



(19)  
Bundesrepublik Deutschland  
Deutsches Patent- und Markenamt

(10) **DE 602 05 662 T2** 2006.06.29

(12) **Übersetzung der europäischen Patentschrift**

(97) **EP 1 456 806 B1**

(21) Deutsches Aktenzeichen: **602 05 662.4**

(86) PCT-Aktenzeichen: **PCT/GB02/05846**

(96) Europäisches Aktenzeichen: **02 788 215.8**

(87) PCT-Veröffentlichungs-Nr.: **WO 03/056505**

(86) PCT-Anmeldetag: **20.12.2002**

(87) Veröffentlichungstag  
der PCT-Anmeldung: **10.07.2003**

(97) Erstveröffentlichung durch das EPA: **15.09.2004**

(97) Veröffentlichungstag  
der Patenterteilung beim EPA: **17.08.2005**

(47) Veröffentlichungstag im Patentblatt: **29.06.2006**

(51) Int Cl.<sup>8</sup>: **G06F 3/037 (2006.01)**  
**G06F 3/033 (2006.01)**

(30) Unionspriorität:  
**01310832 21.12.2001 EP**

(73) Patentinhaber:  
**British Telecommunications p.l.c., London, GB**

(74) Vertreter:  
**BEETZ & PARTNER Patentanwälte, 80538  
München**

(84) Benannte Vertragsstaaten:  
**DE, FR, GB**

(72) Erfinder:  
**HOILE, Richard, Cefn, Ipswich, GB**

(54) Bezeichnung: **Vorrichtung und Verfahren zur Berechnung einer Position einer Anzeige**

Anmerkung: Innerhalb von neun Monaten nach der Bekanntmachung des Hinweises auf die Erteilung des europäischen Patents kann jedermann beim Europäischen Patentamt gegen das erteilte europäische Patent Einspruch einlegen. Der Einspruch ist schriftlich einzureichen und zu begründen. Er gilt erst als eingelegt, wenn die Einspruchsgebühr entrichtet worden ist (Art. 99 (1) Europäisches Patentübereinkommen).

Die Übersetzung ist gemäß Artikel II § 3 Abs. 1 IntPatÜG 1991 vom Patentinhaber eingereicht worden. Sie wurde vom Deutschen Patent- und Markenamt inhaltlich nicht geprüft.

**Beschreibung**

**[0001]** Diese Erfindung bezieht sich auf eine Berechnungseinrichtung für einen Anzeigort und auf Verfahren für die Berechnung eines Anzeigortes, insbesondere, aber nicht ausschließlich, für die Verwendung mit Rechnern.

**[0002]** Benutzer müssen oft mit Maschinen kommunizieren, und manchmal müssen sie sie auch steuern. Es ist zunehmend üblich geworden, dass dies ausgeführt wird, indem der Benutzer auf einen Punkt auf einer Anzeige zeigt. Es ist deshalb erforderlich, den Ort auf dem Bildschirm berechnen zu können, auf den der Benutzer zeigt. Oft informiert die Anzeige den Benutzer auch über den Zustand der Maschine.

**[0003]** Ein frühes Beispiel einer Berechnungseinrichtung für den Anzeigort für einen Rechner verwendete einen Bildschirm mit Kathodenstrahlröhre (CRT, Cathode Ray Tube) für die Anzeige des Zustands und einen Lichtgriffel für das Zeigen. Solche Lichtgriffel waren das Gehäuse für einen Fotodetektor, der ein Signal erzeugte, wenn er einen scharfen Anstieg der empfangenen Lichtintensität erfasste. Dieses Signal wurde zu dem Rechner gesendet, der die CRT-Anzeige steuerte. Um auf einen Punkt auf der Anzeige zu zeigen, drückte der Benutzer die Spitze des Lichtgriffels gegen diesen Punkt. Indem er den Zeitpunkt des Vorbeilaufens des Elektronenstrahls aus der Kathodenstrahlröhre überwachte, konnte der Rechner den Kontaktpunkt berechnen, das heißt, den Anzeigort.

**[0004]** In den letzten Jahren sind Geräte, die keinen Kontakt mit dem Bildschirm erfordern, populär geworden. Die am weitesten verbreitete Einrichtung zum Bezeichnen eines Punktes auf eine Anzeige ist die Computermaus. Eine Computermaus erfordert jedoch für die Verwendung eine ebene, nicht-rutschige Oberfläche und kann Schmutz von dieser Oberfläche aufsammeln, der ihren Betrieb dann mit der Zeit verschlechtert.

**[0005]** Berührungslose Einrichtungen werden für das Bezeichnen eines Punktes auf einer Anzeige in manchen speziellen Schnittstellen verwendet. Beispiele umfassen Licht-Schusswaffen, die als Videospielesteuerungen verwendet werden. Licht-Schusswaffen arbeiten auf dieselbe Weise wie Lichtgriffel, aber in einem Abstand von der Anzeige. Berührungslose Computermäuse sind ebenso verfügbar. Keine dieser berührungslosen Einrichtungen sind jedoch vollständig zufriedenstellend.

**[0006]** US-A-5554980, die die Basis für die Präambel von Anspruch 1 darstellt, veröffentlicht eine Fernsteuerung, die für die Cursorsteuerung in Verbindung mit einer Anzeige verwendet wird. Die Fernsteuerung umfasst eine CCD-Kamera, die ein Senderelement beobachtet, das an die Anzeige angrenzt. Die Position, auf die gezeigt wird, wird aus der beobachteten Position des Senderelementes berechnet.

**[0007]** EP-A-0767443 veröffentlicht eine Fernsteuerung, die für die Cursorsteuerung in Verbindung mit einer Anzeige verwendet wird. Die Fernsteuerung umfasst eine CCD-Kamera, die Bilder der Umgebung aufnimmt. Relative Bewegungen der Fernsteuerung werden aus Änderungen in den Bilddaten bestimmt.

**[0008]** WO-A-00/60534 veröffentlicht eine Fernsteuerung, die für die Cursorsteuerung in Verbindung mit einer Anzeige verwendet wird. Die Fernsteuerung umfasst eine Matrix von fotoempfindlichen Zellen, die ein Bild der Anzeige aufnehmen. Die absolute Position, auf die gezeigt wird, wird aus den empfangenen Bilddaten durch Extraktion von speziellen Bildelementen wie Ecken, Cursors und dergleichen berechnet.

**[0009]** Nach einem ersten Aspekt der vorliegenden Erfindung wird eine Berechnungseinrichtung für den Anzeigort geschaffen, mit:

einem Anzeigegerzeuger, der dazu eingerichtet ist, bei der Verwendung einer Anzeige entsprechend Anzeigedaten zu erzeugen, die von einer Maschine erzeugt werden;

einer Zeigeeinrichtung, die eine Kamera trägt, die betrieben werden kann, um Bilddaten zu erzeugen, die wenigstens einen Teil der Szene in dem Sichtbereich der Kamera darstellen, wobei der Teil bei der Verwendung ein Bild mit wenigstens einem Ausschnitt der Anzeige umfasst;

einer Berechnungseinrichtung, die bei der Verwendung dazu eingerichtet ist, die Bilddaten zu empfangen;

die Anzeigedaten zu empfangen;

die Position und/oder Orientierung der Zeigeeinrichtung relativ zu der Anzeige aus den Bilddaten und den Anzeigedaten zu berechnen;

einen Anzeigort aus der berechneten Position und/oder Orientierung zu berechnen.

**[0010]** Indem man Anzeigedaten erhält, die eine Anzeige darstellen, und Bilddaten, die die Anzeige aus Sicht

einer Kamera darstellen, die in eine Zeigeeinrichtung eingebaut ist, kann die Position und/oder Orientierung der Zeigeeinrichtung relativ zu der Anzeige berechnet werden. Folglich kann ein Anzeigeort berechnet werden. Die Zeigeeinrichtung erfordert keine komplexen mechanischen Teile, arbeitet mit verschiedenen Anzeigeerzeugungstechniken (zum Beispiel CRT, Flüssigkristallanzeigen (LCD), Plasmabildschirmen, usw.) und braucht bei der Verwendung nicht mit dem Bildschirm oder irgend einer anderen Oberfläche in Kontakt gebracht zu werden.

**[0011]** Vorzugsweise hat die Zeigeeinrichtung eine langgestreckte Form und der Anzeigeort ist der Punkt, an dem die Längsachse der Zeigeeinrichtung sich mit der Anzeige schneidet. In bevorzugten Ausführungen wird am Anzeigeort ein Cursor in die Anzeige eingeblendet; wobei die Position des Cursors entsprechend der berechneten Position und/oder Orientierung der Zeigeeinrichtung relativ zu dem Display variiert. Auf diese Weise ist es für den Benutzer leichter, den Punkt zu sehen, den er anzeigt.

**[0012]** Nachdem der Benutzer die Kontrolle über einen Cursor hat, der in eine Anzeige eingeblendet ist, ist es wünschenswert, die Maschine steuern zu können. Heute stellen die meisten Maschinen eine Benutzerschnittstellenanordnung zur Verfügung, um einem Benutzer zu ermöglichen, ihren Betrieb zu steuern. Normalerweise enthält eine Schnittstelle irgendeine Einrichtung für die Eingabe von Benutzerbefehlen und irgendeine Einrichtung, die dem Benutzer den Zustand der Maschine als Reaktion auf diese Befehle anzeigt. Ein Benutzer kann verschiedene Befehle in die Maschine eingeben, indem er auf verschiedene Teile der Anzeige zeigt.

**[0013]** Nach einem zweiten Aspekt der vorliegenden Erfindung wird deshalb eine Schnittstellenanordnung zur Bereitstellung einer Schnittstelle zwischen einem Benutzer und einer Maschine geschaffen, mit: einer Berechnungseinrichtung für den Anzeigeort nach dem ersten Aspekt der vorliegenden Erfindung, bei der die Berechnungseinrichtung weiter dazu eingerichtet ist, bei der Verwendung die Maschine entsprechend der Position des Cursors zu steuern.

**[0014]** Auf diese Weise kann ein Benutzer eine Maschine steuern, ohne dass es erforderlich ist, irgendwelche speziellen Fähigkeiten für die Bedienung zu lernen. Darüber hinaus wird die Notwendigkeit von Unmengen von Knöpfen in der Fernsteuerung und die Notwendigkeit der starken Integration von Fernsteuerungshardware und Schnittstellenauslegung, die in manchen Schnittstellenanordnungen zu finden ist, verringert oder vollständig vermieden. Viele verschiedene Maschinentypen stellen eine solche Schnittstelle bereit, die zum Beispiel PCs, digitale Fernsehempfänger, Videokassettenrecorder (VCR, video cassette recorders) und digitale Videodisk(DVD, digital video disk)-Spieler einschließen.

**[0015]** Vorzugsweise umfasst die Maschine einen Rechner mit einem Prozessor, wobei die Berechnungseinheit den Prozessor umfasst. Indem der Prozessor des Rechners verwendet wird, um die erforderliche Berechnung durchzuführen, wird die Notwendigkeit von Verarbeitungsleistung woanders in der Schnittstellenanordnung verringert oder insgesamt vermieden. Alternativ kann die Berechnungseinheit auch mehr als einen Prozessor enthalten, die allein oder zusammen in der Zeigeeinheit oder davon entfernt angeordnet sein können.

**[0016]** Vorzugsweise ist die Kamera eine digitale Videokamera, da das erfasste Bild dann dasselbe Format wie das Bild der Anzeige hat, und folglich für die Bildlageerfassung (ein Prozess, um die Parameter der Transformation zu bestimmen, die erforderlich ist, um ein Bild zu einem andern Bild auszurichten) leicht verfügbar ist. Wenn die Kamera eine Kamera mit festem Fokus mit einer sehr kleinen Schärfentiefe ist, dann ist das Bild nur in einem bestimmten Abstand von dem Bildschirm fokussiert. Wenn man ein fokussiertes Bild erhält, ist der Abstand zum Bildschirm bekannt. Folglich wird die Verarbeitung, die für die Berechnung erforderlich ist, verringert. Um die Berechnung einfach zu halten, hat die Kamera in bevorzugten Ausführungen eine feste räumliche Beziehung zu der Zeigeeinrichtung.

**[0017]** Nach einem dritten Aspekt der vorliegenden Erfindung wird ein Verfahren zur Berechnung eines Anzeigeortes geschaffen, das aus folgenden Schritten besteht:

- i. Erzeugen einer Anzeige entsprechend Anzeigedaten, die von der Maschine erzeugt werden;
- ii. Erfassen von Bilddaten, die wenigstens einen Teil der Szene in dem Sichtbereich einer Kamera darstellen, die in eine Zeigeeinrichtung eingebaut ist, wobei wenigstens ein Ausschnitt der Anzeige in dem Sichtbereich enthalten ist;
- iii. Berechnen der Position und/oder der Orientierung der Zeigeeinrichtung relativ zu der Anzeige aus den Bilddaten und den Anzeigedaten;
- iv. Berechnen eines Anzeigeortes aus der berechneten Position und/oder Orientierung.

**[0018]** Indem Anzeigedaten, die von einer Maschine erzeugt werden, und die eine Anzeige darstellen, mit

Bilddaten, die die Anzeige aus Sicht einer Kamera darstellen, die ein Teil einer Zeigeeinrichtung ist, verglichen werden, kann die Position und/oder Orientierung der Zeigeeinrichtung relativ zu der Anzeige unabhängig von der Art und Weise, auf die die Anzeige aktualisiert wird, berechnet werden. Auf Basis der Ergebnisse der Berechnung kann dann der Anzeigeort berechnet werden, was einem Benutzer ermöglicht, auf einen Punkt in einer Anzeige zu zeigen.

