



(10) **DE 10 2006 055 641 B4** 2013.01.31

(12) **Patentschrift**

(21) Aktenzeichen: **10 2006 055 641.0**
(22) Anmeldetag: **22.11.2006**
(43) Offenlegungstag: **29.05.2008**
(45) Veröffentlichungstag
der Patenterteilung: **31.01.2013**

(51) Int Cl.: **H04N 13/00 (2006.01)**

Innerhalb von drei Monaten nach Veröffentlichung der Patenterteilung kann nach § 59 Patentgesetz gegen das Patent Einspruch erhoben werden. Der Einspruch ist schriftlich zu erklären und zu begründen. Innerhalb der Einspruchsfrist ist eine Einspruchsgebühr in Höhe von 200 Euro zu entrichten (§ 6 Patentkostengesetz in Verbindung mit der Anlage zu § 2 Abs. 1 Patentkostengesetz).

(73) Patentinhaber:
VisuMotion GmbH, 07745, Jena, DE

(74) Vertreter:
**Patentanwälte Oehmke und Kollegen, 07743,
Jena, DE**

(72) Erfinder:
**Schmidt, Alexander, 07745, Jena, DE; Meichsner,
Jens, 07743, Jena, DE; Billert, Ronny, 99096,
Erfurt, DE**

(56) Für die Beurteilung der Patentfähigkeit in Betracht
gezogene Druckschriften:

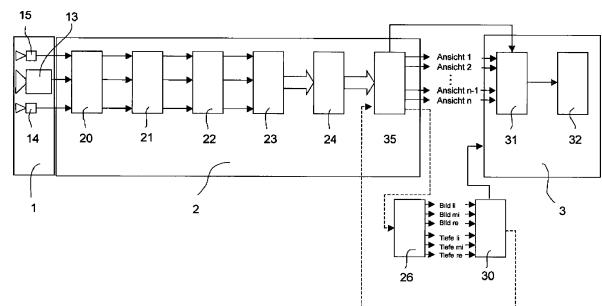
DE	10 2004 061998	A1
US	62 71 876	B1
US	2003/02 31 179	A1
EP	14 18 766	A2
WO	00/13 142	A1
KR10	2002 0 032 954	A

Kim, K.-T., et al., Synthesis of a high-resolution 3D-stereoscopic image pair from a high-resolution monoscopic image and a low-resolution depth map. In: Proceedings of the SPIE, Bd. 3295, Januar 1998, S. 76-86, ISSN: 0277-786X

(54) Bezeichnung: **Anordnung und Verfahren zur Aufnahme und Wiedergabe von Bildern einer Szene und/oder eines Objektes**

(57) Hauptanspruch: Verfahren zur Aufnahme und Wiedergabe von Bildern einer Szene und/oder eines Objektes, die folgenden Schritte umfassend:

- Erstellung von mindestens einem n-Tupel von Bildern, mit $n > 2$, mit mindestens einer Hauptkamera (13) eines ersten Kameratyps und mindestens einer Satellitenkamera (14, 15) eines zweiten Kameratyps,
- Übernahme der Bilddaten in eine Bildkonvertierungseinrichtung (2), in der nachfolgend
- bei Verwendung von einer Hauptkamera (13) und zwei oder vier Satellitenkameras mit von der Auflösung der Hauptkamera (13) verschiedener Auflösung zunächst eine Größenanpassung, so dass sämtliche Bilder nach Größenanpassung jeweils die gleiche Auflösung aufweisen,
- eine Rektifizierung,
- eine Farbjustierung,
- eine Tiefen- oder Disparitätserkennung zur Generation weiterer Ansichten aus dem n-Tupel von Bildern und den Tiefen- bzw. Disparitätserkennungswerten durchgeführt wird, wobei mindestens eine Ansicht generiert wird, die keinem der von dem Kamerasystem (1) aufgenommenen Bildern exakt entspricht,
- anschließend eine Kombination der verschiedenen Ansichten entsprechend...



Beschreibung

[0001] Die Erfindung betrifft eine Anordnung und ein Verfahren zur Aufnahme und Wiedergabe von Bildern einer Szene und/oder eines Objektes. Sie sind insbesondere zur räumlich wahrnehmbaren Wiedergabe der aufgenommenen Bilder geeignet. Fernerhin betrifft die Erfindung ein Verfahren zur Übertragung von räumlich wahrnehmbaren Bildern.

[0002] Zum Aufnehmen von 3D-Bildinformationen gibt es momentan im Wesentlichen drei grundlegend verschiedene Verfahren und die dazugehörigen Anordnungen:

Erstens, die klassische Stereo-Kamera, bestehend aus zwei gleichartigen Kameras für jeweils ein linkes und ein rechtes Bild. Für eine hochauflösende Wiedergabe sind hier jedoch auch hochauflösende Kamerasysteme notwendig. Für Mehrkanal-Systeme wird eine Interpolation der Zwischenansichten notwendig. Dabei werden Artefakte vor allem in den mittleren Ansichten sichtbar.

[0003] Zweitens, die Verwendung eines Multiview-Kamera-Systems. Der Vorteil hier gegenüber der Stereo-Kamera ist die korrekte Bildwiedergabe für Mehrkanal-Systeme. Es werden insbesondere keine Interpolationen notwendig. Nachteilig ist jedoch der hohe Aufwand, der betrieben werden muss, um eine exakte Ausrichtung der Kameras zueinander zu realisieren. Der erhöhte Kostenfaktor durch die Verwendung mehrerer Kameras, die darüber hinaus weitere Probleme nach sich ziehen, wie unterschiedliche Weiß/Farb/Geometriewerte, die wieder entsprechend ausgeglichen werden müssen, ist nachteilig. Als nachteilig ist es ebenfalls anzusehen, dass bei diesem Verfahren eine extrem hohe Datenrate bewältigt werden muss.

[0004] Drittens, die Verwendung einer Tiefen-Kamera. Hierbei kommt eine Farbkamera gemeinsam mit einem Tiefensensor, der die zyklische Tiefeninformation der aufzunehmenden Szene registriert, zum Einsatz. Neben dem, dass ein Tiefensensor relativ teuer ist, ist nachteilig, dass diese oftmals nicht sehr exakt arbeiten und/oder kein vertretbarer Kompromiss zwischen Genauigkeit und Geschwindigkeit erreicht wird. Eine generelle Extrapolation wird notwendig, wobei insbesondere in den äußeren Ansichten Artefakte nicht auszuschließen sind und generell Verdeckungsartefakte nicht kaschiert werden können.

[0005] Mit der DE 10 2004 061 998 A1 wird für eine Stereo-Kamera die Verwendung zweier unterschiedlicher Sensoren in den beiden Teilkameras vorgeschlagen. Dabei gibt es beispielsweise die Variante der Kombination einer Farbkamera mit einer Graustufenkamera in einer Stereokamera. Nicht offenbart wird, wie die Generation von mehreren Ansichten aus

dem gewonnenen Datenmaterial für eine Verwendung in brillenfreien 3D-Displays ermöglicht wird.

[0006] Bei der US 2003/0231179 A1 handelt sich um ein virtuelles Telepräsenzsystem, bei welchem über eine Internetstruktur verschiedene Daten (u. a. Ansichten und Voxel- oder Tiefeninformation) versandt werden, die in aller Regel beim Empfänger in stets eine gewünschte, unter Umständen neue Ansicht konvertiert werden. Dabei kommt keine Auflösungsreduktion zur Übertragung in Frage, lediglich eine verlustbehaftete Komprimierung. Die Verwendung der am Ende erstellten Ansicht zur räumlichen Darstellung ist nicht möglich, weil mit nur einer Ansicht keine stereoskopische Wiedergabe möglich ist.

