



(10) **DE 10 2020 107 868 A1** 2021.03.04

(12)

## Offenlegungsschrift

(21) Aktenzeichen: **10 2020 107 868.4**

(22) Anmeldetag: **23.03.2020**

(43) Offenlegungstag: **04.03.2021**

(51) Int Cl.: **G06K 9/62 (2006.01)**

(30) Unionspriorität:

**10-2019-0104574 26.08.2019 KR**

(71) Anmelder:

**Samsung Electronics Co., Ltd., Suwon-si,  
Gyeonggi-do, KR**

(74) Vertreter:

**KUHNEN & WACKER Patent- und  
Rechtsanwaltsbüro PartG mbB, 85354 Freising,  
DE**

(72) Erfinder:

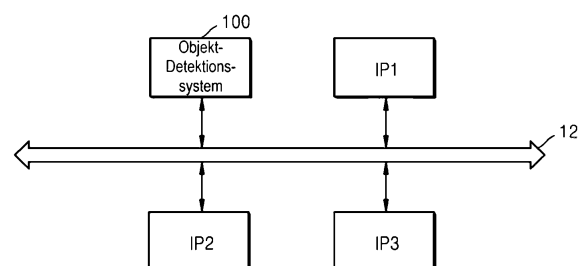
**Kim, Byeongsu, Suwon-si, Gyeonggi-do, KR;  
Ko, Sangsoo, Suwon-si, Gyeonggi-do, KR; Ha,  
Sanghyuck, Suwon-si, Gyeonggi-do, KR**

Prüfungsantrag gemäß § 44 PatG ist gestellt.

**Die folgenden Angaben sind den vom Anmelder eingereichten Unterlagen entnommen.**

(54) Bezeichnung: **OBJEKTDETEKTIONSSYSTEM FÜR DIE DETEKTION EINES OBJEKTS MIT HILFE EINER  
HIERARCHIE-PYRAMIDE UND OBJEKTDETEKTIONSVERFAHREN**

(57) Zusammenfassung: Ein Objektdetektionssystem, das ein Objekt mit Hilfe einer hierarchischen Pyramidenstruktur detektiert, umfasst: einen Pyramidenbildgenerator, der dazu eingerichtet ist, mehrere Eingangsbilder zu empfangen, die mehreren Zeitpunkten entsprechen, und mehrere Pyramidenbilder zu erzeugen, die jedem der mehreren Eingangsbilder entsprechen; einen Objektextraktor, der dazu eingerichtet ist, durch Extrahieren mindestens eines Objekts aus den mehreren Pyramidenbildern mehrere Objektdatenelemente zu erzeugen; und einen Pufferspeicher, der die mehreren Objektdatenelemente auf Objektbasis speichert.



**Beschreibung**

## QUERVERWEIS AUF VERWANDTE ANMELDUNG

**[0001]** Diese Anmeldung beansprucht die Priorität der am 26. August 2019 beim koreanischen Patentamt eingereichten koreanischen Patentanmeldung Nr. 10-2019-0104574, deren Offenbarung durch Inbezugnahme vollumfänglich hier mit aufgenommen wird.

## STAND DER TECHNIK

## Gebiet der Technik

**[0002]** Die Offenbarung bezieht sich auf ein Objektdetektionssystem und insbesondere auf ein Objektdetektionssystem und ein Objektdetektionsverfahren zur Detektion eines Objekts mit Hilfe einer Hierarchie-Pyramide.

## Beschreibung der verwandten Technik

**[0003]** Objektdetektion bezieht sich auf ein Datenverarbeitungsverfahren zur Erfassung eines interessierenden Objekts aus einem Bild oder einem Video und zur Identifizierung oder Klassifizierung des Objekts. Die Objektdetektion spielt in verschiedenen Anwendungen, wie beim automatischen Fahren, bei Fahrassistenten, bei unbemannten Luftfahrzeugen und bei gestenbasierten Interaktionssystemen, eine wesentliche Rolle.

**[0004]** Mit der Entwicklung der Technologie künstlicher Intelligenz wurden Objektdetektionsverfahren und Objektklassifizierungs- und -identifizierungsverfahren entwickelt, die Methoden des tiefgehenden Lernens (Deep-Learning) und Training auf Basis neuronaler Netze verwenden, und diese Verfahren werden weithin eingesetzt.

## KURZFASSUNG DER ERFINDUNG

**[0005]** Ausführungsformen der Offenbarung stellen bereit: ein Objektdetektionssystem, das in der Lage ist, zu mindestens einem mit Hilfe eines Eingangsbildes erzeugten Pyramidenbild eine Zeitinformation hinzuzufügen, die den Zeitpunkt anzeigt, an dem das Eingangsbild aufgenommen wird, und mit Hilfe der hinzugefügten Zeitinformation ein Objekt aus dem Eingangsbild zu detektieren; und ein vom Objektdetektionssystem verwendetes Objektdetektionsverfahren.

**[0006]** Gemäß einem Aspekt der Offenbarung wird ein Objektdetektionssystem bereitgestellt, das umfasst: einen Pyramidenbildgenerator, der eingerichtet ist zum Empfangen eines zu einem ersten Zeitpunkt aufgenommenen ersten Eingangsbildes und eines zu einem zweiten Zeitpunkt aufgenommenen zweiten

Eingangsbildes, und zum Erzeugen eines ersten Pyramidenbildes aus dem ersten Eingangsbild und eines zweiten Pyramidenbildes aus dem zweiten Eingangsbild; einen Objektextraktor, der eingerichtet ist zum Detektieren eines Objekts im ersten Pyramidenbild und im zweiten Pyramidenbild und zum Erzeugen mehrerer Objektdatenelemente, die das Objekt darstellen; und einen Pufferspeicher, der die mehreren Objektdatenelemente speichert, die das im ersten Eingangsbild und im zweiten Eingangsbild detektierte Objekt darstellen.

**[0007]** Gemäß einem anderen Aspekt der Offenbarung wird ein Objektdetektionsverfahren bereitgestellt, das umfasst: Empfangen eines zu einem ersten Zeitpunkt aufgenommenen ersten Eingangsbildes und eines zu einem zweiten Zeitpunkt aufgenommenen zweiten Eingangsbildes; Erzeugen eines mit dem ersten Zeitpunkt verknüpften ersten Pyramidenbildes aus dem ersten Eingangsbild und eines mit dem zweiten Zeitpunkt verknüpften zweiten Pyramidenbildes aus dem zweiten Eingangsbild; und Speichern der mehreren Objektdatenelemente in einem Pufferspeicher.

**[0008]** Gemäß einem weiteren Aspekt der Offenbarung wird ein Fahrassistenzsystem zum Fahren eines Fahrzeugs durch Detektieren eines Objekts bereitgestellt, wobei das Fahrassistenzsystem umfasst: einen Pyramidenbildgenerator, der eingerichtet ist zum Empfangen eines zu einem ersten Zeitpunkt aufgenommenen ersten Eingangsbildes und eines zu einem zweiten Zeitpunkt aufgenommenen zweiten Eingangsbildes und zum Erzeugen eines ersten Pyramidenbildes aus dem ersten Eingangsbild und eines zweiten Pyramidenbildes aus dem zweiten Eingangsbild; einen Objektextraktor, der eingerichtet ist zum Detektieren eines Objekts im ersten Pyramidenbild und im zweiten Pyramidenbild und zum Erzeugen mehrerer Objektdatenelemente, die das Objekt darstellen, durch Verwendung des tiefgehenden Lernens auf Basis eines neuronalen Netzes; einen Pufferspeicher, der die mehreren Objektdatenelemente speichert, die das im ersten Eingangsbild und im zweiten Eingangsbild detektierte Objekt darstellen; und einen Objekttracker, der eingerichtet ist zum Verfolgen des Objekts anhand der im Pufferspeicher gespeicherten mehreren Objektdatenelemente.

## Figurenliste

**[0009]** Ausführungsformen der Offenbarung lassen sich anhand der im Zusammenhang mit den Begleitzeichnungen zu lesenden folgenden Detailbeschreibung besser verstehen, wobei gilt:

**Fig. 1** ist ein Blockschaltbild zur Darstellung eines elektronischen Systems gemäß einer Ausführungsform der Offenbarung;

**Fig. 2** ist ein Blockschaltbild zur Darstellung eines elektronischen Systems gemäß einer Ausführungsform der Offenbarung;

**Fig. 3** ist ein Blockschaltbild zur Darstellung eines Objektdetektionssystems gemäß einer Ausführungsform der Offenbarung;

**Fig. 4** ist ein Flussdiagramm zur Darstellung eines Verfahrens zum Betreiben eines Objektdetektionssystems gemäß einer Ausführungsform der Offenbarung;

**Fig. 5** ist eine Ansicht zur Darstellung eines neuronalen Netzes gemäß einer Ausführungsform der Offenbarung;

**Fig. 6** ist eine Ansicht zur Darstellung eines Verfahrens zum Detektieren eines Objekts gemäß einer Ausführungsform der Offenbarung;

**Fig. 7** ist ein Blockschaltbild zur Darstellung eines Objektdetektionssystems gemäß einer Ausführungsform der Offenbarung;

**Fig. 8** ist ein Flussdiagramm zur Darstellung eines Verfahrens zum Betreiben eines Objektdetektionssystems gemäß einer Ausführungsform der Offenbarung;

**Fig. 9** ist ein Blockschaltbild zur Darstellung eines Objektdetektionssystems gemäß einer Ausführungsform der Offenbarung;

**Fig. 10** ist eine Ansicht zur Darstellung von Objektdaten gemäß einer Ausführungsform der Offenbarung;

**Fig. 11** ist ein Blockschaltbild zur Darstellung eines Objektdetektionssystems gemäß einer Ausführungsform der Offenbarung;

**Fig. 12** ist ein Blockschaltbild zur Darstellung eines Objektdetektionssystems gemäß einer Ausführungsform der Offenbarung;

**Fig. 13** ist eine Ansicht zur Darstellung eines Verfahrens zum Erzeugen eines Pyramidenbildes gemäß einer Ausführungsform der Offenbarung;

**Fig. 14** ist ein Blockschaltbild zur Darstellung eines Objektdetektionssystems gemäß einer Ausführungsform der Offenbarung;

**Fig. 15** ist ein Blockschaltbild zur Darstellung eines Anwendungsprozessors gemäß einer Ausführungsform der Offenbarung; und

**Fig. 16** ist ein Blockschaltbild zur Darstellung eines Fahrassistenzsystems gemäß einer Ausführungsform der Offenbarung.

## DETAILLIERTE BESCHREIBUNG DER AUSFÜHRUNGSFORMEN

**[0010]** Nachfolgend werden Ausführungsformen der Offenbarung unter Verweis auf die begleitenden Zeichnungen im Detail beschrieben.

**[0011]** **Fig. 1** ist ein Blockschaltbild zur Darstellung eines elektronischen Systems gemäß einer Ausführungsform der Offenbarung.

**[0012]** Wie in **Fig. 1** dargestellt, kann ein elektronisches System **10** durch Echtzeitanalyse von Eingangsdaten Informationen extrahieren und anhand der extrahierten Informationen eine Situation bestimmen oder Bauteile einer elektronischen Vorrichtung steuern, in der das elektronische System **10** angeordnet ist. In einem Beispiel kann das elektronische System **10** ein Objekt aus einem Eingangsbild detektieren und das detektierte Objekt verfolgen. Der Ausdruck „Objekt“ kann nach vorliegendem Sprachgebrauch eines oder mehrere der aus der folgenden Aufzählung ausgewählten Elemente bezeichnen, die für einen Benutzer oder das elektronische System **10** von Interesse sind: ein Gebäude, einen Gegenstand, eine Person, ein Tier und eine Pflanze. Das elektronische System **10** kann beispielsweise an einer Drohne, einem fortgeschrittenen Fahrerassistenzsystem (ADAS), einer Robotervorrichtung, einem Smart-TV-Gerät, einem Smartphone, einem medizinischen Gerät, einem mobilen Gerät, einem Bildanzeigergerät, einem Messgerät, einem Gerät des Internets der Dinge (IoT) und so weiter verwendet werden und kann darüber hinaus in einem von verschiedenen elektronischen Geräten installiert sein.

**[0013]** Das elektronische System **10** kann mindestens einen Intellectual-Property-(IP)-Block (IP1, IP2, IP3, ... IPn) und ein Objektdetektionssystem **100** aufweisen. Das elektronische System **10** kann als Beispiel einen ersten bis dritten IP-Block IP1 bis IP3 umfassen, aber das elektronische System kann eine beliebige Zahl von IP-Blöcken aufweisen.

**[0014]** Das elektronische System **10** kann verschiedene IP-Blöcke aufweisen. Die IP-Blöcke können beispielsweise umfassen: eine Verarbeitungseinheit, eine Mehrzahl von in einer Verarbeitungseinheit enthaltenen Kernen, verschiedene Sensormodule, ein Multi-Format-Codec (MFC), ein Video-Modul (zum Beispiel eine Kameraschnittstelle, einen JPEG-(Joint Photographic Experts Group)-Prozessor, einen Videoprozessor, einen Mixer oder dergleichen), einen dreidimensionalen (3D) Grafik-Kern, ein Audiosystem, einen Treiber, einen Anzeigetreiber, flüchtigen Speicher, nichtflüchtigen Speicher, einen Speichercontroller, einen Eingabe-und-Ausgabe-Schnittstellenblock, Cache-Speicher und so weiter. Jeder der ersten bis dritten IP-Blöcke IP1 bis IP3 kann mindes-

tens einen der verschiedenen oben genannten IP-Blöcke umfassen.

**[0015]** Die IP-Blöcke können über mindestens einen Systembus **12** miteinander verbunden sein. Als Standardbusspezifikation kann zum Beispiel ein AMBA-(Advanced Microcontroller Bus Architecture)-Protokoll der Fa. Advanced RISC Machine (ARM) Ltd. verwendet werden. Ein Bustyp des AMBA-Protokolls kann Advanced High-Performance Bus (AHB), Advanced Peripheral Bus (APB), Advanced eXtensible Interface (AXI), AXI4, AXI Coherency Extensions (ACE) und so weiter umfassen.

**[0016]** Das Objektdetektionssystem **100** kann ein Eingangsbild empfangen, kann ein im Eingangsbild enthaltenes Objekt detektieren und kann das detektierte Objekt nachverfolgen oder kann einen Hintergrund extrahieren, indem das detektierte Objekt aus dem Eingangsbild herausgenommen wird. Der Ausdruck „Objekt“ kann nach vorliegendem Sprachgebrauch eines oder mehrere der aus der folgenden Aufzählung ausgewählten Elemente bezeichnen, die für einen Benutzer von Interesse sind und vom Benutzer des elektronischen Systems **10** ausgewählt werden: ein Gebäude, einen Gegenstand, eine Person, ein Tier und eine Pflanze. Der Ausdruck „Hintergrund“ kann sich auf das Restbild beziehen, das übrig bleibt, wenn das Objekt aus einem Einzelbild (Frame) herausgenommen wird. In einer Ausführungsform kann das Objektdetektionssystem **100** ein Objekt, das in Eingangsbilddaten enthalten ist, mit Hilfe eines neuronalen Netzes detektieren und das extrahierte Objekt nachverfolgen. Dies wird im Folgenden unter Verweis auf **Fig. 2** im Detail beschrieben. In einer Ausführungsform kann das Objektdetektionssystem **100** durch Downsampling eines Eingangsbildes mindestens ein Pyramidenbild erzeugen und kann anhand des mindestens einen Pyramidenbildes ein Objekt extrahieren. Hierbei kann eine Struktur, bei der ein Objekt anhand einer Mehrzahl von Pyramidenbildern, die durch Downsampling eines Bilddatenelements erzeugt werden und voneinander abweichende Auflösungen haben, auf hierarchische Weise extrahiert wird, als eine Pyramidenstruktur bezeichnet werden. Dies wird im Folgenden unter Verweis auf **Fig. 6** im Detail beschrieben. Zur Vereinfachung der Beschreibung können sich darüber hinaus die hier gebrauchten Ausdrücke „Eingangsbild“ und „Pyramidenbild“ auf Daten, die dem Eingangsbild entsprechen, beziehungsweise Daten, die dem Pyramidenbild entsprechen, beziehen.

**[0017]** Gemäß der Offenbarung kann bei der Erstellung mindestens eines Pyramidenbildes das Objektdetektionssystem **100** zu dem mindestens einen Pyramidenbild eine Zeitinformation hinzufügen, die einem Zeitpunkt entspricht, an dem das Eingangsbild aufgenommen wird. Das Objektdetektionssystem **100** kann die Zeitinformation zusätzlich zu der räumli-

chen Information entsprechend mindestens einer Pyramide beim Detektieren und Verfolgen eines Objekts zusammen verwenden. Im Ergebnis kann die Detektionsleistung des Objektdetektionssystems **100** verbessert werden.

**[0018]** **Fig. 2** ist ein Blockschaltbild zur Darstellung eines elektronischen Systems gemäß einer Ausführungsform der Offenbarung. **Fig. 2** veranschaulicht insbesondere eine Ausführungsform des in **Fig. 1** dargestellten elektronischen Systems **10**. In Bezug auf das elektronische System **10** von **Fig. 2** wird auf wiederholte Beschreibungen verzichtet, die bereits unter Verweis auf **Fig. 1** gegeben wurden.

**[0019]** Wie in **Fig. 2** dargestellt, kann das elektronische System **10** eine zentrale Verarbeitungseinheit (CPU) **21**, eine neuronale Netzvorrichtung **22**, einen Direktzugriffsspeicher (RAM) **24**, einen Ablagespeicher **25** und ein Sensormodul **26** umfassen. Das elektronische System **10** kann ferner ein Eingabe/Ausgabe-Modul, ein Sicherheitsmodul, eine Leistungsregelungsvorrichtung oder dergleichen umfassen und kann zudem auch verschiedene Berechnungsvorrichtungen umfassen. In einer Ausführungsform können alle oder einige der Bauteile (CPU **21**, neuronale Netzvorrichtung **22**, RAM **24**, Ablagespeicher **25** und Sensormodul **26**) des elektronischen Systems **10** in einem Halbleiterchip verbaut sein. Das elektronische System **10** kann beispielsweise als System-on-Chip (SoC) realisiert werden. Die Bauteile des elektronischen Systems **10** können über mindestens einen Systembus **27** miteinander kommunizieren.

