



(19)
Bundesrepublik Deutschland
Deutsches Patent- und Markenamt

(10) DE 10 2006 001 033 A1 2006.09.07

(12)

Offenlegungsschrift

(21) Aktenzeichen: 10 2006 001 033.7

(22) Anmeldetag: 07.01.2006

(43) Offenlegungstag: 07.09.2006

(51) Int Cl.⁸: G06K 9/62 (2006.01)

(66) Innere Priorität:

10 2005 010 306.5 03.03.2005

(71) Anmelder:

Volkswagen AG, 38440 Wolfsburg, DE

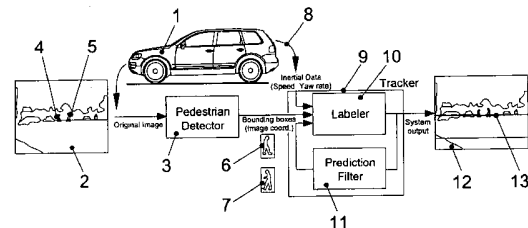
(72) Erfinder:

Graf, Thorsten, Dr., 38518 Gifhorn, DE; Meinecke, Marc-Michael, Dr., 38524 Sassenburg, DE; Broggi, Alberto, Prof.Dr., Parma, IT; Fascioli, Alessandra, Dr., Parma, IT; Bertozzi, Massimo, Parma, IT; Grisleri, Paolo, Roveleto di Cadeo, IT; Ghidoni, Stefano, Parma, IT

Die folgenden Angaben sind den vom Anmelder eingereichten Unterlagen entnommen

(54) Bezeichnung: Vorrichtung und Verfahren zur Erkennung eines Objektes für ein Kraftfahrzeug

(57) Zusammenfassung: Die Erfindung betrifft ein Verfahren und eine Vorrichtung zur Erkennung von Objekten in der Fahrumgebung für ein Kraftfahrzeug. Insbesondere betrifft die Erfindung ein solches Verfahren und eine solche Vorrichtung, mit denen eine Objektidentitätszuordnung von Objekten O_{ij} , die mittels eines Objektselektionsverfahrens in erfassten Bildern der Fahrumgebung selektiert sind, möglich ist. Ferner betrifft die Erfindung ein Verfahren und eine Vorrichtung zur Objektidentitätszuordnung. Geschaffen wird insbesondere ein Verfahren zur Objektidentitätszuordnung von Objekten O_{ij} , die mittels eines Objektselektionsverfahrens in in einer zeitlichen Abfolge $i + 2, i + 1, i, \dots$ erfassten Bildern Bl_i selektiert worden sind, wobei das Objektselektionsverfahren für jedes selektierte Objekt O_{ij} eines Bilds Bl_i eine Bildregion BR_{ij} liefert, in der das jeweilige selektierte Objekt O_{ij} in dem Bild Bl_i abgebildet ist, wobei die Bildregionen BR_{ij} eines Bilds Bl_i jeweils einzeln mit jeder Bildregion $BR_{i+1,L}$ eines zeitlich vorangehend aufgenommenen Bilds Bl_{i+1} in Beziehung gesetzt werden, wobei ein Einzelglaubwürdigkeitswert $p_{ij}(L)$ ermittelt wird, dass die zwei miteinander in Beziehung gesetzten Bildregionen $BR_{ij}, BR_{i+1,L}$ dasselbe Objekt abbilden und den in dem Bild Bl_i selektierten Objekten O_{ij} jeweils ein Label L anhand einer Zuordnungsglaubwürdigkeit $A_{ij,L}^{ij}$ die basierend auf den Einzelglaubwürdigkeitswerten $p_{ij}(L)$ ermittelt ist, zugewiesen wird, ...



Beschreibung

[0001] Die Erfindung betrifft ein Verfahren und eine Vorrichtung zur Erkennung von Objekten in der Fahrumgebung für ein Kraftfahrzeug. Insbesondere betrifft die Erfindung ein solches Verfahren und eine solche Vorrichtung, mit denen eine Objektidentitätszuordnung von Objekten O_{ij} , die mittels eines Objektsелеktionsverfahrens in erfassten Bildern der Fahrumgebung selektiert sind, möglich ist. Ferner betrifft die Erfindung ein Verfahren und eine Vorrichtung zur Objektidentitätszuordnung.

[0002] In zukünftigen Fahrerassistenzsystemen wird elektronischen Sicherheits- und Komfortfunktionen, insbesondere Systemen der aktiven und passiven Fahrzeugsicherheit, eine große Bedeutung zukommen. Die Basis für die Realisierung solcher Systeme stellt die Sensorik dar, die es ermöglicht, die Fahrumgebung eines Fahrzeuges wahrzunehmen. Insbesondere für Fahrzeugsysteme, die eine Auslösung von Schutzmaßnahmen für Fahrzeuginsassen und/oder Fußgänger und/oder Zweiradfahrer umfassen sowie für Fahrzeugbeleuchtungssysteme, die in Abhängigkeit von in der Fahrzeugumgebung vorhandenen Objekten angesteuert werden sowie für Fahrerassistenzsysteme, die eine Precrashfunktion bzw. Vorrichtungen zur Unfallvermeidung bzw. Unfallfolgenminderung umfassen, ist es erforderlich, Objekte, die sich in der Umgebung des Fahrzeuges befinden, zu erfassen und nach der Objektart zu klassifizieren. Es ist bekannt, mittels einer Kamera, eines Bildsensors, einer Videosensorik oder einer Abstandssensorik, wie beispielsweise Radarsensoren, die Fahrzeugumgebung zu erfassen und im Hinblick auf vorhandene Objekte, wie beispielsweise Fußgänger, Zweiradfahrer, Fahrzeuge oder Wild, auszuwerten. Das Auffinden von bestimmten Objekten in einem Bild bzw. das Extrahieren von diesen bestimmten Objekten wird allgemein auch als Objektselektion bezeichnet.

Stand der Technik

[0003] Im Sinne der hier gemachten Ausführungen, ist ein Verfahren, das Bildregionen von 4 abgebildeten Objekten in aufgenommenen Bildern auffindet und diese isoliert und/oder extrahiert, ein Selektionsverfahren. Die Selektion erfolgt in der Regel anhand von Selektionskriterien. Daher wird im Stand der Technik, abweichend von dem hier verwendeten Sprachgebrauch, zuweilen auch ein Selektieren einer Bildregion als „ein Erkennen eines Objekts“ bezeichnet. Ein Erkennen im Sinne der hier gemachten Ausführungen, setzt zusätzlich zu der Selektion zumindest eine Identifizierung des in der selektierten Bildregion abgebildeten Objekts voraus, worunter nicht nur eine bloße Klassifizierung nach der Art des Objekts, beispielsweise Fußgänger, gemeint ist. Gemäß der hier dargelegten Definition gilt ein Objekt

beispielsweise als erkannt, wenn es als ein bereits bekanntes Objekt identifiziert ist.

[0004] Die DE 100 25 678 A1 beschreibt ein kamerabasiertes Precrash-Erkennungssystem zur Detektion von Verkehrsteilnehmern und Hindernissen auf der Grundlage von Kamerabildern. Hierzu werden mittels Klassifikationen mögliche Unfallobjekte ermittelt und deren Entfernung und Relativbewegung ermittelt.

[0005] Aus der DE 10243397 A1 ist ein Sensorsystem zur Auslösung von Schutzmaßnahmen für Kraftfahrzeug-Insassen und Fußgänger bekannt, welches einen Abstandssensor aufweist, der die Abstände und Annäherungsgeschwindigkeit in Winkelsektoren bestimmt und in definierten Abständen über einen Bildsensor Bilder der sich annähernden Objekte aufnimmt und diese Bilder mit der Entfernungsmessung korreliert und mittels der Ergebnisse die Schutzmaßnahmen auswählt und vor dem eigentlichen Aufprall auslöst.

[0006] In der DE 10301468 A1 ist eine Vorrichtung zur Beobachtung der Umgebung eines Fahrzeuges beschrieben, die in der Lage ist, in der Umgebung des Fahrzeuges anwesende Objekte auf der Basis eines Bildes, das durch zumindest ein an dem Fahrzeug vorgesehenes Infrarotkameraelement aufgenommen wird, zu erfassen. Die Vorrichtung umfasst eine Binärobjekt-Extraktionseinheit, die ein Grauwertbild des durch das Infrarotkameraelement aufgenommenen Bildes einem binären Schwellenvergleichsverfahren unterzieht und ein binäres Objekt aus dem Grauwertbild extrahiert, sowie eine Grauwertobjekt-Extraktionseinheit, die ein Grauwertobjekt, wovon ein Bereich das binäre Objekt enthält, auf der Basis einer Leuchtdichteänderung des Grauwertbildes aus dem Grauwertbild extrahiert. Des Weiteren umfasst die Vorrichtung eine Fußgängerbestimmungseinheit, die in einem das Grauwertobjekt enthaltenden Bereich einen Suchbereich einstellt und auf Basis einer Leuchtdichteverteilung in dem Suchbereich einen Fußgänger auf dem Grauwertbild erkennt.

[0007] Aus der DE 10327115 B3 ist ein Fußgängerschutzsystem für Fahrzeuge bekannt, wobei mittels eines optischen Sensors und/oder eines Radarsensors eine Fahr-/Vorausstrecke des Fahrzeuges überwacht wird. Aus den erfassten Sensorsignalen wird eine Ausgangsinformation erzeugt, die an eine erste vorgesehene Datenbank geleitet wird. Mittels einer Mikroprozessor- und/oder Bildprozessoreinheit wird ein elektronischer und/oder optischer Vergleich der Pixel – und/oder Bilddaten aus der ersten Bilddatenbank mit hinterlegten Objektbildklassen/Fußgängerobjektbildklassen aus einer zweiten Datenbank durchgeführt, wobei eine elektronische und optische Identifikation einer Fußgängerkontur ermittelt wird

und nach Identifikation der Fußgängerkontur durch die Mikroprozessor- und/oder Bildprozessoreinheit ein nachgeschalteter Prognoserechner die voraussichtliche Kollisionszeit aus der Abstandsmessung Kraftfahrzeug-Fußgänger und der Kraftfahrzeug-zugengeschwindigkeit ermittelt.

Aufgabenstellung

[0008] Der Erfindung liegt die Aufgabe zugrunde, ein Verfahren und eine Vorrichtung zum Auffinden und zur Erkennung von Objekten in der Fahrumgebung eines Kraftfahrzeuges, derart weiterzubilden, dass eine möglichst fehlerfreie Erkennung der Objekte bzw. der Objektbewegung in mehreren nacheinander aufgenommenen Bildern ermöglicht wird, sowie eine Vorrichtung und ein Verfahren zu schaffen, mit denen eine zuverlässige Identifikation von Objekten in aufeinander folgend aufgenommenen Bildern der Umgebung des Kraftfahrzeugs möglich ist.

