

(12) NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES
PATENTWESENS (PCT) VERÖFFENTLICHTE INTERNATIONALE ANMELDUNG

(19) Weltorganisation für geistiges Eigentum
Internationales Büro

(43) Internationales Veröffentlichungsdatum
05. April 2018 (05.04.2018)



(10) Internationale Veröffentlichungsnummer
WO 2018/059629 A1

(51) Internationale Patentklassifikation:
G06K 9/00 (2006.01)

(21) Internationales Aktenzeichen: PCT/DE2017/200100

(22) Internationales Anmeldedatum:
28. September 2017 (28.09.2017)

(25) Einreichungssprache: Deutsch

(26) Veröffentlichungssprache: Deutsch

(30) Angaben zur Priorität:
10 2016 218 851.8
29. September 2016 (29.09.2016) DE

(71) Anmelder: **CONTI TEMIC MICROELECTRONIC
GMBH** [DE/DE]; Sieboldstraße 19, 90411 Nürnberg (DE).

(72) Erfinder: **WALTER, Michael**; Wiesenstraße 2, 9435
Heerbrugg (CH).

(81) Bestimmungsstaaten (soweit nicht anders angegeben, für
jede verfügbare nationale Schutzrechtsart): AE, AG, AL,
AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY,
BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM,
DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT,
HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN,
KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD,
ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO,
NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW,
SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM,
TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Bestimmungsstaaten (soweit nicht anders angegeben, für
jede verfügbare regionale Schutzrechtsart): ARIPO (BW,
GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST,
SZ, TZ, UG, ZM, ZW), eurasisches (AM, AZ, BY, KG, KZ,

(54) Title: DETECTION AND VALIDATION OF OBJECTS FROM SEQUENTIAL IMAGES FROM A CAMERA BY MEANS
OF HOMOGRAPHS

(54) Bezeichnung: DETEKTION UND VALIDIERUNG VON OBJEKTEN AUS SEQUENTIELLEN BILDERN EINER KAMERA
MITTELS HOMOGRAPHIEN

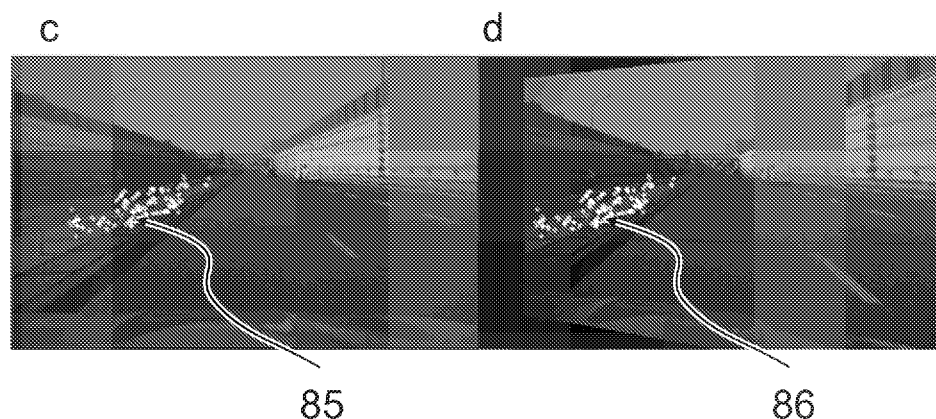


Fig. 8

(57) Abstract: The invention relates to a method and to a device for identifying objects from images from a camera and can be used in particular in camera-based driver assistance systems. The method for detecting objects from a series of images from a vehicle camera comprises the following steps: a) capturing a series of images by means of the vehicle camera, b) determining corresponding features in two consecutive images, d) associating determined adjacent corresponding features in an image region with a plane in space, and f) determining additional corresponding features in the image region while taking into account the associated plane. The invention offers the advantage of compression of the optical flow field.

(57) Zusammenfassung: Die Erfindung betrifft ein Verfahren und eine Vorrichtung zur Erkennung von Objekten aus Bildern einer Kamera und kann insbesondere bei kamerabasierten Fahrerassistenzsystemen verwendet werden. Das Verfahren zur Detektion von Objekten aus einer Folge von Bildern einer Fahrzeugkamera umfasst die Schritte: a) Aufnahme einer Folge von Bildern mit der Fahrzeugkamera, b) Ermittlung von korrespondierenden Merkmalen in zwei aufeinander folgenden Bildern, d) Zuordnung ermittelter benachbarter korrespondierender Merkmale in einem Bildbereich zu einer Ebene im Raum, und f) Bestimmung zusätzlicher korrespondierender Merkmale in dem Bildbereich unter Berücksichtigung der zugeordneten Ebene. Die Erfindung bietet den Vorteil einer Verdichtung des optischen Flussfeldes.

[Fortsetzung auf der nächsten Seite]



WO 2018/059629 A1



RU, TJ, TM), europäisches (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Veröffentlicht:

- mit internationalem Recherchenbericht (Artikel 21 Absatz 3)
- vor Ablauf der für Änderungen der Ansprüche geltenden Frist; Veröffentlichung wird wiederholt, falls Änderungen eingehen (Regel 48 Absatz 2 Buchstabe h)

DETEKTION UND VALIDIERUNG VON OBJEKTEN AUS SEQUENTIELLEN BILDERN EINER KAMERA MITTELS HOMOGRAPHIEN

Die Erfindung betrifft ein Verfahren zur Erkennung von Objekten
5 aus Bildern einer Kamera und kann insbesondere bei kameraba-
sierten Fahrerassistenzsystemen verwendet werden.

Fahrzeugerkennungssysteme nach dem aktuellen Stand der Technik
sind meist klassifikationsbasiert. Klassifikationsbasierte
10 Systeme können Fahrzeuge bzw. Fahrzeugkomponenten wiederer-
kennen, die sie in ihren Trainingsdaten gesehen haben. Neue
Fahrzeugdesigns, sowie sich verändernde Aufbauten können jedoch
zu einer stark reduzierten System Performance führen und fordern
generische Ansätze zur Objekterkennung.

15 EP 2 993 654 A1 zeigt ein Verfahren zur Frontkollisionswarnung
(FCW) aus Kamerabildern. Hierbei wird ein Bildausschnitt
analysiert, wohin das eigene Fahrzeug in einem vorgegebenen
Zeitintervall gelangen wird. Sofern dort ein Objekt erkannt wird,
20 wird eine Kollisionswarnung ausgegeben.

US 2014/0161323 A1 zeigt ein Verfahren zur Erzeugung dichter
dreidimensionaler Strukturen in einer Straßenumgebung aus
Bildern, die mit einer Monokamera aufgenommen werden.

25 Es ist eine Aufgabe der vorliegenden Erfindung, ein verbessertes
Verfahren zur Detektion von Objekten anzugeben.

Ein Ausgangspunkt der Erfindung sind die folgenden Überlegungen:
30 Sind die Kamerapositionen zweier Frames (Einzelbilder) bekannt,
lassen sich Punkt-Korrespondenzen (korrespondierende Merk-
malpunkte) triangulieren, aber es werden keine Objekte ge-
neriert, da die Triangulation über kein Modelwissen verfügt, das
eine Punktwolke in sinnvolle Objekten clustern könnte.

Nachteile monokularer Systeme sind, dass Objekte nahe dem Epipol nur ungenau trianguliert werden können und sich dort kleinste Fehler in der Egomotion (Kamera-Eigenbewegung) bemerkbar machen. Als Epipol bezeichnet man den Bildpunkt in einem ersten Kamerabild, an dem das Zentrum der Kamera zu einem zweiten Zeitpunkt abgebildet wird. Während einer Geradeausfahrt entspricht z.B. der Fluchtpunkt dem Epipol. Dies ist jedoch der relevante Bereich, um Kollisionen mit stehenden bzw. vorrausfahrenden Fahrzeugen zu erkennen. Dynamische Objekte können trianguliert werden, wenn sie sich gemäß der Epipolar-Geometrie bewegen. Sie werden jedoch aufgrund der nicht bekannten Relativgeschwindigkeit zu nah oder zu weit entfernt geschätzt.

Werden anstelle einzelner Korrespondenzen, mehrere (benachbarte) Korrespondenzen (korrespondierende Merkmale) betrachtet, lassen sich Objekte aufgrund unterschiedlicher Geschwindigkeiten, Skalierungen und Deformation segmentieren.

Ein erfindungsgemäßes Verfahren zur Detektion von Objekten aus einer Folge von Bildern einer Fahrzeugkamera umfasst die Schritte:

- a) Aufnahme einer Folge von Bildern mit der Fahrzeugkamera,
- b) Ermittlung von korrespondierenden Merkmalen in zwei aufeinander folgenden Bildern,
- d) Zuordnung (benachbarter) ermittelter korrespondierender Merkmale in einem Bildbereich zu einer Ebene im Raum, und
- f) Bestimmung zusätzlicher korrespondierender Merkmale in dem Bildbereich unter Berücksichtigung der (in Schritt d)) zugeordneten Ebene.

Bevorzugt ist die Fahrzeugkamera zur Aufnahme einer Umgebung eines Fahrzeugs ausgebildet. Bei der Umgebung handelt es sich insbesondere um die vor dem Fahrzeug liegende Umgebung. Vorzugsweise ist die Fahrzeugkamera in eine Fahrerassistenzvor-

richtung integrierbar oder mit dieser verbindbar, wobei die Fahrerassistenzvorrichtung insbesondere zur Objekterkennung aus den von der Fahrzeugkameravorrichtung bereitgestellten Bild-
daten ausgebildet ist. Bevorzugt ist die Fahrzeugkameravor-
richtung eine im Innenraum des Kraftfahrzeugs hinter der
Windschutzscheibe anzuordnende und in Fahrtrichtung gerichtete
Kamera. Besonders bevorzugt ist die Fahrzeugkamera eine mo-
nokulare Kamera.

Bevorzugt werden mit der Fahrzeugkamera zu bestimmten bzw.
bekannten Zeitpunkten Einzelbilder aufgenommen, woraus sich eine
Folge von Bildern ergibt.

Als Korrespondenz wird die Entsprechung eines Merkmals in einem
ersten Bild zu demselben Merkmal in einem zweiten Bild be-
zeichnet. Korrespondierende Merkmale in zwei Bildern können auch
als Flussvektor beschrieben werden, der angibt wie sich das
Merkmal im Bild verschoben hat. Ein Merkmal kann insbesondere ein
Bildausschnitt (bzw. Patch), ein Pixel, eine Kante oder eine Ecke
sein.

Unter Schritt d) wird auch subsummiert, dass eine Mehrzahl von
Ebenen im Raum vorgegeben wird, und eine Zuordnung von (be-
nachbarten) korrespondierenden Merkmalen zu jeweils einer der
vorgegebenen Eben vorgenommen wird (vgl. unten Schritt d2) bzw.
d3)).

Der Begriff „Ebene“ beschreibt im Kontext der vorliegenden
Erfindung folgende Zusammenhänge: einerseits ein Kriterium zur
Akkumulation benachbarter korrespondierender Merkmale. D.h.
diese werden als zusammengehörig angesehen, wenn sie in einer
gemeinsamen Ebene im Raum liegen und sich entsprechend der
Bewegung der Ebene zeitlich entwickeln.