**[0020]** Die CPU **21** kann den Gesamtbetrieb des elektronischen Systems **10** steuern. Die CPU **21** kann einen einzelnen Prozessorkern (Single-Core) oder mehrere Prozessorkerne (Multi-Core) aufweisen. Die CPU **21** kann im Ablagespeicher **25** abgelegte Programme und/oder Daten verarbeiten oder ausführen und in den RAM **24** laden. In einer Ausführungsform kann die CPU **21** durch die Ausführung von im Ablagespeicher **25** gespeicherten Programmen einen Betriebsvorgang des unter Verweis auf **Fig. 1** beschriebenen Objektdetektionssystems **100** ausführen oder kann eine Funktion der neuronalen Netzvorrichtung **22** zur Detektion eines Objekts steuern. Die neuronale Netzvorrichtung **22** kann ein neuronales Netz erzeugen, kann das neuronale Netz trainieren (oder das neuronale Netz zum Lernen veranlassen) oder kann anhand von empfangenen Eingangsdaten eine Berechnung ausführen und kann basierend auf einem Ergebnis der Berechnung ein Informationssignal erzeugen, oder kann das neuronale Netz nachtrainieren. In einem Beispiel kann die neuronale Netzvorrichtung **22** ein Eingangsbild empfangen und kann über im neuronalen Netz enthaltene Berechnungen durch das Extrahieren eines im Eingangsbild enthaltenen Objekts mindestens ein Objektdatenelement erzeugen. Die neuronale Netzvorrichtung **22** kann als

Rechenvorrichtung, Rechenmodul und so weiter bezeichnet werden.

**[0021]** Modelle des neuronalen Netzes können verschiedene Netzmodelle umfassen, wie ein faltendes neuronales Netz (CNN, Convolution Neural Network), das GoogleNet, AlexNet, VGG Network und so weiter umfasst, eine Region mit faltendem neuronalen Netz (R-CNN, Region with Convolution Neural Network), ein RPN (Region Proposal Network), ein RNN (rekurrentes neuronales Netz), ein S-DNN (Stacking-Based Deep Neural Network), ein S-SDNN (State-Space Dynamic Neural Network), ein entfaltetes neuronales Netz, ein DBN (Deep Belief Network), eine RBM (Restricted-Boltzman-Maschine), ein vollfaltendes Netz, ein LSTM-Netz (Long Short-Term Memory), ein Klassifikationsnetz und so weiter, ohne darauf beschränkt zu sein.

**[0022]** Das neuronale Netz **22** kann einen oder mehrere Prozessoren zur Durchführung von Berechnungen entsprechend den Modellen des neuronalen Netzes umfassen. Das neuronale Netz **22** kann außerdem separaten Speicher für die Speicherung von Programmen entsprechend den Modellen des neuronalen Netzes umfassen. Die neuronale Netzvorrichtung **22** kann als neuronaler Netzprozessor, neuronale Netzverarbeitungsvorrichtung, neuronale integrierte Schaltung, Neuronales-Netz-Verarbeitungseinheit (NPU) und so weiter bezeichnet werden.

**[0023]** Die CPU **21** und die neuronale Netzvorrichtung **22** können in dem unter Verweis auf **Fig. 1** beschriebenen Objektdetektionssystem **100** enthalten sein. Das Objektdetektionssystem **100** kann Daten empfangen, die einem Eingangsbild aus einem besonderen IP (zum Beispiel dem RAM **24** oder dem Sensormodul **26**) entsprechen, und kann ein im Eingangsbild enthaltenes Objekt detektieren. In einer Ausführungsform kann die im Objektdetektionssystem **100** enthaltene CPU **21** unter Verwendung eines Eingangsbildes mindestens ein Pyramidenbild erzeugen, das eine Pyramidenstruktur aufweist, und kann bewirken, dass das erzeugte Pyramidenbild eine Zeitinformation enthält, die dem Zeitpunkt entspricht, an dem das Eingangsbild aufgenommen wird. Darüber hinaus kann die im Objektdetektionssystem **100** enthaltene neuronale Netzvorrichtung **22** anhand von räumlichen und zeitlichen Informationen des Pyramidenbildes über netzbasiert trainiertes tiefgehendes Lernen ein im Eingangsbild enthaltenes Objekt extrahieren und kann das extrahierte Objekt verfolgen.

**[0024]** Der RAM **24** kann Programme, Daten oder Anweisungen speichern. Die im Ablagespeicher **25** gespeicherten Programme und/oder Daten können zum Beispiel entsprechend der Ansteuerung durch die CPU **21** oder entsprechend einem Booting-Code in den RAM **24** geladen werden. Der RAM **24**

kann mit Hilfe eines Speicherbausteins wie eines dynamischen RAM (DRAM) oder eines statischen RAM (SRAM) realisiert werden. Der Ablagespeicher **25** ist ein Speicherort zur Datenablage und kann zum Beispiel ein Betriebssystem (OS), verschiedene Programme und verschiedene Datenelemente speichern. Der Ablagespeicher **25** kann flüchtigen Speicher und/oder nichtflüchtigen Speicher umfassen. Das Sensormodul **26** kann Informationen rund um das elektronische System **10** sammeln. Das Sensormodul **26** kann ein Bildsignal aus dem Außenbereich der elektronischen Vorrichtung **10** erfassen oder empfangen und kann das erfasste oder empfangene Bildsignal in Bilddaten, das heißt zu einem Einzelbild (Bildframe), umwandeln. Dazu kann das Sensormodul **26** eine Erfassungsvorrichtung, zum Beispiel mindestens eine von verschiedenen Erfassungsvorrichtungen, wie eine Bildaufnahmevorrichtung, einen Bildsensor, einen LIDAR-(Light Detection and Ranging)-Sensor, einen Ultraschallsensor und einen Infrarotsensor, umfassen oder kann ein Erfassungssignal von der Erfassungsvorrichtung empfangen. In einer Ausführungsform kann das Sensormodul **26** Bilddaten, die ein Einzelbild umfassen, an die CPU **21** oder die neuronale Netzvorrichtung **22** bereitstellen. Das Sensormodul **26** kann beispielsweise einen Bildsensor umfassen, kann durch die Aufnahme von Bildern einer Außenumgebung des elektronischen Systems **10** einen Videostream erzeugen und kann aufeinanderfolgende Einzelbilder des Videostreams nacheinander an die CPU **21** oder die neuronale Netzvorrichtung **22** liefern.

**[0025]** Das elektronische System **10** gemäß einer Ausführungsform der Offenbarung kann beim Erstellen des mindestens einen Pyramidenbildes eine Zeitinformation, die einem Bildaufnahmezeitpunkt der Bilddaten entspricht, zu dem mindestens einen Pyramidenbild hinzufügen und die Zeitinformation zusätzlich zu räumlichen Informationen entsprechend dem mindestens einen Pyramidenbild bei der Detektion und Verfolgung eines Objekts mit Hilfe eines neuronalen Netzes zusammen verwenden. Dadurch kann die Objektdetektionsleistung des Objektdetektionssystems **10** verbessert werden. Der Ausdruck „räumliche Information“ oder „Rauminformation“ kann sich nach vorliegendem Sprachgebrauch auf Pixeldaten eines Eingangsbildes beziehen.

**[0026]** **Fig. 3** ist ein Blockschaltbild zur Darstellung eines Objektdetektionssystems gemäß einer Ausführungsform der Offenbarung.

**[0027]** Wie in **Fig. 3** dargestellt, kann ein Objektdetektionssystem **100** einen Pyramidenbildgenerator **110**, einen Merkmalsextraktor **120**, einen Pufferspeicher **130** und einen Objekttracker **140** aufweisen. Der Pyramidenbildgenerator **110** kann eine Mehrzahl von Eingangsbildern IMs empfangen, die zu mehreren Zeitpunkten aufgenommen wurden, und kann aus

den empfangenen mehreren Eingangsbildern IMs eine Mehrzahl von Pyramidenbildern PIs erzeugen.

**[0028]** Der Pyramidenbildgenerator **110** kann ausgehend von einem Eingangsbild, das einem einzelnen Zeitpunkt entspricht, mehrere Pyramidenbildern erzeugen und kann dafür sorgen, dass jedes der mehreren Pyramidenbilder eine Zeitinformation über die Aufnahmezeit des Eingangsbildes enthält. In einem Beispiel kann der Pyramidenbildgenerator **110** ein erstes Pyramidenbild mit einer ersten Auflösung in Übereinstimmung mit einem ersten Eingangsbild zu einem ersten Zeitpunkt erzeugen, kann ein zweites Pyramidenbild mit einer zweiten Auflösung durch Downsampling des ersten Pyramidenbildes erzeugen, kann ein drittes Pyramidenbild mit einer dritten Auflösung durch Downsampling des zweiten Pyramidenbildes erzeugen und kann Daten, die dem ersten Zeitpunkt entsprechen, an dem das erste Eingangsbild aufgenommen wurde, zu dem ersten Pyramidenbild, dem zweiten Pyramidenbild und dem dritten Pyramidenbild hinzufügen. Die Anzahl der Pyramidenbilder, die durch Downsampling erzeugt werden und unterschiedliche Auflösungen aufweisen, kann adaptiv je nach Menge und/oder Art und Typ der Objekte im Eingangsbild IM bestimmt werden.

**[0029]** Der Pyramidenbildgenerator **110** kann für jedes Eingangsbild, das jeweils einem von mehreren Zeitpunkten entspricht, eine Mehrzahl von Pyramidenbildern erzeugen. In einem Beispiel kann der Pyramidenbildgenerator **110** ein viertes Pyramidenbild mit einer ersten Auflösung in Übereinstimmung mit einem zweiten Eingangsbild zu einem zweiten Zeitpunkt erzeugen, kann ein fünftes Pyramidenbild mit einer zweiten Auflösung durch Downsampling des vierten Pyramidenbildes erzeugen, kann ein sechstes Pyramidenbild mit einer dritten Auflösung durch Downsampling des fünften Pyramidenbildes erzeugen und kann Daten, die dem zweiten Zeitpunkt entsprechen, an dem das zweite Eingangsbild aufgenommen wurde, zu dem vierten Pyramidenbild, dem fünften Pyramidenbild und dem sechsten Pyramidenbild hinzufügen. Mit anderen Worten kann der Einzelzeitpunkt, an dem das Eingangsbild aufgenommen wurde, allen aus dem Eingangsbild erzeugten Pyramidenbildern hinzugefügt werden. In einem Beispiel kann der Pyramidenbildgenerator **110** Informationen über einen Bildaufnahmezeitpunkt aus einem Metabereich eines Eingangsbildes (zum Beispiel IMs) erhalten oder kann den Bildaufnahmezeitpunkt aus einer externen Vorrichtung (zum Beispiel dem Sensor-Modul **26** in **Fig. 2**) beziehen.

**[0030]** In einer Ausführungsform kann der Pyramidenbildgenerator **110** eine Zeitinformation über den Bildaufnahmezeitpunkt eines Eingangsbildes zu einem Header-Bereich jedes der mehreren erzeugten Pyramidenbilder hinzufügen. Die Offenbarung ist hierauf nicht beschränkt, und für den Pyramidenbild-

generator **110** kann ein Bereich zum Hinzufügen der Zeitinformation zu jedem der mehreren Pyramidenbilder verschieden bestimmt werden.

**[0031]** In einer Ausführungsform kann der Pyramidenbildgenerator **110** die Zeitinformation nur zu einigen der mehreren aus dem Eingangsbild erzeugten Pyramidenbildern hinzufügen. In einem Beispiel kann es sein, dass der Pyramidenbildgenerator **110** die Zeitinformation zu dem ersten Pyramidenbild, das eine erste Auflösung aufweist, und zu dem zweiten Pyramidenbild, das eine zweite Auflösung aufweist, aber nicht zu dem dritten Pyramidenbild, das eine dritte Auflösung aufweist, hinzufügt. Mit anderen Worten kann der Einzelzeitpunkt, an dem das Eingangsbild aufgenommen wurde, auch nur zu einigen oder nur zu einer Teilmenge der aus dem Eingangsbild erzeugten Pyramidenbilder hinzugefügt werden.

**[0032]** In einer Ausführungsform kann der Pyramidenbildgenerator **110** die Anzahl an Pyramidenbildern, die aus den mehreren Eingangsbildern, die den mehreren Zeitpunkten entsprechen und eine gleiche Auflösung aufweisen, erzeugt werden, verschieden bestimmen. In einem Beispiel kann der Pyramidenbildgenerator **110** aus den mehreren Eingangsbildern, die den mehreren Zeitpunkten entsprechen, eine erste Menge an ersten Pyramidenbildern mit einer ersten Auflösung erzeugen und eine zweite Menge an zweiten Pyramidenbildern mit einer zweiten Auflösung erzeugen. Das heißt, die erste Menge an ersten Pyramidenbildern kann aus einer ersten Menge von zu unterschiedlichen Zeitpunkten aufgenommenen Eingangsbildern erzeugt werden, und die zweite Menge an zweiten Pyramidenbildern kann aus einer zweiten Menge von zu unterschiedlichen Zeitpunkten aufgenommenen Eingangsbildern erzeugt werden.

**[0033]** Der Merkmalsextraktor **120** kann vom Pyramidenbildgenerator **110** eine Mehrzahl von Pyramidenbildern PIs empfangen und kann aus den mehreren Pyramidenbildern PIs eine Mehrzahl von Objektdatenelementen ODs extrahieren. In einer Ausführungsform kann der Merkmalsextraktor **120** die mehreren Objektdatenelemente ODs über ein auf Basis eines neuronalen Netzes trainiertes tiefgehendes Lernen aus den mehreren Pyramidenbildern PIs extrahieren. In einem Beispiel kann der Merkmalsextraktor **120** durch die neuronale Netzvorrichtung **22** von **Fig. 2** realisiert sein.

**[0034]** Gemäß der Offenbarung kann der Merkmalsextraktor **120** aus einer Mehrzahl von Pyramidenbildern, die einer Mehrzahl von Zeitpunkten entsprechen, Objektdatenelemente extrahieren, die dem gleichen Objekt entsprechen. Der Merkmalsextraktor **120** kann die den mehreren Zeitpunkten entsprechenden mehreren Pyramidenbilder, die einer gleichen Auflösung entsprechen, vom Pyramidenbildgenerator **110** empfangen und kann auf Grundlage von

in den mehreren Pyramidenbildern enthaltenen einzelnen Zeitinformationen ein Objekt detektieren und extrahieren und dadurch die Objektdatenelemente erzeugen. In einem Beispiel kann der Merkmalsextraktor **120** ein erstes Objekt aus mindestens einem ersten Pyramidenbild mit einer ersten Auflösung extrahieren, kann ein zweites Objekt aus mindestens einem zweiten Pyramidenbild mit einer zweiten Auflösung extrahieren und kann ein drittes Objekt aus mindestens einem dritten Pyramidenbild mit einer dritten Auflösung extrahieren. In einer Ausführungsform kann das erste bis dritte Objekt ganz verschiedene Abstände zu einer Bildaufnahme-Position aufweisen und das wird unter Verweis auf **Fig. 6** näher beschrieben. Der Merkmalsextraktor **120** kann mindestens ein extrahiertes Objektdatenelement ODs im Pufferspeicher **130** speichern. In einer Ausführungsform kann der Merkmalsextraktor **120** das mindestens eine Objektdatenelement ODs entsprechend der jeweiligen Objektart in unterschiedlichen Pufferspeichern oder in unterschiedlichen Bereichen eines Pufferspeichers speichern.

**[0035]** In einem Beispiel kann der Merkmalsextraktor **120** Objektdatenelemente ODs, die einem ersten Objekt und entsprechend mehreren Zeitpunkten entsprechen, in einem ersten Bereich des Pufferspeichers **130** speichern, kann Objektdatenelemente ODs, die einem zweiten Objekt und entsprechend mehreren Zeitpunkten entsprechen, in einem zweiten Bereich des Pufferspeichers **130** speichern und kann Objektdatenelemente ODs, die einem dritten Objekt und entsprechend mehreren Zeitpunkten entsprechen, in einem dritten Bereich des Pufferspeichers **130** speichern. In einem Beispiel kann der Merkmalsextraktor **120** auf Grundlage einer Verkettungsoperation eine Mehrzahl von Objektdaten auf Objektbasis im Pufferspeicher **130** speichern.

**[0036]** Der Pufferspeicher **130** kann Objektdatenelemente ODs speichern. Dazu kann der Pufferspeicher **130** flüchtigen Speicher und/oder nichtflüchtigen Speicher umfassen. Gemäß einer Ausführungsform der Offenbarung kann der Pufferspeicher **130** die Objektdatenelemente ODs in unterschiedlichen Speicherbereichen auf Objektbasis speichern. In einer anderen Ausführungsform kann der Pufferspeicher **130** mehrere Speicherbauelemente umfassen und kann die Objektdatenelemente ODs in unterschiedlichen Speicherbauelementen auf Objektbasis speichern.

**[0037]** Der Objekttracker **140** kann die Objektdatenelemente ODs empfangen und kann anhand der Objektdatenelemente ODs ein Objekt verfolgen. In einer Ausführungsform der Offenbarung kann der Objekttracker **140** bei der Verfolgung eines Objekts die mehreren Zeitpunkten entsprechenden Objektdatenelemente ODs verwenden. In einem Beispiel kann der Objekttracker **140** ein erstes Objekt mit Hilfe mehrerer Objektdatenelemente, die einer ersten Auflö-

sung entsprechen, verfolgen und kann ein zweites Objekt mit Hilfe mehrerer Objektdatenelemente, die einer zweiten Auflösung entsprechen, verfolgen.

**[0038]** Gemäß einer Ausführungsform der Offenbarung kann der Objekttracker **140** die mehreren Zeitpunkten entsprechenden Objektdatenelemente ODs zum Extrahieren eines Objekts verwenden. In einem Beispiel kann ein Objekt einen größeren Datenänderungsumfang über die Zeit als ein Hintergrund aufweisen, und der Objekttracker **140** kann das Objekt effizient verfolgen, indem er die mehreren Objektdatenelemente ODs, die mehreren Zeitpunkten entsprechen, miteinander vergleicht.

**[0039]** **Fig. 4** ist ein Flussdiagramm zur Darstellung eines Verfahrens zum Betreiben eines Objektdetektionssystems gemäß einer Ausführungsform der Offenbarung.

**[0040]** Wie in **Fig. 3** und **Fig. 4** dargestellt, kann das Objektdetektionssystem **100** eine Mehrzahl von Eingangsbildern, die einer entsprechenden Mehrzahl von Zeitpunkten entsprechen, empfangen (S110) und kann zu den mehreren Eingangsbildern jeweils einzelne Zeitinformationen über die Bildaufnahmezeitpunkte der mehreren Eingangsbilder hinzufügen (S120). Das Objektdetektionssystem **100** kann unter Verwendung der mehreren Eingangsbilder, denen die einzelnen Zeitinformationen hinzugefügt werden, eine Mehrzahl von Pyramidenbildern erstellen, die jedem der mehreren Eingangsbilder entsprechen. In einer Ausführungsform kann das Objektdetektionssystem **100** die mehreren Pyramidenbilder durch wiederholtes Downsampling des jeweiligen Eingangsbildes, dem eine Zeitinformation hinzugefügt wird, erzeugen.

**[0041]** Das Objektdetektionssystem **100** kann eine Mehrzahl von Objektdatenelementen, die den mehreren Zeitpunkten entsprechen, durch Extraktion eines Objekts aus jedem der mehreren Pyramidenbilder erzeugen (S140). In einer Ausführungsform kann das Objektdetektionssystem **100** die mehreren Objektdatenelemente über ein auf Grundlage eines neuronalen Netzes trainiertes tiefgehendes Lernen aus den mehreren Pyramidenbildern erzeugen. In einem Beispiel kann das Objektdetektionssystem **100** eine Mehrzahl von zeitpunktweisen Objektdatenelementen, die einem Objekt entsprechen, erzeugen.