[0007] In der Schrift Kim K.-T. et al. "Synthesis of a high-resolution 3D-stereoscopic image pair from a high-resolution monoscopic image and a low-resolution depth map", SPIE, Band 3295, Januar 1998, Seiten 76–86, wird ein Verfahren zum Erstellen von zwei oder mehr Bildern vorgestellt, wobei die Ausgangsbasis drei (zunächst gleiche) Bilder sind. Zunächst werden die äußeren Bilder in der Auflösung verringert und monochrom konvertiert. Dann wird eine Disparitätskarte erzeugt, und zwar aus dem linken, rechten Bild und dem mittleren Bild. Bei der Ansichtenrekonstruktion (Interpolation bzw. Synthesizing Left/Right Images) wird jedoch nur eine Tiefenkarte verwendet, und auch nur Bildinformation der mittleren Kamera, nicht aber die der äußeren Kameras. Bezüglich Okklusionen (also Verdeckungen) wird hier offenbart, lediglich die im mittleren hochauflösenden Bild vorhandenen Bildanteile für Interpolationen zu nutzen. Es wird ferner vorgeschlagen, Daten über Okklusionen separat zu ermitteln und auf der Empfängerseite zu nutzen. Nachteilig ist hierbei, dass die so erhaltene Okklusionsinformation rudimentär bleibt, da nicht Bildinformation aus dem linken und rechten Bild einfließen.

[0008] Der Erfindung liegt die Aufgabe zugrunde, eine neue Möglichkeit aufzuzeigen, durch die es mit möglichst geringem Aufwand gelingt, Aufnahmen von realen Szenen und/oder Objekten zu erstellen, um sie nachfolgend dreidimensional in zwei oder mehr Ansichten räumlich wahrnehmbar wiederzugeben. Weiterhin ist es Aufgabe der Erfindung, ein geeignetes Verfahren zur Übertragung von räumlich wahrnehmbaren Bildern aufzuzeigen.

[0009] Erfindungsgemäß wird die Aufgabe bei einer Anordnung zur Aufnahme und Wiedergabe von Bildern einer Szene und/oder eines Objektes dadurch gelöst, dass mindestens eine Hauptkamera eines ersten Kameratyps zur Aufnahme von Bildern, mindestens eine Satellitenkamera eines zweiten Kameratyps zur Aufnahme von Bildern, wobei der erste und der zweite Kameratyp sich in mindestens einem Parameter, bevorzugt der Bildauflösung, unter-

scheiden und eine den Kameras nachgeordnete Bildkonvertierungseinrichtung und ein 3D-Bildwiedergabegerät vorhanden sind. Haupt- und Satellitenkamera unterscheiden sich im allgemeinen, jedoch nicht zwingend, durch ihre Qualität. Die Hauptkamera ist dabei meist eine sogenannte High-Quality-Kamera, wobei Satellitenkameras zum Einsatz kommen können, die sich durch geringere Qualität auszeichnen (z. B. Industriekameras) und damit unter anderen Parametern meist – aber nicht zwingend – auch eine niedrigere Auflösung aufweisen. Der Vorteil der Erfindung besteht im Wesentlichen darin, dass neben dem klassischen Einsatz eines Stereo-Kamerasystems, hier bestehend aus im Wesentlichen zwei identischen hochauflösenden Kameras, ein Drei-Kamerasystem, bestehend aus einer zentralen High-Quality-Kamera und zwei zusätzlichen Kameras mit niedrigerer Auflösung, die links- bzw. rechts der Hauptkamera angeordnet sind, verwendet wird. Die Kameras sind dabei in den üblichen Grenzen bzgl. Abstand und Ausrichtung (parallele Ausrichtung oder auf einen Fokus) variierbar. Die Verwendung weiterer Satellitenkameras kann von Vorteil sein, da insbesondere bei der nachfolgenden Aufbereitung der Bilddaten Fehlinterpretationen weiter reduziert werden können. Alle Kameras können sowohl parallel als auch auf einen Punkt ausgerichtet sein. Ebenfalls möglich ist es, dass nicht alle auf einen Punkt ausgerichtet sind (Konvergenzwinkel). Die optischen Achsen der Kameras können ebenfalls sowohl in einer als auch in unterschiedlichen Ebenen liegen, wobei die Objektivmittelpunkte in einer Linie oder im Dreieck (bevorzugt gleichschenkelig oder gleichseitig) angeordnet sein sollten. Eine Synchronisation der Kameras bzgl. Zoom, Blende, Fokus etc. sollte ebenso wie bzgl. der einzelnen Frames erfolgen. Die Kameras können fest oder beweglich zueinander angeordnet werden, wobei eine automatische Einstellung des Basisabstandes der Kameras ebenso wie die Konvergenzwinkel durchführbar ist.

[0010] Von Vorteil können Adaptersysteme sein, die ein erleichtertes Anbringen, insbesondere der Satellitenkameras, an die Hauptkamera ermöglichen. Damit können gewöhnliche Kameras nachträglich als 3D-Kamera umgerüstet werden. Es ist aber ebenso denkbar, bestehende Stereokamerasysteme durch eine zusätzliche Hauptkamera zu erfindungsgemäßen 3D-Kameras umzurüsten.

[0011] Fernerhin können im Strahlengang zusätzliche optische Elemente, z. B. teildurchlässige Spiegel, vorhanden sein.

[0012] Für besondere Ausgestaltungen kann es außerdem sinnvoll sein, die Satellitenkameras als Schwarz-Weiß-Kameras auszubilden und bevorzugt den von ihnen aufgenommenen Bildern hernach automatisch einen Farbwert zuzuweisen.

[0013] Die Aufgabe wird auch durch ein Verfahren zur Aufnahme und Wiedergabe von Bildern einer Szene und/oder eines Objektes gelöst, welches die folgenden Schritte umfasst:

- Erstellung von mindestens einem n-Tupel von Bildern, mit $n > 1$,
- Übernahme der Bilddaten in eine Bildkonvertierungseinrichtung, in der nachfolgend eine Rektifizierung, eine Farbjustierung, der Aufbau einer Stackstruktur und eine Projektion der Stackstruktur auf eine gewünschte Ansicht durchgeführt wird, wobei mindestens eine Ansicht generiert wird, die keinem der von dem Kamerasystem aufgenommenen Bildern exakt entspricht,
- anschließend eine Kombination der verschiedenen Ansichten entsprechend der Zuordnungsvorschrift des 3D-Displays eines 3D-Bildwiedergabegerätes erstellt wird, und
- abschließend die Darstellung des kombinierten 3D-Bildes auf dem 3D-Display durchgeführt wird.

[0014] Der Aufbau einer Stackstruktur kann auch durch sonstige anwendbare Tiefen- oder Disparitätserkennungsalgorithmen ersetzt werden, wobei dann die erkannten Tiefen- bzw. Disparitätswerte für die Erstellung von gewünschten Ansichten eingesetzt werden. Eine Stackstruktur kann allgemein einer Schichtstruktur von graphischen Elementen in unterschiedlichen Ebenen entsprechen.