[0009] Diese Aufgabe wird erfindungsgemäß durch ein Verfahren zur Objektidentitätszuordnung mit den Merkmalen des Patentanspruchs 1, eine Vorrichtung zur Objektidentitätszuordnung mit den Merkmalen des Patentanspruchs 10, ein Verfahren zum Erkennen von Objekten mit den Merkmalen des Patentanspruchs 19 und eine Vorrichtung zum Erkennen von Objekten mit den Merkmalen des Patentanspruchs 21 gelöst. Vorteilhafte Ausgestaltungen ergeben sich aus den Unteransprüchen.

[0010] Um die Bewegung von Objekten in der realen Welt bestimmen zu können und auch Vorhersagen treffen zu können, müssen die Objekte, die in einem Bild erkannt worden sind, Objekten aus vorangegangenen Bildern zugeordnet werden. Nur so lassen sich aus der Positionshistorie die Bewegungen in der Zukunft schätzen. Dieses ist für viele verschiedene Fahrerassistenzsysteme wichtig, um z.B. Unfallvorhersagen treffen zu können, oder das Fahrzeug automatisch steuern zu können.

[0011] Insbesondere ist daher ein Verfahren zur Objektidentitätszuordnung von Objekten $O_{i,j}$ vorgesehen, die mittels eines Objektselektionsverfahrens in in einer zeitlichen Abfolge $i+2, i+1, i, \dots$ erfassten Bildern BI_i selektiert worden sind, wobei das Objektselektionsverfahren für jedes selektierte Objekt $O_{i,j}$ eines Bilds BI_i eine Bildregion $BR_{i,j}$ liefert, in der das jeweilige selektierte Objekt $O_{i,j}$ in dem Bild BI_i abgebildet ist, wobei die Bildregionen $BR_{i,j}$ eines Bilds BI_i jeweils einzeln mit jeder Bildregion $BR_{i+1,L}$ eines zeitlich vorangehend aufgenommenen Bilds BI_{i+1} in Beziehung gesetzt werden, wobei jeweils ein Einzelglaubwürdigkeitswert $p_{i,j}(L)$ ermittelt wird, dass die zwei miteinander in Beziehung gesetzten Bildregionen $BR_{i,j}, BR_{i+1,L}$ dasselbe Objekt abbilden, und den in dem Bild BI_i selektierten Objekten $O_{i,j}$ jeweils ein Label L anhand einer Zugordnungsglaubwürdigkeit $A_{i,j}^L$,

die basierend auf den Einzelglaubwürdigkeitswerten $p_{i,j}(L)$ ermittelt ist, zugewiesen wird, wobei die Zuordnungsglaubwürdigkeit ein Maß dafür ist, wie glaubwürdig die Zuordnung des Labels L zu dem j -ten selektierten Objekt $O_{i,j}$ im i -ten Bild ist.

[0012] Objektselektionsverfahren, die Bestandteil eines jeden Objekterkennungsverfahrens sind, sind dem Fachmann vom Prinzip her aus dem (zum Teil oben gewürdigten) Stand der Technik bekannt und werden hier nicht näher erläutert. Das genutzte Objektselektionsverfahren ist so ausgestaltet, dass es aus den Bildern Bildregionen selektiert, in denen das selektierte Objekt abgebildet ist. Diese Bildregionen können so ausgestaltet sein, dass die selektierte Bildregion nur das selektierte Objekt umfasst. Es ist jedoch auch möglich, dass die selektierte Bildregion eine vorgegebene geometrische Form, beispielsweise ein Rechteck, besitzt. In diesem Fall umfasst die Bildregion auch Bildbestandteile, die nicht zu dem selektierten Objekt gehören, falls die vorgegebene geometrische Form von der Form der Abbildung des selektierten Objekts abweicht. Indem die selektierten Bildregionen aus zeitlich aufeinander folgend aufgenommenen Bildern miteinander in Beziehung gesetzt werden, ist es möglich, zunächst Einzelglaubwürdigkeiten zu ermitteln, die angeben, wie glaubwürdig es erscheint, dass diese Bildregionen in den aufeinander folgend aufgenommenen Bildern dasselbe Objekt abbilden.

[0013] Die Bildregionen werden vorzugsweise in Beziehung gesetzt, indem ein Überlappungsbereich der Bildregionen ermittelt wird. Die absolute Größe des Überlappungsbereichs stellt ein mögliches Maß für die Einzelglaubwürdigkeit dar. Da die Objekte unterschiedliche Größen aufweisen und/oder aufgrund ihrer unterschiedlichen Entfernung von der Kamera in den Bildern auf unterschiedlich große Bildregionen abgebildet werden, wird es bevorzugt, die Größen der Überlappungsbereiche jeweils auf die Fläche eines der beiden in Beziehung gesetzten Bildregionen, eine Summe dieser Flächen oder eine sich bei der Überlagerung der beiden Bildregionen ergebende Vereinigungsfläche zu normieren. Weitere Kriterien, wie eine Formähnlichkeit, eine Farb- und/oder Helligkeitsähnlichkeit usw. der Objekte, können beim In-Beziehung-Setzen ebenfalls einzeln oder in Kombination verwendet werden, um die Einzelglaubwürdigkeiten zu ermitteln.

[0014] Beim In-Beziehung-Setzen eines Objekt bzw. einer Bildregion mit mehreren Objekten des vorher aufgenommenen Bilds erhält man mehrere Einzelglaubwürdigkeiten. In vielen Fällen weisen mehrere dieser Einzelglaubwürdigkeiten einen nicht verschwindenden Wert auf. Wird die Überlappung zur Beurteilung der Glaubwürdigkeit herangezogen; so tritt dieser Fall immer auf, wenn ein Objekt eine Überlappung mit zwei Objekten aus dem vorherigen Bild

aufweist. Dieser Fall tritt häufig bei Objekten, beispielsweise Fußgängern oder Zweiradfahrern, auf, die sich selbst bewegen. Daher ist es erforderlich, bei der Zuweisung von Labeln mehrere der Einzelglaubwürdigkeiten zu berücksichtigen. Es ist vorgesehen, aus den Einzelglaubwürdigkeiten Zuordnungsglaubwürdigkeiten für das Objekt zu ermitteln. Für das zu identifizierende Objekt werden Zuordnungsglaubwürdigkeiten ermittelt und zwar jeweils eine für jedes der bekannten, identifizierten, mit einem Label versehenen Objekte des vorangehenden Bilds. Das Objekt kann nur mit demjenigen der bekannten Objekte identifiziert werden, dem die höchste Zuordnungsglaubwürdigkeit zugeordnet ist. Das Identifizieren erfolgt, indem dem Objekt das Label des bereits erkannten Objekts zugewiesen wird. Objekte in den unterschiedlichen Bildern, denen dasselbe Label zugewiesen ist, sind als dasselbe Objekt identifiziert. Diese Identifizierung wird für alle Objekte des Bilds ausgeführt, die mit den bekannten Objekten identifiziert werden sollen. Diese Identifizierungsprozesse sind jedoch nicht unabhängig voneinander. Daher wird vorzugsweise das Objekt zuerst identifiziert, für das der höchste Zuordnungsglaubwürdigkeitswert ermittelt wurde. Ist aufgrund der Zuordnungsglaubwürdigkeiten keine sinnvolle Identifizierung eines oder mehrerer Objekte des Bilds mit bereits erkannten Objekten der vorangehenden Bilder möglich, so wird diesen Objekten jeweils eine neues Label zugeordnet. Sie werden somit als neue eigenständige Objekte identifiziert.

[0015] Ein erfindungsgemäßes Verfahren zum Erkennen von Objekten O_{ij} , insbesondere Fußgängern oder Zweiradfahrern, für ein Kraftfahrzeug, welches neben der Selektion der Objekte O_{ij} aus einer Abbildung auch eine Identifizierung der Objekte ermöglicht, wird vorgeschlagen, bei dem mittels einer Kamera von einer Umgebung des Kraftfahrzeugs zeitlich aufeinander folgend Bilder Bl_i erfasst werden. Das Verfahren umfasst ein Objektselektionsverfahren, mit dem in den erfassten Bildern Bl_i die Objekte O_{ij} selektiert werden und für jedes selektierte Objekt O_{ij} eines Bilds Bl_i eine Bildregion BR_{ij} geliefert wird, in der das jeweilige selektierte Objekt O_{ij} in dem Bild Bl_i abgebildet ist, und ein Verfahren zur Objektidentitätszuordnung, wie es oben beschrieben ist. Hierdurch wird ein Verfahren geschaffen, das Informationen für die einzelnen Objekte zur Verfügung stellt, die ausreichend sind, um in Assistenzsystemen des Kraftfahrzeugs, beispielsweise für eine Precrash-Analyse eingesetzt zu werden. Bereits im Vorfeld eines möglichen Zusammenstoßes mit einem Objekt, beispielsweise Fußgänger oder Zweiradfahrer, kann auf diese Gefahr aufmerksam gemacht werden oder sogar in die Fahrzeugsysteme (Lenkung, Bremssystem, Airbag-System, ...) eingegriffen werden oder diese aktiviert werden. Hierdurch lassen sich Unfälle und Gefahrensituationen vermeiden und/oder ihre nachteiligen Folgen vermindern.

[0016] Um Fehlzuordnungen der Identität zu vermeiden, wird es bevorzugt, dass die Zuordnungsglaubwürdigkeit einen Mindestwert, der als vorgegebener Schwellenwert festgelegt ist, überschreiten muss, damit ein Objekt mit einem bereits vorher erkannten Objekt identifiziert wird. Bei einer bevorzugten Ausführungsform ist somit vorgesehen, dass beim Zuweisen der Label L einem selektierten Objekt O_{ij} ein bereits einem der Objekte $O_{i+1,j}$ eines zeitlich vorangehend aufgenommenen Bilds Bl_{i+1} zugewiesenes Label L verwendet wird, wenn die Zuordnungsglaubwürdigkeit A_{ij}^L größer als ein Schwellenwert ist.

[0017] Objekte, die weit von der Kamera entfernt sind, mit der die Bilder aufgenommen und erfasst werden, werden auf eine kleinere Bildregion abgebildet als gleich große Objekte, die sich näher an der Kamera befinden. Dies ist eine Folge der Abbildungsgeometrie. Die Entfernung wirkt sich auch auf die Größe des Überlappungsbereichs von Bildregionen desselben Objekts sowie dessen Auflösung aus. Daher ist bei einer bevorzugten Weiterbildung vorgesehen, dass für die selektierten und/oder erkannten Objekte O_{ij} eine Entfernung ermittelt wird und der Schwellenwert eine Entfernungsabhängigkeit aufweist, wobei der Schwellenwert mit der ermittelten Entfernung abnimmt. Die Entfernung eines Objekts lässt sich beispielsweise schätzen, wenn man annimmt, dass eine Fläche vor der Kamera eben ist und ein Neigungswinkel der Kamera bezüglich der Fläche sowie ein Öffnungswinkel der Kamera bekannt sind. In erster Näherung ist dann ein Abstand von einem unteren Bildrand ein Maß für die Entfernung in der ebenen Fläche von der Kamera. Der geringste Abstand einer Bildregion eines Objekts von dem unteren Bildrand kann als Maß der Entfernung herangezogen werden, wenn ferner angenommen wird, dass sich das Objekt auf der ebenen Fläche bewegt. Diese grobe Abschätzung lässt sich verbessern, indem Informationen über die genaue Abbildungsgeometrie, eine Fahrzeuglage, usw. mit einbezogen werden.

