategien sind vollständig im Detektionsdiagramm eintragbar bzw. ablesbar. [Fig. 14a](#)) zeigt ein Beispiel für sichere Überwachung bei Redundanz r als 1 Schwellenwert. [Fig. 14b](#)) zeigt ein Beispiel mit drei Schwellenwerten f_{\min} , o_{\min} sowie h_{\min} .

[0078] Je nach Strategie können nicht alle Detektionsvektoren eindeutig einem Zustand zugeordnet werden. In [Fig. 15](#) ist ein Beispiel mit einem erweiterten Zustand $Z^+(v)$ dargestellt. Erweiterte Zustände erhalten neben dem Zustand $Z(v)$ den Zustand „unbestimmt“ mit Unbestimmt: v ist nicht eindeutig einem der Zustände Frei, Objekt, Hindernis zuordbar.

[0079] Im Folgenden werden Möglichkeiten aufgezeigt für den Umgang mit Verdeckungen bzw. mit „Unbestimmten Zuständen“. (Die Figuren gehen beispielhaft von einem Differenzbildverfahren aus.)

[0080] Die erste Möglichkeit ist, alle Verdeckungen bzw. unbestimmte Zustände als mögliches Objekt aufzufassen. [Fig. 16](#) zeigt eine solche Situation für Verdeckungen $VB_1 \dots VB_4$. Dies hat den Vorteil, dass diese Situationen einfach zu realisieren sind. Allerdings dürfte der Nutzen gering sein, falls Schutzzonen SZ um Hindernisse zum Beispiel Roboter mit angeschmiegtten Verdeckungen liegen.

[0081] Die zweite Möglichkeit ist anzunehmen, dass alle Verdeckungen, welche vollständig innerhalb einer Schutzzone liegen, keine Objekte enthalten. [Fig. 17](#) zeigt eine solche Situation für Verdeckungen VB_1, VB_2 und Schutzzone SZ . Diese Verdeckungen VB werden dann als Freiraum gewertet. Diese Annahme ist auch für bewegte Hindernisse H einfach realisierbar. Allerdings ist diese Annahme nicht immer pauschal erfüllt.

[0082] Die nächste Möglichkeit ist anzunehmen, dass Verdeckungen v mit Durchmesser $d(v)$ kleiner als d_{\min} kein Objekt vollständig enthalten können ([Fig. 18](#)) Daraus folgt, dass Verdeckungen v als „Frei“ gewertet werden können, falls für die reflexiv-transitive Hülle $N \cdot (v)$ der verdeckten Nachbarvolumina gilt:

$$d(\cup_{v' \in N \cdot (v)} v') < d_{\min}$$

$$\text{mit } d(v) := \max\{\|v_1 - v_2\| \text{ mit } v_1, v_2 \in V\}$$

[0083] Eine weitere Möglichkeit ist anzunehmen, dass Verdeckungen und/oder unbestimmte Zustände ab oder innerhalb eines vorgegebenen Koordinatenabschnitts keine Objekte enthalten können. [Fig. 19](#) zeigt eine solche Situation unter der Annahme, dass Verdeckungen v ab der Höhe z_{\max} nicht von Objekten O erreichbar seien. Daraus folgt, dass Verdeckungen v als „Frei“ gewertet werden, falls gilt:

$$\min\{z \mid \exists x, y: (x, y, z) \in v\} \geq z_{\max}$$

[0084] Analog zur Höhe z kann bezüglich x oder y vorgegangen werden.

[0085] Das erfindungsgemäße Verfahren und das System zur Durchführung des Verfahrens ermöglicht den Einsatz von Kameras C_i für Schutzeinrichtungen der funktionalen Maschinensicherheit, die Arbeitsbereiche von Mensch und Maschine – vorzugsweise dreidimensional – überwachen, um den Menschen vor gefahrbringenden Maschinenbewegungen zu schützen, das heißt Kollisionen zu vermeiden.

[0086] Durch die Erfindung wird die Produktivität von Maschinen verbessert, zum Beispiel der vermehrte Einsatz von Robotern ohne trennenden Schutzraum, um eine direkte Mensch-Maschine-Kooperation zu ermöglichen.

[0087] Auch lassen sich Kosteneinsparungen realisieren, weil häufig auftretende Sicherheitsfunktionen wie „Zutrittsabsicherung“ und „Hintertretschutz“ die heute im Regelfall von zwei Schutzeinrichtungen ausgeführt werden, durch eine einzige Schutzeinrichtung realisieren lassen.

[0088] Auch sind mit dem erfindungsgemäßen System Verbesserungen der Wirksamkeit des Personenschutzes bei Maschinenbetrieb möglich. Bei den Verbesserungen der Personenschutzfunktion werden dabei ausdrücklich auch die Vermeidung des Manipulationsrisikos von Schutzeinrichtungen einbezogen und auch der Missbrauch von Maschinenfunktionalitäten eingeschlossen, zum Beispiel unerlaubterweise auf Förderbändern mitzufahren.

ZITATE ENTHALTEN IN DER BESCHREIBUNG

Diese Liste der vom Anmelder aufgeführten Dokumente wurde automatisiert erzeugt und ist ausschließlich zur besseren Information des Lesers aufgenommen. Die Liste ist nicht Bestandteil der deutschen Patent- bzw. Gebrauchsmusteranmeldung. Das DPMA übernimmt keinerlei Haftung für etwaige Fehler oder Auslassungen.

Zitierte Patentliteratur

- EP 1635108 A1 [0002, 0002]
- EP 1543270 B1 [0004]

Zitierte Nicht-Patentliteratur

- Ebert D., Henrich D.: "Safe Human-Robot-Cooperation: Imagebased Collision Detection for Industrial Robots", IEEE International Conference on Intelligent Robots and Systems, (IROS'02), Lausanne, 30. September bis 04. Oktober 2002 [0003]
- EN IC 61508-1/-7 [0021]

Patentansprüche

1. Verfahren zur Überwachung eines dreidimensionalen Raumbereichs (V), insbesondere einer Arbeitszelle einer automatisiert arbeitenden Maschine wie Roboter, bei dem Bilder (B_i) verwendet werden, die in zeitlicher Folge von wenigstens zwei, richtungs-diversitär bzw. richtungs-disimilar angeordneten Kameras ($C_1, C_2 \dots C_c$) erfasst wurden, deren Sichtstrahlen ($SI_{11}, SI_{12}; SI_{21}, SI_{22}; SI_{31}, SI_{32}$) sich wenigstens teilweise in einem Anfragevolumen (v) in dem Raumbereich (V) überschneiden und bei dem anhand einer Szenenanalyse der Bilder (B_i) in das Anfragevolumen (v) eindringende Objekte (O_j) und/oder gegebenenfalls Hindernisse (H_k) detektiert werden, **dadurch gekennzeichnet**, dass in jeder oder für jede Kamera ($C_1, C_2 \dots C_c$) eine unabhängige Szenenanalyse der Bilder (B_i) durchgeführt wird, dass Ergebnisse der unabhängigen Szenenanalysen der richtungs-diversitären Kameras ($C_1 \dots C_c$) redundant ausgewertet und dadurch ein oder mehrere fehlertolerante Überwachungssignale gebildet werden.

2. Verfahren nach Anspruch 1 oder 2 dadurch gekennzeichnet, dass eine Detektion eines Zustandes (Z) des Anfragevolumens (v) erfolgt, wobei überprüft wird, ob das Anfragevolumens (v) den Zustand „Frei“ aufweist, das heißt weder ein Objekt (O_j) noch ein Hindernis (H_k) enthält, den Zustand „Objekt“ aufweist, das heißt möglicherweise Teile von einem Objekt (O_j) enthält oder den Zustand „Hindernis“ aufweist, das heißt Teile eines Hindernisses (H_k) aber keine Objekte (O_j) enthält.

3. Verfahren nach zumindest einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, dass der Zustand (Z) des Anfragevolumens (v) als „Frei“ detektiert wird, wenn das Anfragevolumen (v) mit mindestens der für die sichere Überwachung geforderten Redundanz (r) von den Kameras (C_i) als „Frei“ von einem Objekt (O_j) erkannt wird.

4. Verfahren nach zumindest einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, dass die Hindernisse (H_k) definiert und in einem Auswertesystem eingelernt und/oder hinterlegt werden.

5. Verfahren nach zumindest einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, dass eine Analyse per Einzelbildverfahren durch Soll-/Ist-Vergleich fortlaufend erfolgt.

6. Verfahren nach zumindest einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, dass Bilder (B_i) von Anfragevolumen (v) bestimmter Kameras (C_i) als nicht redundant und/oder sicherheitstechnisch akzeptabel unberücksichtigt bleiben.

7. Verfahren nach zumindest einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, dass bei einer Kombination der Zustände (Z) einer bestimmten Anzahl von Anfragevolumina (v) eine Plausibilitätskontrolle erfolgt.

8. Verfahren nach zumindest einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, dass für jedes aufgenommene Bild (B_i) unter Berücksichtigung der räumlichen Anordnung der Kamera (C_i) der definierten Hindernisse (H_j) sowie des Anfragevolumens (v) eine Klassifikation von Pixeln (p) erfolgt.

9. Verfahren nach zumindest einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, dass anschließend eine dominante Klasse für das Anfragevolumen (v) bestimmt wird.

10. Verfahren nach zumindest einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, dass unter Berücksichtigung der definierten Redundanz (r) eine kamerabasierte Entscheidung mittels Detektionsdiagramm erfolgt.

11. Verfahren nach zumindest einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, dass unter Berücksichtigung von Randbedingungen wie z. B. Abständen eine anwendungsbasierte Auflösung unbestimmter Zustände (Z) vorgenommen wird.

12. Verfahren nach zumindest einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, dass die Ergebnisse von unabhängigen Szenenanalysen von zumindest drei, vorzugsweise vier Kameras ($C_1 \dots C_c$) redundant ausgewertet werden.

13. System zur Überwachung eines dreidimensionalen Raumbereichs (V), insbesondere einer Arbeitszelle einer automatisiert arbeitenden Maschine wie Roboter, mit zumindest zwei dezentral und richtungs-diversitär bzw. richtungs-disimilar angeordneten einzelnen Kameras ($C_1 \dots C_c$) zum Aufnehmen von zumindest zwei Bil-

den (B_1, B_2) des Raumes (V), wobei die Kameras versetzt zueinander angeordnet sind, und mit einer Auswerteeinheit die mittels einer Szenenanalyse der Bilder Objekte (O_j) und/oder Hindernisse (H_k) in dem Raum detektiert, dadurch gekennzeichnet, dass die Auswerteeinheit Mittel zur redundanten Auswertung von Ergebnissen von unabhängigen Szenenanalysen der einzelnen Kameras ($C_1 \dots C_n$) sowie Mittel zur Erzeugung zumindest eines sicherheitsrelevanten Überwachungssignals auf der Grundlage der Auswertung aufweist.