**[0019]** In bevorzugten Ausführungen wird der Verlauf der Position und/ oder der Orientierung der Zeigeeinrichtung relativ zu der Anzeige als zusätzliche Variable in dem Berechnungsschritt für die Position und/oder die Orientierung verwendet. Dies wird bevorzugt, weil der Beginn mit einer guten Schätzung der richtigen Position und/oder Orientierung die Verarbeitung verringert, die erforderlich ist, um die richtige Position und/oder Orientierung zu berechnen.

**[0020]** Nach einem vierten Aspekt der vorliegenden Erfindung wird eine Berechnungseinrichtung für den Anzeigeort geschaffen, mit:

einem Speichermedium, auf dem prozessorlesbarer Code gespeichert ist, der verarbeitet werden kann, um eine Schnittstelle zwischen einem Benutzer und einer Maschine bereitzustellen, wobei der Code folgendes umfasst:

Kode für die Erfassung von Anzeigedaten, der verarbeitet werden kann, um Anzeigedaten zu erhalten, die eine Anzeige darstellen; Kode für die Erfassung von Bilddaten, der verarbeitet werden kann, um Bilddaten von einer Zeigeeinrichtung zu erhalten, die eine Kamera trägt, wobei die Bilddaten wenigstens einen Teil einer Szene in dem Sichtfeld der Kamera darstellen,

Kode für die Berechnung von Position/Orientierung, der verarbeitet werden kann, um aus den Anzeigedaten und den Bilddaten die Position und/oder die Orientierung der Zeigeeinrichtung relativ zu der Anzeige zu berechnen;

Kode für die Berechnung des Anzeigeortes, der verarbeitet werden kann, um aus der berechneten Position und/oder der Orientierung einen Anzeigeort zu berechnen.

**[0021]** Nach einem fünften Aspekt der vorliegenden Erfindung wird ein digitaler Datenträger geschaffen, der ein Programm aus Anweisungen trägt, die von einer Verarbeitungsvorrichtung ausgeführt werden können, um die Verfahrensschritte, wie sie in dem dritten Aspekt der vorliegenden Erfindung aufgestellt sind, auszuführen.

**[0022]** Nun werden Ausführungen der vorliegenden Erfindung nur als Beispiel mit Bezug auf die Zeichnungen im Anhang beschrieben, in denen sich gleiche Bezugsnummern auf gleiche Teile beziehen, und in denen:

**[0023]** [Fig. 1](#) einen Benutzer darstellt, der eine Zeigeeinrichtung nach einer Ausführung der vorliegenden Erfindung bedient;

**[0024]** [Fig. 2a](#) eine perspektivische Ansicht der Zeigeeinrichtung ist;

**[0025]** [Fig. 2b](#) eine Darstellung der Zeigeeinrichtung im Querschnitt ist;

**[0026]** [Fig. 3a](#) ein anzeigebasiertes, rechtshändiges, rechtwinkliges 3D-Koordinatensystem darstellt;

**[0027]** [Fig. 3b](#) ein bildbasiertes, rechtshändiges, rechtwinkliges 2D-Koordinatensystem darstellt;

**[0028]** [Fig. 3c](#) ein Diagramm ist, das die Projektion von einem kamerabasierten Koordinatensystem auf ein bildbasiertes Koordinatensystem darstellt;

**[0029]** [Fig. 4](#) ein Flussdiagramm ist, das die Arbeitsweise des Steuerprozesses der Zeigeeinrichtung darstellt;

**[0030]** [Fig. 5a](#) ein Flussdiagramm ist, das detaillierter die Stufe in dem Schritt des Aufbaus der Abbildungsparameter in [Fig. 4](#) zeigt, die die Anzeige in dem anzeigebasierten Koordinatensystem transformieren und projizieren, um ein Testbild in dem bildbasierten Koordinatensystem aufzubauen;

**[0031]** [Fig. 5b](#) ein Flussdiagramm ist, das detaillierter die Stufe in dem Schritt des Aufbaus der Abbildungsparameter in [Fig. 4](#) zeigt, die das Testbild und das Kamerabild vergleicht, um die tatsächlichen Abbildungsparameter aufzubauen;

**[0032]** [Fig. 5c](#) ein Diagramm ist, das die Filterfenster und die Schrittweite des Filters für das Kamerabild zeigt;

[0033] [Fig. 6](#) ein Flussdiagramm ist, das detaillierter die Schritte in [Fig. 4](#) zeigt, die den Schnittpunkt berechnen und den Cursor positionieren.

[0034] [Fig. 1](#) zeigt einen Benutzer **101**, der eine Zeigeeinrichtung **103** bedient, die über Kabel **105a/105b** an einen Hub **107** eines universellen seriellen Busses (USB, Universal Serial Bus) angeschlossen ist, der selbst an den USB-Port **109** eines Rechners **111** angeschlossen ist. Die Zeigeeinrichtung **103** ([Fig. 2a](#) und [Fig. 2b](#)) umfasst ein Gehäuse **201**, das eine langgestreckte Form hat, und dazu eingerichtet ist und die Größe hat, um von einem Benutzer in der Hand gehalten zu werden. Das Gehäuse hat ein Fenster **203** an einem vorderen Ende und eine Rückwand **205** am hinteren Ende. Das Gehäuse umschließt eine Kamera **207** (mit einem Zentrum der Apertur **209**), die in Längsrichtung orientiert ist, damit sie ein Sichtfeld durch das Fenster **203** an dem vorderen Ende des Gehäuses **201** hat. Eine geeignete Kamera ist die Logitech Quickcam<sup>®</sup> VC (USB-Version), obwohl irgendeine digitale Videokamera mit einem USB-Anschluss, die ausreichend klein ist, um innerhalb eines solchen Gehäuses angeordnet zu werden, genügt. Der Fokussierungsmechanismus der Kamera ist fest und bewirkt, dass ein fokussiertes Bild eines Objektes auf dem CCD-Detektorfeld erzeugt wird, wenn das Objekt 300 mm von der Kamera entfernt ist. Die Leuchtdiode (LED, Light Emitting Diode) **210** wird bereitgestellt, um dem Benutzer anzuzeigen, dass das aufgenommene Bild im Fokus ist. Ein Fachmann hätte keine Probleme, einen geeigneten Algorithmus zu implementieren, um diese Rückmeldung bereitzustellen.

[0035] Das elektronische Ausgangssignal der Kamera **207**, das die Ansicht durch das Fenster **203** darstellt, wird über eines der USB-Kabel **105a** zu dem Rechner **111** gesendet. Das Kabel **105a** führt von der Kamera **207** durch die Rückwand **205** des Gehäuses **201** und endet in einem USB-Anschluss, der bei der Verwendung an den USB-Hub **107** angeschlossen ist. Der USB-Hub **107** ist bei der Verwendung in den USB-Port **109** des Rechners **111** eingesteckt. Die Kamera **207** entnimmt ihre Stromversorgung aus der USB-Verbindung.

[0036] Zwei Knöpfe **211a/213a** sind jeweils an jeder lateralen Seite des Gehäuses **201** in Richtung seines vorderen Endes angeordnet. Zwei Schalter **211b/213** sind von jedem Knopf **211a/213a** aus nach innen hin angebracht und werden durch das Drücken des Knopfes **211a/213a** gesteuert. Ein Kabel führt von jedem Schalter **211b/213b** zu einer Leiterplatte **215**, die an dem hinteren Ende des Gehäuses **201** angeordnet ist. Der Zustand der Schalter **211b/213b** wird an den Rechner **111** über das andere USB-Kabel **105** gesendet, das von der Leiterplatte **215** durch die Rückwand **205** des Gehäuses **201** führt und in dem USB-Anschluss endet, der bei der Verwendung mit dem USB-Hub **107** verbunden ist.

[0037] Noch einmal mit Bezug auf [Fig. 1](#) umfasst der Rechner **111** eine zentrale Verarbeitungseinheit (CPU, central processing unit) **113** und ein Speichersystem **115**. Das Speichersystem **115** umfasst:

- a) einen Speicher mit beliebigem Zugriff (RAM) **117**;
- b) eine erste Speichereinrichtung **119** (z. B. eine Festplatte), die das Programm für den Betrieb der CPU und Bilddaten speichert;
- c) eine zweite Speichereinrichtung **121** (z. B. ein Diskettenlaufwerk oder ein CD/DVD-Laufwerk), um davon Daten zu lesen und/oder Daten auf ein wechselbares Speichermedium **123** zu schreiben.

[0038] Der Rechner umfasst weiter eine Netzwerkkarte **125**, die eine Schnittstelle zu einem Netzwerk bildet.

[0039] Die CPU **113** kommuniziert mit dem Speichersystem **115**. Die Ausgaben des Rechners **111** werden auf einem Anzeigebildschirm **123** angezeigt. In der aktuellen Ausführung ist der Anzeigebildschirm **123** ein LCD-Bildschirm. Es ist jedoch möglich, einen CRT-Bildschirm statt des LCD-Bildschirms zu verwenden. In einem solchen Fall ist klar, dass sich Flackereffekte, ähnlich denen, die man sieht, wenn CRTs im Fernsehen aufgenommen werden, ergeben können, wenn die Kamera die Bilder zu schnell aufnimmt. Es ist deshalb erforderlich, die Erfassungsdauer auf einen größeren Teil einer Sekunde auszuweiten, um auf diese Weise sicherzustellen, dass die Spitzen der Ausgaben des CRT geglättet werden.

[0040] Obwohl [Fig. 1](#) die Zeigeeinrichtung **102** als mit dem Rechner **111** über Kabel verbunden zeigt, ist es offensichtlich, dass eine schnurlose Verbindung, wie etwa eine Funk- oder Infrarot-Verbindung, verwendet werden kann, was dem Benutzer **101** mehr Bewegungsfreiheit ermöglicht. Dies ist insbesondere nützlich, wenn der Benutzer **101** eine Präsentation macht, und sich in einem Bühnenbereich bewegen muss. In diesem Fall entnimmt die Kamera ihre Leistung aus wiederaufladbaren Batterien, die im Gehäuse angeordnet sind. Die Batterien werden wieder aufgeladen, indem die Zeigeeinrichtung in einen geeignet mit Leistung versorgten Halter für das Wiederaufladen eingesetzt wird. Fachleute haben keine Schwierigkeiten, die geeigneten Modifikationen durchzuführen.

[0041] Verschiedene rechtwinklige Koordinatensysteme können verwendet werden, um den Raum der realen

Welt zu beschreiben, der von dem Anzeigebildschirm, der Kamera und dem Bild eingenommen wird, das in der Kamera erzeugt wird.

**[0042]** Ein anzeigebasiertes, rechtshändiges, rechtwinkliges 3D-Koordinatensystem (**Fig. 3a**) ermöglicht, dass die Position eines Punktes im Bereich der realen Welt durch drei Koordinaten  $[X_s, Y_s, Z_s]$  definiert wird. Die Abstandseinheit auf den Achsen ist mm. Entsprechend dem anzeigebasierten Koordinatensystem nimmt der Anzeigebildschirm **123** einen rechteckigen Ausschnitt in der Ebene  $z_s = 0$  ein. Die  $x_s$ -Achse verläuft horizontal am oberen Rand der Anzeige von dem Ursprung  $(0, 0, 0)$  zu oberen rechten Ecke  $(320, 0, 0)$ , und die  $y_s$ -Achse verläuft die linke Seite der Anzeige herunter zur unteren linken Ecke  $(0, 240, 0)$ . Folglich füllt der Anzeigebildschirm einen Bereich von  $320 \times 240$  mm aus.

**[0043]** Der Anzeigebildschirm bietet VGA (video graphics array, Videografikfeld)  $(640 \times 480 \text{ Pixel})$ -Auflösung. Die Abbildung der Ecken des Anzeigebildschirms (die in dem anzeigebasierten Koordinatensystem definiert sind) auf Pixel des Anzeigebildschirms  $(x_{sp}, y_{sp})$  ist in der Tabelle unten gezeigt:

$(x_s, y_s, z_s)$	$(x_{sp}, y_{sp})$
$(0, 0, 0)$	$(0, 0)$
$(320, 0, 0)$	$(640, 0)$
$(0, 240, 0)$	$(0, 480)$
$(320, 240, 0)$	$(640, 480)$

**[0044]** Folglich ist klar, dass die Beziehung zwischen dem anzeigebasierten Koordinatensystem und den Pixeln in dem Anzeigebildschirm ist:

Beziehung 1:  $(x_s, y_s, z_s) \Leftrightarrow (\frac{1}{2}x_{sp}, \frac{1}{2}y_{sp})$  für  $0 \leq x_{sp} \leq 640$  &  $0 \leq y_{sp} \leq 480$

**[0045]** Ein kamerabasiertes, rechtshändiges, rechtwinkliges 3D-Koordinatensystem (**217, Fig. 2b**) wird durch drei Koordinaten  $[X_c, Y_c, Z_c]$  definiert. In dem kamerabasierten Koordinatensystem nimmt das Zentrum des Kameraobjektivsystems **209** die Position  $(0, 0, 0)$  ein und die Sichtlinie der Kamera zeigt entlang der  $z_c$ -Achse in Richtung steigender Werte von  $z_c$ . Von hinter der Kamera aus gesehen erstreckt sich die  $y_c$ -Achse lateral in Richtung der Oberseite der Kamera, und die  $x_c$ -Achse erstreckt sich lateral und nach links von der Kamera. Die Abstandseinheit auf den Achsen ist mm.

**[0046]** In **Fig. 3c** erzeugt ein Objektpunkt **305** in dem Sichtfeld der Kamera durch das Zentrum des Objektivsystems der Kamera **209** einen Bildpunkt **307** auf dem CCD-Detektorfeld in der Kamera. Der Objektpunkt **305** befindet sich an einer Position  $(x_{co}, y_{co}, z_{co})$  in dem kamerabasierten Koordinatensystem. Der Bildpunkt liegt in der Ebene  $z_c = -n$ , wobei  $n$  der Abstand zwischen dem Zentrum des Linsensystems der Kamera **209** und dem CCD-Detektorfeld in mm ist.

