



(19)
Bundesrepublik Deutschland
Deutsches Patent- und Markenamt

(10) **DE 100 84 525 B4** 2008.04.17

(12)

Patentschrift

(21) Deutsches Aktenzeichen: **100 84 525.8**
(86) PCT-Aktenzeichen: **PCT/US00/10732**
(87) PCT-Veröffentlichungs-Nr.: **WO 2000/067195**
(86) PCT-Anmeldetag: **21.04.2000**
(87) PCT-Veröffentlichungstag: **09.11.2000**
(43) Veröffentlichungstag der PCT Anmeldung
in deutscher Übersetzung: **25.07.2002**
(45) Veröffentlichungstag
der Patenterteilung: **17.04.2008**

(51) Int Cl.⁸: **G06K 9/22 (2006.01)**
G06K 9/03 (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. 2 Patentkostengesetz).

(30) Unionspriorität:
09/301,753 29.04.1999 US

(73) Patentinhaber:
Intel Corporation, Santa Clara, Calif., US

(74) Vertreter:
Zenz, Helber, Hosbach & Partner GbR, 45128 Essen

(72) Erfinder:
Tan, Yap-Peng, Singapore, SG; Acharya, Tinku, Tempe, Ariz., US; Metz, Werner, Chandler, Ariz., US

(56) Für die Beurteilung der Patentfähigkeit in Betracht
gezogene Druckschriften:
US 56 75 672 A
US 55 81 637 A
US 49 49 391 A
US 42 72 756 A

(54) Bezeichnung: **Verwendung einer elektronischen Kamera zum Aufbauen einer Text enthaltenden Datei**

(57) Hauptanspruch: Eine elektronische Kamera, aufweisend:

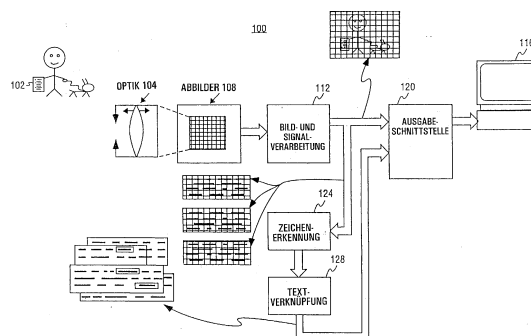
eine Optik mit einer gegenüber einem herkömmlichen Scanner erhöhten Tiefenschärfe zum Erzeugen einer Mehrzahl von Bildern eines Textes anzeigenden räumlichen Objekts;

einen Kameraabbilder, der die Bilder aufnimmt und in Abhängigkeit davon eine Mehrzahl von Bildausschnitten zur Verfügung stellt, die jeweils verschiedene Abschnitte des Textes des Objekts wiedergeben;

eine optische Zeichenerkennungsmaschine, die eine Mehrzahl von Text enthaltenden Textausschnitten, die jeweils aus einem zugehörigen Bildausschnitt erkannt worden sind, zur Verfügung stellt;

einen Textverknüpfer, der in den Textausschnitten nach überlappendem Text sucht; und

eine Einrichtung zum Korrigieren fehlerhaften Textes in der Mehrzahl von Textausschnitten, die eine Neuaufnahme eines Abschnitts des Text anzeigenden Objekts veranlaßt, der dem fehlerhaften Text entspricht, wobei aus dem Bildausschnitt der Neuaufnahme ein entsprechender Textausschnitt erzeugt wird, und die ein Ersetzen des fehlerhaften Textes durch die korrekte Textdaten der erzeugten Textausschnitts bewirkt.



Beschreibung

[0001] Diese Erfindung ist allgemein auf elektronische Kameras und insbesondere auf das Aufnehmen und Aufbauen einer Textdatei unter Verwendung einer elektronischen Kamera gerichtet.

[0002] Elektronische Kameras, wie beispielsweise Digitalkameras und Videokameras, sind populäre Konsumgüter. Die elektronische Kamera weist eine spezielle Optik und eine elektronische Abbildungsschaltung auf, die zusammenwirken, um dreidimensionale Szenen in elektronischer Form aufzunehmen, und werden auf die gleiche Weise wie herkömmliche Kameras mit chemischem Film verwendet. Ein weiterer Bereich der herkömmlichen elektronischen Bildverarbeitung wird durch die Scanner dominiert, welche speziell auf das Einscannen eines Dokumentes in eine Grafikdatei unter Verwendung von Bildzusammenfügetechniken zugeschnitten ist. Die Grafikdatei kann einer optischen Zeichenerkennungsmaschine (OCR) eingespeist werden, welche Text in der Datei erkennt und dann eine Datei erzeugt, die Textzeichenketten enthält. Dies ermöglicht es, daß Papierdokumente in elektronische Dateien zur einfacheren Bearbeitung unter Verwendung eines Computers konvertiert werden. Scanner jedoch können keine Bilder dreidimensionaler Objekte aufnehmen, wie dies unter Verwendung der herkömmlichen elektronischen Kamera ausgeführt werden kann. Die herkömmliche elektronische Kamera kann, obwohl sie ständig verbessert wird, um Bilder höherer Qualität zu liefern, die denen chemischer Filmkameras gleichkommen, nicht Text in eine Datei "scannen". Somit ist der Technologiebenutzer, der elektronische Bilder von Freunden und der Familie aufnehmen möchte und der außerdem Text scannen möchte, gezwungen, sowohl den Scanner als auch die Kamera zu kaufen.

[0003] Aus der US 4949391 ist ein Bilderfassungssystem bekannt, welches einen optischen Scanner oder eine Kamera enthält. Dieses Gerät wird über eine Vorlage geführt und nimmt einen streifen von Bilddaten auf, wobei redundante Informationen beseitigt werden.

[0004] Die US 5675672 beschreibt eine Einrichtung zum Verbinden von Zeichenketten, die sich teilweise Überlappen.

[0005] Aus der US 4272756 ist ein Verfahren zur Mustererkennung bekannt, bei dem ein zweidimensionales Bild in Mustersignale konvertiert wird, wobei das Bild in eine Mehrzahl von überlappenden Teilbildern zerlegt wird.

