



(19)
Bundesrepublik Deutschland
Deutsches Patent- und Markenamt

(10) **DE 694 32 585 T2 2004.04.08**

(12)

Übersetzung der europäischen Patentschrift

(97) **EP 0 660 256 B1**

(51) Int Cl.7: **G06K 9/20**

(21) Deutsches Aktenzeichen: **694 32 585.6**

(96) Europäisches Aktenzeichen: **94 309 673.5**

(96) Europäischer Anmeldetag: **22.12.1994**

(97) Erstveröffentlichung durch das EPA: **28.06.1995**

(97) Veröffentlichungstag

der Patenterteilung beim EPA: **02.05.2003**

(47) Veröffentlichungstag im Patentblatt: **08.04.2004**

(30) Unionspriorität:

171720 22.12.1993 US

(84) Benannte Vertragsstaaten:

DE, FR, GB, IT

(73) Patentinhaber:

Canon K.K., Tokio/Tokyo, JP

(72) Erfinder:

Wang, Shin-Ywan, Tustin, California 92680, US

(74) Vertreter:

**Tiedtke, Bühling, Kinne & Partner GbR, 80336
München**

(54) Bezeichnung: **Verfahren und Gerät zur Auswahl von Text und/oder Non-Text-Blöcken in einem gespeicherten Dokument**

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 Anmeldung ist mit einem Anhang von Computerprogrammauflistungen versehen. Die Computerprogramme sind in der Programmiersprache "C" geschrieben.

[0002] Ein Teil der Offenbarung dieses Dokuments unterliegt dem Copyright. Der Inhaber des Copyrights erhebt keine Ansprüche gegen Vervielfältigung des Dokuments oder der Patentveröffentlichung, wie sie in der Veröffentlichung des Patentamts erscheint; im anderen Fall hat das gesamte Copyright seine Gültigkeit.

[0003] Die vorliegende Erfindung bezieht sich auf ein Verfahren und ein Gerät zur Verarbeitung von Zeichenblöcken gegenüber beispielsweise der Zeichenerkennung, und speziell auf solch ein Verfahren und ein Gerät, bei dem gegenüber der Erkennung Bilddatenblöcke zugeordnet und ausgewählt werden, die auf den Eigenschaften der Bilddaten beruhen. Beispielsweise können Bilddatenblöcke auf der Grundlage ausgewählt und zugeordnet werden, ob die Bilddaten Textbilddaten (horizontal oder vertikal) oder Non-Text-Bilddaten (im folgenden "textfreie Bilddaten" genannt) wie Halbton- (oder Grauwert-)Bilder, Strichzeichnungen, Tabellen, vertikale oder horizontale Linien, schiefe vertikale oder horizontale schräge Linien, Bilder oder ähnliches sind.

[0004] In den vergangenen Jahren wurde es möglich, Bilder von Textdaten zu analysieren, um einzelne Zeichen bei den Textdaten zu erkennen und eine vom Computer lesbare Datei von Zeichencodes zu bilden, die den erkannten Zeichen entsprechen. Derartige Dateien lassen sich dann bei der Textverarbeitung, bei der Datenkompression oder bei Datenverarbeitungsprogrammen bearbeiten. Solche Systeme, die in folgenden als "Zeichenerkennungssysteme" bezeichnet werden, sind von Vorteil, weil sie neues Schreiben oder eine erneue Eingabe der Textdaten überflüssig machen. Beispielsweise ist es möglich, eine Zeichenerkennung eines Dokuments durchzuführen, das durch Bildübertragung übermittelt oder von einem Mikrofilm beziehungsweise einem Photokopierer reproduziert wurde, damit Computer-Textdateien gebildet werden können, die Zeichencodes (beispielsweise ASCII (American Standard Code for Information Interchange) der Zeichen und der Ziffern in dem Dokument enthalten, wobei weitere Textverarbeitung oder Datenverarbeitung des Dokuments zugelassen wird, ohne das Dokument neu schreiben oder neu eingeben zu müssen.

[0005] Dokumente, die Zeichenerkannt werden, enthalten häufig viele verschiedene Arten von Bilddaten, von denen nicht alle erkannt werden können. Während es derzeit beispielsweise möglich ist, Textbilddaten zu erkennen, ist es sehr schwer, textfreie Bilddaten zu erkennen. Typischerweise enthalten Dokumente, die Zeichenerkannt werden sollen, Textbilddatenblöcke und textfreie Bilddatenblöcke wie Halbtonbilder, Strichzeichnungen oder Striche. Zusätzlich

können Dokumente Tabellen oder tabellarisch angeordnete Daten enthalten, die eingerahmt oder nicht eingerahmt sind. Entsprechend ist es vor der Zeichenerkennungsverarbeitung bei einzelnen Blöcken im Dokument erforderlich, daß entsprechend dem Bilddatentyp bei den Blöcken zuzuordnen, und die Blöcke vom Texttyp aus den Bilddaten auszuwählen.

[0006] **Fig. 1** zeigt eine Seite eines typischen Dokuments. In **Fig. 1** ist die Seite **101** des Dokuments zweispaltig angeordnet. Die Seite enthält Titelblöcke **102**, die Textinformationen mit für Titel geeigneter größerer Schriftgröße enthält, Textblöcke **104**, die Textdatenzeilen enthalten, einen Graphikblock **105**, der textfreie graphische Darstellungen enthält, einen Textblock **106**, der eine Texttafel oder eine Zifferninformation enthält, und Bildunterschriftenblöcke **107**, die Textdaten in Kleinbuchstaben enthalten, und die Unterschriften in Bezug auf Graphikblöcke oder tabellarische Informationen sind. Jeder Informationsblock wird entsprechend dem Typ der enthaltenen Information zugeordnet, und die Blöcke werden auf der Grundlage dieser Zuordnung unterteilt.

[0007] Um bisher Bilddatenblöcke vom Texttyp zu erfassen, wurden die Bildpunktdaten in horizontaler und vertikaler Richtung verwischt, indem die dunklen Bildpunkte sowohl horizontal als auch vertikal in ein oder mehrere nebeneinanderliegende weiße Bildpunkte ausgedehnt wurden. Verwischungstechniken sind unbefriedigend, weil sie sich auf die Vorkenntnis der Eigenschaften der Bilddaten vom Texttyp (beispielsweise die Zeichengröße) verlassen, um fähig zu sein, Verwischungsparameter einwandfrei zu wählen. Darüber hinaus können geringfügige Änderungen bei den Verwischungsparametern große Änderungen bei den Auswahlergebnissen hervorrufen. Verwischungstechniken sind auch nicht immer in der Lage, die interne Struktur des ursprünglichen Dokuments zu enthalten. Beispielsweise kann eine Verwischung eines zweispaltigen Originals zu einer einseitigen Seite führen. Eine solche Situation ist nicht zufriedenstellend, da die Anordnung, in der die Textdaten abgespeichert sind, durcheinander gebracht werden, was es unmöglich macht, den ursprünglichen Text genau zu rekonstruieren. Darüber hinaus wurde herausgefunden, daß Verwischungstechniken manchmal Daten vom Texttyp in Daten vom textfreien Typ verschmieren, und als Folge dazu führen, daß der gesamte Bereich als Daten vom Texttyp fehlerhaft interpretiert wird.

[0008] Die Europäische Patentanmeldung EP-A-0567344 schlägt eine andere Technik zur Auswahl von Zeichenblöcken in einem abgespeicherten Dokument vor. Hier wird das abgespeicherte Dokument zuerst durchsucht, um sogenannte "verbundene Teile" aufzufinden, die zwei oder mehrere Bildpunkte enthalten, die zusammen in eine der acht Richtungen, die jeden Bildpunkt umgeben, verbunden sind. Als nächstes werden die verbundenen Textteile von den textfreien Teilen getrennt, und die text-

freien Teile werden beispielsweise Tabellen, Halbtonbildern oder Strichzeichnungen zugeordnet. Als nächstes wird die Richtung jeder Schiefe bei dem Dokument erfaßt, und ist die Schiefe vertikal, wird das Bild um 90° gedreht, und die verbundenen Teile werden erneut gesucht. Nach der Korrektur der Schiefe werden unsichtbare weiße Linien entlang den Rändern der textfreien Teile gesucht, damit zugehörige Textblöcke, beispielsweise Spalten, erkannt werden können. Danach werden die horizontalen Textzeilen und die Titelzeilen gebildet, und die horizontalen Textzeilen werden in rechteckig angeordneten Textblöcken gruppiert. Anschließend wird die Nachverarbeitung durchgeführt, um die identifizierten Textblöcke zur weiteren Zeichenerkennungsverarbeitung aufzubereiten.

[0009] Während die oben beschriebene Blockauswahltechnik für horizontale Dokumente (beispielsweise englischsprachige Dokumente) geeignet sein kann, ist es für eine Seite möglich, die sowohl horizontale als auch vertikale Textblöcke (bidirektional) enthält, zu verwenden. Beispielsweise kann ein japanisches Dokument vertikale Kanji-Zeichen in Verbindung mit horizontalen Zeichen wie Tabellen und Bildlegenden verwenden. Auch bestimmte englischsprachige Dokumente enthalten sich in vertikaler Richtung ausdehnende Dokumente, um bestimmte Informationen hervorzuheben oder bestimmte Effekte zu bilden.

[0010] Darüber hinaus ist häufig die abgetastete Seite schief, und die oben beschriebene Blockauswahltechnik führt dieses Problem derart aus, daß als erstes die Schiefe identifiziert wird, und anschließend das Bild vor der Bildung des Textblocks gedreht wird. Die Geschwindigkeit und die Genauigkeit führen zu zwei praktischen Problemen, wenn diese Technik verwendet wird. Genauer: Bei der oben beschriebenen Blockauswahltechnik wird der Blockbereich durch ein Rechteck dargestellt, und die Grenzen der textfreien Blöcke werden ebenfalls erfaßt. Im Falle eines schiefen Dokuments kann die Trennung zwischen den Textblöcken undeutlich werden, und sie können sich tatsächlich überlappen. Dies führt zu einer Fehlzurordnung der Daten in den Blöcken, und dies kann zu Fehlern bei der Zeichenerkennung führen.

[0011] Ein Verfahren und ein Gerät zum effektiven und wirksamen Auswählen von Textblöcken und textfreien Blöcken bei einem abgespeicherten Dokument, bei dem sowohl vertikale als auch horizontale Textblöcke erkannt werden können, und bei dem ein schiefes Dokument vor der Bildung der Textblöcke gedreht werden muß, würde zu einer flexibleren Blockauswahltechnik führen, wobei dies Verarbeitungszeit einsparen würde und die Erkennungsgenauigkeit erhöhen würde.

[0012] Der Beitrag in der Zeitschrift IBM J. Res. Develop., Bd. 26, No. 6, November 1982 Document Analysis System von K. Y. Wong, et al. beschreibt eine Blockeinteilungstechnik, die einen Lauflän-

genglättungsalgorithmus verwendet, um weiße und schwarze Punktraster zu schaffen. Die Punktraster werden kombiniert und eine weiterer Lauflängenalgorithmus wird zur Herstellung einer entgeltigen Einteilung verwendet.

[0013] Die Europäische Patentanmeldung EP-A-496531 beschreibt ein Verfahren zur Analyse des Entwurfs eines Bildes eines Dokuments, einschließlich von Gegenständen, die in einer hierarchischen Baumstruktur angeordnet sind, bei der Zeichenkettenbereiche vom Eingangsdokumentbild ausgeblendet werden, Trennsegmente erfaßt werden, das Dokumentbild in baumstrukturierte Unterbereiche aufgeteilt wird und mit einem Entwurfsmodell für ein Modelldokument abgeglichen wird.

[0014] Nach einem ersten Aspekt der vorliegende Erfindung vorgesehen ist ein Verfahren zur Auswahl von Pixelblöcken aus Pixelbilddaten, mit den Verfahrensschritten:

Umreißen von Konturen verbundener Komponenten in Pixelbilddaten;

erstes Verbinden, wobei verbundene Komponenten der Breite nach auswählbar auf der Grundlage der Größe und der Nähe zu anderen verbundenen Komponenten verbunden werden, um Zeilen zu bilden; und

zweites Verbinden, wobei Zeilen der Breite nach verbundener Komponenten vertikal auswählbar auf der Grundlage von Größe und Nähe zu anderen Zeilen verbunden und Rechtecke um die vertikal verbundenen Komponenten gebildet werden;

gekennzeichnet durch den Verfahrensschritt Bilden einer unregelmäßigen Kurve um Pixelbilddaten, die von einem Rechteck umgeben sind, das sich mit einem anderen Rechteck überlappt.

[0015] Nach einem zweiten Aspekt der vorliegende Erfindung vorgesehen ist ein Gerät zur Auswahl von Blöcken von Pixeln aus Pixelbilddaten bereitgestellt, mit:

einem Speichermittel, das Pixelbilddaten speichert, einem Verarbeitungsmittel, das Konturen verbundener Komponenten in den gespeicherten Pixelbilddaten umreißt;

einem Mittel zum auswählbaren Verbinden verbundener Komponenten der Breite nach auf der Grundlage von Größe und Nähe zu anderen verbundenen Komponenten, um Zeilen zu bilden; und mit

einem Mittel zum auswählbaren Verbinden von Zeilen verbundener Komponenten vertikal der Breite nach auf der Grundlage der Größe und Nähe zu anderen Zeilen und zum Bilden von Rechtecken um die vertikal verbundenen Komponenten;

dadurch gekennzeichnet, daß das Verarbeitungsmittel ausgestattet ist mit

einem Mittel zu Erzeugen einer unregelmäßig geformten Kurve um jeden Block von Pixelbilddaten, der von einem Rechteck umgeben ist, das ein weiteres Rechteck überlappt.

[0016] Diese und andere Merkmale und Vorteile der vorliegenden Erfindung werden durch Bezugnahme

auf die folgende ausführliche Beschreibung der beispielhaften Ausführungsbeispiele in Verbindung mit der beiliegenden Zeichnung deutlich.

KURZE BESCHREIBUNG DER ZEICHNUNG

[0017] **Fig. 1** zeigt eine typische Seite eines Dokuments, das zeichenerkennbar ist.
 [0018] **Fig. 2** zeigt ein typisches Gerät entsprechend der vorliegenden Erfindung.
 [0019] **Fig. 3** zeigt ein Blockdiagramm eines Geräts entsprechend der vorliegenden Erfindung.
 [0020] **Fig. 4** zeigt ein verallgemeinertes Flußdiagramm, das aufzeigt, wie die Blockauswahleigenschaften der vorliegenden Erfindung in einen Vorgang zur Zeichenerkennung eingebunden werden kann.
 [0021] **Fig. 5** zeigt ein Gesamtflußdiagramm der Blockauswahltechnik entsprechend der vorliegenden Erfindung.
 [0022] **Fig. 6** mit den **Fig. 6A bis 6D** zeigt ein Flußdiagramm, das die Einzelheiten der Schritte S51 bis S53 in **Fig. 5** aufzeigt.
 [0023] **Fig. 7** mit den **Fig. 7A bis 7C** zeigt Ansichten zur Erläuterung des Umrißabtastens.
 [0024] **Fig. 8** mit den **Fig. 8A bis 8C** zeigt Ansichten zur Erläuterung der Zuordnungsverarbeitung für textfreie Einheiten.
 [0025] **Fig. 9** mit den **Fig. 9A bis 9D** zeigt Ansichten zur Erläuterung der weißen Umrißverarbeitung.
 [0026] **Fig. 10** zeigt eine schematische Darstellung der hierarchischen Grund- oder Baumstruktur zur Bestimmung von Textzeichen oder textfreien Zeichen entsprechend der vorliegenden Erfindung.
 [0027] **Fig. 11** zeigt ein Flußdiagramm zur Beschreibung von Schritt S56 in **Fig. 5**.
 [0028] **Fig. 12** zeigt ein Flußdiagramm zur Beschreibung von Schritt S56 in **Fig. 11**.
 [0029] **Fig. 13** mit den **Fig. 13A und 13B** zeigt schematische Darstellungen von Schritt S57 in **Fig. 5**.
 [0030] **Fig. 14** mit den **Fig. 14A und 14B** zeigt ein Flußdiagramm zur Beschreibung von Schritt S58 in **Fig. 5**.
 [0031] **Fig. 15** zeigt ein Flußdiagramm zur Beschreibung von Schritt S59 in **Fig. 5**.
 [0032] **Fig. 16** mit den **Fig. 16A bis 16E** zeigt ein Flußdiagramm zur Beschreibung von Schritt S592 in **Fig. 15**.
 [0033] **Fig. 17** zeigt eine schematische Abbildung unregelmäßig geformter Textblöcke.

GENAUE BESCHREIBUNG DER BEVORZUGTEN AUSFÜHRUNGSBEISPIELE

[0034] Die vorliegende Erfindung kann in einer Reihe von Einheiten, bei denen Kennzeichenerkennungsverarbeitung erwünscht ist, ausgeführt werden, wie beispielsweise Bildverarbeitungs- oder Bildwiedergabegeräten einschließlich Photokopierern, Bildwiedergabemaschinen, Video- oder Standbildvideoka-

meras und Laserstrahldruckern. Bei solchen Einheiten können Bilder, die Zeichenbilder enthalten, derart verarbeitet oder wiedergegeben werden, daß die Zeichenbilder zeichenerkennungsverarbeitet sind. Falls gewünscht können genormte Zeichensätze für erkannte Zeichenbilder eingesetzt werden, und die Normzeichen können in das ursprüngliche Zeichenbild zurückübertragen oder wiederhergestellt werden. Die Erfindung kann auf Universalcomputern und Büroeinheiten wie Arbeitsplatzrechnern, Textverarbeitungs- oder Datenverarbeitungseinheiten und integrierten Büroautomatisierungseinheiten, die eine Reihe von Geschäftszusatzprogrammen in einem einzigen integrierten Paket kombinieren, ausgeführt werden.

[0035] **Fig. 2** zeigt die äußere Erscheinung eines speziellen Ausführungsbeispiels der vorliegenden Erfindung. **Fig. 2** zeigt eine Computeranlage **10**, beispielsweise einen IBM-PC oder einen PC-kompatiblen Computer, der über ein graphisches Anwender-Schnittstellen-Betriebssystem wie das Windows-Betriebssystem der Firma Microsoft verfügt. Die Computeranlage **10** ist mit einem Anzeigebildschirm **11**, beispielsweise einem Farbbildschirm versehen. Die Computeranlage **10** verfügt des weiteren über eine Massenspeichereinheit, beispielsweise ein Computerplattenlaufwerk **14** zur Speicherung von Datendateien wie beispielsweise Schwarzweiß-, Halbton- und farbige Punktrasterbilder, und zur Speicherung von Programmdateien, die gespeicherte Programmbefehlsfolgen enthalten, womit die Computeranlage **10** Datendateien bearbeitet und Daten in solchen Dateien einer Computerbedienperson über den Anzeigebildschirm **11** darstellt. Beispielsweise können die Anwendungsprogramme, die sich auf der Speicherplatte des Computerplattenlaufwerks **14** befinden, ein Anwendungsprogramm enthalten, um auf Daten einer optischen Platte zuzugreifen, oder um andere informationsverarbeitende Programme abzuarbeiten.

[0036] Eine Tastatur **15** wird mit der Computeranlage **10** verbunden, um die Eingabe von Textdaten zu ermöglichen, und um der Bedienperson zu erlauben, die Auswahl und die Bearbeitung von Daten, die auf dem Bildschirm **11** angezeigt werden, vorzunehmen. Ähnlich wird eine Zeigereinheit **16**, beispielsweise eine Maus bereitgestellt, um die Auswahl und die Bearbeitung von Objekten auf dem Bildschirm **11** zu bearbeiten. Ein Abtastgerät **18** tastet Dokumente oder andere Bilder ab, und es liefert beispielsweise Punktrasterbilder dieser Dokumente an die Computeranlage **10**. Diese Bilder können unmittelbar von der Computeranlage **10** verwendet werden, oder die Computeranlage **10** kann diese Bilder auf das Medium des Plattenlaufwerks **14** oder auf eine optische Platte, beispielsweise Platte **19**, des optischen Laufwerks **20** abspeichern. Die Punktrasterbilddaten lassen sich von der optischen Platte **19** oder dem Medium des Plattenlaufwerks **14** für die nachfolgenden Verarbeitung durch die Computeranlage **10** ausle-

sen.

[0037] Andere Mittel zur Lieferung von Punktrasterdaten oder anderen Daten kann durch ein hier nicht gezeigtes Modem in der Computeranlage **10** über eine Telefonverbindung **21** oder eine Netzwerkverbindung **22** erfolgen. Ähnlich kann ein Diskettenlaufwerk oder eine digitale beziehungsweise analoge Videoinformation durch Eingabe in die Computeranlage **10** über eine nicht gezeigte Videoschnittstelle bereitgestellt werden.

[0038] Der Drucker **24** dient zur Ausgabe von Informationen, die von der Computeranlage **10** aufbereitet wurden.

[0039] Entsprechend den Anweisungen der Computerbedienperson und unter der Steuerung des Betriebssystems werden die abgespeicherten Anwendungsprogramme gezielt aktiviert, um Eingangsdaten zu verarbeiten und zu bearbeiten. Beispielsweise, und wie weiter unten genauer beschrieben, kann ein Zeichenerkennungsprogramm aktiviert werden, um der Bedienperson zu ermöglichen, Dokumentenbilder über das Abtastgerät **18** einzulesen, und sie auf der optischen Platte **19** abzuspeichern. Die Bedienperson kann dann auf die eingelesenen Dokumente von der optischen Platte **19** in Verbindung mit dem Betriebssystem und in Verbindung mit den Anwendungsprogrammen wieder zugreifen.

[0040] **Fig. 3** ist ein genaues Blockdiagramm eines Ausführungsbeispiels der vorliegenden Erfindung in Form einer integrierten Büroautomatisierungseinheit, die Abtasten, Bildübertragung, Informationsübertragung und -empfang ausführen kann, und die über Informationsverarbeitungseigenschaften, die Textverarbeitung sowie Datenverarbeitungseigenschaften beinhalten kann, verfügt.

[0041] Bei dem in **Fig. 3** gezeigten Gerät können Bilder über Bildübertragung in das Gerät, indem ein Originaldokument abgetastet wird, oder durch Fernempfang über ein Modem eingegeben werden. Mit der vorliegende Erfindung wird es möglich, die Zeichen beim Bild zu erkennen, um eine Computertextdatei aus erkannten Zeichen zu erstellen, und um Textverarbeitung, Kalkulationstabellenverarbeitung oder andere Informationsverarbeitungseigenschaften des Geräts zu nutzen, um die Textdatei zu verändern. Die veränderte oder unveränderte Textdatei kann nachfolgend zurückübertragen oder anderweitig ausgegeben werden, beispielsweise durch Verwendung einer Sprachsynthesetechnik, um die Textdatei über einen Lautsprecher oder über ein gewöhnliches Telefon auszugeben.

[0042] In **Fig. 3** wird eine Zentraleinheit **30**, beispielsweise ein programmierbarer Mikroprozessor mit einem Bus **31** verbunden. An den Bus angeschlossen sind das Abtastgerät **18**, um ein Bild Bildpunkt-für-Bildpunkt abzutasten und im Bildspeicher (beispielsweise RAM **32**; Random Access Memory, Schreib/Lesespeicher) abzuspeichern, ein Modem **33**, um digitale Daten in analoger Form über die Telefonleitung **21a** zu senden oder zu empfangen, und

eine Bildübertragungseinheit (Faxgerät) **34** (die zusätzlich ein Telefon beinhalten kann; hier nicht gezeigt), um Bilddaten über die Telefonleitung **21b** zu senden oder zu empfangen. Die Telefonleitungen **21a** und **21b** können die gleiche Leitung sein, oder sie können über eine hier nicht gezeigte Netzwerksteuereinheit zugeordnet werden. Darüber hinaus angeschlossen sind an den Bus **31** ein ROM **35** (Read Only Memory, Festwertspeicher) zum Speichern eines oder mehrerer Computerprogramme, die durch die Zentraleinheit **30** ausgeführt werden, eine Zeichentabelle **36** zur Speicherung einer Tabelle von Zeichen, die mit den eingegebenen Zeichen während der Erkennungsverarbeitung verglichen werden, das RAM **32** zur Speicherung von Eingabebilddaten, verarbeiteten Bilddaten, Informationen, die den Bildaufbau betreffen, eine Ausgabeinheit **37** (wie ein Plattenspeicher oder eine Sprachsyntheseinheit mit einem Lautsprecher und (wahlweise) einer Schnittstelle zum Telefon), über die die Identität der erfaßten Zeichen während der Zeichenerkennungsverarbeitung ausgegeben wird, die Drucker/Anzeige-Einheit **24** zur Ausgabe der durch das Gerät verarbeiteten Bilder, und die Tastatur **15**, um der Bedienperson eine Steuerung des Geräts zu ermöglichen.

[0043] Obgleich die an den Bus **31** angeschlossenen Einheiten in einem Büroautomatisierungswerkzeug eingebunden sind, können einige oder sämtliche Einheiten auch als Einzeleinheiten bereitgestellt werden.

[0044] Das Abtastgerät **18**, das Modem **33** und die Bildübertragungseinheit **34** bilden alternative Formen zur Eingabe von Bilddaten in das Gerät. Im Fall des Abtastgeräts **18** werden Originaldokumente Zeile-für-Zeile und Bildpunkt-für-Bildpunkt abgetastet, und unter der Steuerung der Zentraleinheit **30** werden die Bildpunkte der Bilddaten in einem Punktrasterspeicherformat im Bildspeicher (RAM **32**) abgespeichert. Im Fall des Modems **33** werden die Bilddaten über die Telefonleitung **21a** in analoger Form empfangen, und sie werden in eine digitale Bildpunktform durch das Modem **33** umgesetzt, wobei die digitalen Bildpunktdaten im Bildspeicher (RAM **32**) abgespeichert werden. Im Fall der Bildübertragungseinheit **34** werden die Bilddaten in komprimierter oder codierter Form, wie beispielsweise eine modifizierte Huffman-Lauflängencodierung, über die Telefonleitung **21b** empfangen. Die komprimierten Bilddaten werden üblicherweise durch die Bildübertragungseinheit **34** dekomprimiert, und die Zentraleinheit **30** speichert die Bildpunkte der Bilddaten im Punktrasterformat in dem Bildspeicher (RAM **32**) ab. Andere Eingabemittel sind natürlich möglich, und die Bilddaten lassen sich einfach aus einem Massespeichermedium wie einen Plattenspeicher wieder auslesen, oder sie lassen sich von einer Video- oder Standbildvideokamera ableiten.

[0045] Die Bildübertragungseinheit **34** und die Ausgabeinheit **37** stellen alternative Formen von Ausgabemitteln zur Ausgabe der zeichenerkannten Bild-

daten vom Gerät dar. Im Fall der Bildübertragungseinheit **34** können Bilder von Zeichen, die entsprechend der vorliegenden Erfindung erkenntnisverarbeitet wurden, durch Normzeichensätze ersetzt und von dem Gerät übertragen werden. Dies gestattet beispielsweise für den Empfang eines Bildes, das Zeichenbilder enthält, die Zeichenerkennung der Zeichenbilder, und das Ersetzen der erkannten Zeichen durch Normzeichensätze vor der Rückübertragung, wobei die Qualität der Bilder verminderter Qualität verbessert wird.

[0046] Das Modem **33** und die Ausgabeinheit **37** sind alternative Mittel zur Ausgabe und zur Speicherung der Identität der in den Bilddaten erkannten Zeichen, beispielsweise im ASCII-Code. Die Zeichenidentitäten können in dem Gerät (beispielsweise in einem Plattenspeicher) abgespeichert werden, oder sie können zur Übertragung an eine entfernte Stelle über das Modem **33** ausgegeben werden. Mittel können auch zur Umformatierung der Identitäten der Zeichen wie die Umwandlung der ASCII-Zeichen in das bildübertragungskompatible Format bereitgestellt werden, wodurch eine Übertragung an ein entfernt stehende Bildübertragungseinheit über das Modem **33**, ohne die Notwendigkeit der Aktivierung der Bildübertragungseinheit **34**, vorgenommen werden kann.

[0047] Die Drucker-/Anzeigeinheit **24** ist ein Mittel zur Überwachung des Fortschritts der Zeichenerkennungsverarbeitung ebenso wie die Ausgabe und die Bildung einer permanenten Aufzeichnung für jeden der Zeichenerkennungsschritte. Die Tastatur **35** dient der Steuerung der Vorgänge des Geräts in **Fig. 3** durch die Bedienperson.

[0048] **Fig. 4** zeigt in einem Flußdiagramm die allgemeine Zeichenerkennungsverarbeitung, die hilfreich zum Verständnis einer Art ist, wie die Blockauswahltechnik entsprechend der vorliegende Erfindung vorteilhaft eingesetzt werden kann. Die in **Fig. 4** gezeigten Verarbeitungsschritte werden durch die Zentraleinheit **30** mit Hilfe eines Computerprogramms, das im Programm-ROM **35** abgespeichert ist, ausgeführt.