Derart akkumulierte korrespondierende Merkmale werden an-
schließend als z.B. „Bodenebene“ bezeichnet, da sie alle in der

Ebene, die der Fahrbahnebene entspricht liegen. Jedoch erstreckt sich eine solche Bodenebene nicht ins Unendliche, sondern meint einen Teilbereich der Ebene, nämlich den, in dem tatsächlich korrespondierende Merkmale angeordnet sind.

5

In Schritt f) meint die Formulierung „unter Berücksichtigung...“, dass die in Schritt d) korrespondierenden Merkmalen in einem Bildbereich zugeordnete Ebene bei der Bestimmung zusätzlicher korrespondierender Merkmale berücksichtigt wird. Dies kann
10 beispielsweise in der Art geschehen, dass anhand von wenigen korrespondierenden Merkmalen in einem Bildbereich eine Bewegung der zugeordneten Ebene bestimmt bzw. berechnet wird und dann prädiziert wird, wo Merkmale desselben Bildbereichs in einem
ersten Bild infolge der Bewegung der zugeordnete Ebene in einem
15 (nachfolgenden) zweiten Bild aufzufinden sein werden. Mit Hilfe dieser Prädiktion werden zusätzliche korrespondierende Merkmale ermittelt.

Mit der Formulierung „Detektion von Objekten“ kann also bei-
20 spielsweise eine Generierung von Objekthypothesen bzw. Objekten gemeint sein.

Gemäß einer bevorzugten Ausführungsform umfasst das Verfahren den Schritt:

25 c) Berechnung von Homographien für die ermittelten korrespondierende Merkmale in einem Bildbereich, damit diese einer Ebene im Raum zugeordnet werden können.

Eine Homographie beschreibt die Korrespondenz von Punkten auf einer Ebene zwischen zwei Kamerapositionen bzw. die Korres-
30 pondenz zweier Punkte in zwei aufeinanderfolgenden Bildern der Fahrzeugkamera. Durch die Berechnung von Homographien für die ermittelten korrespondierenden Merkmale in einem Bildbereich kann so die Zuordnung zu jeweils einer Ebene im Raum erfolgen (s. Schritt d)).

Die Homographien werden bevorzugt generisch aus dem Bild bzw. aus aufeinanderfolgenden Bildern bestimmt. Eine Vorgabe von einem Abstand einer Ebene zur Kameraposition ist typischerweise nicht erforderlich.

5

Vorteilhaft umfasst das Verfahren die Schritte:

d2) Zuordnung der ermittelten korrespondierenden Merkmale zu einer von einer Mehrzahl von Ebenen vorgegebener Orientierung im Raum, und

10 e) Zuordnung zu der Ebene im Raum, die für die ermittelten korrespondierenden Merkmale den kleinsten Rückprojektionsfehler ergibt, wobei der Rückprojektionsfehler den Unterschied an zwischen der gemessenen Korrespondenz eines Merkmals in zwei aufeinanderfolgenden Bildern und dem aus der berechneten Ho-
15 mographie prädizierten Korrespondenz des Merkmals angibt. Mit anderen Worten bezeichnet der Rückprojektionsfehler einer Ebene die Differenz zwischen einem Punkt x zum Zeitpunkt $t-0$ und dem gemäß der Homographie dieser Ebene abgebildeten korrespondierenden Punktes zum vorangegangenen Zeitpunkt $t-1$ beschrieben
20 wird (siehe unten: Gleichung 4).

Insbesondere anhand der berechneten Homographien können die korrespondierenden Merkmale segmentiert werden, also unterschiedlichen Bildbereichen (bzw. Segmenten) zugeordnet werden. In Schritt f) kann dann eine Bestimmung zusätzlicher korrespondierender Merkmale in einem Bildbereich unter Berücksichtigung der zugeordneten Ebene erfolgen.
25

Eine vorteilhafte Weiterbildung des Verfahrens umfasst den Schritt d3): Zuordnung von (benachbarten) korrespondierenden
30 Merkmalen zu jeweils einer Bodenebene, einer Rückwandebene oder einer Seitenwandebene. Im Falle eines Koordinatensystems, bei dem die x -Richtung horizontal bzw. lateral, die y -Richtung vertikal und die z -Richtung in Fahrzeuglängsrichtung verläuft, kann eine Bodenebene normal zur y -Richtung, eine Rückwandebene

normal zur z-Richtung und eine Seitenwandebene normal zur x-Richtung vorgegeben werden.

Durch eine Berechnung von Homographien einer Bodenebene, einer Rückwandebene und einer Seitenwandebene kann für korrespondierende Merkmale in einem Bildbereich bzw. für jedes korrespondierende Merkmal eine Zuordnung zu einer dieser Ebenen erfolgen. Der Abstand (von der Fahrzeugkamera) insbesondere zu einer Rückwand- oder Seitenwandebene ist vorteilhaft das Ergebnis der Homographieberechnung und nicht etwa eine vorgegebene Annahme.

Bevorzugt können die Homographien für die Rückwandebene nach Gleichung (10) bzw. für die Bodenebene nach Gleichung (9) bzw. für die Seitenwandebene nach Gleichung (11) berechnet werden.

Hierbei sind a , b , c Konstanten, x_0 , y_0 , x_1 , y_1 bezeichnen Korrespondenzen im ersten Bild (Index 0), aufgenommen zu einem Zeitpunkt $t-0$, und zweiten Bild (Index 1), aufgenommen zu einem früheren Zeitpunkt $t-1$, und t_x , t_y , t_z sind die Komponenten des Vektors $\mathbf{t}/d$. $\mathbf{t}$ beschreibt die Translation der Fahrzeugkamera und d die Entfernung zu einer Ebene (senkrecht zu dieser Ebene), also entlang des Normalenvektors dieser Ebene. Die Komponenten t_x , t_y bzw. t_z werden im Folgenden auch als „inverse TTC“ bezeichnet. TTC kommt von ‚Time to collision‘ und ergibt sich in einer Raumrichtung als Abstand geteilt durch Translationsgeschwindigkeit.

Vorteilhaft können ausgehend von einer bereits ermittelten Aufteilung eines Bildes in unterschiedliche Bildbereiche mit jeweils einer zugeordneten Ebene für ein zweites Bild, erste bzw. wenige einander korrespondierende Merkmale in einem Bildbereich im zweiten und einem nachfolgenden dritten Bild bestimmt werden. Aus diesen ersten bzw. wenigen korrespondierenden Merkmalen wird dann eine Homographie für diesen Bildbereich neu berechnet, und

die neu berechnete Homographie wird dazu verwendet, die Position und Form zusätzlicher bzw. weiterer korrespondierender Merkmale im dritten Bild zu präzisieren.

- 5 Bevorzugt können, falls in einem nachfolgenden dritten Bild in einem Bildbereich nicht genügend (erste) korrespondierende Merkmale zur Neuberechnung einer Homographie bestimmbar sind, die für den Bildbereich (anhand des ersten und zweiten Bildes) berechnete Homographie aus dem zweiten Bild dazu verwendet
10 werden, die Position und Form zusätzlicher korrespondierender Merkmale des Bildbereichs im dritten Bild zu präzisieren. Dadurch wird die Korrespondenzfindung robuster gegen Form- und Skalenänderungen gestaltet.
- 15 Gemäß einer bevorzugten Ausgestaltung kann für jeden Bildbereich mit einer zugeordneten Ebene ein aktuelles Bild auf ein vorheriges Bild entsprechend der berechneten Homographie (der zugeordneten Ebene) gewarpt (projiziert bzw. transformiert) werden, um zusätzliche einander im letzten und im aktuellen Bild
20 korrespondierende Merkmale zu bestimmen. Auch diese Ausgestaltung führt zu einer Verdichtung des optischen Flussfeldes.

Gemäß einer vorteilhaften Weiterbildung können, falls mehrere Ebenen mit identischer Orientierung auftreten, die Ebenen mit
25 identischer Orientierung anhand der zugehörigen t_x , t_y , t_z -Werte getrennt werden. Beispielsweise können zwei Rückwandebenen, die in z-Richtung unterschiedlich weit von der Fahrzeugkamera entfernt sind, über unterschiedliche t_z -Werte voneinander unterschieden werden.

30

Bevorzugt kann ein Bild durch ein Gitter in gleichartige Zellen unterteilt werden, und für jede Zelle kann aus den darin ermittelten korrespondierenden Merkmalen eine Homographie be-

rechnet werden. Zellen mit übereinstimmender Homographie können anschließend geclustert werden.

Bevorzugt kann, falls die berechnete Homographie einer ersten Zelle nicht hinreichend mit einer Homographie einer benachbarten Zelle übereinstimmt, zur Ermittlung einer Ebenengrenze vorteilhaft ein sogenannter Rückprojektionsfehler einzelner korrespondierender Merkmale betrachtet werden.

Korrespondierende Merkmale können durch den Rückprojektionsfehler bewertet werden. Der Rückprojektionsfehler gibt den Unterschied an zwischen dem gemessenen Fluss und dem aus der berechneten Homographie prädizierten Fluss an.

Wird der Rückprojektionsfehler eines korrespondierenden Merkmals in einer ersten Zelle mit den Rückprojektionsfehlern der Homographien der benachbarten Zellen verglichen und dieses korrespondierende Merkmal der Homographie mit dem geringstem Fehler zugewiesen werden, kann die Ebenengrenze (bzw. Segmentgrenze bzw. Clustergrenze) innerhalb der ersten Zelle verfeinert werden. Auf diese Weise können verschiedene korrespondierende Merkmale einer Zelle unterschiedlichen Ebenen zugeordnet werden.

Bevorzugt kann die Zuordnung von Ebenen zu benachbarten korrespondierenden Merkmalen im Wesentlichen im gesamten Bild der Fahrzeugkamera (z.B. in mindestens 80% der Bildfläche, bevorzugt mindestens 90%) ermittelt werden. Da das erfindungsgemäße Verfahren sehr schnell ausgestaltet werden kann, ist eine generische Objektdetektion bzw. Szeneninterpretation für nahezu das gesamte Bild in Echtzeit möglich.

Gegenstand der Erfindung ist weiterhin eine Vorrichtung zur Detektion von Objekten aus einer Folge von Bildern einer Fahrzeugkamera umfassend ein Kamerasteuergerät und eine Auswerteelektronik.

Das Kamerasteuergerät ist dazu ausgebildet,

a) eine Folge von Bildern mit der Fahrzeugkamera aufzunehmen.

Die Auswerteelektronik ist dazu ausgebildet,

b) korrespondierende Merkmale in einem Bildbereich in zwei
5 aufeinander folgenden Bildern zu ermitteln,

d) die ermittelten korrespondierender Merkmale in einem
Bildbereich zu einer Ebene im Raum zuzuordnen, und

f) zusätzliche korrespondierende Merkmale in dem Bildbereich
unter Berücksichtigung der zugeordneten Ebene zu bestimmen.

10

Das Kamerasteuergerät bzw. die Auswertungselektronik können
insbesondere einen Mikrocontroller oder -prozessor, einen
Digital Signal Processor (DSP), einen ASIC (Application Specific
Integrated Circuit), einen FPGA (Field Programmable Gate Array)
15 und dergleichen mehr sowie Software zur Durchführung der
entsprechenden Steuerungs- bzw. Auswertungsschritte umfassen.
Die vorliegende Erfindung kann somit in digitalen elektronischen
Schaltkreisen, Computer-Hardware, Firmware oder Software im-
plementiert sein.

