



(19)  
Bundesrepublik Deutschland  
Deutsches Patent- und Markenamt

(10) DE 698 28 909 T2 2006.04.06

(12)

## Übersetzung der europäischen Patentschrift

(97) EP 1 058 909 B1

(21) Deutsches Aktenzeichen: 698 28 909.9

(86) PCT-Aktenzeichen: PCT/US98/24420

(96) Europäisches Aktenzeichen: 98 957 993.3

(87) PCT-Veröffentlichungs-Nr.: WO 99/030270

(86) PCT-Anmeldetag: 16.11.1998

(87) Veröffentlichungstag

der PCT-Anmeldung: 17.06.1999

(97) Erstveröffentlichung durch das EPA: 13.12.2000

(97) Veröffentlichungstag

der Patenterteilung beim EPA: 02.02.2005

(47) Veröffentlichungstag im Patentblatt: 06.04.2006

(51) Int Cl.<sup>8</sup>: G06K 9/48 (2006.01)

G06K 9/40 (2006.01)

G06K 9/38 (2006.01)

G06T 5/00 (2006.01)

(30) Unionspriorität:

986461 08.12.1997 US

(84) Benannte Vertragsstaaten:

DE, GB, NL

(73) Patentinhaber:

Intel Corporation, Santa Clara, Calif., US

(72) Erfinder:

ACHARYA, Tinku, Tempe, US; TSAI, Ping-Sing, Chandler, US

(74) Vertreter:

Patentanwälte Hauck, Graafls, Wehnert, Döring, Siemons, Schildberg, 80339 München

(54) Bezeichnung: NEUE PERZEptive LOKALE GRADIENTENBAsierte GRATBESTIMMUNG

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 eingelebt, 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****STAND DER TECHNIK****1. Gebiet der Erfindung**

**[0001]** Die vorliegende Erfindung betrifft allgemein das Gebiet der Bildverarbeitung. Im Besonderen betrifft die vorliegende Erfindung das Detektieren von Merkmalen in Bildern.

**2. Beschreibung des Stands der Technik**

**[0002]** Auf dem Gebiet der Bildverarbeitung können Bilder eines Objekts/einer Szene durch bestimmte Merkmale bzw. Eigenschaften gekennzeichnet bzw. charakterisiert werden. Ein wichtiges derartiges Bildmerkmal ist das "Kantenmerkmal" eines Bilds. Eine Kante eines Bilds kann durch die Abruptheit der Intensität und/oder der Farbveränderung zwischen einem ersten Pixelbereich und einem zweiten Pixelbereich definiert werden. Entlang der Grenze zwischen zwei Bereichen kann eine Gruppe von Pixeln auf jeder Seite als "Kantenpixel" oder als einem Kantenmerkmal zugehörige Pixel klassifiziert werden. Als ein extremes Beispiel für ein Kantenmerkmal können drei Zeilen von Pixeln mit einer reinen blauen Farbe und zwei Zeilen von Pixeln in weißer Farbe gelten, die direkt auf die drei blauen Zeilen folgen. Die Pixel der dritten und vierten Zeilen der jeweils blauen bzw. weißen Farbe können als Kante angesehen werden. Da zwischen der dritten und der vierten Zeile ein abruper und großer Unterschied in Bezug auf die Farbe und die Intensität besteht, befinden sich diese Pixel an der "Kante" und stellen gemeinsam ein Kanten- bzw. Randmerkmal dar. Im Gegensatz dazu würde eine Zeile von Pixeln in der Farbe weiß und eine folgende, benachbarte Zeile von Pixeln in einem gebrochenen weiß weniger eindeutig als Kantenmerkmal betrachtet werden, da der Intensitätsunterschied nicht so groß und abrupt wäre, um als eine Kante klassifiziert zu werden. Das als Kantenerkennung bzw. Randdetektierung bekannte Gebiet der Technik sieht Techniken zum Detektieren und Unterscheiden von Kanten vor und liefert verschiedenen Datenformen und Statistiken in Bezug auf die Kantenbildung in einem Bild, die in Bildverarbeitungsanwendungen nützlich sind.