[0015] Bei der Verwendung eines 3D-Kamerasystems, bestehend aus Kameras unterschiedlicher Kameratypen mit unterschiedlichen Bildauflösungen, macht es sich erforderlich, dass nach Übernahme der Bilddaten in die Bildkonvertierungseinrichtung zunächst eine Größenanpassung vorgenommen wird. Im Ergebnis dessen liegen Bilder mit jeweils der gleichen Auflösung vor. Daran anschließend erfolgt die Rektifizierung, d. h. eine geometrische Entzerrung der Kamerabilder (Ausgleich von möglichen Linsenverzerrungen, Kameraverdrehungen, Zoomdifferenzen etc.) wird vorgenommen. Die Größenanpassung kann auch im Rahmen des Rektifizierungsprozesses erfolgen. Unmittelbar daran anschließend erfolgt eine Farbjustierung. Es werden insbesondere die Farb-/Helligkeitswerte der Kamerabilder angeglichen, so dass ein einheitliches oder zumindest ein vergleichbares Niveau zu verzeichnen ist. Für die nunmehr vorliegenden Bilddaten wird die Stackstruktur aufgebaut. Dabei werden die Eingangsbilder zeilenweise miteinander verglichen, indem sie im ersten Schritt übereinander gelegt werden. Der Zeilenvergleich kann unter Umständen auch schräg erfolgen, dies wird dann günstig sein, wenn die Kameras nicht horizontal zueinander angeordnet sind. Bei übereinanderliegenden Pixeln mit gleichen Farbwerten wird dieser gespeichert, wenn jedoch übereinanderliegende Pixel unterschiedliche Farbwerte aufweisen, dann wird kein Wert gespeichert. Danach werden die Zeilen gegeneinander in entgegengesetzte

Richtungen in definierten Schritten (z. B. um $\frac{1}{4}$ oder $\frac{1}{2}$ Pixel) verschoben, wobei nach jedem Schritt das Ergebnis des Vergleiches wieder gespeichert wird. Im Ergebnis liegt die dreidimensionale Stackstruktur mit den Koordinaten X, Y und Z vor, wobei X und Y den Pixelkoordinaten des Eingangsbildes entspricht, während Z den Grad der Verschiebung der Ansichten zueinander darstellt. Bei der Verwendung von zwei Kameras werden also jeweils zwei Zeilen verglichen und bei der Verwendung von drei Kameras jeweils drei Zeilen verglichen und gegeneinander verschoben. Möglich ist es, bei der Verwendung von mehreren, z. B. drei, Kameras dennoch jeweils nur zwei Zeilen miteinander zu kombinieren, wobei dann ein Abgleich der Vergleiche nochmals durchgeführt werden muss. Bei einem Vergleich von drei oder mehr Zeilen gibt es weit weniger Mehrdeutungen gegenüber dem Vergleich zwischen den zwei Zeilen lediglich zweier Eingangsbilder. Bei der sich anschließenden Optimierung der Stackstruktur geht es im Wesentlichen darum, bei mehrdeutigen Abbildungen von Bildelementen im Stack die höchst unwahrscheinlichen Kombinationen zu löschen. Dies trägt darüber hinaus zur Datenreduktion bei. Eine weitere Reduktion erhält man, indem eine Höhenprofilinie aus den verbliebenen Elementen erstellt wird, um eine eindeutige Abbildung der Farbwerte in eine diskrete Tiefebene (Z Koordinate) zu erreichen. Normalerweise schließt sich nun die Projektion der Stackstruktur auf die gewünschten Ansichten an. Dabei sollten mindestens zwei Ansichten erstellt werden, wobei eine davon immer noch einem Eingangsbild entsprechen könnte. Dies erfolgt jedoch in der Regel in Kenntnis des sich anschließenden 3D-Bildwiedergabegerätes. Die anschließende Kombination der verschiedenen bereitgestellten Ansichten entspricht der Zuordnungsvorschrift des 3D-Displays.

[0016] Ein weiterer wesentlicher Vorteil der Erfindung besteht jedoch darin, dass nach dem Optimierungsschritt der Stackstruktur die Tiefe pro Originalbild ermittelt wird. Im Ergebnis liegen die Daten in einem äußerst effizienten Datenübertragungsformat vor, nämlich als n Bilder (z. B. Originalbilder oder auch Ansichten) plus n Tiefenbilder (mit vorzugsweise $n = 3$), so dass auch eine deutlich geringere Datenrate als bei der Übertragung sämtlicher Ansichten erzielt wird. In das 3D-Bildwiedergabegerät wären dann folgerichtig eine Einheit zur Rekonstruktion der Stackstruktur und die Einheit zur Projektion der Stackstruktur auf die gewünschte Ansicht zu integrieren.

[0017] Für die vorgenannten Schritte kann jeweils an Stelle der Tiefe auch die Disparität verwendet werden. Überdies sei angemerkt, dass eine Projektion prinzipiell auch eine reine Verschiebung mit einschließt.

[0018] Selbstverständlich können auch andere Tiefen- bzw. Disparitätserkennungsverfahren als das

vorstehend genannte eingesetzt werden, um aus dem n-Tupel von Bildern (mit $n > 1$) die Tiefe- bzw. die Disparitäten zu erkennen und/oder daraus weitere Ansichten zu generieren.

[0019] Die Erfindung soll nachstehend anhand von Ausführungsbeispielen näher erläutert werden. Die Zeichnungen zeigen:

[0020] [Fig. 1](#): eine Prinzipansicht der erfindungsgemäßen Anordnung

[0021] [Fig. 2](#): eine Variante mit einer Hauptkamera und zwei Satellitenkameras

[0022] [Fig. 3](#): schematische Darstellung der schrittweisen Verschiebung zweier Zeilen gegeneinander und Generierung der Z Koordinate

[0023] [Fig. 4](#): Optimierungsschema durch Eliminierung von Mehrdeutigkeiten bzgl. [Fig. 3](#)

[0024] [Fig. 5](#): Optimierungsschema durch Reduktion der Elemente auf eine eindeutige Höhenprofilinie bzgl. [Fig. 4](#)

[0025] [Fig. 6](#) schematische Darstellung der schrittweisen Verschiebung dreier Zeilen gegeneinander und Generierung der Z Koordinate

[0026] [Fig. 7](#): Optimierungsschema durch Eliminierung von Mehrdeutigkeiten bzgl. [Fig. 6](#)

[0027] [Fig. 8](#): Optimierungsschema durch Reduktion der Elemente auf eine eindeutige Höhenprofilinie bzgl. [Fig. 7](#)

[0028] [Fig. 9](#) schematische Darstellung einer Projektion einer Ansicht aus dem Optimierungsschema

[0029] Eine erfindungsgemäße Anordnung besteht im Wesentlichen aus einem Stereo-Kamerasystem **1**, einer Bildkonvertierungseinrichtung **2** und einem 3D-Bildwiedergabegerät **3**, wobei gemäß [Fig. 1](#) das Stereo-Kamerasystem **1** eine rechte Kamera **11** und eine linke Kamera **12**, die Bildkonvertierungseinrichtung **2** eine Rektifizierungseinheit **21**, eine Farbjustierungseinheit **22**, eine Einheit zum Aufbau der Stackstruktur **23**, eine Einheit zur Optimierung der Stackstruktur **24** sowie eine Einheit zur Projektion der Stackstruktur auf die gewünschte Ansicht **25** und das 3D-Bildwiedergabegerät **3** eine Bildkombinationseinheit **31** und ein 3D-Display **32** enthält. Gemäß [Fig. 2](#) ist eine weitere erfindungsgemäße Anordnung aufgezeigt. Hierbei enthält das 3D-Kamerasystem **1** eine Hauptkamera **13**, eine erste Satellitenkamera **14** und eine zweite Satellitenkamera **15**. Die Bildkonvertierungseinrichtung **2** enthält eine Größenanpassungseinheit **20**, eine Rektifizierungseinheit **21**, eine Farbjustierungseinheit **22**, eine Einheit zum Aufbau der Stack-

struktur **23**, eine Einheit zur Optimierung der Stackstruktur **24** sowie eine Einheit zur Ermittlung der Tiefe **26** und das 3D-Bildwiedergabegerät **3** enthält, wie aus [Fig. 2](#) ersichtlich, eine Einheit zur Rekonstruktion der Stackstruktur **30**, eine Einheit zur Projektion der Stackstruktur auf die gewünschte Ansicht **35**, eine Bildkombinationseinheit **31** sowie ein 3D-Display **32**.