[0018] Da sich ein Kraftfahrzeug bewegt, verändert sich die Abbildungsgeometrie bezüglich einer stationären Umgebung des Kraftfahrzeugs abhängig von dieser Bewegung. Bei einer bevorzugten Ausführungsform wird daher beim In-Beziehung-Setzen der Bildregionen BR_{ij} eine Bewegung der die Bilder Bl_i aufnehmenden Kamera zwischen den Zeitpunkten, an denen die zeitlich aufeinander folgenden Bilder Bl_i aufgenommen sind, kompensiert.

[0019] Zu einer sehr verbesserten Identifizierung der Objekte mit bereits erkannten Objekten gelangt man, wenn man auch Informationen berücksichtigt, die im Zusammenhang mit einem vorvorherigen Bild oder weiteren hiervor aufgenommenen Bildern stehen. Eine verbesserte Identifizierung erreicht man bei einer Ausführungsform, bei der zusätzlich zu den für das Bild Bl_i ermittelten Einzelglaubwürdigkeiten

$p_{ij}(L)$ mindestens die für das eine zeitlich vorangehend aufgenommene Bild BI_{i+1} ermittelten Einzelglaubwürdigkeiten $p_{i+1,j}(L)$ zur Ermittlung der Zuordnungsglaubwürdigkeiten A_{ij}^L herangezogen werden. Hierdurch können beispielsweise Fälle besser berücksichtigt werden, bei denen sich zwei Objekte, die zunächst in einem Bild getrennt sichtbar sind, in einem nachfolgenden Bild vollständig überlappen und als ein Objekt erscheinen und in einem späteren Bild erneut als zwei Objekte sichtbar sind.

[0020] Um auch Situationen wie die eben Beschriebene und ferner solche Situationen gut handhaben zu können, in denen eines der Objekte in einem Bild von einem Gegenstand verdeckt wird, ist bei einer Ausführungsform vorgesehen, dass mittels einer Voraussageeinheit für erkannte und zugeordnete Objekte $O_{i+1,j}^{(L)}$ des zeitlich vorangehend aufgenommenen Bilds BI_{i+1} anhand der hiervor aufgenommenen Bilder BI_{i+k} ($k > 2$) vorausgesagte Bildregionen BR'_i bestimmt werden und zusätzlich zu den Bildregionen BR_{ij} der selektierten Objekten O_{ij} bei der Labelzuweisung berücksichtigt werden.

[0021] Die Zuordnungsglaubwürdigkeiten weisen eine besonders hohe Aussagekraft bei einer Ausführungsform auf, bei der die Zuordnungsglaubwürdigkeit A_{ij}^L nach folgender Formel errechnet wird:

$$A_{ij}^L = p_{ij}(L) + \sum_{i=1}^{M-1} \sum_{j=0}^{N_{L,i}} p_{ij}(L) \cdot e^{-i},$$

wobei i die aufgenommenen Bilder BI_i in zeitlich umgekehrter Reihenfolge indexiert, j die in dem jeweiligen Bild BI_i erkannten und gegebenenfalls vorausgesagten Bildregionen BR_{ij} indexiert, von denen es für das jeweilige Bild BI_i $N_{L,i}$ gibt, L jeweils das Label bezeichnet, das den jeweils zugeordneten Objekten $O_{ij}^{(L)}$ in den zeitlich vorangehenden Bildern BI_i ($i > 0$) zugewiesen ist, und e die Euler'sche Zahl ist und M die insgesamt berücksichtigten Bilder BI_i angibt, wobei das eine Bild BI_i für dessen selektierte und gegebenenfalls vorausgesagte Objekte O_{ij} die Zuordnungsglaubwürdigkeiten A_{ij}^L errechnet werden, den Index $i=0$ besitzt. Bei diesem Berechnungsverfahren werden Einzelglaubwürdigkeiten vorheriger Bilder mit einbezogen. Für ein Label L , d.h. eine Identität, wird für das vorher aufgenommene Bild (und entsprechend für die davor aufgenommenen Bilder) jeweils die Glaubwürdigkeit ermittelt, dass das in dem entsprechenden vorvorherigen Bild mit dem Label L identifizierte Objekt in dem vorherigen Bild erneut abgebildet ist. Hierfür werden die Einzelglaubwürdigkeiten für alle Objekte des vorherigen Bilds addiert, dass sie dasselbe Objekt wie das im vorvorherigen Bild mit dem Label L identifizierte Objekt sind. Zu dieser Glaubwürdigkeit für das vorherige Bild (und den entsprechend ermittelten Glaubwürdigkeiten für die davor aufgenommenen Bilder) wird die Einzelwahrscheinlichkeit addiert, dass das in dem vorherigen

Bild mit dem Label L identifizierte Objekt mit dem zuzuordnenden Objekt $O_{0,j}$ identisch ist. Zusätzlich werden die Glaubwürdigkeiten der einzelnen Bilder durch eine Potenz der eulerschen Zahl e dividiert und somit in ihrem Einfluss auf die Zuordnungsglaubwürdigkeit gemindert. Die Minderung ist um so stärker je weiter die Aufnahme des Bilds zeitlich zurückliegt.

[0022] Zur Berücksichtigung von Rotationen der Kamera bzw. des Kraftfahrzeugs ist vorteilhafterweise vorgesehen, dass die Bewegungskompensation einer Drehung der Kamera um einen Rotationswinkel α , der gegebenenfalls aus einer Gierrate κ multipliziert mit der Zeit t errechnet wird, die zwischen den Zeitpunkten, an denen die aufeinander folgenden Bildern BI_i aufgenommen sind, vergangen ist, berücksichtigt wird, indem die zeitlich aufeinander folgend aufgenommenen Bilder BI_i um d Bildpunkte gegeneinander verschoben werden, wobei gilt:

$$d = \frac{\alpha}{\Phi} b$$

und b eine Bildbreite angibt und Φ einen Kameraöffnungswinkel angibt. Dieses stellt eine einfache Vorgehensweise dar, die zu einer guten Kompensation führt.

[0023] Eine Bewegungskompensation einer Translation berücksichtigt man vorteilhafterweise, indem die Bildregionen BI_i eines der in Beziehung zu setzenden Bilder BI_i aus einem zweidimensionalen Bildraum in einen dreidimensionalen mit einem Weltkoordinatensystem versehenen Raum transformiert werden, anschließend in dem dreidimensionalen Raum bezüglich der Translation, die die Kamera zwischen Zeitpunkten, an denen die Bilder BI_i , BR_{i+1} aufgenommen sind, ausgeführt hat, kompensiert (verschoben) werden und anschließend in zweidimensionalen Bildraum zurück transformiert werden.

[0024] Bei einer vorteilhaften Ausführungsform des Verfahrens zum Erkennen von Objekten ist vorgesehen, dass die selektierten Objekte O_{ij} und/oder die zugehörigen Bildregionen BR_{ij} des erfassten Bildes BI_i ausgegeben werden, wobei die selektierten Objekte O_{ij} und/oder die zugehörigen Bildregionen BR_{ij} jeweils mit dem zugewiesenen Label gekennzeichnet sind. Das Ausgeben kann in Form des Bilds so erfolgen, dass das Bild entsprechend modifiziert ist, um das Label und die Bildregion und gegebenenfalls zusätzlich die Entfernung kenntlich zu machen oder hervorzuheben. Ferner können die Objekte selbst, z.B. farblich, hervorgehoben werden. Ebenso können die ermittelten Zusatzinformationen, Identität, Entfernung, Lage im Bild, usw., auch als Daten, gegebenenfalls zusätzlich, ausgegeben werden. Das Ausgeben der selektierten Objekte und zugehörigen Label kann ebenso in parametrisierter Form erfolgen. Ferner können Zusatzinformationen ausgegeben und/oder bereit gestellt werden.

[0025] Bei einer Ausführungsform ist vorgesehen, dass Objekte, die nicht in einer vorgegebenen Mindestanzahl von Bildern nacheinander als ein und dasselbe Objekt identifiziert wurden, nicht ausgegeben werden.

[0026] Die Merkmale der erfindungsgemäßen Vorrichtung zur Objektidentitätszuordnung und einer erfindungsgemäßen Vorrichtung zum Erkennen von Objekten weisen jeweils dieselben Vorteile wie die entsprechenden erfindungsgemäßen Verfahren auf.

Ausführungsbeispiel

[0027] Nachfolgend wird die Erfindung anhand einer Zeichnung näher erläutert. Hierin zeigen:

[0028] Fig. 1 eine Darstellung einer Ausführungsform eines Erkennungsverfahrens für Fußgänger;

[0029] Fig. 2 eine schematische Darstellung zweier aufgenommener Bilder sowie eine Vorgehensweise, um eine Einzelglaubwürdigkeit dafür zu ermitteln, dass darin selektierte Bildregionen dasselbe Objekt abbilden;

[0030] Fig. 3 eine schematische Darstellung von vier aufeinander folgend aufgenommenen Bildern sowie deren entsprechenden Überlagerungen zum Ermitteln von Überlappungsbereichen von Bildregionen der selektierten Objekte in den Bildern;

[0031] Fig. 4 eine Darstellung eines Tracking-Graphen, der mit den Bildern nach Fig. 3 korrespondiert;

[0032] Fig. 5 ein Ausschnitt des Tracking-Graphen nach Fig. 4 für eines der in dem zeitlich zuletzt aufgenommenen Bild selektierten Objekte; und

[0033] Fig. 6a bis Fig. 6c eine Darstellung von Bildern, die von einer Kamera in einer zeitlichen Abfolge aufgenommen sind und von einer Ausführungsform eines Erkennungsverfahrens ausgegeben sind.

[0034] Das erfindungsgemäße Verfahren wird anhand eines Fußgängererkennungssystems beschrieben. Es kann jedoch auf beliebige Objekte angewandt werden. Die erfindungsgemäße Vorrichtung umfasst ein System zur Objektselektion, welches in der Lage ist, in Videobildern Objekte zu selektieren.