**[0003]** Die Klassifizierung eines Bilds in Kantenmerkmale weist viele wichtige Anwendungen auf dem Gebiet der Bildverarbeitung und des Rendering auf. Eine wichtige Anwendung ist die Scharfstellung des Bilds durch das Beseitigen von Unschärfen und die Verbesserung der Kantendarstellung. Eine weitere Anwendung ist die Form- oder Mustererkennung, die beim maschinellen Sehen oder der optischen Zeichenerkennung (OCR als englische Abkürzung von Optical Character Recognition) eingesetzt werden kann: Derartige Anwendungen sind stark von dem

Erfolg des Kantendetektionsprozesses abhängig. Wenn Kanten nicht detektiert werden können, können sie auch nicht verbessert werden, und wenn eine Grenze zwischen einer Nicht-Kante und einer Kante nicht richtig identifiziert wird, bleibt eine Unschärfe auch unscharf. Die Kantendetektierung kann zweifach fehlschlagen: es können "falsche" bzw. "fehlerhafte" Kanten erzeugt werden, das heißt Merkmale, die keine Kanten sind, werden als Kanten klassifiziert, oder Kanten können zerstört werden, das heißt Merkmale, die Kanten sind, werden als Nicht-Kantenmerkmale klassifiziert und so verarbeitet, dass die Kante dezimiert wird.

**[0004]** Das Problem der Kantendetektierung wurde mit verschiedenen Ansätzen behandelt. Ein Ansatz behandelt eine Kante als eine Veränderung des Intensitätswerts von einem Pixel oder einem Bereich zu dem nächsten. Ein Operator (eine bzw. mehrere mathematische Funktionen, die auf eine Gruppe von Daten angewandt werden), wie etwa ein Gradient-Operator, der diese Veränderung messen kann, kann somit bei der Kantendetektierung eingesetzt werden. Zusätzlich zu der Berechnung der Änderung(en) des Intensitätswerts unter Verwendung einer gewissen Form des Operators, muss der Kantendetektor auch einen Schwellenwert aufweisen, der anzeigt, ob eine bestimmte Intensitätswertveränderung so klassifiziert werden kann, dass sie einen Kanten- oder Nicht-Kanten-Bereich oder ein entsprechendes Pixel darstellt. Ein Vergleich zwischen dem Schwellenwert und der gemessenen Intensitätsveränderung ermittelt, ob ein Pixel zu einer Kante gehört oder nicht. Die meisten Kantendetektoren auf der Basis von Schwellenwerten versuchen einen vordefinierten Schwellenwert festzulegen, der für ein ganzes Bild anwendbar ist und ferner auf jedes Bild angewandt werden kann, das durch eine bestimmte bilddarstellende Vorrichtung oder einen entsprechenden Prozess erzeugt wird. Dies ist problematisch, da in einem Bild bestimmte Merkmale wie etwa ein Schatten den Kantendetektionsprozess beeinträchtigen können. Wenn zum Beispiel ein Teil eines Bilds im Schatten einer Lichtquelle liegt, während die anderen Bereiche des Bilds in dem helleren Bereich der Lichtquelle liegen, lassen sich die Intensitätsschwankungen in dem Schattenbereich nicht leicht vergleichen, um die Kanten zu detektieren wie dies in dem helleren Bereich der Fall sein kann. Der Schattenbereich offenbart eine geringfügigere Intensitätsschwankung, da sich die Intensitätswerte in einem niedrigeren und schmalen Bereich als in Bezug auf den helleren Bereich des Bilds für ein identisches Kantenmerkmal befinden. Wenn somit ein hoher Schwellenwert festgelegt wird, werden die in dem Schattenbereich des Bilds vorhandenen Kantenmerkmale falsch als Nicht-Kantenmerkmale klassifiziert.

**[0005]** Somit kann es zur Korrektur des Problems erforderlich sein, einen niedrigeren Schwellenwert

festzulegen. Häufig führt das Festlegen eines niedrigen Schwellenwerts jedoch zu dem entgegengesetzten Problem der Einführung fehlerhafter Kanten. So mit wurde der Schwellenwert häufig von Bild zu Bild manuell festgelegt, wobei ein Benutzer einen bestimmten Schwellenwert für ein bestimmtes Bild auswählt. Die Auswahl eines Schwellenwerts erfolgt häufig durch Ausprobieren, wobei ein Schwellenwert für jedes Bild getestet wird und die Ergebnisse später analysiert werden. Ferner ignorieren die meisten Kantendetektoren das Problem des lokalisierter Intensitätsbereichs, da der Schwellenwert überall in dem ganzen Bild angewandt wird. Eine der bekanntesten und effektivsten bislang entwickelten Kantendetektoren, der Canny-Kantendetektor, verwendet eine Hystereseschwellenwertbildung. Bei der Hystereseschwellenwertbildung werden ein hoher Schwellenwert und ein niedriger Schwellenwert festgelegt, und in einem ersten Durchgang werden alle Pixel mit einer Veränderung des Intensitätswertes (Gradienten), die größer ist als der hohe Schwellenwert, als Kantenpixel angesehen. In dem nächsten Durchlauf werden alle Pixel, die sich in der unmittelbaren Nähe der Kantenpixel befinden oder an diese angrenzen, die größer sind als der niedrige Schwellenwert, als Kantenpixel klassifiziert. Auch bei dem Canny-Detektor werden zurzeit die die Hysterese bildenden hohen und niedrigen Schwellenwerte beide manuell ausgewählt.