14. System nach Anspruch 13, dadurch gekennzeichnet, dass das System zumindest drei Kameras (C_1, C_2, C_3) aufweist.

Es folgen 10 Blatt Zeichnungen

Anhängende Zeichnungen

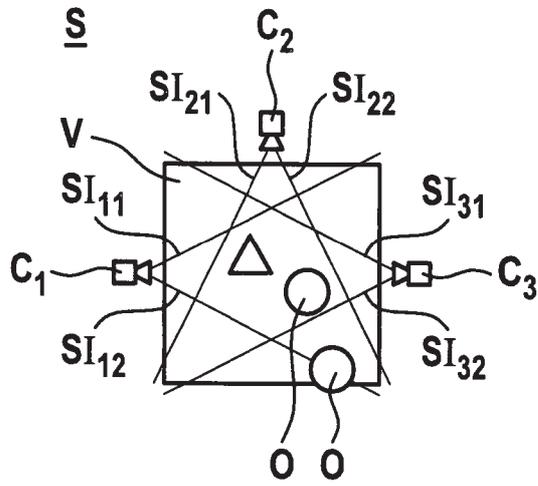


Fig. 1

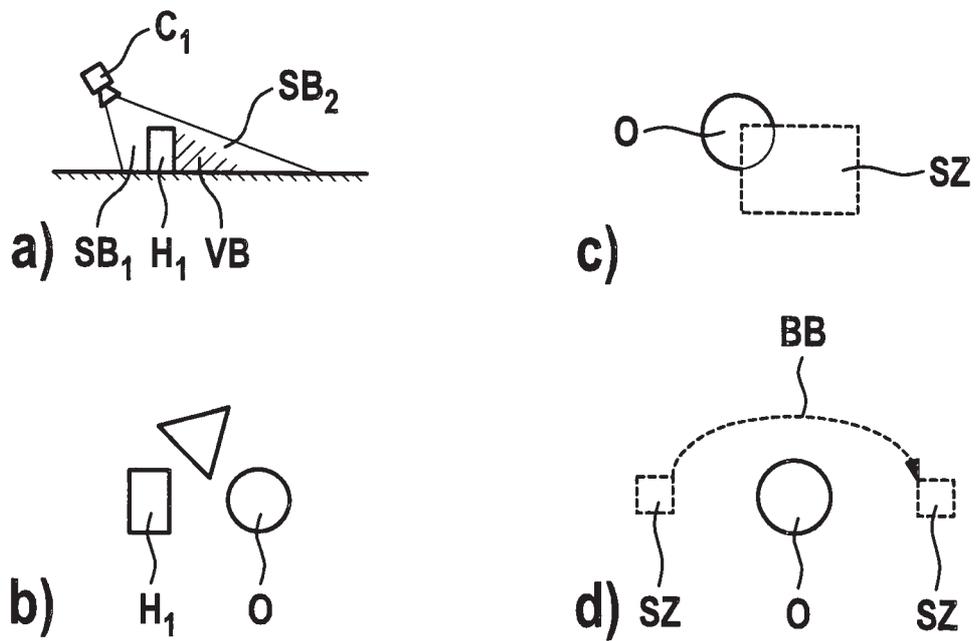


Fig. 2

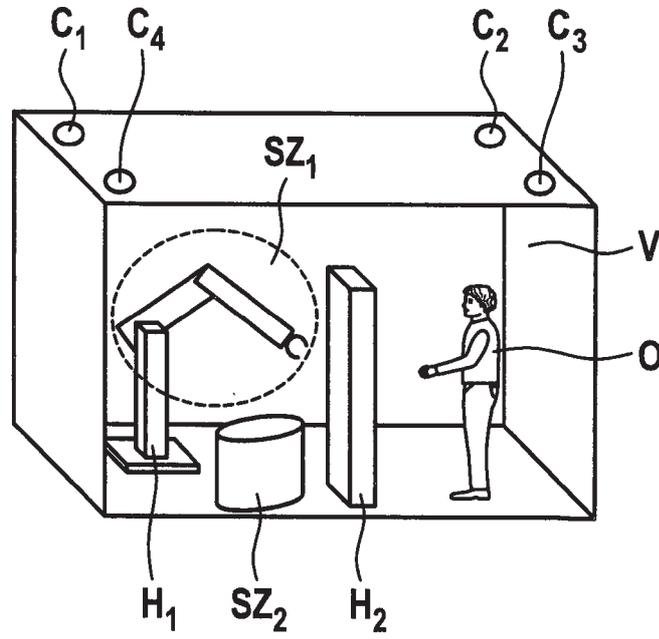


Fig. 3

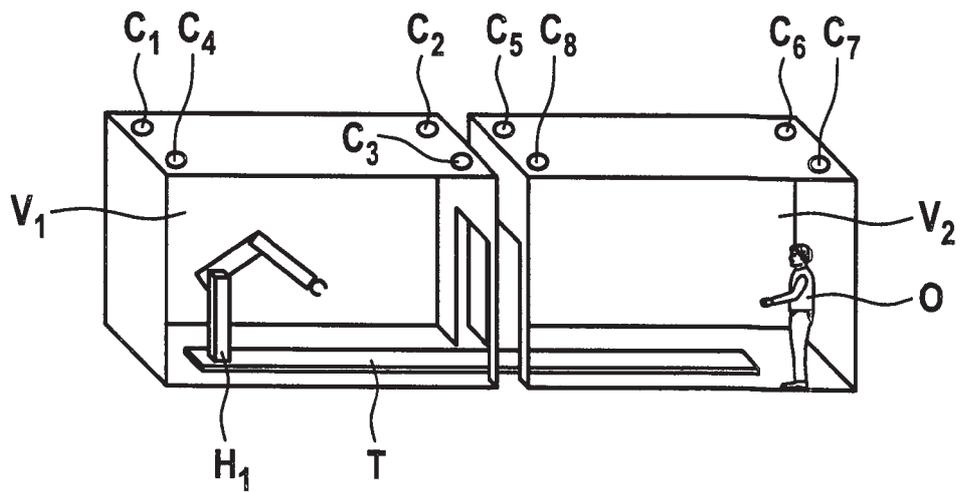


Fig. 4

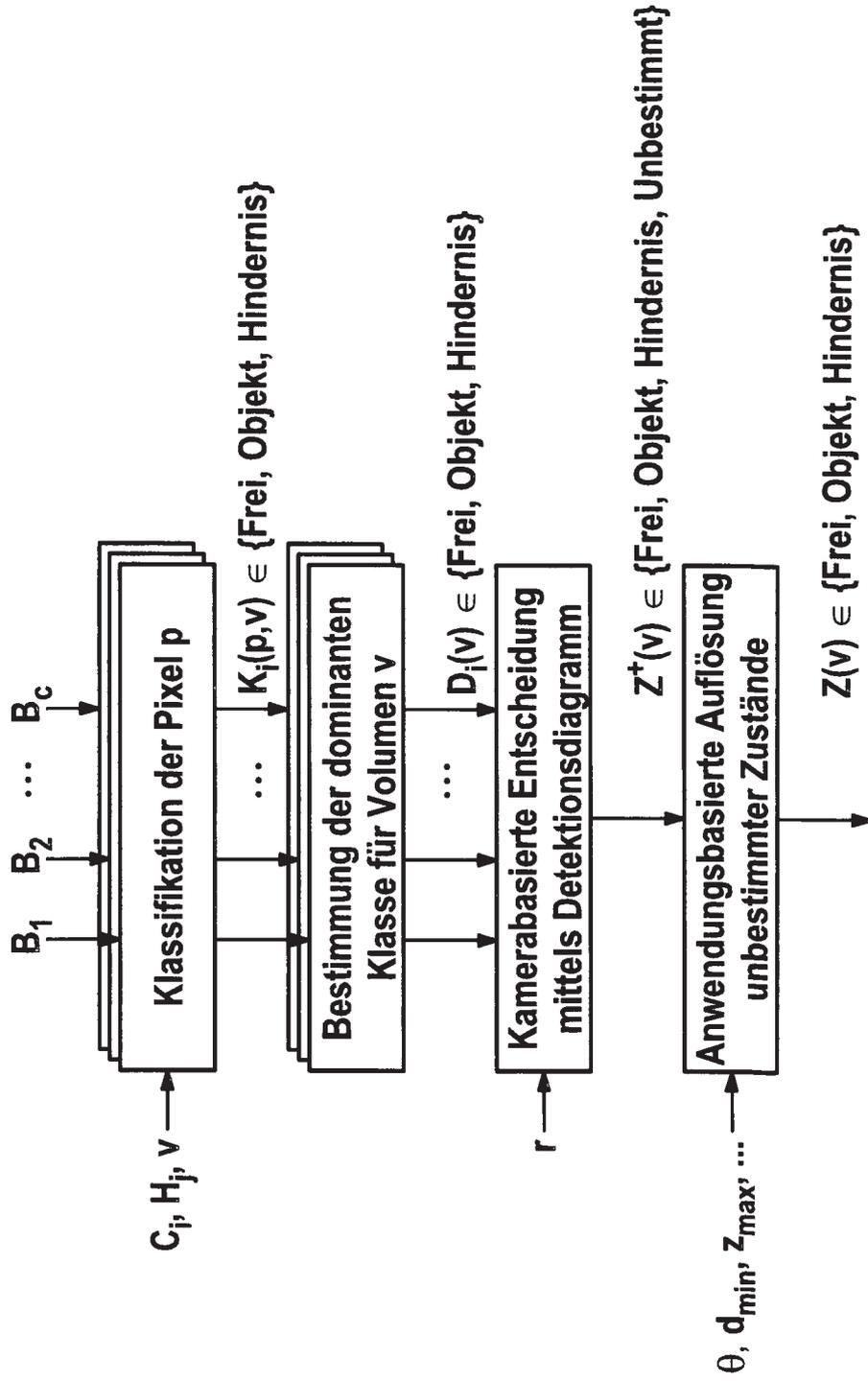


Fig. 5