[0006] Die US 5581637 beschreibt ein System zum Zusammensetzen von Teilbildern, wobei ein Muster auf eine zu erfassende Oberfläche projiziert wird und

mehrere Teilbilder anhand der Musterinformationen zusammengesetzt werden.

[0007] Aufgabe der Erfindung ist es, ein verbessertes System zur Textfassung an Bildern mit dreidimensionalen Objekten bereitzustellen.

[0008] Diese Aufgabe wird erfindungsgemäß durch eine Kamera gemäß Anspruch 1, einen Gegenstand gemäß Anspruch 11 oder ein Verfahren gemäß Anspruch 12 gelöst.

[0009] Ein Ausführungsbeispiel der Erfindung ist auf ein Verfahren gerichtet zum Aufbauen einer elektronischen Datei mit den Schritten des Bildens einer Reihe von Lichtabbildern eines Abschnitts einer Szene auf einem Kamerabildgeber, wobei der Bildgeber Teil einer elektronischen Kamera ist, die dreidimensionale Objekte aufnimmt, und des Erzeugens einer Reihe von Bildausschnitten ("Kacheln" bzw. "Tiles"), die die Bilder darstellen, unter Verwendung der Kamera. Eine Reihe von Textausschnitten, die jeweils Text enthalten, der in einem zugehörigen Bildausschnitt erkannt worden ist, wird erzeugt. Das Verfahren schließt das Durchsuchen nach überlappendem Text in den Textausschnitten und das Ausrichten der Textausschnitte ein.

[0010] Weitere Merkmale und Vorteile der Erfindung werden aus den beigefügten Zeichnungen und aus der unten folgenden detaillierten Beschreibung klar.

[0011] Die Erfindung wird beispielhaft und in keinem einschränkenden Sinne in den Figuren der beigefügten Zeichnungen veranschaulicht, in welchen gleiche Bezugszeichen ähnliche Elemente anzeigen und in welchen:

[0012] [Fig. 1](#) eine Blockdarstellung einer in Übereinstimmung mit einem Ausführungsbeispiel der Erfindung konfigurierten elektronischen Kamera zeigt.

[0013] [Fig. 2](#) zeigt ein Ablaufdiagramm der in Übereinstimmung mit einem Ausführungsbeispiel der Erfindung ausgeführten Operationen.

[0014] [Fig. 3](#) zeigt, wie zwei Textausschnitte ausgerichtet werden.

[0015] [Fig. 4](#) zeigt ein Ablaufdiagramm von Korrekturoperationen.

[0016] Ein Ausführungsbeispiel der Erfindung ist auf eine Einrichtung und ein Verfahren zum Aufbauen einer elektronischen Datei unter Verwendung einer elektronischen Kamera gerichtet. Eine derartige elektronische Kamera spielt somit eine doppelte Rolle, sowohl als eine herkömmliche Kamera als auch als Scanner, wodurch die Kosten des Aufnehmens herkömmlicher Bilder entfernter Objekte und des

Einscannens von Text reduziert werden. Das Konfigurieren einer elektronischen Kamera, so daß sie Text gemäß den unten beschriebenen Techniken einscannen kann, ermöglicht die Erzeugung einer elektronischen Datei, wie beispielsweise einer einfachen Textdatei oder einer zusammengesetzten Text/Grafik-Textverarbeitungsdatei, die Text auf einem dreidimensionalen Objekt, wie beispielsweise einem Verpackungscontainer, oder auf entfernten Objekten mit ungewöhnlich großem Text, wie beispielsweise Wandtafeln, oder weißen Tafeln, zeigt. Das Einscannen von Text von solchen Objekten würde für den herkömmlichen, von Hand gehaltenen Scanner ein Problem darstellen. Ein weiterer von einem Ausführungsbeispiel der Erfindung zur Verfügung gestellter Vorteil besteht darin, daß keine Bildzusammenfügetechniken erforderlich sind, um die Textdatei zu bilden. Bildzusammenfügetechniken, die an Rastergrafikdaten arbeiten, erfordern eine beträchtliche Menge der Datenverarbeitungs- und Speicherressourcen. Im Gegensatz dazu suchen die bei der Erfindung verwendeten Textverknüpfungstechniken nach Textzeichenketten (Buchstaben und Worte), welche einen viel geringeren Freiheitsgrad als Rastergrafikdaten haben und welche folglich weniger rechenintensiv zu bearbeiten sind, und richten diese Textzeichenketten aus. Dies macht die Textverknüpfung zur Verwendung in einer Umgebung mit eingeschränkten Ressourcen, wie beispielsweise einer elektronischen Kamera, geeigneter.

[0017] [Fig. 1](#) veranschaulicht eine elektronische Kamera **100**, die gemäß einem Ausführungsbeispiel der Erfindung konfiguriert ist. Die Kamera **100** enthält eine herkömmliche Optik **104** mit einem optischen Pfad, der zu einem Bildgeber **108** führt, der in der Brennebene der Optik positioniert ist. Die Optik **104** kann eine Zoom-Linse und eine einstellbare Blende haben. Eine derartige Kombination gestattet es der Kamera **100**, Szenen bei einem breiten Bereich von Gesichtsfeldern/Tiefenschärfen, die sowohl Nahaufnahmen von Dokumenten als auch Fernaufnahmen von dreidimensionalen Szenen abdecken, aufzunehmen, die nicht mit herkömmlichen Scannern aufgenommen werden können. Die Optik **104** und der Bildgeber **108** erreichen zusammen ein Gesichtsfeld/eine Tiefenschärfe, die mit herkömmlichen Scannern nicht erreicht werden können. Dies ermöglicht die Aufnahme dreidimensionaler Szenen, die große räumliche Objekte und entfernte Objekte enthalten, wohingegen Scanner nur in der Lage sind, Drucksachen aufzunehmen, die zweidimensional und relativ nah an dem Scanner positioniert sind. Die Optik **104** kann darüber hinaus ein variables Gesichtsfeld/eine variable Tiefenschärfe unterstützen, indem sie beispielsweise eine Zoom-Linse enthält.