20

Weitere Merkmale, Vorteile und Wirkungen der Erfindung ergeben
sich aus der nachfolgenden Beschreibung bevorzugter Ausführ-
ungsbeispiele der Erfindung. Dabei zeigen:

25 Fig. 1 schematisch eine typische Deformation einer sich
nähernden Rückwandebene;

Fig. 2 schematisch eine typische Deformation einer sich
nähernden Bodenebene;

30

Fig. 3 schematisch eine typische Deformation a) einer
sich schnell und b) einer sich langsam nähernden
oder weiter entfernten Rückwandebene;

Fig. 4 schematisch eine Unterteilung eines Bildes mit zwei unterschiedlichen Segmenten in Zellen;

Fig. 5 Segmentierungsergebnisse nach einem dritten Iterationsschritt;

Fig. 6 Ebenen-Orientierung zur Target Validierung (Validierung potentieller Kollisionsobjekte);

Fig. 7 Time to Collision-Betrachtung; und

Fig. 8 Projektion (bzw. Warpen) des Leitplankensegments zum Zeitpunkt $t-0$ (rechts) auf $t-1$ (links).

15

Einander entsprechende Teile sind in der Regel in allen Figuren mit denselben Bezugszeichen versehen.

In Fig. 1 ist schematisch eine Rückwandebene (back plane) dargestellt, die zu einem ersten Zeitpunkt $t-1$ den schraffiert dargestellten Bereich (20, gepunktete Linie) einnimmt. Zu einem darauf folgenden Zeitpunkt t hat sich der Abstand zwischen der Fahrzeugkamera und der Rückwandebene verringert, was zu der durch die Pfeile (d_1) angedeuteten Deformation des Bereichs (21, durchgezogene Linie) der Rückwandebene im Bild führt. Der Bereich (20; 21) skaliert bzw. vergrößert sich infolge der Relativbewegung von der Fahrzeugkamera zu der Rückwandebene.

In Fig. 2 ist schematisch eine Bodenebene (ground plane) dargestellt, die zu einem ersten Zeitpunkt $t-1$ den schraffiert dargestellten Bereich (30, gepunktete Linie) einnimmt. Dies könnte ein Abschnitt einer Fahrbahnoberfläche sein, auf der das Fahrzeug fährt. Infolge der Eigenbewegung der Fahrzeugkamera ändert sich der Bereich (im Bild) zu einem darauf folgenden

Zeitpunkt t , was zu der durch die Pfeile (d_2) skizzierten Deformation des Bereichs (32) der Bodenebene führt. Zum Zeitpunkt t begrenzen die mit 32 bezeichneten Linien den Bereich der Bodenebene. Unter der „Bodenebene“ wird hier also ein abge-
5 grenzter Bereich auf der Fahrbahnoberfläche verstanden. Der Randbereich ergibt sich z.B. aus Signaturen (bzw. Randpunkten) auf der Fahrbahnoberfläche, die in der Bilderfolge getrackt werden können.

10 Fig. 3 veranschaulicht den Unterschied zwischen einer sich schnell (Fig. 3a: 20, 21; Deformation d_1) und einer sich langsam (Fig. 3b) nähernden Rückwandebene (20, 23; Deformation d_3), falls zum Zeitpunkt $t-1$ die Rückwandebene (20) in Fig. 3a denselben Abstand zur Fahrzeugkamera aufweist wie die Rückwandebene (20)
15 in Fig. 3b.

Alternativ könnte Fig. 3 den Unterschied zwischen einer nahen Rückwandebene (Fig. 3a: 20, 21; Deformation d_1) und einer weiter entfernten Rückwandebene (20, 23; Deformation d_3) darstellen, die sich z.B. mit derselben (Relativ-) Geschwindigkeit bewegen,
20 dann wäre das in Fig. 3b dargestellte Objekt (20, 21) im realen Raum größer als das in Fig. 3a dargestellte Objekt (20, 23).

Werden anstelle einzelner Korrespondenzen, mehrere benachbarte Korrespondenzen betrachtet, lassen sich Objekte aufgrund unterschiedlicher Geschwindigkeiten, Skalierungen und Deformation segmentieren.

Geht man davon aus, dass die Welt aus Ebenen besteht, kann man diese durch Homographien beschreiben und wie im Folgenden gezeigt wird über Ihre Distanz, Geschwindigkeit und Orientierung
25 trennen.

Eine Homographie beschreibt die Korrespondenz von Punkten auf einer Ebene zwischen zwei Kamerapositionen bzw. die Korrespondenz zweier Punkte in zwei aufeinanderfolgenden Frames:

$$x_{t0} = H * x_{t1} \text{ mit } x_{t0} = \begin{bmatrix} x_0 \\ y_0 \\ 1 \end{bmatrix}, \quad x_{t1} = \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix}. \quad (1)$$

Hierbei beschreibt der Vektor x_{t0} die 3D-Korrespondenz zum Zeitpunkt t-0 des Vektors x_{t1} zum Zeitpunkt t-1. Eine Homographie kann bildbasiert durch Kenntnis von vier Punkt-Korrespondenzen berechnet werden (vgl. Tutorial: Multiple View Geometry, Hartley, R. and Zisserman, A., CVPR June 1999:

<https://de.scribd.com/document/96810936/Hartley-Tut-4up> ab-

gerufen am 26.09.2016). Die auf Seite 6 des Tutorials links oben (Folie 21) angegebenen Zusammenhänge lassen sich in der Notation von Gleichung 1 wie folgt formulieren:

$$\begin{bmatrix} -x_0 & -y_0 & -1 & 0 & 0 & 0 & x_1x_0 & x_1y_0 & x_1 \\ 0 & 0 & 0 & -x_0 & -y_0 & -1 & y_1x_0 & y_1y_0 & y_1 \\ & & & & \vdots & & & & \end{bmatrix} * \begin{bmatrix} h_1 \\ h_2 \\ h_3 \\ h_4 \\ h_5 \\ h_6 \\ h_7 \\ h_8 \\ h_9 \end{bmatrix} = 0 \quad (2)$$

15 mit $H = \begin{bmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & h_9 \end{bmatrix}.$

Alternativ kann bei Kenntnis der Kameratranslation t , der Rotation R und der Entfernung d entlang des Normalen Vektors n der Ebene die Homographie nach Gleichung 3 berechnet werden.

20 Gleichung 3 verdeutlicht, dass sich bei einer inversen TTC t/d ungleich Null, Ebenen mit unterschiedlichen Orientierung n modellieren lassen und dass sich Ebenen mit identischer Orientierung n über ihre inverse TTC trennen lassen.

$$H = [R - \frac{t * n'}{d}] \quad (3)$$

Eine Homographie lässt sich theoretisch in den Normalenvektor n , die Rotationsmatrix R und die inverse TTC t/d zerlegen. Leider ist diese Zerlegung numerisch äußerst instabil und empfindlich
5 auf Messfehler.

Beschreibt man eine Szene durch Ebenen, lässt sie sich wie im Folgenden angeben segmentieren.

Fig. 4 zeigt schematisch eine Unterteilung in Zellen (Grid, Gitter/-linien). Die Szene wird in $N \times M$ initiale Zellen unterteilt und jeder Punkt-Korrespondenz wird eine eindeutige ID zugewiesen. Diese ID gibt zunächst die Zugehörigkeit zu einer Zelle an. Im weiteren Verlauf kann die ID die Zugehörigkeit zu einem Cluster oder einem Objekt angeben.

Schraffiert ist ein Objekt (insb. eine Rückwandebene) im Vordergrund dargestellt. Der Hintergrund ist weiß dargestellt. Beinhaltet eine Zelle nur ein Objekt (Zellen B3, D3) wird eine Homographie diese Zelle sehr gut beschreiben. Beinhaltet eine Zelle jedoch mehr als ein Objekt (Zelle C3), wird die Homographie keines der beiden Objekte gut beschreiben. Werden die Punkt-Korrespondenzen (schwarzer Punkt bzw. schwarzes Kreuz bzw. x) den Clustern (bzw. Segment) der benachbarten Zellen (B3 bzw. D3) über Ihre Rückprojektionsfehler zugeordnet, wird der schwarze Punkt dem Segment der Zelle B3 und das schwarze Kreuz dem Segment der Zelle D3 zugeordnet, da die Homographie für die Zelle C3 weder den Vordergrund noch den Hintergrund gut beschreibt.

Ist Vorwissen über eine Szene vorhanden, lassen sich die
10 Segmentgrößen an die Szene anpassen, indem z.B. größere Bereiche im Nahbereich des Fahrzeuges oder in Bereichen mit positiver Klassifikations-Antwort generiert werden. Für jedes Segment

wird, wie in den Gleichungen 5 bis 10 gezeigt wird, eine dedizierte Back-/Ground- und Side-Plane-Homographie berechnet.

Die Berechnung der Back-/Ground- und Side-Plane-Homographie, erhöht die Trennschärfe, da eine Homographie mit weniger
 5 Freiheitsgraden Bereiche, die unterschiedliche Ebenen beinhalten, nur schlecht modellieren kann und somit korrespondierende Punkte einen höheren Rückprojektionsfehler aufweisen werden, siehe Fig. 4. Der Rückprojektionsfehler e_i ist also ein Maß dafür, wie gut ein Punkt x zum Zeitpunkt $t-0$ durch die
 10 Homographie einer Ebene i des korrespondierenden Punktes zum Zeitpunkt $t-1$ beschrieben wird:

$$e_i = x_{t0} - H_i x_{t1}. \quad (4)$$

15 Setzt man die statische Einbaulage der Kamera und Kamera Rotation in zwei unterschiedlichen Ansichten als gegeben voraus (z.B. durch Kenntnis der Kamera Kalibration und durch die Berechnung der Fundamental-Matrix in einem monokularen System oder durch Rotationswerte eines Drehratensensor-Clusters), lässt die
 20 inverse TTC t/d mittels der um die statische Kamera-Rotation kompensierten Flussvektoren berechnen, wie im Folgenden exemplarisch für eine Ground Plane $n' = [0 \ 1 \ 0]$ gezeigt wird. Ist die Rotation nicht bekannt, kann sie näherungsweise durch eine Einheitsmatrix ersetzt werden.