[0049] Bei Schritt S401 werden die Bildpunktdaten in das Gerät eingegeben und im RAM **32** abgespeichert. Die Bilddaten sind eine Bildpunkt-zu-Bildpunkt-Darstellung des Bildes. Die Bildpunktdaten sind bevorzugt binäre Bildpunktdaten, das heißt, schwarze und weiße Bilddaten. Es ist jedoch für die Bilddaten auch möglich, Halbtonbilddaten zu sein, wobei jeder Bildpunkt durch einen Grauton aus einer Vielzahl von Grautönen dargestellt wird, oder für die Bilddaten, Farbbilddaten zu sein, wobei jeder Bildpunkt durch ein Mehrbit-Wort dargestellt wird, das die Farbe für den Bildpunkt codiert. In solchen und in anderen Fällen, bei denen die Bildpunktdaten keine binären Bildpunktdaten sind, sollte Schwellwertverarbeitung ausgeführt werden, um die nichtbinären Bildpunktdaten vor der Abspeicherung in das RAM **32** in binäre Bildpunktdaten umzusetzen.

[0050] Die Bildpunktdateneingabe bei Schritt S401

ist typisch für ein Hochformatbild, das heißt, ein Bild, das von der oberen linken Ecke zur unteren rechten Ecke gelesen wird. Ist das Bild nicht in dieser Weise, sondern im Querformat angeordnet, dann können die Bildpunktdaten so umgesetzt werden, daß sie für ein hochformatiges Bild typisch sind. Die Umsetzung des Bildes kann mit Hilfe einer Eingabe der Bedienperson über die Tastatur **15** durchgeführt werden.

[0051] Bei Schritt S402 werden die Bilddaten vorverarbeitet. Normalerweise wird ein Vorverarbeitungsfilter verwendet, um die Bilddaten zu verbessern, beispielsweise durch Verbesserung von Zeichen und von Bilder verminderter Qualität. Eine geeignete Bildverbesserungstechnik ist im Dokument EP-A-0595992 beschrieben.

[0052] Bei Schritt S402 kann auf Wunsch die Anzahl der Bildpunkte bei den Bildpunktdaten reduziert oder komprimiert werden, um die Geschwindigkeit der Erkennungsverarbeitung auf Kosten reduzierter Genauigkeit erhöhen. Beispielsweise kann es erwünscht sein, die Bildpunktwerte für einen $m \times n$ -Block von Bildpunkten (die Werte von m und n können unterschiedlich sein) zu mitteln, und einen einzelnen Bildpunkt, dessen Wert der Mittelwert für den $m \times n$ -Block ist, zu ersetzen.

[0053] Bei Schritt S403 wird eine Blockauswahl durchgeführt, um den Typ der Bilddaten in jedem Abschnitt des Bildes darzustellen, und um Informationsblöcke wie beispielsweise Textinformationen, graphische Informationen, Zeilenbildinformationen oder Bildinformationen zu identifizieren. Hinzu kommt, daß die Blockauswahl von Schritt S403 eine Abweichung einer hierarchischen Baumstruktur des Bildes enthält, wodurch die Teile des Bildes in der richtigen Folge, wie unten in Schritt S412 beschrieben, wiederhergestellt werden können. Beispielsweise enthält die hierarchische Baumstruktur Informationen, die erlauben zweispaltige Bilddaten mit Text von Spalte **1**, der vor dem Text von Spalte **2** steht, wiederherzustellen, damit verhindert wird, daß Text von Spalte **1** in den Text der zweiten Spalte überlesen wird. Die Blockauswahl entsprechend Schritt S403 wird unten genauer beschrieben.

[0054] Ein erster Informationsblock wird vom Bild, das im RAM **32** abgespeichert ist, ausgewählt, und bei Schritt S404 wird bestimmt, ob es sich um einen Textblock entsprechend den Blockidentitäten, die beim Blockauswahlschritt S403 bestimmt wurden, handelt. Ist der erste Block kein Textblock, wird bei Schritt S405 fortgefahren, bei dem der nächste Block ausgewählt wird, und danach wird zu Schritt S404 zurückgesprungen.

[0055] Ist in Schritt S404 der Block ein Textblock, wird bei Schritt S406 fortgefahren, bei dem der Textblock einer Einteilung in Zeilenabschnitte unterworfen wird. Bei der Einteilung in Zeilenabschnitte werden einzelne Textzeilen im Textblock von anderen Textzeilen im Textblock abgetrennt, und die auf diese Weise abgetrennten Zeilen werden hintereinander verarbeitet, wie genauer unten beschrieben wird.

[0056] Bei Schritt S407 wird jedes Zeichen in einer Zeile von anderen Zeichen in der Zeile abgetrennt, um einzelne Zeichen dem Wiederherstellungsverarbeitungsprogramm, wie weiter unten beschrieben, zur Verfügung zu stellen.

[0057] Die Erkennungsverarbeitung wird bei Schritt S408 für jedes Zeichen auf einer Zeichen-zu-Zeichen-Grundlage ausgeführt, bei dem jedes Zeichen mit Techniken, mit einem in der Zeichentabelle **36** abgespeicherten Zeichen verglichen wird. Die Einträge in die Zeichentabelle **36** gelten normalerweise nur für Einzelzeichen. Da jedoch einige Zeichenfolgen nur schwer voneinander zu trennen sind (beispielsweise "fi"), und weil einzelne Zeichen manchmal versehentlich getrennt werden (beispielsweise "j"), kann die Zeichentabelle auch andere Einträge enthalten. Speziell kann die Tabelle Paare sich berührender Zeichen, die sich nur schwer trennen lassen, enthalten. Ähnlich kann die Zeichentabelle Einträge für Zeichenteile enthalten, die häufig versehentlich getrennt werden.

[0058] Auf der Grundlage des Vergleichs wird eine Identität für das Zeichenbild ausgewählt, und die auf diese Weise ausgewählte Zeichenidentität wird im RAM **32** abgespeichert oder an die Ausgabeeinheit **21** ausgegeben. Wahlweise kann das identifizierte Zeichen auch auf der Druck-/Anzeigeeinheit **24** ausgegeben werden.

[0059] Bei Schritt S409 wird bestimmt, ob die Verarbeitung für den Textblock beendet ist. Ist die Bearbeitung noch nicht beendet, wird bei Schritt S406 (oder, wenn geeignet, bei Schritt S407) zur weiteren Verarbeitung der Einteilung in Zeilenabschnitte oder Einteilung in Zeichenabschnitte fortgefahren. Ist die Blockverarbeitung beendet, wird bei Schritt S410 bestimmt, ob die Verarbeitung für eine Seite beendet ist. Ist die Verarbeitung noch nicht beendet, wird zu Schritt S405 verzweigt, bei dem der nächste Block auf der Seite zur Verarbeitung ausgewählt wird.

[0060] Ist die Verarbeitung der Seite beendet, wird bei Schritt **S411** fortgefahren, bei dem eine Nachverarbeitung ausgeführt wird. Die Nachverarbeitung enthält Techniken wie Kontext- oder Rechtschreibprüfung, wobei die Identitäten der Zeichen, die im Erkennungsschritt S408 erkannt wurden, entsprechend dem Kontext, in dem solche Zeichen auftreten (ein ganzheitlicher Standpunkt) eher als auf einer individuellen Zeichengrundlage wie im Fall von Schritt S408, geändert werden.

[0061] Bei Schritt S412 werden die Bilddaten entsprechend der hierarchischen Baumstruktur, die im Blockauswahlschritt S403 definiert wurde, wiederhergestellt. Die Seitenwiederherstellung erlaubt, die Bilddaten in der richtigen Weise anzuordnen, wobei beispielsweise Fußnoten vom Haupttext abgetrennt werden, Spalten folgen anderen Spalten, statt miteinander vermischt zu werden, und graphische Zeichendaten oder Strichzeichnungsdaten können an den entsprechenden Stellen im erkannten Zeichentext der Seite eingefügt werden. Bildunterschriften, die

mit solchen graphischen Bildern oder Strichzeichnungen verbunden sind, können neben solchen Zeichnungen eingefügt werden. Weiter Regeln können angewendet werden, und es ist beispielsweise möglich, daß eine physische Wiederherstellung der Seite nicht erwünscht ist, sondern nur ein Textauszug der Seite. Regeln können von einem Anwender der Wiederherstellung entsprechend den Informationsverarbeitungsmöglichkeiten des Geräts definiert werden.

[0062] Bei Schritt S413 wird die wiederhergestellte Seite beispielsweise auf der Ausgabeeinheit **37** abgespeichert. Die Informationsverarbeitung wie ein Arbeitsblatt oder Textverarbeitung kann dann entsprechend mit anderen Anwendungsprogrammen, die in ROM **35** abgespeichert sind, und durch die Zentraleinheit **30** ausgeführt werden. Die verarbeitete Information (oder unverarbeitete Information, falls erwünscht) kann dann durch eins der verschiedenen Mittel, beispielsweise durch die Bildübertragungseinheit **34**, das Modem **35** oder über ein gewöhnliches Telefon, über eine Sprachsyntheseinheit, die die Computertextdatei "spricht", rückübertragen werden.

[0063] **Fig. 5** ist ein verallgemeinertes Flußdiagramm zur Erläuterung der Blockauswahlverarbeitung bei Schritt S403 in **Fig. 4**. Es sollte darauf hingewiesen werden, daß die Blockauswahlverarbeitung unabhängig von einem Zeichenerkennungssystem verwendet werden kann. Es ist beispielsweise möglich, die Blockauswahlverarbeitung bei der Bildwiederherstellungsausrüstung zu verwenden, um ein erstes Verfahren der Bildwiederherstellung zu veranlassen, für einen Blocktyp und ein zweites Verfahren der Bildwiederherstellung für einen zweiten Blocktyp verwendet zu werden, oder zur Verwendung der Blockauswahlverarbeitung bei Datenkompressionsschemata.

[0064] Bei Schritt S50 kann Bildverbesserung oder Datenreduktion ausgeführt werden, wenn gewünscht wird, die Blockauswahlverarbeitungsgeschwindigkeit zu erhöhen. Wird Bilddatenreduktion ausgeführt, dann wird Blockauswahl bei dem verdichteten Bild ausgeführt. Damit jedoch der Zeichenerkennungsverarbeitung im Rest der **Fig. 4** (das heißt, Schritt S404 bis Schritt S413) nicht beeinflußt wird, werden am Ende der Blockauswahlverarbeitung die ausgewählten Blöcke gegenüber den nicht verdichteten Bildpunktdaten angewendet.

[0065] Die Bilddatenreduktion fährt bei der Auswertung der Vernetzungsfähigkeit der schwarzen Bildpunkte bei den $m \times m$ Blöcken der Bildpunkte fort. Wenn dort beispielsweise zwei verbundene schwarze Bildpunkte bei einem 3×3 -Bildpunktblock verbunden sind, dann wird der 3×3 -Block auf einen schwarzen Bildpunkt reduziert. Sind umgekehrt zwei schwarze Bildpunkte in dem 3×3 -Block nicht verbunden, dann wird der 3×3 -Block zu einem einzigen weißen Bildpunkt reduziert. Bevorzugt werden schwarze Bildpunkte in 4×4 -Blöcken ausgewertet. Das Prinzip ist, die Vernetzungsfähigkeit des Ur-

sprungsbilds aufrechtzuerhalten.

[0066] Bei Schritt S51 wird das Bildpunktdarstellung analysiert, um verbundene Komponenten zu erfassen und um verbundene Komponenten entsprechend ihrer Größe und ihre relative Lage zu anderen verbundene Komponenten einzuordnen. Eine verbundene Komponente ist eine Gruppe von schwarzen Bildpunkten die vollständig von weißen Bildpunkten umgeben ist. Entsprechend ist eine verbundene Komponente eine Gruppe von schwarzen Bildpunkten, die vollständig von anderen Gruppen schwarzer Bildpunkte durch mindestens einen weißen Bildpunkt vollständig getrennt ist. wie weiter unten ausführlicher in Verbindung mit **Fig. 6** erläutert wird, beinhaltet Schritt S51 das Erfassen der verbundene Komponenten (verbunden in jeder der acht Richtungen, die einen Bildpunkt umgeben). Bei Schritt S52 wird auf der Grundlage der Größeninformation und einiger von den verbundenen Komponenten erhaltenen statistischen Werte (eine genauere Beschreibung erfolgt unten) die Zuordnungen jeder verbundenen Komponente zugewiesen. Wie weiter unten genauer erläutert, wird jede verbundene Komponente entweder einer Texteinheit oder einer Nichttexteinheit zugeordnet. Nichttexteinheiten werden bei Schritt S53 einer weiteren Analyse unterzogen, um zu bestimmen, ob es sich um Bilddaten, Halbtonbilder, Strichzeichnungen, Tabellen oder andere tabellarische Anordnungen von Textdaten, vertikalen Zeilen, horizontalen Zeilen, vertikale oder horizontale schiefen Zeilen auf der Grundlage ihrer Stärke und ihrer Fülle oder ob es sich um eine unbekannte Zuordnung handelt. Es wird für jede der verbundene Komponenten eine hierarchische Baumstruktur entwickelt, um organisatorische Daten der verbundene Komponenten zu liefern, und um die Wiederherstellung der digitalen Daten, wie oben unter Bezugnahme auf Schritt S412 beschrieben, zu vereinfachen.

[0067] Bei Schritt S54 werden unsichtbare Zeilen (weiße Zeilen) längs der Ränder von verbundenen Nichttextkomponenten gesucht. Der Grund für die Verwendung solcher unsichtbaren Zeilen ist, daß der Abstand zwischen einigen Textspalten viel enger als der übliche Abstand der Spalten auf der gleichen Seite ist. Die Erfassung derartiger unsichtbaren Zeilen wird später bei der Bestimmung hilfreich sein, ob nebeneinanderliegende verbundene Textkomponenten tatsächlich in den gleichen Textblock gruppiert werden sollen.

[0068] Bei Schritt S55 werden sämtliche Nichttexteinheiten, die in Schritt S51 nicht zugeordnet werden konnten, analysiert, um zu bestimmen, ob es sich um Titel mit großer Schriftgröße handelt. Werden sie als Titel bestimmt, werden die Einheiten entsprechend umbenannt, und die Baumstruktur wird neu aufbereitet. Titel sind hilfreich bei dem Neuaufbau der Seite entsprechend Schritt S412. Titelzeilen werden entweder in horizontaler oder in vertikaler Richtung durch Gruppierung der unbekanntes Nichttext- und Textkomponenten, die in den Schritten S52 und S53

ermittelt wurden, gebildet. Vor dieser Gruppierung wird die Gruppierungsrichtung auf der Grundlage des Abstands zwischen den Komponenten, die gruppiert werden sollen, und der möglichen Länge der Titelzeile in vertikaler oder horizontaler Richtung bestimmt.

[0069] Bei Schritt S56 werden Textblöcke von den verbundenen Textkomponenten, wie weiter unten genauer beschrieben, gebildet. Bei Schritt **57** kann ein gebildeter Textblock aufgeteilt werden, wenn eine sichtbare oder unsichtbare Zeile hindurchläuft. Der Grund für die Verwendung der Zeilen liegt nicht nur in dem extrem engen Spaltenzwischenraum, sondern auch, weil der Abstand zwischen den verbundenen Textkomponenten der verschiedenen Blöcke im Falle der Schiefe der Komponenten, die in einem Block gruppiert werden, sehr verkürzt werden können, insbesondere, wenn die unsichtbaren und die sichtbaren Zeilen die Trennstücke zwischen ihnen sind.

[0070] Bei Schritt S58 wird die Richtung jedes Textblocks erfaßt. Das heißt, entsprechend vorliegender Erfindung wird keine Erfassung der Schiefe oder der Bildrotation ausgeführt, bevor nicht die Textblöcke gebildet sind. Das heißt, im Gegensatz zu den meisten Blockauswahlalgorithmen bildet die vorliegenden Erfindung unmittelbar Textblöcke des Seitenabbilds, ohne zuvor irgendwelche Erfassung der Schiefe und eines Bildrotationsvorgangs anzuwenden. Dies spart Zeit ein, die für die Bildrotation benötigt wird, und vermeidet Ungenauigkeiten aufgrund der Erfassung der Schiefe. Darüber hinaus ist es möglich, daß einige der Textbereiche der gleichen Seite als Kursivschrift speziell aufbereitet werden, um einige speziellen visuellen Effekte zu erzielen. In solch einem Fall trägt die Erfassung der Schiefe und die Bildrotation nicht dazu bei, das Seitenabbild in eine einwandfrei Form zu korrigieren. Das heißt, entsprechend der vorliegenden Erfindung werden die Text- und Nichttextblöcke unmittelbar gesucht, ohne die Schiefe der abgetasteten Seite zu berücksichtigen. Nachdem die Textblöcke gebildet sind, kann dann die Schiefe jedes Textblocks speziell zugeordnet werden.

[0071] Bei Schritt S59 werden die Textzeilen für jeden Textblock auf der Grundlage der Richtung des Textblocks gebildet. Bei Schritt S60 wird eine Nachbearbeitung ausgeführt, um die Text- und Nichttextdarstellung kompakter und klarer zu machen. Beispielsweise erzeugten frühere Blockauswahlalgorithmen Textblöcke, die durch Rechtecke dargestellt wurden. Derartige Rechtecke können jedoch die Trennung von Textblöcken im Fall eines schiefen Dokuments undeutlich werden lassen. Um die Trennung zwischen den Textblöcken hervorzuheben, wenn sich ihre rechteckigen Bereiche überlappen, liefert die vorliegenden Erfindung eine unregelmäßige Kurve, um jeden Textblock dichter zu umschließen. Die umschlossene Kurve ist nicht nur zur Verbesserung der Anzeigeffekte nützlich, sondern auch zum Abtrennen des Blockbilds. Die unregelmäßige Kurve, die den Textblock umhüllt, kann auch im Textblockbildungsschritt S56 erzeugt werden, und zwar immer dann,

wenn ein Textblock neu geschaffen oder aktualisiert wird.

[0072] **Fig. 6**, die aus den **Fig. 6A** bis **6D** besteht, ist ein genaues Flußdiagramm, um aufzuzeigen, wie verbundene Komponenten von Bildpunktdaten erfaßt werden, und wie diese verbundenen Komponenten zugeordnet werden, das heißt, die oben erläuterten Schritte **S51** bis **S53**. Die in **Fig. 6** gezeigten Verfahrensschritte werden von der Zentraleinheit **30** entsprechend den Programmschritten, die im Programm-ROM **35** abgespeichert sind, ausgeführt.

[0073] Bei Schritt **S601** werden die verbundene Komponenten in den Bildpunktdaten durch Konturverfolgung erfaßt. Konturverfolgung erfolgt durch Abtasten der Bilddaten, wie in **Fig. 7a** gezeigt. Das Abtasten erfolgt von der unteren rechten Seite des Bildes nach links, wie durch den Pfeil **A** angezeigt, und erfolgt nach oben, wenn der rechte Rand des Bildes angetroffen wurde. Das Abtasten kann auch in anderen Richtungen erfolgen, beispielsweise von oben links nach unten rechts. Wird ein geschwärtzter Bildpunkt angetroffen, werden nebeneinanderliegende Bildpunkte in dem bei **71** gezeigten sternförmigen Muster geprüft, um zu bestimmen, ob ein neben dem geschwärtzten Bildpunkt angeordneter Bildpunkt ebenfalls schwarz ist. Da das sternförmige Muster **71** acht durchnummerierte Vektoren enthält, die sich von einem gemeinsamen Zentrum strahlenförmig ausbreiten, wird diese Konturverfolgung künftig als eine Verfolgung in "acht Richtungen" bezeichnet. Wird ein benachbarter, geschwärtzter Bildpunkt lokalisiert, fährt die Verarbeitung, wie oben angedeutet, so lange fort, bis die äußere Kontur des Bildes verfolgt wurde. Als Folge, wie in **Fig. 7B** gezeigt, hat das Abtasten in Richtung des Pfeils **A** den Punkt **72** lokalisiert, der zu der Untertlänge des Buchstabens "Q" gehört. Die Untersuchung der benachbarten Bildpunkte wird entsprechend dem sternförmigen Muster **71** fortgesetzt, wodurch die äußere Kontur des Buchstabens "Q" verfolgt wird. Die inneren Teile der geschlossenen Kontur werden nicht verfolgt.

[0074] Nachdem eine verbundene Komponente erfaßt und die zugehörige Kontur durch Verfolgung in acht Richtungen verfolgt ist, wird mit dem Abtasten fortgefahren, bis der nächste geschwärtzte Bildpunkt lokalisiert wird. Als Folge wird beispielsweise das Objekt **74**, das eine vollständig schwarze Fläche darstellen kann, in acht Richtungen verfolgt. Ähnlich wird das Nichttextobjekt **75**, das handgeschrieben das Wort "Nichttext" ist, verfolgt, ebenso wie sich in den Textobjekten **76** die Einzelbuchstaben befinden, die Einzelzeichen sind, die das Wort "TEXT" bilden. Das in **Fig. 7A** gezeigt Abtasten setzt sich so lange fort, bis sämtliche verbundene Komponenten bei den Bildpunktdaten erfaßt wurden, und die entsprechenden Konturen in acht Richtungen erfaßt sind.

[0075] Bei Schritt **S602** wird jede verbundene Komponente in ein Rechteck umgewandelt. Speziell wird das kleinst mögliche umschreibende Rechteck um jede verbundene Komponente gezogen. Als Folge,

wie in **Fig. 7B** gezeigt, wird ein Rechteck **77** um das Objekt **72** gezeichnet, ein Rechteck **79** wird um das Objekt **74** gezeichnet, ein Rechteck **80** wird um das Objekt **75**, und die Rechtecke **81a**, **81b**, **81c** und **81d** werden um die Objekte **76a**, **76b**, **76c** beziehungsweise **76d** gezeichnet.

[0076] Bei Schritt **S603** wird jeder rechteckige Einheit eine Baumstruktur zugewiesen. Für den größten Teil führt die in Schritt **S603** erhaltene Baumstruktur unmittelbar von der Wurzel des Baums zu jedem Objekt in der Bildpunktabbildung. Dies beruht darauf, daß nur die äußeren Konturen der verbundene Komponenten verfolgt werden, und die inneren Teile der geschlossenen Konturen nicht verfolgt werden. Als Folge, wie in **Fig. 7C** gezeigt, führt das Rechteck **77**, das der verbundene Komponente **72** entspricht, unmittelbar von der Wurzel der Seite weg. Für solche verbundene Komponenten, deren Rechtecke ganz innerhalb des Rechtecks einer anderen verbundene Komponente, wie beispielsweise Rechteck **80**, das zu dem Nichttextobjekt **75** gehört, und die Rechtecke **81a** und **81b**, die zu den Textobjekten **76a** und **76b** gehören, werden derartige verbundene Komponenten als Nachkomme der eingeschlossenen verbundene Komponente (in diesem Fall Komponente **74**) bezeichnet. Hinzu kommt: Für jede verbundene Komponente mit mindestens einem Nachkommen, wie beispielsweise die Komponente **74**, wird die Komponente selbst als ein "Hauptnachkomme" von sich selbst bezeichnet. Folglich, wie **Fig. 7c** zeigt, ist die Komponente **79** als ein Hauptnachkomme gegenüber den anderen Nachkommen **80**, **81a** und **81b** der Komponente **79** enthalten.

[0077] Bei Schritt **S604** wird jede verbundene Komponente auf der ersten Ebenen des Baums entweder als eine Texteinheit oder als eine Nichttexteinheit zugeordnet. Die Zuordnung erfolgt in zwei Schritten. Beim ersten Schritt, wird das Rechteck für die verbundene Komponente mit den vorgegeben Größen der Schwellwerte verglichen. Ist die Höhe und die Breite des die verbundene Komponente enthaltenden Rechtecks größer als ein erster vorgegebener Schwellwert, der zu der maximal erwarteten Schriftartgröße gehört, dann wird die verbundene Komponente als eine Nichttexteinheit zugeordnet, und ein "Nichttext"-Merkmal wird der Einheit angefügt.

[0078] Beim zweiten Schritt werden sämtliche verbleibenden Einheiten, das heißt, Einheiten, die bis jetzt nicht als Nichttext zugeordnet sind, mit den Schwellwerten verglichen, die adaptiv auf der Grundlage der gemeinsamen Größe sämtlicher verbleibender verbundener Komponenten bestimmt werden. Speziell die größere der üblichen Textgröße (12 pt) und die mittlere Größe (die kleinere der mittleren Höhe und mittleren Breite) sämtlicher Rechtecke, die noch nicht als Nichttext bestimmt wurden, werden ausgewählt. Die ausgewählten Werte werden mit einem Faktor multipliziert (üblicherweise "1,5"), um adaptiv bestimmte Schwellwerte für sowohl die Höhe als auch die Breite zu erhalten. Bei jeder Einheit, die

größer als sämtliche der bestimmten Schwellwerte ist, wird davon ausgegangen, daß es Nichttext ist, und sie wird entsprechend zugeordnet, während eine jede Einheit, die kleiner als die adaptiv bestimmten Schwellwerte ist, und die die vorgegebenen Texteingenschaften erfüllt, wird davon ausgegangen, daß es Text ist. Die Einheiten werden entsprechend klassifiziert, und das zugehörige Merkmal wird angehängt. Beide dieser Zuordnungen werden der Verfeinerung unterworfen, vorgegeben im Rest der **Fig. 6**, und sie werden unten noch genauer beschrieben.

[0079] Nachdem jede Einheit auf der ersten Ebenen des Baums als Text oder als Nichttext zugeordnet ist, werden sämtliche Texteinheiten, einschließlich der Hauptnachkommen, als Texteinheiten zugeordnet. Die Zuordnung von Hauptnachkommen von Nichttexteinheiten wird als Nichttext zurückgehalten, jedoch sämtliche anderen Nachkommen von Nichttexteinheiten werden als Texteinheiten zugeordnet.

[0080] Bei Schritt S605 wird die erste Einheit ausgewählt. Ist die Einheit bei Schritt S606 eine Texteinheit, wird bei Schritt S607 fortgefahren, bei dem die Texteinheit ausgewählt wird. Die Schritte S606 und **607** werden so lange durchlaufen, bis eine Nichttexteinheit ausgewählt wird, wobei bei Schritt S608 fortgefahren wird.

[0081] Bei Schritt S608 wird die Nichttexteinheit geprüft, um zu bestimmen, ob es Nachkommen von der Einheit gibt. Beispielsweise, wie in **Fig. 7c** gezeigt, enthält die Nichttexteinheit **79** einen Nichttexthauptnachkommen **79** sowie die Textnachkommen **80**, **81a** und **81b**.

[0082] Werden bei Schritt S608 Nachkommen festgestellt, wird bei Schritt S609 fortgefahren, bei dem die Einheit gefiltert wird, um zu bestimmen, ob die Einheit eine Halbton- (oder Grauskalen-) Einheit ist. Bei der Halbtonfilterung werden die Nachkommen der Einheit geprüft, und die Anzahl der Nachkommen, deren Größe kleiner als eine "Stör"-Größe ist, wird bestimmt. Eine "Störgrößen"-Einheit ist eine Einheit, deren Höhe und Breite kleiner ist als die kleinste Zeichensatzgröße, die für Bilddaten erwartet wird. Wenn die Anzahl der Nachkommen, deren Größe kleiner als eine Störgröße ist, größer als die Hälfte der Gesamtzahl von Nachkommen ist, wird entschieden, daß die Einheit ein Halbtonbild ist. Wird in Schritt S610 ein Halbtonbild erkannt, wird zu Schritt S611 verzweigt, wobei ein "Halbton"-Merkmal an die Einheit angehängt wird. Anschließend wird bei Schritt S607 fortgefahren, wobei die nächste Einheit zur Verarbeitung ausgewählt wird.

[0083] Wird bei Schritt S609 bei der Halbtonfilterung festgestellt, daß die Einheit kein Halbtonbild ist, dann wird über den Schritt S610 zum Schritt S613 weitergegangen, bei dem der Hauptnachkomme von der Einheit zur weiteren Verarbeitung ausgewählt wird. Anschließend wird bei Schritt S614 fortgefahren.