[0030] Gemäß der Ausführung nach [Fig. 1](#) enthält das Stereo-Kamerasystem **1** zwei identische Kameras (**11**, **12**) mit gleichem Auflösungsvermögen, so dass die aufgenommenen Bilddaten jeweils unmittelbar der Rektifizierungseinheit **21** zugeleitet werden können. Gemäß der Ausführung nach [Fig. 2](#) besteht das 3D-Kamerasystem **1** aus einer Hauptkamera **13** und zwei Satellitenkameras **14**, **15**, wobei die Hauptkamera **13** eine so genannte High-Quality-Kamera mit hohem Auflösungsvermögen ist, hingegen die beiden Satellitenkameras **14**, **15** mit einem geringeren Auflösungsvermögen ausgestattet sind. Die Kamerapositionen untereinander sind dabei, wie üblich, in bekannten Grenzen bezüglich Abstand und Ausrichtung variierbar, um stereoskopische Bilder aufnehmen zu können. Die aufgenommenen Bilddaten der Kameras (**13**, **14**, **15**) werden gemäß [Fig. 2](#) der Größenanpassungseinheit **20** zugeführt. Hier erfolgt eine Angleichung der Bildauflösung von den Satellitenkameras **14**, **15** an die Auflösung der Hauptkamera **13**, so dass alle Bilder die gleiche Auflösung aufweisen. In der nachgeordneten Rektifizierungseinheit **21** erfolgt nun eine mögliche Entzerrung der Kamerabilder, d. h. es erfolgt ein Ausgleich von Linsenverzerrungen, Kameraverdrehungen, Zoomdifferenzen etc.. An die Rektifizierungseinheit **21** schließt sich die Farbjustierungseinheit **22** an. Hier erfolgt eine Angleichung der Farb-/Helligkeitswerte der aufgenommenen Bilder an ein einheitliches Niveau. Die so korrigierten Bilddaten werden nunmehr der Einheit zum Aufbau der Stackstruktur **23** zugeführt. Es werden die Eingangsbilder zeilenweise miteinander verglichen. Der Vergleich gemäß [Fig. 3](#) basiert dabei auf der Ausgestaltung des Stereo-Kamerasystems **1** gemäß [Fig. 1](#), denn hier gibt es den Vergleich von jeweils nur zwei Zeilen. Im ersten Schritt werden zunächst zwei Zeilen jeweils mit derselben Y Koordinate übereinandergelegt, welches gemäß [Fig. 3](#) der Ebene 0 entspricht. Der Vergleich wird pixelweise durchgeführt und das Ergebnis des Vergleiches wird gemäß [Fig. 3](#) entsprechend der vorliegenden Vergleichsebene als Z Koordinate abgespeichert, wobei übereinanderliegende Pixel mit gleichem Farbwert diesen behalten, hingegen bei Ungleichheit kein Farbwert gespeichert wird. Im zweiten Schritt werden die Zeilen gemäß [Fig. 3](#) jeweils $\frac{1}{2}$ Pixel verschoben und es kommt zur Zuordnung der Ebene 1, bzw. es erfolgt ein nächster Vergleich in der Ebene 1, dessen Ergebnis in der Ebene 1 (Z Koordinate) abgespeichert wird. Die Vergleiche werden im allgemeinen, wie aus [Fig. 3](#) ersichtlich, bis zur Ebene 7 und dann Ebene-1 bis Ebene-7 durchgeführt und entsprechend jeweils

als Z Koordinate in der entsprechenden Ebene abgespeichert. Die Anzahl der Ebenen entspricht hierbei der maximal auftretenden Tiefeninformation und kann je nach Bildinhalt variieren. Die so aufgebaute dreidimensionale Struktur mit den XYZ Koordinaten bedeutet, dass für jedes Pixel über die zugehörige Z Koordinate der Grad der Verschiebung der Ansichten zueinander abgespeichert ist. Der gleiche Vergleich wird gemäß [Fig. 6](#) auf Basis der Ausführungsform von [Fig. 2](#) durchgeführt, nur das hier in entsprechender Weise drei Zeilen verglichen werden. Ein einfacher Vergleich zwischen [Fig. 6](#) und [Fig. 3](#) zeigt, dass durch den Vergleich dreier Zeilen wesentlich weniger Fehlinterpretationen erfolgten. Es ist also vorteilhaft, mehr als zwei Zeilen in den Vergleich einzubeziehen. Die erzeugte Stackstruktur, die sich auch dadurch auszeichnet, dass nunmehr die Eingangsbilder nicht mehr einzeln vorliegen, werden der nachfolgenden Einheit zur Optimierung der Stackstruktur **24** zugeführt. Hier werden mehrdeutige Abbildungen von Bildelementen ermittelt, mit dem Ziel, derartige Fehler aufgrund von unwahrscheinlichen Kombinationen zu löschen, so dass eine korrigierte Datenmenge gemäß [Fig. 4](#) oder auch [Fig. 7](#) erzeugt wird. Im nachfolgenden Schritt wird eine möglichst flache bzw. stetige Höhenprofilinie aus den verbleibenden Elementen erstellt, um eine eindeutige Abbildung der Farbwerte in eine diskrete Tiefenebene (Z Koordinate) zu erreichen. Die Ergebnisse sind in [Fig. 5](#) bzw. [Fig. 8](#) dargestellt. Das Ergebnis gemäß [Fig. 5](#) wird nunmehr gemäß [Fig. 1](#) der Einheit zur Projektion der Stackstruktur auf die gewünschte Ansicht **25** zugeführt. Hier erfolgt die Projektion der Stackstruktur auf eine festgelegte Ebene im Raum. Die zu erzeugende Ansicht wird dabei über den Winkel der Ebene, wie aus [Fig. 9](#) ersichtlich, erzeugt. Dabei wird in der Regel mindestens eine Ansicht erzeugt, die keinem der von dem Kamerasystem **1** aufgenommenen Bildern exakt entspricht. Sämtliche erzeugten Ansichten liegen am Ausgang der Bildkonvertierungseinrichtung **2** an und können somit dem nachfolgenden 3D-Bildwiedergabegerät **3** zur stereoskopischen Wiedergabe übergeben werden, wobei mittels der enthaltenen Bildkombinationseinheit **31** zunächst die Kombination der verschiedenen Ansichten entsprechend der vorgegebenen Zuordnungsvorschrift des 3D-Displays **32** erfolgt.

[0031] Gemäß [Fig. 2](#) wird ein anderer Weg zur Übertragung der aufbereiteten Daten an das 3D-Bildwiedergabegerät **3** gewählt. Der Einheit zur Optimierung der Stackstruktur **24** folgt hier die Einheit zur Ermittlung der Tiefe **26**. Durch die Ermittlung der Tiefe der Bilder wird ein besonders effizientes Datenübertragungsformat geschaffen. Übertragen werden hier nämlich nur drei Bilder und drei Tiefenbilder, und zwar bevorzugt im MPEG-4 Format. Gemäß [Fig. 2](#) ist in dem 3D-Bildwiedergabegerät **3** eingangsseitig eine Einheit zur Rekonstruktion der Stackstruktur **30** mit nachfolgender Bildkombinationseinheit **31** und einem 3D-Display **32** vorhanden. Die Übernahme der

Bilder und die Tiefen können in der Einheit zur Rekonstruktion der Stackstruktur **30** besonders effizient durch inverse Projektion in die Stackstruktur wieder gewandelt werden, so dass die Stackstruktur der nachfolgenden Einheit zur Projektion der Stackstruktur auf die gewünschte Ansicht **35** bereitgestellt werden kann. Der weitere Weg ist dann identisch zur Variante gemäß [Fig. 1](#), nur das nunmehr vorteilhaft keine Übertragung sämtlicher Ansichten erfolgen muss.