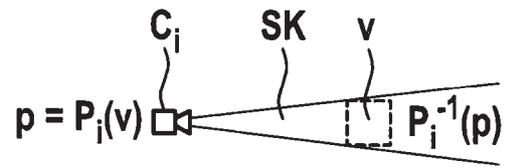


Fig. 6

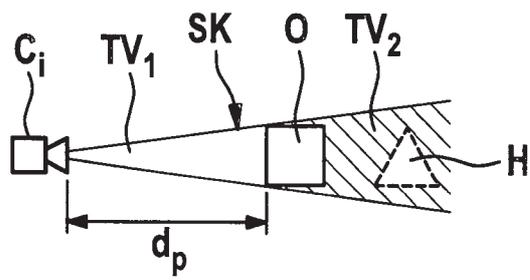
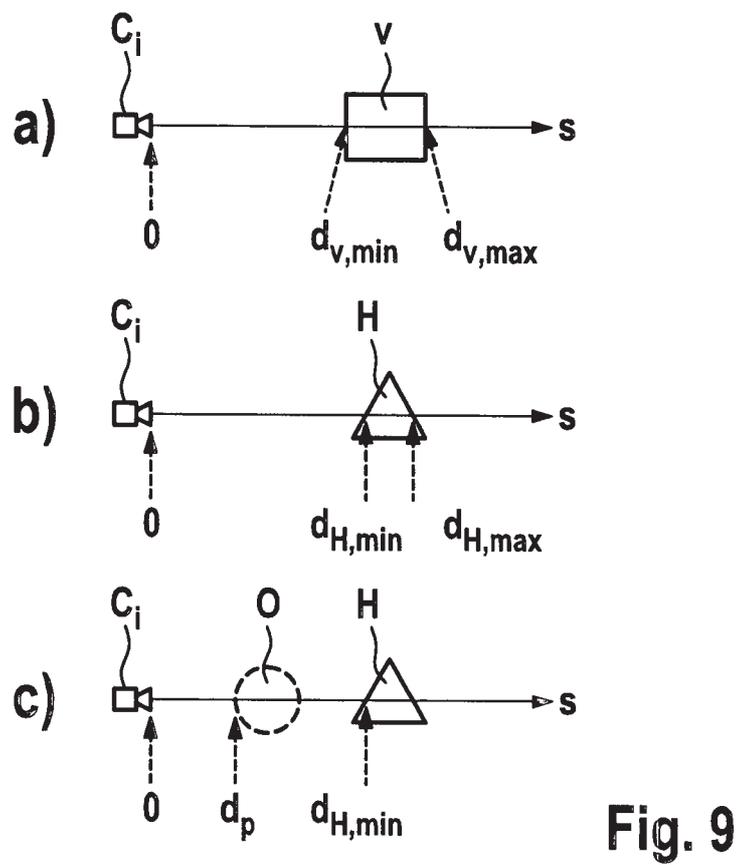
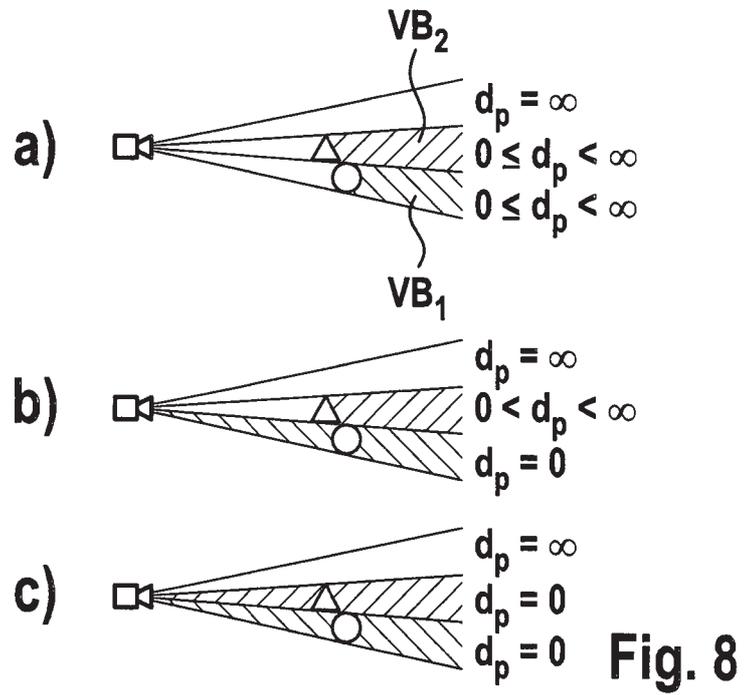


Fig. 7



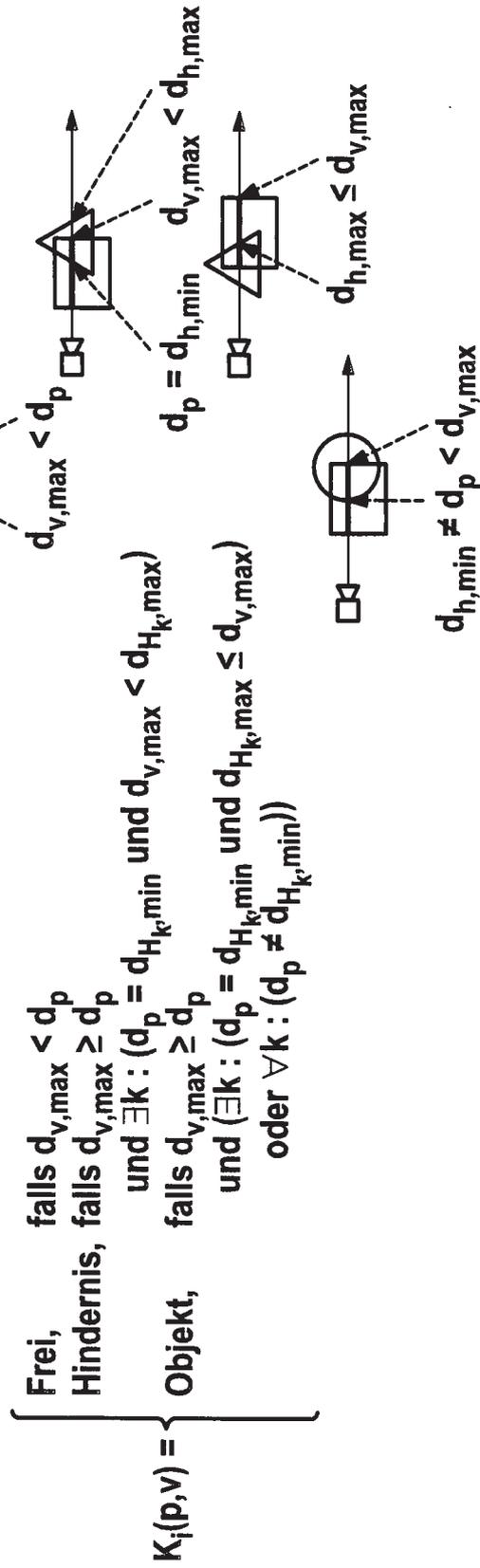


Fig. 10

$K_i(p,v):$	$D_i(v):$
Hindernis, Objekt	Objekt
Hindernis, Frei	Hindernis
Objekt, Frei	Objekt
Hindernis, Objekt, Frei	Objekt

Fig. 11

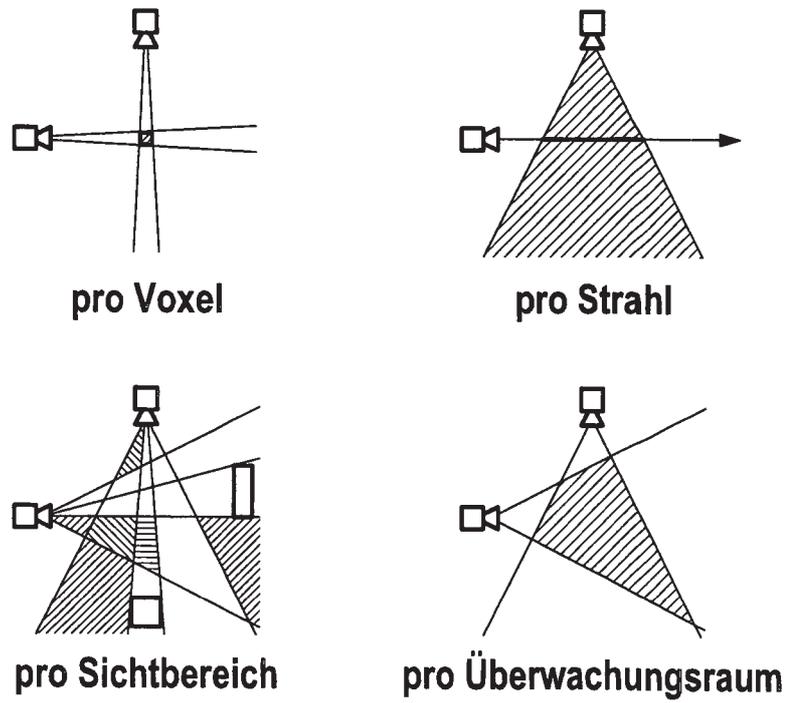
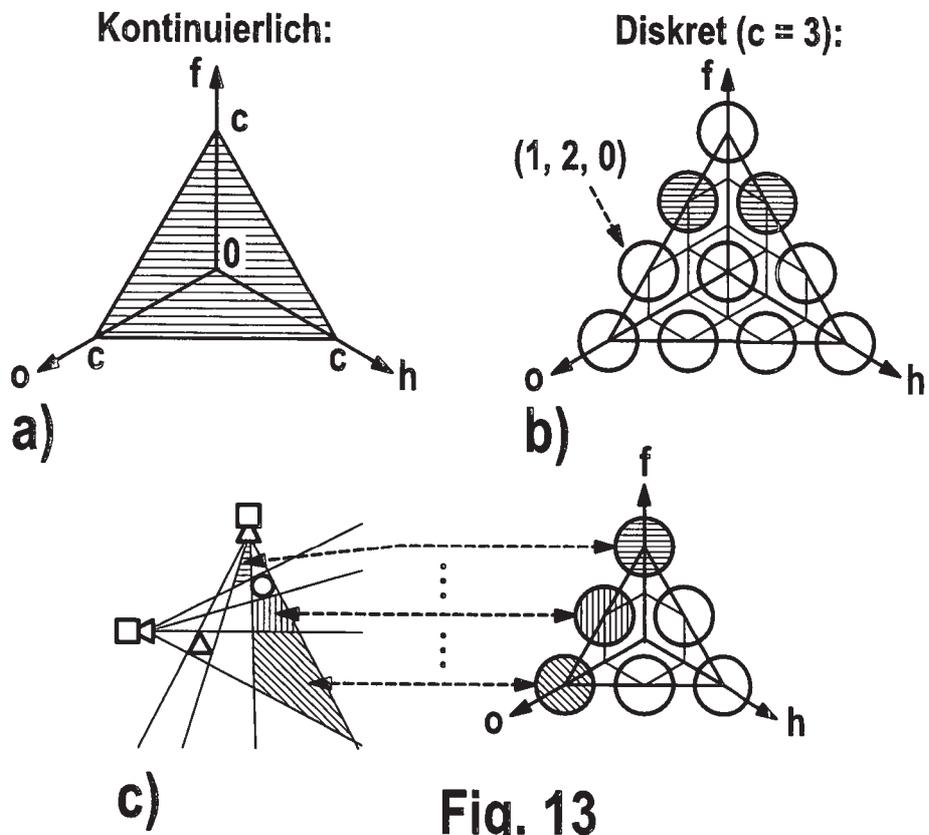