[0018] Der Bildgeber **108** kann in Übereinstimmung mit einem breiten Bereich herkömmlicher Techniken implementiert werden, beispielsweise unter Verwen-

dung von Technologien, wie dem ladungsgekoppelten Bauelement (CCD) oder dem Komplementär-Metall-Oxid-Halbleiter (CMOS). Die von dem Bildgeber **104** erzeugten digitalisierten Rohbilddaten werden von einer herkömmlichen Digitalbild- und Signalverarbeitungsschaltung **112** verarbeitet, um die digitalen Bildausschnitte (z.B. Pixelarrays) der aufgenommenen Szene hervorzubringen. Diese Ausschnitte (Flächenelemente) können die vollständige räumliche Auflösung des Bildsensorarrays aufweisen und sie können skalierte oder beschnittene Abschnitte derselben sein. Die Ausschnitte können Abschnitte eines Dokuments **102** repräsentieren und werden normalerweise gewonnen, indem die Kamera **100** um das Dokument **102** herum bewegt wird, während die Kamera eine Serie von einander überlappenden Ausschnitte aufnimmt, um das vollständige Dokument abzudecken.

[0019] Die Ausschnitte können zu einer externen Datenverarbeitungseinrichtung **116** in einem digitalen Bilddateiformat übertragen werden, wobei die externe Datenverarbeitungseinrichtung nicht ausschließlich ein Stand-alone-Computer, wie beispielsweise ein Personalcomputer (PC), ist. Eine Schnittstelle **120** in der Kamera **100** kann eine zu einem herkömmlichen Computerperipheriebus sein, der die Kamera mit dem PC oder mit einer PC-Peripherieeinrichtung verbindet. Eine Software in dem PC kann dann die digitalen Dateien wiedergeben oder weiterverarbeiten, um sie als Standbilder oder als bewegtes Video anzuzeigen. Der PC kann darüber hinaus mit zusätzlicher Software konfiguriert sein, um die Texterkennungs- und Textverknüpfungsschritte in Übereinstimmung mit bestimmten Ausführungsbeispielen der Erfindung auszuführen. Die Schnittstelle **120** kann auch verwendet werden, um eine elektronische Datei, die den verknüpften Text enthält, zu dem PC zu übertragen.

[0020] Um die verknüpfte Textdatei zu erzeugen, weist die Kamera **104** eine Zeichenerkennungsmaschine **124** auf, die mit einem Textverknüpfher **128** arbeitet. Die Zeichenerkennungsmaschine **124** kann auf einer herkömmlichen optischen Erkennungsmaschine (OCR), die gedruckten Text in einer Grafikbilddatei erkennt und daraufhin den erkannten Text in einem durch einen Computer leicht nutzbaren Format ausgibt, basieren. Die Zeichenerkennungsmaschine **124** stellt eine Textdaten-Tile zur Verfügung, die Ketten erkannten Textes enthält, der in einem zugehörigen Bildausschnitt erscheint. Diese Textausschnitte werden gesammelt und miteinander über den Textverknüpfher **128** verknüpft, um somit eine "verknüpfte Textdatei" aufzubauen.

[0021] Der Textverknüpfher **128** sucht nach einer übereinstimmenden Textzeichenkette in zwei verschiedenen Textausschnitten und "klebt" die Textausschnitte in geeigneter Ausrichtung in die Textdatei

ein. Um die Wahrscheinlichkeit der richtigen Ausrichtung auf ein Maximum zu bringen, sollte die übereinstimmende Zeichenkette relativ lang sein und darüber hinaus nicht häufig in einer Textdatei vorkommen. Dieser Verknüpfungsprozeß wiederholt sich mit nachfolgenden Textausschnitten, um die zusammengefügte Textdatei aufzubauen, die schließlich das Dokument **102** repräsentiert. Sowohl der Textverknüpfer als auch die Zeichenerkennungsmaschine können entweder in der Kamera **100** oder in der digitalen Verarbeitungseinrichtung **116** als Prozessor implementiert werden, der Befehle aus einem Speicher ausführt.

[0022] [Fig. 2](#) zeigt eine Texterkennungs- und Verknüpfungsprozedur gemäß einem Ausführungsbeispiel der Erfindung, und [Fig. 3](#) zeigt zwei Beispiel-Textausschnitte, die verknüpft werden. Die Operation beginnt mit einem ersten digitalen Bildausschnitt, `first_scan_img`, die einem OCR im Schritt **204** eingegeben wird, um einen ersten Textausschnitt, `first_scan_tex`, hervorzubringen. Jede dieser Ausschnitte kann als zweidimensionales Array von Zeichenketten dargestellt werden, wobei jedes Element des Arrays ein Satz, Buchstaben oder gar ein einzelnes Zeichen sein kann. Normalerweise wird die OCR so konfiguriert, daß sie sich nur auf Text in "Bereichen hohen Vertrauens" eines Bildausschnitts konzentriert, um die Möglichkeit von Erkennungsfehlern zu reduzieren. Die OCR kann einen Vertrauensindexwert zur Verfügung stellen, der eine Anzeige des Vertrauensniveaus der OCR beim Erkennen einer bestimmten Textzeichenkette in dem Bildausschnitt gibt. Es kann sein, daß ein von einer Digitalkamera gewonnenes Bild voller Auflösung eine geometrische Verzerrung in den Ecken und an den Kanten zeigt. So kann eine Ausschnitt oder Tile, die die volle Auflösung einer Kamera aufweist, beschnitten werden, bevor sie der OCR eingespeist wird, um den Bereich hohen Vertrauens und den Vertrauensindex auf ein Maximum zu bringen. Die Operation wird dann mit dem Schritt **208** fortgesetzt.