[0035] Der schematische Ablauf einer Ausführungsform eines Verfahrens zum Erkennen von Objekten ist in Fig. 1 beschrieben. Ein Fahrzeug **1** umfasst zumindest eine oder mehrere Kameras (nicht dargestellt), die die Fahrzeugumgebung wahrnehmen. Hierzu werden Bilder **2** aufgenommen. Die ursprünglich erfassten Bilder werden auch als Originalbilder bezeichnet. Die Originalbilder werden einem Objektselektionssystem **3** zugeführt. Solche Objektselekti-

onssysteme werden zum Teil auch als Objekterkennungssystem im engeren Sinne bezeichnet. Im Sinne der hier gemachten Ausführungen ist ein Objektselektionssystem, das lediglich Objekte in einem Bild auffindet und selektiert, kein Objekterkennungssystem. Ein Objekterkennungssystem im Sinne der hier gemachten Ausführungen ist ein übergeordnetes System, das eine über ein reines Auffinden, das heißt ein Selektieren, von Objekten hinausgehende zusätzliche Identifizierung der Objekte ausführt.

[0036] Das Objektselektionssystem selektiert in der vorgestellten Ausführungsform Fußgänger als Objekte **4, 5**. Es werden Bildregionen **6, 7** ermittelt und deren Koordinaten ausgegeben, in denen die selektierte Fußgänger in den Bildern abgebildet sind. Die so selektierten Bildregionen **6, 7** der Objekte **4, 5** werden auch als Objekthypothesen bezeichnet.

[0037] Anschließend werden die selektierten Bildregionen **6, 7** sowie zusätzliche Fahrzeugdaten **8** (z.B. Geschwindigkeit, Gierrate, etc.) dem so genannten Tracker **9** übergeben, der die Zuordnung der Bildregionen **6, 7** der selektierten Objekte **4, 5** zu den bereits bekannten Objekten der vorhergehenden Bilder vornimmt.

[0038] Der Tracker **9** enthält zwei Module, einen Labeler **10** (Kennzeichner) und eine als Prediction Filter (Vorhersagefilter) **11** ausgestaltete Voraussageeinheit. Der Labeler **10** und die Voraussageeinheit werden ebenso wie ihre Funktionsweisen weiter unten genauer beschrieben.

[0039] Ausgegeben wird ein Bild **12**, in dem die selektierten Bildregionen **6, 7** der Objekte **4, 5** mit einem Label versehen sind, das heißt, erkannt sind, wobei das Label jeweils eine Identifikation mit den bereits bekannten Objekten ermöglicht bzw. angibt. Ferner sind in der Regel Zusatzinformationen in dem ausgegebenen Bild enthalten, hier Entfernungslinien **13**.

[0040] In anderen Ausführungsformen können auch lediglich Objektinformationen, wie z. B. Entfernung, ausgegeben werden, um z. B. andere Fahrerassistenzsysteme zu steuern.

[0041] In wieder anderen Ausführungsformen können Informationen über identifizierte Objekte gemeinsam mit ihrer Zuordnung in parametrisierter Form und gegebenenfalls mit Zusatzinformationen bereitgestellt und/oder ausgegeben werden.

[0042] Der Labeler **10** stellt ein zentrales Modul dar, welches die Zuordnung der neu selektierten Objekte zu den bereits erkannten (selektierten und mit vorbekannten Objekten identifizierten) Objekten in den vorhergehenden Bildern vornimmt. Hierzu erhält der Labeler **10** von dem Selektionssystem eine Liste mit den im aktuellen Bild selektierten Objekten in Form

von Bildregionen und vergleicht diese mit den bereits selektierten und identifizierten sowie gespeicherten Bildregionen der vergangenen Bilder. Überlappt sich die Bildregion eines neu selektierten Objektes in ausreichendem Maße mit einer Bildregion der bereits erkannten Objekte, so wird das neu selektierte Objekt mit dem bereits erkannten identifiziert. Diese Zuordnung geschieht über einen Schwellenwert, der in Abhängigkeit mit der Entfernung zum Objekt variiert (je weiter das Objekt von der Kamera entfernt ist, desto geringer ist der Schwellenwert). Hierbei kann es vorkommen, dass ein neu selektiertes Objekt zu mehreren bereits erkannten Objekten zugeordnet wird.

[0043] Dieser Vergleich geschieht direkt im zweidimensionalen Bildraum, da eine Entfernungsschätzung von Mono-Kamerasystemen üblicher Weise recht ungenau ist. Im Folgenden wird die Funktionsweise des Labelers **10** beispielhaft beschrieben.

[0044] In **Fig. 2** sind schematisch ein Bilder **20** ($i=0$) und ein zeitlich vorangehend aufgenommenes Bild **21** ($i=1$) dargestellt. In beiden Bildern **20**, **21** ist jeweils ein Objekt A in Bildregionen **22**, **23** abgebildet. Das Objekt A hat sich zwischen den Zeitpunkten, an denen die Bilder **20**, **21** aufgenommen wurden, bewegt. Daher befindet sich die Bildregionen **22**, **23** nicht an derselben Position in den Bildern **20**, **21**.

[0045] Überlagert man die beiden Bilder **20**, **21**, so erhält man ein Überlagerungsbild **24**. Betrachtet man die Überlagerung der Bildregionen **22**, **23**, so erhält man einen Überlappungsbereich **25** und einen Vereinigungsbereich **26**. Der Überlappungsbereich **25** ist eine Schnittmenge der Bildregionen **22**, **23**. Der Vereinigungsbereich ist eine Vereinigungsmenge der Bildregionen **22**, **23**.

[0046] Um auszudrücken, wie glaubwürdig es ist, dass die Bildregion **22** aus Bild **20** dasselbe Objekt A, das in der Bildregion **23** von dem zeitlich vorangehend aufgenommenen Bild **21** dargestellt ist und mit einem Label $L=1$ bezeichnet ist, abbildet, wird eine Einzelglaubwürdigkeit $p_{i=0,j=0}(L=1)$ errechnet, die sich aus dem Quotienten der Flächen des Überlappungsbereichs **25** und des Vereinigungsbereichs **26** ergibt, wie unten rechts in **Fig. 2** dargestellt ist. Der Vereinigungsbereich **26** wird zur Normierung verwendet. Ebenso könnte eine Normierung auf eine der Flächen der Bildregionen **22**, **23** erfolgen oder ganz unterbleiben.

[0047] In diesem Fall ist in den Bildern **20**, **21** jeweils nur das eine Objekt A abgebildet. Es wird somit nur eine Einzelglaubwürdigkeit $p_{ij}(L)$ bestimmt, die in diesem Falle zugleich auch eine Zuordnungsglaubwürdigkeit A_{ij}^L ist. Die Zuordnungsglaubwürdigkeit A_{ij}^L wird mit einem Schwellenwert verglichen. Übersteigt die Zuordnungsglaubwürdigkeit A_{ij}^L diesen Schwellenwert, so wird der Bildregion **22** des Bilds **20** und

somit dem in dem Bild **20** selektierten Objekt A das Label $L=1$ der Bildregion **23** des zeitlich vorausgehend aufgenommenen Bilds **21** zugewiesen. Das in den Bildern **20**, **21** abgebildete Objekt wird als dasselbe identifiziert und ist mit dem Label $L=1$ gekennzeichnet.

[0048] Der Schwellenwert ist vorzugsweise abhängig von einer Entfernung, die für das Objekt bestimmt wird, dessen Identität ermittelt werden soll. Unter der Annahme, dass die Fläche vor der Kamera eben ist und ein Neigungswinkel der Kamera bezüglich der ebenen Fläche sowie ein Öffnungswinkel der Kamera bekannt sind und der weiteren Annahme, dass sich die Objekte auf dieser ebenen Fläche bewegen, ist ein minimaler Abstand der jeweiligen Bildregion **22**, **23** von einem entsprechenden unteren Rand **27**, **28** der Bilder **20**, **21** ein Maß für die Entfernung der Objekte von der Kamera, wie diese durch Skalen **29**, **30** neben den Bildern **20**, **21** angedeutet ist. Eine horizontale Linie **31** im Bild **20** berührt die Bildregion **22** an einem unteren Ende **32** und schneidet die Skala **29** bei dem Wert 3. Das der Bildregion **22** zugeordnete Objekt A befindet sich somit 3 Längeneinheiten entfernt. Auf ähnliche Weise lässt sich die Entfernung des Objekts A in dem zeitlich vorangehend aufgenommenen Bild **21** ermitteln. Diese ergibt sich aus einem Schnittpunkt einer weiteren horizontalen Linie **33** mit der Skala **30** zu 1 Längeneinheit.

[0049] In **Fig. 3** sind vier Bilder **41**, **42**, **43**, **44** dargestellt. Oben ist das zeitlich zuerst aufgenommene Bild **44**, $i=3$, und unten das zuletzt aufgenommene Bild **41**, $i=0$, dargestellt.

[0050] Von einem Selektionssystem sind in den Bildern **41** (BI_0), **42** (BI_1), **43** (BI_2), **44** (BI_3) Bildregionen **41-1** ($BR_{i=0,j=0}$), **41-2** ($BR_{0,1}$); **42-1** ($BR_{1,0}$), **42-2** ($BR_{1,1}$); **43-1** ($BR_{2,0}$), **43-2** ($BR_{2,1}$), **43-3** ($BR_{2,2}$) und **44-1** ($BR_{3,0}$), **44-2** ($BR_{3,1}$), **44-3** ($BR_{3,2}$) selektiert worden, die entsprechenden Objekten O_{ij} zugeordnet sind.

[0051] In einer rechten Spalte der **Fig. 3** sind entsprechende Überlagerungsbilder **45** (Überlagerung der Bilder **44** und **43**), **46** (Überlagerung der Bilder **43** und **42**) und **47** (Überlagerung der Bilder **42** und **41**) dargestellt. Die Bildbereiche des jeweils vorangehend aufgenommenen Bilds sind gestrichelt dargestellt. Die sich ergebenden Überlappungsbereiche **45-1**, **45-2**, **45-3**, **45-4**, **45-5**; **46-1**, **46-2**, **46-3**; **47-1**, **47-2**, **47-3** korrespondieren mit Kanten in einem so genannten Tracking-Graphen, wie er in **Fig. 4** dargestellt ist. Wie oben im Zusammenhang mit **Fig. 3** erläutert, lassen sich aus den Überlappungsbereichen **45-1**, **45-2**, **45-3**, **45-4**, **45-5**; **46-1**, **46-2**; **47-1**, **47-2**, **47-3**, **47-4** die Einzelglaubwürdigkeiten $p_{ij}(L)$ ermitteln. Diese sind in **Fig. 3** und **Fig. 4** jeweils graphisch zugeordnet.