**[0047]** Ein bildbasiertes, rechtshändiges, rechtwinkliges 2D-Koordinatensystem (**Fig. 3b**) ist durch zwei Koordinaten  $[X_i, Y_i]$  definiert. Die Abstandseinheit auf den Achsen ist wieder mm. In dem bildbasierten Koordinatensystem liegt der Ursprung  $(0, 0)$  im Zentrum des CCD-Detektorfeldes. Von hinter dem Feld aus betrachtet verläuft die  $x_i$ -Achse horizontal und nach links von dem Ursprung  $(0, 0)$  zu dem Punkt  $(1,96585, 0)$ , und die  $y_i$ -Achse verläuft vertikal und nach oben von dem Ursprung  $(0, 0)$  zu dem Punkt  $(0, 1,6084)$ . Auf diese Weise füllt das CCD-Detektorfeld einen Bereich von  $3,9317 \times 3,2168$ mm aus. Der Punkt  $(x_c, y_c, -n)$  in dem kamerabasierten Koordinatensystem entspricht dem Punkt  $(x_i, y_i)$  in dem bildbasierten Koordinatensystem.

**[0048]** Das CCD-Detektorfeld bietet  $352 \times 288$  Pixel Auflösung. Die Abbildung der Ecken des Feldes (definiert in dem bildbasierten Koordinatensystem) auf Bildpixel  $(x_{ip}, y_{ip})$  ist in der Tabelle unten gezeigt:

$(x_i, y_i)$	$(x_{ip}, y_{ip})$
$(-1,96585, -1.6084)$	$(0, 0)$
$(1,96585, -1.6084)$	$(352, 0)$
$(-1,96585, 1.6084)$	$(0, 288)$
$(1,96585, 1.6084)$	$(352, 288)$

**[0049]** Folglich wird klar, dass die Beziehung zwischen dem bildbasierten Koordinatensystem und dem Bildpixeln ist:

Beziehung 2:  $(x_i, y_i) \Leftrightarrow (0,01117[x_{ip}-176], 0,01117[y_{ip}-144])$  für  $0 \leq x_{ip} \leq 352$  &  $0 \leq y_{ip} \leq 288$

**[0050]** Die allgemeine Beziehung zwischen dem anzeigebasierten Koordinatensystem und dem kamerabasierten Koordinatensystem kann durch eine Starrkörpertransformation ausgedrückt werden. Eine Starrkörpertransformation ist eine beliebige Zusammenfassung von Transformationen bzgl. Maßstab, Verschiebung und Rotation.

**[0051]** Die x-Koordinate eines Objektpunktes **305**, der in dem anzeigebasierten Koordinatensystem definiert ist, kann in die x-Koordinate eines Punktes, der in dem kamerabasierten Koordinatensystem definiert ist, nach der folgenden Gleichung transformiert werden:

$$x_{co} = m_{00}x_s + m_{01}y_s + m_{02}z_s + m_{03}$$

wobei  $m_{ab}$  Koeffizienten der Transformation sind.

**[0052]** Ähnlich gilt für die y- und z-Koordinaten:

$$y_{co} = m_{10}x_s + m_{11}y_s + m_{12}z_s + m_{13}$$

$$z_{co} = m_{20}x_s + m_{21}y_s + m_{22}z_s + m_{23}$$

**[0053]** Diese drei Gleichungen können in der folgenden Matrix Gleichung geschickt kombiniert werden:

Beziehung 3:

$$\begin{bmatrix} x_{co} \\ y_{co} \\ z_{co} \\ 1 \end{bmatrix} = T \begin{bmatrix} x_s \\ y_s \\ z_s \\ 1 \end{bmatrix} = \begin{bmatrix} m_{00} & m_{01} & m_{02} & m_{03} \\ m_{10} & m_{11} & m_{12} & m_{13} \\ m_{20} & m_{21} & m_{22} & m_{23} \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_s \\ y_s \\ z_s \\ 1 \end{bmatrix} = \begin{bmatrix} m_{00}x_s + m_{01}y_s + m_{02}z_s + m_{03} \\ m_{10}x_s + m_{11}y_s + m_{12}z_s + m_{13} \\ m_{20}x_s + m_{21}y_s + m_{22}z_s + m_{23} \\ 1 \end{bmatrix}$$

wobei T die Transformationsmatrix ist.

**[0054]** Es ist auch möglich, eine Beziehung zwischen den kamerabasierten Koordinaten eines Objektpunkts in dem Sichtfeld der Kamera und den bildbasierten Koordinaten des Bildes des Objektpunkts, der auf das CCD-Detektorfeld der Kamera abgebildet wird, abzuleiten.

**[0055]** Ein Dreieck Q wird vor dem System der Kamera durch folgendes ([Fig. 3c](#)) definiert:

1. Die Ebene  $y_c = 0$  (die, wie man sich erinnert, eine Ebene ist, die Kamera längs unter rechtem Winkel zu der Achse  $y_c$  schneidet)
2. Die Projektion der Linie, die durch den Objektpunkt **305** und das Zentrum des Objektivsystems der Kamera **209** verläuft, auf die Ebene  $x_c = 0$  (eine Ebene, die auch die Kamera längs auf der halben Höhe des CCD-Detektorfeldes schneidet)
3. Die Ebene  $z_c = z_{co}$  (eine Ebene im rechten Winkel zur Längsachse der Kamera, die durch den Objektpunkt **305** verläuft).

**[0056]** Ein weiteres Dreieck R, das (bezüglich der relativen Längen seiner Seiten) dem Dreieck Q ähnlich ist, ist hinter dem Linsensystem durch folgendes definiert:

4. Die Ebene  $y_c = 0$  (wie oben),
5. Linie 2 wie oben,
6. die Ebene des CCD-Detektorfeldes.

**[0057]** Da die Dreiecke Q und R ähnlich sind, folgt, dass die Beziehung zwischen dem kamerabasierten Koordinatensystem und dem bildbasierten Koordinatensystem (bei Ignorieren von Vorzeichenwechseln) als folgendes ausgedrückt werden kann:

$$\frac{y_{co}}{z_{co}} = \frac{y_i}{n}$$

**[0058]** Ebenso:

$$\frac{x_{co}}{z_{co}} = \frac{x_i}{n}$$

**[0059]** Oder, indem umgeformt wird und Vorzeichenwechsel berücksichtigt werden:

Beziehung 4: 
$$y_i = -n \frac{y_{co}}{z_{co}}$$

Beziehung 5: 
$$x_i = -n \frac{x_{co}}{z_{co}}$$

**[0060]** Die Kombination dieser zwei Beziehungen und die obige Matrixgleichung definieren eine Beziehung:  $(x_s, x_s, 0) \Leftrightarrow (x_i, y_i)$  die irgendeinen Punkt auf der Anzeige, der im Sichtfeld der Kamera liegt, auf einen entsprechenden Punkt in dem Bild abbildet, das auf dem CCD-Detektorfeld abgebildet wird.

**[0061]** Der Rechner **111** ([Fig. 1](#)) arbeitet unter der Steuerung von Software, die ausgeführt werden kann, um einen Berechnungsprozess für den Anzeigeort durchzuführen. Wie Fachleuten klar ist, kann irgendeine oder die gesamte Software, die verwendet wird, um die Erfindung zu implementieren, auf verschiedenen Übertragungs- und/oder Speichermedien **123** wie etwa Disketten, CD-ROM oder Magnetband enthalten sein, sodass sie auf den Rechner **111** geladen werden kann, oder über ein Rechnernetzwerk mit einem geeigneten Übertragungsmedium heruntergeladen werden kann. Der Berechnungsprozess für den Anzeigeort wird nun detaillierter mit Bezug auf das Flussdiagramm in [Fig. 4](#) beschrieben.

**[0062]** Der Berechnungsprozess für den Anzeigeort beginnt mit der Erfassung (Schritt **401**) der Anzeigedaten, die auf dem Anzeigebildschirm **123** von dem Rechner **111** erzeugt werden, und deren Speicherung in den konstanten Datenspeicher **121**. Dann werden die Bilddaten, die das letzte Kamerabild definieren, von dem Rechner **111** erfasst (Schritt **403**) und in den konstanten Datenspeicher **121** gespeichert. Die Anzeigedaten und Kameradaten, die soeben erfasst wurden, werden dann verglichen, um die Transformationsparameter ( $m_{ab}$  oben) abzuleiten, die die Transformation von dem anzeigebasierten Koordinatensystem auf das kamerabasierte Koordinatensystem darstellen (Schritt **405**). Danach werden die Parameter berechnet, die die 3D-Beziehung der Position und/oder Orientierung der Kamera **207** (und folglich der Zeigeeinrichtung **103**) relativ zu dem Anzeigebildschirm **123** definieren (Schritt **407**). Dann wird der Schnittpunkt zwischen der Längsachse der Kamera und der Oberfläche des Anzeigebildschirms **123** berechnet (Schritt **409**). Die Software steuert dann den Rechner **111** und fügt in der Anzeige an dem Schnittpunkt, der in Schritt **409** berechnet wurde, einen Cursor hinzu (Schritt **411**). Danach werden alle obigen Schritte **401–411** wiederholt, um eine feine Steuerung des Cursors durch die Bewegung der Zeigeeinrichtung zu gewährleisten. Die erneute Berechnung der Cursorposition sollte so oft wie möglich stattfinden, wobei die Reserve-Verarbeitungsleistung der CPU verwendet wird, idealerweise sollte sie jedoch so oft wie bei einer Standard-Computermaus stattfinden; d. h. ungefähr 40 mal pro Sekunde.

**[0063]** Eine optionale Erweiterung des Prozesses oben umfasst einen Schritt **413**, der den Verlauf der Position und/oder Orientierung der Kamera relativ zu dem Anzeigebildschirm (in der Form der zuvor verwendeten Transformationen) dem Schritt **405** bereitstellt, der die Parameter der Abbildung von dem anzeigebasierten Koordinatensystem auf das kamerabasierten Koordinatensystem ableitet. Dies wird bevorzugt, weil das Beginnen mit einem guten Schätzwert der richtigen Position und/oder Orientierung die erforderliche Verarbeitung verringert.

**[0064]** Die Schritte des Berechnungsalgorithmus für den Anzeigort, die oben beschrieben wurden, werden nun in weiteren Details beschrieben.

**[0065]** Das Erfassen der Anzeigedaten ([Fig. 4](#), Schritt **401**) enthält die Erfassung der digital quantisierten Intensitätsniveaus der ausgegebenen Pixel, die auf dem Anzeigebildschirm **123** in einem oder mehreren Farbbändern  $c$  angezeigt werden. Ein Intensitäts-Bild (schwarz und weiß) hat nur ein Band, wogegen ein RGB-Bild (Echtfarben) drei Farbbänder hat; das heißt, rot, grün und blau.

**[0066]** Die Intensitätswerte der Pixel der Anzeige in jedem Farbband  $c$  sind Abtastwerte der Intensitätsfunktion der Anzeige  $S_c(x_{sp}, y_{sp})$ . Die Anzahl der Abtastwerte hängt von der Auflösung des Anzeigebildschirms **123** ab. Wie oben beschrieben bietet der Anzeigebildschirm **123** VGA-Auflösung ( $640 \times 480$  Pixel). Die Position des Pixels im Anzeigebildschirm (306, 204) in anzeigebasierten Koordinaten (153, 102, 0) ist in [Fig. 3a](#) gezeigt. Der Intensitätswert des Pixels im Anzeigebildschirm für das Pixel (306, 204) im Anzeigebildschirm im grünen Band ist der Abtastwert von  $S_g(306, 204)$ .

**[0067]** Die Intensitätswerte der Pixel im Anzeigebildschirm für alle Abtastwerte werden als 2D-Array gespeichert, auf das sich als  $S_{c\text{sample}}$  bezogen wird. Zum Beispiel wird der Abtastwert von  $S_g(306, 204)$  in  $S_{c\text{sample}}$  an der Position (306, 204) gespeichert.

**[0068]** Die Farbwerte der Pixel in der Anzeige sind Abtastwerte der Farbfunktion der Anzeige  $S_{\text{colour}}(x_{sp}, y_{sp})$ , die eine Funktion der Intensitätswerte der Pixel in der Anzeige in jedem Farbband ist, sodass  $S_{\text{colour}}(x_{sp}, y_{sp}) = S_R(x_{sp}, y_{sp})S_G(x_{sp}, y_{sp})S_B(x_{sp}, y_{sp})$  ist. Die Farbwerte der Pixel in der Anzeige werden für alle Abtastwerte in einem 2D-Array gespeichert, auf das sich als  $S_{\text{colour}}$  bezogen wird. Zum Beispiel werden die Farbwerte für das Pixel in der Anzeige (306, 204) in  $S_{\text{colour}}$  an der Position (306, 204) gespeichert.

**[0069]** Das Erfassen der Bilddaten ([Fig. 4](#), Schritt **403**) enthält das Erfassen der digital quantisierten Intensitätsniveaus der Pixel des Bildes, die von dem CCD-Detektorfeld der Kamera in einem oder mehreren Farbbändern  $c$  aufgenommen werden.