[0084] Wird bei Schritt **608** eine Nichttexteinheit erkannt, die keine Nachkommen hat, oder wurde bei Schritt S613 kein Hauptnachkomme zur weiteren

Verarbeitung ausgewählt, dann wird bei Schritt S614 die in Frage kommende Einheit der Rahmenfilterung unterworfen. Rahmenfilterung wird entwickelt, um zu bestimmen, ob die in Frage kommende Einheit ein Rahmen ist, und die Erfassung von parallelen horizontalen und parallelen vertikalen Zeilen enthält, die annähernd die gleiche Breite und Höhe wie das Rechteck, das die Einheit umgibt, haben. Spezielle wird die verbundene Komponente geprüft, um für jede Zeile beim Bildpunkt den längsten Abstand, der sich über den inneren Teil der verbundene Komponente spannt, bei der Einheit zu bestimmen. Folglich, wie in **Fig. 8A** gezeigt, enthält die Nichttexteinheit **82** die verbundene Komponente **83**, deren Kontur durch eine Verfolgung in acht Richtungen verfolgt wurde, wie bei **83** gezeigt. Für die i-te Zeile ist der längste Abstand, der sich über den inneren Teil der verbundene Komponente spannt, das heißt, der Abstand von dem äußersten linken Rand **85a** der Kontur bis zum äußersten rechten Rand **85b**. Bei der j-ten Zeile existieren zwei Abstände, die über das Innere der verbundene Komponente spannen; der Abstand zwischen den Punkten **86a** und **86b** auf dem Rand der verbundene Komponente und dem Abstand zwischen den Punkten **87a** und **87b**. Da der Abstand zwischen den Punkten **86a** und **86b** größer wie der Abstand zwischen den Punkten **87a** und **87b** ist, ist der Abstand x_j der längste Abstand, der sich im Inneren der verbundene Komponente für die j-te Zeile spannt.

[0085] Für jede der n Zeilen bei der Nichttexteinheit **82** wird ein "x"-Abstand hergeleitet, und die folgende Ungleichung wird zur Bestimmung, ob die Nichttexteinheit ein Rahmen ist, geprüft:

$$\sum_{i=1}^N \frac{(X_k - W)^2}{N} < \text{Schwellwert}$$

wobei X_k der längste Abstand ist, der sich über das Innere der verbundene Komponente für die k-te Zeile (wie oben beschrieben) spannt, W ist die Breite der rechteckigen Einheit **82**, N ist die Anzahl der Zeilen, und der Schwellwert wird im Voraus berechnet, um die Erfassung der Rahmen zu erlauben, selbst wenn die Rahmen bei den Bilddaten schief oder gekippt sind. Um einen schiefen Winkel oder einen Kippwinkel von 1° zuzulassen, wurde ein Schwellwert von $\sin(1^\circ)$ mit L multipliziert und mit einem Versatz addiert, der gleich der in Schritt S604 berechneten mittleren Texthöhe ist, herausgefunden, der zufriedenstellende Ergebnisse erzielt.

[0086] Ist die oben angeführten Ungleichung erfüllt, dann wird die Einheit als Rahmendatum bestimmt, und im Flußdiagramm wird bei den Schritten S615 und S616 fortgefahren, bei denen ein "Rahmen"-Merkmal an die Einheit angehängt wird.

[0087] Im Anschluß an Schritt S616 wird die Möglichkeit betrachtet, daß die gerahmten Daten eine Tabelle oder tabellarisch angeordnete Daten enthält. Folglich wird in Schritt S617 (**Fig. 6B**) das Innere der

verbundenen Komponente geprüft, um weiße Konturen zu erhalten.

[0088] Weiße Konturen sind ähnlich den Konturen, die in Schritt S601 erfaßt wurden, jedoch weiße Bildpunkte werden eher als schwarze Bildpunkte geprüft. Folglich, wie in **Fig. 9A** gezeigt, wird das Innere der Nichttexteinheit in Richtung des Pfeils B vom unteren rechten Teil des Inneren der Nichttexteinheit zum oberen linken Teil abgetastet. Wird ein erster weißer Bildpunkt angetroffen, werden benachbarte Bildpunkte in der Reihenfolge, wie im sternförmigen Muster **91** gezeigt, geprüft. Es sei angemerkt daß die sternförmigen Muster **91** durchnummerierte Vektoren von **1** bis **4** enthalten. Entsprechend wird weiße Konturverfolgung bei diesem Schritt im folgenden als weiße Konturverfolgung in "vier Richtungen" bezeichnet. Die weiße Konturverfolgung in vier Richtungen wird fortgeführt, bis sämtliche weißen Konturen, die in schwarzen Bildpunkten eingeschlossen sind, verfolgt sind. Beispielsweise folgt die weiße Konturverfolgung Bildpunkten, die die innere Kontur der schwarzen Bildpunktsegmente **92, 93, 94** und **95** ebenso wie andere schwarze Bildpunkte im Inneren wie beispielsweise die schwarzen Bildpunkte, die allgemein bei **96** bezeichnet werden. Nachdem jeder weißen Kontur geortet ist, fährt die Abtastung, wie oben, in Richtung des Pfeils B fort, bis sämtliche weißen Konturen, die in den Nichttextobjekten eingeschlossen waren, verfolgt wurden.

[0089] Bei Schritt S618 wird die Dichte der Nichttexteinheiten berechnet. Die Dichte wird durch Zählen der Anzahl der schwarzen Bildpunkte innerhalb der verbundene Komponente und durch Division der Anzahl der schwarzen Bildpunkte durch die Gesamtzahl der im Rechteck eingeschlossenen Bildpunkte berechnet.

[0090] Bei Schritt S619 wird die Anzahl der weißen Konturen innerhalb der Nichttexteinheit geprüft. Ist die Anzahl der weißen Konturen größer oder gleich dem Wert 4, ist die Möglichkeit gegeben, daß das Nichttextbild in Wirklichkeit eine Tabelle oder eine tabellarische Anordnung von Textblöcken ist. Entsprechend wird in Schritt S620 (s. **Fig. 6D**) der Füllanteil der weißen Konturen bestimmt. Der weiße Konturfüllanteil ist der Grad, zu dem die weißen Konturen den Bereich, der beim Nichttextbild eingeschlossen ist, ausfüllen. Folglich, wie in **Fig. 9A** gezeigt, enthalten die von links oben nach rechts unten gestreiften Bereiche weiße konturausfüllende Anteile wie **97** und **99**, die vollkommen leere, weiße Räume sind, ebenso wie beispielsweise die Bereiche **100** und **101**, die weiße Räume, in denen sich schwarze Bildpunkte befinden, sind. Ist der Füllanteil groß, dann besteht die Möglichkeit, daß das Nichttextbild eine Tabelle oder eine tabellarisch angeordnete Folge von Textdaten ist. Entsprechend wird bei Schritt S621 der Füllanteil geprüft. Ist der Füllanteil groß, ist es wahrscheinlich, daß das Nichttextbild eine Tabelle oder eine tabellarisch angeordnete Folge von Textdaten ist. Um die Aussage dieser Festlegung zu erhöhen,

werden die weißen Konturen geprüft, um zu bestimmen, ob sie eine rasterförmige Struktur haben, die sich sowohl horizontal als auch vertikal ausdehnt. Speziell bei Schritt S622 werden nichttrasterförmig angeordnete weiße Strukturen wiedervereint, wenn sich ihre Begrenzungen nicht in horizontaler und vertikaler Richtung über mindestens zwei Konturen ausdehnen. Beispielsweise, wie in **Fig. 9A** gezeigt, dehnt sich die linke Begrenzung **102** und die rechte Begrenzung **103** der weißen Kontur **99** vertikal aus, wodurch sie mit der linken Begrenzung **104** und der rechten Begrenzung **105** der weißen Kontur **100** zusammenfällt. Entsprechend, da diese weißen Konturen in einer Gitterstruktur angeordnet sind, werden diese weißen Konturen nicht wiedervereint. Ähnlich dehnen sich die obere Begrenzung **106** und die untere Begrenzung **107** der weißen Kontur in horizontaler Richtung aus, wodurch sie mit der oberen Begrenzung **108** und der unteren Begrenzung **109** der weißen Kontur **110** zusammenfallen. Da diese weißen Konturen entsprechend in einer rasterähnlichen Struktur angeordnet sind, werden die weißen Konturen nicht wiedervereint.

[0091] Die **Fig. 9B** und **9C** dienen zur Erklärung einer Situation, bei der weiße Konturen wiedervereint werden. **Fig. 9B** zeigt eine Nichttexteinheit **111**, die beispielsweise durch Schwellwertbildung eines Halbtonbilds in ein Binärbild, wie bei Schritt S401 beschrieben, gebildet werden kann. Das Nichttextbild **111** enthält einen schwarzen Bereich **112** ebenso wie die weißen Bereiche **114, 115, 116, 117, 118** und **119**. Vermutlich ist der Füllanteil dieser weißen Bereiche ausreichend hoch, wodurch bei Schritt S621 bei Wiedervereinigungsschritt S622 fortgefahren wird. Als erstes, wie in **Fig. 9B** gezeigt, werden die oberen und die unteren Begrenzungen der weißen Kontur **115** mit den oberen und den unteren Begrenzungen der weißen Kontur **117** verglichen. Da diese oberen und unteren Begrenzungen nicht zusammenfallen, wird die weiße Kontur **115** mit der weißen Kontur **116**, wie in **Fig. 9C** gezeigt, kombiniert, um die weiße Kontur **116'** zu schaffen.

[0092] In **Fig. 9C** werden die linken und die rechten Begrenzungen der weißen Kontur **117** mit den linken und den rechten Begrenzungen der weißen Kontur **118** verglichen. Da diese Begrenzungen nicht die gleichen sind, werden die weißen Konturen **117** und **119** zu einer einzigen weißen Kontur **117'** rekombiniert, wie in **Fig. 9D** zu sehen ist.

[0093] Dieser Vorgang wird in horizontaler und in vertikaler Richtung so lange wiederholt, bis keine Rekombinationen mehr auftreten.

[0094] Folglich, wie oben erklärt, lassen sich weiße Konturen für eine Tabelle wahrscheinlich nicht rekombinieren, wohingegen sich weiße Konturen für keine vorliegenden Tabelle, beispielsweise ein Halbtonbild oder eine Strichzeichnung mit größerer Wahrscheinlichkeit rekombinieren lassen. Entsprechend wird bei Schritt S623 der Rekombinationsanteil geprüft. Ist der Rekombinationsanteil hoch oder ist der

Füllanteil bei Schritt S621 niedrig, ist es wahrscheinlich, daß die Nichttextrahmeneinheit eine schiefe Linie, ein unbekanntes Halbtonbild oder eine Strichzeichnung ist, und im Flußdiagramm wird bei Schritt S628 fortgefahren, bei dem die Nichttexteinheit nicht als eine Tabelle bezeichnet wird, und es wird bei Schritt S642A (**Fig. 6B**), wie unten beschrieben, fortgefahren.

[0095] Ist bei Schritt S623 der Wiedervereinigungsanteil nicht hoch, wird bei Schritt S624 fortgefahren, bei dem der Nichttextrahmen als eine "Tabelle" bezeichnet wird. Bei Schritt S625 wird das Innere der neu geschaffenen Tabelle geprüft, um die verbundenen Komponenten in acht Richtungen zu erfassen und zu klassifizieren. Bei Schritt S626 wird die hierarchische Struktur entsprechend mit den neuen, intern verbundenen Komponenten auf den neuesten Stand gebracht. Bei Schritt S627 werden die intern verbundenen Komponenten neu als Text oder Nichttext klassifiziert, und das entsprechende Merkmal wird angehängt, wie es oben bezogen auf die Schritte S602 bis S604 geschehen ist. Im Anschluß an Schritt S627 können die nichttextverbundenen Komponenten innerhalb einer Tabelle weiter klassifiziert als Strichzeichnungsbilder, Halbtonbilder oder unbekannte Bilder bezüglich Größe oder Dichte sein; die nichttextverbundenen Komponenten innerhalb eines Rahmens und innerhalb eines Strichzeichnungsbilds können weiter klassifiziert werden, wie schon bei den Schritten S608 ff diskutiert. Der Programmablauf verzweigt anschließend zu Schritt S607, bei dem die nächste Einheit ausgewählt wird.

[0096] Wenn, bei Rückkehr zu den Schritten S621 und S623, bei Schritt S621 der Füllanteil nicht hoch ist oder wenn bei Schritt **S623** der Rekombinationsanteil hoch ist, ist es wahrscheinlich, daß die Nichttextbildeinheit eine schiefe Zeile, eine Unbekannte, ein Halbtonbild oder eine Strichzeichnung, das heißt, keine Tabelle, ist.

[0097] Wenn, bei Rückkehr zu Schritt S619, die Anzahl der weißen Konturen nicht größer als vier ist, wird die Bildeinheit nicht als Tabelle betrachtet. Entsprechend wird im Programm bei Schritt S642A fortgefahren, bei dem bestimmt wird, ob der Rahmen und die Dichte kleiner als ein Schwellwert von etwa 0,5 ist. Der Schwellwert wird auf der Grundlage der Erwartung ausgewählt, daß die Texteinheiten oder die Strichzeichnungen innerhalb eines Rahmens weniger als die Hälfte der Bildpunkte belegen soll. Sind Rahmen und Dichte kleiner als der Schwellwert, wird zu den Schritten S625 bis S627 zurückgekehrt und danach verzweigt das Programm zu Schritt S607, um eine nächste Einheit auszuwählen.

[0098] Wenn bei Schritt S642A der Rahmen und die Dichte nicht kleiner als ein vorgegebene Schwellwert sind, dann wird bei Schritt S642 fortgefahren, um festzustellen, ob die Rahmeneinheit als eine Strichzeichnung oder als ein Halbtonbild klassifiziert werden kann, oder ob das Bild nicht klassifiziert werden kann (das heißt, der Rahmen ist "unbekannt"), wie

weiter unten genauer beschrieben wird.

[0099] Wenn bei Rückkehr zu Schritt S615 (**Fig. 6A**) die Rahmenfilterung bei Schritt S614 keinen Rahmen in der Nichttexteinheit erfaßt, wird bei Schritt S635 (**Fig. 6B**) fortgefahren, um zu bestimmen, ob die Nichttexteinheit eine Zeile enthält. Zeilen sind nützliche Nichttexteinheiten, um Textbegrenzungen aufzuzeigen. Da jedoch Text, der durch solche Zeilen gebunden ist, häufig in sehr enger Nachbarschaft zu den Zeilen erscheint, ist es für den Text möglich, der Zeile zugeordnet zu werden. Entsprechend wird Zeilenerfassung entworfen, um Zeilen mit oder ohne Textzuordnung zu erfassen.

[0100] Um eine Zeile ohne Textzuordnung zu erfassen, wird ein Histogramm der Nichttexteinheit in der längsseitigen Richtung der Einheit berechnet. Wie **Fig. 8B** zeigt, sollte das Histogramm **88** einer Zeile einen ziemlich einheitlichen Verlauf haben, wobei deren Höhe in etwa gleich der Breite der Zeile ist. Die Breite der Zeile ist etwa gleich der Breite ("W") der Nichttexteinheit; jeder Unterschied aufgrund eines schiefen Winkels θ_s , der auftritt, wenn das Originaldokument schief ist, wenn das Bild der Bildpunkte gebildet wird. Um entsprechend zu bestimmen, ob die Nichttexteinheit eine Linie enthält wird die Höhe **89** jeder Zelle_k im Histogramm mit der Breite w der Nichttexteinheit verglichen. Die quadratische Mittelwertabweichung zwischen diesen Werten wird mit einem Schwellwert folgendermaßen verglichen:

$$\sum_{k=1}^N \frac{(Zelle_k - W)^2}{N} < \text{Schwellwert}.$$

[0101] Der Schwellwert wird berechnet, um eine Schiefe oder eine Neigung θ_s der Linie der innerhalb Nichttexteinheit zuzulassen. Für eine Schiefe oder eine Neigung von 1° wurde ein Schwellwert von

$$\sum_{k=1}^N \left[\frac{k \sin(1^\circ)}{N} \right]^2$$

gefunden, um ein zufriedenstellendes Ergebnis zu erzielen.

[0102] Wenn eine Zeile ohne Anschlüsse entsprechend der oben angeführten Ungleichung nicht gefunden wird, wird eine Bestimmung vorgenommen, ob die Einheit eine Zeile mit Anschlüssen enthält. Um zu bestimmen, ob eine Zeile mit Anschlüssen in der Nichttexteinheit enthalten ist, wird die Texteinheit geprüft, um zu bestimmen, ob sich eine Zeile in Längsrichtung entlang des Randes der Einheit ausdehnt. Wenn sich speziell eine Zeile in Längsrichtung durch die Einheit ausdehnt, dann wird der Rand der rechtwinkligen Umschreibung der Einheit sehr nahe an einer anderen Zeile liegen, wie **Fig. 8C** zeigt. Entsprechend wird die Gleichmäßigkeit, mit der der erste schwarze Bildpunkt innerhalb der Grenze des Rechtecks liegt, geprüft, indem die Summe der Quadrate der Entfernung vom Rand berechnet wird. Folglich,

unter Bezugnahme auf **Fig. 8C**, wird die folgende Ungleichheit geprüft:

$$\sum_{k=1}^N \frac{X_k^2}{N} < \text{Schwellwert.}$$

[0103] Ist die Summe der Quadrate kleiner als der vorgegebene Schwellwert, dann wird eine Zeileneinheit mit Anschlüssen gefunden. Der gleiche Schwellwert wie der oben vorgegebene für Zeilen ohne Schlüsse wurde gefunden, um ein zufriedenstellendes Ergebnis zu erzielen.

[0104] Ist in Schritt S635 eine Zeile erfaßt, wird im Programmablauf über Schritt S636 bei Schritt S637A fortgefahren, bei dem eine "Zeilen"-Merkmal an die Nichttexteinheit angefügt wird. Anschließend wird bei Schritt S607 (**Fig. 6A**) fortgefahren, bei dem die nächst Texteinheit ausgewählt wird.

[0105] Wenn andererseits eine Zeile bei Schritt S635 nicht erfaßt wird, wird über Schritt S636 bei Schritt S637 fortgefahren, bei dem die Größe der Nichttexteinheit geprüft wird. Übersteigt sie in der Größe den vorgegebenen Schwellwert nicht, dann kann die Klassifizierung der Nichttexteinheit nicht bestimmt werden. Der Schwellwert wird in Abhängigkeit von der maximalen Schriftgröße und der Art des Dokuments (horizontal, vertikal oder gemischt) eingestellt; die halbe Schriftgröße erzielt zufriedenstellende Ergebnisse. Es wird zu Schritt S638 verzweigt, bei dem ein Merkmal "unbekannt" an die Nichttexteinheit hinzugefügt wird. Danach wird das Programm bei Schritt S607 (**Fig. 6A**) fortgesetzt, bei dem die nächste Einheit ausgewählt wird.

[0106] Übersteigt bei Schritt S637 die Größe den Schwellwert, dann werden die Schritte S617, S618 und S619 bearbeitet, bei denen die interne weißen Konturen der Nichttexteinheit verfolgt werden, die Dichte der Nichttexteinheit wird berechnet und die Anzahl der weißen Konturen werden, wie vorgegeben, geprüft.

[0107] Ist bei Schritt S619 die Anzahl der weißen Konturen nicht größer oder gleich 4, dann wird bei den Schritten S642A und S642 fortgefahren, wie oben ausgeführt. Bei Schritt S642, bei dem die Größe der Einheit berechnet wird, um zu bestimmen, ob sie groß genug ist, um entweder eine Strichzeichnung oder ein Halbtonbild darzustellen. Diese Größenbestimmung beruht auf der Grundlage der Höhe und der Breite der Nichttexteinheit, der Art des Dokuments, ebenso wie der maximalen Lauflänge der Bildpunkte. Wenn beispielsweise die Höhe und die Breite der Nichttexteinheit nicht größer als die maximale Schriftgröße ist, dann ist die Nichttexteinheit nicht groß genug, um entweder ein Halbtonbild oder eine Strichzeichnung zu sein, und es wird bei Schritt S643 fortgefahren, bei dem das Merkmal "unbekannt" hinzugefügt wird.

[0108] Wenn jedoch in ähnlicher Weise die Breite der Nichttexteinheit größer als die maximale Schriftgröße ist, jedoch die maximale Lauflänge der schwar-

zen Bildpunkte nicht größer als die maximale Schriftgröße ist, wird zu Schritt S643 verzweigt, bei dem das Merkmal "unbekannt" hinzugefügt wird. Anschließend wird bei Schritt S607 (**Fig. 6A**) eine neue Einheit ausgewählt.

[0109] Ist bei Schritt S642 die Nichttexteinheit groß genug, um entweder ein Strichzeichnung oder ein Halbtonbild zu sein, dann wird bei Schritt S643A fortgefahren, bei dem bestimmt wird, ob die Nichttexteinheit eine horizontale oder vertikale schiefe Linie ist. Bei Schritt S644 wird bestimmt, ob die Nichttexteinheit eine Strichzeichnung oder ein Halbtonbild ist. Ob die Einheit als ein Halbtonbild oder als eine Strichzeichnung klassifiziert wird, wird auf der Grundlage der mittleren horizontalen Lauflänge der schwarzen Bildpunkte in der Einheit, der mittleren horizontalen Lauflänge der weißen Bildpunkte in der Einheit und dem Verhältnis zwischen weißen und schwarzen Bildpunkten und der Dichte bestimmt. Im allgemeinen werden sehr dunkle Bilder als Halbtonbilder und helle Bilder als Strichzeichnungen angesehen.

[0110] Wenn speziell die mittlere Lauflänge der weißen Bildpunkte ungefähr gleich Null (das heißt, ein vorwiegend dunkles oder gesprenkeltes Bild) ist, dann, wenn die bei Schritt S618 berechnete Dichte anzeigt, daß die Einheit eher schwarz als weiß ist (das heißt, der Wert der Dichte ist größer als der erste Schwellwert von etwa 1/2), und die meisten Zeilen enthalten lange schwarze Lauflängen, wird die Bildeinheit als Halbtoneinheit festgelegt. Ist der Wert der Dichte nicht größer als der erste Schwellwert, wird die Einheit als Schnittzeichnung festgelegt.

[0111] Ist die mittlere Lauflänge der weißen Bildpunkte nicht ungefähr Null, und ist die mittlere Lauflänge der weißen Bildpunkte größer als die mittlere Lauflänge der schwarzen Bildpunkte, und enthalten die meisten Zeilen keine lange schwarzen Lauflängen, dann wird die Rahmeneinheit als eine Strichzeichnung festgelegt. Ist jedoch die mittlere Lauflänge der weißen Bildpunkte nicht größer als die mittlere Lauflänge der schwarzen Bildpunkte (das heißt, wieder ein vorwiegend schwarzes Bild), dann ist weiteres Prüfen erforderlich.

[0112] Wenn speziell die Anzahl der schwarzen Bildpunkte sehr viel kleiner als die Anzahl der weißen Bildpunkte (das heißt, die Anzahl der schwarzen Bildpunkte dividiert durch die Anzahl der weißen Bildpunkte größer als der zweite Schwellwert mit etwa dem Wert 2) ist, dann wird die Rahmeneinheit als Halbtoneinheit festgelegt. Wenn andererseits die Anzahl der schwarzen Bildpunkte dividiert durch die Anzahl der weißen Bildpunkte nicht größer als der zweite Schwellwert ist, jedoch die bei Schritt S618 bestimmte Dichte größer als der erste Schwellwert ist, dann wird die Einheit als Halbtonbild festgelegt. Andernfalls wird die Einheit als Strichzeichnung festgelegt.

[0113] Ist entsprechend Schritt S644 die Einheit als eine Strichzeichnung festgelegt, wird bei Schritt S645 fortgefahren, bei dem ein "Strichzeichnungs"-Merk-

mal angehängt wird, um dann bei Schritt S646 sämtliche Nachkommen von den derzeitigen Eltern zu verschieben, um unmittelbare Nachkommen der Großeltern zu werden. Speziell, wenn einmal ein Einheit als Strichzeichnung festgelegt wurde, dann werden keine Blöcke von der Strichzeichnungseinheit zur möglichen Zeichenerkennung ausgewählt. Danach wird im Programm bei den Schritten S625, S626 und S627 und danach bei Schritt S607 fortgefahren, um die nächste Einheit auszuwählen.

[0114] Wenn andererseits bei Schritt S644 die Einheit nicht als Strichzeichnung festgelegt wird, dann wird bei Schritt S647 fortgefahren, bei dem ein "Halbton"-Merkmal angehängt wird, und bei Schritt S648, um sämtliche Nachkommen von den derzeitigen Eltern zu verschieben, um unmittelbare Nachkommen der Großeltern zu werden. Das Programm kehrt dann zu Schritt S607 zurück, bei dem die nächste Einheit ausgewählt wird.

[0115] Zusätzlich zu den oben ausgeführten Nichttexteinheiten, die oben ermittelt wurden, identifiziert die vorliegenden Erfindung auch schiefe vertikale und horizontale Linien auf der Grundlage der Dicke und der Fülle. Auf einen genauen Programmablaufplan für diese Bestimmung wird aus Gründen der Klarheit verzichtet, es kann jedoch Bezug auf den Anhang der Computerauflistung genommen werden.

[0116] Zwei weitere Funktionen können in Schritt S53 hinzugefügt werden. Speziell kann eine Zeilenteilungszusatzfunktion ausgeführt werden, bei der Textzeichen von den unterlegten Zeilen abgespalten werden. Eine zweite Funktion kann ebenfalls ausgeführt werden, bei der das Vorhandensein einer punktierten Linie, einer gestrichelten Linie oder einiger Schmucklinien (entweder gerade oder schief) erfaßt werden kann.

[0117] Die Funktion des Teilens eines Zeilenzusatzes kann enthalten sein, beidem der Text (oder ein anderer Zusatz in Textgröße) einer Nichttextkomponente (wie einer vertikalen Zeile, einer horizontalen Zeile, einem Rahmen) zugeordnet zu sein scheint. Beispielsweise, wo das Textzeichen von einer unterlegten horizontalen Zeile umgeben zu sein scheint, ist es erforderlich das Textzeichen von der Zeile abzuspalten, damit die Textzeichen als textverbundene Komponenten charakterisiert werden kann, während die Zeilenbildpunkte als nichttextverbundene Komponenten erfaßt werden können.

[0118] Das Verfahren zum Abtrennen des Zusatzes zwischen dem Text und der Linie kann ausgeführt werden, indem als erstes das Vorhandensein eines solchen Zusatzes erfaßt wird. Im Fall einer Linie, wenn ein Seitenrand der Linie als gleichmäßig bestimmt wird, jedoch die andere Seite der Linie nicht als gleichmäßig bestimmt wird, dann besteht ein Zusatz auf der nichtgleichmäßigen Seite. Wenn im Fall eines Rahmens einige der äußeren Ränder nicht gleichmäßig sind, kann ein Zusatz zu diesen Rahmenrändern ebenfalls vorhanden sein.

[0119] Als nächstes werden die beiden Endpunkte

der zugefügten Linie (oder des Rahmenrands) berechnet, wobei die Lage der beiden Endpunkte auf der Bildaußenlinie der nichtgleichmäßigen Seite beruht. Ein "Schnitt", der durch die beiden berechneten Endpunkte verläuft, wird dann vorgenommen, und der Zusatz wird von der Linie entlang des "Schnitts" getrennt. Der Zusatz ist dann von der Linie entlang des "Schritts" getrennt, und der Zusatz sowie die Linie können dann getrennt verarbeitet werden, um zu bestimmen, ob die zusatzverbundenen Komponenten textverbundene Komponenten oder nichttextverbundene Komponenten sind, obgleich in der bevorzugten Form, der Teilungszusatz als eine textverbundene Komponente klassifiziert wird.

[0120] Die Funktion zur Erfassung des Vorhandenseins punktierter Linien, gestrichelter Linien oder Schmucklinien kann ausgeführt werden, indem jeder erfaßten Linie ein Merkmal entweder einer vertikalen Linie, einer horizontalen Linie oder einer geneigten Linie zugeteilt wird. Um solche nichtkontinuierlichen Linien zu erfassen, werden textverbundene Komponenten, die geringe horizontale oder vertikale Größe (ähnlich der Größe eines Punktes) haben, erfaßt. Als nächstes werden die erfaßten Komponenten in unterschiedliche Gruppen auf der Grundlage ihrer relativen Abstände unterteilt. Jede Gruppe wird dann geprüft, um zu bestimmen, ob jede Größe eine horizontale, keine vertikale oder keine schräge Linie beschreibt. Trifft die nicht zu, wird geprüft, ob die Gruppe eine punktierte Linie beschreibt oder nicht beschreibt. Wenn bestimmt ist, daß die Größe der betrachteten Gruppe eine Linie beschreibt, werden die Gruppen auf der Grundlage der Größe und der Sauberkeit der Zwischenräume zwischen den Gruppen reorganisiert. Danach wird die Gruppe erneut geprüft, um zu sehen, daß es sich um eine punktierte Linie handelt.