Bezugszeichenliste

1	Kamerasystem
11	rechte Kamera
12	linke Kamera
13	Hauptkamera
14	erste Satellitenkamera
15	zweite Satellitenkamera
2	Bildkonvertierungseinrichtung
20	Größenanpassungseinheit
21	Rektifizierungseinheit
22	Farbjustierungseinheit
23	Einheit zum Aufbau der Stackstruktur
24	Einheit zur Optimierung der Stackstruktur
25	Einheit zur Projektion der Stackstruktur auf die gewünschte Ansicht
26	Einheit zur Ermittlung der Tiefe
3	3D-Bildwiedergabegerät
30	Einheit zur Rekonstruktion der Stackstruktur
31	Bildkombinationseinheit
32	3D-Display
35	Einheit zur Projektion der Stackstruktur auf die gewünschte Ansicht

Patentansprüche

1. Verfahren zur Aufnahme und Wiedergabe von Bildern einer Szene und/oder eines Objektes, die folgenden Schritte umfassend:

- Erstellung von mindestens einem n-Tupel von Bildern, mit $n > 2$, mit mindestens einer Hauptkamera (**13**) eines ersten Kameratyps und mindestens einer Satellitenkamera (**14**, **15**) eines zweiten Kameratyps,
- Übernahme der Bilddaten in eine Bildkonvertierungseinrichtung (**2**), in der nachfolgend
- bei Verwendung von einer Hauptkamera (**13**) und zwei oder vier Satellitenkameras mit von der Auflösung der Hauptkamera (**13**) verschiedener Auflösung zunächst eine Größenanpassung, so dass sämtliche Bilder nach Größenanpassung jeweils die gleiche Auflösung aufweisen,
- eine Rektifizierung,
- eine Farbjustierung,
- eine Tiefen- oder Disparitätserkennung zur Generation weiterer Ansichten aus dem n-Tupel von Bildern und den Tiefen- bzw. Disparitätserkennungswerten durchgeführt wird, wobei mindestens eine Ansicht generiert wird, die keinem der von dem Kamerasystem (**1**) aufgenommenen Bildern exakt entspricht,

- anschließend eine Kombination der verschiedenen Ansichten entsprechend der Zuordnungsvorschrift des 3D-Displays (**32**) eines 3D-Bildwiedergabegerätes (**3**) erstellt wird, und
- abschließend wird die hilfsmittelfreie Darstellung des kombinierten 3D-Bildes in mindestens drei Ansichten auf dem 3D-Display (**32**) durchgeführt.

2. Verfahren nach Anspruch 1, dadurch gekennzeichnet, dass die erstellten Bilder an die Bildkonvertierungseinrichtung (**2**) übertragen werden.

3. Verfahren nach Anspruch 1, dadurch gekennzeichnet, dass sämtliche von der Bildkonvertierungseinrichtung (**2**) generierten Ansichten jedes Bildes an das 3D-Bildwiedergabegerät (**3**) übertragen werden.

4. Verfahren nach Anspruch 1, dadurch gekennzeichnet, dass die Bilder mittels Computer erstellt werden.

5. Verfahren zur Übertragung von 3D-Informationen, die mittels des Verfahrens nach Anspruch 1 gewonnen wurden, zum Zweck der späteren räumlich wahrnehmbaren Wiedergabe, dadurch gekennzeichnet, dass ausgehend von mindestens einem Tripel von Bildern, die unterschiedliche Blickwinkel eines Objektes oder einer Szene charakterisieren und mindestens zwei der drei Bilder eine paarweise voneinander unterschiedliche Bildauflösung aufweisen, für jedes Bild die Tiefe ermittelt wird und hernach in einem Übertragungskanal das Tripel von Bildern zusammen mit den jeweiligen Tiefeninformationen übertragen wird.

6. Verfahren nach Anspruch 5, dadurch gekennzeichnet, dass die Bilddaten und die Tiefeninformationen im MPEG-4 Format generiert werden.