**[0042]** Das Objektdetektionssystem **100** kann die mehreren Objektdatenelemente auf Objektbasis speichern (S150). In einer Ausführungsform kann das Objektdetektionssystem **100** die mehreren zeitpunktweisen Objektdatenelemente in unterschiedlichen Bereichen eines Pufferspeichers auf Objektbasis speichern und kann die mehreren zeitpunkt-spezifischen Objektdaten auch unter Verwendung einer Verkettungsoperation im Pufferspeicher abspeichern. Das Objektdetektionssystem **100** kann die Po-

sition und/oder Bewegung eines Objekts sowohl mit Hilfe der mehreren auf Objektbasis gespeicherten Objektdatenelemente als auch mit Hilfe der einzelnen Zeitinformationen verfolgen (S160).

**[0043]** Fig. 5 ist eine Ansicht zur Darstellung eines neuronalen Netzes gemäß einer Ausführungsform der Offenbarung. Fig. 5 veranschaulicht als Beispiel einer neuronalen Netzstruktur ein faltendes neuronales Netz. Fig. 5 veranschaulicht ein Beispiel eines vom Merkmalsextraktor 120 aus Fig. 3 verwendeten neuronalen Netzes.

**[0044]** Wie in Fig. 5 dargestellt, kann ein neuronales Netz NN eine Mehrzahl von Schichten (Layern) L1, L2 ... bis Ln umfassen. Jede der mehreren Schichten L1, L2 ... bis Ln kann eine lineare oder nichtlineare Schicht sein. In einer Ausführungsform können mindestens eine lineare Schicht und mindestens eine nichtlineare Schicht miteinander gekoppelt sein und damit als eine Schicht bezeichnet werden. Eine lineare Schicht kann zum Beispiel eine faltende Schicht und eine vollverbundene Schicht umfassen, und eine nichtlineare Schicht kann Pooling- und Aktivierungsschichten umfassen.

**[0045]** Zum Beispiel kann eine erste Schicht L1 eine faltende Schicht sein, eine zweite Schicht L2 kann eine Pooling-Schicht sein, und eine n-te Schicht Ln kann eine vollverbundene Schicht als Ausgangsschicht sein. Das neuronale Netz NN kann ferner eine Aktivierungsschicht und eine Schicht, die eine andere Art von Operation abweichend von den oben erörterten Vorgängen ausführt, umfassen.

**[0046]** Jede der mehreren Schichten L1 bis Ln kann als Input-Merkmalskarte Eingangsdaten (zum Beispiel ein Einzelbild) oder eine von einer vorhergehenden Schicht erzeugte Merkmalskarte empfangen und kann die Input-Merkmalskarte berechnen, wodurch eine Output-Merkmalskarte oder ein Erkennungssignal REC erzeugt wird. Hierbei kann sich die Merkmalskarte auf Daten beziehen, in denen verschiedene Merkmale der Eingangsdaten dargestellt sind. Merkmalskarten FM1 bis FMn können beispielsweise jeweils die Form einer zweidimensionalen Matrix oder einer dreidimensionalen Matrix (oder eines Tensors) aufweisen. Die Merkmalskarten FM1 bis FMn weisen jeweils eine Breite W (oder Spalte), eine Höhe H (oder Zeile) und eine Tiefe D auf, die einer x-Achse, einer y-Achse beziehungsweise einer z-Achse in einem Koordinatensystem entsprechen. Hierbei kann die Tiefe D als Anzahl der Kanäle bezeichnet werden.

**[0047]** Die erste Schicht L1 kann eine zweite Merkmalskarte FM2 erzeugen, indem sie eine erste Merkmalskarte FM1 mit einer Gewichtungskarte WM faltet. Die Gewichtungskarte WM kann die erste Merkmalskarte FM1 filtern und kann auch als Filter oder

Kernel bezeichnet werden. Die Tiefe, das heißt die Anzahl der Kanäle, der Gewichtungskarte WM kann gleich der Tiefe, das heißt der Anzahl der Kanäle, der ersten Merkmalskarte FM1 sein, und die Faltung kann zwischen den gleichen Kanälen der Gewichtungskarte WM und der ersten Merkmalskarte FM1 ausgeführt werden. Die Gewichtungskarte WM kann so verschoben werden, dass die Gewichtungskarte WM die erste Merkmalskarte FM1 traversiert, indem die erste Merkmalskarte FM1 als Gleitfenster genommen wird. Der Verschiebungsbetrag kann mit dem Ausdruck „Schrittlänge“ oder „Schritt“ bezeichnet werden. In der Gewichtungskarte WM enthaltene Gewichtungswerte können während jeder Verschiebung jeweils mit allen Pixeldatenelementen in einem die erste Merkmalskarte überlappenden Bereich multipliziert werden. Dann können die Ergebnisse durch Summierung aufaddiert werden. Datenelemente der ersten Merkmalskarte FM1 in dem Bereich, wo jeder der in der Gewichtungskarte WM enthaltenen Gewichtungswerte die erste Merkmalskarte FM1 überlappt, können als Extraktionsdaten bezeichnet werden. Bei Durchführung der Faltung zwischen der ersten Merkmalskarte FM1 und der Gewichtungskarte WM kann ein Kanal der zweiten Merkmalskarte FM2 erzeugt werden. Obwohl in Fig. 3 nur eine Gewichtungskarte WM abgebildet ist, kann mit der ersten Merkmalskarte FM1 eine Faltung mehrerer Gewichtungskarten erfolgen, wodurch mehrere Kanäle der zweiten Merkmalskarte FM2 erzeugt werden können. Mit anderen Worten kann die Anzahl der Kanäle der zweiten Merkmalskarte FM2 der Anzahl der Gewichtungskarten entsprechen.

**[0048]** Die zweite Schicht L2 kann durch Änderung einer räumlichen Größe der zweiten Merkmalskarte FM2 per Pooling eine dritte Merkmalskarte FM3 erzeugen. Der Ausdruck "Pooling" kann mit dem Ausdruck „Sampling“ oder „Downsampling“ bezeichnet werden. Ein zweidimensionales Pooling-Fenster PW kann auf der zweiten Merkmalskarte FM2 in Einheiten der Größe des Pooling-Fensters PW verschoben werden, und es kann ein Maximalwert aus den Pixeldatenelementen (oder ein Durchschnittswert der Pixeldatenelemente) in einem das Pooling-Fenster PW überlappenden Bereich ausgewählt werden. So kann die dritte Merkmalskarte FM3 erzeugt werden, die eine von einer Größe der zweiten Merkmalskarte FM2 abweichende räumliche Größe aufweist. Die Anzahl der Kanäle der dritten Merkmalskarte FM3 ist gleich der Anzahl der Kanäle der zweiten Merkmalskarte FM2. In einer Ausführungsform der Offenbarung kann das Pooling-Fenster PW während seiner Verschiebung auf den Merkmalskarten FM2 und FM3 verschiedene Operationen zur Extraktion eines Objekts ausführen.

**[0049]** Die n-te Schicht Ln kann durch Erstellen einer Kombination von Merkmalen einer n-ten Merkmalskarte FMn eine Klasse (CL) der Eingangsda-



ten abgrenzen. Darüber hinaus kann die n-te Schicht Ln das der Klasse entsprechende Erkennungssignal REC erzeugen. In einer Ausführungsform können die Eingangsdaten einem mit Hilfe eines Eingangsbildes erzeugten Pyramidenbild entsprechen, und die n-te Schicht Ln kann ein Objekt erkennen, indem sie die Klasse, die dem Objekt, das in einem durch Einzelbilddaten dargestellten Bild enthalten ist, entspricht, auf Grundlage der von einer vorhergehenden Schicht bereitgestellten n-ten Merkmalskarte FMn extrahiert. Dementsprechend kann ein dem erkannten Objekt entsprechendes Erkennungssignal REC ausgegeben werden. In einer Ausführungsform kann der Merkmalsextraktor **120** (Fig. 3) im Pufferspeicher **130** (Fig. 3) das Erkennungssignal REC als Objektdaten ODs abspeichern oder kann im Pufferspeicher **130** (Fig. 3) die mit Hilfe des Erkennungssignals REC erzeugten Objektdaten ODs speichern.

**[0050]** Fig. 6 ist eine Ansicht zur Darstellung eines Verfahrens zum Detektieren eines Objekts gemäß einer Ausführungsform der Offenbarung.

**[0051]** Wie in Fig. 3 und Fig. 6 dargestellt, kann der Pyramidenbildgenerator **110** basierend auf einem zu einem ersten Zeitpunkt t1 aufgenommenen Eingangsbild ein erstes Pyramidenbild PI1\_1 mit einer ersten Auflösung erzeugen. Der Pyramidenbildgenerator **110** kann per Downsampling des ersten Pyramidenbildes PI1\_1 ein zweites Pyramidenbild PI1\_2 mit einer zweiten Auflösung erzeugen. Der Pyramidenbildgenerator **110** kann per Downsampling des zweiten Pyramidenbildes PI1\_2 ein drittes Pyramidenbild PI1\_3 mit einer dritten Auflösung erzeugen. Der Pyramidenbildgenerator **110** kann ein Downsampling basierend auf einem voreingestellten Ganzzahlverhältnis ausführen, und in einem Beispiel kann der Pyramidenbildgenerator **110** ein Downsampling durchführen, indem die Auflösung eines bestehenden Bildes

mit  $\frac{1}{2}$  oder  $\frac{1}{\sqrt{2}}$  multipliziert wird.

**[0052]** Der Pyramidenbildgenerator **110** kann basierend auf dem zu einem zweiten Zeitpunkt t2 aufgenommenen Eingangsbild ein viertes Pyramidenbild PI2\_1 mit der ersten Auflösung erzeugen. Der Pyramidenbildgenerator **110** kann per Downsampling des vierten Pyramidenbildes PI2\_1 ein fünftes Pyramidenbild PI2\_2 mit der zweiten Auflösung erzeugen. Der Pyramidenbildgenerator **110** kann per Downsampling des fünften Pyramidenbildes PI2\_2 ein sechstes Pyramidenbild PI2\_3 mit der dritten Auflösung erzeugen.

**[0053]** Gemäß einer Ausführungsform der Offenbarung kann der Pyramidenbildgenerator **110** eine dem ersten Zeitpunkt t1 entsprechende Zeitinformation zu dem ersten bis dritten Pyramidenbild PI1\_1 bis PI1\_3 hinzufügen und kann eine dem zweiten Zeitpunkt

t2 entsprechende Zeitinformation zu dem vierten bis sechsten Pyramidenbild PI2\_1 bis PI2\_3 hinzufügen.

**[0054]** Der Merkmalsextraktor **120** kann aus unterschiedlichen Pyramidenbildern mehrere Objekte extrahieren. In einem Beispiel kann der Merkmalsextraktor **120** mit Hilfe des dritten Pyramidenbildes PI1\_3 und des sechsten Pyramidenbildes P2\_3, die beide die dritte Auflösung, das heißt die niedrigste Auflösung, aufweisen, ein erstes Objekt **O1** extrahieren, das einem Bildaufnahmegerät, welches das Eingangsbild aufgenommen hat, am nächsten liegt. Gleichmaßen kann der Merkmalsextraktor **120** nach dem ersten Objekt **O1** mit Hilfe des zweiten Pyramidenbildes PI1\_2 und des fünften Pyramidenbildes P2\_2, die beide die zweite Auflösung, das heißt die zweitniedrigste Auflösung, aufweisen, ein zweites Objekt **O2** extrahieren, das einem Bildaufnahmegerät, welches das Eingangsbild aufgenommen hat, am zweitnächsten liegt. Darüber hinaus kann der Merkmalsextraktor **120** nach dem zweiten Objekt **O2** mit Hilfe des ersten Pyramidenbildes PI1\_1 und des vierten Pyramidenbildes P2\_1, die beide die erste Auflösung, das heißt die höchste Auflösung, aufweisen, ein drittes Objekt **O3** extrahieren, das einem Bildaufnahmegerät, welches das Eingangsbild aufgenommen hat, am drittnächsten liegt.

**[0055]** Der Objekttracker **140** kann ein Objekt basierend auf mehreren vom Merkmalsextraktor **120** erzeugten Objektdatenelementen verfolgen. Gemäß einer Ausführungsform der Offenbarung kann der Objekttracker **140** für die Verfolgung eines Objekts Objektdatenelemente, die mehreren Zeitpunkten entsprechen, unter Zuhilfenahme von einzelnen Zeitinformationen verwenden. In einem Beispiel kann der Objekttracker **140** zur Verfolgung des dritten Objekts **O3**, zusätzlich zu den aus dem ersten Pyramidenbild PI1\_1 erzeugten Objektdaten, ferner Objektdaten, die aus dem vierten Pyramidenbild PI2\_1 erzeugt werden, und eine Zeitdifferenz zwischen dem ersten Zeitpunkt t1 und dem zweiten Zeitpunkt t2 verwenden.

**[0056]** Obwohl Fig. 6 ein Beispiel für die Extraktion von drei Objekten mit Hilfe von Pyramidenbildern mit drei Auflösungen zeigt, handelt es sich hierbei um ein reines Beispiel; die Anzahl der für die Objektextraktion verwendeten Pyramidenbilder kann verschieden bestimmt werden, und auch die Anzahl der mit Hilfe dieser Pyramidenbilder extrahierten Objekte kann verschieden bestimmt werden. Darüber hinaus ist davon auszugehen, dass die Offenbarung auch auf eine Ausführungsform angewendet werden kann, bei der zwei oder mehrere Objekte mit Hilfe eines einzigen Pyramidenbildes extrahiert werden können.

**[0057]** Fig. 7 ist ein Blockschaltbild zur Darstellung eines Objektdetektionssystems gemäß einer Ausführungsform der Offenbarung. Fig. 7 ist insbesondere

re ein Blockschaltbild zur Darstellung einer Ausführungsform, in der ein Objekt anhand von Eingangsbildern, die zu zwei Zeitpunkten aufgenommen werden, extrahiert wird. Auf Beschreibungen, die sich in Bezug auf **Fig. 3** doppeln, wird verzichtet.

**[0058]** Wie in **Fig. 6** und **Fig. 7** dargestellt, kann das Objektdetektionssystem **100** den Pyramidenbildgenerator **110**, einen ersten Merkmalsextraktor **121**, einen zweiten Merkmalsextraktor **122**, den Pufferspeicher **130** und den Objekttracker **140** aufweisen. Der Pyramidenbildgenerator **110** kann ein zum ersten Zeitpunkt  $t_1$  aufgenommenes erstes Eingangsbild  $IM_1$  und ein zum zweiten Zeitpunkt  $t_2$  aufgenommenes zweites Eingangsbild  $IM_2$  empfangen. Der Pyramidenbildgenerator **110** kann einen Datenmanager **111** und einen Downsampler **112** umfassen. Der Datenmanager **111** kann das erste Pyramidenbild  $PI1_1$  erzeugen, indem er dem ersten Eingangsbild  $IM_1$  eine dem ersten Zeitpunkt entsprechende Zeitinformation hinzufügt, und kann das vierte Pyramidenbild  $PI2_1$  erzeugen, indem er dem zweiten Eingangsbild  $IM_2$  eine dem ersten Zeitpunkt entsprechende Zeitinformation hinzufügt.

**[0059]** Der Downsampler **112** kann per Downsampling des ersten Pyramidenbildes  $PI1_1$  das zweite Pyramidenbild  $PI1_2$  und das dritte Pyramidenbild  $PI1_3$  erzeugen. Darüber hinaus kann der Downsampler **112** per Downsampling des vierten Pyramidenbildes  $PI2_1$  das fünfte Pyramidenbild  $PI2_2$  und das sechste Pyramidenbild  $PI2_3$  erzeugen.

**[0060]** Der Pyramidenbildgenerator **110** kann das erzeugte erste bis dritte Pyramidenbild  $PI1_1$  bis  $PI1_3$  zum ersten Merkmalsextraktor **121** hinzufügen und kann das erzeugte vierte bis sechste Pyramidenbild  $PI2_1$  bis  $PI2_3$  zum zweiten Merkmalsextraktor **122** hinzufügen. Der erste Merkmalsextraktor **121** kann die dem ersten Zeitpunkt  $t_1$  entsprechenden ersten bis dritten Pyramidenbilder  $PI1_1$  bis  $PI1_3$  empfangen und kann, durch Extraktion von entsprechenden Objekten aus den empfangenen ersten bis dritten Pyramidenbildern  $PI1_1$  bis  $PI1_3$ , erste bis dritte Objektdaten  $OD1_1$ ,  $OD1_2$  und  $OD1_3$  erzeugen. Im Beispiel von **Fig. 6** kann der erste Merkmalsextraktor **121** die ersten Objektdaten  $OD1_1$  durch Extraktion eines ersten Objekts aus dem ersten Pyramidenbild  $PI1_1$  erzeugen, kann die zweiten Objektdaten  $OD1_2$  durch Extraktion eines zweiten Objekts aus dem zweiten Pyramidenbild  $PI1_2$  erzeugen, und kann die dritten Objektdaten  $OD1_3$  durch Extraktion eines dritten Objekts aus dem dritten Pyramidenbild  $PI1_3$  erzeugen. Der zweite Merkmalsextraktor **122** kann die dem zweiten Zeitpunkt  $t_2$  entsprechenden vierten bis sechsten Pyramidenbilder  $PI2_1$ ,  $PI2_2$  und  $PI2_3$  empfangen und kann, durch Extraktion von entsprechenden Objekten aus den empfangenen vierten bis sechsten Pyramidenbildern  $PI2_1$  bis  $PI2_3$ ,

3, vierte bis sechste Objektdaten  $OD2_1$  bis  $OD2_3$  erzeugen.

**[0061]** Der erste Merkmalsextraktor **121** kann die erzeugten ersten Objektdaten  $OD1_1$  in einem ersten Bereich  $Ar_1$  des Pufferspeichers **130** speichern, kann die erzeugten zweiten Objektdaten  $OD1_2$  in einem zweiten Bereich  $Ar_2$  des Pufferspeichers **130** speichern und kann die erzeugten dritten Objektdaten  $OD1_3$  in einem dritten Bereich  $Ar_3$  des Pufferspeichers **130** speichern. Der zweite Merkmalsextraktor **122** kann die erzeugten vierten Objektdaten  $OD2_1$  im ersten Bereich  $Ar_1$  des Pufferspeichers **130** speichern, kann die erzeugten fünften Objektdaten  $OD2_2$  im zweiten Bereich  $Ar_2$  des Pufferspeichers **130** speichern und kann die erzeugten sechsten Objektdaten  $OD2_3$  im dritten Bereich  $Ar_3$  des Pufferspeichers **130** speichern.