[0121] Wenn die Breite der Gruppe ähnlich einer vertikalen Linie ist, und wenn sich eine vertikale Linie in der Nachbarschaft der betrachteten Linie befindet, und der Zwischenraum zwischen dieser Gruppe und der Linie sauber ist, dann wird entschieden, daß es sich bei der betrachteten Gruppe um eine vertikale Linie handelt. Ist dies nicht der Fall, muß bestimmt werden, ob die Länge der Gruppe ähnlich einer horizontalen Linie ist, und ob in der Nachbarschaft eine horizontale Linie ist, wobei der Zwischenraum zwischen der benachbarten Linie und der betrachteten Linie sauber ist. In diesem Fall wird entschieden, daß es sich bei der erfaßten nichtkontinuierlichen Linie um eine horizontale Linie handelt.

[0122] Nachdem sämtliche verbundenen Komponenten in dem Bild der Bildpunkte erfaßt und in der oben beschriebenen Weise bezüglich der **Fig. 6** (Schritt S53 in **Fig. 5**) klassifiziert wurde, wird eine modifizierte Baumstruktur, wie beispielsweise in **Fig. 10** gezeigt, erhalten. Wie dort gezeigt, entspricht die Wurzel des Baums der Seite der Bildpunktdaten. Von der Wurzel verzweigen Textblöcke, Nichttextblöcke, deren Inhalte unbekannt sind, Rahmen, Bilder

und Linien. Von den Bildern verzweigen Textblöcke, unbekannte Nichttextdaten, Tabellen, die abwechselnd Textblöcke von Bildern und Linien enthalten.

[0123] Nachdem, um auf **Fig. 5** zurückzukommen, sämtliche der nichttextverbundenen Komponenten bei Schritt S53 klassifiziert wurden, sucht Schritt S54 nach unsichtbaren (weiße Zeilen) entlang der Ränder der nichttextverbundenen Komponenten. Derartige unsichtbare Zeilen (zusammen mit den in Schritt S52 bestimmten unsichtbaren Zeilen) werden bei Schritt S57 dazu verwendet, Textblöcke zu spalten, wenn eine sichtbare oder unsichtbare Zeile durch diesen Textblock verläuft. Dieser Vorgang wird später genauer beschrieben.

[0124] Bei Schritt S55 werden horizontale und vertikale Titelzeilen unter Verwendung unbekannter nichttextverbundener Komponenten, die bei Schritt S53 klassifiziert wurden, gebildet. Diese Unbekannten werden verwendet, um die Richtung des Titels, der üblicherweise vertikal, horizontal und schief sein kann, auf der Grundlage des Abstands zwischen den Komponenten, die als gruppiert betrachtet werden, und der möglichen Titelzeilenlänge, wenn der Titel horizontal oder vertikal ist, zu erfassen. Mit oben beschriebenen Verfahren wird jede titelverbundene Komponente mit einem unmittelbar benachbarten Block in den horizontalen und den vertikalen Richtungen verglichen, und welcher Block auch immer enger an den horizontalen und vertikalen Titel angefügt ist.

[0125] Bei Schritt S56 werden die textverbundenen Komponenten dazu verwendet, die Textblöcke zu bilden. In Kürze: Für jede textverbundene Komponente werden die in der Nähe befindlichen horizontalen und vertikalen Nachbarn gesucht und in einem Ansammlungsverfahren in Textblöcken gruppiert. Dieses Verfahren wird nun im einzelnen unter Bezugnahme auf **Fig. 11** beschrieben.

[0126] In **Fig. 11** werden die bei Schritt S52 identifizierten textverbundenen Komponenten bei Schritt S561 verwendet, um die Werte der Variablen H_GAP und V_GAP zu berechnen. Diese Größen sind der mittlere horizontale und vertikale Zwischenraum zwischen den benachbarten, textverbundenen Komponenten. Bei Schritt S562 werden dann sämtliche textverbundenen Komponenten nach ihrer vertikalen Lage sortiert. Bei Schritt S563 wird die Bildung des derzeitigen Textblocks begonnen. Bei Schritt S564 wird eine verbundene Komponente aus der Sortierreihenfolge herausgeholt, und die Werte der Variablen V_DIST und H_DIST werden berechnet. Diese Größen sind die vertikalen und horizontalen Abstände zwischen den geholten verbundenen Komponenten und dem derzeitigen Textblock. Steht keine verbundene Komponente zur Verfügung, wird zu Schritt S563 zurückgekehrt, und der nächste derzeitige Textblock wird gebildet.

[0127] Bei Schritt S565 wird bestimmt ob für die geholte verbundene Komponente V_DIST größer als V_GAP oder H_DIST größer als H_GAP ist. Lautet bei Schritt S565 die Antwort JA, zeigt dies an, daß die

geholt verbundene Komponente nicht im derzeitigen Textblock sein sollte, es wird bei Schritt S564 fortgefahren, und die nächste verbundene Komponente in der Sortierreihenfolge wird geholt. Ist die Antwort bei Schritt S565 NEIN, zeigt dies an, daß die geholte verbundene Komponente für den derzeitigen Textblock geeignet sein kann; es wird bei Schritt S566 fortgefahren.

[0128] Bei Schritt S566 wird bestimmt, ob die geholte verbundene Komponente und jede andere verbundene Komponente schon im derzeitigen Textblock enthalten ist, und eine vorbestimmte Bedingung erfüllt. Beim bevorzugten Ausführungsbeispiel wird die vorbestimmte Bedingung in **Fig. 12** vorgegeben.

[0129] Bei **Fig. 12** wird bei Schritt S5661 die geholte verbundene Komponente mit einer verbundene Komponente, die schon im dem derzeitigen Textblock enthalten ist, verglichen. Bei Schritt S5662 werden die Werte der Variablen h_dist und v_dist berechnet. Diese Werte stellen den horizontalen und vertikalen Abstand zwischen der geholten verbundenen Komponente und der bei Schritt S5661 erkannten verbundene Komponente dar.

[0130] Bei Schritt S5663 wird bestimmt, ob v_dist kleiner V_GAP oder ob h_dist kleiner H_GAP ist. Lautet die Antwort bei Schritt S5663 NEIN, zeigt dies an, daß eine verbundene Komponente, die bei Schritt S5661 erkannt wurde, sich nicht nahe genug an der geholten verbundenen Komponente liegt. Es wird bei Schritt S5661 fortgefahren, bei dem die nächste verbundene Komponente im derzeitigen Textblock mit der geholten verbundenen Komponente verglichen wird.

[0131] Lautet die Antwort bei Schritt S5663 JA, wird bei Schritt **S5664** fortgefahren, bei dem bestimmt wird, welche Art von vertikaler Beziehung die geholte verbundene Komponente zu der einen verbundenen Komponente hat. Drei Beziehungen sind möglich: (1) Die geholte verbundene Komponente liegt vertikal über der einen verbundenen Komponente; und (2) die geholte verbundene Komponente liegt nicht oberhalb der einen verbundenen Komponente. Trifft Bedingung (1) zu wird bei Schritt S5665 fortgefahren, wobei im anderen Fall bei Schritt S5667 fortgefahren wird.

[0132] Bei Schritt S5665 wird bestimmt, ob eine andere verbundene Komponente bei dem derzeitigen Textblock, der in horizontaler Richtung nahe an der geholten verbundenen Komponente ist, befindet. Die wird ermittelt, indem bestimmt wird, ob, h_dist kleiner als W_GAP ist (wobei W_GAP ein klein wenig breiter als H_GAP ist). Lautet die Antwort NEIN, wird bei Schritt S5661 mit der nächsten verbundenen Komponente beim derzeitigen Textblock fortgefahren. Lautet die Antwort hingegen JA, wird bei Schritt S5667 fortgefahren, bei dem bestimmt wird, ob ein Strichzeichnungsobjekt (beispielsweise eine Linie oder ein Strichzeichnungsbild) zwischen die geholte verbundene Komponente und die andere bei Schritt S5665 erkannte verbundene Komponente fällt. Existiert

solch eine Strichzeichnung, wird bei Schritt S564 fortgefahren, bei dem die nächste verbundene Komponente in der Bestandteileihenfolge geholt wird. Lauter hingegen bei Schritt S5667 NEIN, wird zu Schritt S567 in **Fig. 11** verzweigt.

[0133] Bei Rückkehr zu Schritt S567 in **Fig. 11** wird bestimmt, ob die gerade verarbeitete geholte Komponente die letzte textverbundenen Komponente war, die bei Schritt S52 erkannt wurde. Ist die geholte verbundene Komponente nicht die letzte Komponente, wird die geholte verbundene Komponente in den derzeitigen Textblock bei Schritt S568 eingebracht, und es wird bei Schritt S564 fortgefahren, bei dem die nächste verbundene Komponente in der sortierten Reihenfolge geholt wird. Lautet andererseits die Antwort bei Schritt S567 JA, wurde die letzte geholte verbundene Komponente in den derzeitigen Textblock eingebracht, und es wird bei Schritt S57 in **Fig. 5** fortgefahren.

[0134] Bei Rückkehr zu Schritt S57 in **Fig. 5**, wird bestimmt, ob ein bei Schritt S56 gebildeter Textblock ein einzelner Textblock sein sollte, oder ob er aufgrund einer sichtbaren oder unsichtbaren Zeile, die durch den gebildeten Textblock hindurchtritt, in zwei oder mehr Textblöcke unterteilt werden sollte. Beispielsweise wird in **Fig. 13A** ein Textblock **131** abgebildet, der eine Vielzahl von darin angeordneten verbundene Komponenten hat. Benachbart zu Textblock **131** sind zwei Nichttextblöcke **133** und **134**. Zwischen den Nichttextblöcke befindet sich eine unsichtbare (weiße) Zeile **135**, die bei Schritt S54 erkannt wurde. Wenn die beiden Nichttextblöcke **133** und **134** beispielsweise jeweils ein Bild enthalten mit einer unsichtbaren Zeile **135**, die sich dazwischen ausdehnt, können die textverbundenen Komponenten des benachbarten Nichttextblocks **133** nur zu diesem Nichttextblock gehören, und die textverbundenen Komponenten des benachbarten Nichttextblocks **134** können nur zu diesem Nichttextblock gehören. Ist der bei Schritt S57 gebildete Textblock **131** ein einzelner Textblock, kann künftige Zeichenerkennungsverarbeitung von beiden der Nichttextblöcke **133** und **134** falsch dargestellt werden.

[0135] Bei der Situation, wie in **Fig. 13B** gezeigt, bei der die unsichtbare Zeile **135** (oder eine sichtbare Zeile) durch den zuvor gebildeten Textblock **131** tritt, wird der Textblock in die beiden Textblöcke **136** und **137** gespalten. Nun werden die Nichttextblöcke **133** und **134** ihre entsprechenden textverbundenen Komponenten haben, die entsprechend verarbeitet werden. In Anschluß an Schritt S57 wird bei Schritt S58 fortgefahren, bei dem die Richtung jedes Textblocks erfaßt wird. Die Erfassung der Textblockrichtung hängt von der Länge und der Breite der textverbundenen Komponenten im Textblock, den vertikalen und horizontalen Zwischenräumen und ihrer relativen Lage zueinander ab. **Fig. 14** mit den **Fig. 14A** und **14B** ist ein genauer Programmablaufplan, der diese Verarbeitung beschreibt.

[0136] Bei **Fig. 14A** werden sämtliche Textkompo-

nenten innerhalb eines bei Schritt S56 gebildeten Textblöcke zuerst Schritt S5801 unterworfen, bei dem eine Anzahl von Werten berechnet werden. Es sollte angemerkt werden, daß sich die Breite auf die horizontale Ausdehnung bezieht, während sich die Länge auf die vertikale Ausdehnung bezieht. Die Werte der Variablen MAX_LEN und MAX_WID werden berechnet und entsprechend stellen sie die maximale Länge und die maximale Breite unter den textverbundenen Komponenten innerhalb des Textblocks dar. Die Werte der Variablen AVG_LEN und AVG_WID werden ebenfalls berechnet, und sie stellen entsprechend die mittlere Länge und die mittlere Breite sämtlicher verbundener Komponenten innerhalb des Textblocks dar. Schließlich werden die Werte der Variablen STD_LEN und STD_WID berechnet, und sie stellen entsprechend die Standardabweichung der Länge und die Standardabweichung der Breite sämtlicher verbundener Komponenten innerhalb des Textblocks dar.

[0137] Danach werden eine Reihe von Verarbeitungsschritten unternommen, um zu bestimmen, ob der Textblock ein horizontaler oder ein vertikaler Textblock ist. Folglich wird bei Schritt **S5802** bestimmt, ob STD_LEN kleiner oder gleich STD_WID ist, MAX_WID größer als MAX_LEN ist, und MAX_WID größer als MAX ist (hier festgelegt, 12 pt Druckgröße oder $1,5 \times \text{MAX_LEN}$ zu sein; der Wert 12 pt beruht auf einer üblichen Textgröße. Natürlich werden größere oder kleinere Textgrößen unterschiedliche Werte erforderlich machen, was bevorzugt dynamisch eingestellt werden kann.). Sind die Beziehungen wahr, wird entschieden, daß der Textblock ein horizontaler Textblock ist. Sind diese Beziehungen falsch, wird bei Schritt S5803 fortgefahren.

[0138] Schritt S5803 bestimmt, ob STD_WID kleiner oder gleich STD_LEN ist, MAX_LEN größer als MAX_WID ist und MAX_LEN größer als MAX ist (12 pt oder $1,5 \times \text{MAX_WID}$). Sind diese Beziehungen wahr, wird entschieden, daß der Textblock ein vertikaler Textblock ist. Sind diese Beziehungen falsch, wird bei Schritt S5804 fortgefahren.

[0139] Bei Schritt S5804 werden die Werte der Variablen H_SHORT und V_SHORT für jede verbundene Komponente bestimmt. Diese Werte stellen den kürzesten horizontalen und den kürzesten vertikalen Zwischenraum von jeder verbundene Komponente und ihren nachbarverbundenen Komponenten dar.

[0140] Es wird bei Schritt S5805 (**Fig. 14B**) fortgefahren, bei dem bestimmt wird, ob H_SHORT kleiner als V_SHORT für die meisten der verbundenen Komponenten in einem Textblock ist. Ist diese Beziehung wahr, wird entschieden, daß der Textblock ein horizontaler Textblock ist. Ist diese Beziehung falsch, wird bei Schritt S5806 fortgefahren, bei dem bestimmt wird, ob V_SHORT kleiner als H_SHORT für die meisten der verbundenen Komponenten in dem Textblock ist. Ist diese Beziehung wahr, wird entschieden, daß der Textblock ein vertikaler Textblock ist. Ist diese Beziehung falsch, wird bei Schritt S5807

fortgefahren Bei Schritt S5807 wird die Anzahl der horizontalen und der vertikalen Zwischenräume beim Textblock berechnet, Danach wird bei Schritt S508 bestimmt, ob die Anzahl der horizontalen Zwischenräume größer als die Anzahl der vertikalen Zwischenräume ist, und ob die Breite des Textblocks größer als $12 \text{ pt} \times 2$ ist.

[0141] Ist diese Beziehung wahr, wird entschieden, daß der Textblock ein horizontaler Textblock ist. Ist diese Beziehung falsch, wird bei Schritt S5809 fortgefahren, bei dem bestimmt wird, ob die Anzahl der vertikalen Zwischenräume größer ist als die Anzahl der horizontalen Zwischenräume, und ob die Länge des Textblocks größer als $12 \text{ pt} \times 2$ ist. Ist diese Beziehung wahr, wird entschieden, daß der Textblock ein vertikaler Textblock ist. Ist die Beziehung falsch, wird bei Schritt S5810 fortgefahren.

[0142] Schritt S5810 bestimmt, ob die Länge der Textblocks kleiner als $12 \text{ pt} \times 2$ ist und ob sich sämtliche verbundene Komponente auf der gleichen horizontalen Zeile befinden. Ist die Bestimmung wahr, dann ist der Textblock ein horizontaler Textblock. Ist die Bestimmung falsch, wird bei Schritt S5811 fortgefahren. Bei Schritt S5811 wird bestimmt, ob die Breite des Textblock kleiner als $12 \text{ pt} \times 2$, und ob sich sämtliche verbundene Komponente auf der gleichen vertikalen Linie befinden. Ist die Beziehung wahr, wird entschieden, daß der Textblock einer vertikalen Textblock ist. Ist die Beziehung falsch, wird bei Schritt S5812 fortgefahren.

[0143] Bei Schritt S5812 werden sämtliche verbundene Komponente im Textblock in vertikaler Richtung sortiert, und es wird erfaßt, ob die erste verbundene Komponente und die letzte verbundene Komponente sich vertikal überlappende Bereiche haben. Überlappen sie sich, wird entschieden, daß der Textblock ein horizontaler Textblock ist. Überlappen sie sich nicht, wird bei Schritt S5813 fortgefahren.

[0144] Bei Schritt S5813 werden sämtliche verbundene Komponente im Textblock in horizontaler Richtung sortiert, und es wird erfaßt, ob die erste verbundene Komponente und die letzte verbundene Komponente sich horizontal überlappende Bereiche haben. Überlappen sie sich, wird entschieden, daß der Textblock ein vertikaler Textblock ist. Überlappen sie sich nicht, wird entschieden, daß die Richtung dieses Textblocks unbekannt ist.

[0145] Ist einmal die Richtung sämtlicher erkannten Textblöcke bestimmt, wird bei Schritt S59 fortgefahren, bei dem die Textzeilen bei jedem Textblock aufgrund der bei Schritt S58 erkannten Richtung gebildet werden. **Fig. 15** zeigt einen genaues Flußdiagramm dieser Verarbeitung.

[0146] **Fig. 15** zeigt die Verarbeitung von textverbundenen Komponenten in einem horizontalen Textblock. Die Verarbeitung verbundener Komponenten innerhalb eines vertikalen Textblocks ist ähnlich, und sie wird deshalb nicht genauer beschrieben. Bei Schritt S591 werden die Textkomponenten innerhalb des horizontalen Textblock entsprechend ihrer hori-

zontalen Lage sortiert. Danach wird bei Schritt S592 die nächste verbundene Komponente aus der Sortierliste geholt, und es wird erfaßt, ob die geholte verbundene Komponente in jede schon bestehende Textzeile paßt. Die genaue Verarbeitung für diese Bestimmung wird nun unter Bezugnahme auf **Fig. 16**, die aus den **Fig. 16A** bis **16E** besteht, beschrieben.

[0147] In **Fig. 16A** wird die bei Schritt S592 geholte, verbundene Komponente dazu verwendet, die C_LINE1 und den C_BLOCK1 bei Schritt S5921 zu berechnen. C_LINE1 ist die in Prozent ausgedrückte vertikale Bereichsüberlappung zwischen der geholten verbundene Komponente und der derzeitigen Textzeile, während C_BLOCK1 die in Prozent ausgedrückte vertikale Bereichsüberlappung zwischen der geholten verbundene Komponente und dem hinteren Ende der verbundenen Komponente der derzeitigen Textzeile ist. Die Wiedergabe innerhalb des Schritts S5921 in **Fig. 16A** zeigt die Beziehung zwischen der derzeitigen Textzeile und dem darin befindlichen hinteren Ende der verbundenen Komponente, der geholten verbundenen Komponente und ihrer vertikalen Bereichsüberlappung.

[0148] Bei Schritt S5922 wird bestimmt, ob C_LINE1 größer als Null ist oder nicht. Wenn nicht, wird bei Schritt S5936 fortgefahren, der genauer unter Bezugnahme auf **Fig. 16E** beschrieben wird. Wenn andererseits bei Schritt S5922 bestimmt wird, daß C_LINE1 größer Null ist, werden die drei Werte h_dist1, v_dist1 und NEXT_CLOSEST_LINE (nächst naheliegende Zeile) bestimmt. Der Wert der Variablen h_dist1 ist der horizontale Abstand zwischen der geholten verbundenen Komponente und dem hinteren Ende der verbundenen Komponente der derzeitigen Textzeile. Der Wert der Variablen NEXT_CLOSEST_LINE ist eine Textzeile, die (i) unterhalb der derzeitigen Zeile liegt, und die (ii) eine verbundene Komponente mit einem hinteren Ende hat, das sich sehr nahe an der geholten verbundenen Komponente befindet, und der Abstand zwischen der geholten verbundenen Komponente und der verbundenen Komponente mit einem hinteren Ende innerhalb eines vorgegebenen Abstands liegt. Fortgefahren wird bei Schritt S5924 in **Fig. 16B**.

[0149] Bei Schritt S5924 werden die Werte der Variablen C_LINE2, C_BLOCK2, h_dist2 und v_dist2 bestimmt. Der Wert von C_LINE2 ist die in Prozent ausgedrückte vertikale Bereichsüberlappung zwischen der geholten verbundenen Komponente und der NEXT_CLOSEST_LINE. Der Wert von C_BLOCK2 ist die in Prozent ausgedrückte vertikalen Bereichsüberlappung zwischen der geholten verbundenen Komponente und dem hinteren Ende der verbundenen Komponente der NEXT_CLOSEST_LINE. Der Wert von h_dist2 ist der horizontale Abstand zwischen der geholten verbundenen Komponente und dem hinteren Ende der verbundenen Komponente von NEXT_CLOSEST_LINE. Der Wert von v_dist2 ist der vertikale Abstand zwischen der geholten verbundenen Komponente und dem hinteren Ende der

verbundenen Komponente der
NEXT_CLOSEST_LINE.

[0150] Danach wird bei Schritt S5925 fortgefahren, bei dem bestimmt wird, ob C_LINE1 kleiner als 50% und C_LINE2 größer oder gleich 50% ist, und die derzeitige Zeile und NEXT_CLOSEST_LINE sich aufwärts neigen. Treffe die Beziehungen der Schritt S5926 zusammen, wird bei Schritt S5926 fortgefahren, bei dem die nächste horizontale Textzeile ausgewählt wird und zu Schritt S592 in **Fig. 15** zurückgekehrt wird. Wenn andererseits die Beziehungen bei Schritt S5925 nicht zusammenfallen, wird bei Schritt S5927 in **Fig. 16C** fortgefahren.

[0151] In **Fig. 16C** wird bei Schritt S5927 entschieden, ob acht Beziehungen vorkommen. Diese Beziehungen sind folgende:

- (1) $C_LINE2 = 0\%$
- (2) $C_LINE1 \geq 50\%$ UND $C_LINE2 < 50\%$ UND $C_BLOCK1 > 0\%$ UND $C_BLOCK1 > C_BLOCK2$
- (3) $h_dist1 > 0$ UND $h_dist2 > 0$ UND ($h_dist1 < 12\text{pt}$ ODER $h_dist1 < h_dist2$) UND $C_BLOCK1 > 0\%$ UND $C_BLOCK1 > C_BLOCK2$
- (4) $h_dist1 > 0$ UND $h_dist2 > 0$ UND ($h_dist1 < 12\text{pt}$ ODER $h_dist1 < h_dist2$) UND $C_BLOCK1 > 0\%$ UND DERZEITIGE ZEILENNEIGUNG UND NEXT_CLOSEST_LINE NEIGEN SICH NACH UNTEN
- (5) $h_dist1 > 0$ UND $h_dist2 > 0$ UND ($h_dist1 < 12\text{pt}$ ODER $h_dist1 < h_dist2$) UND $C_BLOCK1 = 0\%$ UND $C_BLOCK2 = 0\%$ UND $v_dist1 < v_dist2$
- (6) $h_dist1 \geq 0$ UND $h_dist1 < 12\text{pt}$ UND $C_BLOCK1 > 0\%$ UND $h_dist2 < 0$
- (7) $h_dist1 < 0$ UND $h_dist2 < 0$ UND $v_dist1 < v_dist2$
- (8) $C_BLOCK1 > C_BLOCK2$

[0152] Es wird anschließend bei Schritt S5928 fortgefahren, um zu entscheiden, welche der Beziehungen (1) bis (8) erfüllt sind. Ist keine dieser Beziehungen erfüllt, wird bei Schritt S5929 fortgefahren, bei dem die nächste horizontale Textzeile gewählt wird, und es wird zu Schritt S592 in **Fig. 15** zurückgekehrt. Wenn andererseits irgendeine der Beziehungen (1) bis (8) bei Schritt S5928 erfüllt ist, wird bei Schritt S5930 fortgefahren, bei dem entschieden wird, ob die horizontale Lage der geholten verbundenen Komponente das hinteren Ende der verbundenen Komponente der derzeitigen Textzeile überlappt. Tritt Überlappung auf, wird bei Schritt S5931, bei Nichtüberlappung bei Schritt S5932 fortgefahren.

[0153] In **Fig. 16D** entscheidet Schritt S5931, ob die kombinierte Länge der geholten verbundenen Komponente und das hintere Ende der verbundenen Komponente der derzeitigen Textzeile ähnlich einer Zeichenlänge ist. Ist die kombinierte Länge nicht ähnlich einer Zeichenlänge, wird bei Schritt S3935 fortgefahren, bei dem die geholte verbundene Komponente in die derzeitigen Textzeile paßt. Wenn andererseits die kombinierte Länge einer Zeichenlänge nicht ähn-

lich ist, wird bei Schritt S5931' fortgefahren, bei dem bestimmt wird, ob C_BLOCK1 größer als 50% ist oder v_dist1 kleiner als ein Punktgröße ist. Ist eine dieser Beziehungen wahr, wird bei Schritt S5935 fortgefahren, bei dem die geholte verbundene Komponente in die derzeitige horizontale Textzeile paßt. Wird jedoch eine der Beziehungen von Schritt S5931 nicht erfüllt, wird bei Schritt S5933 fortgefahren, bei dem die nächste horizontale Textzeile ausgewählt wird, und es wird anschließend zu Schritt S592 in **Fig. 15** zurückgekehrt.

[0154] Bei Schritt S5932 wird entschieden, ob h_dist1 kleiner als $12\text{pt} \times 2$ ist und (C_LINE1 größer als 0% ist oder C_BLOCK1 größer als 0% ist) oder NEXT_CLOSEST_LINE existiert nicht oder v_dist1 ist kleiner als v_dist2. Sind diese Bedingungen erfüllt, wird erneut bestimmt, ob die geholte verbundene Komponente in die derzeitigen Textzeile paßt; es wird über Schritt S3935 bei Schritt S592 in **Fig. 15** fortgefahren. Werden die Beziehungen bei Schritt S5932 nicht erfüllt, wird eine nächste Textzeile bei Schritt S5934 ausgewählt; es wird erneut bei Schritt S592 in **Fig. 15** zurückgekehrt.

[0155] Wenn bei Rückkehr zu **Fig. 16A** in Schritt S5922 bestimmt wird, daß C_LINE1 nicht größer Null ist, wird bei Schritt S5936 in **Fig. 16E** fortgefahren. Bei Schritt S3936 wird der Wert der Variablen NEXT_CLOSEST_LINE, h_dist1, h_dist2, v_dist1 und dist2, wie oben bezüglich der Schritte S923 und S924 ermittelt.

[0156] Bei Schritt S3937 wird bestimmt, ob NEXT_CLOSEST_LINE existiert, und v_dist1 größer als v_dist2 ist. Lautet die Antwort JA, wird bei Schritt S5939 fortgefahren, bei dem die nächste horizontale Textzeile gewählt wird; anschließend wird zu Schritt S592 in **Fig. 15** zurückgekehrt. Wenn jedoch die Beziehungen von Schritt S5937 nicht existieren, wird zu Schritt S5938 verzweigt, bei dem entschieden, ob h_dist1 kleiner Null und ob v_dist1 kleiner oder gleich eine Punktgröße und die kombinierte Länge der geholten verbundenen Komponente und das hinteren Ende der verbundenen Komponente ähnlich der Länge eines Zeichens ist. Sind diese Beziehungen wahr, wird bestimmt, ob die geholte verbundene Komponente in die derzeitige horizontale Textzeile bei Schritt S5941 hineinpaßt, und es wird zu Schritt S592 in **Fig. 15** zurückgekehrt. Sind die Beziehungen von Schritt S5938 falsch, wird bei Schritt S5940 fortgefahren, wo bestimmt wird, ob h_dist1 größer oder gleich Null ist, ob die geholte verbundene Komponente oberhalb dem hinteren Ende der verbundenen Komponente liegt, und ob die vertikale Lage der geholten verbundenen Komponente nahe bei dem hinteren Ende der verbundenen Komponente ist. Ist diese Beziehung wahr, wird wiederum bei Schritt S5941 bestimmt, daß die geholte verbundene Komponente in die derzeitige horizontale Textzeile hineinpaßt, und es wird wieder zu Schritt S592 in **Fig. 15** zurückgekehrt. Sind diese Beziehungen falsch, wird bei Schritt S5942 fortgefahren, bei dem die nächste horizontale

Textzeile gewählt wird; anschließend wird zu Schritt S592 in **Fig. 15** zurückgekehrt.