**[0070]** Die Intensitätswerte der Bildpixel jedem Farbband  $c$  sind Abtastwerte der Funktion  $I_c(x_{ip}, y_{ip})$ . Die Anzahl von Abtastwerten hängt von der Auflösung des CCD-Detektorfeldes ab. Wie oben beschrieben bietet das CCD-Detektorfeld  $352 \times 228$  Pixel Auflösung. Die Position des Bildpixels (90, 70) in bildbasierten Koordinaten  $(-0,96062, -0,82658)$  ist in [Fig. 3b](#) gezeigt. Der Intensitätswert des Bildpixels für das Pixel (90, 70) in dem roten Band ist der Abtastwert von  $I_r(90, 70)$ .

**[0071]** Die Intensitätswerte der Bildpixel für alle Abtastwerte werden als 2D-Feld gespeichert, auf das sich als  $I_{c\text{sample}}$  bezogen wird. Zum Beispiel wird der obige Abtastwert von  $I_r(90, 70)$  in  $I_{c\text{sample}}$  an der Position **(90, 70)** gespeichert.

**[0072]** Die Farbwerte der Bildpixel sind Abtastwerte der Bildfarbfunktion  $I_{\text{colour}}(x_{ip}, y_{ip})$ , die eine Funktion der Intensitätswerte der Anzeigepixel in jedem Farbband ist, sodass  $I_{\text{colour}}(x_{ip}, y_{ip}) = I_R(x_{ip}, y_{ip})I_G(x_{ip}, y_{ip})I_B(x_{ip}, y_{ip})$  ist. Die Farbwerte der Anzeigepixel für alle Abtastwerte werden als 2D-Feld gespeichert, auf das sich als  $I_{\text{colour}}$  bezogen wird. Zum Beispiel wird der Farbwert für das Anzeigepixel **(90, 70)** in  $I_{\text{colour}}$  an der Position **(90, 70)** gespeichert.

**[0073]** Mit Bezug auf die [Fig. 5a](#) und [Fig. 5b](#) wird nun die Ableitung der Transformationsparameter, die die Transformation von dem anzeigebasierten Koordinatensystem in das kamerabasierte Koordinatensystem ([Fig. 4](#), Schritt **405**) darstellen, nun detaillierter offengelegt. Der Zweck der Ableitung der Abbildungsparameter ist, die Transformation zu bestimmen, die die Beziehung zwischen dem anzeigebasierten Koordinatensystem und dem kamerabasierten Koordinatensystem am nächsten beschreibt.

**[0074]** Nach dem eine Transformation mit optimaler Anpassung identifiziert wurde, kann die mathematische Beziehung zwischen dem anzeigebasierten Koordinatensystem und dem kamerabasierten Koordinatensystem bestimmt werden. In anderen Worten kann die 3D-Position der Kamera zu dem Anzeigebildschirm **123** und die Orientierung der Kameraachse bestimmt werden ([Fig. 4](#), Schritt **407**), da jede Position und/oder Orientierung der Kamera (abgesehen von den Positionen, an denen der Anzeigebildschirm nicht sichtbar ist) zu einem eindeutigen Bild des Anzeigebildschirms **123** in der Kamera **207** führt.

**[0075]** Mit Bezug auf [Fig. 5a](#) wird ein Testbild erzeugt, indem ein hypothetische Transformation  $T$  auf das aktuelle Bild in der Anzeige auf dem Anzeigebildschirm **123** angewendet wird. Dieses Testbild ist eine sehr gute

Annäherung an die Art und Weise, wie das Kamerabild aussehen sollte, unter der Annahme, dass T die richtige Starrkörpertransformation war, die das Anzeigebild transformiert, damit es wie das Kamerabild aussieht.

**[0076]** Bezüglich der Beziehung des Anzeigebildschirms und der Kamera wird eine anfängliche Hypothese aufgestellt (Schritt **501**) und durch die Starrkörpertransformation T dargestellt, die Punkte in dem anzeigebasierten Koordinatensystem auf das kamerabasierte Koordinatensystem abbildet. Die Informationen von Schritt **413** bezüglich des Verlaufs der Position und/oder Orientierung der Kamera relativ zum dem Anzeigebildschirm **123** liefern wichtige Anhaltspunkte und können bei der Aufstellung der Hypothese verwendet werden. Zum Beispiel kann die letzte Position des Cursors in der anfänglichen Hypothese verwendet werden. Die "Rate der Positionsänderung" oder Geschwindigkeit über die letzten zwei Positionen kann die Schätzung verbessern, und die Beschleunigung kann die Schätzung noch weiter verbessern. Wenn keine geeignete Hypothese existiert kann eine Standardbeziehung für senkrechte Orientierung zum Anzeigebildschirm als Startpunkt angenommen werden.

**[0077]** Ein Farbband  $c$  wird aus einem der verfügbaren Farbbänder ausgewählt (Schritt **503**). Dann werden alle Werte in einem Feld  $I_{c_{test}}$  auf Null gesetzt und das Pixel  $(0, 0)$ , das auf dem Anzeigebildschirm **123** angezeigt wird, wird ausgewählt (Schritt **505**). Die Koordinaten  $(x_s, y_s)$  des Punktes, der auf dieses Pixel abgebildet wird, werden dann mit Beziehung **1** berechnet (Schritt **506**), und dieser Punkt wird dann zuerst mit Beziehung **3** von dem anzeigebasierten Koordinatensystem in das kamerabasierte Koordinatensystem transformiert (Schritt **507**) und nachfolgend mit den Gleichungen **4** und **5** von dem kamerabasierten Koordinatensystem in das bildbasierte Koordinatensystem (Schritt **509**). Danach werden diese bildbasierten Koordinaten mit Beziehung **2** in Bildpixelpositionen  $(x_{ip}, y_{ip})$  umgewandelt (Schritt **510**). Es kann erforderlich sein, die Bildpixelpositionen auf die nächste ganze Zahl zu runden (Schritt **511**). Dann wird der Abtastwert der Anzeigeeintensitätsfunktion  $S_c(x_{sp}, y_{sp})$  aus dem Feld  $S_{c_{sample}}$  erfasst (Schritt **513**). Dieser Wert wird an der Position  $(x_{ip}, y_{ip})$  zu dem laufenden Mittelwert in dem Feld  $I_{c_{test}}$  addiert, in dem die vorhergesagten Bildintensitätswerte gespeichert sind. Ein laufender Mittelwert wird gebraucht, da das Feld  $I_{c_{test}}$  die gleiche Größe wie  $I_{c_{sample}}$  hat und folglich kleiner als  $S_{c_{sample}}$  ist. Der Wert, der in irgendeiner Position in  $I_{c_{test}}$  gespeichert ist, kann deshalb der Mittelwert einer Anzahl von Anzeigepixeln sein, die auf die Bildpixel abgebildet werden, die von der Position dargestellt werden.

**[0078]** Es werden Tests ausgeführt (Schritte **517** und **519**), um zu überprüfen, ob der Intensitätsbeitrag aller Anzeigepixel in  $I_{c_{test}}$  berücksichtigt wurde. Wenn nicht, werden die Schritte **506** bis **515** für das nächste Anzeigepixel wiederholt. Wenn alle Anzeigepixel berücksichtigt wurden, dann wird eine weitere Prüfung durchgeführt (Schritt **521**), um festzustellen, ob  $I_{c_{test}}$  für alle Farbbänder  $c$  geschätzt wurde oder nicht. Wenn nicht, werden die Schritte **503** bis **519** für das nächste Farbband wiederholt.

**[0079]** Das Testbild, das wie oben beschrieben erzeugt und in dem Feld  $I_{c_{test}}$  gespeichert wurde, und das Kamerabild, das von der Kamera erfasst und in dem Feld  $I_{c_{sample}}$  gespeichert ist, können verglichen werden, um eine Zahl mit einem Real-Wert zu erzeugen, die die Diskrepanz zwischen den Bildern angibt. Diese Diskrepanz kann durch iteratives Prüfen nach irgendwelchen wohlbekannten Fehlerminimierungsalgorithmen minimiert werden.

**[0080]** Mit Bezug auf [Fig. 5b](#) wird nun die Prozedur, das Testbild und das Kamerabild zu vergleichen, um die Diskrepanz zu minimieren (und folglich die Transformationsparameter zu bestimmen), nun detaillierter beschrieben.

**[0081]** Zu Beginn wird der Fehler auf Null gesetzt, und der Index Anzfenster wird initialisiert (z. B. auf Null) (Schritt **530**). Dann wird ein Farbband  $c$  aus einem der verfügbaren Farbbänder ausgewählt (Schritt **531**). Nun werden mit Bezug auf [Fig. 6c](#) das reale Bild (und ähnlich, aber nicht gezeigt, das Testbild) in Fenster **580** mit der Größe von  $w \times w$  Pixeln aufgeteilt. Jedes Filterfenster überlagert das benachbarte Filterfenster um  $s$  Pixel sowohl in die horizontale als auch die vertikale Richtung, und auf  $s$  wird sich als die Filterfensterschrittweite bezogen. Schmalere Filterfenster liefern eine genauere Lösung, erfordern jedoch mehr Verarbeitung. Gute Kompromisswerte sind  $w = 128$  und  $s = 32$ .

**[0082]** Nochmals mit Bezug auf [Fig. 5b](#) werden die Indizes  $u$  und  $v$  initialisiert (z. B. auf Null). Dann wird die Summe der Intensitätsbeiträge, die in  $I_{c_{sample}}$  gespeichert sind, von allen  $w \times w$  Pixeln in dem ersten Filterfenster für reale Bilder berechnet, und als eine Variable  $I_{c_{samplesum}}$  gespeichert (Schritt **535**), und die Summe der Intensitätsbeiträge, die in  $I_{c_{test}}$  gespeichert sind, aller  $w \times w$  Pixel in dem ersten Filterfenster für Testbilder wird berechnet und als Variable  $I_{c_{testsum}}$  gespeichert (Schritt **537**).

**[0083]** Eine laufende Gesamtsumme des Quadrierten\_Fehlers wird mit der Formel

$$\text{Quadrierter\_Fehler} = \text{Quadrierter\_Fehler} + (I_{\text{csamplesum}} - I_{\text{ctestsum}})^2$$

berechnet (Schritt **539**), und dann wird der Index Anzfenster um eins weitergesetzt (Schritt **541**). Danach werden Prüfungen ausgeführt (Schritte **543** und **545**), um festzustellen, ob der Beitrag von allen Pixeln in den zwei Pixelbildern (dem Kamerabild und dem Testbild) in der Fehlerberechnung berücksichtigt wurde. Wenn nicht, werden die Schritte **535** bis **541** für das nächste Filterfenster wiederholt. Sonst wird eine weitere Prüfung ausgeführt (Schritt **547**), um festzustellen, ob alle Farbbänder  $c$  berücksichtigt wurden oder nicht. Wenn nicht, werden die Schritte **531** bis **545** wiederholt. Wenn alle Farbbänder berücksichtigt wurden, wird der mittlere Fehler aus der laufenden Gesamtsumme des Fehlers und dem Index Anzfenster berechnet (Schritt **549**), sodass:

$$\text{Mittlerer\_quadrierter\_Fehler} = \text{Quadrierter\_Fehler}/\text{Anzfenster}$$

**[0084]** Dann wird die Wurzel aus dem Mittleren quadrierten Fehler berechnet (Schritt **551**). Der RMS (Root Mean Squared\_Error)-Fehler stellt die Diskrepanz zwischen den zwei Bildern für die Transformation dar, die in Schritt **501** geschätzt wurde. Dieser RMS-Fehler wird dann zusammen mit anderen RMS-Fehlern, die die Diskrepanz zwischen zwei Bildern für eine andere Transformation  $T$  darstellen, iterativ entsprechend irgendwelchen wohlbekanntem Fehlerminimierungsalgorithmen überprüft (z. B. Gradientensteigung), um festzustellen, welche Transformation  $T$  die minimale Diskrepanz ergibt.

**[0085]** Wenn der minimale RMS-Fehler gefunden wurde, wird für die (bekannten) Transformationsparameter (einschließlich der Transformation  $T$ ), die das Anzeigebild in dem anzeigebasierten Koordinatensystem auf das Testbild in dem bildbasierten Koordinatensystem abbilden, angenommen, dass sie die (unbekannten) Transformationsparameter sind, die das Anzeigebild in dem anzeigebasierten Koordinatensystem auf das Kamerabild in dem bildbasierten Koordinatensystem abbilden. Folglich wird die 3D-Beziehung abgeleitet, die die Position und/oder Orientierung der Kamera **207** (und folglich der Zeigeeinrichtung **103**) relativ zu dem Anzeigebildschirm **123** ([Fig. 4](#), Schritt **407**) definiert.

**[0086]** Wenn jedoch in dieser Stufe der Fehler nicht minimiert wurde, wird die Hypothese der Transformation  $T$  verfeinert und entsprechend der vorangegangenen Prozedur erneut geprüft.

**[0087]** In einer bevorzugten Ausführung wird statt des Beginnens mit einem Schätzwert für die Transformation  $T$  wie in Schritt **501** in [Fig. 5a](#) ein Versuch gemacht, vier Punkte, die in dem Anzeigebild sichtbar sind, mit vier entsprechenden Punkte in dem Kamerabild in Übereinstimmung zu bringen. Da vier Beziehungen der Entsprechungen 3, 4 und 5 oben sich zu einem System mit 8 Gleichungen und 8 Unbekannten vereinfachen, ermöglicht die Kenntnis dieser vier entsprechenden Punkte, eine exakte Lösung für die Transformation zu finden.

**[0088]** Der Prozess, entsprechende Punkte in Übereinstimmung zu bringen und eine Lösung für die Transformation  $T$  zu finden, wird nun beschrieben. In dieser bevorzugten Ausführung werden Bereiche mit homogener Farbe identifiziert, um entsprechende Punkte an den Rändern dieser Bereiche zu finden. Fachleute erkennen jedoch, dass solche Bereiche nur ein Beispiel für ein Merkmal sind, das in dem Kamerabild erfasst werden kann. Alle anderen Merkmale, die in dem Kamerabild trotz der Transformation, die das Anzeigebild durchläuft, wenn es von der Kamera erfasst wird, detektiert werden können, können verwendet werden.