25 Ersetzt man den Quotienten t/d in Gleichung 3 durch die inverse

Time to Collision $\begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix}$ folgt:

$$\begin{bmatrix} x_0 \\ y_0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 & t_x & 0 \\ R & 0 & 0 \\ 0 & t_z & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix} \Rightarrow \begin{bmatrix} x_0 \\ y_0 \\ 1 \end{bmatrix} - R \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix} = - \begin{bmatrix} 0 & t_x & 0 \\ 0 & t_y & 0 \\ 0 & t_z & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix}. \quad (5)$$

Durch Einführung der Konstanten a, b, c , wobei $\begin{bmatrix} a \\ b \\ c \end{bmatrix} := R \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix}$ ist, ergibt Gleichung 5 die vereinfachte Form:

$$\begin{bmatrix} x_0 \\ y_0 \\ 1 \end{bmatrix} - \begin{bmatrix} a \\ b \\ c \end{bmatrix} = - \begin{bmatrix} t_x y_1 \\ t_y y_1 \\ t_z y_1 \end{bmatrix} \Rightarrow \begin{bmatrix} x_0 \\ y_0 \\ 1 \end{bmatrix} = \begin{bmatrix} a \\ b \\ c \end{bmatrix} - \begin{bmatrix} t_x y_1 \\ t_y y_1 \\ t_z y_1 \end{bmatrix} \quad (6)$$

5

Durch Normierung der homogenen Koordinaten ergibt sich:

$$x_0(c - t_z y_1) = a - t_x y_1 \quad (7)$$

10

$$y_0(c - t_z y_1) = b - t_y y_1 \quad (8)$$

15

Für mehr als eine Messung ergibt sich ein Gleichungssystem der Form $Mx = v$, mit einem zu bestimmenden Vektor x , einer Matrix M und einem Vektor v (siehe Gleichung 9), das sich für mindestens drei Bild-Korrespondenzen als Stützstellen durch z.B. eine Singular Value Decomposition (Singulärwertzerlegung der Matrix) oder ein Least-Square-Verfahren lösen lässt.

$$\begin{bmatrix} x_0 c - a \\ y_0 c - b \\ \vdots \end{bmatrix} = \begin{bmatrix} -y_1 & 0 & y_1 x_0 \\ 0 & -y_1 & y_1 y_0 \\ \vdots & \vdots & \vdots \end{bmatrix} * \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix} \quad (9)$$

20

Die Herleitung der Back- und Side-Plane-Homographien erfolgt analog und ergibt:

$$\begin{bmatrix} x_0 c - a \\ y_0 c - b \\ \vdots \end{bmatrix} = \begin{bmatrix} -x_1 & 0 & x_1 x_0 \\ 0 & -x_1 & x_1 y_0 \\ \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix} \quad (10)$$

25

$$\text{bzw.} \begin{bmatrix} x_0 c - a \\ y_0 c - b \\ \vdots \end{bmatrix} = \begin{bmatrix} -1 & 0 & x_0 \\ 0 & -1 & y_0 \\ \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix}. \quad (11)$$

Um größere, aus mehreren Zellen bestehende Objekte zu segmentieren, lassen sich in einem weiteren Schritt benachbarte Zellen zusammenfassen, indem die Rückprojektionsfehler $\sum x_{t_0}^i - H_j x_{t_1}^i$ bzw. $\sum x_{t_0}^j - H_i x_{t_1}^j$ über Stützstellen (siehe unten Punkt 1.: RANSAC) der benachbarten Segmente j und i und deren Homographien berechnet werden. Zwei benachbarte Cluster werden zusammengefasst, falls $\sum x_{t_0}^i - H_j x_{t_1}^i$ kleiner $\sum x_{t_0}^j - H_i x_{t_1}^j$ ist oder z.B. der auf die prädizierte Flusslänge normierte Rückprojektionsfehler unter einer einstellbaren Schwelle liegt. Insbesondere können zwei benachbarte Cluster zusammengefasst werden, falls $\sum x_{t_0}^i - H_j x_{t_1}^i$ kleiner $\sum x_{t_0}^j - H_i x_{t_1}^j$ ist und die beiden Rückprojektionsfehler $\sum x_{t_0}^i - H_j x_{t_1}^i$ und $\sum x_{t_0}^j - H_i x_{t_1}^j$ eine auf die Flusslänge normierte Schwelle unterschreiten. Alternativ lassen sich Rückprojektionsfehler als Potentiale in einem Graph verwenden und eine globale Lösung berechnen. Die Kompaktheit der Cluster lässt sich hierbei über die Kantenpotentiale im Graphen bestimmen.

Wurden die Segmente zusammengefasst, werden die Homographien neu berechnet und die Punkt-Korrespondenzen den Clustern mit geringstem Rückprojektionsfehler zugeordnet. Betrachtet man nur direkt angrenzende Cluster, lassen sich sehr kompakte Objekte generieren. Überschreitet der minimale Fehler eine einstellbare Schwelle, werden den Korrespondenzen neue (Cluster-/Objekt-)IDs zugewiesen, um teilverdeckte Objekte oder Objekte mit leicht unterschiedlicher TTC erkennen zu können. Durch die Einstellung der Schwelle kann die Auflösung (leicht) unterschiedlicher Objekte angepasst werden.

Die Rückprojektionsfehler lassen sich mit einem Bias versehen, der Kosten für zusammenhängende Bereiche reduziert oder einem Bias, der die Kosten für einen ID-Wechsel erhöht, falls Punkt-Korrespondenzen über eine längere Zeit dieselbe ID-Zugehörigkeit hatten.

Fig. 5 zeigt ein Beispiel einer Szenensegmentierung:

Fig. 5a zeigt ein Bild, das von einer Fahrzeugkamera aufgenommen wurde, die im Inneren des Fahrzeugs angeordnet ist und die vorausliegende Umgebung durch die Windschutzscheibe erfasst. Zu sehen ist eine dreispurige Fahrbahn (51), z.B. eine Autobahn. Die Fahrspuren sind durch entsprechende Fahrspurmarkierungen getrennt. Auf allen drei Fahrspuren fahren Fahrzeuge. Das auf der eigenen Fahrspur vorausfahrende Fahrzeug (53) verdeckt möglicherweise weitere auf der eigenen Fahrspur befindliche vorausfahrende Fahrzeuge. Links der dreispurigen Fahrbahn befindet sich eine bauliche erhabene Begrenzung (52) zur Gegenfahrbahn. Rechts der dreispurigen Fahrbahn (51) befindet sich ein Rand- bzw. Standstreifen, der nach rechts von einer Leitplanke begrenzt wird, hinter der sich ein Waldgebiet anschließt. In einiger Entfernung vor dem eigenen Fahrzeug sind Schilderbrücken (54) zu erkennen, von denen eine die dreispurige Fahrbahn (51) überspannt.

Analog dem anhand von Fig. 4 beschriebenen Verfahren, kann diese Szene segmentiert werden. In Fig. 5b bis 5d sind Zellen (56) zu erkennen. In den Zellen sind Punktkorrespondenzen (55) dargestellt. Die Zuordnung einer Zelle (56) zu einem Segment ist über die Farbe des Zellrahmens bzw. der Punktkorrespondenzen (55) dargestellt.

Fig. 5b zeigt den roten Kanal des segmentierten Bildes, Fig. 5c den grünen Kanal und Fig. 5d den blauen Kanal.

Unterschiedliche Segmente wurden mit unterschiedlichen Farben versehen. Ein Segment, welches im Original grün ist, erstreckt sich über die untersten fünf bis sechs Zeilen (in Fig. 5b und 5d entsprechend weiss dargestellt und ohne Zellrahmen). Dieses Segment entspricht der Bodenebene, also der Oberfläche der Fahrbahn (51), auf der das eigene Auto fährt.

Ein weiteres Segment ist in der Mitte des Bildes zu erkennen, im Original ist es pink. Daher weist es in Fig. 5b hohe Rotwerte auf, in Fig. 5d schwächere Blauwerte und in Fig. 5c keine Grünwerte. Dieses Segment entspricht der Rückwandebene des auf der eigenen Fahrspur vorausfahrenden (Transporter-) Fahrzeugs (53).

Das gezeigte Segmentierungsergebnis wurde ohne Vorwissen über die Szene in nur drei Iterationsschritten ermittelt. Das zeigt die enorme Schnelligkeit und Leistungsfähigkeit einer Ausführungsform der Erfindung durch zeitliche Integration.

Fig. 6 zeigt eine Ermittlung der Orientierung von Ebenen in der bereits bei Fig. 5 beschriebenen Szene. Fig. 6a zeigt zur Orientierung nochmals die Umgebungssituation gemäß Fig. 5a.

Sämtliche Korrespondenzen, die einer Seitenwandebene zugeordnet zeigt Fig. 6b. Die Korrespondenzen am linken Rand wurden einer rechten Seitenwandebene zugeordnet, was zutreffend ist, da dort im Bild die rechte Seite der baulichen Begrenzung (52) zur Gegenfahrbahn befindlich ist. Die Korrespondenzen in der rechten Bildhälfte wurden linken Seitenwandebenen zugeordnet, was ebenfalls zutreffend ist, da dort die „linke Seite“ der Fahrbahnrandbebauung bzw. -bepflanzung im Bild befindlich ist. Fig. 6c zeigt, welche Korrespondenzen einer Bodenebene zugeordnet werden, was zutreffend ist, da dort im Bild die Oberfläche der Fahrbahn (51) zu sehen ist.

Fig. 6d zeigt, welche Korrespondenzen einer Rückwandebene zugeordnet werden. Das ist weitestgehend zutreffend. Aus dieser Ermittlung alleine können unterschiedliche Rückwandebenen noch nicht hinreichend unterschieden werden, z.B. die von dem auf derselben Fahrspur vorausfahrenden Lieferwagen (53) von den im Bild darüber angeordneten Schildern der Schilderbrücke (54). Dieser Darstellung können aber bereits wichtige Hinweise entnommen werden, wo erhabene Objekte in der Umgebung des Fahrzeugs auftreten.

Wie in Fig. 7 veranschaulicht wird, kann zur Erkennung dynamischer Objekte die inverse TTC (t_x , t_y , t_z) verwendet werden. Fig. 7a zeigt wiederum das Bild der Fahrzeugsituation (identisch mit Fig. 6a). Das auf der eigenen Fahrspur vorausfahrende Fahrzeug (73) ist ein Lieferwagen. Auf der linken Spur fahren zwei Fahrzeuge (71 und 72) und auf der rechten Spur zwei weitere Fahrzeuge (74 und 75).

Fig. 7b zeigt Korrespondenzen, die wiederum der Bodenebene entsprechen (im Original violett) und als einzige einen Rotanteil aufweisen.

Fig. 7c zeigt Korrespondenzen, die bewegten Objekten zugeordnet werden. Diese sind im Original grün, wenn sie sich vom eigenen Fahrzeug entfernen (also schneller fahren) bzw. türkis, wenn sie langsamer fahren.

Fig. 7d zeigt Korrespondenzen mit Blauanteil, also solche, die der Bodenebene entsprechen (vgl. Fig. 7b), bewegte Objekte, die sich dem eigenen Fahrzeug nähern (vgl. Fig. 7c) und solche, die statischen erhabenen Objekten entsprechen, diese sind nur in Fig. 7d dargestellt, wie z.B. Waldbereiche links und rechts der Autobahn und die Schilderbrücken. Aus Fig. 7c und 7d gemeinsam, ist zu erkennen, dass sich das Fahrzeug auf der eigenen Fahrspur (73) nähert. Dasselbe gilt für das vordere Fahrzeug auf der rechten Fahrspur (75). Dagegen entfernen sich die übrigen Fahrzeuge (71, 72 und 74).

Der Bereich, der im Bild dem Himmel entspricht, führt mangels Struktur im Bild zu keinen Korrespondenzen (weiß in Fig. 7b bis 7d).

Wird die Eigenrotation in den Korrespondenzen vor der Berechnung der Homographie berücksichtigt, bzw. wird die Eigenrotation in der Rotationsmatrix R berücksichtigt, lassen sich überholende Fahrzeuge aufgrund Ihrer negativen t_z Komponente erkennen bzw. ausscherende oder in einer Kurve fahrende Fahrzeuge durch eine laterale t_x Komponente ungleich Null erkennen. Werden die dynamischen Segmente über ihre Homographien prädiziert (siehe

unten „Verdichtung des optischen Flusses basierend auf Homographien“), kann über die Zeit eine dynamische Karte aufgebaut werden.