[0157] Nachdem sämtliche Textzeilen für jeden Textblock bei Schritt S59 in **Fig. 5** gebildet wurden, wird bei Schritt S60 fortgefahren, bei dem Nachverarbeitung der Blöcke ausgeführt werden kann. Bei der Nachbearbeitung kann abgestimmt werden, ob eine spezielle anschließende Bearbeitung wie Zeichenerkennung oder Datenkompression ausgeführt werden soll. Zum Zweck dieser Anwendung wird die Nachbearbeitung der Blöcke für die nachfolgende Zeichenerkennung beschrieben. Zweck einer solchen Nachbearbeitung ist es, die Text- und die Nichttext-Blockdarstellung kompakter und deutlicher darzustellen. Bei der Nachbearbeitung können Textblöcke mit anderen Textblöcken oder Nichttextblöcken kombiniert werden, und Nichttextblöcke können je nach Anforderung mit anderen Text- oder Nichttextblöcken kombiniert werden, um die weitere Verarbeitung wie beispielsweise Zeichenerkennung zu vereinfachen. Als erstes wird die Schiefe der gesamten Seite berechnet, indem beispielsweise eine Approximation der kleinsten Quadrate verwendet wird, um den schiefen Winkel der gebildeten Textzeilen (wird bei Schritt S59 bestimmt) auf der Grundlage der verbundenen Komponenten, die in einer solche Zeile enthalten sind, näherungsweise zu bestimmen. Die Schiefe der Seite wird dann der Mittelwert der schiefen Winkel der Textzeilen auf der Seite sein.

[0158] Wird bestimmt, daß eine Seite nur leicht schief ist (beispielsweise, wo der Zwischenraum zwischen der rechteckigen oder der ungleichmäßigen Kurve, die jeden Textblock umgibt, deutlich wie in **Fig. 17** ist), kann eine Zahl zur weiteren Verarbeitungsschritten realisiert werden. Erstens kann der Textoder Titel-)Block mit einem Titelblock in Abhängigkeit davon, ob sie umgebende Rechtecke oder unregelmäßige Kurvenüberlappung sind oder seine Ausrichtung kombiniert werden. Zweitens kann bestimmt werden, ob einige Halbtonbilder eine quadratische Fläche voll ausfüllen, um zu beurteilen, ob irgendein kleiner Textblock tatsächlich ein Textblock oder nur Teil eines Nichttextblocks, wie beispielsweise ein Bild, ist. Drittens kann ein Spaltenanalyse durchgeführt werden, und aufgrund einer Spalteninformation können die Textblöcke kombiniert werden. Eine derartige Kombination hängt grundsätzlich davon ab, ob sich die Blöcke innerhalb der gleichen Spalte befinden, wenn sie eng genug nebeneinander stehen, wenn die Kombination andere Blöcke überlappt, und wenn sie die gleiche Ausrichtung haben.

[0159] Ist bei der Seite eine größere Schiefe ermittelt, werden die drei Verarbeitungsschritte übersprungen. Es wird dann eine Schritt ausgeführt, bei dem ein kleiner Textblock mit einem nahe befindlichen Nachbarn kombiniert wird. Diese Kombination könnte vom Merkmal des Nachbars, der Geschlossenheit des Nachbarn und den Eigenschaften des Textblocks selbst abhängen. Nach einer solchen Kombination kann das Merkmal des kombinierten Textblocks ver-

ändert werden, um mit dem kombinierten Nachbar-textblock übereinzustimmen.

[0160] Ebenso können Textblöcke mit Titelblöcken oder Blöcken, die nur eine unbekannte verbundene Komponente enthält, verbunden werden. Diese Kombination hängt von der relativen Lage, der Richtung der Blöcke oder der relativen Größe der Blöcke ab.

[0161] Schließlich kann die Nachbearbeitung der Blöcke die Kombination bestimmter Nichttextblöcke enthalten. Solch eine Kombination hängt von der relativen Lage der Nichttextblöcke und den Merkmalen dieser Blöcke ab. Beispielsweise können einige Bildblöcke kombiniert werden, wenn erfaßt wird, daß sie dem Gesamtbild zugehörig sind. Derartige Nachbearbeitungsschritte können durchgeführt werden, um die nachfolgende Kennzeichenerkennungsverarbeitung zu vereinfachen. Auf ein genaues Flußdiagramm der Nachbearbeitungsschritte wird aus Gründen der Klarheit verzichtet, es kann jedoch Bezug auf den anliegenden Anhang der Computerprogramm-auflistung genommen werden.

[0162] Es wurde ein Verfahren und ein Gerät zur wirksamen und leistungsfähigen Auswahl von Text- und Nichttextblöcken in einem abgespeicherten Dokument beschrieben, das sowohl vertikale als auch horizontale Textblöcke zu erkennen vermag, und bei dem ein schiefes Dokument vor der Bildung des Textblocks nicht gedreht werden muß.

[0163] Die einzelnen in der Kontur gezeigten oder in der Zeichnung durch Blöcke dargestellten Teile sind sämtlich bei der Kennzeichenerkennungstechnik bekannt, und ihr spezifischer Aufbau und ihre spezifische Arbeitsweise sind bezüglich der Arbeitsweise nicht kritisch, oder sie sind die beste Art die vorliegenden Erfindung auszuführen.

[0164] Während die vorliegenden Erfindung im Hinblick darauf beschrieben wurde, was derzeit als bevorzugtes Ausführungsbeispiel betrachtet wird, kann verstanden werden, daß die vorliegenden Erfindung nicht auf das offenbarte Ausführungsbeispiel beschränkt ist. Im Gegenteil ist bei der Erfindung beabsichtigt, verschiedene Modifikationen und äquivalente Anordnungen, die innerhalb des Schutzbereichs der anliegenden Patentansprüche enthalten sind, abzudecken.

Patentansprüche

1. Verfahren zur Auswahl von Pixelblöcken aus Pixelbilddaten, mit den Verfahrensschritten:
Umreißen von Konturen verbundener Komponenten in Pixelbilddaten (S56);
erstes Verbinden, wobei verbundene Komponenten der Breite nach auswählbar auf der Grundlage der Größe und der Nähe zu anderen verbundenen Komponenten verbunden werden, um Zeilen zu bilden (S593); und
zweites Verbinden, wobei Zeilen der Breite nach verbundener Komponenten vertikal auswählbar auf der Grundlage von Größe und Nähe zu anderen Zeilen

verbunden und Rechtecke um die vertikal verbundenen Komponenten gebildet werden; gekennzeichnet durch den Verfahrensschritt Bilden einer unregelmäßigen Kurve um Pixelbilddaten, die von einem Rechteck umgeben sind, das sich mit einem anderen Rechteck überlappt.

2. Verfahren nach Anspruch 1, mit dem weiteren Verfahrensschritt des Eingebens (S401) von Bilddaten, der den Verfahrensschritt des Binärumsetzens der Bilddaten in Pixelbilddaten an der Stelle umfaßt, bei der die Pixelbilddaten keine binären Pixelbilddaten sind.

3. Verfahren nach Anspruch 1, bei dem beim der Verfahrensschritt des Erzeugens die kleinste unregelmäßige Kurve bildet.

4. Verfahren nach Anspruch 1, mit dem weiteren Verfahrensschritt des Bildens eines hierarchischen Baumes von im Verfahrensschritt des Umreißen umrissener verbundener Komponenten, wobei der hierarchische Baum wenigstens einen Pixelblock von Bilddaten darstellt, die von der unregelmäßigen Kurve umgeben sind, die der Erzeugungsschritt erzeugt.

5. Verfahren nach Anspruch 1, mit dem weiteren Verfahrensschritt des Klassifizierens der im Bildungsschritt erzeugten unregelmäßigen Kurven in Text- und Non-Textblöcke (**S604**).

6. Verfahren nach Anspruch 5, mit dem weiteren Verfahrensschritt des Erkennens von Zeichenbildern in den Textblöcken (S52).

7. Verfahren nach Anspruch 6, mit dem weiteren Verfahrensschritt des Herleitens weißer Konturen für non-Textumrissene verbundene Komponenten (S54).

8. Verfahren nach Anspruch 7, das Non-Text-Blöcke als Tabelleneinheiten gemäß der Anzahl weißer Konturen bestimmt (**S624**).

9. Verfahren nach Anspruch 8, mit dem weiteren Verfahrensschritt des Errechnens der Weißkonturfüllrate (S620).

10. Verfahren nach Anspruch 9, bei dem der Non-Text-Block dann nicht als Bilddaten bestimmt ist, wenn die Weißkonturfüllrate hoch ist.

11. Verfahren nach Anspruch 9, mit dem weiteren Verfahrensschritt des Kombinierens von Non-Gitter-angeordneten Weißkonturen (S622).

12. Verfahren nach Anspruch 11, bei dem der Non-Text-Block dann als Tabelle bestimmt ist, wenn die Rekombinationsrate nicht hoch ist (S624).

13. Verfahren nach Anspruch 7, das die Weißkonturen in vier Richtungen errechnet (S617).

14. Verfahren nach Anspruch 1, das Konturen verbundener Komponenten in wenigstens 8 Richtungen umreißt (S625).

15. Verfahren nach Anspruch 1, das im Verfahrensschritt des Umreißen Konturen verbundener Komponenten nur an äußeren Abschnitten der verbundenen Komponenten umreißt.

16. Verfahren nach Anspruch 1, mit dem weiteren Verfahrensschritt des Feststellens von Spalten in Pixelbilddaten, und wobei im ersten Verbindungsschritt verbundene Komponenten dann nicht der Breite nach in Zeilen im Falle verbunden werden, wenn ein Spalt die verbundenen Komponenten trennt (S57).

17. Verfahren nach Anspruch 16, das gemäß vertikal sich erstreckender Spalte Spalten zwischen verbundenen Komponenten feststellt.

18. Verfahren nach Anspruch 1, bei dem der zweite Verbindungsschritt den Verfahrensschritt des Bestimmens von Non-Text-Grenzen zwischen Zeilen von im ersten Verbindungsschritt verbundenen Textdaten umschließt, und wobei der zweite Verbindungsschritt Vertikalzeilen in Blöcken dann nicht verbindet, wenn es eine dazwischentretende Non-Text-Grenze gibt.

19. Verfahren nach Anspruch 1, mit dem weiteren Verfahrensschritt des Komprimierens der Pixelbilddaten vor dem Umreißen.

20. Gerät zur Auswahl von Blöcken von Pixeln aus Pixelbilddaten, mit:
einem Speichermittel (**32**), das Pixelbilddaten speichert;

einem Verarbeitungsmittel (**30**), das Konturen verbundener Komponenten in den gespeicherten Pixelbilddaten umreißt;

einem Mittel zum auswählbaren Verbinden verbundener Komponenten der Breite nach auf der Grundlage von Größe und Nähe zu anderen verbundenen Komponenten, um Zeilen zu bilden; und mit

einem Mittel zum auswählbaren Verbinden von Zeilen verbundener Komponenten vertikal der Breite nach auf der Grundlage von Größe und Nähe zu anderen Zeilen und zum Bilden von Rechtecken um die vertikal verbundenen Komponenten;

dadurchgekennzeichnet, daß das Verarbeitungsmittel ausgestattet ist mit

einem Mittel zu Erzeugen einer unregelmäßig geformten Kurve um jeden Block von Pixelbilddaten, der von einem Rechteck umgeben ist, das ein weiteres Rechteck überlappt.

21. Gerät nach Anspruch 20, das des weiteren

über ein Eingabemittel zur Eingabe von Pixelbilddaten in das Speichermittel **(32)** verfügt, wobei das Eingabemittel ein Umsetzmittel enthält, das die Pixeldaten dann in binäre Pixelbilddaten in dem Falle umsetzt, wenn die eingegebenen Pixeldaten keine binären Pixelbilddaten sind.

22. Gerät nach Anspruch 20, dessen Verarbeitungsmittel **(30)** eingerichtet ist, die kleinste unregelmäßig geformte Kurve um jeden Block von Pixelbilddaten zu erzeugen.

23. Gerät nach Anspruch 20, dessen Verarbeitungsmittel **(30)** eingerichtet ist, einen hierarchischen Baum der umrissenen verbundenen Komponenten zu bilden, wobei der hierarchische Baum wenigstens ein Block von Pixelbilddaten darstellt, der von der gebildeten unregelmäßig geformten Kurve umgeben ist.

24. Gerät nach Anspruch 20, dessen Verarbeitungsmittel **(30)** eingerichtet ist, die erzeugten unregelmäßig gestalteten Kurven in Textblöcke und Non-Text-Blöcke zu klassifizieren.

25. Gerät nach Anspruch 24, dessen Verarbeitungsmittel **(30)** eingerichtet ist, Zeichenbilder in Textblöcken zu erkennen.

26. Gerät nach Anspruch 24, dessen Verarbeitungsmittel **(30)** eingerichtet ist, weiße Konturen für umrissene verbundene Non-Text-Komponenten herzuweisen.

27. Gerät nach Anspruch 26, dessen Verarbeitungsmittel **(30)** eingerichtet ist, Non-Text-Blöcke als Tabelleneinheiten gemäß der Anzahl weißer Konturen zu bestimmen.

28. Gerät nach Anspruch 26, dessen Verarbeitungsmittel **(30)** eingerichtet ist, eine Weißkonturfüllrate zu errechnen.

29. Gerät nach Anspruch 28, dessen Verarbeitungsmittel **(30)** eingerichtet ist, den Non-Text-Block nicht als Bilddaten dann zu bestimmen, wenn die Weißkonturfüllrate hoch ist.

30. Gerät nach Anspruch 28, dessen Verarbeitungsmittel **(30)** eingerichtet ist, angeordnete Nicht-Gitter-Weißkonturen zu rekombinieren.

31. Gerät nach Anspruch 30, dessen Verarbeitungsmittel **(30)** eingerichtet ist, den Non-Text-Block als eine Tabelle dann zu bestimmen, wenn die Rekombinationsrate nicht hoch ist.

32. Gerät nach Anspruch 26, dessen Verarbeitungsmittel **(30)** eingerichtet ist, Weißkonturen in vier Richtungen zu errechnen.

33. Gerät nach Anspruch 20, dessen Verarbeitungsmittel **(30)** eingerichtet ist, Konturen verbundener Komponenten in wenigstens 8 Richtungen zu umreißen.

34. Gerät nach Anspruch 20, dessen Verarbeitungsmittel **(30)** eingerichtet ist, die Konturen verbundener Komponenten nur an äußeren Abschnitten der verbundenen Komponenten zu umreißen.

35. Gerät nach Anspruch 20, dessen Verarbeitungsmittel **(30)** eingerichtet ist, Spalten in den Pixelbilddaten festzustellen und verbundene Komponenten der Breite nach in Zeilen dann nicht zu verbinden, wenn ein Spalt die Komponenten trennt.

36. Gerät nach Anspruch 35, dessen Verarbeitungsmittel **(30)** eingerichtet ist, Spalten gemäß vertikal sich erstreckender Spalte zwischen verbundenen Komponenten festzustellen.

37. Gerät nach Anspruch 20, dessen Verarbeitungsmittel **(30)** eingerichtet ist, Non-Text-Grenzen zwischen Zeilen verbundener Textdaten zu bestimmen und vertikale Zeilen in Blöcke dann nicht zu verbinden, wenn es eine zwischentretende Non-Text-Grenze gibt.

38. Gerät nach Anspruch 20, dessen Verarbeitungsmittel **(30)** eingerichtet ist, die Pixelbilddaten vor Umreißen zu komprimieren.

39. Verfahren nach Anspruch 1, mit dem weiteren Verfahrensschritt:

Bestimmen, ob die umrissenen verbundenen Komponenten Texteinheiten oder Non-Text-Einheiten enthalten;

wobei der erste Verbindungsschritt selektives Verbinden von Texteinheiten der Breite nach umfaßt, um Textzeilen auf der Grundlage von Nähe angrenzender Texteinheiten zu bilden;

wobei der zweite Verbindungsschritt selektives Verbinden von vertikalen Textzeilen umfaßt, um Blöcke von Texteinheiten auf der Grundlage von Nähe und angrenzenden Textzeilen und der Position von Non-Text-Einheiten zwischen Textzeilen zu bilden; wobei der Bildungsschritt das Umgeben eines jeden Blockes von Texteinheiten in den Blöcken von Texteinheiten mit einer unregelmäßig gestalteten Kurve umfaßt, wenn Rechtecke um die Blöcke von Texteinheiten sich überlappen;

Bilden eines hierarchischen Baumes von Blöcken von Texteinheiten auf der Grundlage der verbundenen Komponenten, die der Verfahrensschritt des Umreißen umreißt (S603), und

Segmentieren der baumgruppierten Blöcke von Texteinheiten in Zeilen von Pixelbilddaten, die Blöcke von Texteinheiten in wenigstens einer Zeile adaptiv teilen.

40. Verfahren nach Anspruch 39, mit dem weiteren Verfahrensschritt des Vorverarbeitens von Pixelbilddaten (S402).

41. Verfahren nach Anspruch 40, dessen Vorverarbeitungsschritt den Verfahrensschritt des Bildkomprimierens umfaßt.

42. Verfahren nach Anspruch 41, bei dem der Vorverarbeitungsschritt den Verfahrensschritt des Erhörens von Pixelbilddaten umfaßt.

43. Verfahren nach Anspruch 39, mit dem weiteren Verfahrensschritt der Weißkonturspurführung des Inneren von Non-Text-Einheiten.

44. Verfahren nach Anspruch 43, mit dem weiteren Verfahrensschritt des Anfügens von Markierern an Non-Text-Einheiten (S645).

45. Verfahren nach Anspruch 43, das die Tabellenmarkierer auf der Grundlage der Anzahl enthaltener Weißkonturen an Non-Text-Einheiten anhängt (S628).

46. Gerät nach Anspruch 20, dessen Prozessormittel des weiteren Verwendung findet zum:

Bestimmen, ob umrissene Komponenten Texteinheiten oder Non-Text-Einheiten enthalten, Ausführen des Umreißschrittes, auswählbaren Verbinden von Komponenten, auswählbaren Verbinden von Zeilen und Bilden bezüglich verbundener Komponenten, die als Texteinheiten bestimmt sind, wobei Texteinheiten der Breite nach verbunden sind, um Textzeilen auf der Grundlage von Nähe zu anderen Textzeilen zu bilden, wobei Textzeilen auf der Grundlage von Nähe benachbarter Außenzeilen und bezüglich der Lage von Non-Text-Einheiten zwischen Textzeilen zum Bilden von Rechtecken um Blöcke von Texteinheiten vertikal verbunden werden, und wobei jeder Block einer Texteinheit mit einer unregelmäßig gestalteten Kurve umgeben wird, wenn Rechtecke um die Blöcke von Texteinheiten sich überlappen, zum Bilden eines hierarchischen Baumes von Blöcken von Texteinheiten auf der Grundlage der umrissenen verbundenen Komponenten und zum Segmentieren der baumgestalteten Textblöcke in Zeilen von Pixelbilddaten durch adaptives Teilen von Blöcken von Texteinheiten in wenigstens eine Zeile.

47. Gerät nach Anspruch 46, dessen Verarbeitungsmittel (30) eingerichtet ist, die eingegebenen Pixelbilddaten vorzuverarbeiten.

48. Verfahren nach Anspruch 39, mit den weiteren Verfahrensschritten
Feststellen einer Textausrichtung für jeden gebildeten Textblock, wobei die Textausrichtung horizontale Textausrichtung und vertikale Textausrichtung umfaßt;

Bilden von Textzeilen aus dem gebildeten Textblock basierend auf der festgestellten Textausrichtung; und danach

Feststellen eines Schräglaufwinkels für jeden der gebildeten Textblöcke.

49. Verfahren nach Anspruch 48, bei dem der Verfahrensschritt des Gruppierens den Verfahrensschritt des Trennens von Text und verbundenen Non-Text-Komponenten umfaßt (S52).

50. Verfahren nach Anspruch 49, das des weiteren über den Verfahrensschritt des Klassifizierens von Non-Text-Komponenten in wenigstens eine einer vorbestimmten Vielzahl von Non-Text-Klassifizierungen umfaßt (S53).

51. Verfahren nach Anspruch 49, das den weiteren Verfahrensschritt des Suchens vom gespeicherten Dokument für unsichtbare Zeilen unter Kanten der verbundenen Non-Text-Komponenten umfaßt.

52. Verfahren nach Anspruch 48, mit dem weiteren Verfahrensschritt des Splittens eines gebildeten Textblockes, wenn eine sichtbare Zeile oder eine unsichtbare Zeile durchläuft.

53. Verfahren nach Anspruch 48, mit dem weiteren Verfahrensschritt des Bildens von Textzeilen für einen jeden Textblock auf der Grundlage des festgestellten Schrägstellungswinkels.

54. Verfahren nach Anspruch 48, mit dem weiteren Verfahrensschritt der Zeichenerkennung des Pixels, das Komponenten innerhalb der gebildeten Textblöcke verbindet.

55. Gerät nach Anspruch 46, dessen Prozessormittel des weiteren ausgestattet ist mit:
einem Mittel zum Feststellen einer Textausrichtung für jeden gebildeten Textblock, wobei die Textausrichtung horizontale Textausrichtung und vertikale Textausrichtung umfaßt, zum Bilden von Textzeilen für den gebildeten Textblock auf der Grundlage der festgestellten Textausrichtung und danach zum Feststellen eines Schrägstellungswinkels für jeden gebildeten Textblock.

56. Gerät nach Anspruch 55, dessen Verarbeitungsmittel (30) eingerichtet ist zum Trennen von Text- und Non-Text-Pixeln, die Pixelkomponenten verbinden.

57. Gerät nach Anspruch 55, dessen Verarbeitungsmittel (30) eingerichtet ist, die Non-Text-Komponenten in wenigstens eine einer Vielzahl von Non-Text-Klassifizierungen zu klassifizieren.

58. Gerät nach Anspruch 56, dessen Verarbeitungsmittel (30) eingerichtet ist, das gespeicherte

Dokument für unsichtbare Zeilen entlang Kanten von verbundenen Non-Text-Komponenten zu suchen.

54 programmiert ist.

Es folgen 25 Blatt Zeichnungen

59. Gerät nach Anspruch 55, dessen Verarbeitungsmittel (30) eingerichtet ist, einen gebildeten Textblock zu splitten, wenn eine sichtbare Zeile oder eine unsichtbare Zeile durchläuft.

60. Gerät nach Anspruch 55, dessen Verarbeitungsmittel (30) eingerichtet ist zum Bilden von Textzeilen für jeden Textblock auf der Grundlage des festgestellten Schrägstellwinkels.

61. Gerät nach Anspruch 55, dessen Verarbeitungsmittel (30) eingerichtet ist zur Zeichenerkennung der verbundenen Pixelkomponenten in einem jeden der gebildeten Textblöcke.

62. Verfahren nach Anspruch 39, das des Weiteren über die Verfahrensschritte verfügt:
Suchen des Dokuments für sichtbare Zeilen und für unsichtbare Zeilen längs Kanten von Non-Text-Komponenten;
Bilden von Non-Text-Blöcken und dem Textblock unter Verwendung der sichtbaren und unsichtbaren Zeilen ohne Ausführen der Schrägstellfeststellung;
Feststellen einer Textausrichtung für jeden gebildeten Textblock, wobei die Textausrichtung horizontale Textausrichtung und vertikale Textausrichtung umfaßt;
Auslesen von Textzeilen aus dem Textblock auf der Grundlage festgestellter Textausrichtung;
danach Feststellen eines Schrägstellwinkels für einen jeden der gebildeten Textblöcke; und
Modifizieren der gebildeten Textblöcke auf der Grundlage des festgestellten Schrägstellwinkels.

63. Gerät nach Anspruch 46, dessen Prozessormittel ausgestattet ist mit:
einem Mittel zum Suchen des gespeicherten Dokuments für sichtbare Zeilen und für unsichtbare Zeilen entlang Kanten von Non-Text-Komponenten, zum Bilden von Non-Text-Blöcken und des Textblockes der Textblöcke unter Verwendung der sichtbaren und unsichtbaren Zeilen ohne die Ausführung der Schrägstellfeststellung, zum Feststellen einer Textausrichtung für jeden gebildeten Textblock, wobei die Textausrichtung horizontale Textausrichtung und vertikale Textausrichtung umfaßt, zum Auslesen von Textzeilen aus den Textblöcken auf der Grundlage der festgestellten Textausrichtung, danach Feststellen eines Schrägstellwinkels für jeden der gebildeten Textblöcke und zum Modifizieren der gebildeten Textblöcke auf der Grundlage des festgestellten Schrägstellwinkels.

64. Computerprogrammprodukt, das mit prozessorbearbeitbaren Befehlen zum Ausführen eines Verfahrens nach einem jeden der Verfahrensschritte eines der Patentansprüche 1–19, 39 bis 45 und 48 bis

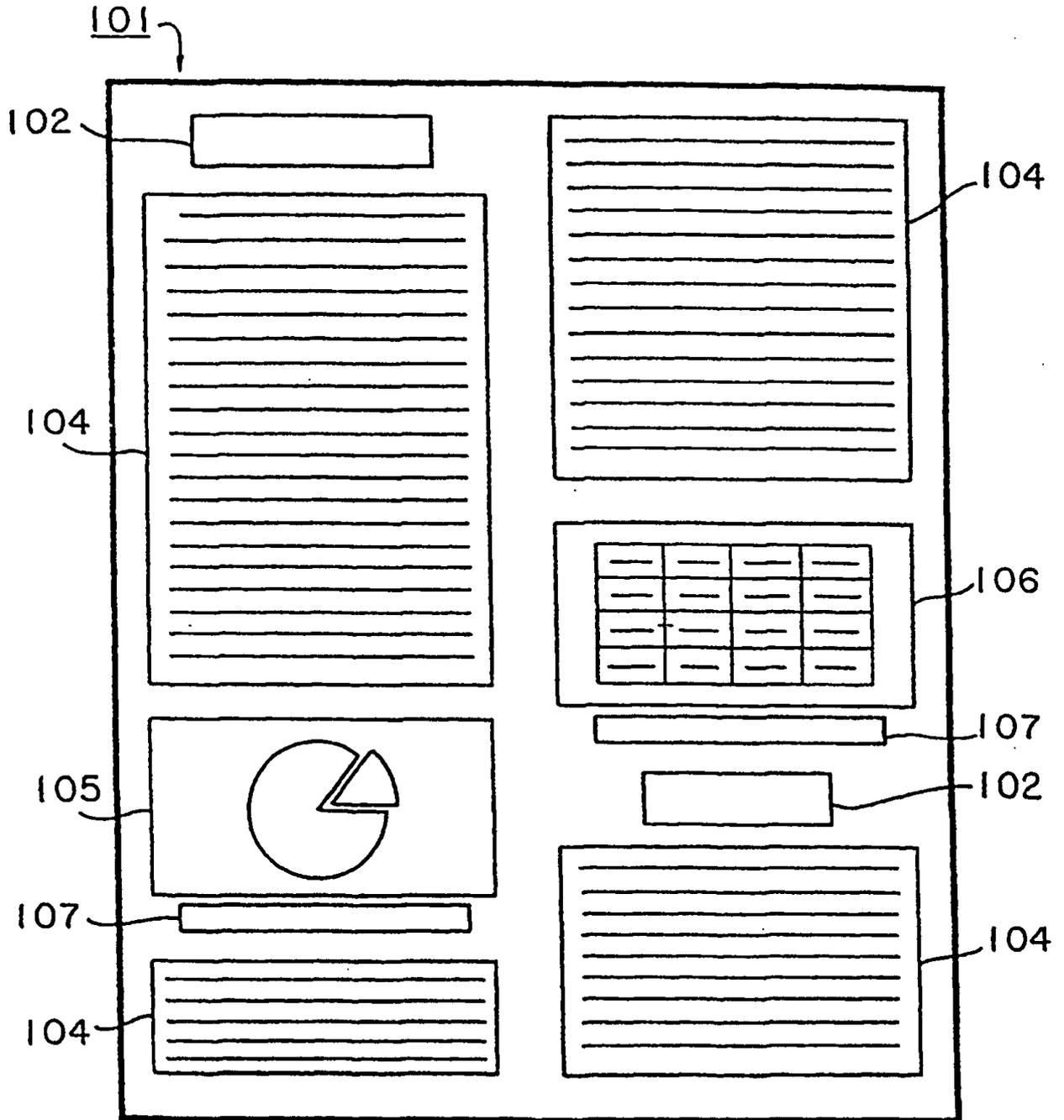
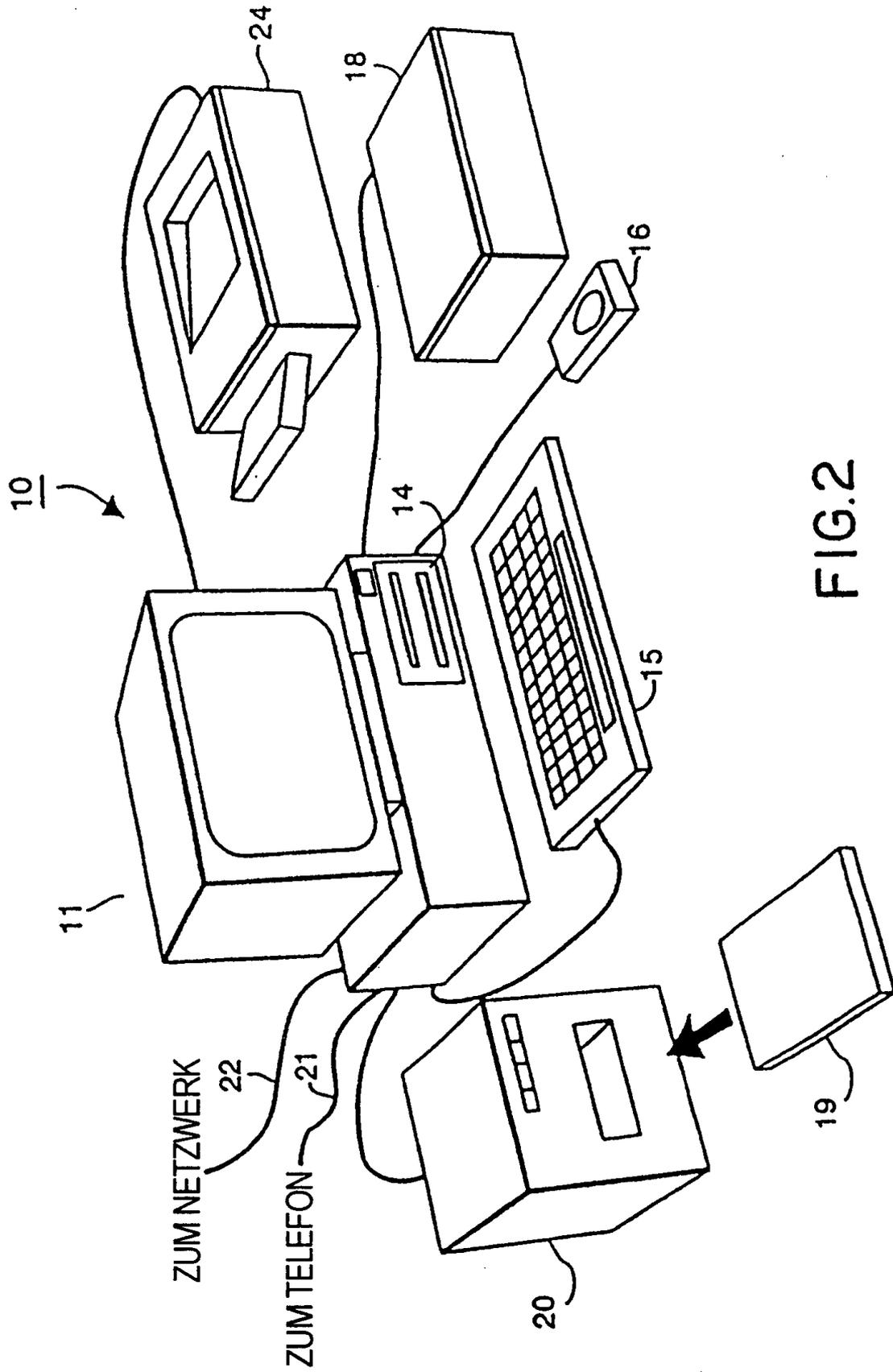