**[0089]** Die erste Stufe in dem Prozess ist, jedes Anzeigepixel und jedes Bildpixel auf Basis des Farbwertes des Pixels, der, wie man sich erinnert, ein Abtastwert der Farbfunktion an diesem Pixel ist, exakt einem identifizierbaren Bereich zuzuordnen. An dieser Stelle lohnt es sich, zwei zusätzliche Felder zu definieren:  $S_{\text{region}}$  und  $I_{\text{region}}$ , die sich auf die Anzeigepixel beziehungsweise Bildpixel beziehen. Diese 2D-Felder speichern die Identität des Bereichs, dem Pixel zugeordnet sind (oder speichern eine Angabe, dass Pixel noch einem Bereich zugeordnet werden müssen). Zum Beispiel gibt der Bereichsbezeichner  $R1$ , der an der Position (34, 78) in dem Feld  $S_{\text{region}}$  gespeichert ist, an, dass das Anzeigepixel (34, 78) dem Bereich  $R1$  zugeordnet ist. Als ein weiteres Beispiel gibt 0 (Null), gespeichert an der Position (10, 23) in dem Feld  $I_{\text{region}}$ , an, dass das Bildpixel (10, 23) unzugeordnet bleibt.

**[0090]** Die Zuordnung der Anzeigepixel zu Bereichen wird wie folgt erreicht. Ein Anzeigepixel  $(x_{\text{sp}}, y_{\text{sp}})$ , das noch einem Bereich zugeordnet werden muss, wird zu einem Bereich  $R1$  zugeordnet, wobei  $R1$  ein eindeutiger Bezeichner für diese neue Region ist. Der Bereichsbezeichner  $R1$  wird deshalb in dem Feld  $S_{\text{region}}$  an der Position  $(x_{\text{sp}}, y_{\text{sp}})$  gespeichert. Der Farbwert für Pixel  $(x_{\text{sp}}, y_{\text{sp}})$  wird dann aus dem Feld  $S_{\text{colour}}$  gelesen. Die Nachbarn von Pixel  $(x_{\text{sp}}, y_{\text{sp}})$  (d. h., die Pixel  $(x_{\text{sp}} - 1, y_{\text{sp}})$ ,  $(x_{\text{sp}} + 1, y_{\text{sp}})$ ,  $(x_{\text{sp}}, y_{\text{sp}} - 1)$ ,  $(x_{\text{sp}}, y_{\text{sp}} + 1)$ ) werden dann auch

der Region R 1 zugeordnet (unter der Annahme, dass sie nicht schon einem anderen Bereich zugeordnet wurden), wenn ihre Farbwerte die gleichen wie die (oder innerhalb eines vorher festgelegten Grenzwertes der) Farbwerte des Pixels  $(x_{sp}, y_{sp})$  sind. Wenn irgend einer der Nachbarn von Pixel  $(x_{sp}, y_{sp})$  einen Farbwert hat, der diese Kriterien nicht erfüllt, wird er unzugeordnet belassen. Dieser Prozess, unzugeordnete Nachbarpixel zu überprüfen und die entweder dem Bereich R1 zuzuordnen, oder sie unzugeordnet zu belassen, wird fortgesetzt, bis die Grenzen des Bereichs R1 gefunden wurden. Dann wird eines der Pixel, das unzugeordnet belassen wurde (vorzugsweise das nächste unzugeordnete Pixel zu Pixel  $(x_{sp}, y_{sp})$ ), einem zweiten Bereich R2 zugeordnet, und der Prozess wird wiederholt. Schließlich ist jedes Anzeigepixel genau einem identifizierbaren Bereich zugeordnet.

**[0091]** Indem die Anzahl des Auftretens jedes Bereichsbezeichners (z. B. R1, R2 ...) in dem Feld  $S_{region}$  gezählt wird, ist es möglich, den größten Bereich in dem Bild zu identifizieren. Es ist dann möglich, alle Pixel an der Grenze dieses größten Bereichs zu identifizieren. Dies kann erreicht werden, indem alle Pixel identifiziert werden, die dem größten Bereich zugeordnet sind, die wenigstens ein Nachbarpixel haben, das einem anderen Bereich zugeordnet wurde. Die Krümmung der Grenze an jedem dieser Grenzpixel (d. h., die Änderungsrate der Steigung der Grenze an jedem dieser Pixelorte) kann dann berechnet werden, und die vier Ecken des Bereichs (d. h., Punkte mit starker Krümmung, an denen die Krümmung oberhalb eines vorher festgelegten Grenzwertes liegt) können identifiziert werden. Des Ergebnis ist ein Satz von eindeutigen, identifizierbaren Punkten: die vier Ecken des größten Bereiches im Anzeigebild.

**[0092]** Die Zuordnung der Bildpixel zu Bereichen und die Ableitung der Eckpunkte des größten Bereiches in dem Kamerabild wird auf ähnliche Weise wie die erreicht, die oben in Bezug auf die Anzeigepixel beschrieben wurde.

**[0093]** Die nächste Stufe in dem Prozess ist, die vier Ecken des größten Bereiches in dem Anzeigebild mit den vier entsprechenden Ecken des größten Bereiches in dem Kamerabild in Übereinstimmung zu bringen. Mit den Pixelkoordinaten dieser vier sich entsprechenden Paare von Punkten kann dann eine Lösung für die Transformation T gefunden werden, indem die Gleichungen gelöst werden, die sich aus der Vereinfachung der Gleichungen 3, 4 und 5 oben ergeben.

**[0094]** Die obige bevorzugte Ausführung nutzt einen Prozess, der nach dem Stand der Technik als Merkmalsvergleich bekannt ist. Der oben beschriebene Prozess ist ein einfaches Beispiel eines solchen Prozesses, und Fachleute erkennen, dass ein hochentwickelterer Ansatz möglich ist.

**[0095]** Vorausgesetzt, dass die Transformation T im allgemeinen jeden Punkt in dem anzeigebasierten Koordinatensystem in einen Punkt in dem kamerabasierten Koordinatensystem übersetzen kann, kann sie dadurch anzeigebasierte Koordinaten der Ecken des Anzeigebildschirms in Positionen in kamerabasierte Koordinaten übersetzen ([Fig. 6](#), Schritte **601** und **603**).

**[0096]** Dann wird eine Prüfung durchgeführt (Schritt **605**), um zu überprüfen, ob die Koordinaten von 3 Ecken des Anzeigebildschirms in den Kameraraum übersetzt wurden. Wenn nicht, werden die Schritte **601** bis **603** wiederholt.

**[0097]** Drei Punkte in einem Euklidischen Raum definieren eine Ebene und wenn folglich einmal die Koordinaten von 3 Ecken des Anzeigebildschirms in kamerabasierte Koordinaten übersetzt wurden, kann die Formel für die Ebene des Anzeigebildschirms  $z_s = 0$  in kamerabasierten Koordinaten bestimmt werden (Schritt **607**). Danach kann der Schnittpunkt zwischen der Linie  $x_c = 0, y_c = 0$  (der Achse des Sichtbereichs der Kamera, die aus dem Zentrum der Apertur der Kamera **209** in  $[0, 0, 0]$  entlang der  $z_c$ -Achse herausläuft) und der Ebene des Anzeigebildschirms berechnet werden (Schritt **609** und [Fig. 4](#), Schritt **409**). Da Starrkörpertransformationen reversibel sind, wird dieser Punkt dann in anzeigebasierte Koordinaten zurückübersetzt (Schritt **611**), was die Position bestimmt, die der Cursor in dem anzeigebasierten Koordinatensystem einnehmen soll. Diese hat die Form  $(x_s, y_s, 0)$ , da die  $z_s$ -Komponente jedes Punktes auf der Ebene des Anzeigebildschirms  $z_s = 0$  ist. Danach werden diese anzeigebasierten Koordinaten mit der Beziehung 1 in Pixelpositionen des Anzeigebildschirms  $(x_{sp}, y_{sp})$  umgewandelt (Schritt **612**). Es kann erforderlich sein, die Positionen der Pixel des Anzeigebildschirms auf den nächsten ganzen Wert zu runden (Schritt **613**). Schließlich steuert die Software den Computer, der den Cursor auf  $(x_s, y_s, 0)$  positioniert (Schritt **615** und [Fig. 4](#), Schritt **411**).

**[0098]** Dies schließt eine einzelne Interaktion des Berechnungsalgorithmus zur Berechnung des Anzeigeportes ab, der iteriert wird, um eine feine Steuerung des Cursors durch die Bewegung der Zeigeeinrichtung zu gewährleisten.

**[0099]** Noch einmal mit Bezug auf [Fig. 1](#) zeigt der Benutzer **101** mit der Zeigeeinrichtung **103** derart, dass seine Längsachse sich mit dem Anzeigebildschirm **123** schneidet. Der Benutzer **101** kann dann den Cursor steuern, der auf dem Anzeigebildschirm erscheint, indem die Zeigeeinrichtung **103** derart bewegt wird, dass er auf den Teil des Anzeigebildschirms zeigt, an dem der Benutzer den Cursor wünscht. Der Benutzer **101** kann die Knöpfe **211a** und **213a** verwenden, um mit den Ausgaben des Rechners **111**, die auf dem Anzeigebildschirm **123** angezeigt werden, zu interagieren. Auf diese Weise wird eine Schnittstellenanordnung zwischen dem Benutzer und dem Rechner bereitgestellt. Zum Beispiel kann der Benutzer auf verschiedene Icons in der Anzeige zeigen. Er kann speziellen Text in einem Dokument hervorheben. Er kann dann sogar die Erscheinung von Text verändern, indem er ihn unterstreicht, ihn fett oder kursiv macht. Er kann Klappmenüs bedienen, um verschiedene Programme aufzurufen oder Dateien anzusehen. Er kann Dias in einer digitalen Dia-Präsentation wechseln. Er kann auf Hyperlinks in einem Internet- oder Intranet-Dokument klicken. Fachleuten ist klar, dass dem Benutzer durch eine solche Zeigeeinrichtung viel mehr Funktionen zur Verfügung stehen.

**[0100]** Durch die vorangegangene Beschreibung wird offensichtlich, dass viele Modifikationen oder Variationen mit den oben beschriebenen Ausführungen gemacht werden können, ohne von der Erfindung abzuweichen. Solche Modifikationen am Variationen schließen ein:

Für ein 1024 × 768, 24-Bit-Echtfarbenbild auf dem Anzeigebildschirm wird der Speicher mit nahezu 2,4MB belegt. Die Verarbeitung aller dieser Daten ist intensiv. Um die Verarbeitung des Bildes auf dem Anzeigebildschirm zu optimieren, kann es nicht in jedem Zyklus des Berechnungsalgorithmus für den Anzeigeort aktualisiert werden. Die vorverarbeiteten Daten aus dem letzten Zyklus sind oft ausreichend, und dies vermeidet intensive Berechnungen, die erforderlich sind, um den großen Satz von Abtastwerten zu verarbeiten. Darüber hinaus wird angemerkt, dass die Auflösung des Bildes, das von der Kamera aufgenommen wird, in erheblichem Maße geringer als die Auflösung des Anzeigebildschirms ist. Dieser Unterschied erlaubt die Vorverarbeitung der Bilddaten des Anzeigebildschirms auf eine verringerte Auflösung, um die Anzahl von Abtastwerten zu minimieren, die später in dem Algorithmus bearbeitet werden müssen, ohne dass ein wesentlicher Verlust von Genauigkeit eintritt. Diese Vorverarbeitung kann durch irgend eine wohlbekannt Methode wie etwa bilineare Interpolation durchgeführt werden. Um die Verarbeitungszeit noch weiter zu verringern, kann die Zeigeeinrichtung deaktiviert werden, wenn sie nicht in Richtung des Anzeigebildschirms zeigt. Dies kann durchgeführt werden, indem zu dem Gehäuse der Zeigeeinrichtung ein weiterer Knopf hinzugefügt wird, der gedrückt werden muss, wenn Betrieb der Einrichtung erforderlich ist, und losgelassen werden kann, um die Einrichtung zu deaktivieren.

**[0101]** Ein weiterer Weg, um die Anforderungen an die Verarbeitung zu verringern, ist, Abtastwerte aus dem Bildschirm heranzuziehen, statt die gesamten Informationen des Bildschirms heranzuziehen. Dies ist sowohl auf die Erzeugung des Testbildes (das heißt, nur eine Untermenge der verfügbaren Pixel aus der Anzeige aufzunehmen, und sie auf das Testbild abzubilden), und bezüglich der Prüfungen auf die Korrelation mit dem realen Bild anwendbar (wobei nur versucht wird, eine Untermenge der Pixel zu korrelieren, da die anderen leer sind – da sie zuerst nicht abgebildet wurden).

**[0102]** Wenn die Transformation für eine sehr grundlegende Position und Orientierung der Zeigeeinrichtung relativ zum Bildschirm bekannt war, dann kann in den Stufen zu Beginn des Betriebs die Zeigeeinrichtung in dieser Position gehalten werden, um die Berechnungseinrichtung mit einer genauen Anfangshypothese für die Transformation zu versorgen, und folglich die Anforderungen an die Verarbeitung noch weiter zu verringern. Eine solche Kalibrierungsposition kann z. B. senkrecht auf ein ungenutztes Zentrum der Anzeige in einem Abstand von 300 mm sein. In anderen Ausführungen können bekannte künstliche und unterscheidbare Merkmale (z. B. eine bekannte Form in einer bekannten Farbe) in die Anzeige eingeblendet werden, um die Kalibrierung zu unterstützen. Dies ist besonders nützlich, wenn die Zeigeeinrichtung verwendet wird, um auf komplexe Anzeigedaten zu zeigen, wie etwa Fernsehbilder.