**[0062]** In einer Ausführungsform können der erste Merkmalsextraktor **121** und der zweite Merkmalsextraktor **122** unter Verwendung einer Verkettungsoperation die erzeugten ersten bis sechsten Objektdaten  $OD1_1$ ,  $OD1_2$ ,  $OD1_3$ ,  $OD2_1$ ,  $OD2_2$  und  $OD2_3$  im Pufferspeicher **130** speichern. Obwohl **Fig. 7** eine Ausführungsform darstellt, bei der Objektdatenelemente (zum Beispiel  $OD1_1$  bis  $OD2_3$ ) in unterschiedlichen Bereichen (zum Beispiel  $Ar_1$  bis  $Ar_3$ ) eines Pufferspeichers **130** auf Objektbasis gespeichert werden, kann die Offenbarung außerdem auch auf eine Ausführungsform angewendet werden, bei der die Objektdatenelemente (zum Beispiel  $OD1_1$  bis  $OD2_3$ ) in unterschiedlichen Pufferspeichern auf Objektbasis gespeichert werden, wie oben beschrieben.

**[0063]** Der Objekttracker **140** kann Objekte mit Hilfe der auf Objektbasis gespeicherten ersten bis sechsten Objektdaten  $OD1_1$  bis  $OD2_3$  verfolgen. In einem Beispiel kann der Objekttracker **140** die im ersten Bereich  $Ar_1$  des Pufferspeichers **130** gespeicherten ersten Objektdaten  $OD1_1$  und vierten Objektdaten  $OD2_1$  lesen und kann das erste Objekt mit Hilfe der ersten Objektdaten  $OD1_1$  und vierten Objektdaten  $OD2_1$  verfolgen. Obwohl **Fig. 7** eine Ausführungsform darstellt, bei der die Extraktion von Objekten anhand von Eingangsbildern erfolgt, die zwei Zeitpunkten entsprechen, ist dies ein bloßes Beispiel, und es ist davon auszugehen, dass die Offenbarung auch auf eine Ausführungsform angewendet werden kann, bei der die Extraktion von Objekten anhand von Eingangsbildern erfolgt, die mehr als zwei Zeitpunkten entsprechen.

**[0064]** **Fig. 8** ist ein Flussdiagramm zur Darstellung eines Verfahrens zum Betreiben eines Objektdetektionssystems gemäß einer Ausführungsform der Offenbarung. **Fig. 8** zeigt insbesondere ein Verfahren zum Betreiben eines Objektdetektionssystems, das ein Objekt für jede Auflösung mit Hilfe einer unterschiedlichen Anzahl von Pyramidenbildern detektiert.

**[0065]** Wie in **Fig. 3** und **Fig. 8** dargestellt, kann das Objektdetektionssystem **100** mit Hilfe mehrerer Eingangsbilder, die mehreren Zeitpunkten entsprechen, einen ersten Pyramidenbildsatz mit einer ersten Auflösung erzeugen (S210). Das Objektdetektionssystem **100** kann durch Downsampling mindestens einiger der mehreren im ersten Pyramidenbildsatz enthaltenen Pyramidenbilder einen zweiten Pyramidenbildsatz erzeugen (S220).

**[0066]** Das Objektdetektionssystem **100** kann N Elemente erster Objektdaten (wobei N eine natürliche Zahl ist), die N Zeitpunkten entsprechen, durch Extraktion eines ersten Objekts aus dem ersten Pyramidenbildsatz erzeugen (S230). Das Objektdetektionssystem **100** kann M Elemente zweiter Objektdaten (wobei M eine von N verschiedene natürliche Zahl ist), die M Zeitpunkten entsprechen, durch Extraktion eines zweiten Objekts aus dem zweiten Pyramidenbildsatz erzeugen (S240). Das Objektdetektionssystem **100** kann die N Elemente der ersten Objektdaten in einem ersten Bereich des Pufferspeichers **130** speichern (S250) und kann die M Elemente der zweiten Objektdaten in einem zweiten Bereich des Pufferspeichers **130** speichern (S260). In einer Ausführungsform kann die Anzahl N der ersten Objektdatenelemente größer als die Anzahl M der zweiten Objektdatenelemente sein. Gemäß einer Ausführungsform der Offenbarung kann das Objektdetektionssystem **100** Objektdaten durch Extraktion eines Objekts mit Hilfe einer für jede Auflösung unterschiedlichen Anzahl von Pyramidenbildern erzeugen. In einem Beispiel kann das Objektdetektionssystem **100** Objektdaten erzeugen, indem es für den Fall, dass das erste Objekt unzureichende räumliche Informationen aufweist, mehr Pyramidenbilder verwendet als für den Fall, dass das zweite Objekt relativ viele räumliche Informationen aufweist. Das heißt mit anderen Worten, dass ein Objekt an einer Position, die entfernt zu einer Position liegt, an der das Bild aufgenommen wird, im Bild klein erscheinen kann. Deshalb kann das Objekt durch eine entsprechend geringe Menge an Informationen und/oder Pixeln dargestellt werden. Bei einem Objekt mit unzureichenden räumlichen Informationen im Bild kann somit das Objekt mit Hilfe einer erhöhten Menge von Pyramidenbildern extrahiert werden. Im Ergebnis können zusätzliche räumliche Informationen und Pixelinformationen über das Objekt aus den zusätzlichen Pyramidenbildern bezogen werden, wodurch sich die Leistung der Objektextraktion erhöht.

**[0067]** **Fig. 9** ist ein Blockschaltbild zur Darstellung eines Objektdetektionssystems gemäß einer Ausführungsform der Offenbarung. **Fig. 9** zeigt insbesondere ein Objektdetektionssystem, das ein Objekt mit Hilfe einer für jede Auflösung unterschiedlichen Anzahl von Pyramidenbildern detektiert. Auf Beschreibungen, die sich in Bezug auf **Fig. 7** doppeln, wird verzichtet.

**[0068]** Wie in **Fig. 9** dargestellt, kann das Objektdetektionssystem **100** den Pyramidenbildgenerator **110**, den ersten Merkmalsextraktor **121**, den zweiten Merkmalsextraktor **122**, einen dritten Merkmalsextraktor **123**, den Pufferspeicher **130** und den Objekttracker **140** aufweisen. Der Pyramidenbildgenerator **110** kann das zum ersten Zeitpunkt t1 aufgenommene erste Eingangsbild IM1, das zum zweiten Zeitpunkt t2 aufgenommene zweite Eingangsbild IM2 und ein zu einem dritten Zeitpunkt t3 aufgenommenes drittes Eingangsbild IM3 empfangen.

**[0069]** Der Pyramidenbildgenerator **110** kann das erste Pyramidenbild PI1\_1 erzeugen, indem er dem ersten Eingangsbild IM1 eine dem ersten Zeitpunkt t1 entsprechende Zeitinformation hinzufügt, kann durch Downsampling des ersten Pyramidenbildes PI1\_1 das zweite Pyramidenbild PI1\_2 erzeugen und kann durch Downsampling des zweiten Pyramidenbildes PI1\_2 das dritte Pyramidenbild PI1\_3 erzeugen. Der Pyramidenbildgenerator **110** kann das erste bis dritte Pyramidenbild PI1\_1 bis PI1\_3 als einen ersten Pyramidenbildsatz PS1 an den ersten Merkmalsextraktor **121** ausgeben.

**[0070]** Der Pyramidenbildgenerator **110** kann das vierte Pyramidenbild PI2\_1 erzeugen, indem er dem zweiten Eingangsbild IM2 eine dem zweiten Zeitpunkt t2 entsprechende Zeitinformation hinzufügt, und kann durch Downsampling des vierten Pyramidenbildes PI2\_1 das fünfte Pyramidenbild PI2\_2 erzeugen. Der Pyramidenbildgenerator **110** kann das vierte und fünfte Pyramidenbild PI2\_1 und PI2\_2 als einen zweiten Pyramidenbildsatz PS2 an den zweiten Merkmalsextraktor **122** ausgeben. Der Pyramidenbildgenerator **110** kann ein sechstes Pyramidenbild PI3\_1 erzeugen, indem er dem dritten Eingangsbild IM3 eine dem dritten Zeitpunkt t3 entsprechende Zeitinformation hinzufügt, und kann das sechste Pyramidenbild PI3\_1 als einen dritten Pyramidenbildsatz PS3 an den dritten Merkmalsextraktor **123** ausgeben.

**[0071]** Der erste Merkmalsextraktor **121** kann die dem ersten Zeitpunkt t1 entsprechenden ersten bis dritten Pyramidenbilder PI1\_1 bis PI1\_3 empfangen und kann, durch entsprechende Extraktion von Objekten aus den empfangenen ersten bis dritten Pyramidenbildern PI1\_1 bis PI1\_3, die ersten bis dritten Objektdaten OD1\_1, OD1\_2 und OD1\_3 erzeugen. Im Beispiel von **Fig. 9** kann der erste Merkmalsextraktor **121** die ersten Objektdaten OD1\_1 durch Extraktion eines ersten Objekts aus dem ersten Pyramidenbild PI1\_1 erzeugen, kann die zweiten Objektdaten OD1\_2 durch Extraktion eines zweiten Objekts aus dem zweiten Pyramidenbild PI1\_2 erzeugen und kann die dritten Objektdaten OD1\_3 durch Extraktion eines dritten Objekts aus dem dritten Pyramidenbild PI1\_3 erzeugen. Der erste Merkmalsextraktor **121** kann die erzeugten ersten Objektdaten

OD1\_1 im ersten Bereich Ar1 des Pufferspeichers **130** speichern, kann die erzeugten zweiten Objektdaten OD1\_2 im zweiten Bereich Ar2 des Pufferspeichers **130** speichern und kann die erzeugten dritten Objektdaten OD1\_3 im dritten Bereich Ar3 des Pufferspeichers **130** speichern.

**[0072]** Der zweite Merkmalsextraktor **122** kann das dem zweiten Zeitpunkt t2 entsprechende vierte und fünfte Pyramidenbild PI2\_1 und PI2 2 empfangen und kann, durch Extraktion von entsprechenden Objekten aus dem empfangenen vierten und fünften Pyramidenbild PI2\_1 und PI2 2, die vierten und fünften Objektdaten OD2\_1 und OD2 2 erzeugen. Der zweite Merkmalsextraktor **122** kann die erzeugten vierten Objektdaten OD2\_1 im ersten Bereich Ar1 des Pufferspeichers **130** speichern und kann die erzeugten fünften Objektdaten OD2\_2 im zweiten Bereich Ar2 des Pufferspeichers **130** speichern.

**[0073]** Der dritte Merkmalsextraktor **123** kann das dem dritten Zeitpunkt t3 entsprechende sechste Pyramidenbild PI3\_1 empfangen und kann, durch Extraktion des dritten Objekts aus dem empfangenen sechsten Pyramidenbild PI3\_1, die sechsten Objektdaten OD3\_1 erzeugen. Der dritte Merkmalsextraktor **123** kann die erzeugten sechsten Objektdaten OD3\_1 im ersten Bereich Ar1 des Pufferspeichers **130** speichern.

**[0074]** Der Objekttracker **140** kann Objekte mit Hilfe der auf Objektbasis gespeicherten ersten bis sechsten Objektdaten OD1\_1 bis OD3\_1 verfolgen. In einem Beispiel kann der Objekttracker **140** das erste Objekt mit Hilfe der im ersten Bereich Ar1 des Pufferspeichers **130** gespeicherten ersten Objektdaten OD1\_1, vierten Objektdaten OD2\_1 und sechsten Objektdaten OD3\_1 verfolgen.

**[0075]** Gemäß einer Ausführungsform der Offenbarung kann das Objektdetektionssystem **100** ein Objekt detektieren, indem es für jedes Objekt eine unterschiedliche Anzahl von Pyramidenbildern verwendet. In einem Beispiel kann das Objektdetektionssystem **100** das dritte Objekt mit Hilfe von drei Pyramidenbildern (zum Beispiel PI1\_1, PI2\_1 und PI3\_1) detektieren, kann das zweite Objekt mit Hilfe von zwei Pyramidenbildern (zum Beispiel PI1\_2 und PI2 2) detektieren und kann das erste Objekt mit Hilfe eines Pyramidenbildes (zum Beispiel PI1\_3) detektieren. In einer Ausführungsform kann das Objektdetektionssystem **100** ein Objekt mit weniger Pyramidenbildern detektieren, wenn das Objekt von einer Bildaufnahme-position, an der das Bild aufgenommen wird, weiter entfernt ist.

**[0076]** Fig. 10 ist eine Ansicht zur Darstellung von Objektdaten gemäß einer Ausführungsform der Offenbarung. Fig. 10 zeigt insbesondere eine Ausführungsform, bei der ein Objektdetektionssystem für je-

des Objekt eine unterschiedliche Anzahl Objektdatenelementen erzeugt.

**[0077]** Wie in Fig. 9 und Fig. 10 dargestellt, kann das Objektdetektionssystem **100** die dem ersten Objekt **O1** entsprechenden dritten Objektdaten OD1\_3 im dritten Bereich Ar3 des Pufferspeichers **130** speichern, kann die dem zweiten Objekt **O2** entsprechenden zweiten und fünften Objektdaten OD1\_2 und OD2 2 im zweiten Bereich Ar2 des Pufferspeichers **130** speichern und kann die dem dritten Objekt **O3** entsprechenden ersten, vierten und sechsten Objektdaten OD1\_1, OD2\_1 und OD3\_1 im ersten Bereich Ar1 des Pufferspeichers **130** speichern.

**[0078]** Das erste Objekt **O1** kann ein Objekt in relativer Nähe zu einer Bildaufnahmevorrichtung sein, und über das erste Objekt **O1** kann eine relativ große Menge räumlicher Informationen vorliegen. Mit anderen Worten kann ein Objekt an einer Position in der Nähe zu einer Position, an der das Bild aufgenommen wird, im Bild groß erscheinen. Deshalb kann das Objekt durch eine entsprechend große Menge an Informationen und/oder Pixeln dargestellt werden. Dementsprechend kann das Objektdetektionssystem **100** das erste Objekt **O1** unter alleiniger Zuhilfenahme der dritten Objektdaten OD1\_3, die einem ersten Zeitpunkt t1 entsprechen, detektieren. Das dritte Objekt **O3** hingegen kann ein Objekt in relativer Entfernung zur Bildaufnahmevorrichtung sein, und über das dritte Objekt **O3** kann eine relativ kleine Menge räumlicher Informationen vorliegen. Mit Hilfe der ersten, vierten und sechsten Objektdaten OD1\_1, OD2\_1 und OD3\_1, die mehreren Zeitpunkten (zum Beispiel dem ersten bis dritten Zeitpunkt t1 bis t3) entsprechen, kann somit das Objektdetektionssystem **100** eine relativ kleine Menge räumlicher Informationen mit Objektdatenelementen ergänzen, die mehreren Zeitpunkten entsprechen, und so kann eine effiziente Objektdetektion erfolgen.

**[0079]** Fig. 11 ist ein Blockschaltbild zur Darstellung eines Objektdetektionssystems gemäß einer Ausführungsform der Offenbarung. Fig. 11 zeigt insbesondere eine Ausführungsform, bei welcher der Objekttracker **140** die für die Objektverfolgung erforderliche Anzahl von Objektdatenelementen selektiv bestimmt. Auf Beschreibungen, die sich in Bezug auf Fig. 7 doppeln, wird verzichtet.

**[0080]** Wie in Fig. 11 dargestellt, kann das Objektdetektionssystem **100** den Pyramidenbildgenerator **110**, den ersten Merkmalsextraktor **121**, den zweiten Merkmalsextraktor **122**, den dritten Merkmalsextraktor **123**, den Pufferspeicher **130** und den Objekttracker **140** aufweisen. Der Pyramidenbildgenerator **110** kann das zum ersten Zeitpunkt t1 aufgenommene erste Eingangsbild IM1, das zum zweiten Zeitpunkt t2 aufgenommene zweite Eingangsbild IM2 und das

zum dritten Zeitpunkt t3 aufgenommene dritte Eingangsbild IM3 empfangen.

**[0081]** Der Pyramidenbildgenerator **110** kann das erste bis dritte Pyramidenbild PI1\_1, PI1\_2 und PI1\_3 als ersten Pyramidenbildsatz PS1 an den ersten Merkmalsextraktor **121** ausgeben, wobei das erste bis dritte Pyramidenbild PI1\_1 bis PI1\_3 durch das oben beschriebene Verfahren erzeugt werden. Der Pyramidenbildgenerator **110** kann gleichermaßen das vierte bis sechste Pyramidenbild PI2\_1, PI2\_2 und PI2\_3 als zweiten Pyramidenbildsatz PS2 an den zweiten Merkmalsextraktor **122** ausgeben und kann ein siebtes bis neuntes Pyramidenbild PI3\_1, PI3\_2 und PI3\_3 als einen dritten Pyramidenbildsatz PS3 an den dritten Merkmalsextraktor **123** ausgeben.

**[0082]** Der erste Merkmalsextraktor **121** kann durch entsprechende Extraktion von Objekten aus dem ersten bis dritten Pyramidenbild PI1\_1 bis PI1\_3, die dem ersten Zeitpunkt t1 entsprechen, die ersten bis dritten Objektdaten OD1\_1 bis OD1\_3 erzeugen. Der erste Merkmalsextraktor **121** kann die erzeugten ersten Objektdaten OD1\_1 im ersten Bereich Ar1 des Pufferspeichers **130** speichern, kann die erzeugten zweiten Objektdaten OD1\_2 im zweiten Bereich Ar2 des Pufferspeichers **130** speichern und kann die erzeugten dritten Objektdaten OD1\_3 im dritten Bereich Ar3 des Pufferspeichers **130** speichern.

**[0083]** Der zweite Merkmalsextraktor **122** kann gleichermaßen durch entsprechende Extraktion von Objekten aus dem vierten bis sechsten Pyramidenbild PI2\_1 bis PI2\_3, die dem zweiten Zeitpunkt t2 entsprechen, die vierten bis sechsten Objektdaten OD2\_1 bis OD2\_3 erzeugen. Der zweite Merkmalsextraktor **122** kann die vierten Objektdaten OD2\_1 im ersten Bereich Ar1 des Pufferspeichers **130** speichern, kann die fünften Objektdaten OD2\_2 im zweiten Bereich Ar2 des Pufferspeichers **130** speichern und kann die sechsten Objektdaten OD2\_3 im dritten Bereich Ar3 des Pufferspeichers **130** speichern.

**[0084]** Der dritte Merkmalsextraktor **123** kann durch entsprechende Extraktion von Objekten aus dem siebten bis neunten Pyramidenbild PI3\_1 bis PI3\_3, die dem dritten Zeitpunkt t3 entsprechen, die siebten bis neunten Objektdaten OD3\_1 bis OD3\_3 erzeugen. Der dritte Merkmalsextraktor **123** kann die siebten Objektdaten OD3\_1 im ersten Bereich Ar1 des Pufferspeichers **130** speichern, kann die achten Objektdaten OD3\_2 im zweiten Bereich Ar2 des Pufferspeichers **130** speichern und kann die neunten Objektdaten OD3\_3 im dritten Bereich Ar3 des Pufferspeichers **130** speichern.