5 Betrachtet man Gleichung 3, erkennt man, dass Segmente mit einer inversen TTC gleich Null die Rotationsmatrix beschreiben und man kann sie durch Berechnung einer Homographie mit vollen Freiheitsgrad (Gleichung 2) aus Segmenten mit t/d gleich Null bestimmen. Geht man davon aus, dass sich die translatorische
10 Komponenten in der Nähe des Epipols nicht bemerkbar machen, kann man die Pitch- und Gierrate auch bestimmen, indem die Koordinaten des Epipols (x_e , y_e) durch die Homographie statischer Segmente prädiziert werden und der $\text{atan}((x_{e0} - x_{e1})/f)$ bzw. $\text{atan}((y_{e0} - y_{e1})/f)$ mit der auf einen Pixel bezogenen Brennweite f berechnet wird.

15

Wird für jedes Cluster eine Homographie mit allen Freiheitsgraden berechnet, können diese auch zur Rekonstruktion der 3D-Umgebung verwendet werden, indem anstelle der gemessenen Position x_{t0} , die prädizierte Position $H * x_{t1}$ zur Triangulation verwendet wird. Dies reduziert nicht nur den Einfluss von Messfehlern, sondern ermöglicht es auch Objekte nahe des Epipols zu rekonstruieren.

Im Folgenden wird ein Ausführungsbeispiel zur Verdichtung des optischen Flusses basierend auf Homographien beschrieben.

Ist die Segmentierung zum Zeitpunkt $t-1$ bekannt, kann sie sowohl
20 zur Prädiktion der Objekte als auch zur Generierung eines dichten Flussfeldes verwendet werden. Signaturbasierte Flussverfahren erzeugen Signaturen und versuchen diese in aufeinanderfolgenden Frames eindeutig zuzuordnen. Meist werden die Signaturen aus einem Patch (Bildausschnitt bzw. Bildbereich) definierter Größe
25 berechnet. Verändern sich jedoch Größe und Form eines Patches, ist eine Korrespondenzfindung mit einem festen Template

(Vorlage, Muster, gemeint ist z.B. ein Bildausschnitt eines Bildes der Bildfolge, der einem Objekt entspricht - beispielsweise ein Fahrzeug-Template) nicht mehr möglich. Wenn man sich z.B. einer Back-Plane annähert, verändert sich die Größe eines Patches. Oder wenn man sich über einer Ground-Plane oder parallel zu einer Side-Plane bewegt, verändern sich sowohl Größe als auch Form eines Patches, siehe Fig. 1 und 2). Ist die Segmentierung zum Zeitpunkt $t-1$ vorhanden, können die Homographien über bereits gefundenen Flussvektoren neu berechnet und dazu verwendet werden die Position und Form bereits etablierter Korrespondenzen von $t-1$ auf $t-0$ zu präzisieren.

Alternativ lässt sich der aktuelle Frame zum Zeitpunkt $t-0$ auf den Zeitpunkt $t-1$ transformieren, um Skalen- und Formänderungen zu kompensieren.

Fig. 8 veranschaulicht ein solches Vorgehen.

Fig. 8a zeigt ein Bild einer anderen Fahrsituation, das von der Fahrzeugkamera zu einem Zeitpunkt $t-1$ aufgenommen wurde. Zu sehen ist eine Autobahn mit drei Fahrspuren je Fahrtrichtung. Links der eigenen dreispurigen Fahrbahn befindet sich eine Leitplanke (81) als erhabene Begrenzung zur Gegenfahrbahn. Rechts der Fahrbahn befindet sich eine Lärmschutzwand (82).

Fig. 8b zeigt ein Bild, das zum darauffolgenden Zeitpunkt t aufgenommen wurde und über die Homographie der Leitplanke derart transformiert („gewarpt“, englisch: to warp) wurde, dass die infolge der Bewegung des Fahrzeuges und damit der Fahrzeugkamera zwischen den beiden Aufnahmezeitpunkten auftretenden Änderungen im Bild im Bereich der Leitplanke kompensiert werden. Die Vorwärtsbewegung des eigenen Fahrzeugs führt in Fig. 8b dazu, dass der naheliegendste Teilstrich der Fahrspurmarkierung näher am eigenen Fahrzeug ist als in Fig. 8a. Die Transformation führt

zu der trapezförmigen Versetzung des Bildes, welche in Fig. 8f durch eine gestrichelte Linie veranschaulicht ist.

Fig. 8c zeigt nun korrespondierende Merkmale (85), die im Bereich
5 der Leitplanke (81, vgl. Fig. 8a) ermittelt wurden, als weiße Punkte.

Fig. 8d zeigt, wo diese korrespondierenden Merkmale im nächsten Bild zu erwarten sind (86), nachdem dieses wie zu Fig. 8b beschrieben transformiert worden ist.

10 In Fig. 8e und 8f ist dieser Sachverhalt nochmals in einer schwarz-weiß Darstellung gezeigt, wobei die korrespondierenden Merkmale (85) nun den schwarzen Punkten auf der Leitplanke (81) in der linken Bildhälfte entsprechen.

15 Zur Generierung eines dichten Flussfeldes kann also für jedes Segment das aktuelle Bild auf das vorherige Bild gewarpt werden, um bereits bestehende Korrespondenzen, die sich in ihrer Skale oder Form veränderte haben, wieder zu finden bzw. um neue
20 Korrespondenzen mittels deckungsgleicher Templates zu etablieren.

Sind in einem aktuellen Frame nicht genügend Flussvektoren zur Neuberechnung einer Homographie vorhanden, lassen sich näherungsweise die Homographie aus dem letzten Frame verwenden um die
25 Korrespondenzfindung robuster gegen Form und Skalenänderungen zu gestalten.

Folgende Ausgestaltungsformen bzw. -aspekte sind vorteilhaft und können einzeln oder in Kombination vorgesehen werden:

- 30
1. Das Bild wird in $N \times M$ Zellen unterteilt und den Punkt-Korrespondenzen einer Zelle wird eine eindeutige Zellen-ID zugewiesen. Aus den Korrespondenzen mit gleichen

IDs werden mittels RANSAC die Back-/Ground- und Side-Plane-Homographien (Gleichung 9, 10 und 10) berechnet und sowohl die Homographie mit dem geringsten Rückprojektionsfehler, als auch die zur Berechnung der Homographie verwendeten Stützstellen gespeichert. Bei RANSAC (RANdom SAmple Consensus) Verfahren wird üblicherweise bei jeder Iteration eine minimale Anzahl an zufällig ausgewählten Korrespondenzen verwendet, um eine Hypothese zu bilden. Für jedes korrespondierende Merkmal wird anschließend ein Wert berechnet, der beschreibt, ob das korrespondierende Merkmal die Hypothese unterstützt. Wenn die Hypothese eine hinreichende Unterstützung durch die korrespondierenden Merkmale erreicht, können die nicht-unterstützenden korrespondierenden Merkmale als Ausreißer verworfen werden. Andernfalls wird erneut eine minimale Anzahl an Korrespondenzen zufällig ausgewählt.

2. Für benachbarte Zellen i, j werden die Rückprojektionsfehler $\sum x_{t0}^i - H_j x_{t1}^i$ bzw. $\sum x_{t0}^j - H_i x_{t1}^j$ über die Stützstellen der benachbarten Homographie berechnet. Ist der Rückprojektionsfehler $\sum x_{t0}^i - H_j x_{t1}^i$ kleiner $\sum x_{t0}^j - H_i x_{t1}^j$ bzw. unterschreiten die Fehler eine auf die Flusslänge normierte Schwelle, werden die IDs zusammengefasst und die Homographien neu berechnet. Insbesondere können zwei benachbarte Zellen als zur selben Ebene (bzw. zum selben Segment oder zum selben Objekt) zugehörig geclustert werden, wenn der Rückprojektionsfehler $\sum_i (x_{t0}^i - H_j x_{t1}^i)$ kleiner als $\sum_i (x_{t0}^i - H_i x_{t1}^i)$ ist und wenn beide Rückprojektionsfehler $\sum_i (x_{t0}^i - H_j x_{t1}^i)$ und $\sum_i (x_{t0}^i - H_i x_{t1}^i)$ eine auf die Flusslänge normierte Schwelle unterschreiten.

3. Die Rückprojektionsfehler $x_{t0} - H_i x_{t1}$ sämtlicher Punkt-Korrespondenzen werden für die angrenzenden Segmente

berechnet und eine Punkt-Korrespondenz wird dem Segment mit geringstem Rückprojektionsfehler zugeordnet. Überschreitet der minimale Fehler eine Schwelle, werden die Korrespondenzen mit einer neuen Objekt ID versehen um auch kleinere bzw. teilverdeckte Objekte erkennen zu können.

4. Die Homographien der zum Zeitpunkt $t-1$ extrahierten Segmente werden zu Beginn eines neuen Frames ($t-0$) über die bereits gefunden Bild Korrespondenzen neu berechnet und die bereits bestehenden Segment IDs in den aktuellen Frame prädiiziert. Sind im aktuellen Frame nicht genügend Flussvektoren zur Neuberechnung einer Homographie vorhanden, lassen sich näherungsweise die Homographien aus dem letzten Frame verwenden.

5. Zur Generierung eines dichten Flussfeldes wird für jedes Segment der aktuelle Frame ($t-0$) auf den letzten Frame ($t-1$) gewarpt um bereits bestehende Korrespondenzen, die sich in ihrer Skale oder Form veränderte haben, wieder zu finden bzw. um neue Korrespondenzen zu etablieren.

6. Die Rückprojektionsfehler der Back-/Ground- und Side-Plane können zur Validierung erhabener Ziele verwendet werden, siehe Fig. 6.

7. Ist z.B. bei einer Fahrzeugstereokamera eine Disparitätskarte vorhanden, können die absoluten Geschwindigkeiten aus der inversen TTC t/d berechnet werden, da dann die absoluten Entfernungen d für einzelne Pixel in der Disparitätskarte vorliegen.

8. Wird für jedes Segment eine vollständige Homographie mit allen Freiheitsgraden berechnet, kann aus Segmenten mit

einer TTC nahe unendlich (bzw. inverse TTCs annähernd Null)
die Rotationsmatrix R bestimmt werden.

- 5 9. Die 3D-Umgebung kann aus der prädizierten Position
 (Hx_{t1}, x_{t1}) anstelle der gemessenen Position (x_{t0}, x_{t1})
 rekonstruiert werden und erlaubt es, auch Objekte am Epipol
 zu rekonstruieren.

Patentansprüche

1. Verfahren zur Detektion von Objekten aus einer Folge von Bildern einer Fahrzeugkamera umfassend die Schritte:

- a) Aufnahme einer Folge von Bildern mit der Fahrzeugkamera,
- b) Ermittlung von korrespondierenden Merkmalen in zwei aufeinander folgenden Bildern,
- d) Zuordnung ermittelter korrespondierender Merkmale in einem Bildbereich zu einer Ebene im Raum, und
- f) Bestimmung zusätzlicher korrespondierender Merkmale in dem Bildbereich unter Berücksichtigung der zugeordneten Ebene.

2. Verfahren nach Anspruch 1, umfassend den Schritt:

- c) Berechnung von Homographien für die ermittelten korrespondierenden Merkmale in einem Bildbereich, damit diese einer Ebene im Raum zugeordnet werden können.