**[0006]** Somit wäre es wünschenswert, einen verbesserten Kantendetektor vorzusehen, der den Schwellenwert automatisch von Bild zu Bild ohne Benutzereingriffe auswählt. Wie dies vorstehend im Text erwähnt worden ist, wäre es ferner wünschenswert, potenzielle lokale Intensitätsmerkmale in einem Bild zu berücksichtigen, so dass der Schwellenwert lokal in Bereichen des gleichen Bilds angepasst werden kann. Da die Kantendetektierung aufgrund der Kosten für die Berechnung für gewöhnlich in Software implementiert ist, wäre es wünschenswert, eine Technik zu entwickeln, die sich für eine schnelle Hardware-Implementierung eignet, so dass eine digitale Kamera oder ein Roboterauge bzw. ein maschinelles Auge eine automatisierte Kantendetektierungsfunktion auf Hardwarebasis aufweist.

**[0007]** 'Automatic edge extraction using locally adaptive threshold', von Jung G. S. et al., Electronics Letters, IEE Stevenage, GB, Band 24, Nr. 11, 26. Mai 1988 (1988-05-26), Seiten 711–712, XP000002656 ISSN: 0013-5194, offenbart einen automatischen Kantenextraktionsalgorithmus, der einen lokalen adaptiven Schwellenwert verwendet, der Helligkeits- und Gradientenwerten in einem Bild unterliegt. Kanten mit einer Dicke von einem Pixel können extrahiert werden, so dass ein optisch ansprechendere Darstellung eines ursprünglichen Graustufenbilds resultiert.

**[0008]** 'Nonlinear gradient-based edge detection al-

gorithms in the telesign system', von Tian-Hu Yu, APCCAS 94. 1994 IEEE Asia-Pacific Conference on Circuits and Systems (Cat. No. 94TH8029), Proceedings of APCCAS 94–1994 Asia Pacific Conference on Circuits and Systems, Taipee, Taiwan, 5.–8. Dezember 1994, Seiten 661–666, XP002196389 1994, New York, NY, USA, IEEE, USA ISBN: 0-7803-2440-4, offenbart einen nichtlinearen Gradient zum Detektieren wahrnehmbarer Kanten in einem Bild, mit dem ein entsprechender Schwellenwert verhältnismäßig leicht bestimmt werden kann, und wobei eine Anwendung auf ein Telesign-System vorgesehen wird.

## ZUSAMMENFASSUNG DER ERFINDUNG

**[0009]** Vorgesehen ist gemäß einem ersten Aspekt der vorliegenden Erfindung ein Verfahren gemäß dem gegenständlichen Anspruch 1.

**[0010]** Vorgesehen ist gemäß einem zweiten Aspekt der vorliegenden Erfindung ein Computerprogramm gemäß dem gegenständlichen Anspruch 6.

## KURZE BESCHREIBUNG DER ZEICHNUNGEN

**[0011]** Die Aufgaben, Merkmale und Vorteile des erfindungsgemäßen Verfahrens und der erfindungsgemäßen Vorrichtung werden aus der folgenden Beschreibung deutlich. In den Zeichnungen zeigen:

**[0012]** [Fig. 1](#) ein Flussdiagramm eines Ausführungsbeispiels der Erfindung;

**[0013]** [Fig. 2](#) eine näherungsweise Darstellung des Ansprechverhaltens des menschlichen Sehvermögens an Änderungen der Intensität;

**[0014]** [Fig. 3](#) eine Kantendetektierung auf Gradient-Basis gemäß einem Ausführungsbeispiel der Erfindung;

**[0015]** [Fig. 4\(a\)](#) eine Maske zur Ermittlung eines ersten Richtungsdifferentials für die Gradient-Bestimmung;

**[0016]** [Fig. 4\(b\)](#) eine Maske zur Ermittlung eines ersten Richtungsdifferentials für die Gradient-Bestimmung;

**[0017]** [Fig. 5](#) ein Systemdiagramm eines Ausführungsbeispiels der Erfindung; und

**[0018]** [Fig. 6](#) ein Flussdiagramm der Modifikationen der Gradient-basierten Kantendetektierung aus [Fig. 3](#), die auf eine direkte Hardwareimplementierung anwendbar ist.