[0052] **Fig. 4** zeigt einen Tracking-Graphen. Die

Kreise bezeichnen Objekte, wobei die Zahlen in den Kreisen bereits zugeordnete Label bzw. vorläufig zugeordnete Label angeben. Die Verbindungslinien bezeichnen mögliche Zuordnungen von einem Bild zum nächsten. Das heißt sie geben die Beziehungen zwischen Objekten bzw. deren Bildregionen an, für die eine Einzelglaubwürdigkeit $p_{i,j}(L)$, dass die Bildregionen dasselbe Objekt abbilden, nicht gleich null ist. Die Kreise, die untereinander stehen, beschreiben jeweils die Objekte ein und desselben Bildes, während die Kreise, die nebeneinander stehen, die selektierten Objekte in verschiedenen Bildern darstellen. Zu jeder Linie gehört noch eine Einzelglaubwürdigkeit $p_{i,j}(L)$, wobei i das Bild von vor i Zeitschritten bezeichnet und L das Label des Objekts bzw. der Bildregion des jeweils zeitlich vorangehend aufgenommenen Bilds $i+1$ ist und j das entsprechende selektierte (und gegebenenfalls erkannte) Objekt des i -ten Bilds indexiert. Diese Einzelglaubwürdigkeiten ergeben sich aus dem Umfang der Überlappungen der zugeordneten Bildregionen.

[0053] Jedes Objekt $O_{i,j}$ bzw. jede Bildregion $BR_{i,j}$ ($j=0,1$) in dem aktuellen Bild **41**, BI_0 , (vgl. **Fig. 3**) erhält zunächst eine temporäres Label, das auch als Objektnummer bezeichnet wird. Dem Objekt $O_{i=0,j=0}$, welches mit einem Index $j=0$ assoziiert ist, ist das temporäre Label **4** und dem anderen Objekt $O_{i=0,j=1}$, welches mit einem Index $j=1$ assoziiert ist, ist das temporäre Label **5** zugewiesen. Wenn es eine Überlappung der Bildregionen **41-1** ($BR_{i=0,j=0}$), **41-2** ($BR_{0,1}$) (vgl. **Fig. 3**) der aktuellen Objekte mit einem oder mehreren Bildregionen **42-1** ($BR_{1,0}$), **42-2** ($BR_{1,1}$) von Objekten des vorherigen Bildes **42**, BI_1 , gibt, gibt es auch eine mögliche Zuordnung mit einer dazugehörigen Glaubwürdigkeit, die Zuordnungsglaubwürdigkeit $A_{i=0,j}^{i,j}$ genannt wird. Die Zuordnungsglaubwürdigkeit $A_{i=0,j}^{i,j}$ gibt für die j -te selektierte Bildregion $BR_{i=0,j}$ bzw. vorausgesagte Bildregion $BR'_{i=0,j}$ des aktuellen Bilds **41**, BI_0 , für das $i=0$ gilt, die Glaubwürdigkeit an, dass in der Bildregion $BR_{i=0,j}$, das bereits mit dem Label L identifizierte Objekt abgebildet ist. Diese Zuordnungsglaubwürdigkeit wird basierend auf den Einzelglaubwürdigkeiten $p_{i,j}(L)$ ermittelt. Das temporäre Label kann durch das Label ersetzt werden, zu dem die höchste Zuordnungsglaubwürdigkeit $A_{i,j}^{i,j}$ gehört.

[0054] In **Fig. 5** ist ein Ausschnitt des Tracking-Graphen nach **Fig. 4** für die mit dem temporären Label **5** versehene Bildregion $BR_{0,1}$ in einer anderen Darstellungsweise gezeigt. Alle möglichen Zuordnungszweige, die eine ununterbrochene Einzelglaubwürdigkeitskette zu der mit dem Label **5** gekennzeichneten Bildregion $BR_{0,1}$ darstellen sind in einer Baumstruktur dargestellt.

[0055] Die Zuordnungsglaubwürdigkeit $A_{i,j}^{i,j}$ ergibt sich bei einer bevorzugten Ausführungsform aus:

$$A_{i,j}^{i,j} = p_{i,j}(L) + \sum_{l=1}^{M-1} \sum_{j=0}^{N_{L,i}} p_{i,j}(L) \cdot e^{-l},$$

wobei i die aufgenommenen Bilder BI_i in zeitlich umgekehrter Reihenfolge indexiert, j die in dem jeweiligen Bild BI_i erkannten und gegebenenfalls vorausgesagten Bildregionen $BR_{i,j}$ indexiert, von denen es für das jeweilige Bild BI_i $N_{L,i}$ gibt, L jeweils das Label bezeichnet, das den jeweils zugeordneten Objekten $O_{i,j}(L)$ in den zeitlich vorangehenden Bildern BI_i ($i>0$) zugewiesen ist, e die Euler'sche Zahl ist und M die insgesamt berücksichtigten Bilder BI_i angibt, wobei das eine Bild $BI_{i=0}$, für dessen selektierte Objekte $O_{i,j}$ die Zuordnungsglaubwürdigkeiten $A_{i,j}^{i,j}$ errechnet werden, den Index $i=0$ besitzt. Bei diesem Berechnungsverfahren werden Einzelglaubwürdigkeiten vorheriger Bilder mit einbezogen. Für ein Label L , d.h. eine Identität, wird für das vorher aufgenommene Bild (und entsprechend für die davor aufgenommenen Bilder) die Glaubwürdigkeit ermittelt, dass das in dem vorherigen Bild mit dem Label L identifizierte Objekt in dem vorherigen Bild erneut abgebildet ist. Hierfür werden für alle Objekte des vorherigen Bilds die Einzelglaubwürdigkeiten dafür addiert, dass sie dasselbe Objekt wie das im vorherigen Bild mit dem Label L identifizierte Objekt sind. Zu dieser Glaubwürdigkeit für das vorherige Bild (und den entsprechend ermittelten Glaubwürdigkeiten für die davor aufgenommenen Bilder) wird die Einzelglaubwürdigkeit dafür addiert, dass die Bildregion $BR_{i=0,j}$ des aktuellen Bild $BI_{i=0}$ dasselbe Objekt abbildet, das in dem vorherigen Bild mit dem Label L identifiziert ist. Zusätzlich werden die Glaubwürdigkeiten jeweils durch eine Potenz der eulerschen Zahl e dividiert und somit in ihrem Einfluss auf die Zuordnungsglaubwürdigkeit gemindert. Die Minderung ist um so stärker, je weiter die Aufnahme des Bilds zeitlich zurückliegt. Das bedeutet, dass der Einfluss vergangener Zuordnungen mit der verstrichenen Zeit abnimmt.

[0056] Prinzipiell können auch andere Berechnungsvorschriften verwendet werden.

[0057] In dem obigen Fall würde nun die temporäre Objektnummer **5** durch die glaubwürdigste (in diesem Fall **2**) ersetzt werden. Wird keine Zuordnung gefunden, so behält das Objekt die temporäre Objektnummer. Ein als neu identifiziertes Objekt gilt im Sinne dieser Ausführungen für eine Auswertung eines nachfolgend aufgenommenen nächsten Bilds $i-1$ als erkanntes Objekt. Es kann jedoch vorgesehen werden, dass ein Objekt erst als erkanntes Objekt ausgegeben wird, wenn es in einer vorgegebenen Anzahl m von Bildern erkannt worden ist.

[0058] Einen großen Einfluss auf die Stabilität des Trackings, das heißt der Objektidentifizierung, hat eine Eigenbewegung des Kraftfahrzeugs, speziell in Kurvenfahrten, da sich hierdurch die Position der Objekte in den Bildern sehr stark unterscheiden kann.

[0059] Aus diesem Grund hat es sich als hilfreich erwiesen, die Eigenbewegung des Kraftfahrzeugs zu berücksichtigen.

[0060] Rotationen können berücksichtigt werden, in dem man die Gierrate κ des Kraftfahrzeugs verwendet. Der Rotationswinkel ist dann gegeben durch:

$$\alpha = \kappa \cdot t$$

wobei t die zwischen zwei Bildern vergangene Zeit ist. Die Folge einer Rotation ist, dass alle Objekte (unabhängig von deren Entfernung) um d Punkte im Bild verschoben erscheinen:

$$d = \frac{\alpha}{\phi} b$$

wobei b die Bildbreite und ϕ der Öffnungswinkel der Kamera ist.

[0061] Translationen können berücksichtigt werden, in dem die Bildregionen in einen dreidimensionalen Raum mit Weltkoordinaten transformiert werden, die Eigenbewegung der Kamera abgezogen wird und die korrigierten transformierten Bildregionen zurück ins zweidimensionale Bild projiziert werden. Hierbei ist allerdings zu beachten, dass die Genauigkeit der bestimmten Bildregion sehr stark von der geschätzten Entfernung des zugehörigen Objektes abhängt.

[0062] Um Objekte auch dann weiterverfolgen zu können, wenn die Erkennung in einem Bild fehlschlägt (vgl. [Fig. 6a](#) bis [Fig. 6c](#)), benötigt das System die Möglichkeit aus der Historie die Bewegung der Objekte, beispielsweise der Fußgänger, über mehrere Bilder hinweg vorherzusagen zu können.

[0063] In [Fig. 6a](#) bis [Fig. 6c](#) sind drei in einer zeitlichen Abfolge aufgenommene Bilder **61a** bis **61c** dargestellt, die von einer Ausführungsform eines Erkennungsverfahrens ausgegeben sind. In den Bildern **61a** und **61c** ist jeweils ein Fußgänger **62a**, **62c** zu erkennen. Das Selektionsverfahren hat jeweils eine Bildregion **63a**, **63c** festgelegt, in der der Fußgänger in den Bildern **61a**, **61c** abgebildet ist. Im Bild **61b** ist der Fußgänger von einem Gegenstand **64b** verdeckt. Um hier eine kontinuierliche Identifizierung des Objekts Fußgänger **62a**, **62c** zu ermöglichen, wird für das Bild **61b** eine Bildregion **63b** vorausgesagt, in der der Fußgänger **62b** abgebildet sein sollte, wenn er nicht durch den Gegenstand **64b** verdeckt wäre.

[0064] Dieses übernimmt eine Bewegungsvoraussage (Motion Prediction) der als Prediction Filter (Voraussagefilter) **11** ausgebildeten Voraussageeinheit (vgl. [Fig. 1](#)). Mit Hilfe eines bekannten α - β -Filters oder eines Kalman-Filters kann die Bewegung in bekannter Weise vorhergesagt werden.

[0065] Diese Bewegungen werden im zweidimensi-

onalen Raum bestimmt und können zusätzlich zu den durch das Objektselektionssystem selektierten Objekten an den Labeler weitergegeben werden, d.h., dass nicht nur die Bildregionen aus dem Selektionssystem, sondern auch die Bildregionen aus der Bewegungsvoraussage (Prädiktion) für eine Zuordnung im Labeler verwendet werden.