**[0085]** Der Objekttracker **140** kann ein Objekt durch Lesen von mindestens einigen der auf Objektbasis gespeicherten Objektdatenelemente (zum Beispiel OD1\_1 bis OD3\_3) verfolgen. Gemäß einer Ausführungsform der Offenbarung kann der Objekttracker **140** ein Objekt verfolgen, indem er nur einige der auf Objektbasis gespeicherten Objektdatenelemente (zum Beispiel OD1\_1 bis OD3\_3) liest. In einem Beispiel kann der Objekttracker **140** ein erstes Objekt verfolgen, indem er nur die ersten Objektdaten OD1\_1 und die vierten Objektdaten OD2\_1, die einige der ersten Objektdaten OD1\_1 sind, die vierten Objektdaten OD2\_1 und die siebten Objektdaten OD3\_1, die dem ersten Objekt entsprechen, verwendet.

**[0086]** Fig. 12 ist ein Blockschaltbild zur Darstellung eines Objektdetektionssystems gemäß einer Ausführungsform der Offenbarung. Fig. 12 zeigt insbesondere ein Objektdetektionssystem, das ein Objekt anhand eines interessierenden Bereichs (Region of Interest, ROI) detektiert. Auf Beschreibungen, die sich in Bezug auf Fig. 3 doppelten, wird verzichtet.

**[0087]** Wie in Fig. 12 dargestellt, kann eine Objektdetektionsvorrichtung **100a** den Pyramidenbildgenerator **110**, den Merkmalsextraktor **120**, den Pufferspeicher **130**, den Objekttracker **140** und einen ROI-Manager **150** aufweisen. Der ROI-Manager **150** kann einen in einem Eingangsbild IMs enthaltenen Teilbereich ausgehend vom Eingangsbild IMs als ROI identifizieren und kann ROI-Informationen RI, welche die den ROI anzeigenden Daten enthalten, ausgeben. Wenn die Objektdetektionsvorrichtung **100a** in einem Fahrassistenzsystem enthalten ist, kann der ROI-Manager **150** durch Analyse des Eingangsbildes IMs als ROI eine Zone identifizieren, die Informationen enthält, die für das Fahren des Fahrzeugs erforderlich sind. Der ROI kann zum Beispiel ein Bereich sein, der eine Verkehrsstraße, ein anderes Fahrzeug, eine Verkehrsampel, einen Fußgängerübergang oder dergleichen umfasst.

**[0088]** Der ROI-Manager **150** kann einen Tiefengenerator **151** aufweisen. Der Tiefengenerator **151** kann eine Tiefenkarte erzeugen, die Tiefendaten bezüglich eines im Eingangsbild IMs enthaltenen Objekts und Hintergrunds aufweist. In einem Beispiel kann das Eingangsbild IMs ein linksäugiges Bild und ein rechtsäugiges Bild aufweisen, und der Tiefengenerator **151** kann mit Hilfe des linksäugigen Bildes und des rechtsäugigen Bildes eine Tiefe berechnen und kann eine auf der berechneten Tiefe basierende Tiefenkarte erzeugen. In einem anderen Beispiel kann der Tiefengenerator **151** die Tiefenkarte bezüglich des im Eingangsbild IMs enthaltenen Objekts und Hintergrunds mit Hilfe der von einem Entfernungssensor generierten dreidimensionalen Informationen erzeugen.

**[0089]** Der ROI-Manager **150** kann die ROI-Informationen RI unter Verwendung der vom Tiefengenerator **151** generierten Tiefenkarte erzeugen. In einem Beispiel kann der ROI-Manager **150** auf Grundlage der

Tiefenkarte einen Bereich innerhalb einer bestimmten Entfernung als ROI einstellen.

**[0090]** Der ROI-Manager **150** kann die erzeugten ROI-Informationen RI an den Pyramidenbildgenerator **110** ausgeben, und der Pyramidenbildgenerator **110** kann anhand der ROI-Informationen RI Pyramidenbilder PIs erzeugen. In einer Ausführungsform kann der Pyramidenbildgenerator **110** anhand der ROI-Informationen RI einen Teilabschnitt des Eingangsbildes IMs, der kein ROI ist, maskieren und kann die Pyramidenbilder PIs durch alleinige Verwendung eines unmaskierten Bildabschnitts erzeugen. Mit anderen Worten kann der Pyramidenbildgenerator **110** Zonen des Eingangsbildes außerhalb des von den ROI-Informationen angezeigten interessierenden Bereichs ROI unberücksichtigt lassen. Dies ermöglicht eine Verbesserung der Leistung des Eingangsbildes IMs durch den Pyramidenbildgenerator **110**.

**[0091]** Fig. **13** ist eine Ansicht zur Darstellung eines Verfahrens zum Erzeugen eines Pyramidenbildes gemäß einer Ausführungsform der Offenbarung.

**[0092]** Wie in Fig. **12** und Fig. **13** dargestellt, kann der Pyramidenbildgenerator **110** das dem ersten Zeitpunkt t1 entsprechende erste Eingangsbild IM1 empfangen und kann eine dem ersten Zeitpunkt t1 entsprechende Zeitinformation zu dem ersten Eingangsbild IM **1** hinzufügen. Darüber hinaus kann der Pyramidenbildgenerator **110** das erste Pyramidenbild PI1\_1 dadurch erzeugen, dass er einen außerhalb des ROI liegenden Bereich anhand der ROI-Informationen RI maskiert. Der ROI kann das erste bis dritte Objekt **O1** bis **O3** enthalten.

**[0093]** Der Pyramidenbildgenerator **110** kann durch Downsampling des maskierten ersten Pyramidenbildes PI1\_1 das zweite Pyramidenbild PI1\_2 erzeugen und kann durch Downsampling des zweiten Pyramidenbildes PI1\_2 das dritte Pyramidenbild PI1\_3 erzeugen. Das Objektdetektionssystem **100** kann das dritte Objekt **O3** mit Hilfe des maskierten ersten Pyramidenbildes PI1\_1 detektieren, kann das zweite Objekt **O2** mit Hilfe des maskierten zweiten Pyramidenbildes PI1\_2 detektieren und kann das dritte Objekt **O3** mit Hilfe des maskierten dritten Pyramidenbildes PI1\_3 detektieren. Gemäß einer Ausführungsform der Offenbarung kann durch die Detektion eines Objekts nach Maskierung eines Eingangsbildes ein außerhalb des ROI liegender maskierter Bereich unberücksichtigt bleiben und die Detektionsleistung verbessert werden.

**[0094]** Fig. **14** ist ein Blockschaltbild zur Darstellung eines Objektdetektionssystems gemäß einer Ausführungsform der Offenbarung. Fig. **14** zeigt insbesondere ein Objektdetektionssystem, das mit Hilfe von Objektdaten einen Hintergrund detektiert. Auf Be-

schreibungen, die sich in Bezug auf Fig. **3** doppelten, wird verzichtet.

**[0095]** Wie in Fig. **14** dargestellt, kann ein Objektdetektionssystem **100b** den Pyramidenbildgenerator **110**, den Merkmalsextraktor **120**, den Pufferspeicher **130**, den Objekttracker **140** und einen Hintergrundextraktor **160** aufweisen. Der Hintergrundextraktor **160** kann mehrere auf Objektbasis gespeicherte Objektdatenelemente ODs aus einem Pufferspeicher empfangen und kann anhand der mehreren Objektdatenelemente ODs einen Hintergrund des Eingangsbildes IMs extrahieren. In einem Beispiel kann der Hintergrundextraktor **160** den Hintergrund extrahieren, indem er ausgehend von den Objektdatenelementen ODs mindestens ein Objekt aus dem Eingangsbild IMs entfernt. Gemäß einer Ausführungsform der Offenbarung kann der Hintergrundextraktor **160** die Herausnahme von Objekten aus dem Hintergrund anhand der Objektdatenelemente ODs, die mehreren Zeitpunkten entsprechen, vollziehen.

**[0096]** Fig. **15** ist ein Blockschaltbild zur Darstellung eines Anwendungsprozessors gemäß einer Ausführungsform der Offenbarung. Ein in Fig. **15** dargestellter Anwendungsprozessor kann ein Halbleiterchip sein und kann durch ein System-on-Chip (SoC) realisiert sein.

**[0097]** Der Anwendungsprozessor **1000** kann einen Prozessor **1010** und einen Betriebssystem **1020** aufweisen. Der Anwendungsprozessor **1000** kann ferner ein oder mehrere mit einem Systembus verbundene IP-Module umfassen. Der Betriebssystem **1020** kann Software wie verschiedene Programme und Anweisungen speichern, die sich auf Betriebsvorgänge eines Systems beziehen, in dem der Anwendungsprozessor **1000** verwendet wird. Der Betriebssystem **1020** kann beispielsweise ein Betriebssystem **1021**, ein neuronales Netzmodul **1022** und ein Objektdetektionsmodul **1023** umfassen. Der Prozessor **1010** kann das in den Betriebssystem **1020** geladene Objektdetektionsmodul **1023** ausführen und kann gemäß der oben beschriebenen Ausführungsform eine Funktion zur Detektion eines Objekts aus einem Eingangsbild anhand von Zeitinformationen ausführen.

**[0098]** Ein oder mehrere Hardware-Teile können den Prozessor **1010** umfassen und können Operationen eines neuronalen Netzes durch Ausführung des neuronalen Netzmoduls **1022** vollziehen, und das eine oder die mehreren Hardware-Teile können Objektdaten aus einem Pyramidenbild gemäß den oben beschriebenen Ausführungsformen erzeugen.

**[0099]** Fig. **16** ist ein Blockschaltbild zur Darstellung eines Fahrassistenzsystems gemäß einer Ausführungsform der Offenbarung.

**[0100]** Wie in **Fig. 16** dargestellt, kann ein Fahrassistenzsystem **2000** einen Prozessor **2010**, eine Sensoreinheit **2040**, ein Kommunikationsmodul **2050**, eine Fahrsteuerungseinheit **2060**, eine autonome Fahreinheit **2070** und eine Benutzerschnittstelle **2080** umfassen. Der Prozessor **2010** kann den Gesamtbetrieb des Fahrassistenzsystems **2000** steuern und kann gemäß den oben beschriebenen Ausführungsformen aus einem Eingangsbild, das mit Bezug auf zeitliche Informationen aus der Sensoreinheit **2040** empfangen wurde, ein Objekt detektieren.

**[0101]** Die Sensoreinheit **2040** kann Informationen über ein vom Fahrassistenzsystem **2000** erfasstes Objekt sammeln. Die Sensoreinheit **2040** kann in einem Beispiel eine Bildsensoreinheit sein und mindestens einen Bildsensor umfassen. Die Sensoreinheit **2040** kann ein Bildsignal aus dem Außenbereich des Fahrassistenzsystems **2000** erfassen oder empfangen und kann das Bildsignal in Bilddaten, das heißt zu einem Einzelbild (Bildframe), umwandeln.

**[0102]** In einem anderen Beispiel kann die Sensoreinheit **2040** eine Entfernungsmesseinheit sein und kann mindestens einen Entfernungssensor umfassen. Der Entfernungssensor kann beispielsweise mindestens eine von verschiedenen Erfassungsvorrichtungen, wie einen LIDAR-(Light Detection and Ranging)-Sensor, einen RADAR-(Radio Detection and Ranging)-Sensor, einen Lichtlaufzeit-(ToF)-Sensor, einen Ultraschallsensor, einen Infrarotsensor und dergleichen, umfassen. Der LIDAR-Sensor und der RADAR-Sensor können je nach effektiver Messentfernung in Klassen eingeteilt werden. Zum Beispiel kann der LIDAR-Sensor in einen LIDAR-Weitbereichssensor und einen LIDAR-Nahbereichssensor eingeteilt werden, und der RADAR-Sensor kann in einen RADAR-Weitbereichssensor und einen RADAR-Nahbereichssensor eingeteilt werden. Die Offenbarung ist hierauf nicht beschränkt, und die Sensoreinheit **2040** kann einen Geomagnetsensor, einen Positionssensor (zum Beispiel einen Globalpositionbestimmungssensor (GPS)), einen Beschleunigungssensor, einen Luftdrucksensor, einen Temperatur-/Feuchtigkeitssensor, einen Näherungssensor und/oder ein Gyroskop umfassen, ohne darauf beschränkt zu sein.

**[0103]** Das Kommunikationsmodul **2050** kann Daten zum Fahrassistenzsystem **2000** senden und von diesem empfangen. In einem Beispiel kann das Kommunikationsmodul **2050** eine Kommunikation der Art Fahrzeug zu Allem (V2X) ausführen. Als ein Beispiel kann das Kommunikationsmodul **2050** eine Kommunikation der Art Fahrzeug zu Fahrzeug (V2V), Fahrzeug zu Infrastruktur (V2I), Fahrzeug zu Fußgänger (V2P) und Fahrzeug zu Nomadengerät (V2N) ausführen. Die Offenbarung ist jedoch hierauf nicht beschränkt, und das Kommunikationsmodul **2050** kann Daten in verschiedenen öffentlich bekannten Kom-

munikationsarten senden und empfangen. Das Kommunikationsmodul **2050** kann beispielsweise eine Kommunikation über ein Kommunikationsverfahren wie 3G, Long Term Evolution (LTE), Wi-Fi, Bluetooth, Bluetooth Low Energy (BLE), Zigbee, Nahfeldkommunikation (NFC) oder Ultraschallkommunikation durchführen und kann sowohl eine Kurzstrecken- als auch Langstreckenkommunikation umfassen.

**[0104]** Der Sensor **2040** kann ein Eingangsbild erzeugen, indem er ein Bild des Umfeldes oder Umgebungsbereichs des Fahrassistenzsystems **2000** aufnimmt, und kann das Eingangsbild an den Prozessor **2010** senden. Der Prozessor **2010** kann auf Grundlage des Eingangsbildes und des Zeitpunkts, an dem das Eingangsbild aufgenommen wird, ein Objekt (zum Beispiel ein anderes Fahrzeug) detektieren und kann die Fahrsteuerungseinheit **2060** und die autonome Fahreinheit **2070** steuern. Obwohl ein Beispiel dargestellt ist, bei dem der Prozessor **2010** ein Objekt anhand des Eingangsbildes detektiert, kann der Prozessor **2010** in einem anderen Beispiel ein Objekt anhand von Tiefeninformationen detektieren, die von einem Entfernungssensor ausgegeben werden.

**[0105]** Die Fahrsteuerungseinheit **2060** kann umfassen: eine Fahrzeuglenkvorrichtung, die dazu eingerichtet ist, eine Richtung eines Fahrzeugs zu steuern; eine Drosselvorrichtung, die dazu eingerichtet ist, durch Ansteuerung eines Verbrennungs- oder anderen Motors des Fahrzeugs eine Beschleunigung und/oder Verzögerung zu steuern; ein Bremssystem, das dazu eingerichtet ist, die Bremsung des Fahrzeugs zu steuern; eine Außenbeleuchtungseinrichtung; und dergleichen. Die autonome Fahreinheit **2070** kann eine Rechenvorrichtung umfassen, die dazu eingerichtet ist, die autonome Steuerung der Fahrsteuerungseinheit **2060** zu realisieren. Die autonome Fahreinheit **2070** kann beispielsweise mindestens eine der Komponenten des Fahrassistenzsystems **2000** umfassen. Die autonome Fahreinheit **2070** kann Speicher, der mehrere Programmanweisungen abspeichert, und einen oder mehrere Prozessoren, die die Programmanweisungen ausführen, umfassen. Die autonome Fahreinheit **2070** kann dazu eingerichtet sein, die Fahrsteuerungseinheit **2060** auf Grundlage eines von der Sensoreinheit **2040** ausgegebenen Erfassungssignals zu steuern. Die Benutzerschnittstelle **2080** kann verschiedene elektronische Vorrichtungen und mechanische Ausrüstungen wie ein in einem Fahrersitz, einem Mitfahrersitz und dergleichen enthaltenes Display, das ein Armaturenbrett des Fahrzeugs anzeigt, aufweisen.

**[0106]** Der Prozessor **2010** kann bei der Detektierung eines Objekts verschiedene Erfassungsdatenelemente wie das Eingangsbild, die Tiefeninformationen und so weiter verwenden. In diesem Fall kann der Prozessor **2010** ein künstliches neuronales Netz zur

effizienten Verarbeitung von Betriebsvorgängen verwenden und kann eines der in der Offenbarung beschriebenen Objektdetektionsverfahren ausführen.

**[0107]** Während die Offenbarung unter Bezugnahme auf entsprechende Ausführungsformen besonders dargestellt und beschrieben wurde, versteht es sich, dass hier verschiedene Änderungen an Form und Einzelheiten vorgenommen werden können, ohne vom Wesen und Umfang der nachfolgenden Ansprüche abzuweichen.



**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**

- KR 1020190104574 [0001]

## Patentansprüche

1. Objektdetektionssystem umfassend:  
 einen Pyramidenbildgenerator, der eingerichtet ist zum Empfangen eines zu einem ersten Zeitpunkt aufgenommenen ersten Eingangsbildes und eines zu einem zweiten Zeitpunkt aufgenommenen zweiten Eingangsbildes und zum Erzeugen eines ersten Pyramidenbildes aus dem ersten Eingangsbild und eines zweiten Pyramidenbildes aus dem zweiten Eingangsbild;  
 einen Objektextraktor, der eingerichtet ist zum Detektieren eines Objekts im ersten Pyramidenbild und im zweiten Pyramidenbild und zum Erzeugen mehrerer Objektdatenelemente, die das Objekt darstellen; und  
 einen Pufferspeicher, der die mehreren Objektdatenelemente speichert, die das im ersten Eingangsbild und im zweiten Eingangsbild detektierte Objekt darstellen.

2. Objektdetektionssystem nach Anspruch 1, wobei der Pyramidenbildgenerator ferner eingerichtet ist zum:  
 Erzeugen des ersten Pyramidenbildes mit einer ersten Auflösung durch Hinzufügen erster Zeitinformationen, die dem ersten Zeitpunkt entsprechen, zum ersten Eingangsbild;  
 Erzeugen eines dritten Pyramidenbildes mit einer zweiten Auflösung durch Downsampling des ersten Pyramidenbildes;  
 Erzeugen des zweiten Pyramidenbildes mit der ersten Auflösung durch Hinzufügen zweiter Zeitinformationen, die dem zweiten Zeitpunkt entsprechen, zum zweiten Eingangsbild; und  
 Erzeugen eines vierten Pyramidenbildes mit der zweiten Auflösung durch Downsampling des dritten Pyramidenbildes.

3. Objektdetektionssystem nach Anspruch 2, wobei das Objekt ein erstes Objekt und ein zweites Objekt umfasst, und wobei der Objektextraktor ferner eingerichtet ist zum:  
 Erzeugen erster Objektdaten des ersten Objekts von den mehreren Objektdatenelementen durch Extrahieren des ersten Objekts aus dem ersten Pyramidenbild unter Verwendung eines auf Grundlage eines neuronalen Netzes trainierten Modells des tiefgehenden Lernens (Deep-Learning);  
 Erzeugen zweiter Objektdaten des zweiten Objekts von den mehreren Objektdatenelementen durch Extrahieren des zweiten Objekts aus dem zweiten Pyramidenbild unter Verwendung des Modells des tiefgehenden Lernens;  
 Erzeugen dritter Objektdaten des ersten Objekts von den mehreren Objektdatenelementen durch Extrahieren des ersten Objekts aus dem dritten Pyramidenbild unter Verwendung des Modells des tiefgehenden Lernens; und  
 Erzeugen vierter Objektdaten des zweiten Objekts von den mehreren Objektdatenelementen durch Ex-

trahieren des zweiten Objekts aus dem vierten Pyramidenbild unter Verwendung des Modells des tiefgehenden Lernens.