[0023] Im Schritt **208** wird die elektronische Datei `doc_tex` mit dem ersten Textausschnitt initialisiert. Die Textdatei `doc_tex` kann anfänglich als ein ausreichend großes zweidimensionales Array von Zeichenketten dargestellt werden. Wenn sämtliche der gewünschten Bildausschnitte verarbeitet worden sind und die Verknüpfung abgeschlossen ist, kann das Array dann in irgendeines einer Reihe bekannter Textverarbeitungsformate konvertiert werden. Ein `prev_scan_tex` entsprechendes Array wird im Schritt **208** auf `first_scan_tex` initialisiert. Die Operation wird dann mit einer Schleife fortgesetzt, die im Schritt **212** beginnt. Solange es Standbildausschnitte gibt, die zu verarbeiten sind, wird ein aktueller Textausschnitt, `curr_scan_tex`, aus der OCR im Schritt **216** gewonnen, und es wird ein zu vergleichendes Merkmal, `str_ftr`, daraus im Schritt **220** extrahiert. Um die Wahr-

scheinlichkeit des Ausrichtens des Textes in `prev_scan_tex` mit dem in `curr_scan_tex` auf ein Maximum zu bringen, sollte `str_ftr` ein relativ langes Wort oder eine relativ lange Sequenz von Zeichen enthalten. Sofern kein langes Wort in `prev_scan_tex` verfügbar ist, so sollte dann eine Reihe sequentiell auftretender Worte als `str_ftr` ausgewählt werden. Alternativ oder zusätzlich kann der Verknüpfer so konfiguriert werden, daß er bestimmte "Stop-Worte", wie beispielsweise "the", "of", "an" und "and" ignoriert, so daß die ausgewählte `str_ftr` keine solchen Stop-Worte enthalten sollte. Sobald `str_ftr` extrahiert worden ist, fährt die Operation mit Schritt **224** fort.

[0024] Im Schritt **224** durchsucht der Verknüpfer den `curr_scan_tex` und `prev_scan_tex` nach `str_ftr`. Sofern `str_ftr` in beiden Ausschnitten gefunden wird, wird der `curr_scan_tex` an `doc_tex` im Schritt **228** angehängt, sofern er richtig zu `prev_scan_tex` ausgerichtet ist. Die Ausrichtungs- und Anhängschritte sind anhand eines Beispiels in [Fig. 3](#) gezeigt. Das Wort "cameras" (das bei der 2. Zeile und 3. Spalte von `curr_scan_tex` beginnt) wird als `str_ftr` ausgewählt und sein zugehöriger Ort in `prev_scan_tex` ist an der 3. Zeile und 17. Spalte. Die Gültigkeit dieses zugehörigen Orts, d.h. die Ausrichtung, kann näher durch seine Nachbartext-Zeichenketten, wie beispielsweise "suc", "ts. The", "elec", etc. bestätigt werden. Diese zwei Textausschnitte können in eine verknüpft werden, indem `curr_scan` an `prev_scan` in Übereinstimmung mit ihrer Ortsdifferenz, welche $(3 - 2 \text{ Zeilen}, 17 - 3 \text{ Spalten}) = (1 \text{ Spalte}, 14 \text{ Spalten})$ ist. Spezieller ausgedrückt, kann die Verknüpfung dieser beiden Tiles formuliert werden als:

$\text{prev_scan}(i + 1, j + 14) = \text{curr_scan}(i, j)$ für alle (i, j) .

[0025] Das Ergebnis dieses Textverknüpfungsschritts ist in [Fig. 3](#) als `doc_tex` gezeigt. Die Operation kehrt dann in einer Schleife zurück zum Schritt **212**. Wenn sämtliche Bildausschnitte auf diese Weise verarbeitet worden sind, erzeugt der Verknüpfer **128** den `doc_tex` im Schritt **232** entweder als herkömmliche ASCII-Datei oder als ein Zeiger in ein Zeichenkettenarray. Sofern die Kamera **100** außerdem mit einem Computer verbunden ist, dann kann die verknüpfte Textdatei gleichzeitig angezeigt werden, während sie aufgebaut wird.

[0026] Ein zusätzliches Merkmal der Erfindung ist in [Fig. 4](#) als Technik zum Korrigieren von Fehlern in der verknüpften Textdatei gezeigt. Solche Fehler können falsch erkannten Text oder Fehlausrichtungen der Textausschnitte einschließen. Diese Fehler können durch einen Benutzer erfaßt werden, während er die verknüpfte Textdatei, die aufgebaut wird, beobachtet und sie mit dem gescannten tatsächlichen Dokument vergleicht. Alternativ oder zusätzlich können die Fehler unter Verwendung einer automatisierten Methodik erfaßt werden, wie beispielsweise eines Aussprache-

überprüfers, der in die Kamera eingebaut ist, und der falsch ausgesprochene Worte oder Sequenzen von Worten in der verknüpften Textdatei erfassen kann. Bei einer weiteren Alternative kann der OCR-Vertrauensindex einen Erkennungsfehler anzeigen, welcher sofort dem Benutzer unter Verwendung eines audiovisuellen Alarms, wie beispielsweise eines warnenden Piepens, signalisiert wird. Sobald ein Fehler erfaßt worden ist, kann die normale Erkennungs- und Verknüpfungsprozedur unterbrochen werden, um eine Korrekturroutine beim Antreffen des Fehlers aufzurufen. Dies kann effektiver sein als darauf zu warten, daß das gesamte Dokument verarbeitet wird, bevor die sich ergebende verknüpfte Textdatei nach Fehlern durchgesehen wird. Im allgemeinen kann die Korrekturtechnik gemäß [Fig. 4](#) zu einem beliebigen Zeitpunkt während oder nach der Prozedur gemäß [Fig. 2](#) angewendet werden.

[0027] Die Operation beginnt im Schritt **404**, bei welchem die Korrekturroutine bestimmt, ob es keine weiteren zu korrigierenden Fehler gibt. Wenn es irgendwelche gibt, dann wird die Operation im Schritt **408** fortgesetzt, bei welchem die Routine darauf wartet, daß die Kamera über einem Bereich der Szene positioniert wird, die einen Text enthält, der mit dem Fehler korrespondiert. Die Operation fährt dann beim Schritt **412** fort, bei welchem ein neuer Bildausschnitt, `new_scan_img`, des Bereichs durch die Kamera gewonnen wird und in die OCR eingegeben wird. Die OCR fährt mit der Texterkennung fort und erzeugt einen neuen Textausschnitt, `new_scan_tex`, auf der Grundlage von `new_scan_img`.