[0066] Auf diese Weise können Objekte trotz Verdeckungen etc. verfolgt werden. Zusätzlich besteht die Möglichkeit, nur die Objekte an ein Fahrerassistenzsystem weiter zu reichen, die mit Hilfe des Objektidentifizierungsverfahrens in mindestens m Bildern erkannt worden sind. Hierdurch ergibt sich eine zusätzliche Robustheit gegenüber fälschlicher Weise erkannten Objekten.

[0067] In [Fig. 6a](#) bis [Fig. 6c](#) sind die Bildregionen **63a** bis **63c** jeweils mit einem Label **65a** bis **65c** gekennzeichnet, das hier eine **2** ist. Ferner ist die Entfernung des Fußgängers **61a** bis **61c** jeweils mittels einer Entfernungslinie **66a** bis **66c** und einer Entfernungsangabe **67a** bis **67c** angegeben, die in das jeweilige Bild **61a** bis **61c** eingeblendet sind.

Bezugszeichenliste

1	Fahrzeug
2	Bilder
3	Objektselektionssystem
4, 5	Objekte (hier Fußgänger)
6, 7	Bildregionen
8	Fahrzeugdaten
9	Tracker
10	Labeler
12	Bild
13	Entfernungslinien
20	Bild ($i=0$)
21	Bild ($i=1$)
22, 23	Bildregionen
24	Überlagerungsbild
25	Überlappungsbereich
26	Vereinigungsbereich
27, 28	unterer Rand der Bilder 20 , 21
29, 30	Skalen
31, 33	horizontale Linien
32	unteres Ende
41–44	Bilder
45	Überlagerungsbild (Überlagerung der Bilder 43 und 44)
46	Überlagerungsbild (Überlagerung der Bilder 42 und 43)
47	Überlagerungsbild (Überlagerung der Bilder 41 und 42)

- 61a-c in zeitlicher Reihenfolge aufgenommene Bilder
 62a-c Fußgänger
 63a-c Bildregionen
 64b Gegenstand
 65a-c Label
 66a-c Entfernungslinie
 67a-c Entfernungsangabe

Patentansprüche

1. Verfahren zur Objektidentitätszuordnung von Objekten $O_{i,j}$, die mittels eines Objektselektionsverfahrens in in einer zeitlichen Abfolge $i+2, i+1, i, \dots$ erfassten Bildern BI_i selektiert worden sind, wobei das Objektselektionsverfahren für jedes selektierte Objekt $O_{i,j}$ eines Bilds BI_i eine Bildregion $BR_{i,j}$ liefert, in der das jeweilige selektierte Objekt $O_{i,j}$ in dem Bild BI_i abgebildet ist,

dadurch gekennzeichnet, dass

die Bildregionen $BR_{i,j}$ eines Bilds BI_i jeweils einzeln mit jeder Bildregion $BR_{i+1,L}$ eines zeitlich vorangehend aufgenommenen Bilds BI_{i+1} in Beziehung gesetzt werden, wobei ein Einzelglaubwürdigkeitswert $p_{i,j}(L)$ ermittelt wird, dass die zwei miteinander in Beziehung gesetzten Bildregionen $BR_{i,j}, BR_{i+1,L}$ dasselbe Objekt abbilden, und

den in dem Bild BI_i selektierten Objekten $O_{i,j}$ jeweils ein Label L anhand einer Zugordnungsglaubwürdigkeit $A^{i,j}_L$, die basierend auf den Einzelglaubwürdigkeitswerten $p_{i,j}(L)$ ermittelt ist, zugewiesen wird, wobei die Zuordnungsglaubwürdigkeit ein Maß dafür ist, wie glaubwürdig die Zuordnung des Labels L zu dem j -ten erkannten Objekt $O_{i,j}$ im i -ten Bild ist.

2. Verfahren nach Anspruch 1, dadurch gekennzeichnet, dass beim Zuweisen der Label L einem selektierten Objekt $O_{i,j}$ ein bereits einem der erkannten Objekte $O_{i+1,j}$ eines zeitlich vorangehend aufgenommenen Bilds BI_{i+1} zugewiesenes Label L verwendet wird, wenn die Zuordnungsglaubwürdigkeit $A^{i,j}_L$ größer als ein Schwellenwert ist.

3. Verfahren nach Anspruch 2, dadurch gekennzeichnet, dass für die selektierten oder/und erkannten Objekte $O_{i,j}$ eine Entfernung ermittelt wird und der Schwellenwert eine Entfernungsabhängigkeit aufweist, wobei der Schwellenwert mit der ermittelten Entfernung abnimmt.

4. Verfahren nach einem der vorangehenden Ansprüche, dadurch gekennzeichnet, dass beim In-Beziehung-Setzen der Bildregionen $BR_{i,j}$ eine Bewegung der die Bilder BI_i aufnehmenden Kamera zwischen den Zeitpunkten, an denen die zeitlich aufeinander folgenden Bilder BI_i aufgenommen sind, kompensiert wird.

5. Verfahren nach einem der vorangehenden Ansprüche dadurch gekennzeichnet, dass zusätzlich zu

den für das Bild BI_i ermittelten Einzelglaubwürdigkeiten $p_{i,j}(L)$ mindestens die für das eine zeitlich vorangehend aufgenommene Bild BI_{i+1} ermittelten Einzelglaubwürdigkeiten $p_{i+1,j}(L)$ zur Ermittlung der Zuordnungsglaubwürdigkeiten $A^{i,j}_L$ herangezogen werden.

6. Verfahren nach einem der vorangehenden Ansprüche dadurch gekennzeichnet, dass mittels einer Voraussageeinheit für erkannte und zugeordnete Objekte $O_{i+1,j}$ des zeitlich vorangehend aufgenommenen Bilds BI_{i+1} anhand der hiervor aufgenommenen Bilder BI_{i+k} ($k > 2$) vorausgesagte Bildregionen BR'_i bestimmt werden und zusätzlich zu den Bildregionen $BR_{i,j}$ der selektierten Objekten $O_{i,j}$ bei der Labelzuweisung berücksichtigt werden.

7. Verfahren nach einem der vorangehenden Ansprüche, dadurch gekennzeichnet, dass die Zuordnungsglaubwürdigkeit $A^{i,j}_L$ nach folgender Formel errechnet wird:

$$A^{i,j}_L = p_{i,j}(L) + \sum_{i=1}^{M-1} \sum_{j=0}^{N_{L,i}} p_{i,j}(L) \cdot e^{-i},$$

wobei i die aufgenommenen Bilder BI_i in zeitlich umgekehrter Reihenfolge indexiert, j die in dem jeweiligen Bild BI_i erkannten und gegebenenfalls vorausgesagten Bildregionen $BR_{i,j}$ indexiert, von denen es für das jeweilige Bild BI_i $N_{L,i}$ gibt, L jeweils das Label bezeichnet, das den jeweils erkannten Objekten $O_{i,j}$ in den zeitlich vorangehenden Bildern BI_i ($i > 0$) zugewiesen ist, und e die Euler'sche Zahl ist und M die insgesamt berücksichtigten Bilder BI_i angibt, wobei das eine Bild BI_i für dessen selektierte Objekte $O_{i,j}$ die Zuordnungsglaubwürdigkeiten $A^{i,j}_L$ errechnet werden, den Index $i=0$ besitzt.

8. Verfahren nach einem der Ansprüche 4 bis 7, dadurch gekennzeichnet, dass die Bewegungskompensation einer Drehung der Kamera um einen Rotationswinkel α , der gegebenenfalls aus einer Gierrate κ multipliziert mit der Zeit t errechnet wird, die zwischen den Zeitpunkten, an denen die aufeinander folgenden Bilder BI_i aufgenommen sind, vergangen ist, berücksichtigt wird, indem die zeitlich aufeinander folgend aufgenommenen Bilder BI_i um d Bildpunkte gegeneinander verschoben werden, wobei gilt:

$$d = \frac{\alpha}{\Phi} b$$

und b eine Bildbreite angibt und Φ einen Kameraöffnungswinkel angibt.

9. Verfahren nach einem der Ansprüche 4 bis 8, dadurch gekennzeichnet, dass eine Bewegungskompensation einer Translation berücksichtigt wird, indem die Bildregionen BI_i eines der Bilder BI_i aus einem zweidimensionalen Bildraum in einen dreidimensionalen mit einem Weltkoordinatensystem versehenen Raum transformiert werden, anschließend

in dem Raum bezüglich der Translation, die die Kamera zwischen Zeitpunkten, an denen die Bilder Bl_i, Bl_{i+1} aufgenommen sind, ausgeführt hat, kompensiert, insbesondere verschoben, werden und anschließend in den zweidimensionalen Bildraum zurück transformiert werden.

10. Vorrichtung (V) zur Objektidentitätszuordnung von Objekten $O_{i,j}$, die mittels einer Objektselektionsvorrichtung in in einer zeitlichen Abfolge $i+2, i+1, i, \dots$ mittels einer Kamera (K) erfassten Bildern Bl_i selektiert worden sind, wobei die Objektselektionsvorrichtung für jedes selektierte Objekt $O_{i,j}$ eines Bilds Bl_i eine Bildregion $BR_{i,j}$ liefert, in der das jeweilige selektierte Objekt $O_{i,j}$ in dem Bild Bl_i abgebildet ist, umfassend einen Tracker (9), dadurch gekennzeichnet, dass der Tracker (9) eine Vergleichseinheit umfasst, um die Bildregionen $BR_{i,j}$ eines Bilds Bl_i jeweils einzeln mit jeder Bildregion $BR_{i+1,L}$ eines zeitlich vorangehend aufgenommenen Bilds Bl_{i+1} in Beziehung zu setzen, um hierbei einen Einzelglaubwürdigkeitswert $p_{i,j}(L)$ zu ermitteln, dass die zwei miteinander in Beziehung gesetzten Bildregionen $BR_{i,j}, BR_{i+1,L}$ dasselbe Objekt abbilden, und einen Labeler (10) umfasst, um den in dem Bild Bl_i selektierten Objekten $O_{i,j}$ jeweils ein Label L anhand einer Zuordnungsglaubwürdigkeit $A_{i,j}^L$, die basierend auf den auf den Einzelglaubwürdigkeitswerten $p_{i,j}(L)$ ermittelt ist, zuzuweisen, wobei die Zuordnungsglaubwürdigkeit ein Maß dafür ist, wie glaubwürdig die Zuordnung des Label L zu dem j -ten erkannten Objekt $O_{i,j}$ im i -ten Bild ist.

11. Vorrichtung nach Anspruch 10, dadurch gekennzeichnet, dass der Labeler (10) eine weitere Vergleichseinheit zum Vergleichen der Zuordnungsglaubwürdigkeit $A_{i,j}^L$ mit einem Schwellenwert umfasst und der Labeler (10) so ausgestaltet ist, dass beim Zuweisen der Label L einem selektierten Objekt $O_{i,j}$ ein bereits einem der erkannten Objekte $O_{i+1,j}$ eines zeitlich vorangehend aufgenommenen Bilds Bl_{i+1} zugewiesenes Label L verwendet wird, wenn die Zuordnungsglaubwürdigkeit $A_{i,j}^L$ größer als der Schwellenwert ist.