3. Verfahren nach Anspruch 2, umfassend die Schritte:

- d2) Zuordnung der ermittelten korrespondierenden Merkmale zu einer von einer Mehrzahl von Ebenen vorgegebener Orientierung im Raum, und
- e) Zuordnung zu der Ebene im Raum, die für die ermittelten korrespondierenden Merkmale den kleinsten Rückprojektionsfehler ergibt, wobei der Rückprojektionsfehler den Unterschied an zwischen der gemessenen Korrespondenz eines Merkmals in zwei aufeinanderfolgenden Bildern und dem aus der berechneten Homographie prädizierten Korrespondenz des Merkmals angibt.

4. Verfahren nach einem der vorhergehenden Ansprüche, umfassend den Schritt:

d3) Zuordnung der ermittelten korrespondierenden Merkmalen in einem Bildbereich zu jeweils einer Bodenebene, einer Rückwandebene oder einer Seitenwandebene.

5. Verfahren nach Anspruch 4, wobei die mindestens eine Rückwandebene berechnet wird nach

$$\begin{bmatrix} x_0 c - a \\ y_0 c - b \\ \vdots \end{bmatrix} = \begin{bmatrix} -x_1 & 0 & x_1 x_0 \\ 0 & -x_1 & x_1 y_0 \\ \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix}$$

wobei a , b , c Konstanten sind, x_0 , y_0 , x_1 , y_1 Korrespondenzen in einem ersten Bild (Index 0) und einem zweiten Bild (Index 1) und t_x , t_y , t_z , die Komponenten des Vektors $\mathbf{t}/d$ sind. $\mathbf{t}$ beschreibt die Translation der Fahrzeugkamera und d die Entfernung zu einer Ebene.

6. Verfahren nach Anspruch 4 oder 5, wobei die mindestens eine Bodenebene berechnet wird nach

$$\begin{bmatrix} x_0 c - a \\ y_0 c - b \\ \vdots \end{bmatrix} = \begin{bmatrix} -y_1 & 0 & y_1 x_0 \\ 0 & -y_1 & y_1 y_0 \\ \vdots & \vdots & \vdots \end{bmatrix} * \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix}$$

7. Verfahren nach Anspruch 4, 5 oder 6, wobei die mindestens eine Seitenwandebene berechnet wird nach

$$\begin{bmatrix} x_0 c - a \\ y_0 c - b \\ \vdots \end{bmatrix} = \begin{bmatrix} -1 & 0 & x_0 \\ 0 & -1 & y_0 \\ \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix}$$

8. Verfahren nach einem der vorhergehenden Ansprüche, wobei ausgehend von einer ermittelten Aufteilung eines Bildes in unterschiedliche Bildbereiche mit zugeordneten Ebenen für ein zweites Bild, werden erste einander korrespondierende

Merkmale in einem Bildbereich im zweiten und einem nachfolgenden dritten Bild bestimmt,
aus diesen ersten korrespondierenden Merkmalen wird eine Homographie für diesen Bildbereich neu berechnet, und
die neu berechnete Homographie wird dazu verwendet, die Position und Form zusätzlicher korrespondierender Merkmale im dritten Bild zu präzisieren.

9. Verfahren nach einem der vorhergehenden Ansprüche, wobei
falls in einem nachfolgenden dritten Bild in einem Bildbereich nicht genügend erste korrespondierende Merkmale zur Neuberechnung einer Homographie bestimmbar sind,
die für den Bildbereich berechnete Homographie aus dem zweiten Bild dazu verwendet wird, die Position und Form zusätzlicher korrespondierender Merkmale des Bildbereichs im dritten Bild zu präzisieren.

10. Verfahren nach einem der vorhergehenden Ansprüche, wobei
zur für jeden Bildbereich mit einer zugeordneten Ebene ein aktuelles Bild auf ein vorheriges Bild entsprechend der berechneten Homographie gewarpt wird, um zusätzliche einander im letzten und im aktuellen Bild korrespondierende Merkmale zu bestimmen.

11. Vorrichtung zur Detektion von Objekten aus einer Folge von Bildern einer Fahrzeugkamera umfassend:

ein Kamerasteuergerät, welches dazu ausgebildet ist,
a) eine Folge von Bildern mit der Fahrzeugkamera aufzunehmen;
und eine Auswerteelektronik, welche dazu ausgebildet ist,
b) korrespondierende Merkmale in einem Bildbereich in zwei aufeinander folgenden Bildern zu ermitteln,

- d) ermittelte korrespondierende Merkmale in einem Bildbereich zu einer Ebene im Raum zuzuordnen, und
- f) zusätzliche korrespondierende Merkmale in dem Bildbereich unter Berücksichtigung der zugeordneten Ebene zu bestimmen.