FIG. 1



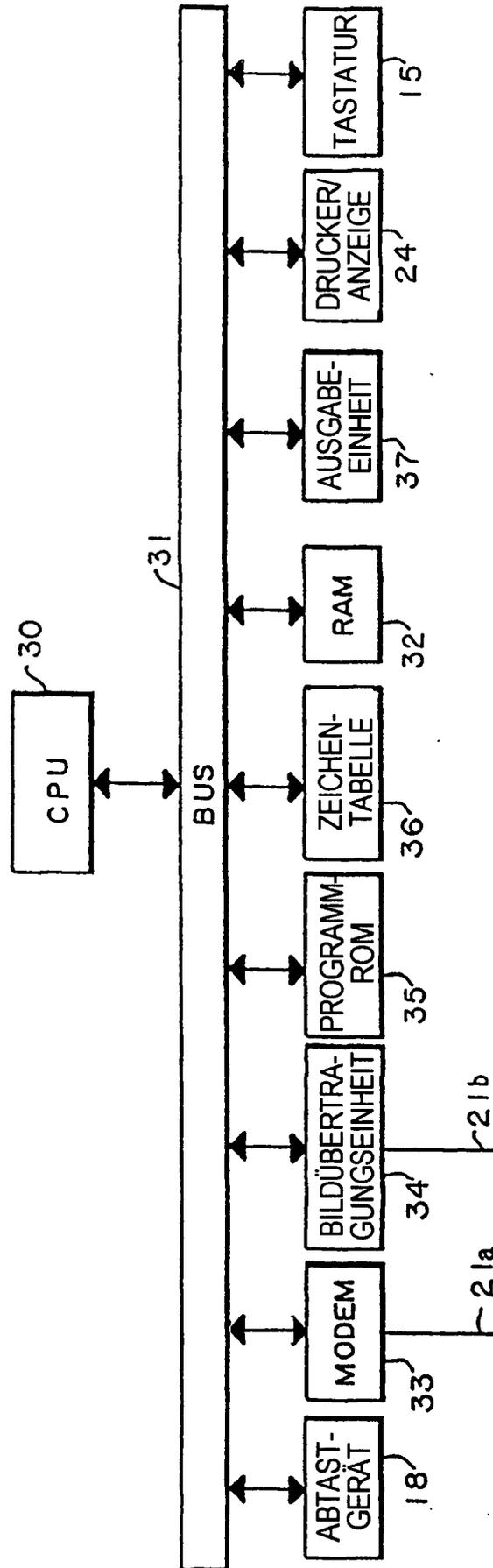


FIG. 3

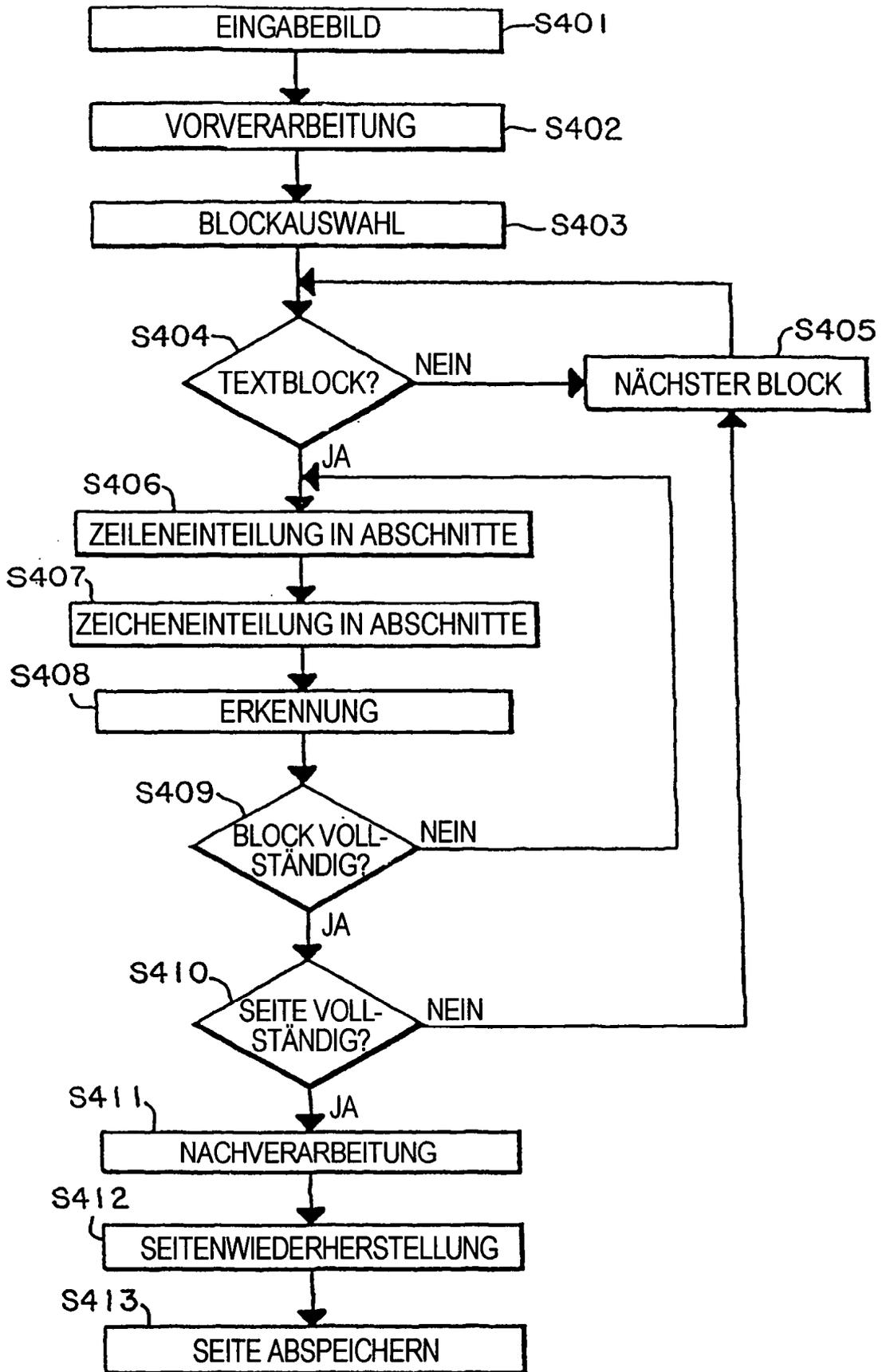
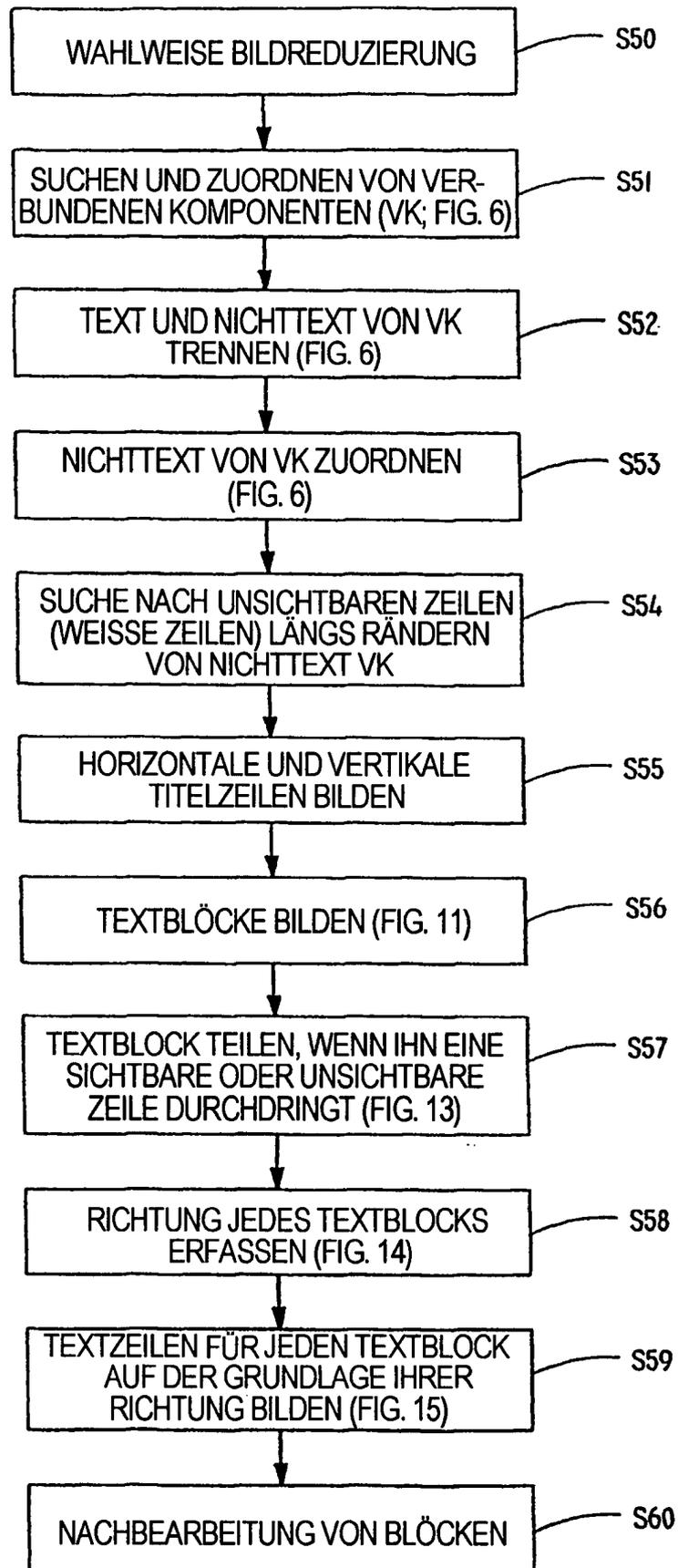


FIG. 4

**FIG. 5**

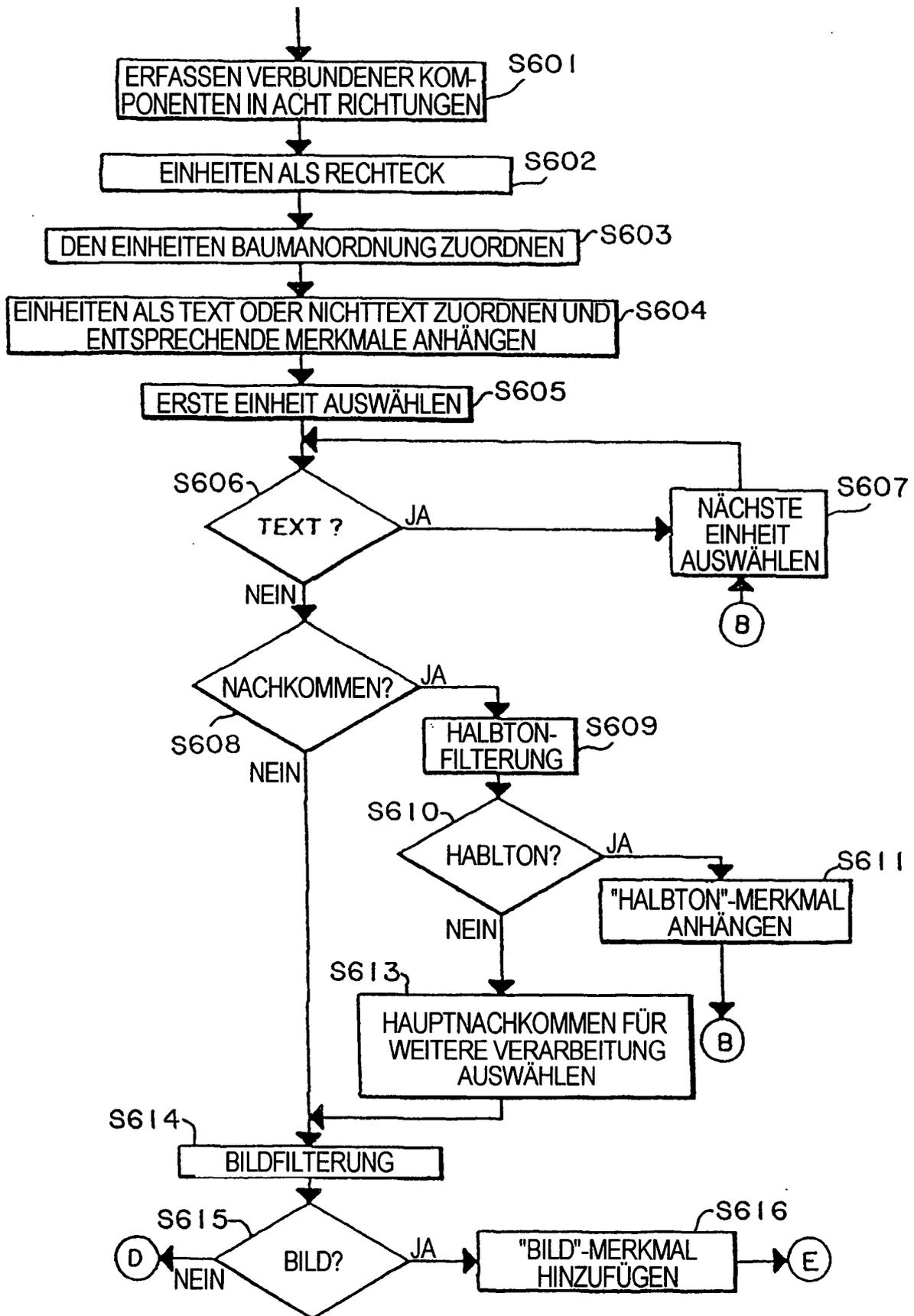


FIG. 6A

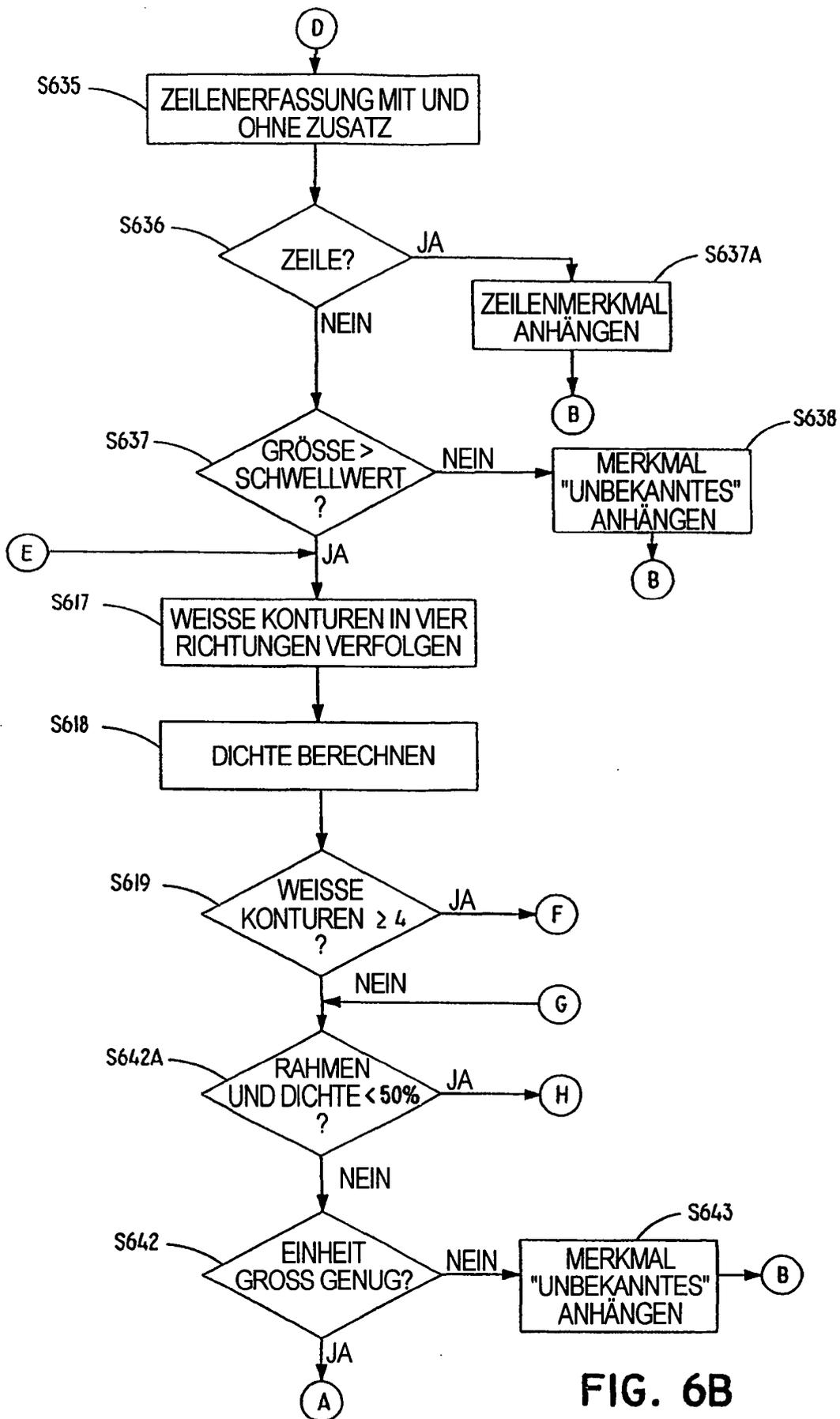


FIG. 6B

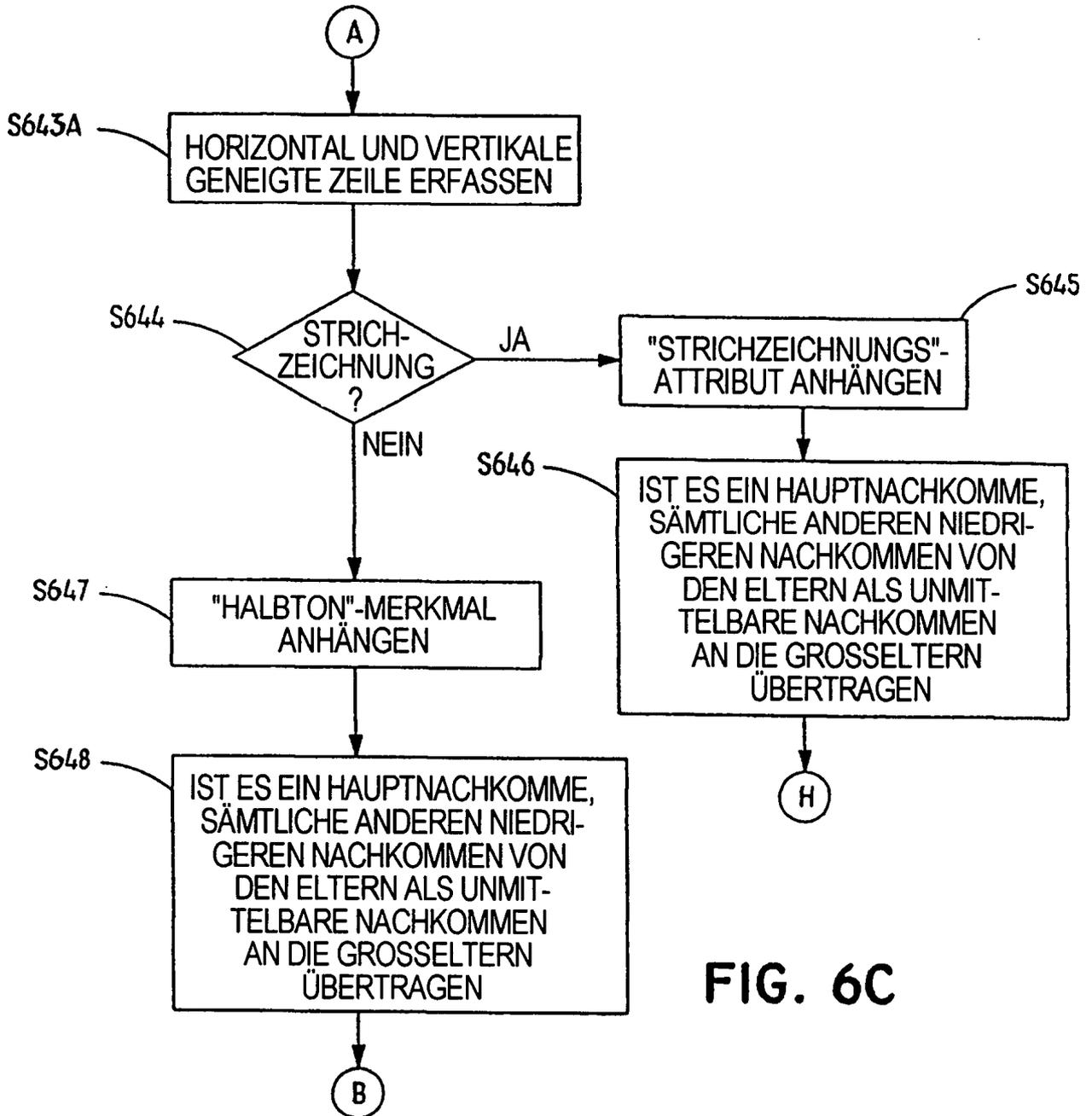


FIG. 6C

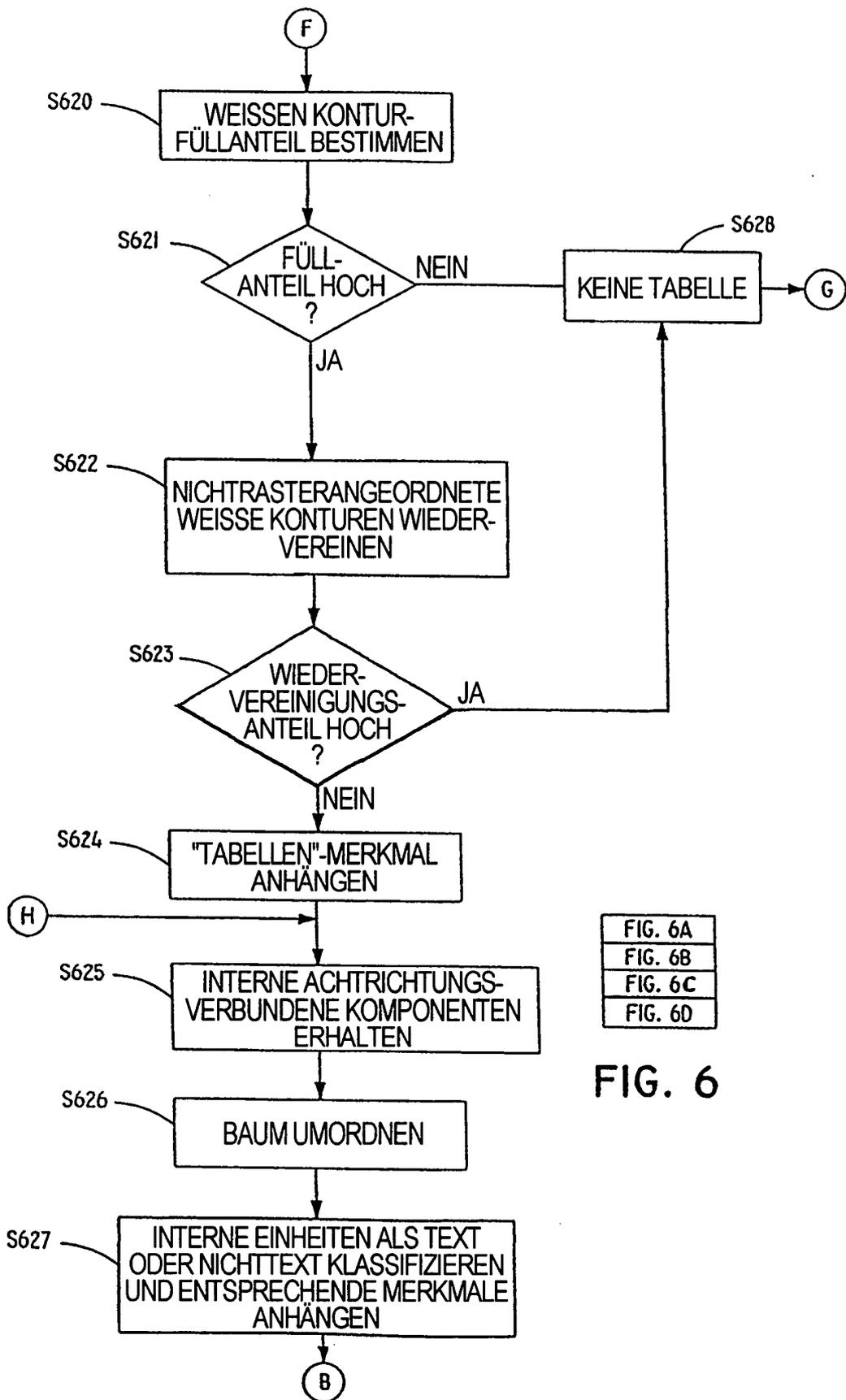
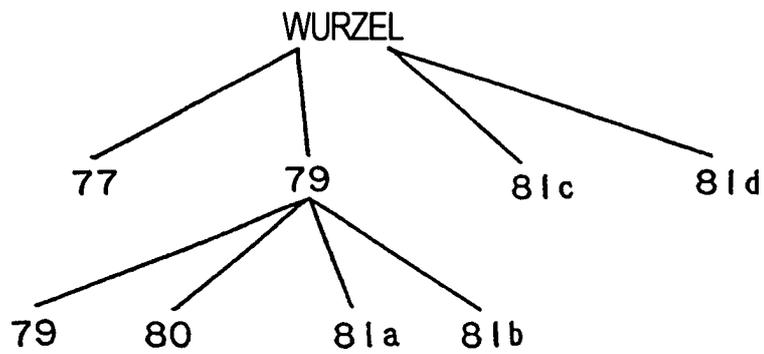
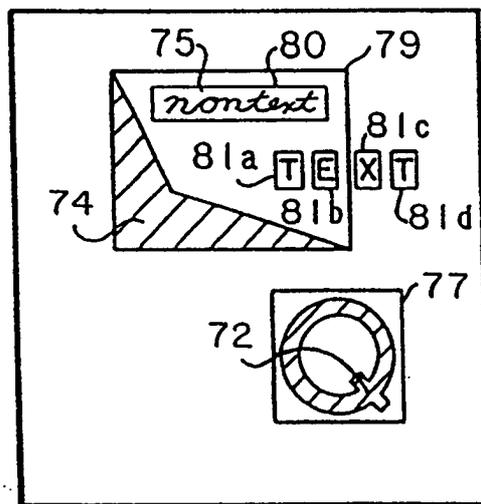
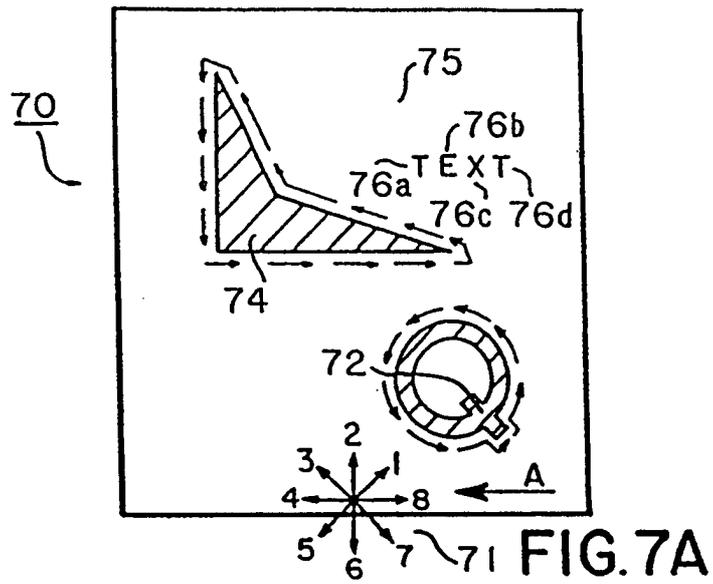