12. Vorrichtung nach Anspruch 10 oder 11, dadurch gekennzeichnet, dass der Tracker (9) eine Entfernungsermittlungseinheit zum Ermitteln einer Entfernung für die selektierten oder/und erkannten Objekte $O_{i,j}$ umfasst, wobei der Schwellenwert eine Entfernungsabhängigkeit aufweist und der Schwellenwert mit der ermittelten Entfernung abnimmt.

13. Vorrichtung nach einem der Ansprüche 10 bis 12, dadurch gekennzeichnet, dass der Tracker eine Bewegungskompensationseinheit zum Kompensieren einer Bewegung der Kamera umfasst, die die Kamera zwischen Zeitpunkten, zu denen die zeitlich aufeinander folgenden Bilder Bl_i aufgenommen sind,

ausführt.

14. Vorrichtung nach einem der Ansprüche 10 bis 13, dadurch gekennzeichnet, dass der Labeler (10) eine Speichereinheit umfasst, in der die ermittelten Einzelglaubwürdigkeiten gespeichert sind, und der Labeler so ausgestaltet ist, dass zusätzlich zu den für das Bild Bl_i ermittelten Einzelglaubwürdigkeiten $p_{i,j}(L)$ mindestens die für das eine zeitlich vorangehend aufgenommene Bild Bl_{i+1} ermittelten Einzelglaubwürdigkeiten $p_{i+1,j}(L)$ zur Ermittlung der Zuordnungsglaubwürdigkeiten $A_{i,j}^L$ herangezogen werden.

15. Vorrichtung nach einem der Ansprüche 10 bis 14, dadurch gekennzeichnet, dass der Tracker (9) eine Voraussageeinheit umfasst, um für erkannte und zugeordnete Objekte $O_{i+1,j}$ des zeitlich vorangehend aufgenommenen Bilds Bl_{i+1} anhand der hiervor aufgenommenen Bilder Bl_{i+k} ($k > 2$) vorausgesagte Bildregionen $BR'_{i,j}$ zusätzlich zu bestimmen, wobei der Labeler (10) die vorausgesagten Bildregionen $BR'_{i,j}$ und die selektierten Bildregionen $BR_{i,j}$ der selektierten Objekten $O_{i,j}$ in dem einen Bild Bl_i bei der Labelzuweisung berücksichtigt.

16. Vorrichtung nach einem der Ansprüche 10 bis 15, dadurch gekennzeichnet, dass der Labeler eine Recheneinheit zum Errechnen der Zuordnungsglaubwürdigkeit $A_{i,j}^L$ nach folgender Formel umfasst:

$$A_{i,j}^L = p_{i,j}(L) + \sum_{i=1}^{M-1} \sum_{j=0}^{N_{L,i}} p_{i,j}(L) \cdot e^{-i},$$

wobei i die aufgenommenen Bilder Bl_i in zeitlich umgekehrter Reihenfolge indexiert, j die in dem jeweiligen Bild Bl_i erkannten und gegebenenfalls vorausgesagten Bildregionen $BR_{i,j}$ indexiert, von denen es für das jeweilige Bild Bl_i $N_{L,i}$ gibt, L jeweils das Label bezeichnet, das den jeweils zugeordneten Objekten $O_{i,j}^{(L)}$ in den zeitlich vorangehenden Bildern Bl_i ($i > 0$) zugewiesen ist, e die Euler'sche Zahl ist und M die insgesamt berücksichtigten Bilder Bl_i angibt, wobei das eine Bild Bl_i , für dessen selektierte Objekte $O_{i,j}$ die Zuordnungsglaubwürdigkeiten $A_{i,j}^L$ errechnet werden, den Index $i=0$ besitzt.

17. Vorrichtung nach einem der Ansprüche 10 bis 16, dadurch gekennzeichnet, dass eine Bewegungskompensation einer Drehung der Kamera um einen Rotationswinkel α , der gegebenenfalls aus einer Gierrate κ multipliziert mit der Zeit t errechnet wird, die zwischen den Zeitpunkten, an denen die aufeinander folgenden Bildern Bl_i aufgenommen sind, vergangen ist, berücksichtigbar ist, indem die zeitlich aufeinander folgend aufgenommenen Bilder Bl_i um d Bildpunkte gegeneinander verschoben werden, wobei gilt:

$$d = \frac{\alpha}{\Phi} b$$

und b eine Bildbreite angibt und Φ einen Kameraöffnungswinkel angibt.

18. Vorrichtung nach einem der Ansprüche 10 bis 17, dadurch gekennzeichnet, dass der Tracker (9) eine Translationskompensationseinheit umfasst, mittels der eine Bewegungskompensation einer Translation berücksichtigt ist, indem die Bildregionen BR_i eines der Bilder BI_i aus einem zweidimensionalen Bildraum in einen dreidimensionalen mit einem Weltkoordinatensystem versehenen Raum transformiert werden, anschließend in dem dreidimensionalen Raum bezüglich der Translation, die die Kamera zwischen Zeitpunkten, an denen die Bilder BI_i , BI_{i+1} aufgenommen sind, ausgeführt hat, kompensiert, insbesondere verschoben, werden und anschließend in den zweidimensionalen Bildraum zurück transformiert werden.

19. Verfahren zum Erkennen von Objekten $O_{i,j}$, insbesondere Fußgängern oder Zweiradfahrern, für ein Kraftfahrzeug, bei dem mittels einer Kamera von einer Umgebung des Kraftfahrzeugs zeitlich aufeinander folgend Bilder BI_i erfasst werden, umfassend ein Objektselektionsverfahren, mit dem in den erfassten Bildern BI_i die Objekte $O_{i,j}$ selektiert werden und für jedes selektierte Objekt $O_{i,j}$ eines Bilds BI_i eine Bildregion $BR_{i,j}$ geliefert wird, in der das jeweilige selektierte Objekt $O_{i,j}$ in dem Bild BI_i abgebildet ist, und ein Verfahren gemäß einem der Ansprüche 1 bis 9.

20. Verfahren nach Anspruch 19, dadurch gekennzeichnet, dass die erkannten Objekt $O_{i,j}$ und/oder die zugehörigen Bildregionen $BR_{i,j}$ des erfassten Bild BI_i ausgegeben wird, wobei die erkannten Objekt $O_{i,j}$ und/oder die zugehörigen Bildregionen $BR_{i,j}$ mit dem zugewiesenen Label gekennzeichnet sind.

21. Vorrichtung zum Erkennen von Objekten $O_{i,j}$, insbesondere Fußgängern oder Zweiradfahrern, für ein Kraftfahrzeug, umfassend eine Kamera zum Aufnehmen von Bildern BI_i von einer Umgebung des Kraftfahrzeugs in einer zeitlichen Abfolge $i+2$, $i+1$, i , ..., ein Objektselektionssystem zum Erkennen von Objekten $O_{i,j}$ in den mittels der Kamera erfassten Bildern BI_i und Liefern von Bildregionen $BR_{i,j}$ des Bilds BI_i , in denen das jeweilige selektierte Objekt $O_{i,j}$ in dem Bild BI_i abgebildet ist, und eine Vorrichtung zur Objektidentitätszuordnung gemäß einem der Merkmale 10 bis 18.

22. Vorrichtung nach Anspruch 21, gekennzeichnet durch eine Ausgabeeinheit zum Ausgeben der erkannten Objekt $O_{i,j}$ und/oder der zugehörigen Bildregionen $BR_{i,j}$ des erfassten Bilds BI_i , wobei die erkannten Objekt $O_{i,j}$ und/oder die zugehörigen Bildregionen $BR_{i,j}$ mit dem zugewiesenen Label gekennzeichnet sind.