4. Objektdetektionssystem nach Anspruch 3, wobei der Objektextraktor ferner eingerichtet ist zum:  
 Speichern der ersten Objektdaten des ersten Objekts und der dritten Objektdaten des ersten Objekts in einem ersten Bereich des Pufferspeichers; und  
 Speichern der zweiten Objektdaten des zweiten Objekts und der vierten Objektdaten des zweiten Objekts in einem zweiten Bereich des Pufferspeichers.

5. Objektdetektionssystem nach Anspruch 4, ferner umfassend:  
 einen Objekttracker, der eingerichtet ist zum Verfolgen des ersten Objekts unter Verwendung mindestens eines der Objektdaten von den mehreren Objektdatenelementen, die ausgewählt sind aus den im ersten Bereich gespeicherten ersten Objektdaten des ersten Objekts und dritten Objektdaten des ersten Objekts, und der ersten Zeitinformationen und Verfolgen des zweiten Objekts unter Verwendung mindestens eines der Objektdaten von den mehreren Objektdatenelementen, die ausgewählt sind aus den im zweiten Bereich gespeicherten zweiten Objektdaten des zweiten Objekts und vierten Objektdaten des zweiten Objekts, und der zweiten Zeitinformationen.

6. Objektdetektionssystem nach Anspruch 4, ferner umfassend:  
 einen Hintergrundextraktor, der eingerichtet ist zum Detektieren eines ersten Hintergrunds des ersten Eingangsbildes außerhalb eines das erste Objekt umfassenden Bereichs auf Grundlage der ersten Objektdaten und/oder der dritten Objektdaten und zum Detektieren eines zweiten Hintergrunds des zweiten Eingangsbildes außerhalb eines das zweite Objekt umfassenden Bereichs auf Grundlage der zweiten Objektdaten und/oder der vierten Objektdaten.

7. Objektdetektionssystem nach Anspruch 3, wobei das erste Objekt relativ weit von einem das erste Eingangsbild und das zweite Eingangsbild aufnehmenden Bildaufnahmegerät angeordnet ist und das zweite Objekt relativ näher am Bildaufnahmegerät als das erste Objekt angeordnet ist.

8. Objektdetektionssystem nach Anspruch 1, wobei der Pyramidenbildgenerator ferner eingerichtet ist zum:  
 Erzeugen eines ersten Pyramidenbildsatzes, der das erste Pyramidenbild und das zweite Pyramidenbild umfasst; und  
 Erzeugen eines zweiten Pyramidenbildsatzes durch Downsampling des ersten Pyramidenbildes und/oder zweiten Pyramidenbildes im ersten Pyramidenbildsatz, und  
 wobei der Objektextraktor ferner eingerichtet ist zum:

Erzeugen von N Elementen erster Objektdaten, die N Zeitpunkten entsprechen, der mehreren Objektdatenelemente aus dem ersten Pyramidenbildsatz, wobei N eine natürliche Zahl ist; und

Erzeugen von M Elementen zweiter Objektdaten, die M Zeitpunkten entsprechen, der mehreren Objektdatenelemente aus dem zweiten Pyramidenbildsatz, wobei M eine natürliche Zahl ist.

9. Objektdetektionssystem nach Anspruch 8, wobei N größer als M ist.

10. Objektdetektionssystem nach Anspruch 8, wobei das Objekt ein erstes Objekt und ein zweites Objekt umfasst, und

wobei das Verfahren ferner umfasst:

einen Objekttracker, der eingerichtet ist zum Verfolgen des ersten Objekts durch Verwendung von P Elementen der ersten Objektdaten aus den N Elementen der ersten Objektdaten, wobei P eine natürliche Zahl kleiner als oder gleich N ist, und

Verfolgen des zweiten Objekts durch Verwendung von Q Elementen der zweiten Objektdaten aus den M Elementen der zweiten Objektdaten, wobei Q eine natürliche Zahl ist, die kleiner als oder gleich M ist.

11. Objektdetektionssystem nach Anspruch 1, ferner umfassend:

einen Region-of-Interest-Manager, der eingerichtet ist zum Einstellen eines ersten interessierenden Bereichs für das erste Eingangsbild und eines zweiten interessierenden Bereichs für das zweite Eingangsbild,

wobei der Objektextraktor ferner eingerichtet ist zum: Extrahieren des Objekts aus einem ersten Bereich des ersten Eingangsbildes und

einem zweiten Bereich des zweiten Eingangsbildes, wobei der erste Bereich und

der zweite Bereich dem ersten interessierenden Bereich und dem zweiten interessierenden Bereich entsprechen.

12. Verfahren zum Detektieren eines Objekts, wobei das Verfahren umfasst:

Empfangen eines zu einem ersten Zeitpunkt aufgenommenen ersten Eingangsbildes und eines zu einem zweiten Zeitpunkt aufgenommenen zweiten Eingangsbildes;

Erzeugen eines mit dem ersten Zeitpunkt verknüpften ersten Pyramidenbildes aus dem ersten Eingangsbild und eines mit dem zweiten Zeitpunkt verknüpften zweiten Pyramidenbildes aus dem zweiten Eingangsbild;

Erzeugen mehrerer Objektdatenelemente, die das im ersten Eingangsbild und im zweiten Eingangsbild detektierte Objekt darstellen, auf Grundlage des ersten Pyramidenbildes und des zweiten Pyramidenbildes; und

Speichern der mehreren Objektdatenelemente in einem Pufferspeicher.

13. Verfahren nach Anspruch 12, wobei das Erzeugen des ersten Pyramidenbildes und des zweiten Pyramidenbildes umfasst:

Erzeugen des ersten Pyramidenbildes mit einer ersten Auflösung durch Hinzufügen erster Zeitinformationen, die dem ersten Zeitpunkt entsprechen, zum ersten Eingangsbild;

Erzeugen eines dritten Pyramidenbildes mit einer zweiten Auflösung durch Downsampling des ersten Pyramidenbildes;

Erzeugen des zweiten Pyramidenbildes mit der ersten Auflösung durch Hinzufügen zweiter Zeitinformationen, die dem zweiten Zeitpunkt entsprechen, zum zweiten Eingangsbild; und

Erzeugen eines vierten Pyramidenbildes mit der zweiten Auflösung durch Downsampling des dritten Pyramidenbildes.

14. Verfahren nach Anspruch 13, wobei das Objekt ein erstes Objekt und ein zweites Objekt umfasst, und wobei das Erzeugen der mehreren Objektdatenelemente umfasst:

Erzeugen erster Objektdaten des ersten Objekts von den mehreren Objektdatenelementen durch Extrahieren des ersten Objekts aus dem ersten Pyramidenbild unter Verwendung eines auf Grundlage eines neuronalen Netzes trainierten Modells des tiefgehenden Lernens;

Erzeugen zweiter Objektdaten des zweiten Objekts von den mehreren Objektdatenelementen durch Extrahieren des zweiten Objekts aus dem zweiten Pyramidenbild unter Verwendung des Modells des tiefgehenden Lernens;

Erzeugen dritter Objektdaten des ersten Objekts von den mehreren Objektdatenelementen durch Extrahieren des ersten Objekts aus dem dritten Pyramidenbild unter Verwendung des Modells des tiefgehenden Lernens; und

Erzeugen vierter Objektdaten des zweiten Objekts von den mehreren Objektdatenelementen durch Extrahieren des zweiten Objekts aus dem vierten Pyramidenbild unter Verwendung des Modells des tiefgehenden Lernens.

15. Verfahren nach Anspruch 14, wobei das Speichern umfasst:

Speichern der ersten Objektdaten des ersten Objekts und der dritten Objektdaten des ersten Objekts in einem ersten Bereich des Pufferspeichers; und

Speichern der zweiten Objektdaten des zweiten Objekts und der vierten Objektdaten des zweiten Objekts in einem zweiten Bereich des Pufferspeichers.

16. Verfahren nach Anspruch 15, ferner umfassend:

Verfolgen des ersten Objekts unter Verwendung mindestens eines der Objektdaten von den mehreren Objektdatenelementen, die ausgewählt sind aus den im ersten Bereich gespeicherten ersten Objektdaten

des ersten Objekts und dritten Objektdaten des ersten Objekts, und der ersten Zeitinformationen; und Verfolgen des zweiten Objekts unter Verwendung mindestens eines der Objektdaten von den mehreren Objektdatenelementen, die ausgewählt sind aus den im zweiten Bereich gespeicherten zweiten Objektdaten des zweiten Objekts und vierten Objektdaten des zweiten Objekts.

17. Verfahren nach Anspruch 15, ferner umfassend:

Extrahieren eines ersten Hintergrunds des ersten Eingangsbildes außerhalb eines das erste Objekt umfassenden Bereichs auf Grundlage der ersten Objektdaten und/oder der dritten Objektdaten; und Extrahieren eines zweiten Hintergrunds des zweiten Eingangsbildes außerhalb eines das zweite Objekt umfassenden Bereichs auf Grundlage der zweiten Objektdaten und/oder der vierten Objektdaten.

18. Verfahren nach Anspruch 14, wobei das erste Objekt relativ weit von einem das erste Eingangsbild und das zweite Eingangsbild aufnehmenden Bildaufnahmegerät angeordnet ist, und das zweite Objekt relativ näher am Bildaufnahmegerät als das erste Objekt angeordnet ist.

19. Verfahren nach Anspruch 12, wobei das Erzeugen des ersten Pyramidenbildes und des zweiten Pyramidenbildes umfasst:

Erzeugen eines ersten Pyramidenbildsatzes, der das erste Pyramidenbild und das zweite Pyramidenbild umfasst; und

Erzeugen eines zweiten Pyramidenbildsatzes durch Downsampling des ersten Pyramidenbildes und/oder zweiten Pyramidenbildes im ersten Pyramidenbildsatz, und

wobei das Erzeugen der mehreren Objektdatenelemente umfasst:

Erzeugen von N Elementen erster Objektdaten, die N Zeitpunkten entsprechen, der mehreren Objektdatenelemente aus dem ersten Pyramidenbildsatz, wobei N eine natürliche Zahl ist; und

Erzeugen von M Elementen zweiter Objektdaten, die M Zeitpunkten entsprechen, der mehreren Objektdatenelemente aus dem zweiten Pyramidenbildsatz, wobei M eine natürliche Zahl ist.

20. Verfahren nach Anspruch 19, wobei N größer als M ist.

21. Verfahren nach Anspruch 19, wobei das Objekt ein erstes Objekt und ein zweites Objekt umfasst, und wobei das Verfahren ferner umfasst:

Verfolgen des ersten Objekts unter Verwendung von P Elementen der ersten Objektdaten aus den N Elementen der ersten Objektdaten, wobei P eine natürliche Zahl kleiner als oder gleich N ist; und

Verfolgen des zweiten Objekts unter Verwendung von Q Elementen der zweiten Objektdaten aus den

M Elementen der zweiten Objektdaten, wobei Q eine natürliche Zahl ist, die kleiner als oder gleich M ist.

22. Verfahren nach Anspruch 12, ferner umfassend:

Einstellen eines ersten interessierenden Bereichs für das erste Eingangsbild und

eines zweiten interessierenden Bereichs für das zweite Eingangsbild,

wobei das Erzeugen der mehreren Objektdatenelemente umfasst:

Extrahieren des Objekts aus einem ersten Bereich des ersten Eingangsbildes und

einem zweiten Bereich des zweiten Eingangsbildes, wobei der erste Bereich und

der zweite Bereich dem ersten interessierenden Bereich und dem zweiten interessierenden Bereich entsprechen.

23. Fahrassistenzsystem zum Fahren eines Fahrzeugs durch Detektieren eines Objekts, wobei das Fahrassistenzsystem umfasst:

einen Pyramidenbildgenerator, der eingerichtet ist zum Empfangen eines zu einem ersten Zeitpunkt aufgenommenen ersten Eingangsbildes und eines zu einem zweiten Zeitpunkt aufgenommenen zweiten Eingangsbildes und zum Erzeugen eines ersten Pyramidenbildes aus dem ersten Eingangsbild und eines zweiten Pyramidenbildes aus dem zweiten Eingangsbild;

einen Objektextraktor, der eingerichtet ist zum Detektieren des Objekts im ersten Pyramidenbild und im zweiten Pyramidenbild und zum Erzeugen mehrerer Objektdatenelemente, die das Objekt darstellen, durch Verwendung des tiefgehenden Lernens auf Basis eines neuronalen Netzes;

einen Pufferspeicher, der die mehreren Objektdatenelemente speichert, die das im ersten Eingangsbild und im zweiten Eingangsbild detektierte Objekt darstellen; und

einen Objekttracker, der eingerichtet ist zum Verfolgen des Objekts auf Grundlage der im Pufferspeicher gespeicherten mehreren Objektdatenelemente.

24. Fahrassistenzsystem nach Anspruch 23, wobei der Pyramidenbildgenerator ferner eingerichtet ist zum:

Erzeugen eines ersten Pyramidenbildsatzes, der das erste Pyramidenbild und das zweite Pyramidenbild umfasst; und

Erzeugen eines zweiten Pyramidenbildsatzes durch Downsampling des ersten Pyramidenbildes und/oder zweiten Pyramidenbildes, und

wobei der Objektextraktor ferner eingerichtet ist zum:

Erzeugen von N Elementen erster Objektdaten, die N Zeitpunkten entsprechen, der mehreren Objektdatenelemente aus dem ersten Pyramidenbildsatz, wobei N eine natürliche Zahl ist; und

Erzeugen von M Elementen zweiter Objektdaten, die M Zeitpunkten entsprechen, der mehreren Objekt-

datenelemente aus dem zweiten Pyramidenbildsatz, wobei  $M$  eine natürliche Zahl ist.

25. Verfahren zum Detektieren eines Objekts, wobei das Verfahren umfasst:

Erzeugen eines ersten Pyramidenbildsatzes, der eine erste Mehrzahl von Pyramidenbildern umfasst, aus einer Mehrzahl von Eingangsbildern, die entsprechend zu mehreren Zeitpunkten aufgenommen werden;

Erzeugen eines zweiten Pyramidenbildsatzes durch Downsampling mindestens eines Pyramidenbildes aus der ersten Mehrzahl von Pyramidenbildern im ersten Pyramidenbildsatz;

Erzeugen von  $N$  Elementen erster Objektdaten, die  $N$  Zeitpunkten entsprechen, aus dem ersten Pyramidenbildsatz, wobei  $N$  eine natürliche Zahl ist; und

Erzeugen von  $M$  Elementen zweiter Objektdaten, die  $M$  Zeitpunkten entsprechen, aus dem zweiten Pyramidenbildsatz, wobei  $M$  eine natürliche Zahl ist, wobei  $N$  größer als  $M$  ist.