FIG. 6A
FIG. 6B
FIG. 6C
FIG. 6D

FIG. 6

FIG. 6D



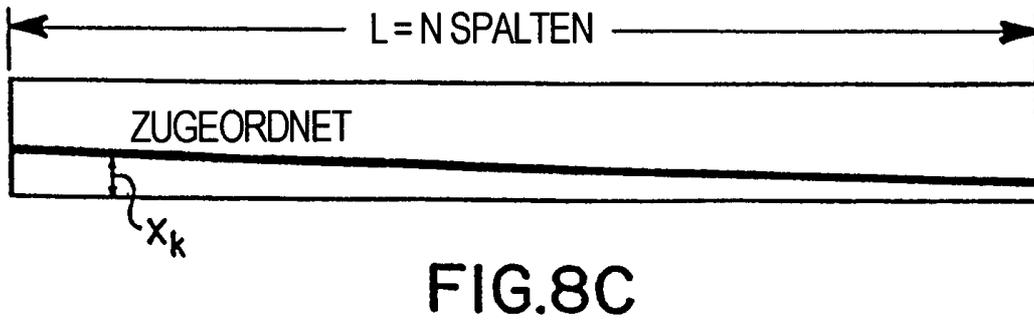
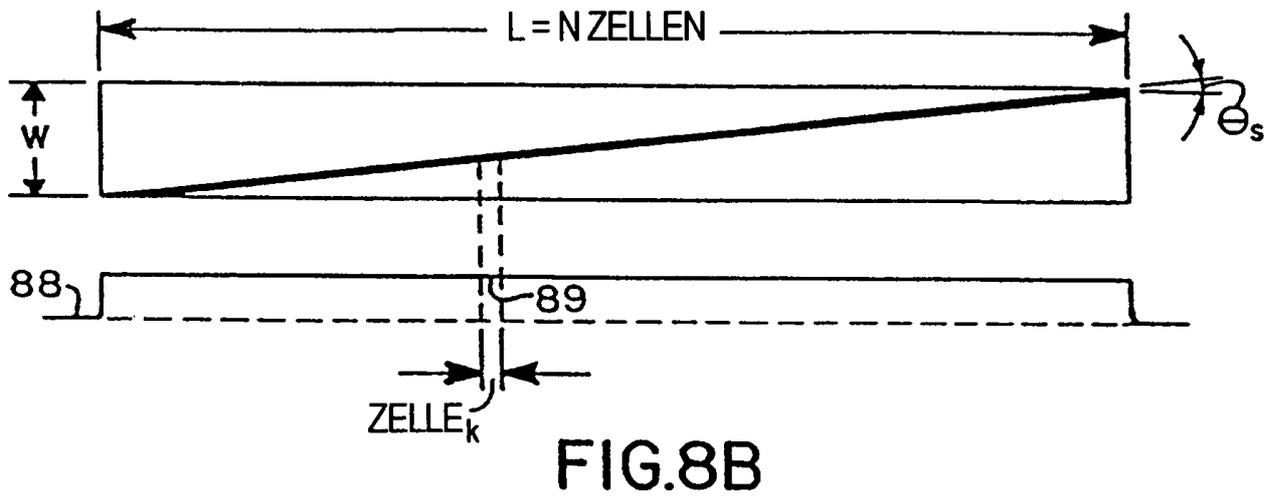
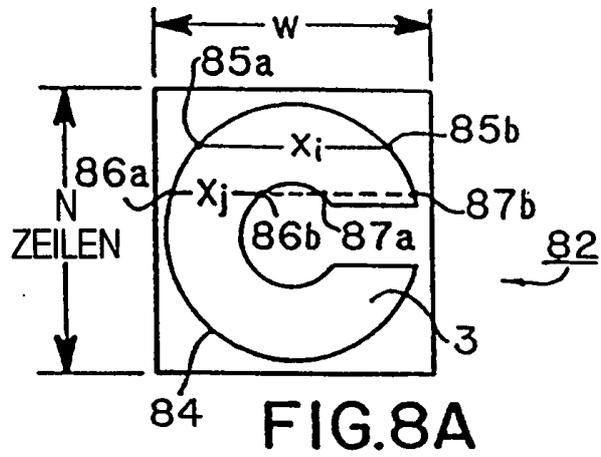


FIG.9A

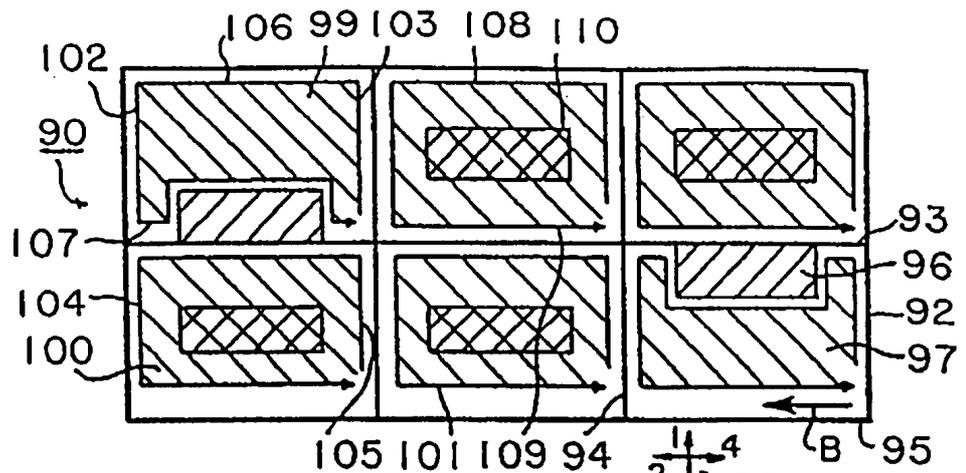


FIG.9B

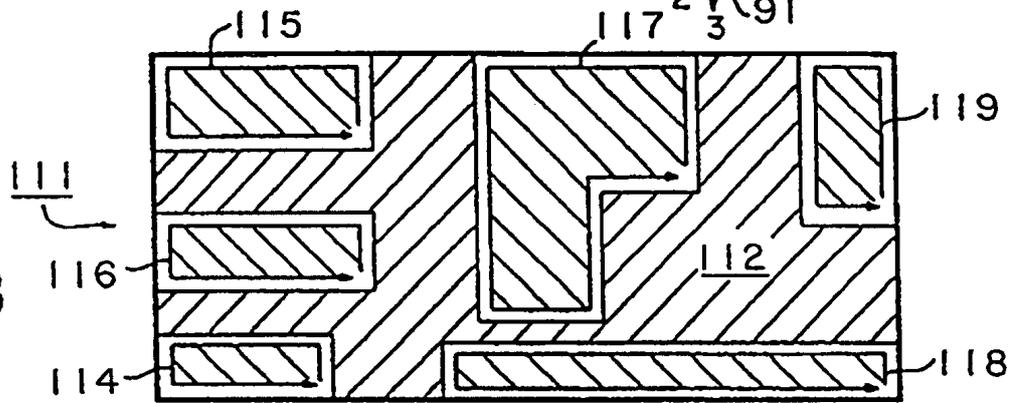


FIG.9C

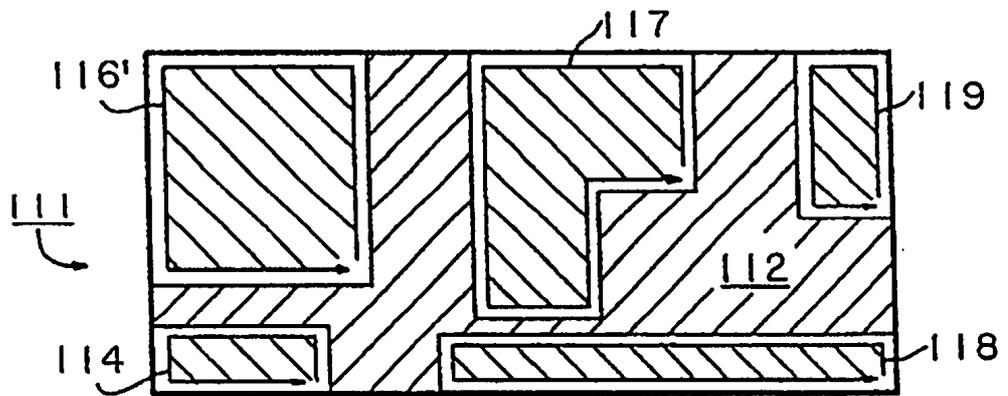
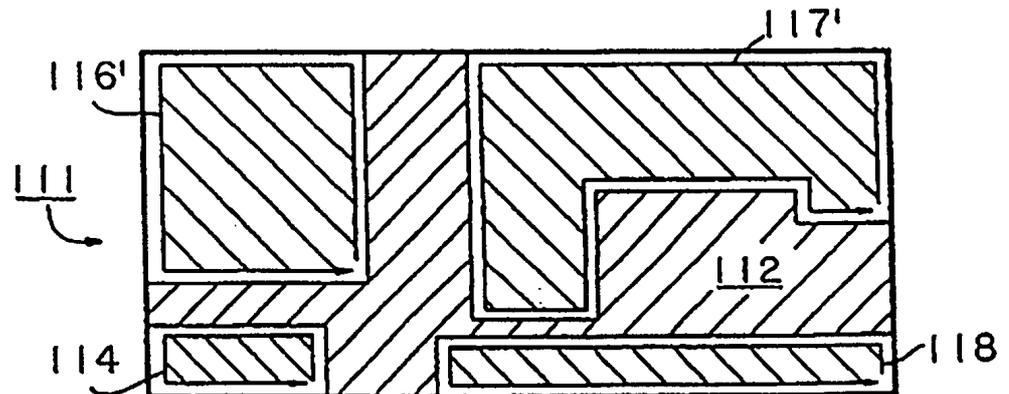