**[0103]** Obwohl in den oben beschriebenen Ausführungen die Zeigeeinrichtung verwendet wurde, um einen Cursor in zwei Dimensionen zu steuern, ist es möglich, einen 3D-Zeiger zu erschaffen, da die 3D-Position der Kamera relativ zu dem Anzeigebildschirm und die Orientierung der Kameraachse bekannt ist. Der 3D-Zeiger kann auf solche Weise arbeiten, dass jede Bewegung der Zeigeeinrichtung in einem 3D-Modell aufgezeichnet werden kann, um 3D-Schnittstellen zu steuern, wie etwa die, die in Programmpaketen für rechnergestützte Konstruktion (CAD, computer aided design) zu finden sind.

**[0104]** Obwohl in den oben beschriebenen Ausführungen ein Cursor an dem berechneten Anzeigeort eingeblendet wurde, ist es auch eine Möglichkeit, keinen Cursor einzublenden. Ein solcher Umstand kann z. B. in einem Programmpaket für rechnergestütztes Konstruieren (CAD, computer aided design) entstehen, wenn der Benutzer mit der Einrichtung auf die Anzeige zeigt und die Einrichtung bewegt, um die perspektivische Ansicht

des Objektes zu verändern, das konstruiert/gezeichnet wird. Zum Beispiel kann ein Benutzer wünschen, von unten auf das Objekt zu sehen, und indem er die Einrichtung in Abwärtsrichtung bewegt, kann er die Ansicht in eine drehen, die gewünscht wird.

**[0105]** Obwohl in den oben beschriebenen Ausführungen die Position und Orientierung der Zeigeeinrichtung relativ zur Anzeige berechnet wurde, ist es auch möglich, dass nur die Position oder die Orientierung berechnet werden. Im Fall der Position kann der Benutzer der Zeigeeinrichtung die Einrichtung zu allen Zeiten senkrecht auf die Anzeige halten, und deshalb ändert sich die Orientierung nicht. Nur die Position ändert sich. Im Fall der Orientierung kann der Benutzer die Einrichtung in derselben Position halten, und bloß die Richtung ändern, in die sie zeigt, und folglich nur ihre Orientierung verändern.

**[0106]** Obwohl in den oben beschriebenen Ausführungen die Position und Orientierung der Zeigeeinrichtung relativ zu der Anzeige verwendet wurde, um den Cursor zu steuern, ist es auch möglich, dass nur die Position oder die Orientierung verwendet wird. Wenn zum Beispiel im Fall eines CAD-Programmpakets ein Benutzer eine gerade Linie zeichnen möchte, ist es üblich, dass er nur an den zwei Endpunkten der Linie klicken muss. Ein Benutzer der Einrichtung kann an einem Endpunkt klicken und die Einrichtung bewegen, um den Cursor zu dem anderen Ende zu bewegen, und wieder klicken. Die Orientierung der Einrichtung wäre in diesem Fall nicht relevant. Die oben beschriebene Situation in Bezug auf das Ändern der perspektivischen Ansicht eines Objektes, das konstruiert/gezeichnet wird, ist ein Fall, in dem nur die Orientierung der Einrichtung relevant ist.

**[0107]** Obwohl die Anzeige in den oben beschriebenen Ausführungen von einem LCD-Bildschirm erzeugt wurde, ist es möglich, dass die Anzeige von einem LCD- oder digitalen Lichtverarbeitungs(DLP, digital light processing)-Projektor erzeugt und auf eine große Leinwand oder sogar auf eine geeignete Wand projiziert wird. Dies ist insbesondere vor dem Hintergrund eines Benutzers nützlich, der auf eine Anzeige zeigen muss, die von einem großen Publikum gesehen wird. Die Kamera nimmt immer noch ein Bild der Anzeige auf, wenn der Benutzer die Einrichtung auf die Leinwand oder die Wand richtet, und die Bilddaten, die auf diese Weise erfasst werden, können dann mit den Anzeigedaten, die von der Maschine erzeugt werden (an die der Projektor angeschlossen ist), verwendet werden, um den Schnittpunkt zu berechnen.

**[0108]** Obwohl die Zeigeeinrichtung in den oben beschriebenen Ausführungen verwendet wurde, um eine Maschine zu steuern, ist es auch möglich, sie bloß zu verwenden, um auf ein Objekt in der Anzeige zu zeigen. In diesem Fall wären keine Aktivierungseinrichtungen erforderlich. Dies ist wiederum vor dem Hintergrund eines Benutzers angebracht, der einem Publikum einer Präsentation gibt.

**[0109]** Obwohl in den Ausführungen oben eine digitale Videokamera verwendet wurde, ist es möglich, dass stattdessen eine Kamera verwendet wird, die eine Momentaufnahme der Anzeige aufnimmt. Die Momentaufnahme des Bildes würde die Bilddaten liefern, die zusammen mit den Anzeigedaten für die Bildregistrierung verwendet werden können. Die Momentaufnahme würde mit einer neuen Momentaufnahme sooft wie möglich aktualisiert, wobei alle Reserveverarbeitungsleistung verwendet wird.

**[0110]** Obwohl in den oben beschriebenen Ausführungen die Kamera eine manuelle Kamera mit festem Fokus war, ist es auch möglich, eine Kamera zu verwenden, die einen Autofokus-Mechanismus enthält. Der Abstand als Kriterium für den Betrieb kann dann entschärft werden, da die Kamera konstant einen adäquaten Fokus aufrechterhält. Dieses Abstandskriterium kann auch in dem Fall eines Objektivs mit festem Fokus entschärft werden, wenn irgendeine Einrichtung für die Eingabe der Vergrößerung des optischen Systems in den Algorithmus vorgesehen wird.

**[0111]** Obwohl die Zeigeeinrichtung in den oben beschriebenen Ausführungen nur eine Kamera verwendet, ist es möglich, mehr als eine Kamera zu verwenden. Auf diese Weise kann der Schnittpunkt der Achse von jeder Kamera und dem Bildschirm separat identifiziert werden, und die Kombination aller Kamerapositionen und -orientierungen als Eingabe für die Steuerung genommen werden. Zum Beispiel wäre es möglich, einen Handschuh mit einer Kamera in dem Ende jedes Fingers zu bauen. Folglich kann eine natürliche Form von physikalischer Bedienung angewendet werden, zum Beispiel Greifen und Fallenlassen, wobei die fünf Fingerspitzen auf dem Bildschirm verwendet werden.

**[0112]** In alternativen Ausführungen können die Zeigeeinrichtungen irgendeine Berechnungseinrichtung enthalten (in der Form einer CPU, Speicher und Speichereinrichtungen), sodass manche oder alle der Anforderungen an die Verarbeitung von der Berechnungseinrichtung innerhalb der Zeigeeinrichtung ausgeführt werden können. Zum Beispiel kann die Zeigeeinrichtung mit einer handgehaltenen Rechneinrichtung wie etwa einem persönlichen digitalen Assistenten (PDA, personal digital assistant) kombiniert werden.

**Patentansprüche**

1. Berechnungseinrichtung für den Anzeigort mit:  
einem Anzeigegerzeuger, der dazu eingerichtet ist, bei der Verwendung eine Anzeige als Reaktion auf Anzeigedaten zu erzeugen, die von einer Maschine (111) erzeugt werden;  
einer Zeigeeinrichtung (103), die eine Kamera (207) trägt, die betrieben werden kann, um Bilddaten zu erzeugen, die wenigstens einen Teil der Szene in dem Sichtbereich der Kamera darstellen, wobei der Teil ein Bild mit wenigstens einem Ausschnitt der Anzeige (123) umfasst;  
einer Berechnungseinrichtung, die bei der Verwendung dazu eingerichtet ist, die Bilddaten zu empfangen;  
**dadurch gekennzeichnet**, dass die Berechnungseinrichtung weiter dazu eingerichtet ist, bei der Verwendung die Anzeigedaten zu empfangen;  
die Position und/oder Orientierung der Zeigeeinrichtung relativ zu der Anzeige aus den Bilddaten und den Anzeigedaten zu berechnen;  
einen Anzeigort aus der berechneten Position und/oder Orientierung zu berechnen.
2. Berechnungseinrichtung für den Anzeigort nach Anspruch 1, bei der die Zeigeeinrichtung eine gestreckte Form hat, und bei der der Anzeigort der Punkt ist, an dem die Längsachse der Zeigeeinrichtung sich mit der Anzeige schneidet.
3. Berechnungseinrichtung für den Anzeigort nach Anspruch 1 oder 2, bei der die Anzeige an dem Anzeigort einen Cursor enthält.
4. Berechnungseinrichtung für den Anzeigort nach Anspruch 3, bei der die Position des Cursors entsprechend der berechneten Position und/oder Orientierung variiert.
5. Schnittstellenanordnung zur Bereitstellung einer Schnittstelle zwischen einem Benutzer und einer Maschine mit  
einer Berechnungseinrichtung für den Anzeigort nach einem der Ansprüche 2 bis 4;  
bei der die Berechnungseinrichtung weiter dazu eingerichtet ist, bei der Verwendung die Maschine entsprechend der Position des Cursors zu steuern.
6. Schnittstellenanordnung nach Anspruch 5, bei der die Maschine einen Rechner mit einem Prozessor umfasst, und bei der die Berechnungseinrichtung den Prozessor umfasst.
7. Schnittstellenanordnung nach Anspruch 5, bei der die Zeigeeinrichtung weiter einen Rechner mit einem Prozessor trägt, und bei der die Berechnungseinrichtung den Prozessor umfasst.
8. Schnittstellenanordnung nach Anspruch 5, bei der die Maschine einen ersten Rechner mit einem ersten Prozessor umfasst, und bei der die Zeigeeinrichtung weiter einen zweiten Rechner mit einem zweiten Prozessor trägt, und bei der die Berechnungseinrichtung den ersten Prozessor und den zweiten Prozessor umfasst.
9. Schnittstellenanordnung nach einem der Ansprüche 5 bis 8, bei der die Zeigeeinrichtung weiter wenigstens eine Aktivierungseinrichtung zur Verwendung bei der Steuerung der Maschine umfasst.
10. Schnittstellenanordnung nach einem der Ansprüche 5 bis 9, bei der der Anzeigegerzeuger einen Projektor umfasst.
11. Schnittstellenanordnung nach einem der Ansprüche 5 bis 10, bei der die Kamera eine digitale Videokamera ist.
12. Schnittstellenanordnung nach einem der Ansprüche 5 bis 11, bei der die Kamera eine Kamera mit konstanter Brennweite ist.
13. Schnittstellenanordnung nach Anspruch 12, die weiter eine Anzeigevorrichtung umfasst, die bei der Verwendung dazu eingerichtet ist, anzuzeigen, dass das von der Kamera erfasste Bild fokussiert ist.
14. Schnittstellenanordnung nach einem der Ansprüche 5 bis 13, bei der die Kamera eine feste räumliche Beziehung zu der Zeigeeinrichtung aufweist.

15. Verfahren zur Berechnung eines Anzeigeortes, wobei das Verfahren die folgenden Schritte umfasst:

- Erzeugen einer Anzeige entsprechend Anzeigedaten, die von einer Maschine (**111**) erzeugt werden;
- Erfassen von Bilddaten, die wenigstens einen Teil der Szene in dem Sichtbereich einer Kamera (**207**) darstellen, die von einer Zeigeeinrichtung (**103**) getragen wird, wobei wenigstens ein Ausschnitt der Anzeige in dem Sichtbereich enthalten ist;
- Berechnen der Position und/oder der Orientierung der Zeigeeinrichtung relativ zu der Anzeige aus den Bilddaten und den Anzeigedaten;
- Berechnen eines Anzeigeortes aus der berechneten Position und/oder Orientierung.

16. Verfahren nach Anspruch 15 mit dem zusätzlichen Schritt der Erzeugung eines Cursors am Anzeigeort in der Anzeige.

17. Verfahren nach Anspruch 16, bei dem die Position des Cursors entsprechend der berechneten Position und/oder Orientierung variiert.

18. Verfahren nach Anspruch 17, bei dem die Zeigeeinrichtung eine gestreckte Form hat, und bei dem der Anzeigeort der Punkt ist, an dem sich die Längsachse der Zeigeeinrichtung mit der Anzeige schneidet.

19. Verfahren nach einem der Ansprüche 15 bis 18 mit dem zusätzlichen Schritt des Speicherns von Daten, die den Verlauf der Position und/ oder der Orientierung der Zeigeeinrichtung relativ zu der Anzeige darstellen, und bei dem der Berechnungsschritt (iii) weiter diese Daten berücksichtigt.

20. Berechnungseinrichtung für den Anzeigeort mit:  
einem Speichermedium, auf dem prozessorlesbarer Code gespeichert ist, der verarbeitet werden kann, um eine Schnittstelle zwischen einem Benutzer und einer Maschine (**111**) bereitzustellen, wobei der Code folgendes umfasst:  
Kode für die Erfassung von Anzeigedaten, der verarbeitet werden kann, um Anzeigedaten zu erhalten, die eine Anzeige darstellen;  
Kode für die Erfassung von Bilddaten, der verarbeitet werden kann, um Bilddaten von einer Zeigeeinrichtung (**103**), die eine Kamera (**207**) trägt, zu erhalten, die wenigstens einen Teil einer Szene in dem Sichtfeld der Kamera darstellen,  
Kode für die Berechnung von Position/Orientierung, der verarbeitet werden kann, um aus den Anzeigedaten und den Bilddaten die Position und/oder die Orientierung der Zeigeeinrichtung relativ zu der Anzeige zu berechnen;  
Kode für die Berechnung des Anzeigeortes, der verarbeitet werden kann, um aus der berechneten Position und/oder der Orientierung einen Anzeigeort zu berechnen.