Es folgen 7 Blatt Zeichnungen

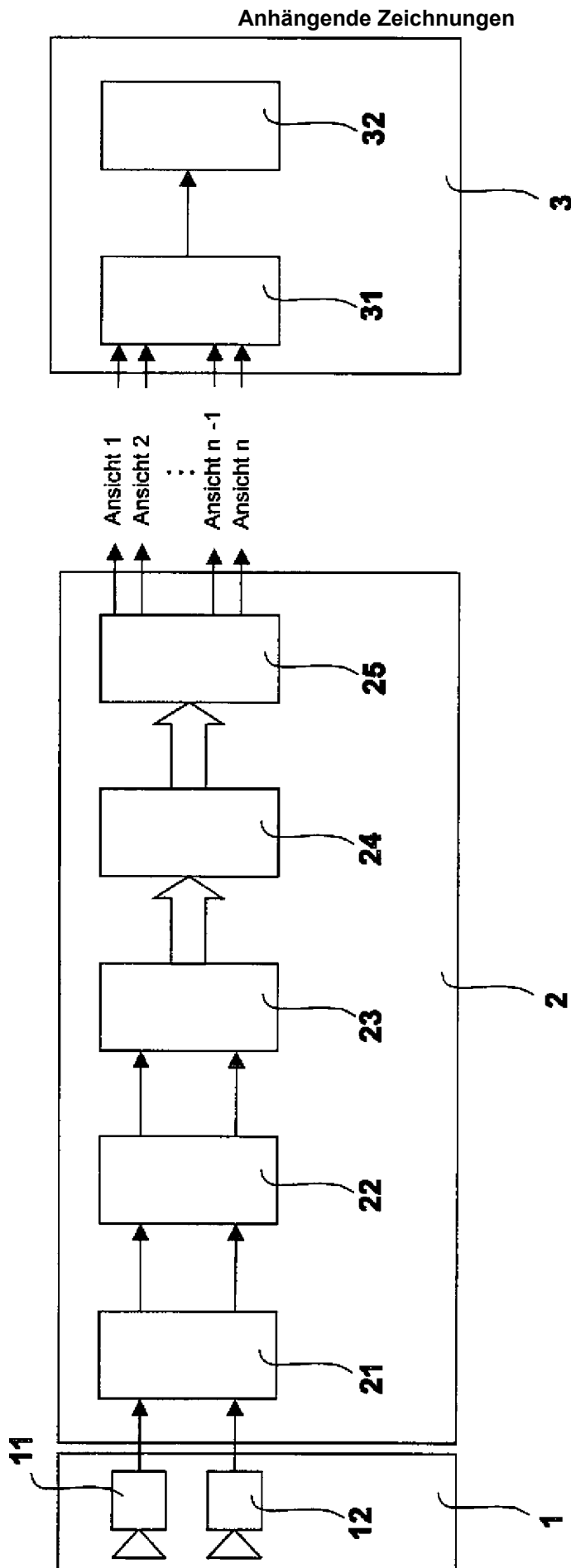


Fig. 1

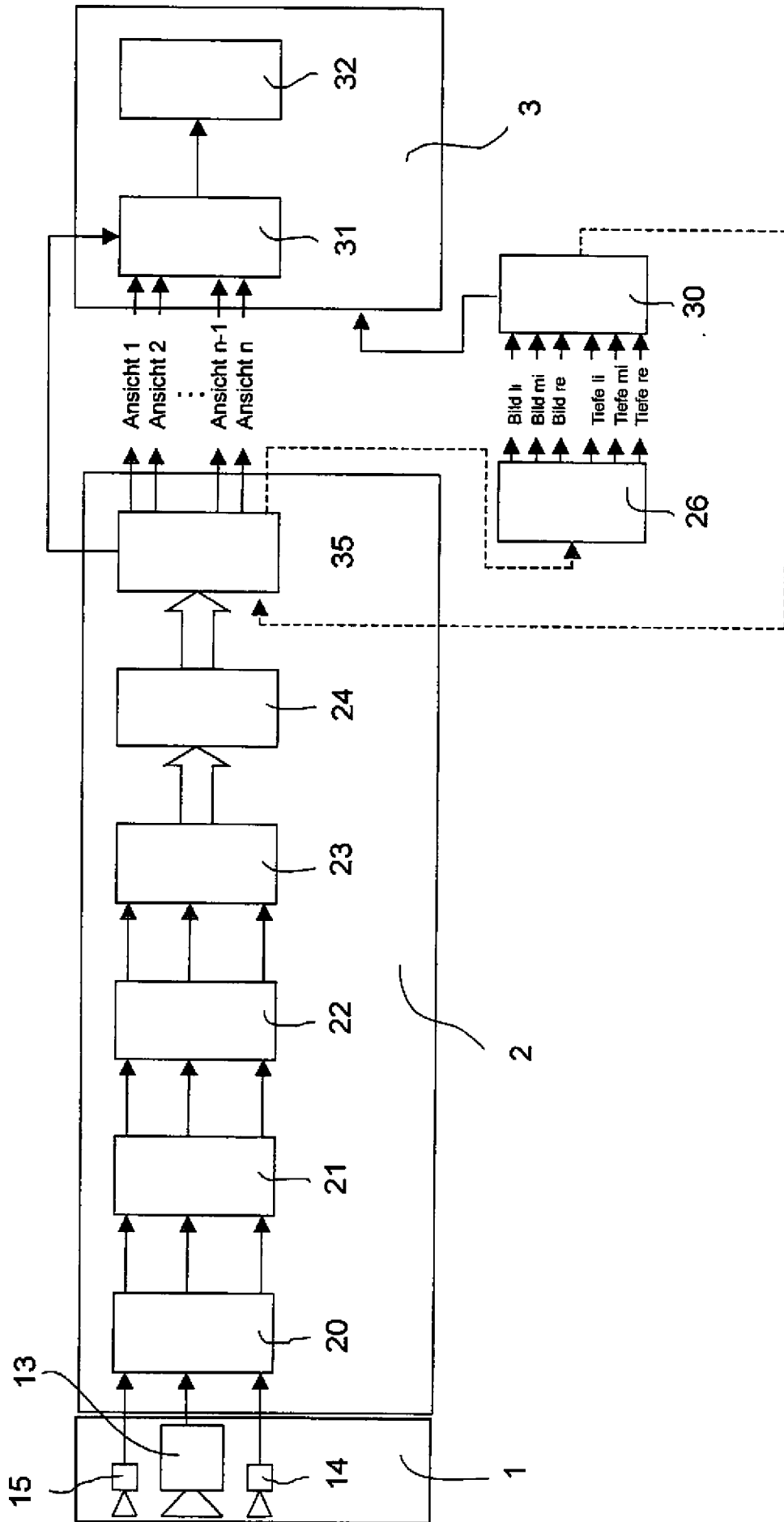


Fig. 2

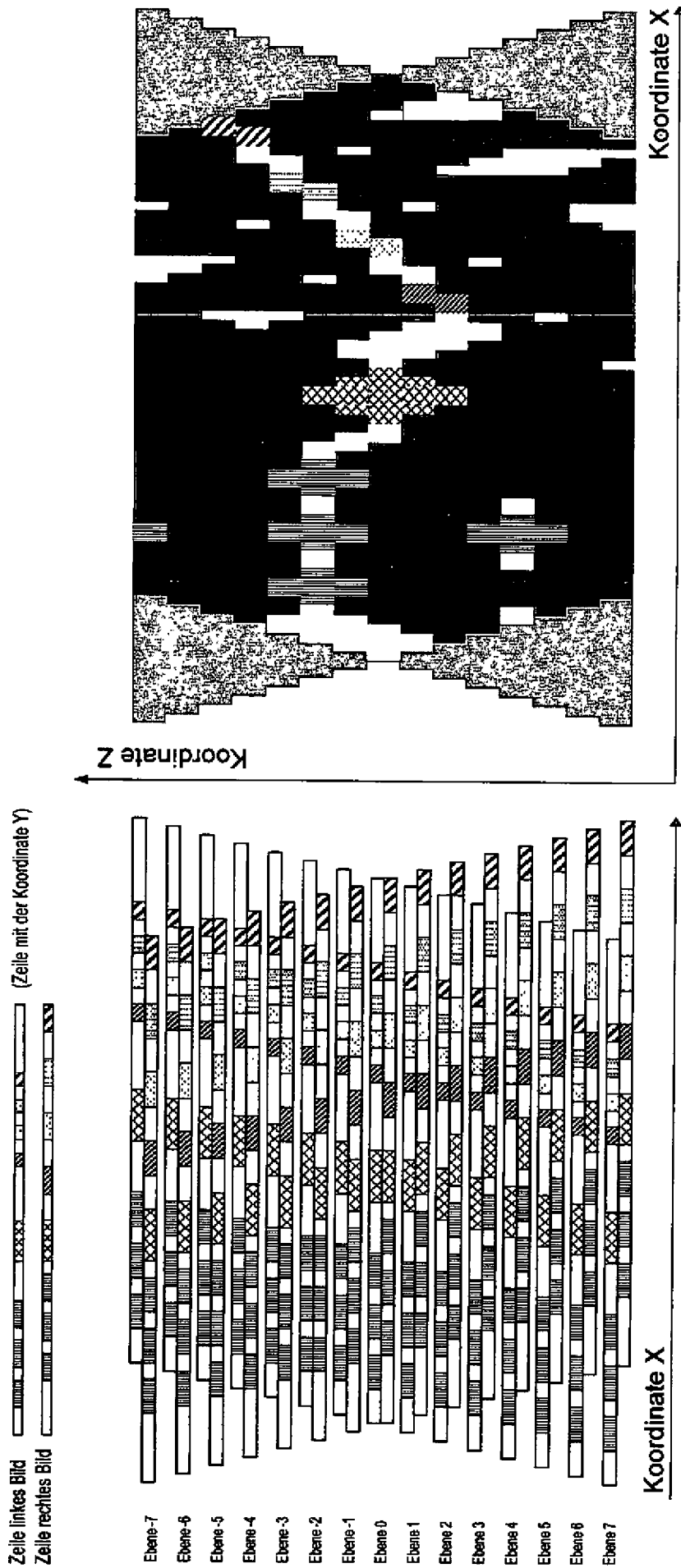


Fig. 3