## GENAUE BESCHREIBUNG DER VORLIEGENDEN ERFINDUNG

**[0019]** In Bezug auf die Abbildungen werden nachstehend beispielhafte Ausführungsbeispiele der vorliegenden Erfindung beschrieben. Die Ausführungsbeispiele veranschaulichen die Aspekte der vorliegenden Erfindung und schränken den Umfang der Erfindung nicht ein. Die Ausführungsbeispiele werden hauptsächlich in Bezug auf Blockdiagramme oder Flussdiagramme beschrieben. In Bezug auf die Flussdiagramme stellt jeder Block in den Flussdiagrammen sowohl einen Verfahrensschritt als auch ein Vorrichtungselement zur Ausführung des Verfahrensschritts dar. Abhängig von der Implementierung kann das entsprechende Vorrichtungselement in Hardware, Software, Firmware oder Kombinationen dieser konfiguriert werden.

**[0020]** Wenn ein durch einen Sensor, eine Kamera oder eine andere bilddarstellende Vorrichtung oder anderweitig erfasstes Bild analysiert wird, kann man sagen, dass es Merkmale aufweist, die Kantenmerkmale sind, während alle anderen Merkmale als Nicht-Kantenmerkmale bezeichnet werden. Es fällt schwer zu bestimmen, was eine Kante ist, und es ist kein präzise vorgegebenes Verfahren. Die Kantendetektierung basiert häufig auf der Messung einer Intensitätsveränderung und dem folgenden Vergleich der Intensitätsveränderung mit einem Schwellenwert. Die Auswahl des Schwellenwerts ist ein fundamentales Problem der Kantendetektierung, da ein zu niedriger Schwellenwert fehlerhafte Kanten erzeugt, und wobei ein zu hoher Schwellenwert Kanten auslässt. Da kein einzelner Schwellenwert für jedes Bild richtig ist, muss der Schwellenwert von Bild zu Bild angepasst werden. Häufig erfolgt dies manuell durch einen Anwender. Wenn dies nicht der Fall ist, wird der Schwellenwert voreingestellt, was nicht wünschenswert ist, da bestimmte Bilder auf einen bestimmten Schwellenwert zum Zweck der Kantendetektierung nicht gut ansprechen können. Selbst wenn ein Schwellenwert bei der traditionellen Kantendetektierung in gewisser Weise Bild für Bild bestimmt wird, wird der Schwellenwert in dem ganzen Bild einheitlich angewandt. Bei Bildern mit unterschiedlichen Lichtquellen, die das Felde oder den Bereich der Intensität in einem Bereich eines Bilds verändern können, der sich von den anderen Bereichen unterscheidet, ist ein einheitlicher Schwellenwert nicht geeignet und führt zu ungenauen Kantendetektierungsergebnissen.

**[0021]** Eine präzise Kantendetektierung ist für die Bildmustererkennung, das maschinelle Sehen, Bildverbesserungen und viele andere Gebiete der Bildverarbeitung von Bedeutung. Wünschenswert ist es, auf Benutzereingriffe beim Festlegen eines Schwellenwerts verzichten zu können und einen Schwellenwert auf der Basis der Intensitätsmerkmale eines

Bilds zu bestimmen. Vorzugsweise sollten die Techniken auf einen lokalen Bereich angewandt werden, so dass Kanten mit geringer Intensitätsschwankung weiter präzise erkannt werden. Ferner sollte(n) sich die Technik(en) schnell in Hardware implementieren lassen, so dass sie bei Bedarf in einer Digitalkamera oder in einer bilddarstellenden Vorrichtung implementiert werden können, in der das Bildmuster erzeugt wird. Ein derartige mögliche Methode zur Kantendetektierung ist in der Abbildung aus [Fig. 1](#) dargestellt sowie in verschiedenen weiteren nachstehend beschriebenen Ausführungsbeispielen.