[0028] Angenommen, daß `new_scan_tex` keine Erkennungsfehler enthält, so wird dann ein Zeichenkettenmerkmal `str_ftr` aus dem `new_scan_tex` im Schritt **416** extrahiert. Wiederum sollte `str_ftr` so einzigartig wie möglich sein und einen hohen Erkennungsvertrauensindex, wie er durch die OCR gegeben ist, aufweisen, so daß die Zeichenkette schnell gefunden werden kann, während das zuvor erzeugte `doc_tex` um den Bereich herum, bei dem der Fehler auftrat, durchsucht wird, wie im Schritt **420**. Wenn die `str_ftr` sowohl in dem `doc_tex` als auch in dem `new_scan_tex` gefunden wird, wie im Schritt **424**, so wird die Ausrichtung der Ausschnitte überprüft, und, sofern sie ausgerichtet sind, die `new_scan_tex` über den zuvor wiederhergestellten Text, welcher den Fehler in `doc_tex` enthielt, aufgeheftet. Diese Korrektur kann weiter verifiziert werden, indem die Korrektur von `new_scan_tex` auf einen von dem Benutzer betrachteten Monitor angezeigt wird. Die Operation kehrt dann in einer Schleife zum Schritt **404** zurück, um irgendwelche weiteren Fehler zu korrigieren.

[0029] Wenn der Benutzer, während er einen Anzeigemonitor betrachtet, der die `doc_tex` anzeigt, sieht, daß der Fehler nicht korrigiert worden ist, dann unternimmt er einen weiteren Versuch, den Fehler zu kor-

rigieren, ändert aber diesmal entweder den Kamerabrennpunkt, indem er die Optik **104** einstellt oder ändert den Abstand zwischen der Kamera und dem zu scannenden Objekt. Dies kann dazu führen, daß die Kamera ein schärferes Bild des Fehlerbereichs gewinnt, um die Wahrscheinlichkeit eines weiteren Erkennungsfehlers zu reduzieren. Alternativ kann ein Bild, das einen größeren Bereich abdeckt, und das einen zuverlässigeren Ausrichtungsschritt gestattet, gewonnen werden. Als eine Verbesserung der oben beschriebenen Ausführungsbeispiele kann das System so konfiguriert werden, daß es den Benutzer warnt, daß die aus dem Kameraabbilder empfangenen Textausschnitte nicht ausreichend klar sind und daß sie wahrscheinlich zu Texterkennungsfehlern führen oder daß sie sich nicht ausreichend überlappen, um eine zuverlässige Ausrichtung zu erbringen. Dies kann beispielsweise ausgeführt werden, indem eine Warnung an ein Kameraanzeigenfenster oder ein Fenster der externen Datenverarbeitungseinrichtung **116** (siehe [Fig. 1](#)) gesendet wird.

[0030] Bei einem Ausführungsbeispiel der Erfindung kann, wobei auch auf [Fig. 1](#) Bezug genommen wird, eine herkömmliche Bildzusammenfügeroutine in die separate Datenverarbeitungseinrichtung **116** geladen werden, um das Abtasten der grafischen Bilder zusätzlich zu der oben beschriebenen Textverknüpfung zu ermöglichen. Sofern die OCR keinen Text in einigen der Bildausschnitte erkennt, so können dann diese Ausschnitte der (nicht gezeigten) Bildzusammenfügeroutine anstelle des Textverknüpfers **128** eingespeist werden, so daß eine grafische Abbildung erzeugt wird. Dem Benutzer kann dann eine Option signalisiert werden, ob diese grafische Figur an die elektronische Datei angehängt werden sollte. Selbstverständlich könnte der Benutzer später die zusammengefügte grafische Figur durch ein Original höherer Qualität ersetzen, sofern dies gewünscht wird.

Patentansprüche

1. Eine elektronische Kamera, aufweisend:
eine Optik mit einer gegenüber einem herkömmlichen Scanner erhöhten Tiefenschärfe zum Erzeugen einer Mehrzahl von Bildern eines Textes anzeigenden räumlichen Objekts;
einen Kameraabbilder, der die Bilder aufnimmt und in Abhängigkeit davon eine Mehrzahl von Bildausschnitten zur Verfügung stellt, die jeweils verschiedene Abschnitte des Textes des Objekts wiedergeben;
eine optische Zeichenerkennungsmaschine, die eine Mehrzahl von Text enthaltenden Textausschnitten, die jeweils aus einem zugehörigen Bildausschnitt erkannt worden sind, zur Verfügung stellt;
einen Textverknüpfer, der in den Textausschnitten nach überlappendem Text sucht; und
eine Einrichtung zum Korrigieren fehlerhaften Textes in der Mehrzahl von Textausschnitten, die eine Neu-

aufnahme eines Abschnitts des Text anzeigenden Objekts veranlaßt, der dem fehlerhaften Text entspricht, wobei aus dem Bildausschnitt der Neuaufnahme ein entsprechender Textausschnitt erzeugt wird, und die ein Ersetzen des fehlerhaften Textes durch die korrekte Textdaten der erzeugten Textausschnitts bewirkt.

2. Die Kamera nach Anspruch 1, wobei die Optik so einstellbar ist, daß sie eine variable Tiefenschärfe zur Verfügung stellt.

3. Die Kamera nach Anspruch 1, wobei der Textverknüpfer die Textausschnitte in richtiger Ausrichtung an eine elektronische Datei anhängt.

4. Die Kamera nach Anspruch 1, wobei die Bildausschnitte von unterschiedlicher räumlicher Auflösung sind.

5. Die Kamera nach Anspruch 3, wobei der Text eine relativ lange Zeichenkette enthält, die relativ selten in einem ersten Textausschnitt erscheint.

6. Die Kamera nach Anspruch 1, wobei der Bildausschnitt die vollständige räumliche Auflösung des Kameraabbilders aufweist.

7. Die Kamera nach Anspruch 1, wobei die Zeichenerkennungsmaschine und der Textverknüpfer als vom Prozessor ausgeführte Befehle implementiert sind.

8. Die Kamera nach Anspruch 1, ferner aufweisend eine Ausgabeschnittstelle zu einer Datenverarbeitungseinrichtung, welche nicht ausschließlich eine Stand-Alone-Kamera ist, zum Übertragen der elektronischen Datei zu der Einrichtung.