FIG.9D



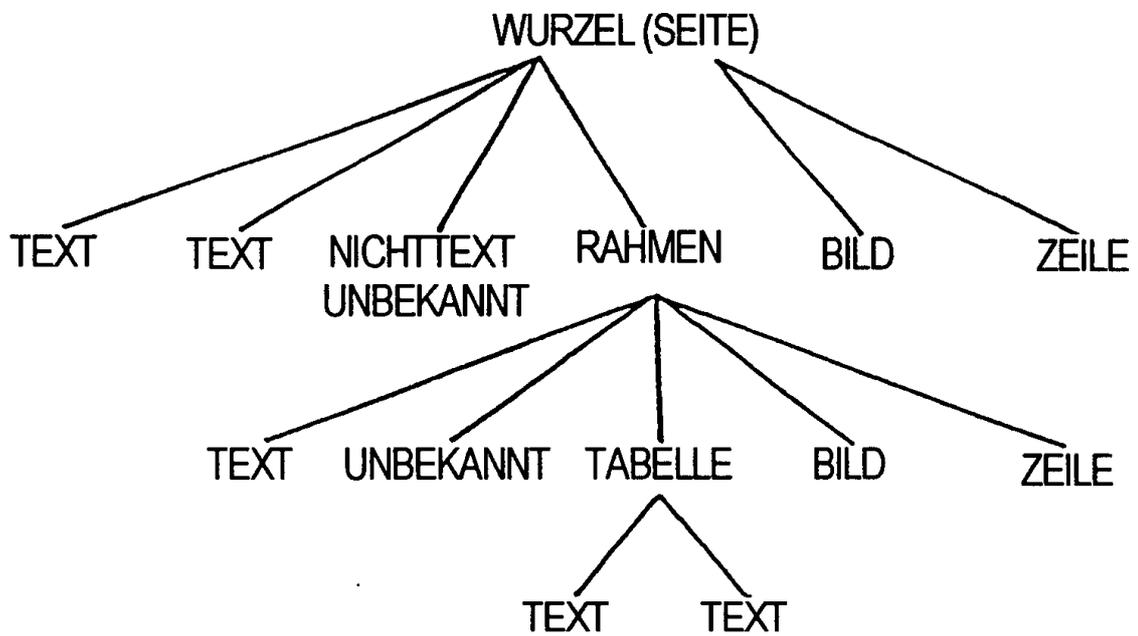


FIG.10

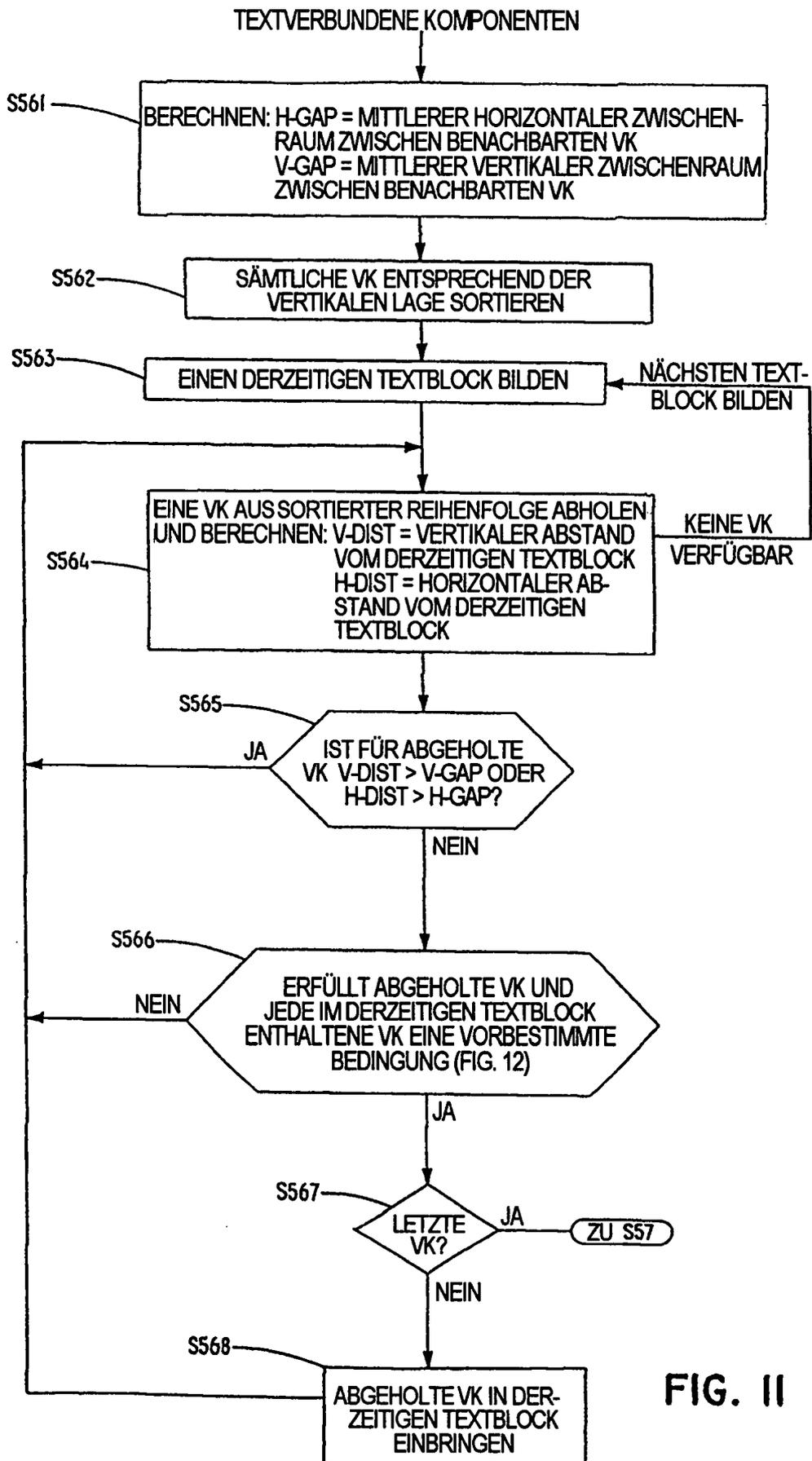


FIG. II

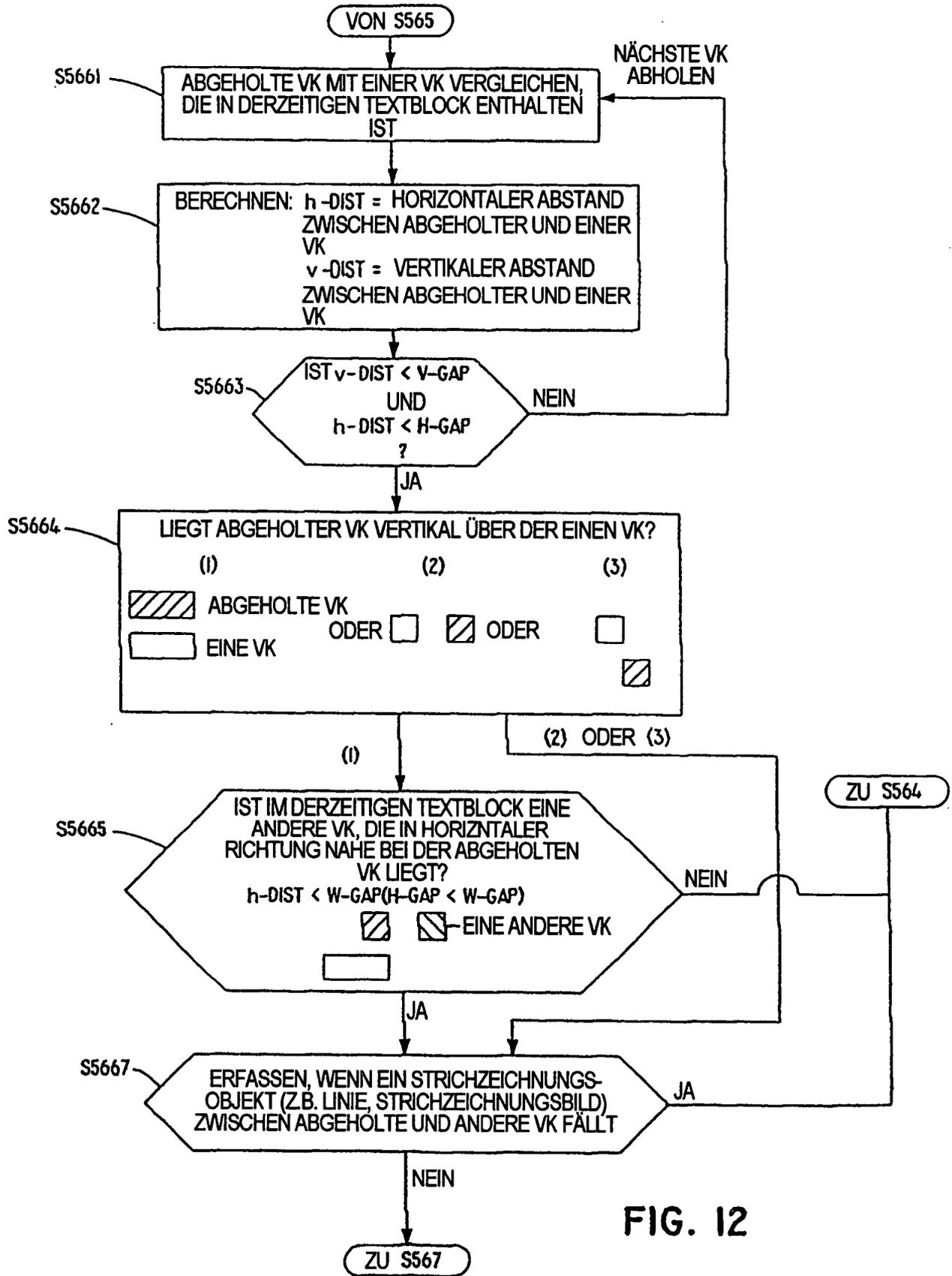


FIG. 12

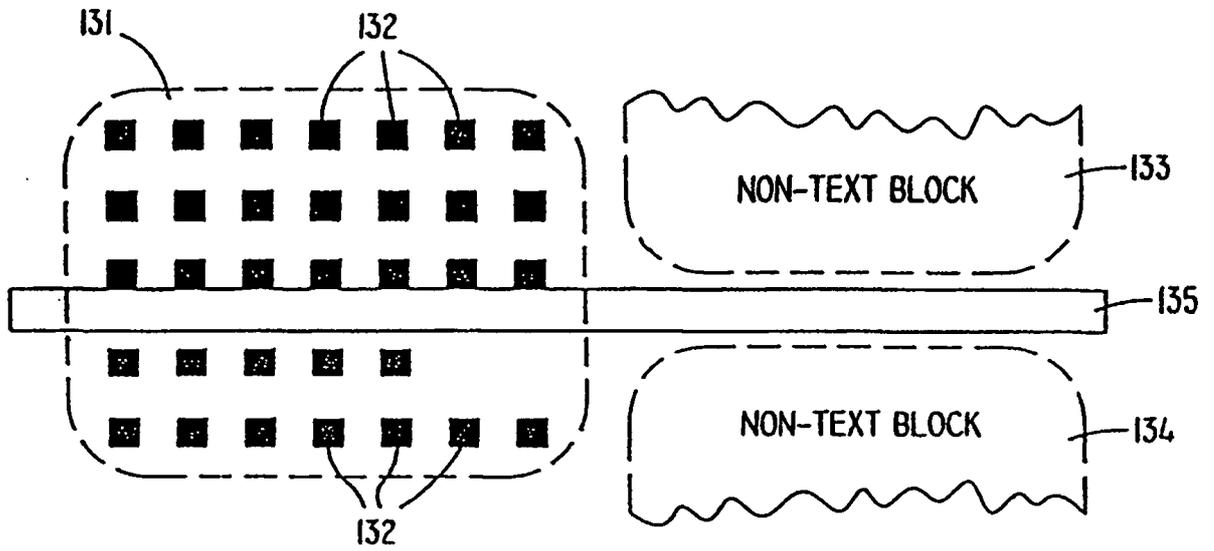


FIG. 13A

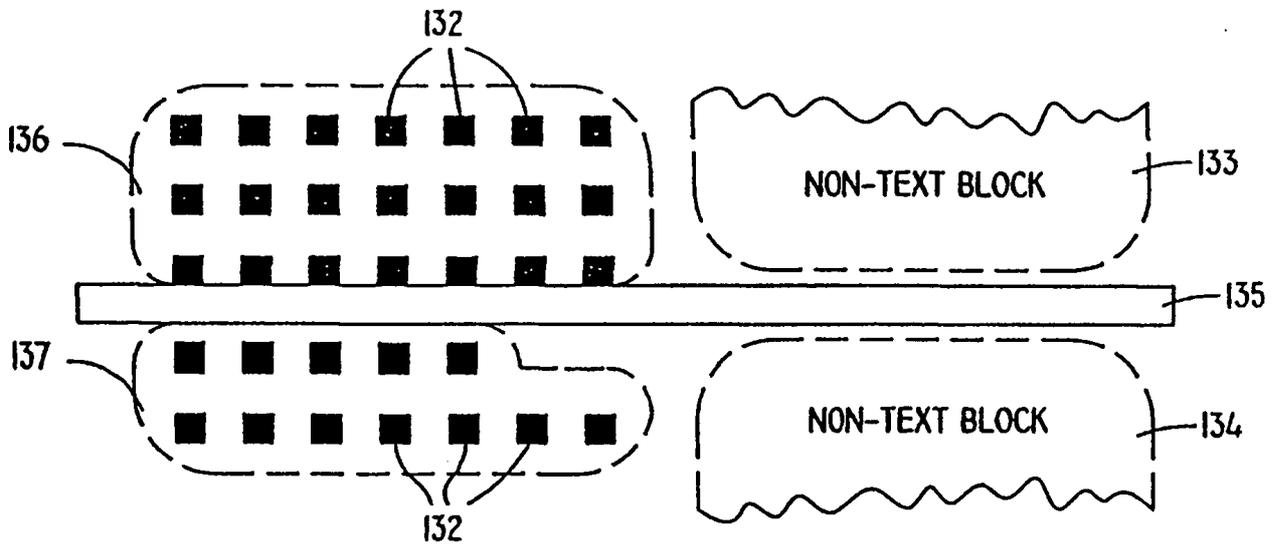


FIG. 13B

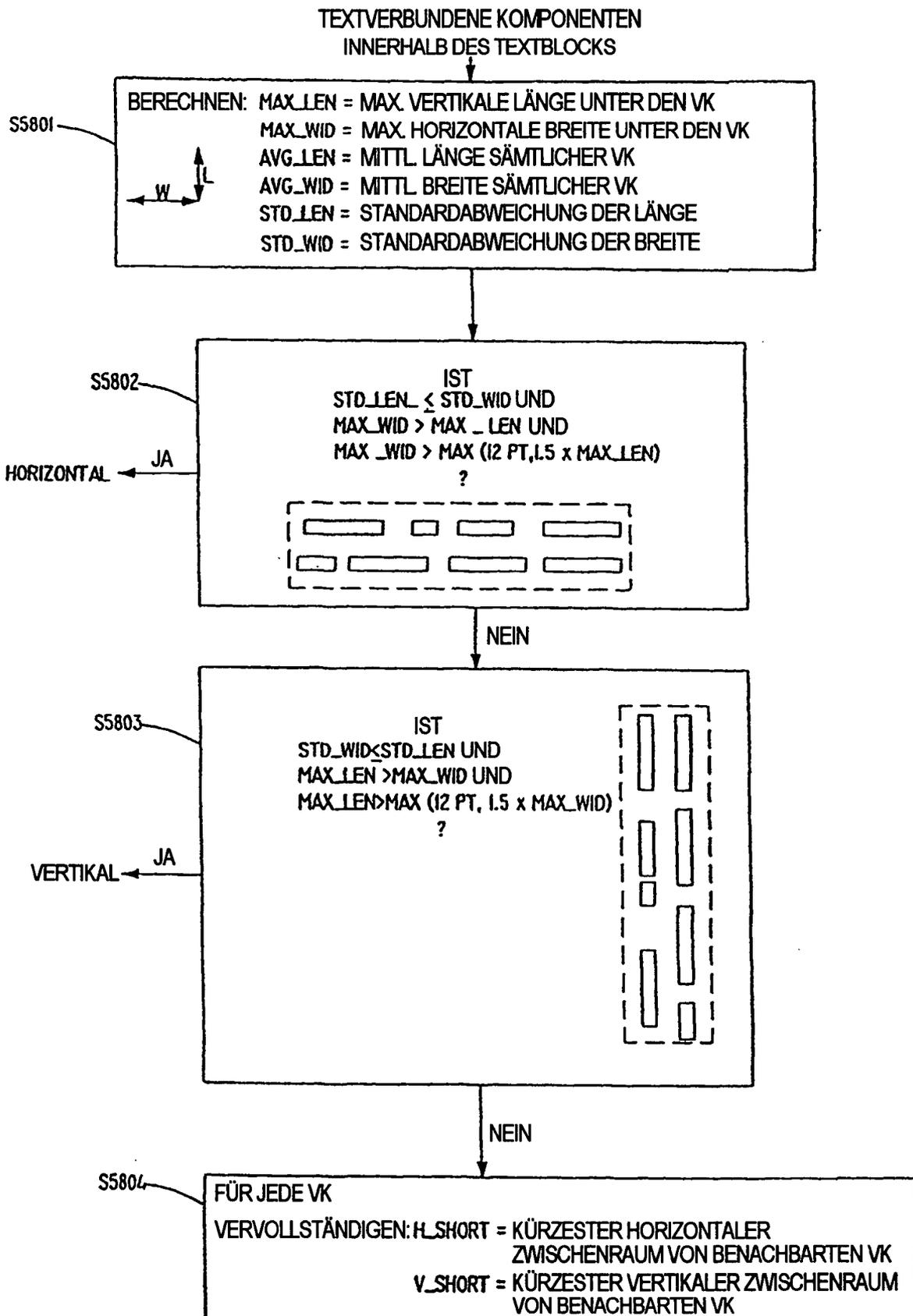


FIG. 14A

A

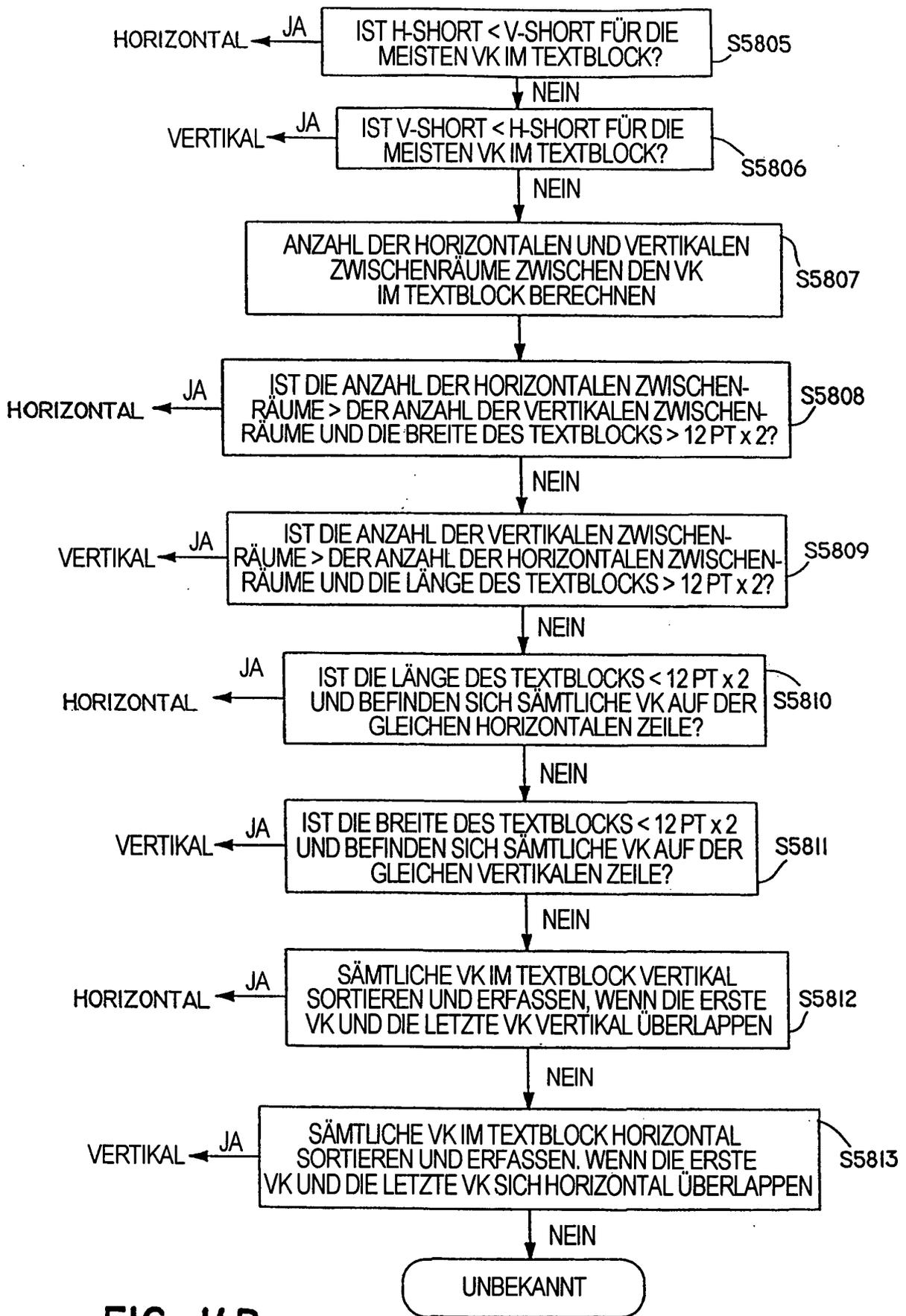


FIG. 14B

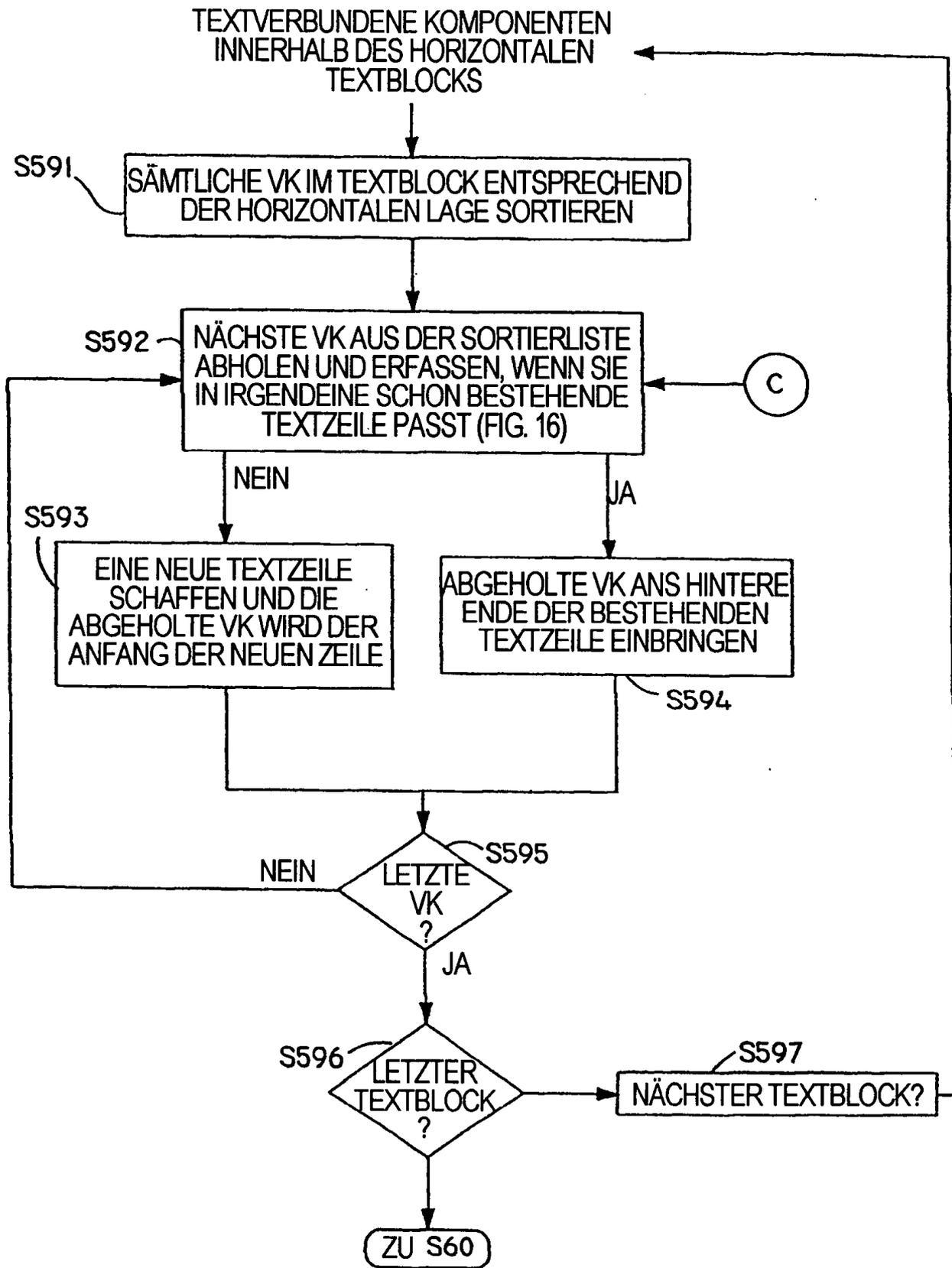


FIG. 15

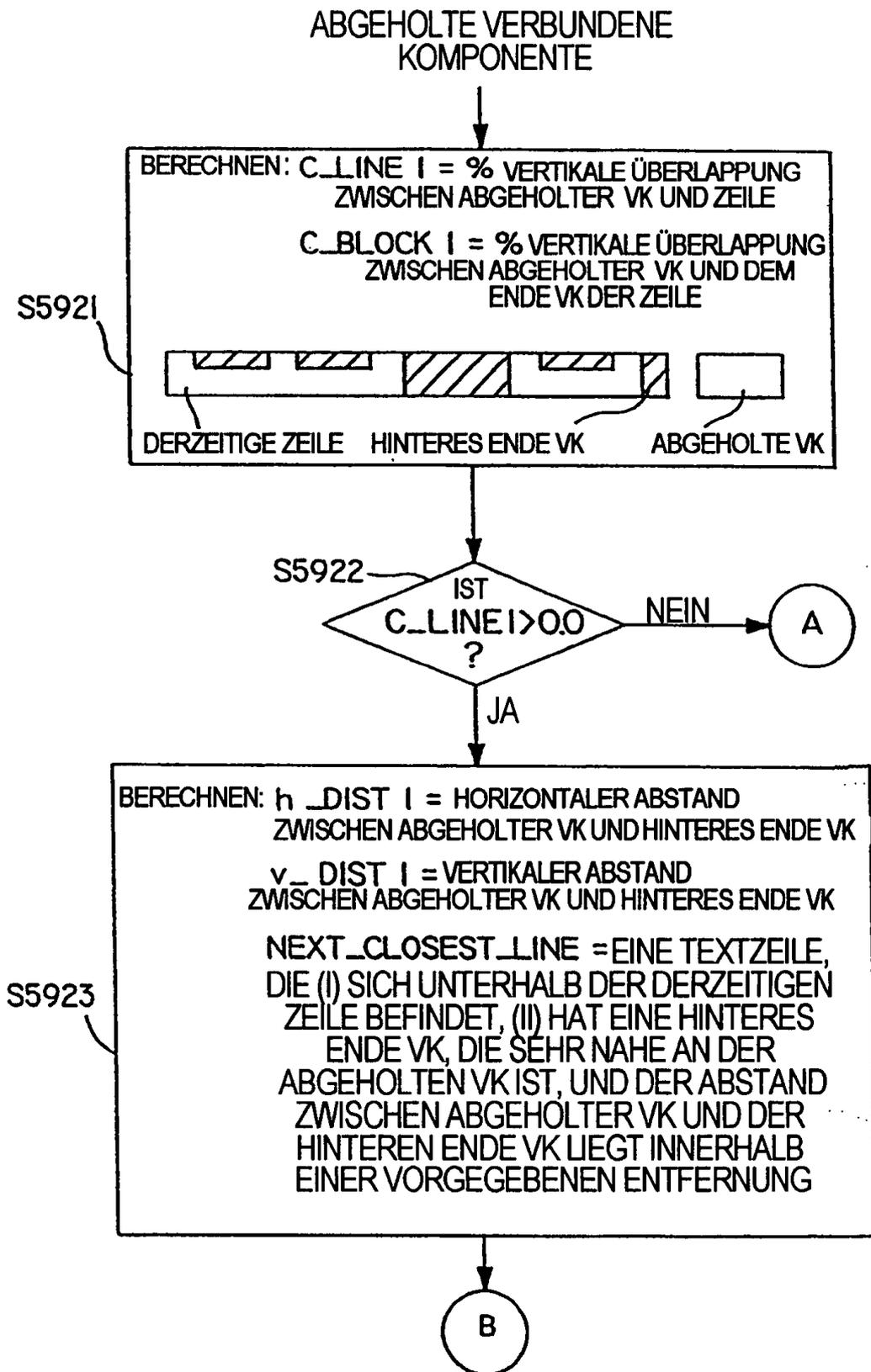


FIG. 16A

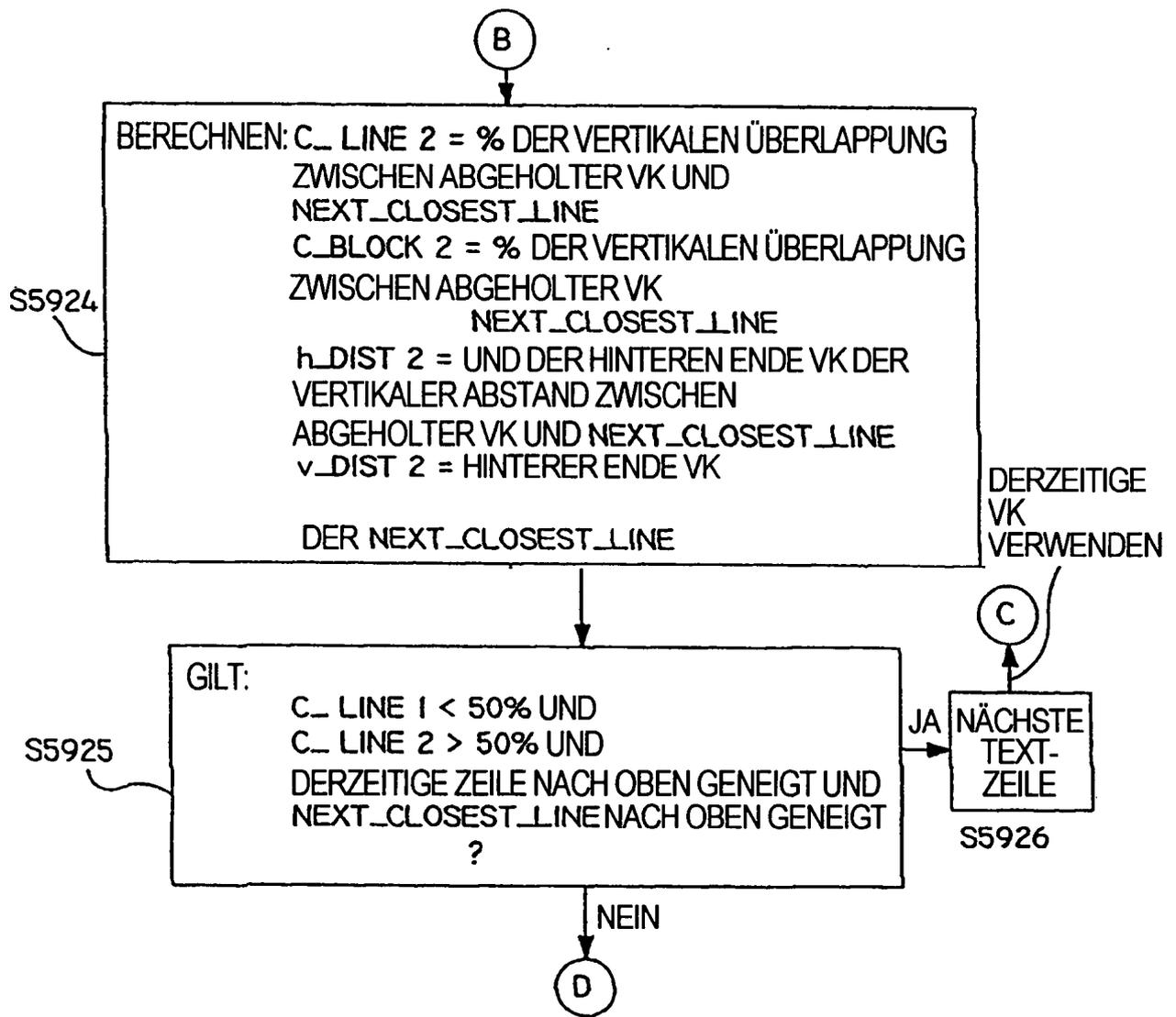


FIG. 16B

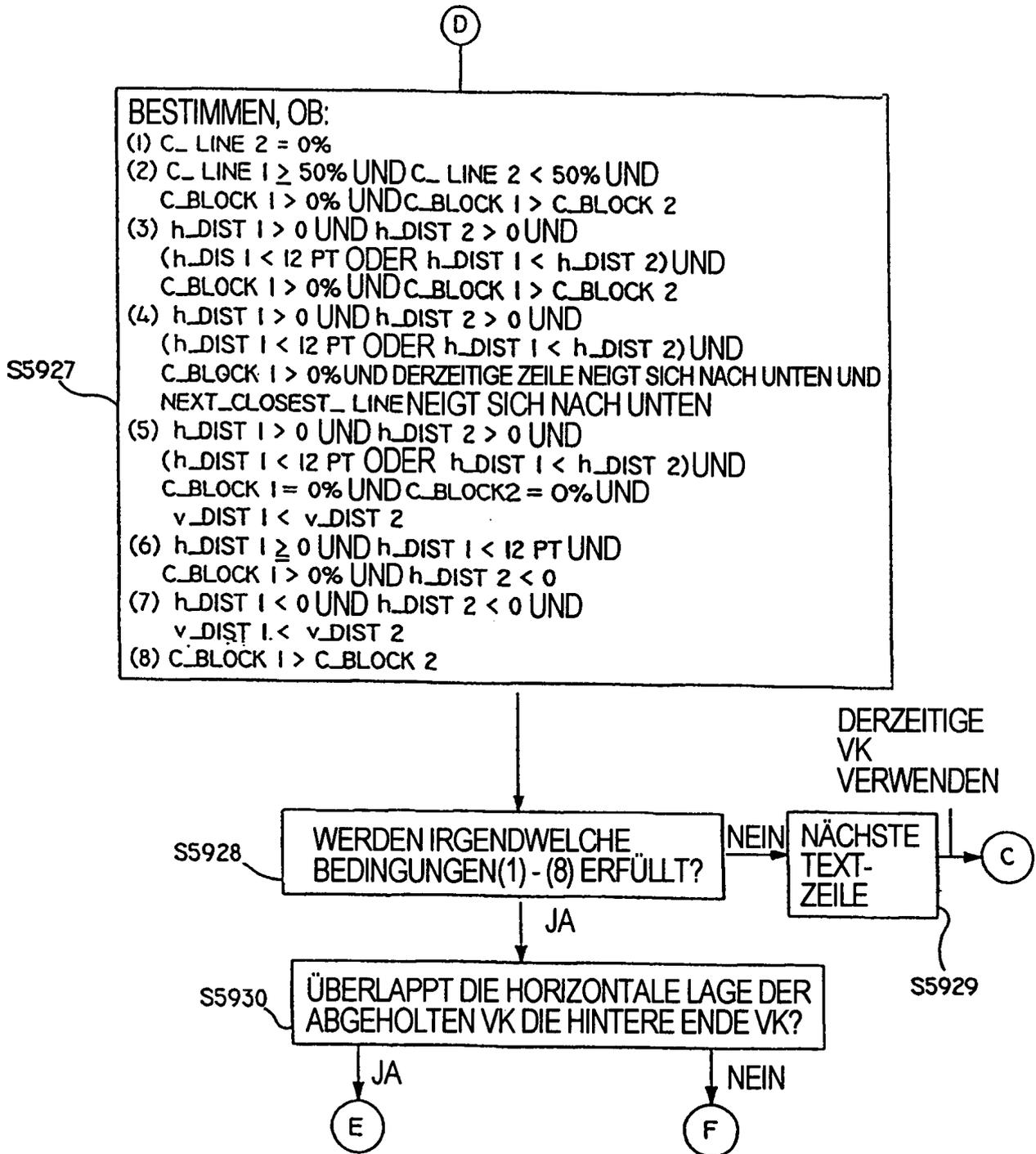


FIG. 16C

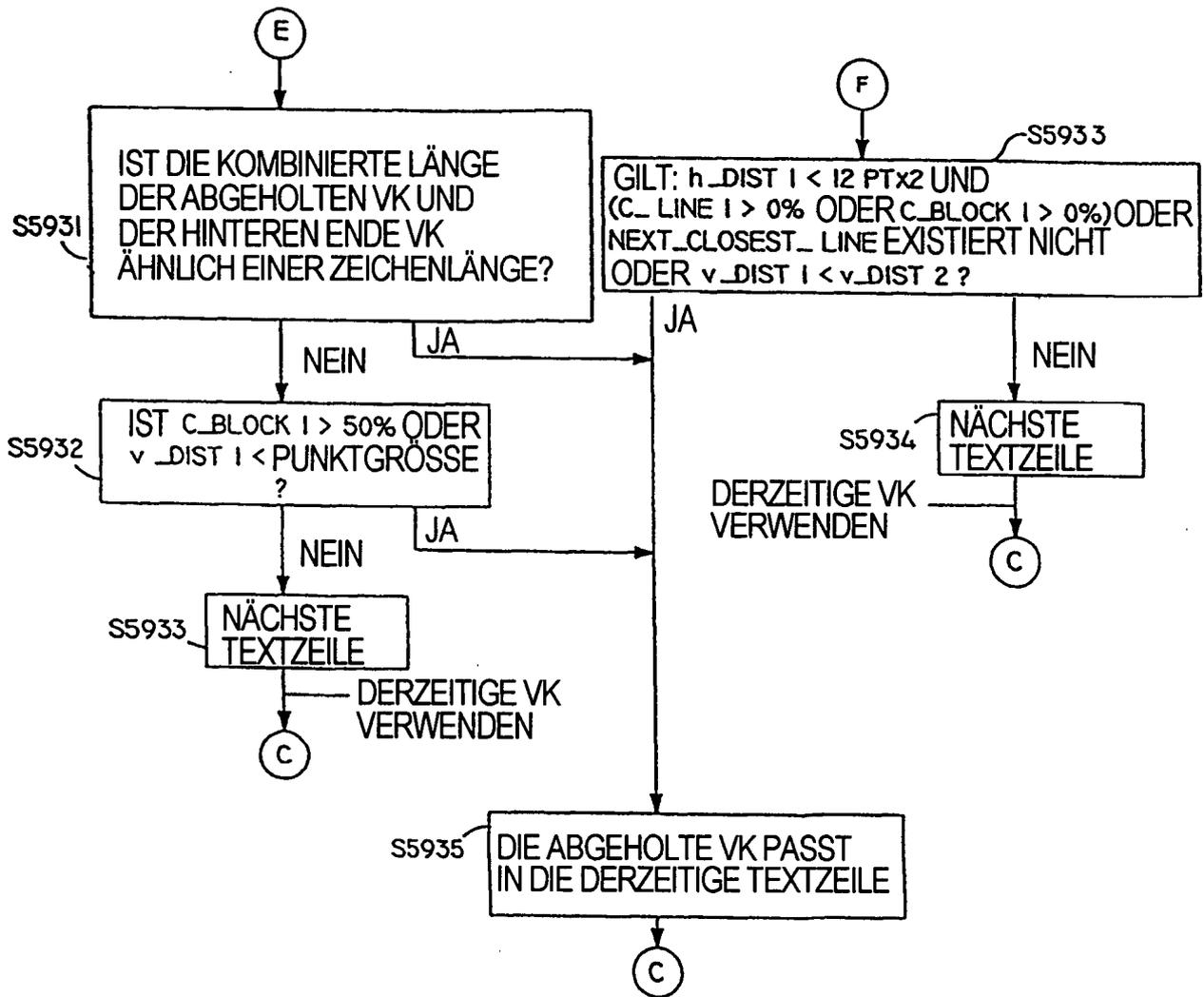


FIG. 16D

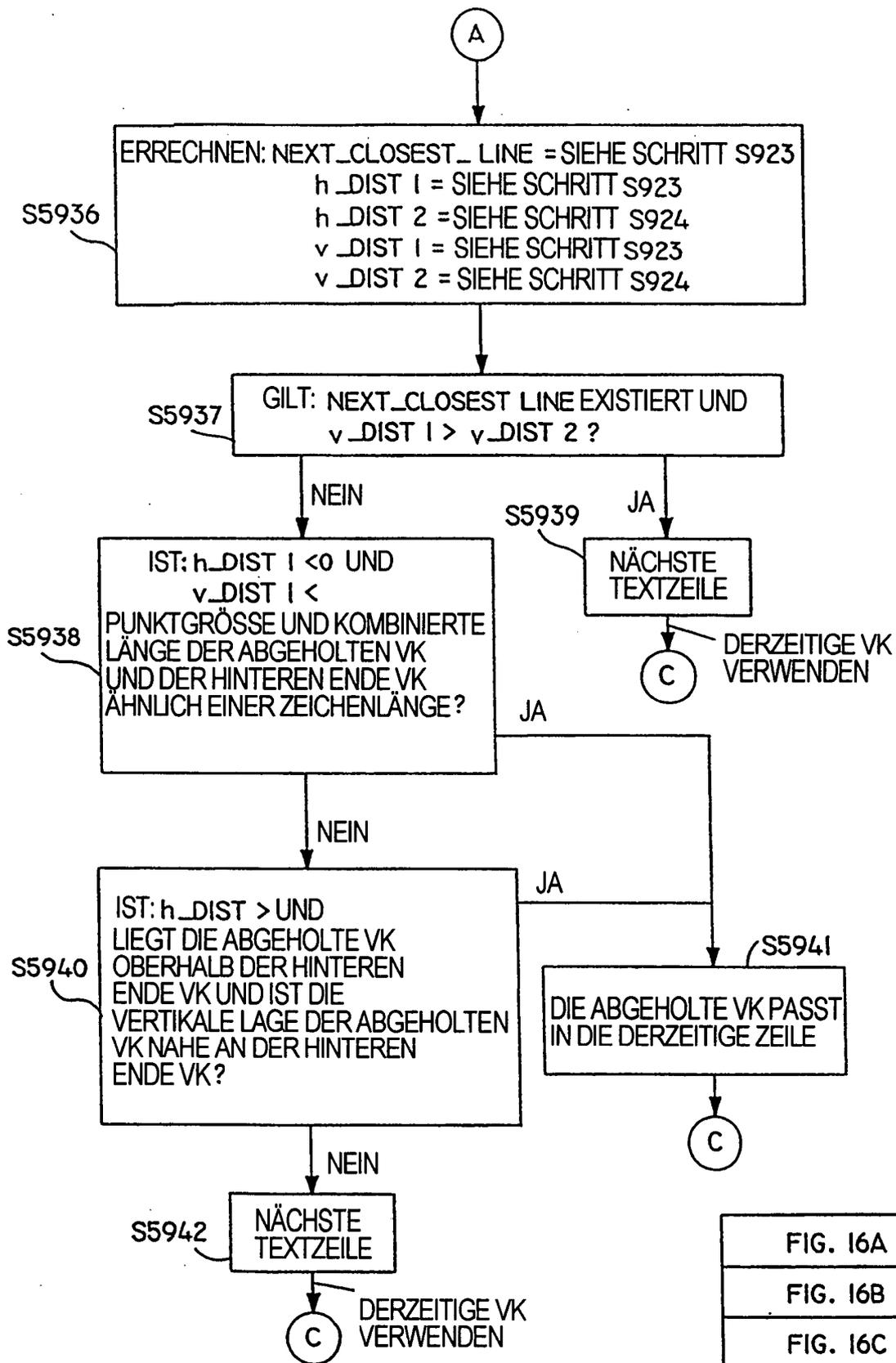


FIG. 16E

FIG. 16A
FIG. 16B
FIG. 16C
FIG. 16D
FIG. 16E

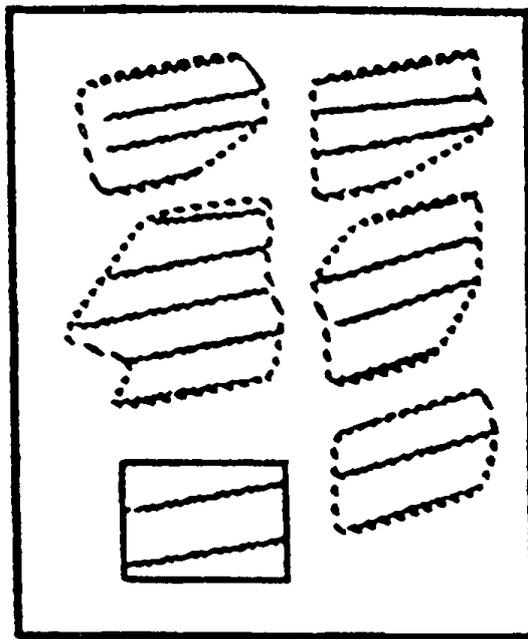


FIG.17