Fig. 12



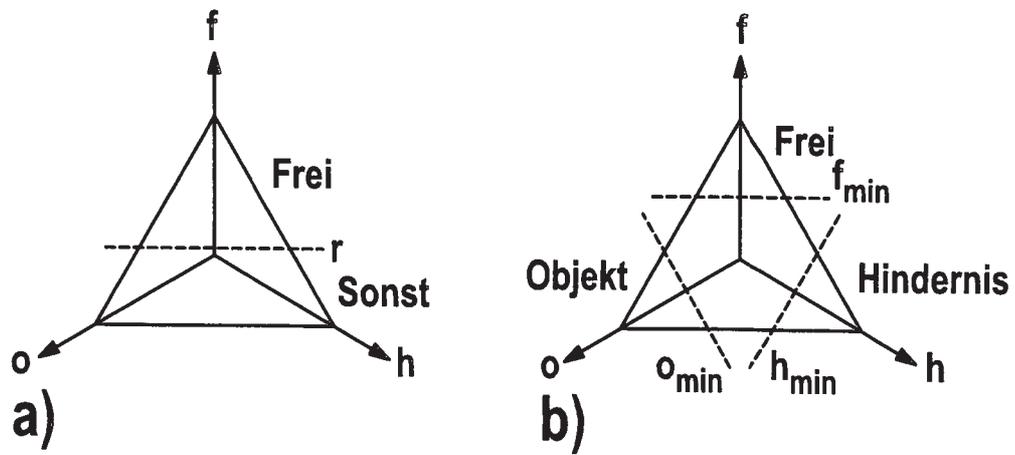


Fig. 14

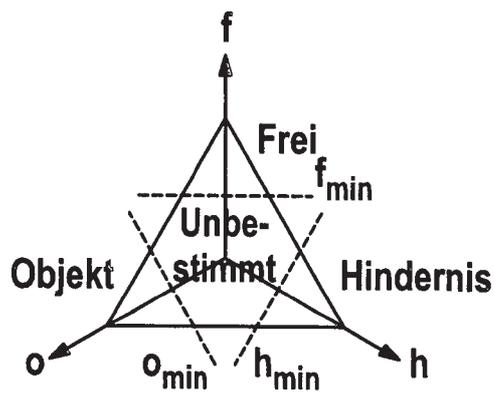


Fig. 15

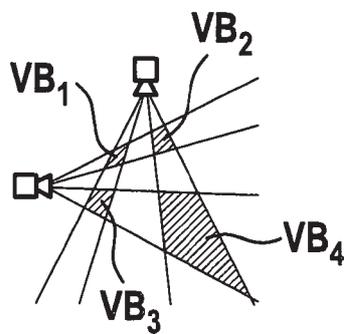


Fig. 16

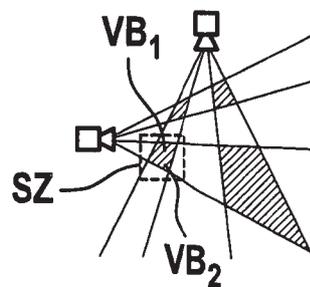


Fig. 17

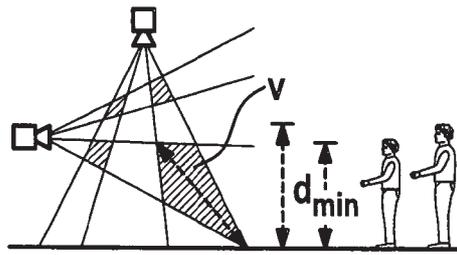


Fig. 18

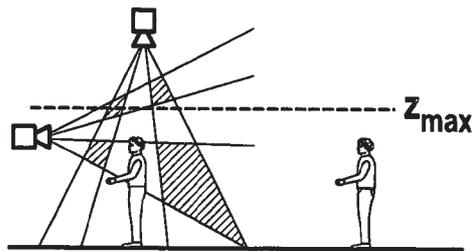


Fig. 19

DE 102009026091 A1

Anmeldeland:DE

Anmeldenummer:102009026091

Anmeldedatum:02.07.2009

Veröffentlichungsdatum:01.04.2010

Priorität:DE 102008030938 02.07.2008

Hauptklasse:B25J 19/06(2006.01,A)

MCD-Hauptklasse:B25J 19/06(2006.01,A)

ECLA:B25J 19/06

ECLA:B25J 21/00

ECLA:F16P 3/14

Entgegenhaltung (PL):EP 000001543270 B1

Entgegenhaltung (PL):EP 000001635108 A1

Entgegenhaltung (NPL):Ebert D., Henrich D.: "Safe Human-Robot-Cooperation: Imagebased Collision Detection for Industrial Robots", IEEE International Conference on Intelligent Robots and Systems, (IROS'02), Lausanne, 30. September bis 04. Oktober 2002

Entgegenhaltung (NPL):EN IC 61508-1/-7

Erfinder:Henrich, Dominik, Prof. Dr., 91126 Schwabach, DE

Anmelder:ELAN Schaltelemente GmbH & Co. KG, 35435 Wettenberg, DE

[EN]Method for monitoring three-dimensional spatial area, particularly work cell of automatic working machine, for e.g. robot, involves accomplishing independent scene analysis of images in or for each camera

[DE]Verfahren und System zur Überwachung eines dreidimensionalen Raumbereichs mit mehreren Kameras

[EN]

The method involves accomplishing an independent scene analysis of images in or for each camera (C-1 to C-3). The result of the independent scene analysis of the directional cameras is redundantly evaluated. One or multiple fault-tolerant monitoring signals are formed. An independent claim is included for a system for monitoring a three-dimensional spatial area, particularly a work cell of automatic working machine, for e.g. robot.

[DE]

Die Erfindung bezieht sich auf ein Verfahren zur Überwachung eines dreidimensionalen Raumbereichs (V), insbesondere einer Arbeitszelle einer automatisiert arbeitenden Maschine wie Roboter, bei dem Bilder (Bi) verwendet werden, die in zeitlicher Folge von wenigstens zwei, richtungsdiversitär bzw. richtungs-disimilar angeordneten Kameras (C1, C2 ... C) erfasst wurden, deren Sichtstrahlen (SI11, SI12; SI21, SI22; SI31, SI32) sich wenigstens teilweise in einem Anfragevolumen (v) in dem Raumbereich (V) überschneiden und bei dem anhand einer Szenenanalyse der Bilder (Bi) in das Anfragevolumen (v) eindringende Objekte (Oj) und/oder gegebenenfalls Hindernisse (Hk) detektiert werden. Um die Sicherheit zu erhöhen und um nicht einsehbare Bereiche hinter Hindernissen (Verdeckungen) auch bei der Überwachung von beliebigen Anfragevolumina zu minimieren wird vorgeschlagen, dass in jeder oder für jede Kamera (C1, C2 ... CC) eine unabhängige Szenenanalyse der Bilder (Bi) durchgeführt wird, dass Ergebnisse der unabhängigen Szenenanalysen der richtungsdiversitären Kameras (C1 ... CC) redundant ausgewertet und dadurch ein oder mehrere fehlertolerante Überwachungssignale gebildet werden.

Seite 1 --- ()

Seite 2 --- ()

[0001] Die Erfindung bezieht sich auf ein Verfahren zur Überwachung eines dreidimensionalen Raumbereichs, insbesondere einer Arbeitszelle einer automatisiert arbeitenden Maschine wie Roboter, bei dem Bilder verwendet werden, die in zeitlicher Folge von wenigstens zwei, richtungsdiversitär bzw. richtungs-disimilar angeordneten Kameras erfasst wurden, deren Sichtstrahlen sich wenigstens teilweise in einem Anfragevolumen in dem Raumbereich überschneiden und bei dem anhand einer Szenenanalyse der Bilder in das Anfragevolumen eindringende Objekte und/oder gegebenenfalls Hindernisse detektiert werden, sowie auf ein System zur Überwachung eines dreidimensionalen Raumbereichs, insbesondere einer Arbeitszelle einer automatisiert arbeitenden Maschine wie Roboter, mit zumindest zwei dezentral und richtungsdiversitär bzw. richtungs-disimilar angeordneten einzelnen Kameras zum Aufnehmen von zumindest zwei Bildern des Raumes, wobei die Kameras versetzt zueinander angeordnet sind, und mit einer Auswerteeinheit die mittels einer Szenenanalyse der Bilder Objekte und/oder Hindernisse in dem Raum detektiert.

[0002] Ein Verfahren sowie ein System der eingangs genannten Art ist in der EP 1 635 108 A1 beschrieben. Die EP 1 635 108 A1 beschreibt ein Verfahren und eine Vorrichtung zum Erfassen eines Objektes innerhalb eines Überwachungsbereichs, bei dem zumindest zwei Sensoren jeweils mindestens einen Teil des Überwachungsbereichs überwachen. Jeder Sensor erzeugt ein digitales Einzelbild des gesamten Überwachungsbereichs, wobei die Einzelbilder aller Sensoren in einem einheitlichen Koordinatensystem erzeugt werden. Von jedem Sensor werden diejenigen Koordinatenpunkte der Einzelabbilder als "sicher" gekennzeichnet, die er als frei von dem Objekt erkennt, wobei alle anderen Koordinatenpunkte der Einzelabbilder als "unsicher" gekennzeichnet werden. Die derart erzeugten Einzelabbilder werden jeweils zu einer Auswerteeinheit übertragen und die übertragenen Einzelabbilder werden von der Auswerteeinheit entsprechend ihren gemeinsamen Koordinatensystemen zu einem resultierenden Gesamtbild überlagert, in dem diejenigen Koordinatenpunkte des Gesamtbildes als "sicher" gekennzeichnet werden, die zumindest in einem der von den Sensoren übertragenen Einzelabbildern als "sicher" gekennzeichnet sind, während alle anderen Koordinatenpunkte des Gesamtbildes als "unsicher" gekennzeichnet werden. Durch das Verfahren soll eine genauere Erfassung von Objekten möglich sein, insbesondere soll bei der Erfassung eines Objektes auch der "Schattenbereich" des Objektes als solcher erkannt und als "sicher" eingestuft werden können, so lange wenigstens ein Sensor diesen Schattenbereich einsehen kann. Eine redundante Auswertung der Sensoren ist in dieser Druckschrift nicht angesprochen.