1/6

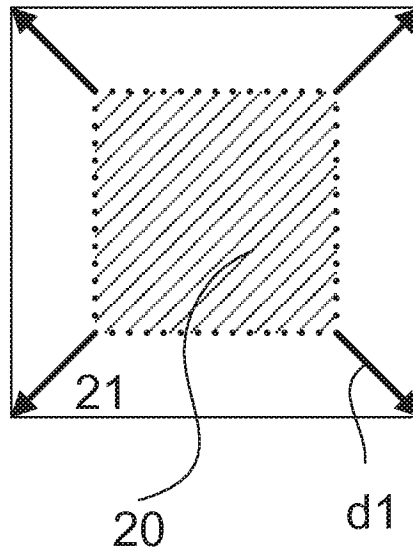


Fig. 1

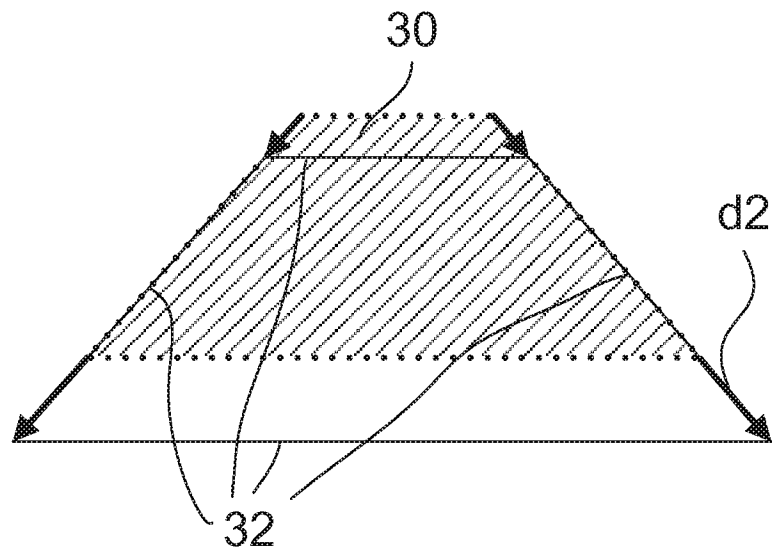


Fig. 2

2/6

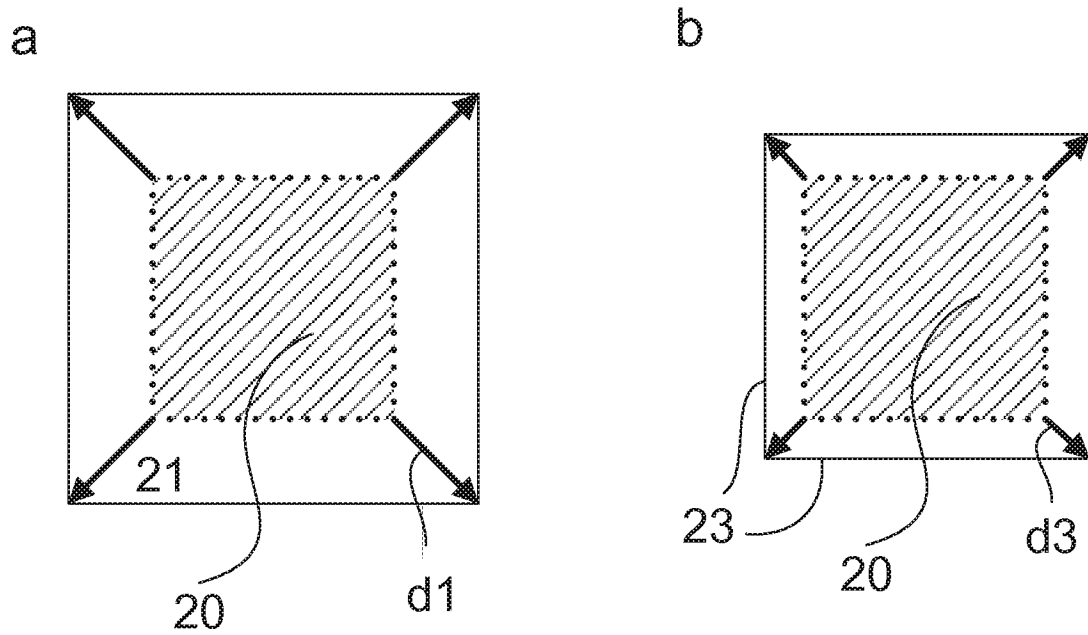


Fig. 3

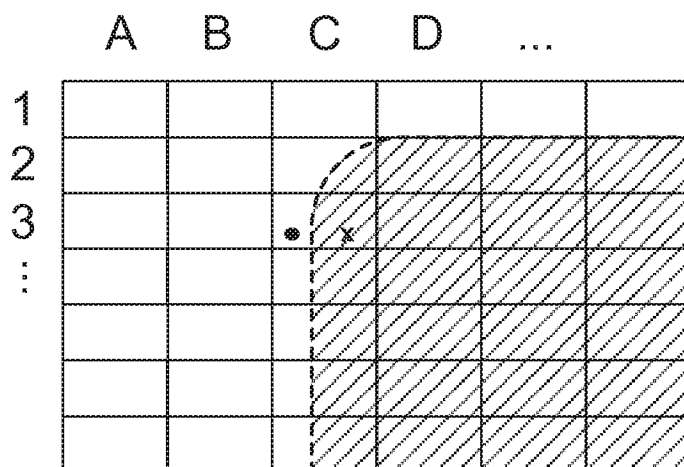


Fig. 4

3/6

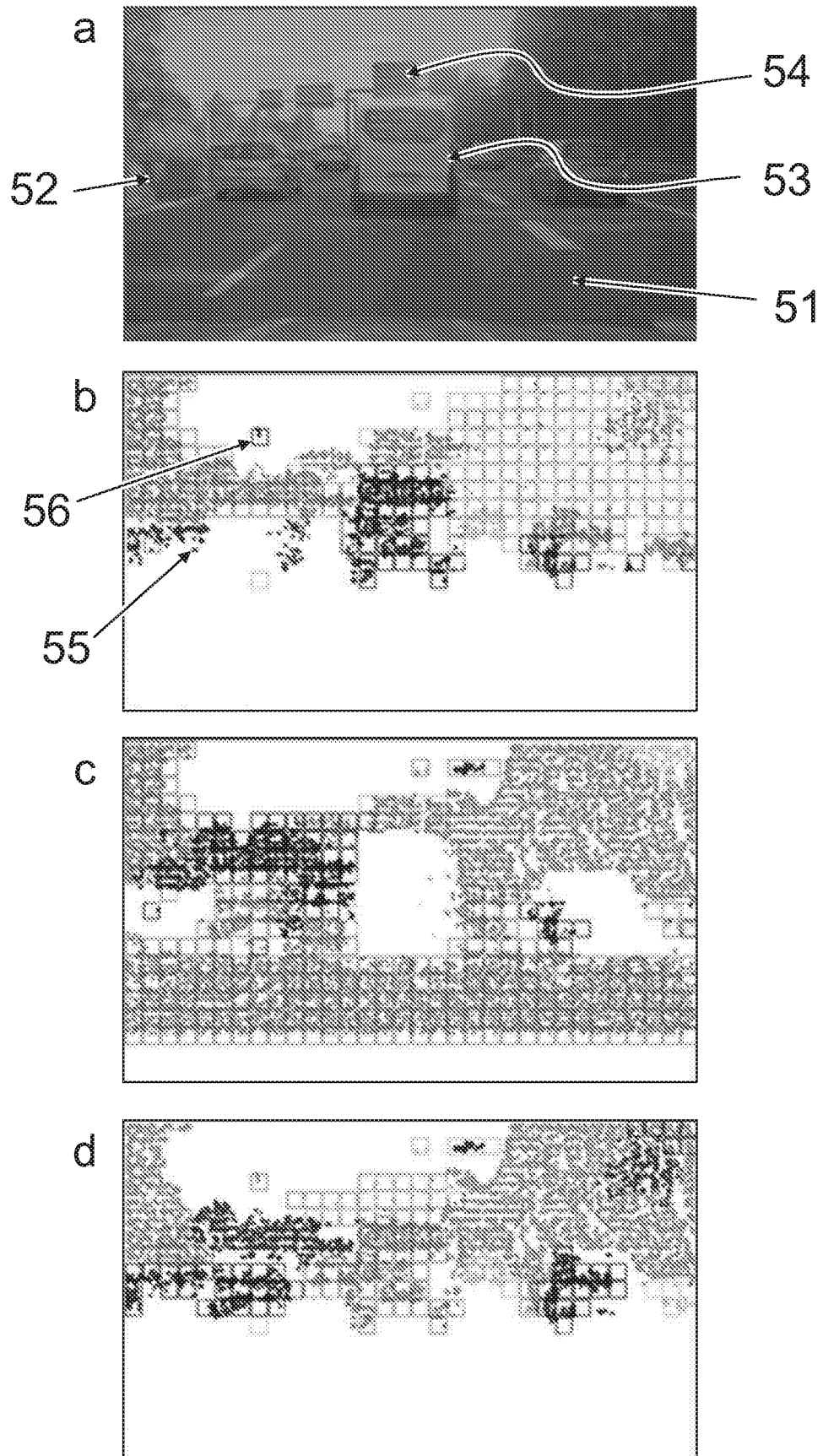
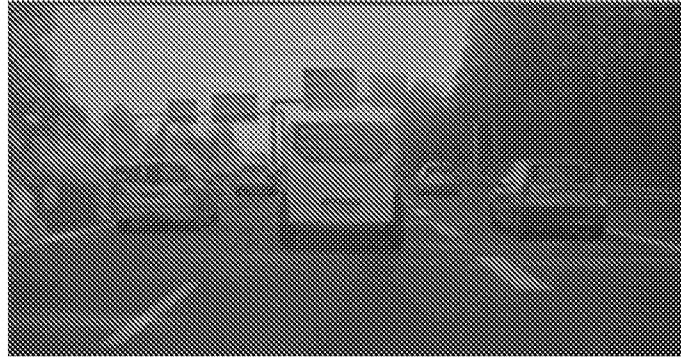