Es folgen 16 Seiten Zeichnungen

Anhängende Zeichnungen

FIG. 1

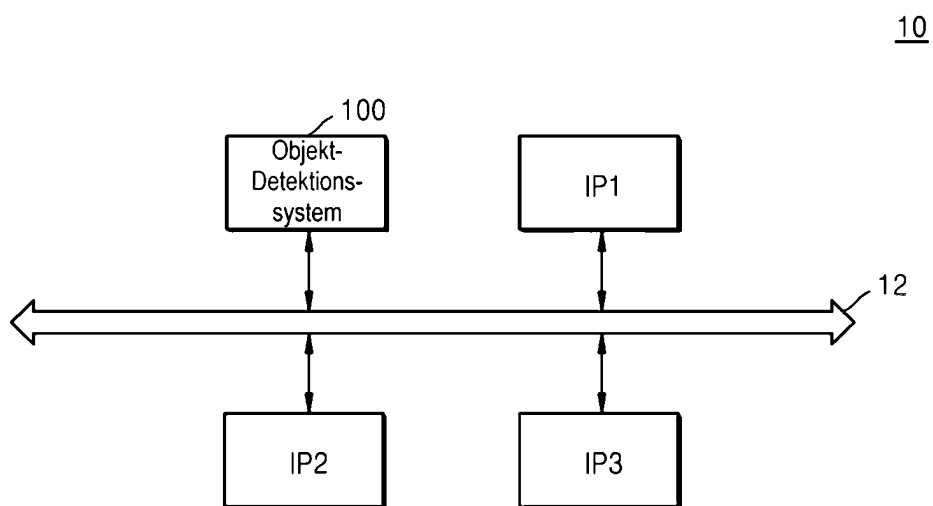


FIG. 2

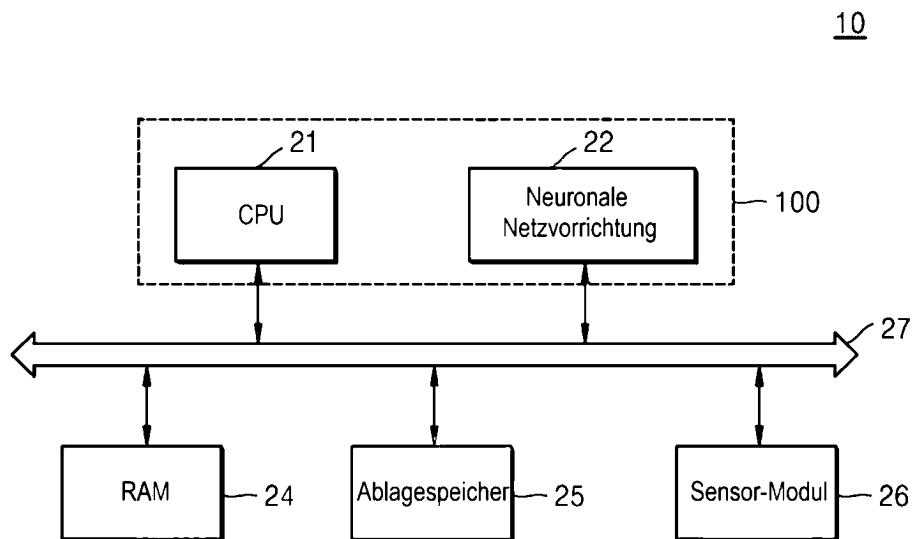


FIG. 3

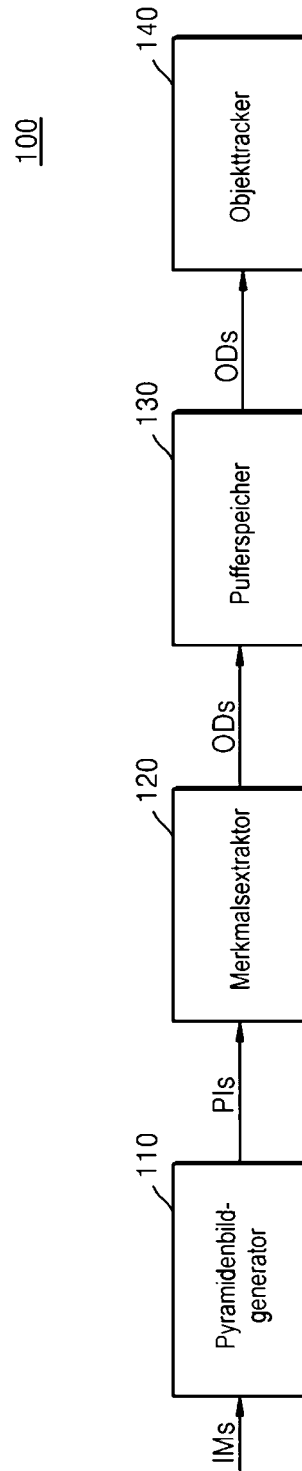




FIG. 4

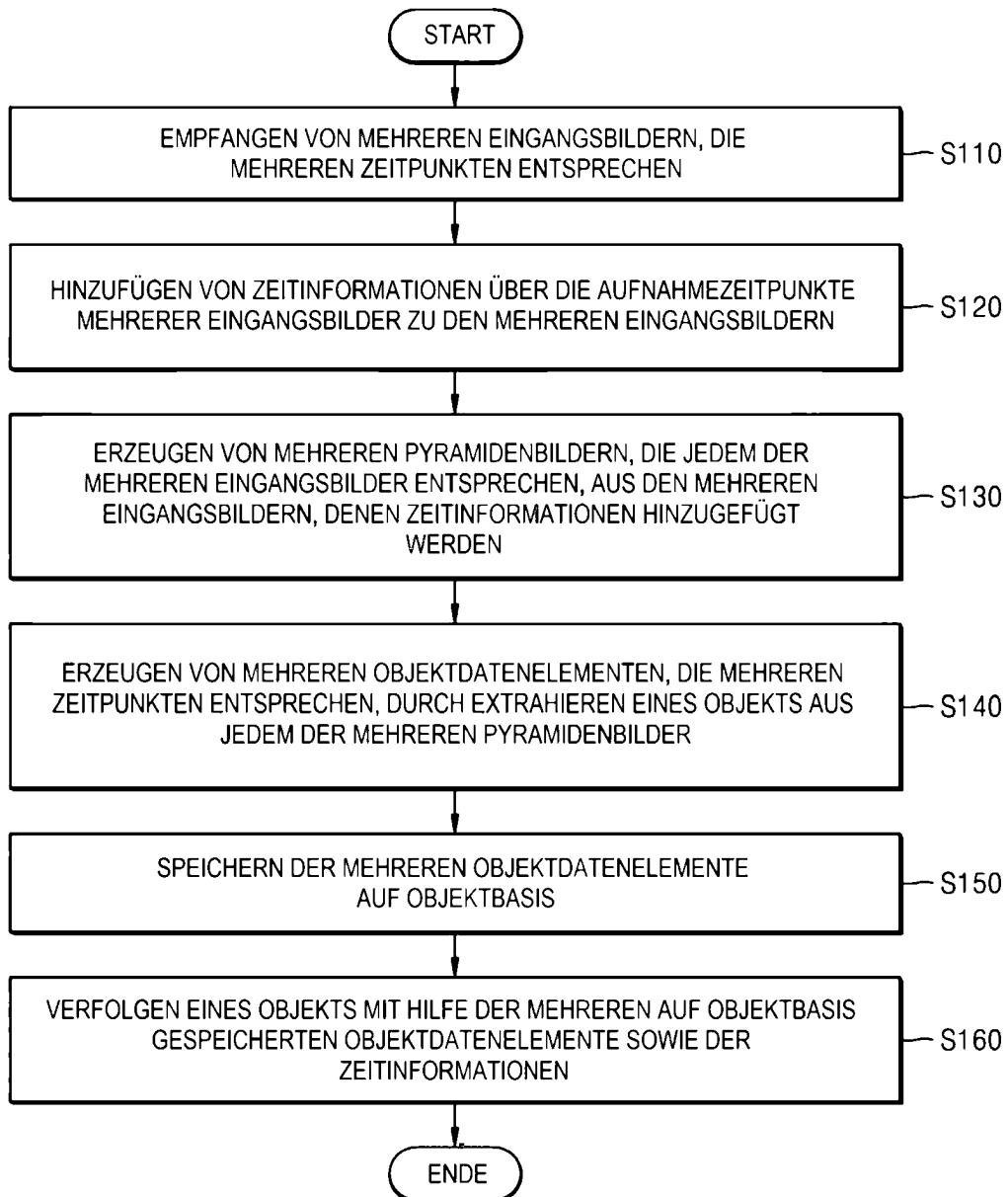


FIG. 5

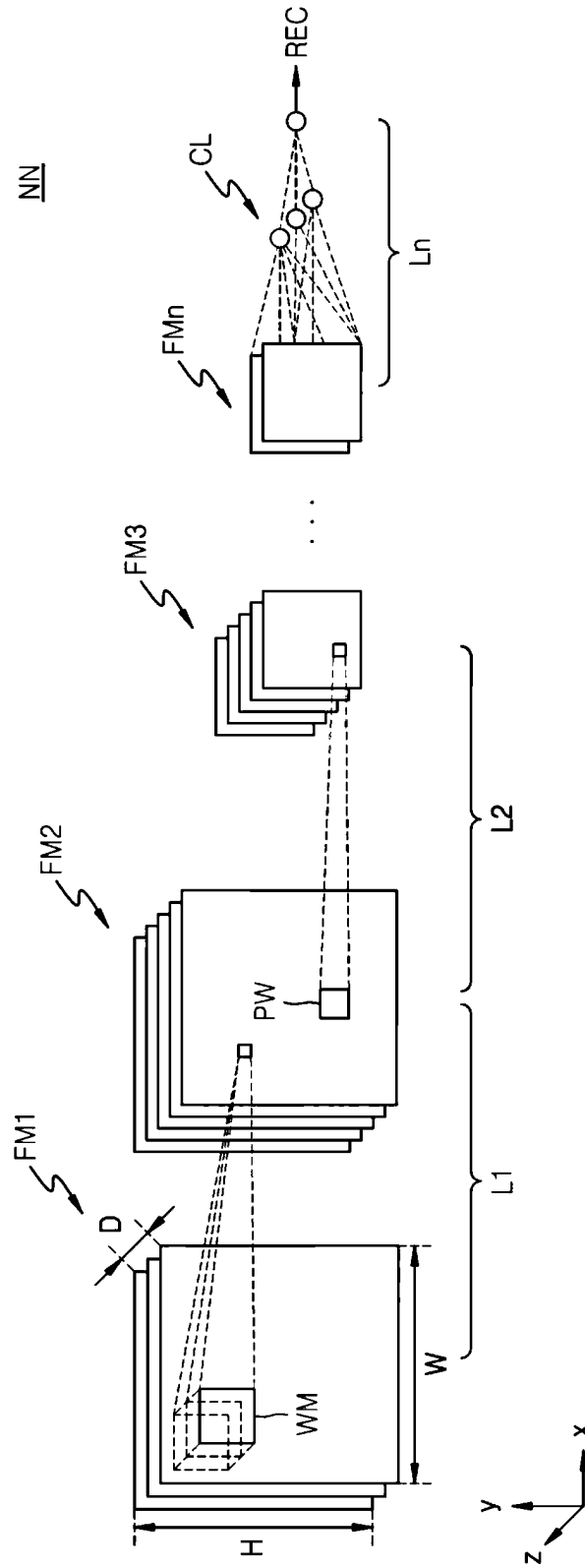


FIG. 6

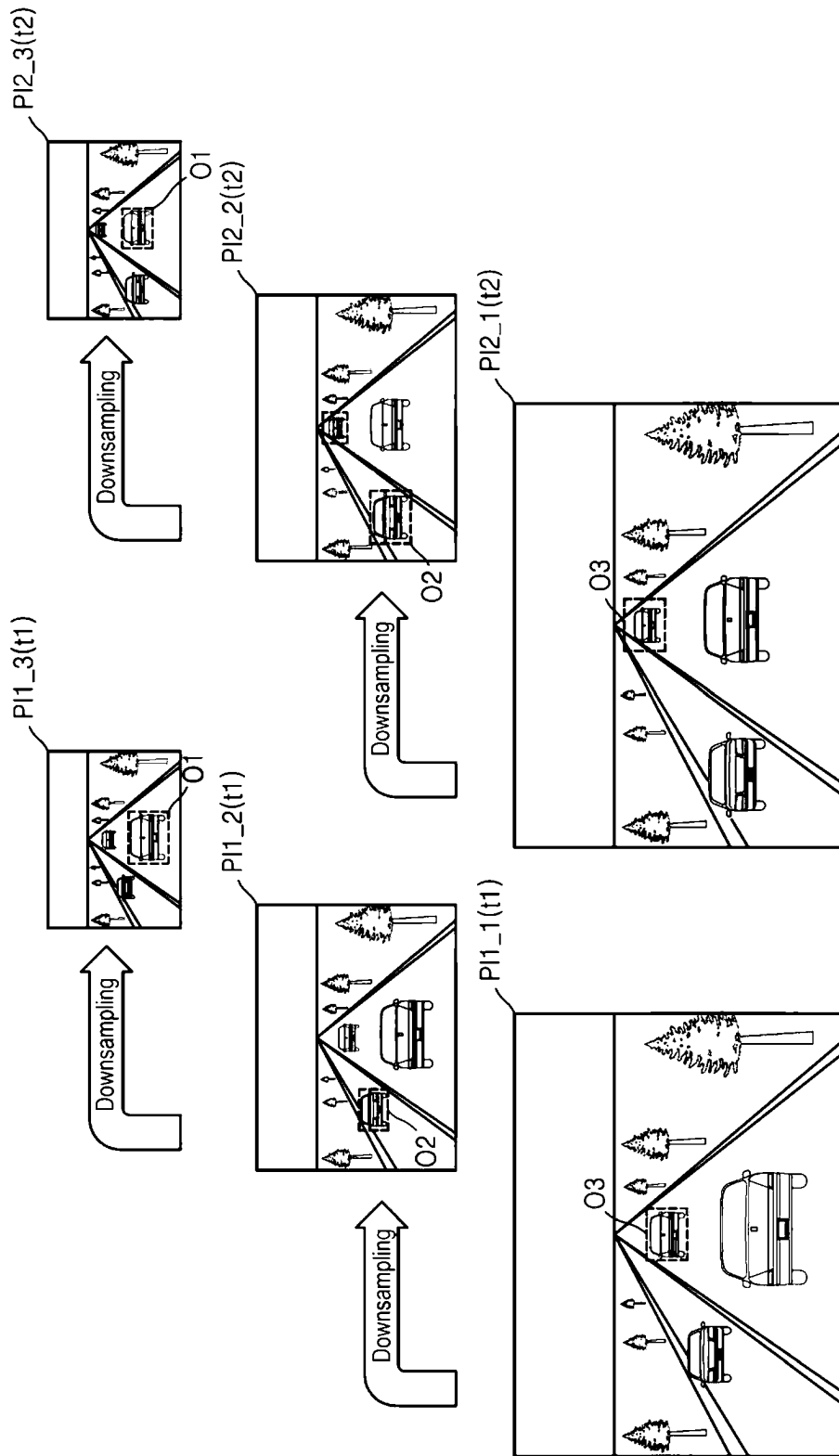


FIG. 7

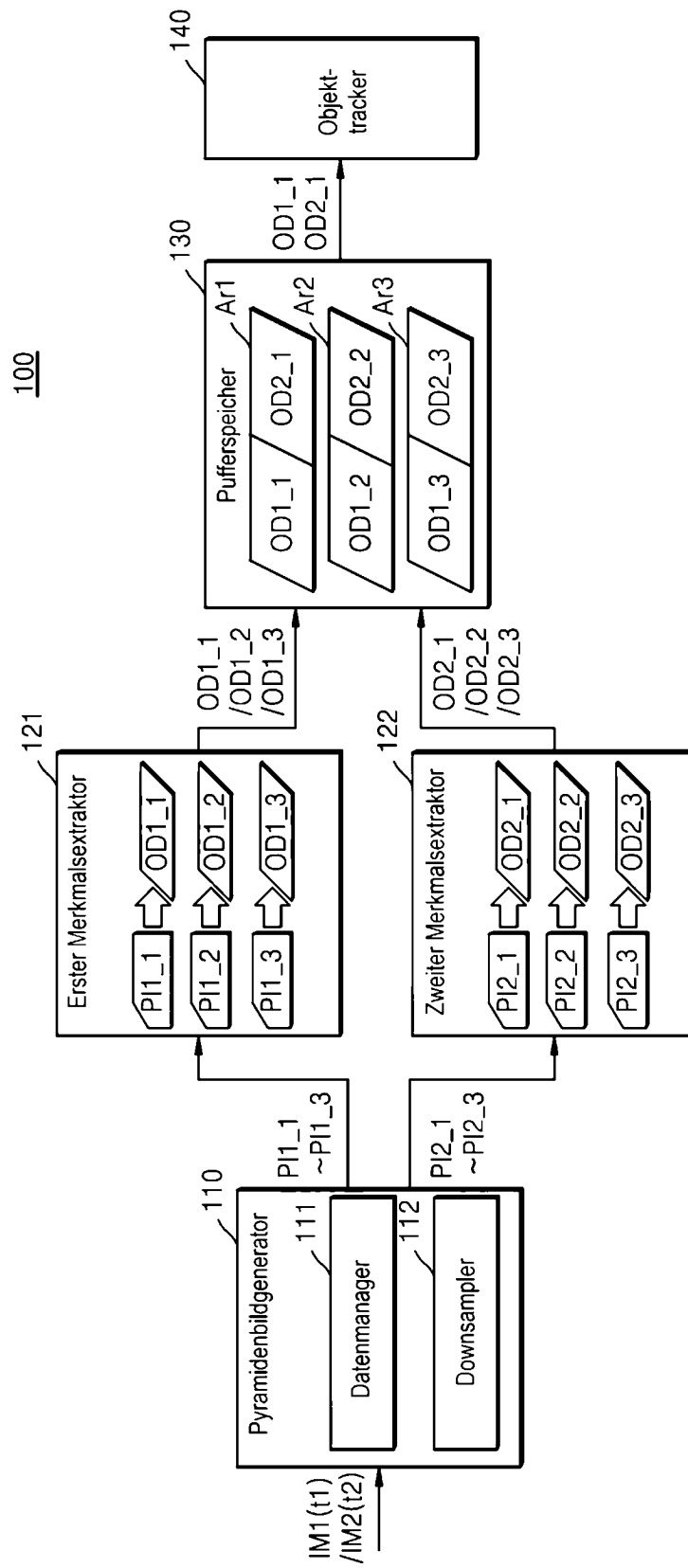


FIG. 8