[0003] In einem Dokument von Ebert D., Henrich D.: "Safe Human-Robot-Cooperation: Imagebased Collision Detection for Industrial Robots", IEEE International Conference on Intelligent Robots and Systems, (IROS'02), Lausanne, 30. September bis 04. Oktober 2002 wird die Überwachung eines dreidimensionalen Arbeitsraumes mittels mehrerer stationärer Videokameras beschrieben. In dem zu überwachenden

Arbeitsraum befindet sich ein Roboter, bekannte statische Hindernisse sowie zu detektierende, dynamische, 3-dimensionale Hindernisse, die sich vom Hintergrund unterscheiden. Ein Lösungsansatz zur bildbasierten Kollisionserkennung zwischen dynamischen Hindernissen und einer beliebigen Testkonfiguration durch Roboter wird beschrieben, ohne dass allerdings sicherheitsgerichtete Aspekte betreffend die Auswertung der Videokameras angesprochen werden.

[0004] Die EP 1 543 270 B1 bezieht sich auf ein Verfahren und eine Vorrichtung zum Absichern eines Gefahrenbereichs, insbesondere des Gefahrenbereichs einer automatisiert arbeitenden Maschine, bei dem mit zumindest einer ersten und einer zweiten Bildaufnahmeeinheit zumindest zwei Bilder des Gefahrenbereichs aufgenommen werden, wobei die Bildaufnahmeeinheiten versetzt zueinander angeordnet sind und bei den anhand einer Sequenzanalyse der Bilder Fremdobjekte in dem Gefahrenbereich detektiert werden. Der dort beschriebenen Lösung liegt der Gedanke zugrunde, eine dreidimensionale Auswertung von mindestens zwei Bildern auszuführen, die mit zu einander versetzten Bildaufnahmeeinheiten von dem Gefahrenbereich aufgenommen werden. Es erfolgt eine Szenenanalyse anhand der mindestens zwei Bilder unter Anwendung mindestens zweier algorithmisch unterschiedlicher Verfahren. Der Gefahrenbereich wird abgesichert, wenn mindestens eines der Analyseverfahren eine Fremdobjektdetektion liefert. Bei diesem Verfahren sind die Kameras räumlich nicht diversitär angeordnet. Eine Auswertung eines Raumvolumens ist ebenfalls nicht angesprochen. Bei dem Verfahren werden Kamerasysteme eingesetzt, welche eine bauliche Einheit mit drei bzw. vier zentral angeordneten Kameras bilden. Die zentrale Anordnung der Kameras ist dabei notwendig, um mittels der Stereoskopie Entfernungen ermitteln zu können. “Blicken” solche Kamerasysteme jedoch auf ein Hindernis im Überwachungsraum, kann der dahinter liegende, verdeckte Raum nicht mehr überwacht werden. Daraus ergeben sich Verdeckungen (“Tot-Zonen”), die entweder den Systemnutzen einschränken, beispielsweise frei Schutzräume bzw. Schutzräume ohne Hindernisse erfordern oder den Einsatz trotz allem nicht angezeigt erscheinen lassen. Eine Kaskadierung ist dagegen unter Kostengesichtspunkten kritisch.

Seite 3 --- ()

[0005] Weitere Ansätze, die auf einen sicherheitsgerichteten Einsatz von Kameras bzw. Kamerasystemen beziehen, jedoch eine zweidimensionale Arbeitsweise zugrunde legen, existieren beispielsweise auf dem Gebiet der Normung “Vision based protective devices”. Allerdings können derartige Systeme nicht ohne Weiteres einen Raum dreidimensional überwachen bzw. verdeckte Bereiche einsehen.

[0006] Davon ausgehend liegt der Erfindung die Aufgabe zugrunde, ein Verfahren und ein System der eingangs genannten Art derart weiterzubilden, dass die Sicherheit erhöht wird und nicht einsehbare Bereiche hinter Hindernissen (Verdeckungen) minimiert werden. Auch soll eine sichere Raumüberwachung für beliebige Anfragevolumina ermöglicht werden.

[0007] Die Aufgabe wird erfindungsgemäß dadurch gelöst, dass in jeder oder für jede Kamera eine unabhängige Szenenanalyse der Bilder durchgeführt wird, dass Ergebnisse der unabhängigen Szenenanalysen der richtungs-diversitären Kameras redundant ausgewertet und dadurch ein oder mehrere fehlertolerante Überwachungssignale gebildet werden.

[0008] Die Erfindung basiert auf der Idee, zumindest zwei vorzugsweise zumindest drei einzelne Kameras dezentral und dabei richtungs-diversitär (bzw. richtungs-disimilar) zur Überwachung des dreidimensionalen Raumbereichs anzuordnen. Dadurch können Verdeckungen, das heißt nicht einsehbare Bereiche hinter Hindernissen minimiert bzw. besser eingesehen werden. In bevorzugter Ausführungsform kann ein Sichtstrahl einzelner oder mehrerer Kameras aufgeteilt werden.

[0009] Zunächst erfolgt eine Kalibrierung der Kameras, sowohl bezogen auf deren intrinsische Parameter (z. B. Verzerrung, Fokus) als auch bezogen auf extrinsische Parameter (z. B. Ort und Richtung ihrer Installation). Dadurch lässt sich relativ genau bestimmen, wo sich Sichtstrahlen oder aufgeteilte Sichtstrahlen der einzelnen Kameras im dreidimensionalen Raum kreuzen bzw. überlappen. Ergänzend kann noch ermittelt werden, welche Sichtstrahlen und wie viele Sichtstrahlen sich kreuzen bzw. überlappen. Zusätzlich wird der zu überwachende Raum in frei wählbare so genannte Anfragevolumina unterteilt, welche zwar vorzugsweise konvex aber nicht notwendigerweise disjunkt sind oder den Überwachungsraum vollständig abdecken.

[0010] Sofern eine bestimmte Anzahl an Kameras ein freies Anfragevolumen “sieht”, kann angenommen werden, dass der betreffende Raum frei ist. Sofern darin Hindernisse “erlaubt” bzw. gegeben vorhanden sind, vermindern sich so durch die richtungs-diversitäre Anordnung der Kameras die verdeckten Bereiche in einem Überwachungsraum.

[0011] Des Weiteren zeichnet sich die Erfindung durch einen eigenerfinderischen Gedanken dadurch aus, dass nicht alle Kameras das Anfragevolumen gleichermaßen “frei” sehen müssen, sondern ein Voting beispielsweise über eine Auswahlhaltung möglich ist, wenn beispielsweise aus Sicht einer Kamera das Anfragevolumen hinter einem Hindernis liegt und somit verdeckt ist. In diesem Fall müssen nun seitens der anderen Kameras konsistente Informationen in redundanter bzw. sicherheitstechnisch akzeptabler Form vorliegen, zum Beispiel eine Auswahl von drei (zwei out of drei, 2 o o 3) um in der bzw. in den nachfolgenden Stufe(n) der Signalverarbeitung der Auswerteelektronik eine Anfrageeinheit dennoch als “frei” zu bewerten, andernfalls würde ein Signal zur Warnung oder Abschaltung oder Umsteuerung der gefahrbringenden Bewegung gegeben werden.

[0012] Vorzugsweise werden Hindernisse wie statische oder dynamische Hindernisse vor Inbetriebnahme des Systems in dieses “eingelernt” oder in diesem hinterlegt, so dass der den Sichtbereich einer Kamera beschränkende verdeckte Raum und die somit dahinter liegenden Verdeckungsbereiche dem System bekannt sind. Somit kann fortlaufend, vorzugsweise per Einzelbildverfahren, eine Art von Soll-/Ist-Vergleich erfolgen, der in das System-Voting intelligent eingebunden werden kann. Die Hindernisse, definiert bzw. interpretiert als erlaubte Objekte im Raum, können dabei nahezu beliebige Geometrie haben.

[0013] Ferner besteht gemäß einer bevorzugten Verfahrensweise die Möglichkeit, im Rahmen des Votings Bilder von Anfragevolumina bestimmter Kameras als nicht redundant bzw. sicherheitstechnisch akzeptabel auszuklammern, zum Beispiel wenn die Richtungs-Diversität der Kameras nicht gegeben wäre.

[0014] Eine weitere bevorzugte Ausführungsform der Erfindung mit eigenständigem Erfindungsgedanken besteht darin, Zustände einer bestimmten Anzahl von Anfragevolumina zu kombinieren bzw. zusammen zu fügen und aus der Kombination (aus dem Zusammenfügen) der Zustände einer bestimmten Anzahl von Anfragevolumina Plausibilitätskontrollen, das heißt Schutzfeldanalysen vorzunehmen.

Seite 4 --- ()

[0015] So können beispielsweise einzelne im Raum schwebende Anfragevolumina, die als nicht “frei” erkannt werden, dennoch toleriert und – bedingt durch eine Störung – als unkritisch uminterpretiert werden, wenn es sich beispielsweise aufgrund der geometrischen Anordnung nicht um einen Menschen handeln kann.

[0016] Ebenso erlaubt diese Verfahrensweise auch eine Mustererkennung bzw. Texturerkennung der Plausibilitätskontrolle, zum Beispiel um im Rahmen einer Muting-Funktion die Schutzfunktion der Schutzeinrichtung zeitlich begrenzt und automatisch zu überbrücken. So können beispielsweise Waren in einem gefahrbringenden Bereich hinein – oder aus ihm heraus – befördert werden, ohne dass die Schutzeinrichtung die Schutzfunktion auslöst.

[0017] Auch besteht die Möglichkeit Menschen zu detektieren, die unerlaubter Weise auf einem Transportband mitfahren. Auch in diesem Fall würden “erlaubte” und/oder “nicht erlaubte” Geometrien dem System vorher eingelernt.