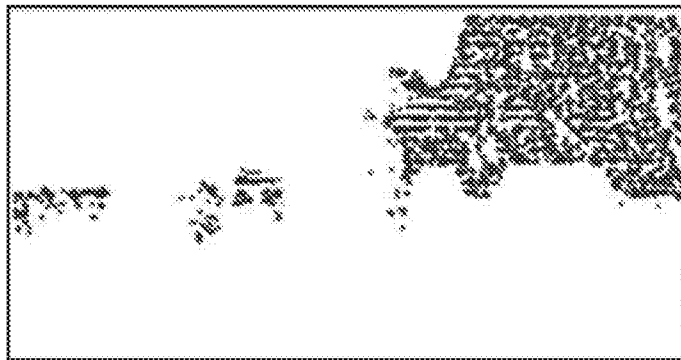
Fig. 5

4/6

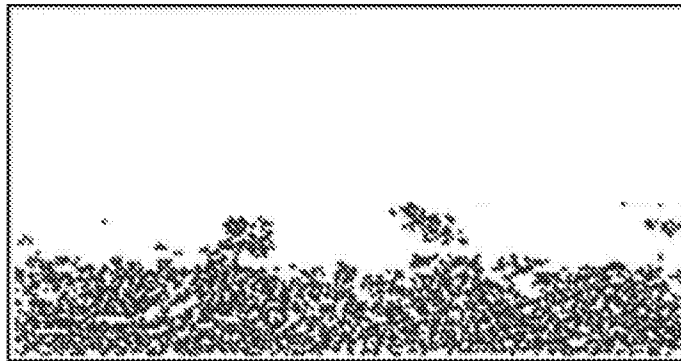
a



b



c



d

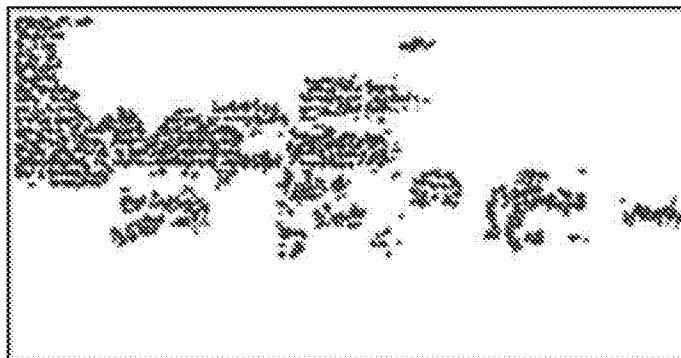


Fig. 6

5/6

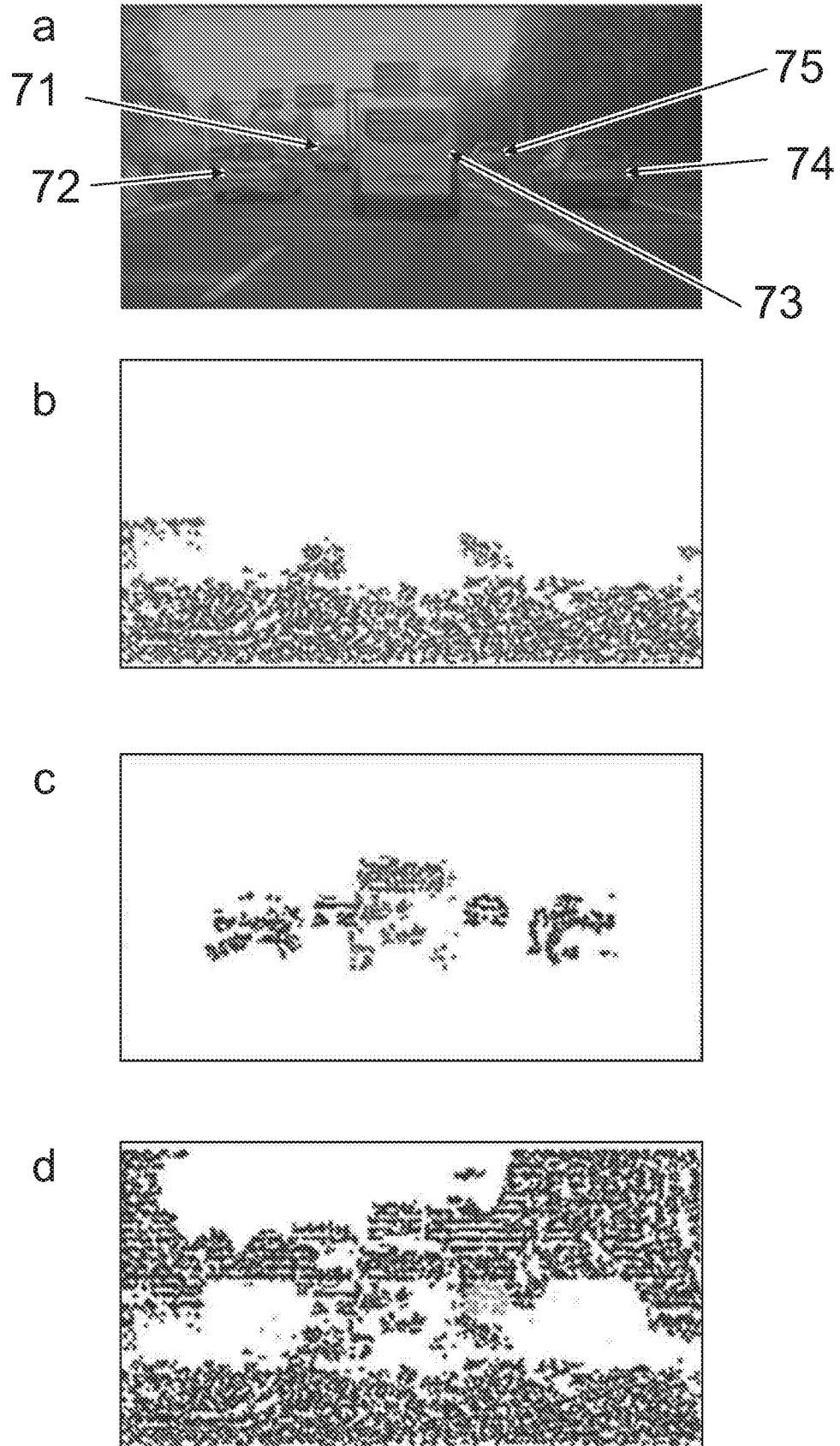


Fig. 7

6/6

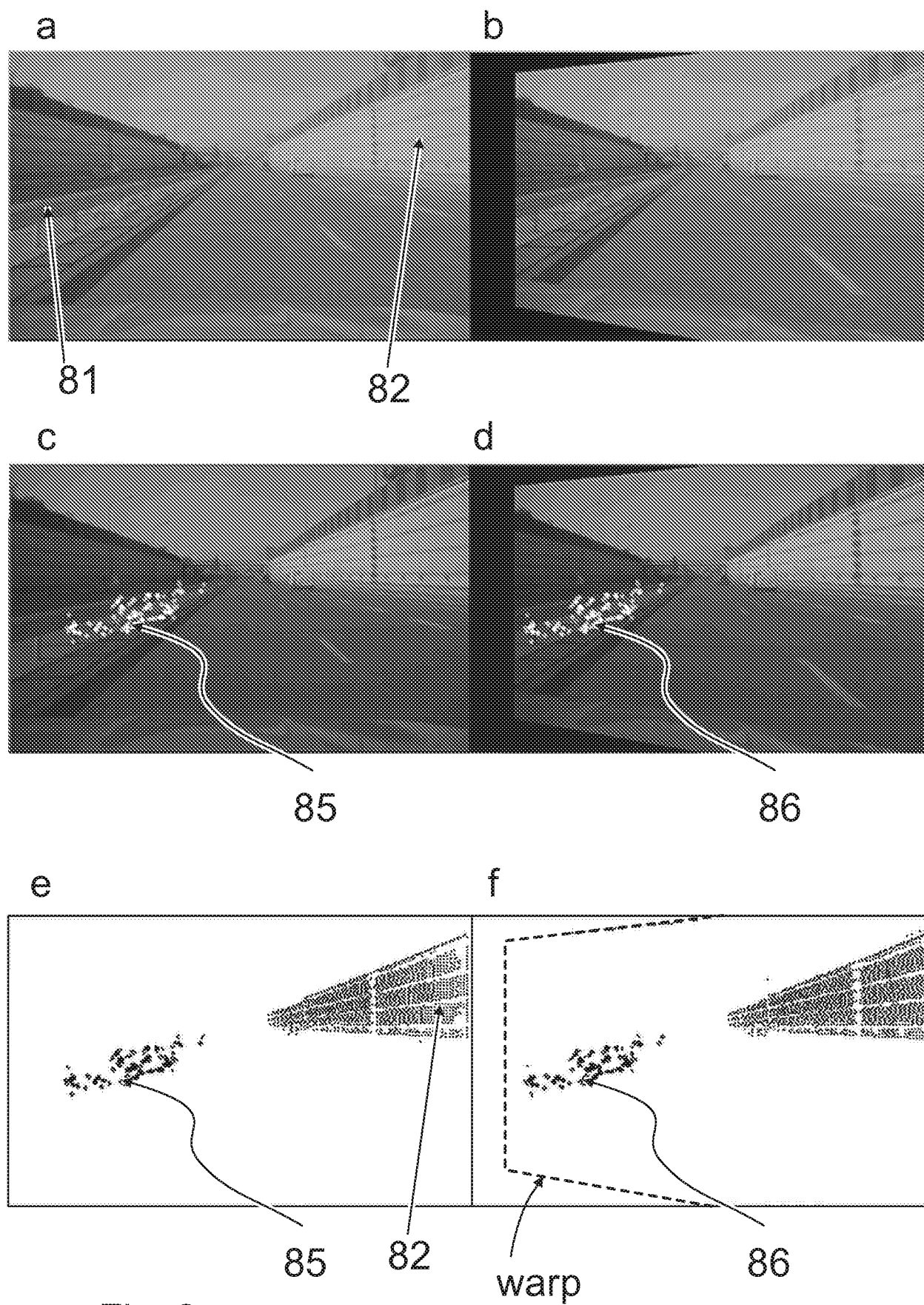


Fig. 8

INTERNATIONAL SEARCH REPORT

International application No

PCT/DE2017/200100

A. CLASSIFICATION OF SUBJECT MATTER

INV. G06K9/00

ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-Internal

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	M.I.A. LOURAKIS ET AL: "Detecting Planes In An Uncalibrated Image Pair", PROCEEDINGS OF THE 13TH BRITISH MACHINE VISION CONFERENCE : 2ND - 5TH SEPTEMBER 2002, CARDIFF UNIVERSITY, 1 January 2002 (2002-01-01), pages 57.1-57.10, XP055439691, Cardiff DOI: 10.5244/C.16.57 ISBN: 978-1-901725-19-3 Sektion 4, 5 Zeilen 6-27; figure 2 ----- -/-	1-11



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

17 January 2018

Date of mailing of the international search report

29/01/2018

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Authorized officer

Meurisse, Wim

INTERNATIONAL SEARCH REPORT

International application No
PCT/DE2017/200100

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	PRATEEK SINGHAL ET AL: "Top Down Approach to Detect Multiple Planes from Pair of Images", PROCEEDINGS OF THE 2014 INDIAN CONFERENCE ON COMPUTER VISION GRAPHICS AND IMAGE PROCESSING, ICVGIP '14, 1 January 2014 (2014-01-01), pages 1-8, XP055439700, New York, New York, USA DOI: 10.1145/2683483.2683536 ISBN: 978-1-4503-3061-9 Sektion I, II, III A-E -----	1-11
A	US 2015/332114 A1 (SPRINGER OFER [IL]) 19 November 2015 (2015-11-19) paragraphs [0131] - [0132], [0146] - [0164]; figures 9-11 -----	1-11
X	Olaf Kähler ET AL: "DETECTING COPLANAR FEATURE POINTS IN HANDHELD IMAGE SEQUENCES", 1 January 2007 (2007-01-01), XP055439695, Retrieved from the Internet: URL: http://www.robots.ox.ac.uk/~olaf/bib/kahler07detecting.pdf [retrieved on 2018-01-10] Sektion 2-4 -----	1-11
A	EP 2 993 654 A1 (MOBILEYE VISION TECHNOLOGIES LTD [IL]) 9 March 2016 (2016-03-09) cited in the application abstract paragraphs [0024], [0032] - [0033] -----	1-11
A	US 2014/161323 A1 (LIVYATAN HAREL [IL] ET AL) 12 June 2014 (2014-06-12) cited in the application paragraphs [0010] - [0013], [0036] - [0039], [0060] - [0072], [0092] -----	1-11

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/DE2017/200100

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2015332114 A1	19-11-2015	CN 106663193 A	10-05-2017
		EP 3143607 A1	22-03-2017
		US 2015332114 A1	19-11-2015
		US 2016371550 A1	22-12-2016
		WO 2015177648 A1	26-11-2015

EP 2993654 A1	09-03-2016	EP 2463843 A2	13-06-2012
		EP 2993654 A1	09-03-2016
		US 2012140076 A1	07-06-2012

US 2014161323 A1	12-06-2014	NONE	

A. KLASSIFIZIERUNG DES ANMELDUNGSGEGENSTANDES
INV. G06K9/00
ADD.

Nach der Internationalen Patentklassifikation (IPC) oder nach der nationalen Klassifikation und der IPC

B. RECHERCHIERTE GEBIETE

Recherchierter Mindestprüfstoff (Klassifikationssystem und Klassifikationssymbole)
G06K

Recherchierte, aber nicht zum Mindestprüfstoff gehörende Veröffentlichungen, soweit diese unter die recherchierten Gebiete fallen

Während der internationalen Recherche konsultierte elektronische Datenbank (Name der Datenbank und evtl. verwendete Suchbegriffe)

EPO-Internal

C. ALS WESENTLICH ANGESEHENE UNTERLAGEN

Kategorie*	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile	Betr. Anspruch Nr.
X	M.I.A. LOURAKIS ET AL: "Detecting Planes In An Uncalibrated Image Pair", PROCEEDINGS OF THE 13TH BRITISH MACHINE VISION CONFERENCE : 2ND - 5TH SEPTEMBER 2002, CARDIFF UNIVERSITY, 1. Januar 2002 (2002-01-01), Seiten 57.1-57.10, XP055439691, Cardiff DOI: 10.5244/C.16.57 ISBN: 978-1-901725-19-3 Sektion 4, 5 Zeilen 6-27; Abbildung 2 ----- -/-	1-11



Weitere Veröffentlichungen sind der Fortsetzung von Feld C zu entnehmen



Siehe Anhang Patentfamilie

* Besondere Kategorien von angegebenen Veröffentlichungen :

"A" Veröffentlichung, die den allgemeinen Stand der Technik definiert, aber nicht als besonders bedeutsam anzusehen ist

"E" frühere Anmeldung oder Patent, die bzw. das jedoch erst am oder nach dem internationalen Anmeldedatum veröffentlicht worden ist

"L" Veröffentlichung, die geeignet ist, einen Prioritätsanspruch zweifelhaft erscheinen zu lassen, oder durch die das Veröffentlichungsdatum einer anderen im Recherchenbericht genannten Veröffentlichung belegt werden soll oder die aus einem anderen besonderen Grund angegeben ist (wie ausgeführt)

"O" Veröffentlichung, die sich auf eine mündliche Offenbarung, eine Benutzung, eine Ausstellung oder andere Maßnahmen bezieht

"P" Veröffentlichung, die vor dem internationalen Anmeldedatum, aber nach dem beanspruchten Prioritätsdatum veröffentlicht worden ist

"T" Spätere Veröffentlichung, die nach dem internationalen Anmeldedatum oder dem Prioritätsdatum veröffentlicht worden ist und mit der Anmeldung nicht kollidiert, sondern nur zum Verständnis des der Erfindung zugrundeliegenden Prinzips oder der ihr zugrundeliegenden Theorie angegeben ist

"X" Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann allein aufgrund dieser Veröffentlichung nicht als neu oder auf erfinderischer Tätigkeit beruhend betrachtet werden

"Y" Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann nicht als auf erfinderischer Tätigkeit beruhend betrachtet werden, wenn die Veröffentlichung mit einer oder mehreren Veröffentlichungen dieser Kategorie in Verbindung gebracht wird und diese Verbindung für einen Fachmann naheliegend ist

"&" Veröffentlichung, die Mitglied derselben Patentfamilie ist

Datum des Abschlusses der internationalen Recherche

17. Januar 2018

Absendedatum des internationalen Recherchenberichts

29/01/2018

Name und Postanschrift der Internationalen Recherchenbehörde

Europäisches Patentamt, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Bevollmächtigter Bediensteter

Meurisse, Wim

C. (Fortsetzung) ALS WESENTLICH ANGESEHENE UNTERLAGEN		
Kategorie*	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile	Betr. Anspruch Nr.
X	PRATEEK SINGHAL ET AL: "Top Down Approach to Detect Multiple Planes from Pair of Images", PROCEEDINGS OF THE 2014 INDIAN CONFERENCE ON COMPUTER VISION GRAPHICS AND IMAGE PROCESSING, ICVGIP '14, 1. Januar 2014 (2014-01-01), Seiten 1-8, XP055439700, New York, New York, USA DOI: 10.1145/2683483.2683536 ISBN: 978-1-4503-3061-9 Sektion I, II, III A-E	1-11
A	US 2015/332114 A1 (SPRINGER OFER [IL]) 19. November 2015 (2015-11-19) Absätze [0131] - [0132], [0146] - [0164]; Abbildungen 9-11	1-11
X	Olaf Kähler ET AL: "DETECTING COPLANAR FEATURE POINTS IN HANDHELD IMAGE SEQUENCES", 1. Januar 2007 (2007-01-01), XP055439695, Gefunden im Internet: URL: http://www.robots.ox.ac.uk/~olaf/bib/kahler07detecting.pdf [gefunden am 2018-01-10] Sektion 2-4	1-11
A	EP 2 993 654 A1 (MOBILEYE VISION TECHNOLOGIES LTD [IL]) 9. März 2016 (2016-03-09) in der Anmeldung erwähnt Zusammenfassung Absätze [0024], [0032] - [0033]	1-11
A	US 2014/161323 A1 (LIVYATAN HAREL [IL] ET AL) 12. Juni 2014 (2014-06-12) in der Anmeldung erwähnt Absätze [0010] - [0013], [0036] - [0039], [0060] - [0072], [0092]	1-11

INTERNATIONALER RECHERCHENBERICHT

Angaben zu Veröffentlichungen, die zur selben Patentfamilie gehören

Internationales Aktenzeichen

PCT/DE2017/200100

Im Recherchenbericht angeführtes Patentdokument		Datum der Veröffentlichung	Mitglied(er) der Patentfamilie		Datum der Veröffentlichung
US 2015332114	A1	19-11-2015	CN	106663193 A	10-05-2017
			EP	3143607 A1	22-03-2017
			US	2015332114 A1	19-11-2015
			US	2016371550 A1	22-12-2016
			WO	2015177648 A1	26-11-2015

EP 2993654	A1	09-03-2016	EP	2463843 A2	13-06-2012
			EP	2993654 A1	09-03-2016
			US	2012140076 A1	07-06-2012

US 2014161323	A1	12-06-2014	KEINE		