9. Die Kamera nach Anspruch 8, wobei die Ausgabeschnittstelle einer Computerperipheriebuspezifikation entspricht.

10. Die Kamera nach Anspruch 3, ferner aufweisend, eine Bildzusammenfügemaschine zum Erzeugen einer grafischen Figur in Abhängigkeit vom Zusammenfügen einiger der mehreren Bildausschnitte, die von der optischen Zeichenerkennungsmaschine unerkannte Abschnitte enthalten, und wobei der Verknüpfer ferner so konfiguriert ist, daß er die Grafikfigur an die elektronische Datei anhängt.

11. Ein Gegenstand, aufweisend:
ein maschinenlesbares Medium, das Befehle enthält, die, wenn sie von einem Prozessor ausgeführt werden, ein den Prozessor enthaltendes System veranlassen:
eine Mehrzahl von Bildausschnitten zu empfangen, die eine Mehrzahl von Bildern von Abschnitten einer Szene darstellen, die von einer elektronischen Ka-

mera aufgenommen worden sind, die eine Optik mit einer Tiefenschärfe aufweist, die größer ist als die eines herkömmlichen Scanners;
eine Mehrzahl von Textausschnitten zu erzeugen, die jeweils Text enthalten, der jeweils in einem zugehörigen der Bildausschnitte erkannt worden ist; und
nach überlappendem Text in den Textausschnitten zu suchen, den Text in geeigneter Ausrichtung an eine elektronische Datei anzuhängen und Fehler in der elektronischen Datei zu korrigieren, indem ein korrekter Text in einem weiteren Bildausschnitt erkannt und in einen Korrektur-Textausschnitt eingebracht wird, wobei die Bildausschnitte einen zuvor von der Kamera aufgenommenen Abschnitt, der den dem Textfehler entsprechenden Text enthält, entspricht, und indem der Fehler durch den korrekten Text ersetzt wird.

12. Ein Verfahren, umfassend:

Erzeugen einer Mehrzahl von Bildern eines Abschnittes einer Text enthaltenden Szene auf einem Kameraabbilder, wobei der Abbilder Teil einer elektronischen Kamera ist, die dreidimensionale Objekte aufnimmt;

Erzeugen einer Mehrzahl von Bildausschnitten, die die Bilder repräsentieren, unter Verwendung der Kamera;

Erzeugen einer Mehrzahl von Textausschnitten, die jeweils Text enthalten, der in einem zugehörigen der Bildausschnitte erkannt worden ist;

Suchen nach überlappendem Text in den Textausschnitten und Erzeugen einer elektronischen Datei, an welche die Textausschnitte auf der Basis des überlappenden Textes angehängt werden;

Feststellen, ob fehlerhafter Text in den mehreren Textausschnitten vorhanden ist, und, sofern dies der Fall ist:

- Neu-Aufnehmen eines Bereichs der Szene, der den dem fehlerhaften Text entsprechenden Text enthält, und Gewinnen eines neuen Bildausschnitts,

- Erzeugen eines neuen Textausschnitts, die anstelle des fehlerhaften Textes den korrekten Text enthält, wie er aus dem neuen Bildausschnitt erkannt worden ist, und

- Einsetzen des erkannten, korrekten Textes in die elektronische Datei.

13. Ein Verfahren nach Anspruch 12, wobei das Erzeugen der Textausschnitte von einer von der Kamera getrennten Datenverarbeitungseinrichtung ausgeführt wird.

14. Ein Verfahren nach Anspruch 12, ferner umfassend das Übertragen der elektronischen Datei in eine von der Kamera getrennte Datenverarbeitungseinrichtung unter Verwendung eines Computerperipheriebusses.

15. Ein Verfahren nach Anspruch 12, ferner umfassend

das Erzeugen einer Grafikfigur in Abhängigkeit von einem Zusammenfügen einiger der Mehrzahl von Bildausschnitten, die Nicht-Text-Abschnitte enthalten und von der optischen Zeichenerkennungsmaschine un erkannt sind; und
das Anfügen der Grafikfiguren an die elektronische Datei.