[0018] Die besonders bevorzugten Aspekte des Verfahrens können wie folgt zusammengefasst werden: – Mehrere kalibrierte Kameras (z. B. Tiefen- oder Intensitätskameras) beobachten von unterschiedlichen Richtungen ein Überwachungsvolumen (Anfragevolumen).– Es liegen ausreichend Informationen über die Verdeckungsbereiche im Überwachungsvolumen vor (z. B. der Ort und die Geometrie der Hindernisse oder die maximale Größe der Verdeckungen).– Aus den aktuellen Kamerabildern und der Verdeckungsinformation können unbekannte Objekte in beliebiger Anzahl sicher detektiert werden.– Die Detektion der Objekte wird dadurch sicher, dass eine vorgegebene Anzahl von Fehlern durch redundante Auswertung der Kamerabilder kompensiert werden (z. B. pro Voxel, Strahl, Sichtbereich, Überwachungsraum).

[0019] Das Verfahren kann zur Raumrekonstruktion, Sicht-/Verdeckungsanalyse, Schutzzonenüberwachung, Kollisionserkennung, Kollisionsvermeidung oder als virtueller Schalter angewendet werden.

[0020] Ferner ist das Verfahren auf mehrere Überwachungsräume, welche jeweils von einer Menge an Kameras überwacht werden, erweiterbar. Diese Mengen können diskunkt, teilweise oder vollständig überlappend sein.

[0021] Die Aufgabe wird erfindungsgemäß durch ein System dadurch gelöst, dass die Auswerteeinheit Mittel zur redundanten Auswertung von Ergebnissen von unabhängigen Szenenanalysen der einzelnen Kameras sowie Mittel zur Erzeugung zumindest eines sicherheitsrelevanten Überwachungssignals auf der Grundlage der Auswertung aufweist. Das System besteht dabei aus einer Anzahl von zumindest zwei, vorzugsweise zumindest drei Kameras mit zugehöriger Auswerteelektronik. Sicherheitstechnisch soll dabei eine Hardwarefehltoleranz von mindestens 1 und alle weiteren Maßnahmen realisiert werden, die zu einem Safety-Integrity-Level von mindestens 2 führen. (Siehe hierzu auch EN IC 61508-1/-7.)

[0022] In dem System können sowohl handelsübliche Kameras als auch speziell ertüchtigte Kameras eingesetzt werden.

[0023] Weitere Einzelheiten, Vorteile und Merkmale der Erfindung ergeben sich nicht nur aus den Ansprüchen, den diesen zu entnehmenden Merkmalen – für sich und/oder in Kombination –, sondern auch aus der nachfolgenden Beschreibung von den Zeichnungen zu entnehmenden bevorzugten Ausführungsformen.

[0024] Es zeigen:

[0025] Fig. 1 eine schematische Darstellung eines Systems zur sicherheitsgerichteten Überwachung eines Raumes umfassend Kameras, Objekte sowie Hindernisse,

[0026] Fig. 2a)–e) eine schematische Darstellung verschiedener Anwendungen wie Sicht-/Verdeckungsanalyse, Raumrekonstruktion, Zonenüberwachung, sowie Kollisionsvermeidung,

[0027] Fig. 3 eine schematische Darstellung einer Arbeitszelle mit Industrieroboter, statischen Hindernissen, dynamischen Hindernissen sowie unbekanntem Objekten (z. B. Mensch),

[0028] Fig. 4 eine schematische Darstellung mehrerer Überwachungsräume mit mobilem Industrieroboter,

Seite 5 --- ()

[0029] Fig. 5 ein Ablaufdiagramm des erfindungsgemäßen Verfahrens,

[0030] Fig. 6 Detektion mit einer Kamera: (Rück-)Projektion,

[0031] Fig. 7 Detektion mit einer Kamera: verallgemeinerte Abstandsbilder,

[0032] Fig. 8a)–c) Detektion mit einer Kamera: abgedeckte Technologien wie Abstandsbilder, Hybridbilder sowie Differenzbilder,

[0033] Fig. 9a)–c) Detektion mit einer Kamera: Definition verschiedener Abstände zur Kamera,

[0034] Fig. 10 Detektion mit einer Kamera: Klassifikation von Pixeln,

[0035] Fig. 11 Tabelle,

[0036] Fig. 12 Detektion mit mehreren Kameras: Formen der Redundanz,

[0037] Fig. 13a), b) Detektion mit mehreren Kameras: Definition von Detektionsdiagrammen: (a) Kontinuierliche Darstellung, (b) Diskrete Darstellung

[0038] Fig. 13c) Detektion mit mehreren Kameras: Beispiel eines Detektionsdiagramms

[0039] Fig. 14a), b) Detektion mit mehreren Kameras: Beispiele für Entscheidungsstrategien,

[0040] Fig. 15 Detektion mit mehreren Kameras: Der Zustand “unbestimmt”,,

[0041] Fig. 16 Umgang mit Verdeckungen: Verdeckung als Objekt,

[0042] Fig. 17 Umgang mit Verdeckungen: Maskierung,

[0043] Fig. 18 Umgang mit Verdeckungen: minimaler Durchmesser, und

[0044] Fig. 19 Umgang mit Verdeckungen: Koordinaten-Abschnitt.

[0045] Fig. 1 zeigt rein schematisch ein System S zur sicherheitsgerichteten Überwachung eines Volumens V im dreidimensionalen Raum, insbesondere einer Arbeitszelle durch mehrere Kameras C_i mit $i = 1 \dots c$. Im vorliegenden Ausführungsbeispiel sind drei kalibrierte Kameras C_1, C_2, C_3 , beispielsweise Zeilen-, Grauwert-, Farb- oder Tiefenkamera ausgebildet, die an einem bekannten Ort und mit bekannter Orientierung (extrinsische Parameter) sowie Fokus und Verzerrung (intrinsische Parameter) angeordnet sind.

[0046] Das überwachte Volumen $V \subseteq R_3$ wird teilweise von mehreren Kameras C_i überwacht. In dem Volumen V befinden sich opake Objekte O_i wie Menschen mit unbekannter Anzahl und Geometrie. Gegebenenfalls befinden sich in dem Volumen V noch vorzugsweise opake, konvexe Hindernisse H_k wie Pfeiler, Industrieroboter mit zu jedem Zeitpunkt in dem Volumen V bekannter Geometrie.

[0047] Jede Kamera C_i liefert ein aufgenommenes (Teil-)Bild B_i des Volumens V gegebenenfalls mit detektierten Objekten O_i bzw. Hindernissen H_k .

[0048] Eine für die sichere Überwachung geforderte Redundanz r ist vorgegeben. Aufgrund der zuvor erfolgten Kalibrierung der Kameras C_i lässt sich relativ genau bestimmen, wo sich Sichtstrahlen $SI_{11}, SI_{12}, SI_{21}, SI_{22}, SI_{31}, SI_{32}$ im dreidimensionalen Raum V kreuzen bzw. überlappen.

[0049] Mittels des Systems S erfolgt für alle konvexen Anfragevolumina $v \subseteq V$; V eine sichere Detektion des Zustandes $Z(V)$ aus der Menge der möglichen Zustände {Frei, Objekte, Hindernis} mit

Frei: v enthält weder O_i noch H_k

Objekt: v enthält möglicherweise Teile von O_j

Hindernis: v enthält Teile von H_k aber keine O_i

[0050] Die Detektion ist sicher, falls Zustand $Z(v)$ nur dann frei ist, wenn das Anfragevolumina v mit mindes

Seite 6 --- ()

stens r -facher Redundanz von den Kameras C_i als “frei”; von Objekten O_i erkannt wird.

[0051] Mit dem oben genannten Ansatz sind verschieden Anwendungen denkbar. Dazu gehören die Raumrekonstruktion, die Zonenüberwachung und die Kollisionserkennung. Fig. 2a) zeigt ein Beispiel der Raumrekonstruktion, bei welcher mit einer Kamera C_1 sowie einem Hindernis H_1 , der von

der Kamera C_1 überwachte Bereich in Sichtbereich SB_1, SB_2 sowie in einen Verdeckungsgebiet VB unterteilt werden kann. Ein weiteres Beispiel der Raumrekonstruktion ist die Bestimmung der möglichen Aufenthaltsorte von eindringenden Objekten O (**Fig. 2b**).

[0052] In **Fig. 2c** ist ein Beispiel der Zonenüberwachung dargestellt, wobei das Eindringen von (möglichen) Objekten O in eine Warn- bzw. Schutzzone SZ dargestellt ist. Das Eindringen des Objektes O (z. B. eine Hand) in die Schutzzone SZ kann beispielsweise auch als das Betätigen eines virtuellen Schalters interpretiert werden.

[0053] **Fig. 2d** zeigt ein Beispiel der Kollisionsvermeidung, wobei eine Schutzzone SZ , beispielsweise ein Roboterarm, auf einer Bewegungsbahn BB kollisionsfrei um ein Objekt O zu einem Zielpunkt bewegt wird.

[0054] **Fig. 3** zeigt ebenfalls rein schematisch ein dreidimensionales Volumen V in Form einer Arbeitszelle, umfassend ein beispielsweise bewegtes Hindernis H_1 wie Industrieroboter mit einem ebenfalls bewegten Hindernis evtl. als Schutzzone SZ_1, SZ_2 sowie bekannte, statische Hindernisse H_2 in Form von Säulen, Wänden, Schränken und Tischen. Der Zugang des Objektes O wie Mensch kann mittels der Kameras C_1, C_2, C_3, C_4 erfasst werden.

[0055] In **Fig. 4** ist ebenfalls rein schematisch ein System zur sicherheitsgerichteten Überwachung mehrerer Überwachungsräume V_1, V_2 dargestellt, wobei zum Beispiel ein Hindernis H_1 wie mobiler Industrieroboter über eine Transportvorrichtung T zwischen den Überwachungsräumen V_1, V_2 transportierbar ist. Der Überwachungsraum V_1 wird durch die Kameras C_1 bis C_4 und der Überwachungsraum V_2 durch die Kameras C_5 bis C_8 überwacht.