Es folgen 4 Blatt Zeichnungen

Anhängende Zeichnungen

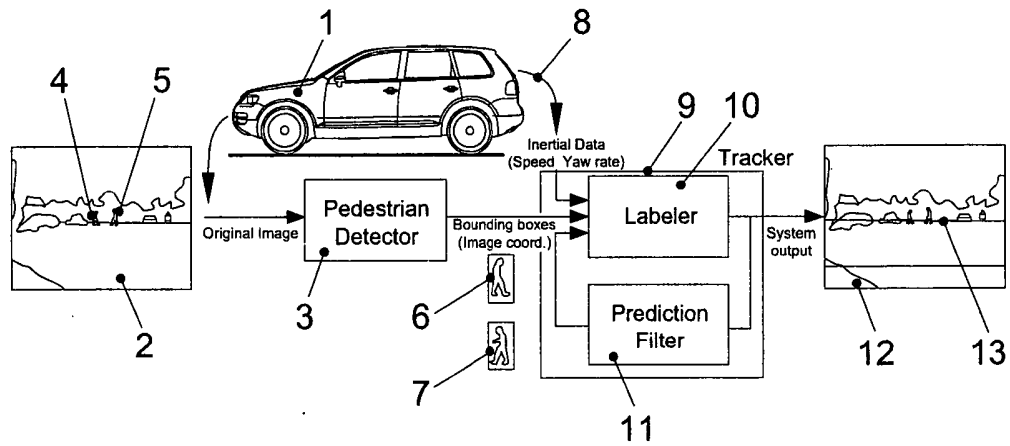


FIG. 1

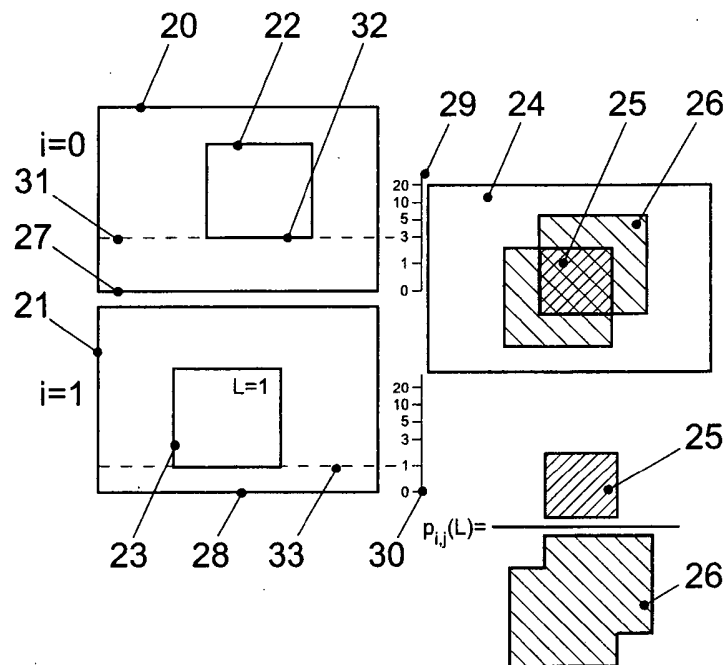


FIG. 2

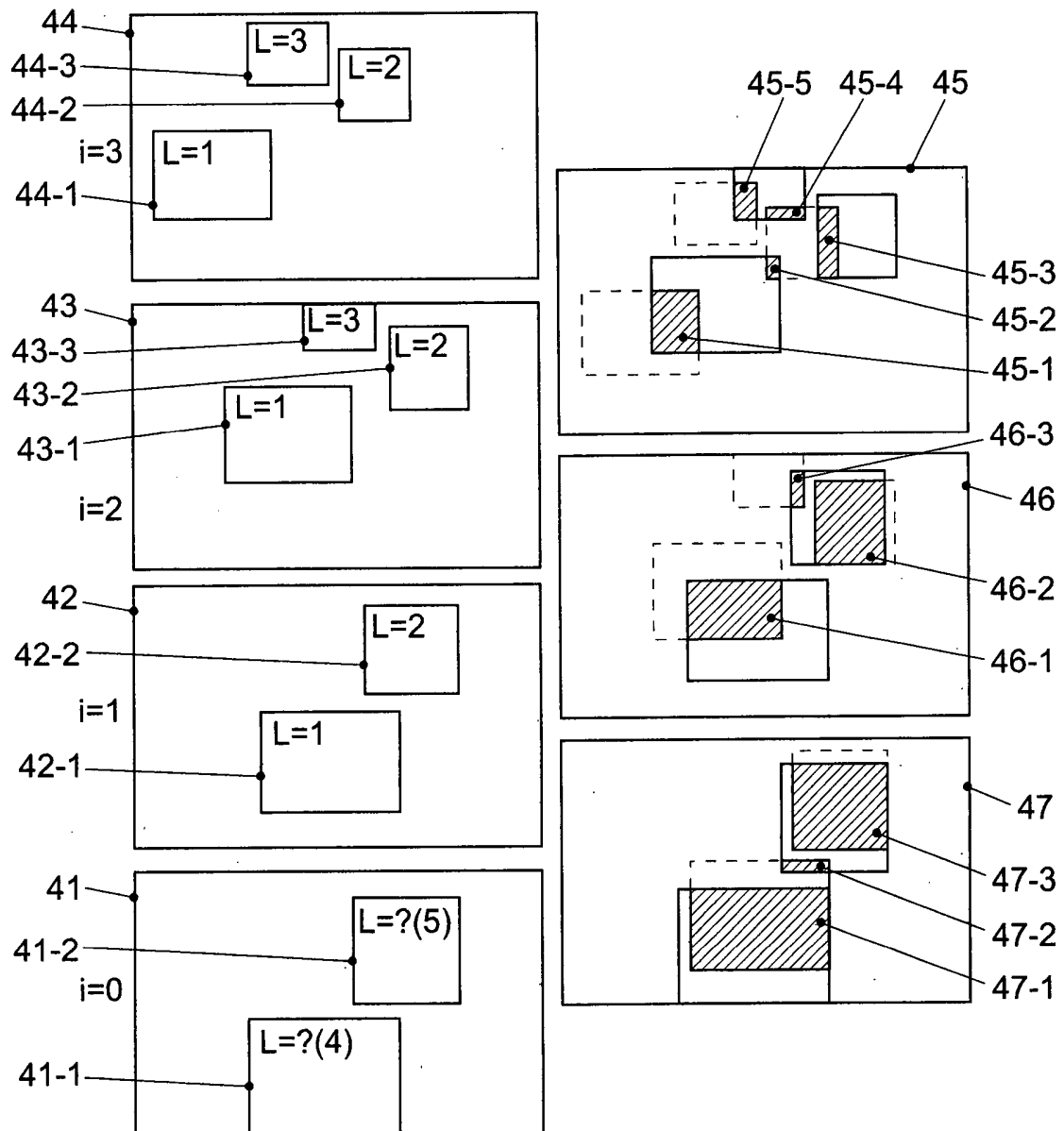


FIG. 3

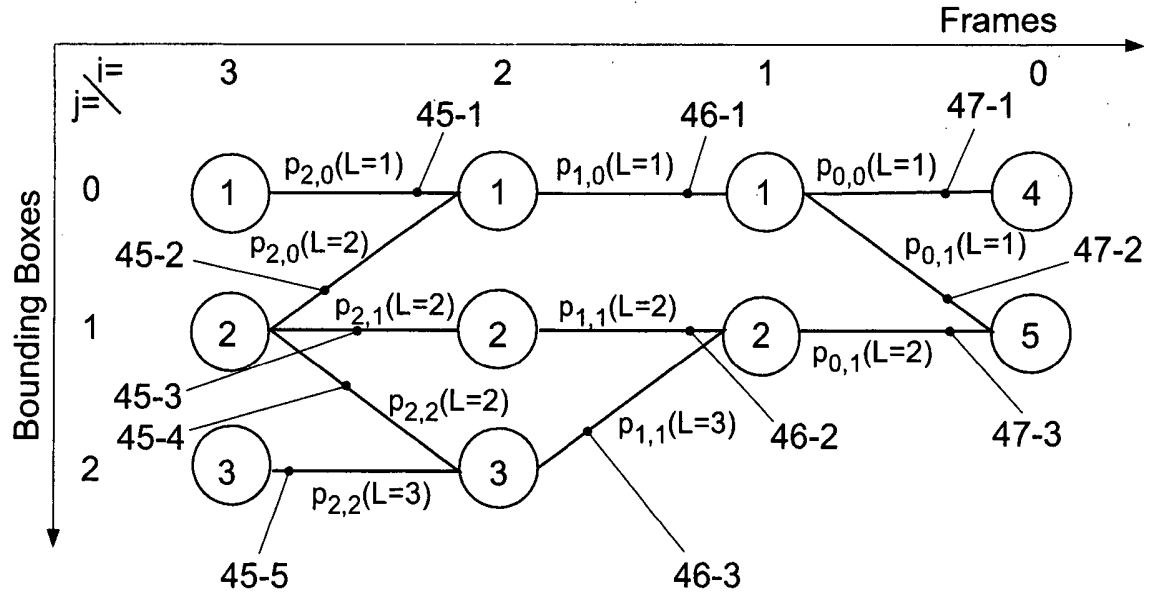


FIG. 4

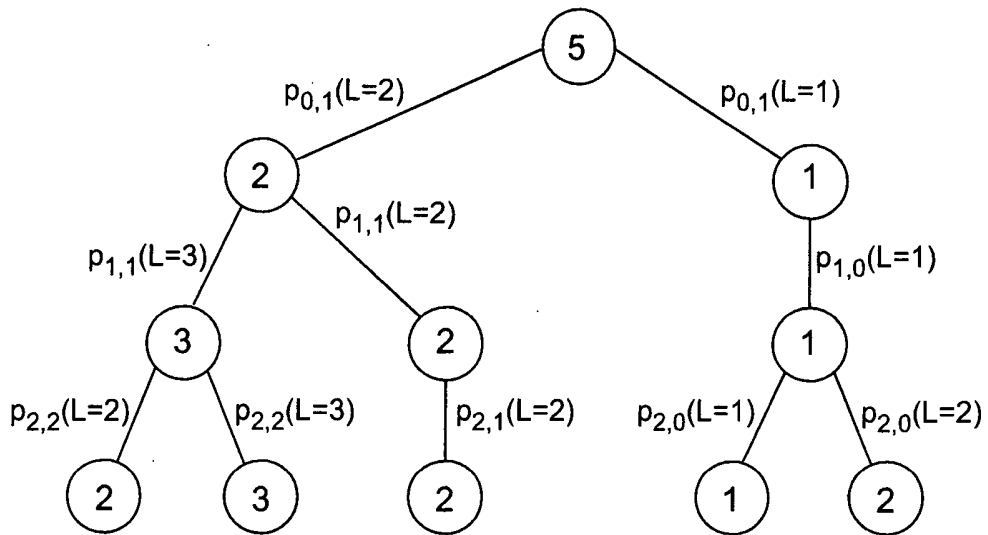


FIG. 5

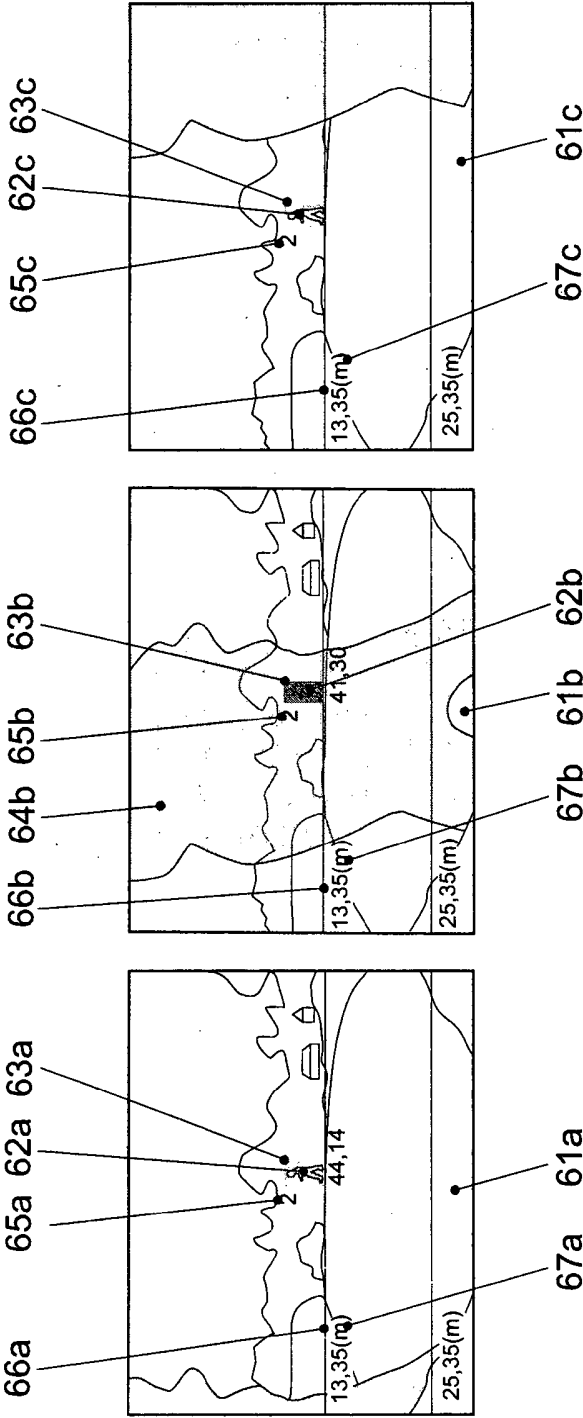


FIG. 6a

FIG. 6b

FIG. 6c