Es folgen 4 Blatt Zeichnungen

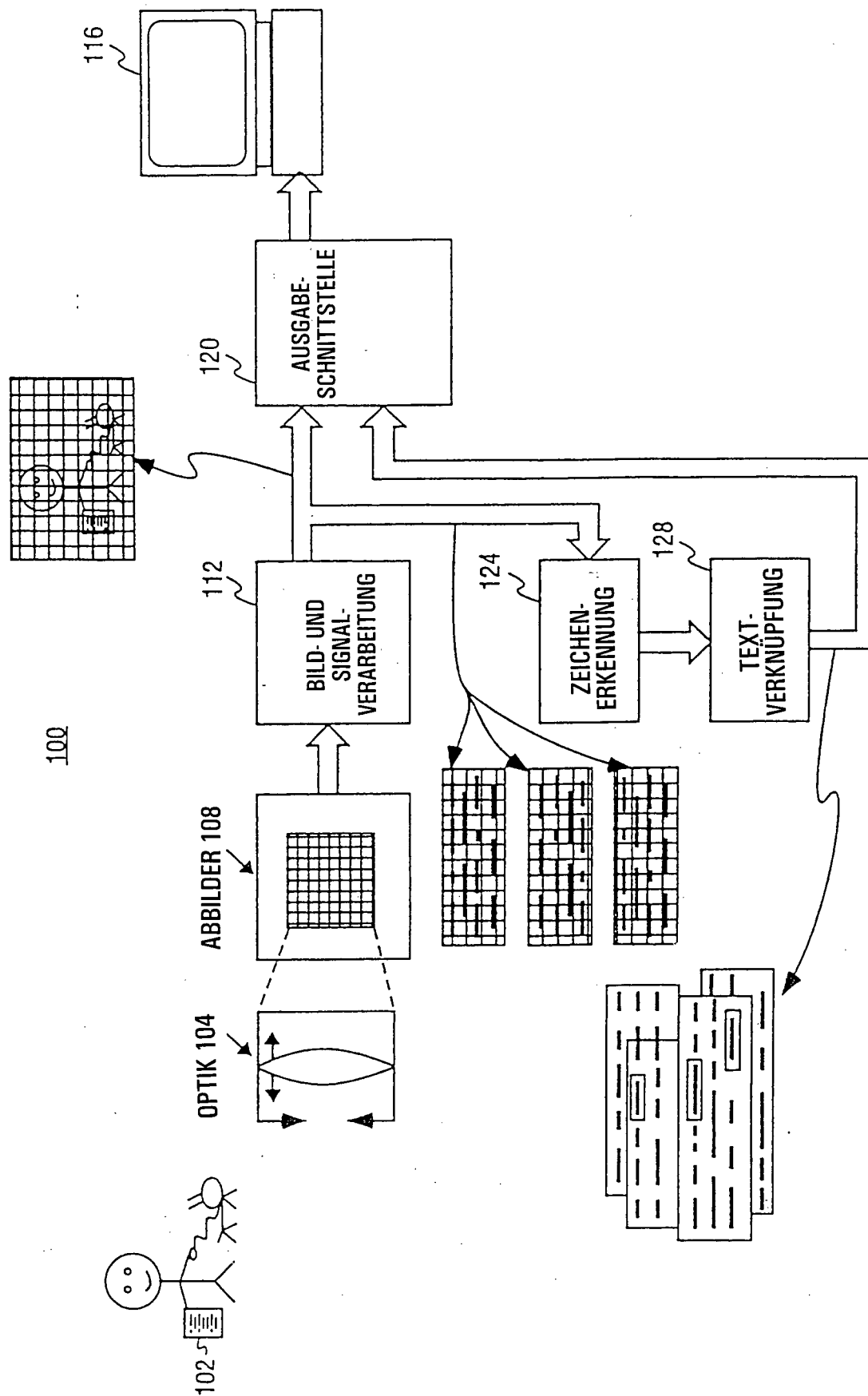


FIG. 1

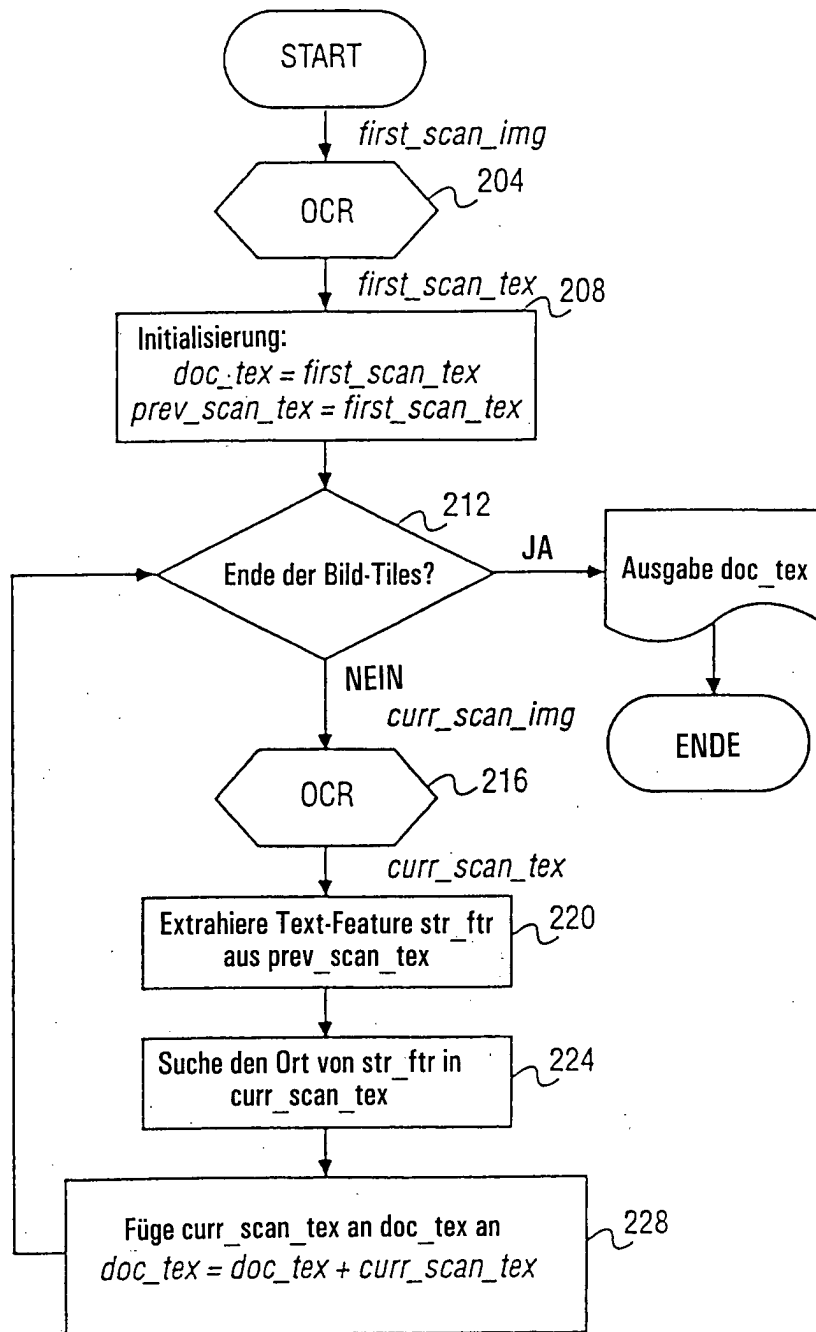