[0056] **Fig. 5** zeigt ebenfalls rein schematisch ein Ablaufdiagramm des erfindungsgemäßen Verfahrens. Wie bereits zuvor erwähnt erfolgt zunächst eine Kalibrierung der Kameras C_i sowohl bezogen auf deren intrinsischen und extrinsischen Parameter. Weiterhin werden vorab Ort und Richtung, Hindernisse H_j sowie das Anfragevolumen v definiert. Jede Kamera C_i liefert ein Bild bzw. Teilbild B_i des Volumens V mit detektierten Objekten O bzw. Hindernissen H . In einem ersten Verfahrensschritt erfolgt eine Klassifikation der Pixel p . Die Klassifikation erfolgt für das Anfragevolumen jedes Bild B_1, \dots, B_c separat. Mit dem Ergebnis der Klassifikationen $K_i(p, v)$ ∈ {Frei, Objekt, Hindernis} wird die dominante Klasse $D_i(v)$ ∈ {Frei, Objekt, Hindernis} für das Anfragevolumen v bestimmt. Unter Berücksichtigung der Redundanz r erfolgt sodann eine kamerabasierte Entscheidung mittels eines Detektionsdiagramms, d. h. die Detektion des Zustands $Z+(v)$ ∈ {Frei, Objekt, Hindernis, Unbestimmt}.

[0057] Unter Berücksichtigung von weiteren Parameter wie Verdeckungsanzahl $\&thgr;$, Abstände d_{min}, z_{max} usw. erfolgt dann eine anwendungsorientierte Auflösung unbestimmter Zustände. Als Ergebnis wird der Zustand $Z(v)$ ∈ {Frei, Objekt, Hindernis} zur Verfügung gestellt.

[0058] Zur Erläuterung ist die Detektion eines Volumens mit einer Kamera in **Fig. 6** mittels (Rück-)Projektion erläutert. Für ein Volumen v enthält die Projektion $P_i(v)$ in der Kamera C_i alle Pixel p ∈ B_i auf welche das Volumen v durch eine Optik der Kamera C_i abgebildet wird. Für Pixel p ∈ B_i der Kamera C_i ist die Rückprojektion $P_{i, \‐1}(p)$ die Vereinigung aller Teilvolumina v' mit $p = P_i(v')$. $P_{i, \‐1}(p)$ wird Sichtkegel SK von p genannt. Es gilt immer $v \⊂ P_{i, \‐1}(P_i(v))$ und häufig auch noch $v \⁣ P_{i, \‐1}(P_i(v))$.

[0059] Ferner können die Bilder der Kamera C_i auf allgemeine Abstandsbilder generalisiert werden, wie dies in **Fig. 7** dargestellt ist. In einem verallgemeinerten Abstandsbild enthält ein Kamerabild B_i pro Pixel p ∈ B_i den Abstand d_p von Kamera C_i zu dem nächsten detektierbaren Objekt O bzw. Hindernis H innerhalb des Sichtkegels SK von Pixel p . Daraus folgt, dass ein Teilvolumen TV im Sichtkegel SK von Kamera C_i bis d_p frei ist, während ein Teilvolumen TV_2 im Sichtkegel SK ab d_p verdeckt ist und gegebenenfalls weitere Objekte O bzw. Hindernisse H enthalten kann.

[0060] Die **Fig. 8a** bis **c**) zeigen Beispiele von Technologien, die mit dem verallgemeinerten Abstand d_p pro Pixel p ∈ B_i erfasst werden können: Abstandsbilder aus Tiefenkameras (**Fig. 8a**) mit Abstand $0 \leq d_p < \infty$; zu Hindernissen H bzw. eventuell Objekten

Seite 7 --- ()

O und $d_p = \infty$; ansonsten.

[0061] Hybridbilder aus Intensitätskameras (**Fig. 8b**) mit Abstand $d_p = 0$ für Objekte (Vordergrund) bzw. $0 < d_p < \infty$; für Hindernisse und $d_p = \infty$; ansonsten (Abstand). Bei bewegten Hindernissen ist hierbei allerdings ein so genanntes "Background Modelling" nötig. Bei dem Background Modelling wird der Hintergrund der Szene so modelliert (z. B. über Referenzbilder oder Geometriemodelle), dass bekannte Hindernisse H_i nicht als Vordergrund detektiert werden und somit unbekannte Objekte O_j vor diesen Hindernissen als Objekte erkannt werden können.

[0062] **Fig. 8c**) zeigt Differenzbilder aus Intensitätskameras mit Abstand $d_p = 0$ für Objekte bzw. Hindernisse (Vordergrund) und $d_p = \infty$; sonst (Hintergrund).

[0063] Die **Fig. 9a** bis **Fig. 9c**) zeigen weitere wichtige Abstandsinformationen. Dazu sei Pixel p ∈ $B_i(v)$ aus der Projektion des Anfangsvolumens $v \⊂ E; V$ in Kamera C_i mit Sichtstrahl $s = P_{i, \‐1}(p)$ gegeben.

[0064] Aus der Kamerakalibrierung und der gegebenen Hindernis-Geometrie sind dann folgende Abstände $d_p, d_{v,min}, d_{v,max}, d_{H,min}$ sowie $d_{H,max}$ messbar bzw. berechenbar.

[0065] **Fig. 9a**) zeigt die Abstände $d_{v,min}$ bzw. $d_{v,max}$ als minimalen bzw. maximalen Abstand des Anfragevolumens v zur Kamera C_i entlang des Sichtstrahls s .

[0066] **Fig. 9b**) zeigt die Abstände $d_{H,min}$ bzw. $d_{H,max}$ als minimalen bzw. maximalen Abstand eines Hindernisses H zur Kamera C_i entlang des Sichtstrahls s .

[0067] **Fig. 9c**) zeigt den gemessenen Abstand d_p zu dem nächsten Objekt, welches auch mit dem minimalen Hindernisabstand $d_{H,min}$ zusammenfallen kann. Wegen der Konvexität der Hindernisse befindet sich zwischen $d_{H,min}$ und $d_{H,max}$ nur Hindernisvolumen.

[0068] **Fig. 10** zeigt den Verfahrensschritt der Klassifikation von einzelnen Pixeln p einer Kamera C_i . Sei ein Pixel p ∈ $B_i(v)$ aus der Projektion des Anfragevolumens $v \⊂ E; V$ in der Kamera C_i mit Sichtstrahl $s = P_{i, \‐1}(p)$ gegeben. Dann liefert die sensorbasierte Klassifikation $K_i(p, v)$ von Pixel p einen Zustand aus {Frei, Objekt, Hindernisse} bezüglich des Teilvolumens $v \⋂ s$ mit

[0069] $K_i(p, v)$ ist undefiniert für $p \in P_i(v)$, da dann $v \⋂ s = \emptyset$; gilt.

[0070] Die in dem Kamerabild B_i projizierten Pixel p ∈ $B_i(v)$ können pro Anfragevolumen v Teilmenge von V unterschiedlich klassifiziert sein. Gemischte Klassifikationen $K_i(p, v)$ sind pro Kamera C_i auf eine dominante Klasse $D_i(v)$ zurückführbar wie dies in der Tabelle gemäß **Fig. 11** dargestellt ist.

[0071] Zusammen mit den einheitlichen Klassifikationen erfolgt die Bestimmung der dominanten Klasse $D_i(v)$ von Volumen V mit

[0072] Um einen oder mehrere Fehler im System zu tolerieren, wird die Information über mehrere Kameras hinweg redundant ausgewertet. In **Fig. 12** sind verschiedene Formen der Redundanz r dargestellt. Sie unterscheiden sich in der konkreten Ausprägung des Anfragevolumens v . Dieses kann beispielsweise ein vorzugs

Seite 8 --- ()

weise würfelförmiges Volumenelement (Voxel) (a), ein Sichtstrahl (b), ein Sichtbereich, welcher von allen Kameras identisch klassifiziert wird (c) oder das gesamte, von allen Kameras prinzipiell überdeckte Überwachungsvolumen (d) sein. Dieses Anfragevolumen v wird von mehreren Kameras gleichzeitig überwacht, so dass mehrere Detektionsaussagen über v vorliegen.

[0073] Diese redundanten Detektionsaussagen sind nicht unproblematisch. Liefert beispielsweise eine Kamera C_i für ein beliebiges Teilvolumen v die (pixelbasierte) Detektion $D_i(v) \in \{\text{Frei, Objekt, Hindernis}\}$ kann dies zu widersprüchlichen Aussagen führen, wie beispielsweise $D_i(v) = \text{Frei}$ oder $D_i(v) = \text{Objekt}$.

[0074] Diese Problematik kann mit den in **Fig. 13a**, b) eingeführten Detektionsdiagrammen behoben werden. Es sei (f, o, h) der Detektionsvektor für Volumen $v \in E; V$ mit $f := \text{card}\{|D_i(v) = \text{Frei}\}|$ $o := \text{card}\{|D_i(v) = \text{Objekt}\}|$ $h := \text{card}\{|D_i(v) = \text{Hindernis}\}|$

[0075] Für eine Anzahl von c Kameras gilt immer $f + o + h = c$.

[0076] Das Detektionsdiagramm entspricht dabei einer Dreiecksfläche im Koordinatensystem von (f, o, h) wie dies in **Fig. 13a**) kontinuierlich dargestellt ist. Für eine diskrete Anzahl von Kameras entsteht das diskrete Detektionsdiagramm in **Fig. 13b**). Beispiele für drei unterschiedliche Detektionsvektoren sind in **Fig. 13c**) gegeben. Oben in **Fig. 13c**) ist der Detektionsvektor $(f, o, h) = (2, 0, 0)$, in der Mitte $(f, o, h) = (1, 1, 0)$ und unten $(f, o, h) = (0, 2, 0)$ gegeben.

[0077] In dem Detektionsdiagramm bestimmt nun die Entscheidungsstrategie aus der Information mehrerer Kameras C_i den Zustand $Z(v)$ des Anfragevolumens v . Kamerabasierte Entscheidungsstrategien sind vollständig im Detektionsdiagramm eintragbar bzw. ablesbar. **Fig. 14a**) zeigt ein Beispiel für sichere Überwachung bei Redundanz r als 1 Schwellenwert. **Fig. 14b**) zeigt ein Beispiel mit drei Schwellenwerten f_{\min}, o_{\min} sowie h_{\min} .

[0078] Je nach Strategie können nicht alle Detektionsvektoren eindeutig einem Zustand zugeordnet werden. In **Fig. 15** ist ein Beispiel mit einem erweiterten Zustand $Z_e(v)$ dargestellt. Erweiterte Zustände erhalten neben dem Zustand $Z(v)$ den Zustand "unbestimmt"; mit Unbestimmt: v ist nicht eindeutig einem der Zustände Frei, Objekt, Hindernis zuordbar.