21. Digitaler Datenträger, der ein Programm aus Anweisungen trägt, die von einer Verarbeitungsvorrichtung ausgeführt werden können, um die Verfahrensschritte wie in einem der Ansprüche 15 bis 19 aufgestellt auszuführen.

Es folgen 10 Blatt Zeichnungen

Anhängende Zeichnungen

Fig.1

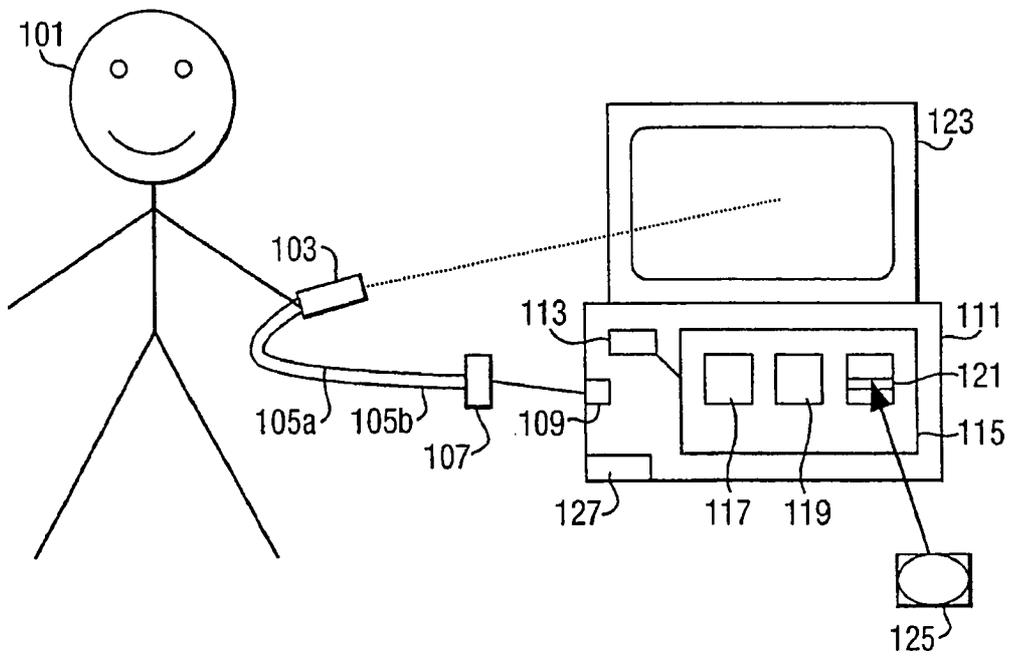


Fig.2a

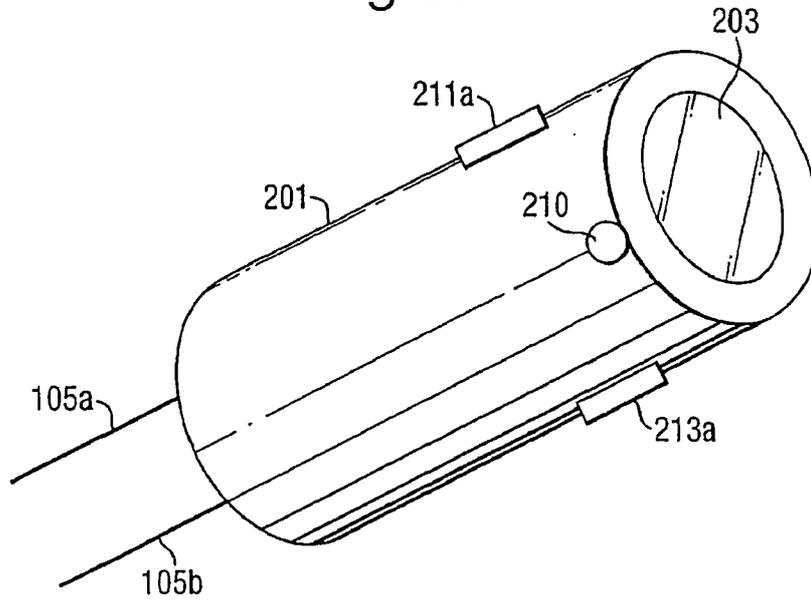


Fig.2b

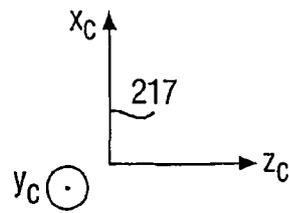
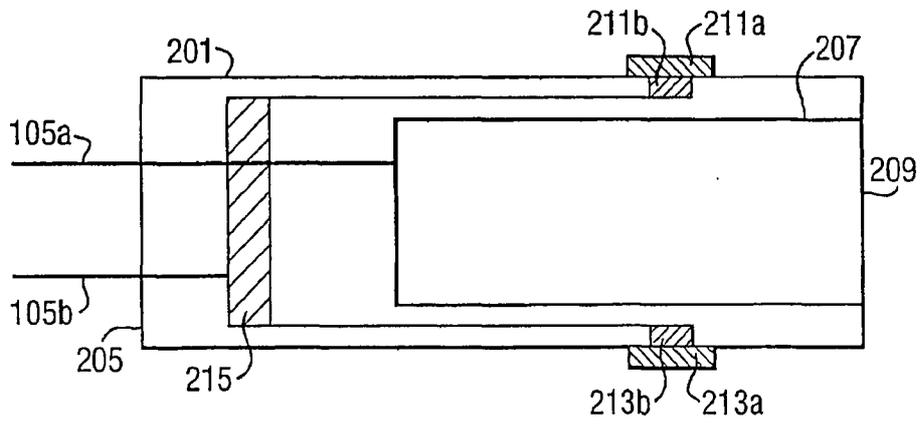