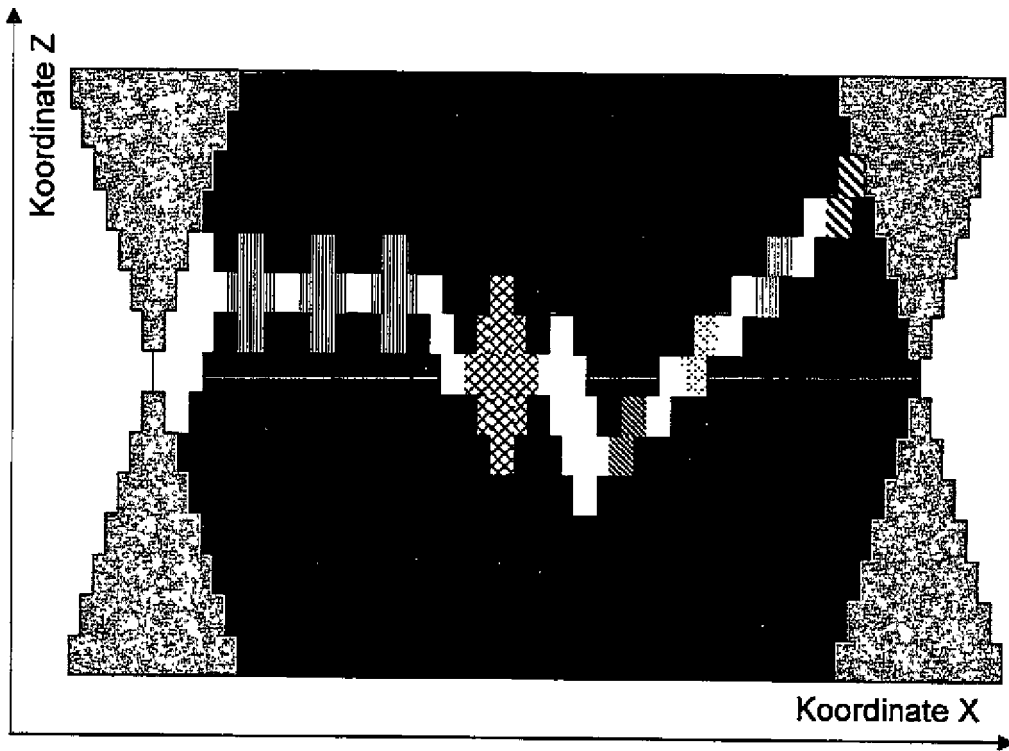


Fig. 4

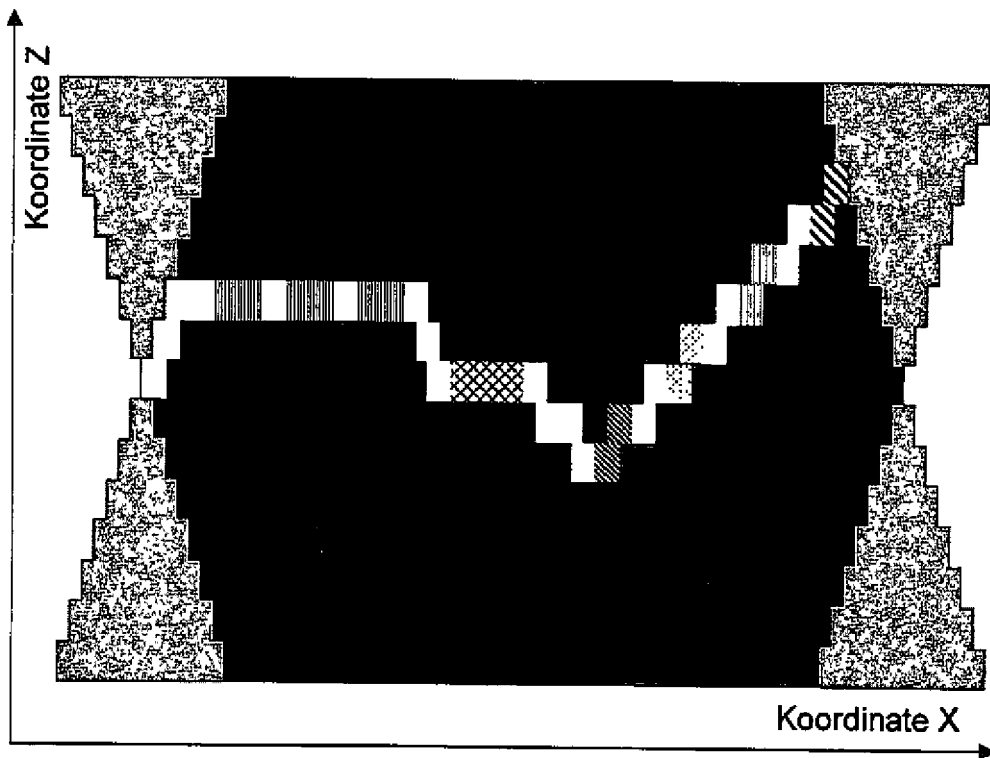


Fig. 5

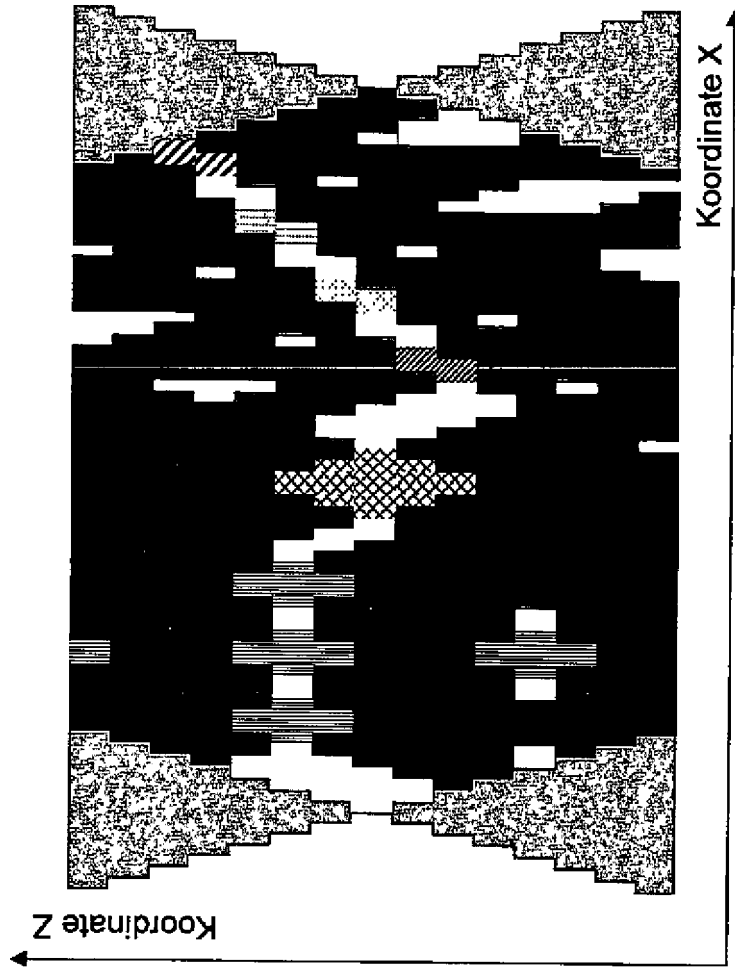
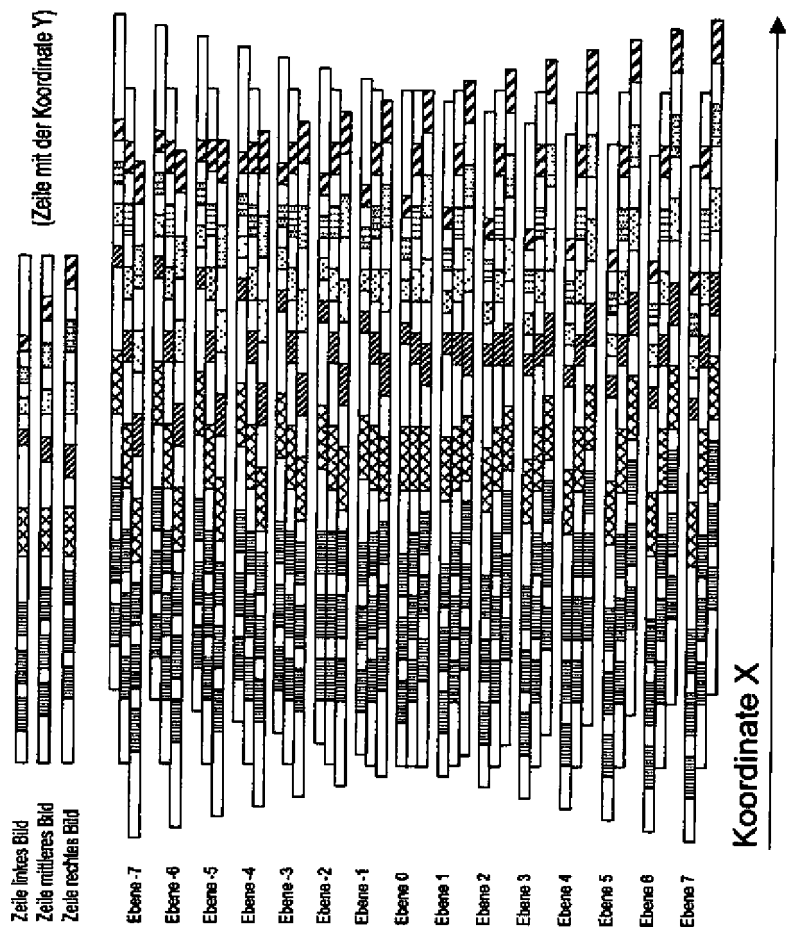