[0079] Im Folgenden werden Möglichkeiten aufgezeigt für den Umgang mit Verdeckungen bzw. mit "Unbestimmten Zuständen". (Die Figuren gehen beispielhaft von einem Differenzbildverfahren aus.)

[0080] Die erste Möglichkeit ist, alle Verdeckungen bzw. unbestimmte Zustände als mögliches Objekt aufzufassen. **Fig. 16** zeigt eine solche Situation für Verdeckungen $VB_1 \dots VB_4$. Dies hat den Vorteil, dass diese Situationen einfach zu realisieren sind. Allerdings dürfte der Nutzen gering sein, falls Schutzzonen SZ um Hindernisse zum Beispiel Roboter mit angeschmiegtten Verdeckungen liegen.

[0081] Die zweite Möglichkeit ist anzunehmen, dass alle Verdeckungen, welche vollständig innerhalb einer Schutzzone liegen, keine Objekte enthalten. **Fig. 17** zeigt eine solche Situation für Verdeckungen VB_1, VB_2 und Schutzzone SZ . Diese Verdeckungen VB werden dann als Freiraum gewertet. Diese Annahme ist auch für bewegte Hindernisse H einfach realisierbar. Allerdings ist diese Annahme nicht immer pauschal erfüllt.

[0082] Die nächste Möglichkeit ist anzunehmen, dass Verdeckungen v mit Durchmesser $d(v)$ kleiner als d_{\min} kein Objekt vollständig enthalten können (**Fig. 18**) Daraus folgt, dass Verdeckungen v als "Frei" gewertet werden können, falls für die reflexiv-transitive Hülle $N^+(v)$ der verdeckten Nachbarvolumina gilt: $d(\cup_{v \in N^+(v)} v) < d_{\min}$ mit $d(v) := \max\{r_1, r_2\}$; mit $r_1, r_2 \in V$

[0083] Eine weitere Möglichkeit ist anzunehmen, dass Verdeckungen und/oder unbestimmte Zustände ab oder innerhalb eines vorgegebenen Koordinatenabschnitts keine Objekte enthalten können. **Fig. 19** zeigt eine solche Situation unter der Annahme, dass Verdeckungen v ab der Höhe z_{\max} nicht von Objekten O erreichbar seien. Daraus folgt, dass Verdeckungen v als "Frei" gewertet werden, falls gilt:

Seite 9 --- ()

$\min\{z \mid \exists x, y: (x, y, z) \in N; N \geq z_{\max}$

[0084] Analog zur Höhe z kann bezüglich x oder y vorgegangen werden.

[0085] Das erfindungsgemäße Verfahren und das System zur Durchführung des Verfahrens ermöglicht den Einsatz von Kameras C_i für Schutzeinrichtungen der funktionalen Maschinensicherheit, die Arbeitsbereiche von Mensch und Maschine "vorzugsweise dreidimensional" überwachen, um den Menschen vor gefahrbringenden Maschinenbewegungen zu schützen, das heißt Kollisionen zu vermeiden.

[0086] Durch die Erfindung wird die Produktivität von Maschinen verbessert, zum Beispiel der vermehrte Einsatz von Robotern ohne trennenden Schutzraum, um eine direkte Mensch-Maschine-Kooperation zu ermöglichen.

[0087] Auch lassen sich Kosteneinsparungen realisieren, weil häufig auftretende Sicherheitsfunktionen wie "Zutrittsabsicherung" und "Hindertrettschutz"; die heute im Regelfall von zwei Schutzeinrichtungen ausgeführt werden, durch eine einzige Schutzeinrichtung realisieren lassen.

[0088] Auch sind mit dem erfindungsgemäßen System Verbesserungen der Wirksamkeit des Personenschutzes bei Maschinenbetrieb möglich. Bei den Verbesserungen der Personenschutzfunktion werden dabei ausdrücklich auch die Vermeidung des Manipulationsrisikos von Schutzeinrichtungen einbezogen und auch der Missbrauch von Maschinenfunktionalitäten eingeschlossen, zum Beispiel unerlaubterweise auf Förderbändern mitzufahren.

Seite 10 --- ()

ZITATE ENHALTEN IN DER BESCHREIBUNG

[0089] Diese Liste der vom Anmelder aufgeführten Dokumente wurde automatisiert erzeugt und ist ausschließlich zur besseren Information des Lesers aufgenommen. Die Liste ist nicht Bestandteil der deutschen Patent- bzw. Gebrauchsmusteranmeldung. Das DPMA übernimmt keinerlei Haftung für etwaige Fehler oder Auslassungen.

Zitierte Patentliteratur

[0090] - EP 1635108 A1 **[0002, 0002]** - EP 1543270 B1 **[0004]**

Zitierte Nicht-Patentliteratur

[0091] - Ebert D., Henrich D.: "Safe Human-Robot-Cooperation: Imagebased Collision Detection for Industrial Robots"; IEEE International Conference on Intelligent Robots and Systems, (IROS'02), Lausanne, 30. September bis 04. Oktober 2002 **[0003]** - EN IC 61508-1/-7 **[0021]**

Seite 11 --- ()

[1] Verfahren zur Überwachung eines dreidimensionalen Raumbereichs (V), insbesondere einer Arbeitszelle einer automatisiert arbeitenden Maschine wie Roboter, bei dem Bilder (B_i) verwendet werden, die in zeitlicher Folge von wenigstens zwei, richtungs-diversitär bzw. richtungs-disimilar angeordneten Kameras ($C_1, C_2 \dots C_c$) erfasst wurden, deren Sichtstrahlen ($SI_{11}, SI_{12}; SI_{21}, SI_{22}; SI_{31}, SI_{32}$) sich wenigstens teilweise in einem Anfragevolumen (v) in dem Raumbereich (V) überschneiden und bei dem anhand einer Szenenanalyse der Bilder (B_i) in das Anfragevolumen (v)

eindringende Objekte (O_i) und/oder gegebenenfalls Hindernisse (H_k) detektiert werden, **dadurch gekennzeichnet**, dass in jeder oder für jede Kamera (C_1, C_2, \dots, C_c) eine unabhängige Szenenanalyse der Bilder (B_i) durchgeführt wird, dass Ergebnisse der unabhängigen Szenenanalysen der richtungsdiversitären Kameras (C_1, \dots, C_c) redundant ausgewertet und dadurch ein oder mehrere fehlertolerante Überwachungssignale gebildet werden.

[2] Verfahren nach Anspruch 1 oder 2 dadurch gekennzeichnet, dass eine Detektion eines Zustandes (Z) des Anfragevolumens (v) erfolgt, wobei überprüft wird, ob das Anfragevolumens (v) den Zustand "Frei"; aufweist, das heißt weder ein Objekt (O_i) noch ein Hindernis (H_k) enthält, den Zustand "Objekt"; aufweist, das heißt möglicherweise Teile von einem Objekt (O_i) enthält oder den Zustand "Hindernis"; aufweist, das heißt Teile eines Hindernisses (H_k) aber keine Objekte (O_i) enthält.

[3] Verfahren nach zumindest einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, dass der Zustand (Z) des Anfragevolumens (v) als "Frei"; detektiert wird, wenn das Anfragevolumen (v) mit mindestens der für die sichere Überwachung geforderten Redundanz (r) von den Kameras (C_i) als "Frei"; von einem Objekt (O_i) erkannt wird.

[4] Verfahren nach zumindest einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, dass die Hindernisse (H_k) definiert und in einem Auswertesystem eingelesen und/oder hinterlegt werden.

[5] Verfahren nach zumindest einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, dass eine Analyse per Einzelbildverfahren durch Soll-/Ist-Vergleich fortlaufend erfolgt.

[6] Verfahren nach zumindest einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, dass Bilder (B_i) von Anfragevolumen (v) bestimmter Kameras (C_i) als nicht redundant und/oder sicherheitstechnisch akzeptabel unberücksichtigt bleiben.

[7] Verfahren nach zumindest einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, dass bei einer Kombination der Zustände (Z) einer bestimmten Anzahl von Anfragevolumina (v) eine Plausibilitätskontrolle erfolgt.

[8] Verfahren nach zumindest einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, dass für jedes aufgenommene Bild (B_i) unter Berücksichtigung der räumlichen Anordnung der Kamera (C_i) der definierten Hindernisse (H_i) sowie des Anfragevolumens (v) eine Klassifikation von Pixeln (p) erfolgt.

[9] Verfahren nach zumindest einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, dass anschließend eine dominante Klasse für das Anfragevolumen (v) bestimmt wird.

[10] Verfahren nach zumindest einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, dass unter Berücksichtigung der definierten Redundanz (r) eine kamerabasierte Entscheidung mittels Detektionsdiagramm erfolgt.

[11] Verfahren nach zumindest einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, dass unter Berücksichtigung von Randbedingungen wie z. B. Abständen eine anwendungsbasierte Auflösung unbestimmter Zustände (Z) vorgenommen wird.

[12] Verfahren nach zumindest einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, dass die Ergebnisse von unabhängigen Szenenanalysen von zumindest drei, vorzugsweise vier Kameras (C_1, \dots, C_c) redundant ausgewertet werden.

[13] System zur Überwachung eines dreidimensionalen Raumbereichs (V), insbesondere einer Arbeitszelle einer automatisiert arbeitenden Maschine wie Roboter, mit zumindest zwei dezentral und richtungs-diversitär bzw. richtungs-disimilar angeordneten einzelnen Kameras (C_1, \dots, C_c) zum Aufnehmen von zumindest zwei Bil

Seite 12 --- ()

dern (B_1, B_2) des Raumes (V), wobei die Kameras versetzt zueinander angeordnet sind, und mit einer Auswerteeinheit die mittels einer Szenenanalyse der Bilder Objekte (O_i) und/oder Hindernisse (H_k) in dem Raum detektiert, dadurch gekennzeichnet, dass die Auswerteeinheit Mittel zur redundanten Auswertung von Ergebnissen von unabhängigen Szenenanalysen der einzelnen Kameras (C_1, \dots, C_c) sowie Mittel zur Erzeugung zumindest eines sicherheitsrelevanten Überwachungssignals auf der Grundlage der Auswertung aufweist.

[14] System nach Anspruch 13, dadurch gekennzeichnet, dass das System zumindest drei Kameras (C_1, C_2, C_3) aufweist.

Seite 13 --- ()

Seite 14 --- ()

Seite 15 --- ()

Seite 16 --- ()

Seite 17 --- ()

Seite 18 --- ()

Seite 19 --- ()

Seite 20 --- ()

Seite 21 --- ()

Seite 22 --- ()