Fig.3a

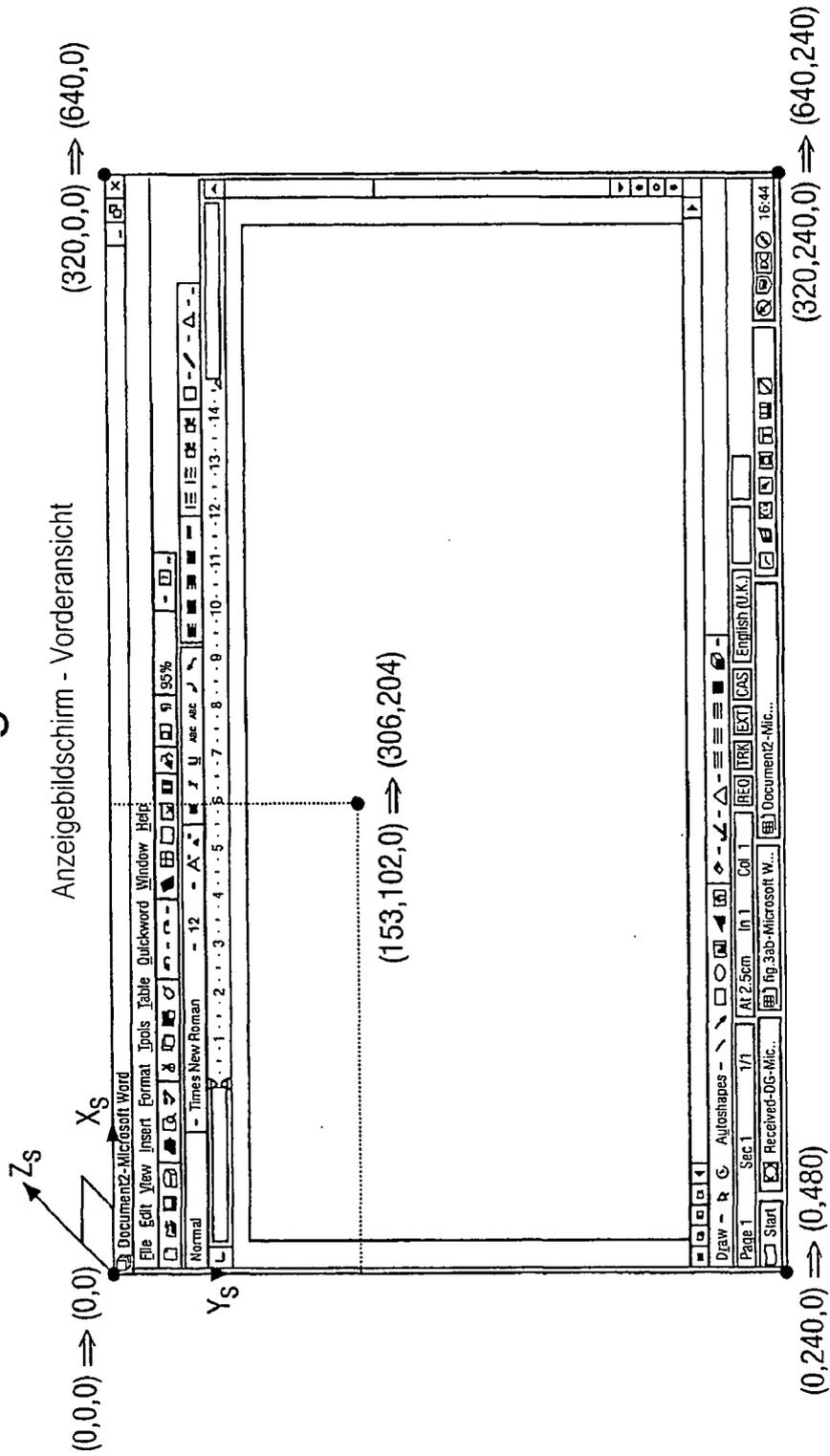


Fig.3b

CCD-Detektorfeld - von hinten gesehen

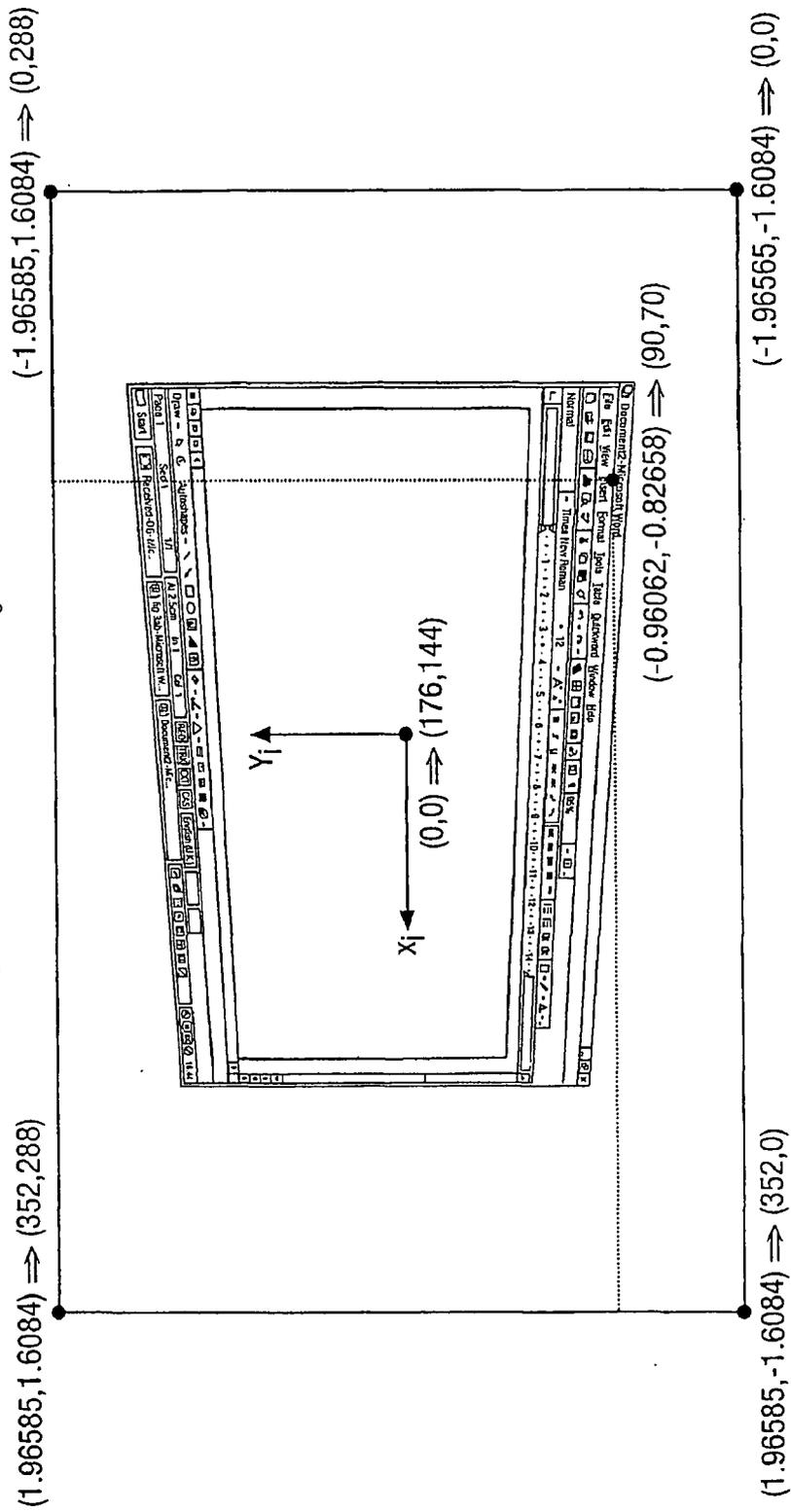




Fig.4

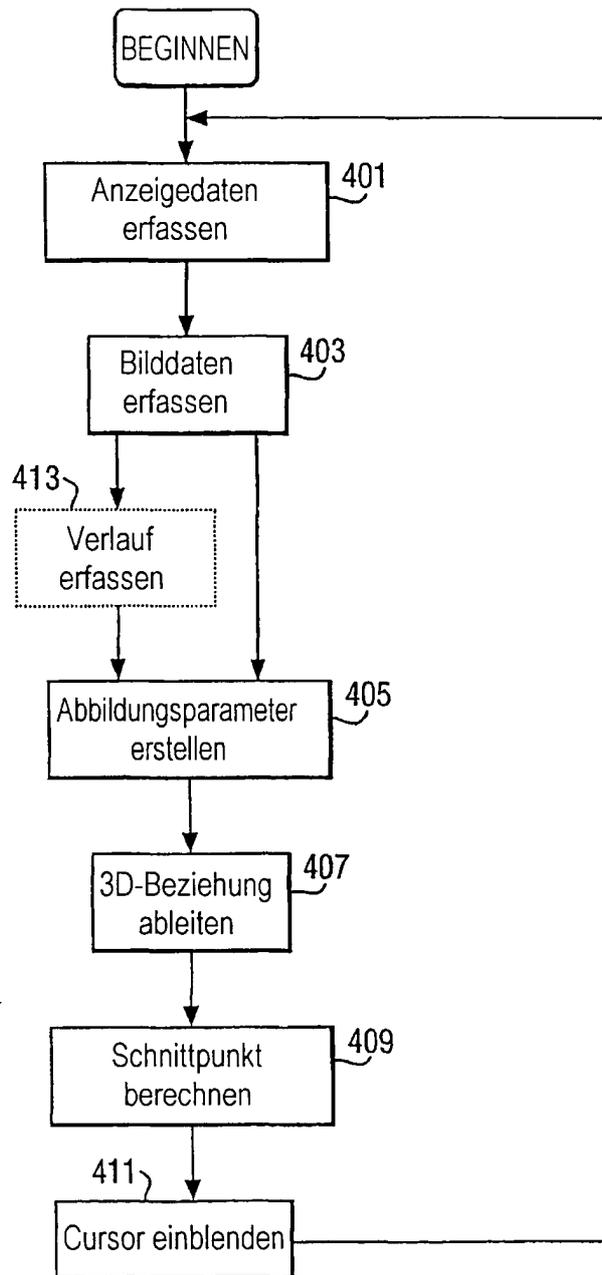


Fig.5a

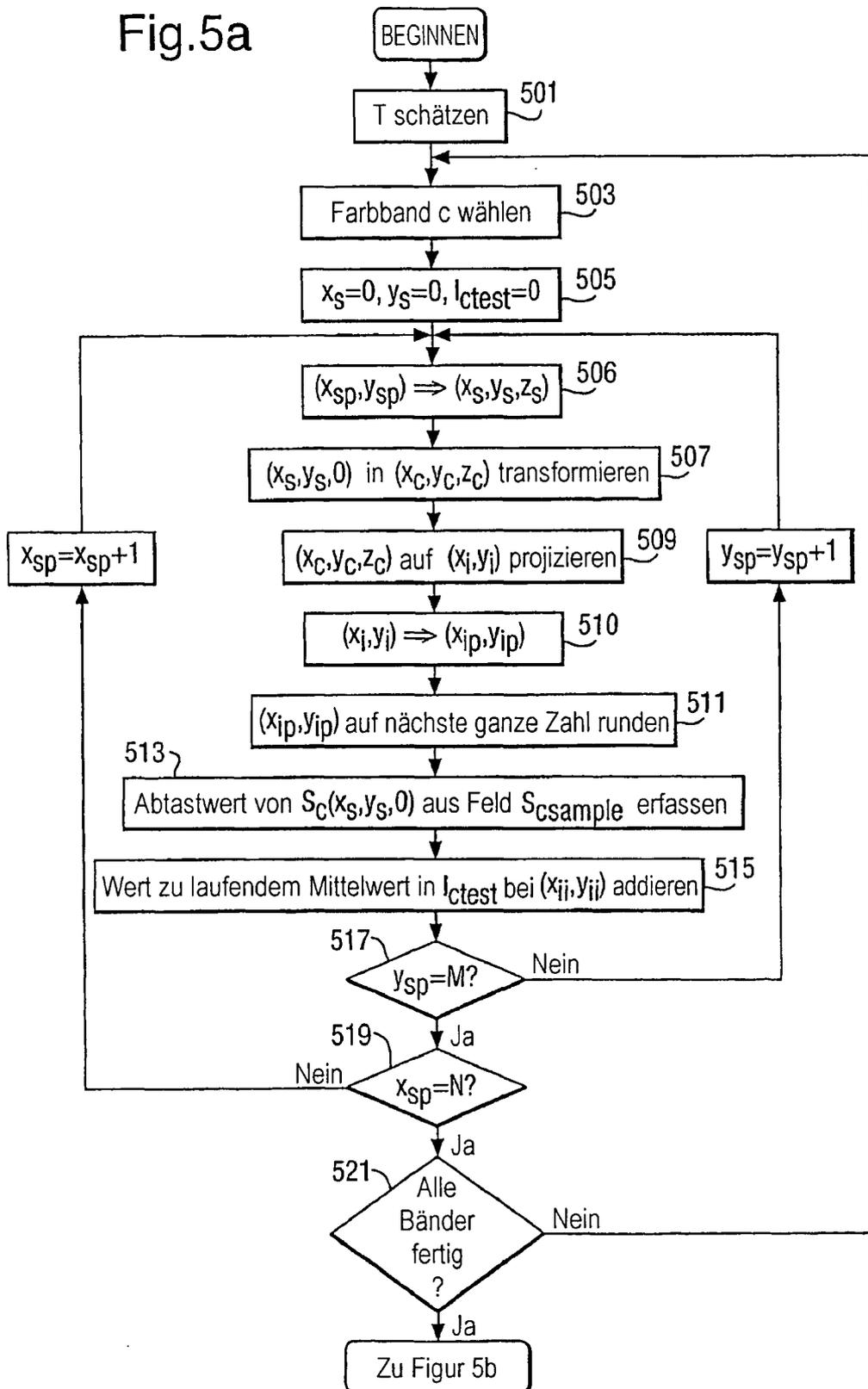


Fig.5b

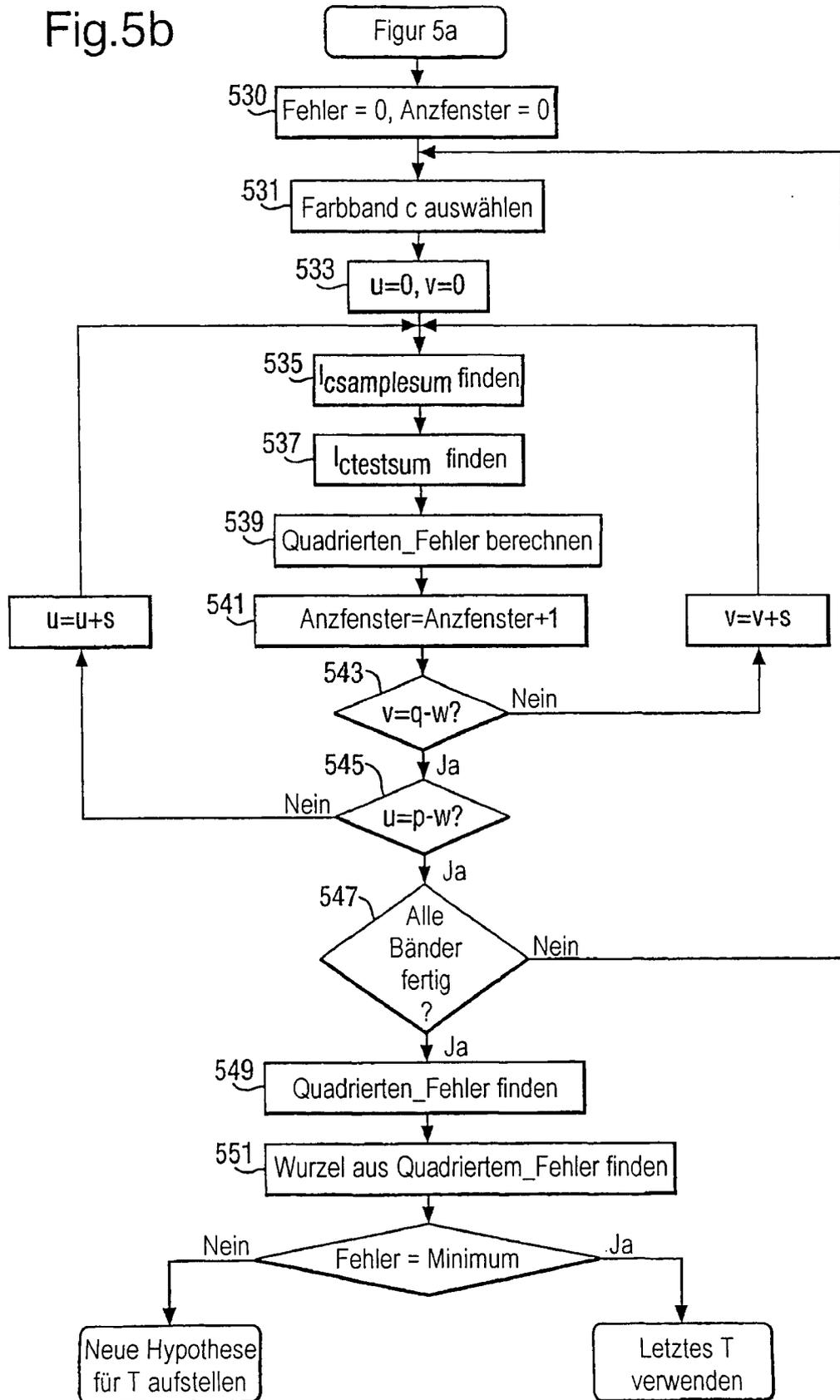


Fig.5c

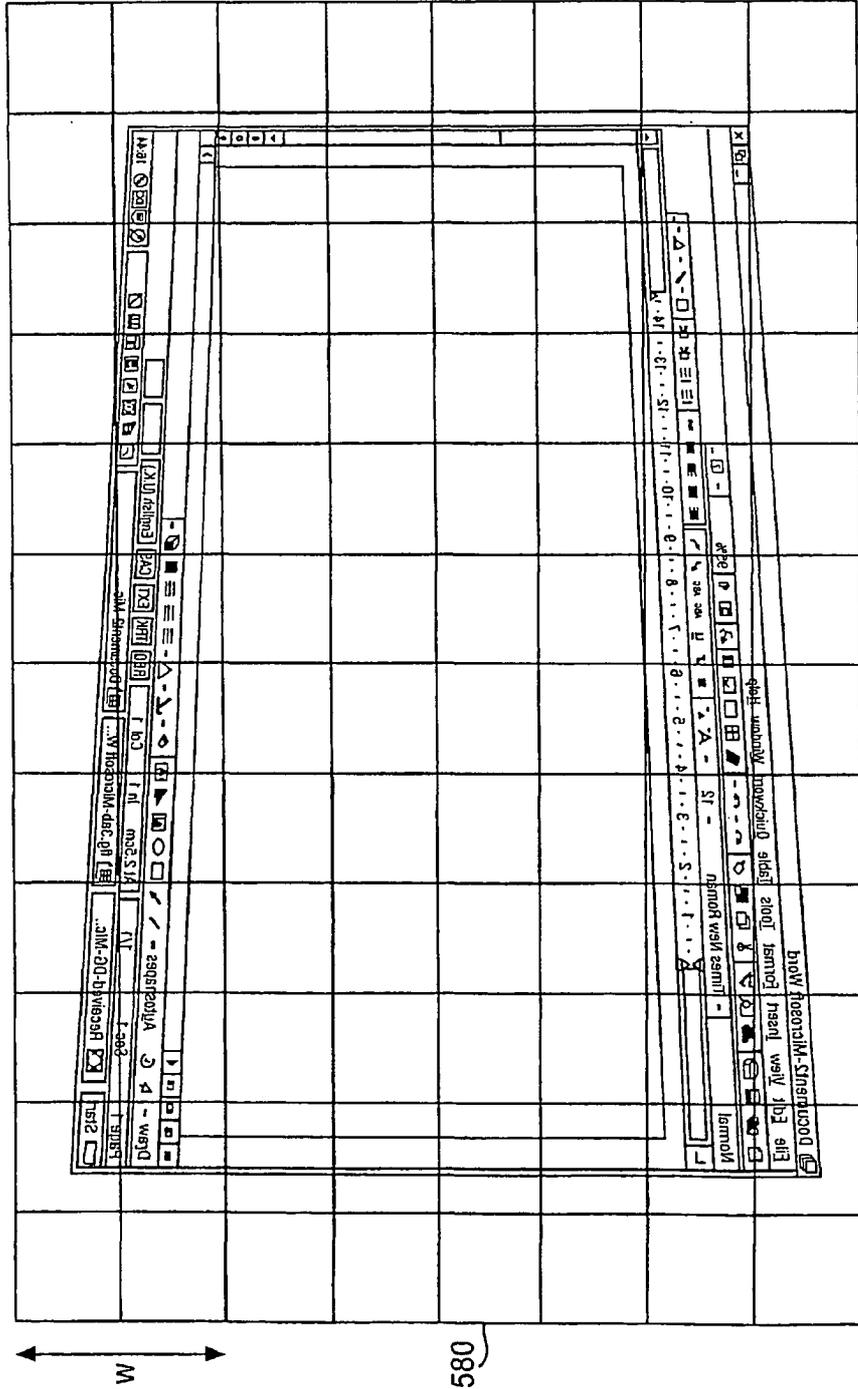


Fig.6