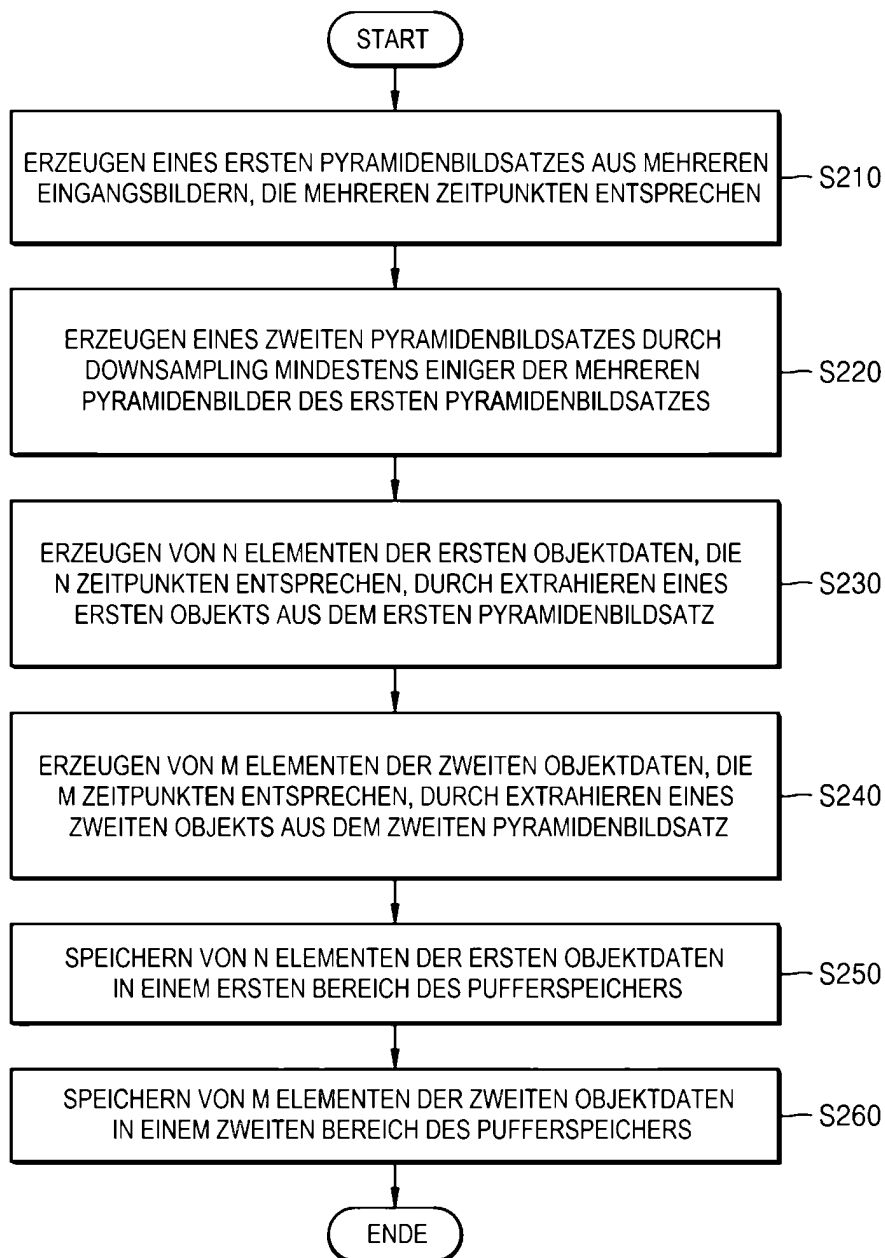


FIG. 9

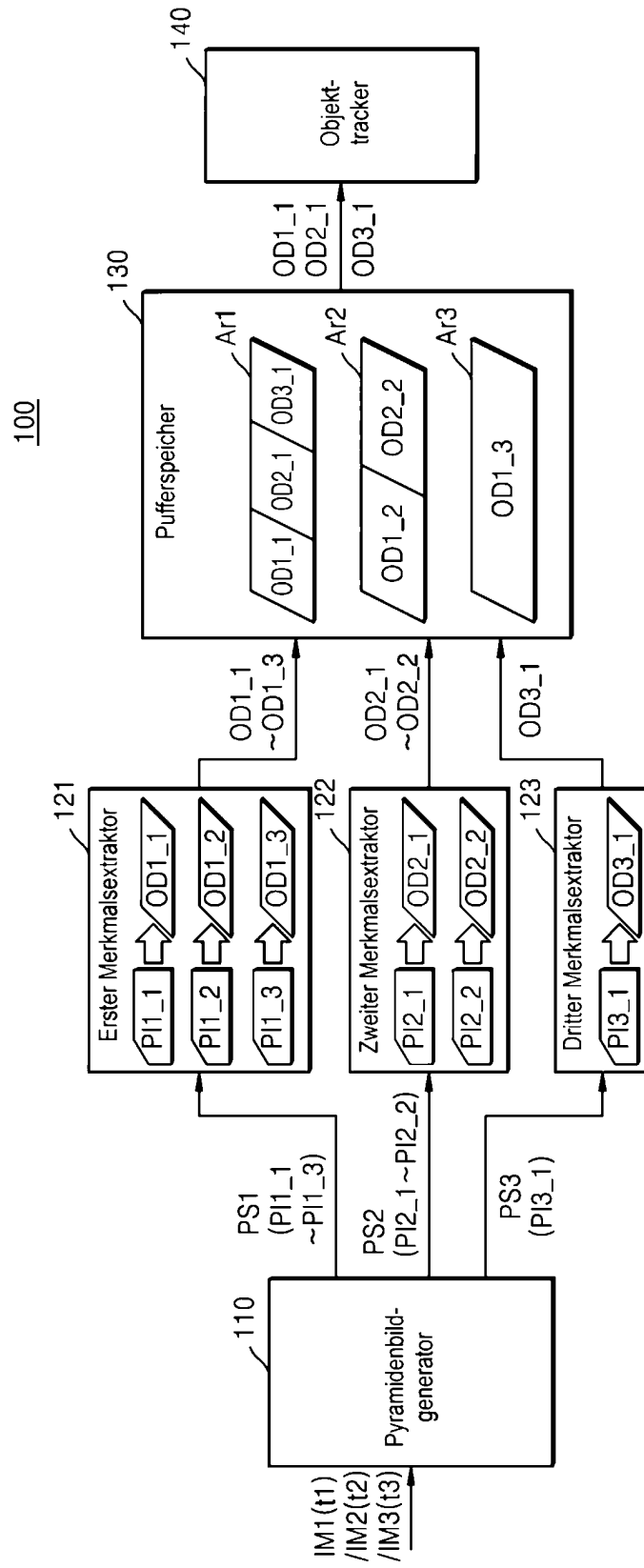


FIG. 10

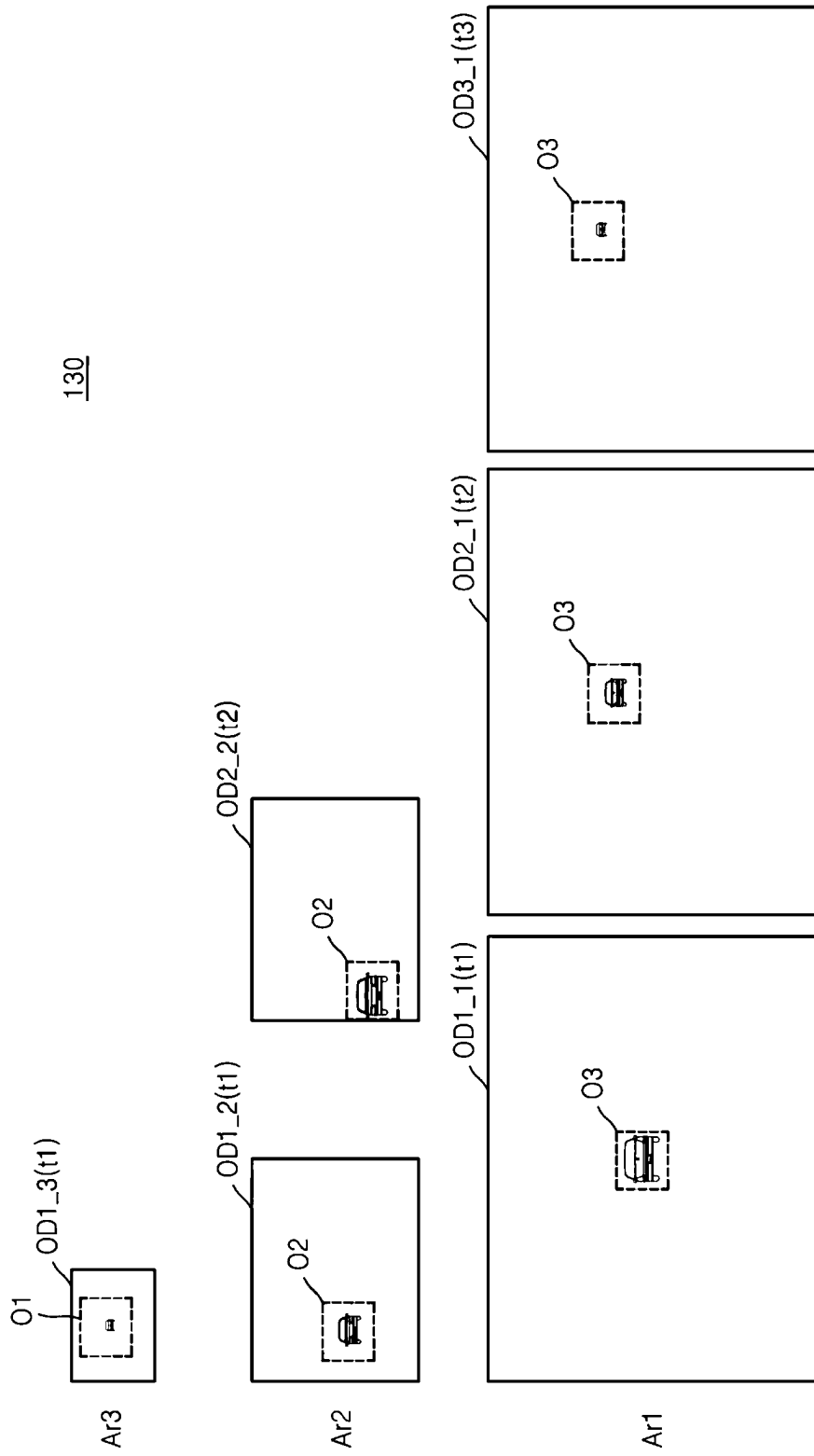


FIG. 11

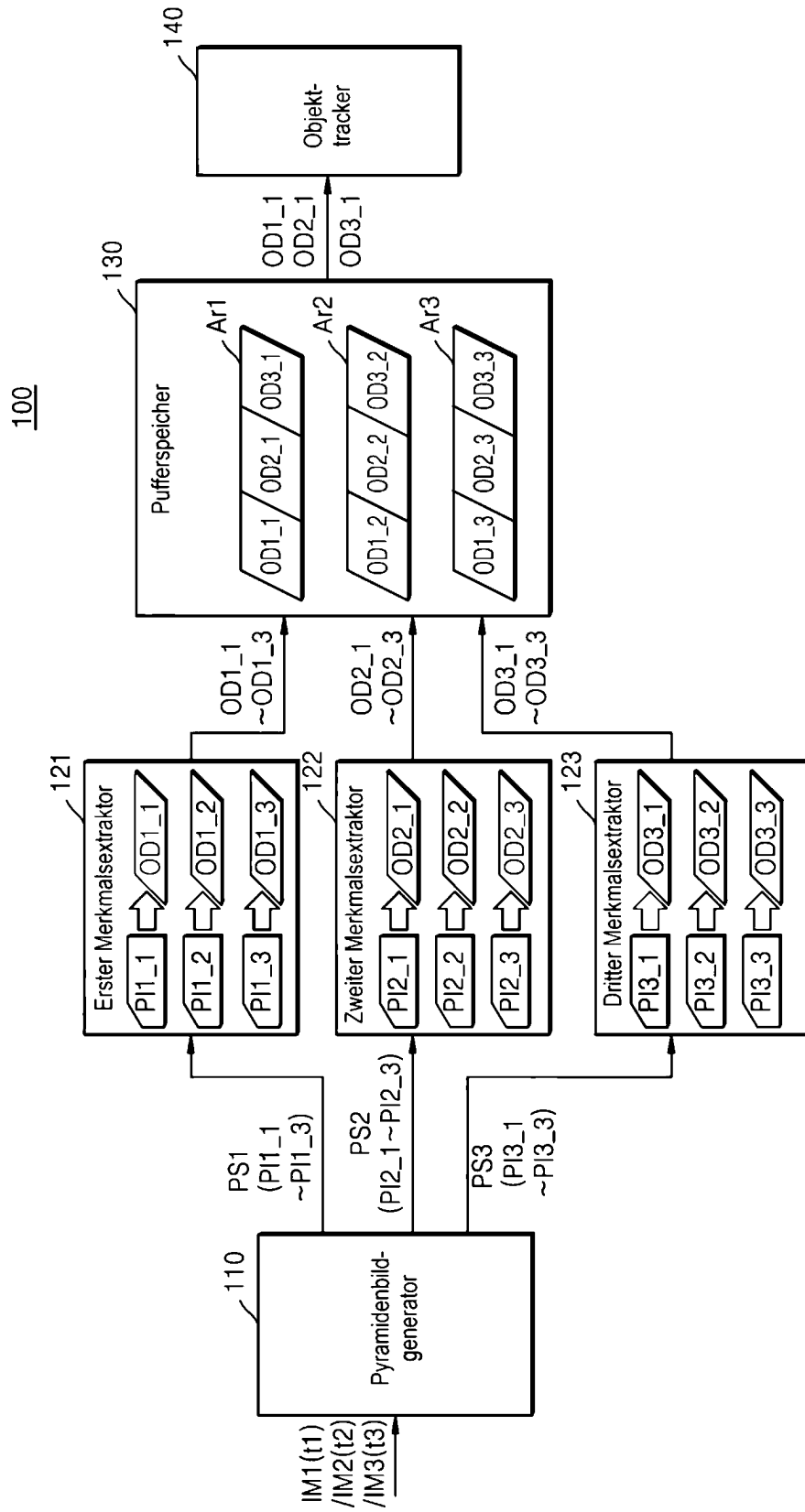




FIG. 12

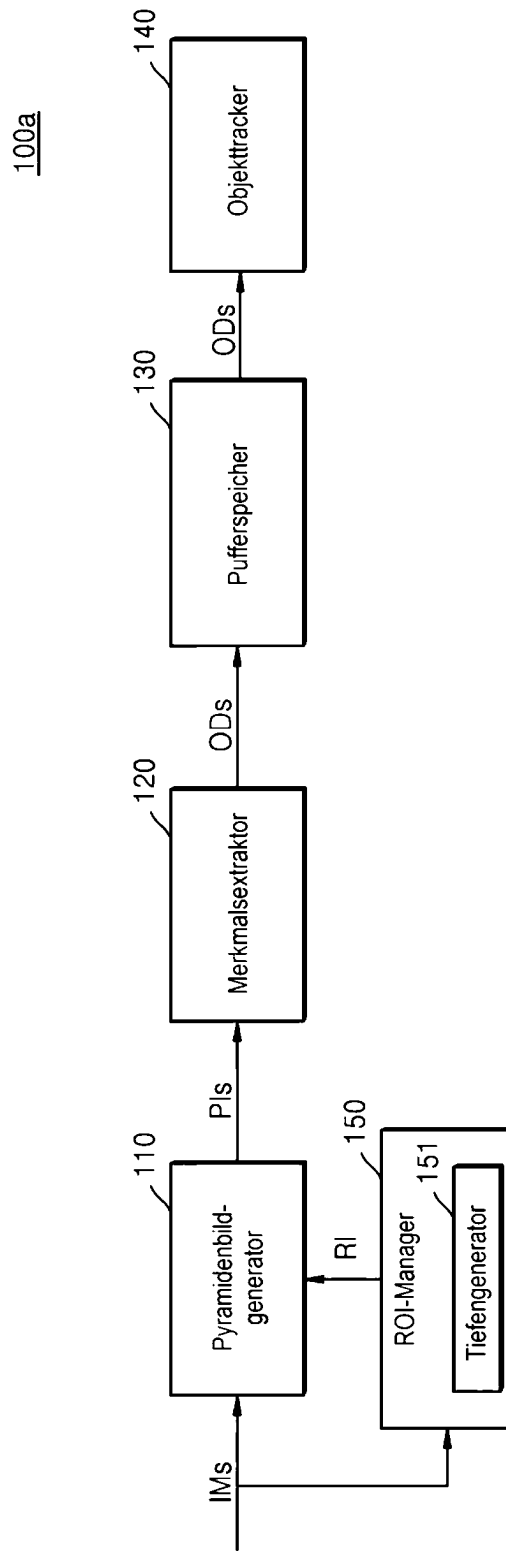


FIG. 13

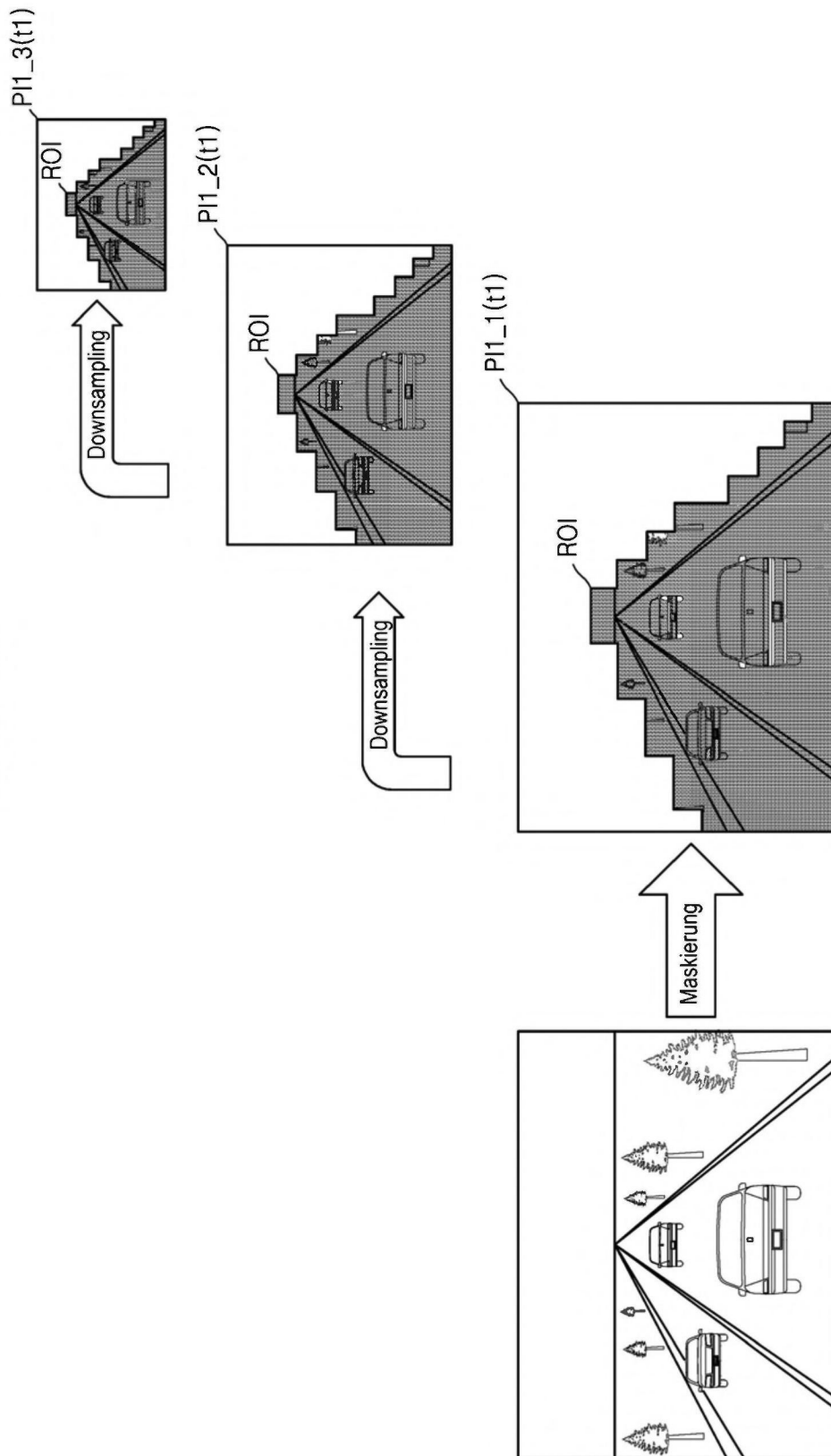


FIG. 14

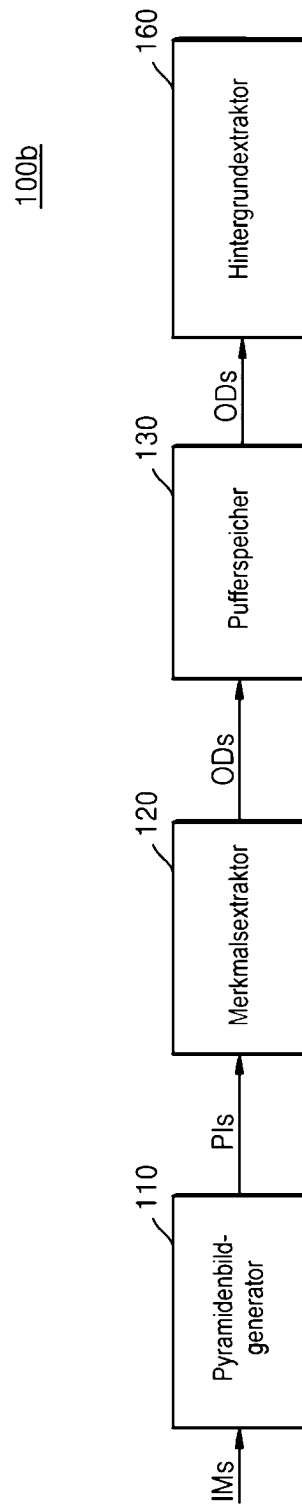


FIG. 15

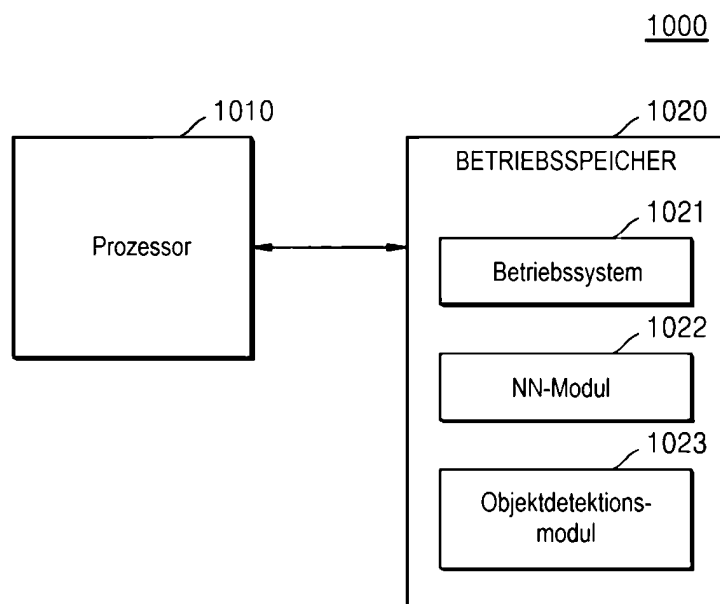


FIG. 16

2000