FIG. 2

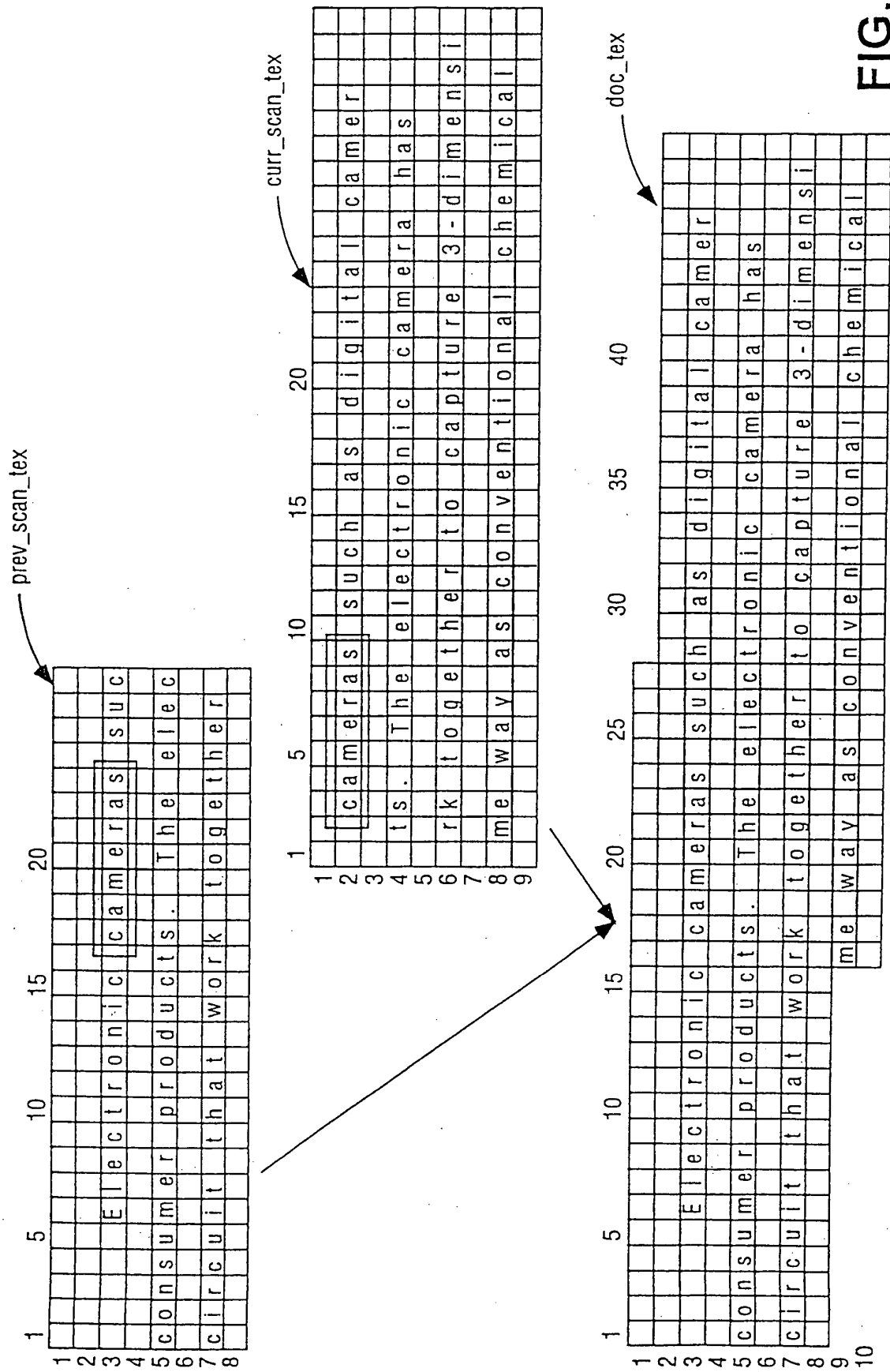


FIG. 3

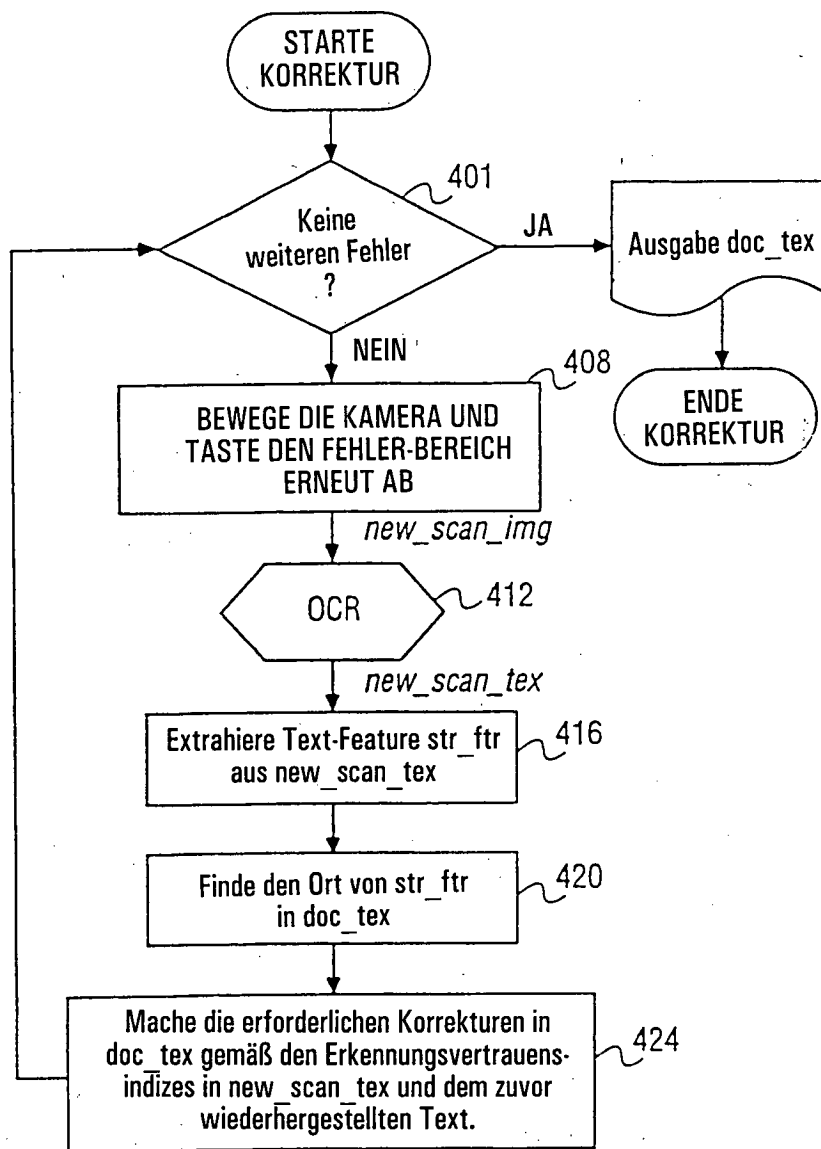


FIG. 4