Fig. 6

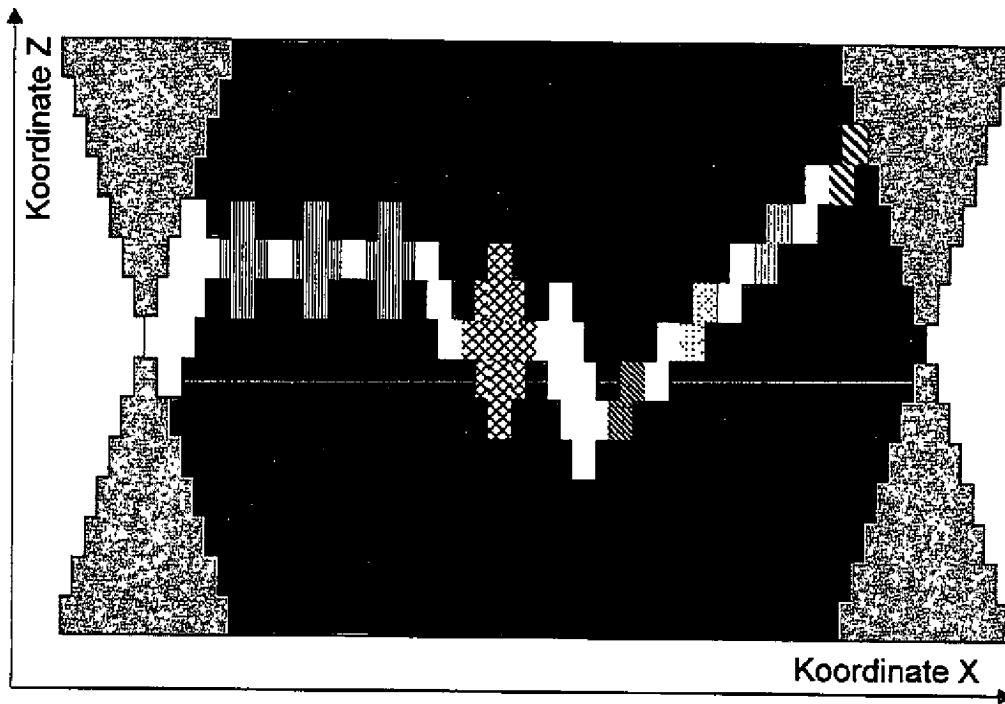


Fig. 7

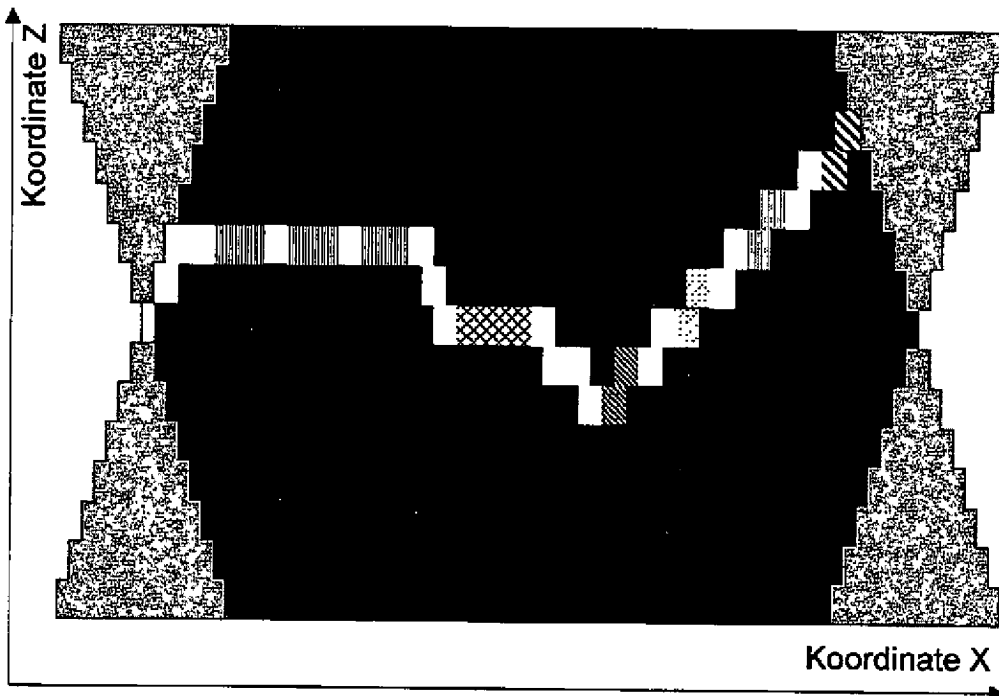


Fig. 8

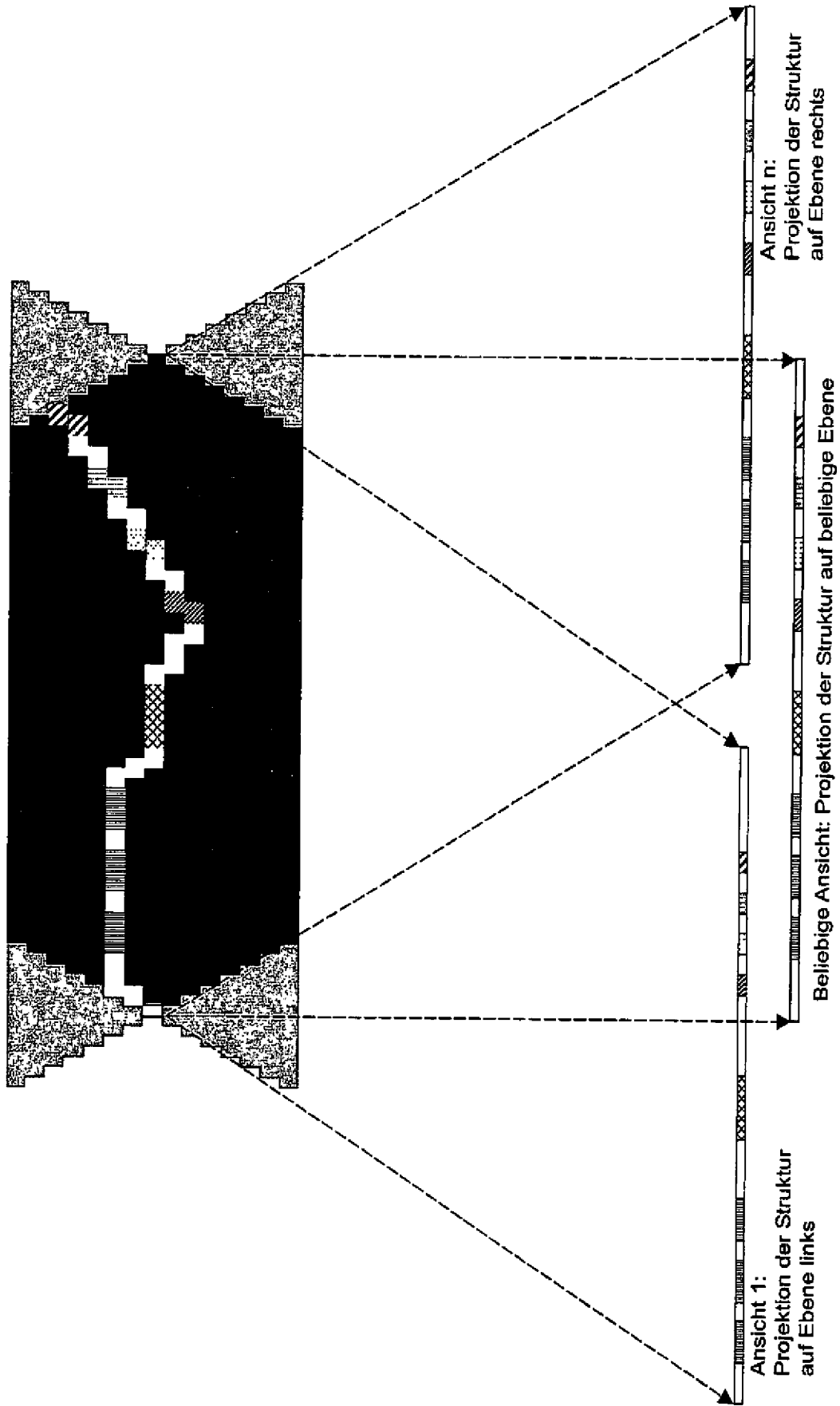


Fig. 9