**[0022]** Die Abbildung aus [Fig. 1](#) zeigt ein Flussdiagramm eines Ausführungsbeispiels der vorliegenden Erfindung.

**[0023]** Zuerst wird ein Lokalisierungsbereich in dem Bild definiert (Schritt **110**). Bei dem Lokalisierungsbereich handelt es sich um einen rechteckigen Bereich  $M \times N$  des Bilds, wobei M und N identisch sein können, wodurch ein quadratischer Bereich definiert wird. Der Lokalisierungsbereich kann in Prozent der insgesamt erfassten Bildgröße ausgedrückt abhängig von der gewünschten Implementierung sowie der verfügbaren Hardware-/Software-Ressourcen und Geschwindigkeitsbeschränkungen in Bezug auf die gewünschte Bildverarbeitungsgeschwindigkeit groß oder klein ausfallen. Zur Verwendung von Hardware und Software kann es in den meisten Computersystemen nützlich sein, einen Lokalisierungsbereich von  $8 \times 8$  oder  $16 \times 16$  vorzusehen. Der Lokalisierungsbereich ist für gewöhnlich um ein bestimmtes Pixel in dem Bild definiert. Wenn das Bild über einen größeren Bereich einheitlichere Merkmale aufweist, kann der Lokalisierungsbereich groß definiert werden. Wenn jedoch eine feine Detailunterscheidung wichtig ist oder das Bild große Schwankungen in Bezug auf die Eigenschaften in kleinen Bereichen aufweist, so kann der Lokalisierungsbereich kleiner definiert werden.

**[0024]** Nachdem der Lokalisierungsbereich definiert worden ist, werden in dem nächsten Schritt drei Werte bestimmt: der Mittelwert bzw. Median, das Minimum (Min.) und das Maximum (Max.) aller Intensitätswerte in dem Lokalisierungsbereich (Schritt **120**). Bei einer Umgebung des Lokalisierungsbereichs in Form eines Quadrats mit  $M \times M$  sind  $M \cdot M - 1$  Vergleiche erforderlich. Das Minimum, Maximum und der Median der Intensitätswerte in einem Lokalisierungsbereich ermöglichen eine feinere Unterscheidung in einem Bild bei der Auswahl eines Schwellenwerts. Die Intensitätsmerkmale eines Lokalisierungsbereichs in dem Bild ermöglichen es, dass die Kantendetektierungsanwendung bestimmt, welche Bedeutung einer bestimmten Intensitätsänderung eines Pixels in dem Bereich zugemessen werden soll. Im Gegensatz dazu berücksichtigt die kennzeichnende Kantendetektierung die lokalisierte Intensitätsände-

rung nicht.

**[0025]** Nachdem die Gradienteninformationen bestimmt worden sind, sollte ein Schwellenwert bestimmt werden (Schritt 130). Der Schwellenwert wird mit einer Intensitätsänderung verglichen (normalerweise als Verlauf bzw. Gradient gemessen), und wenn die Intensitätsänderung über dem Schwellenwert liegt, kann das Pixel oder der Bereich als eine Kante klassifiziert werden. Bei dem traditionellen Kantendetektor ist der Schwellenwert bereits voreingestellt und variiert nicht von Bild zu Bild. Wenn eine Variation des Schwellenwerts von Bild zu Bild gewünscht wird, muss dieser häufig manuell durch einen Benutzer angepasst werden. Ferner versuchen die traditionellen Kantendetektoren auf der Basis von Schwellenwerten nicht, den Schwellenwert dynamisch in einem lokalen Bereich eines bestimmten Bilds anzupassen, vielmehr verwenden sie den gleichen Schwellenwert in einem ganzen gegebenen Bild, ob angepasst auf ein bestimmtes Bild oder nicht.

**[0026]** Gemäß dem Schritt 130 wird der Schwellenwert für die Kantendetektierung gemäß einem Ausführungsbeispiel der vorliegenden Erfindung im Gegensatz dazu automatisch bestimmt. Diese Bestimmung basiert auf dem Ansprechverhalten des menschlichen Sehens (Optik des menschlichen Auges und mentale Wahrnehmung durch das menschliche Gehirn) auf Intensitätsveränderungen. Die nachstehend im Text beschriebene Abbildung aus [Fig. 2](#) veranschaulicht grafisch eine näherungsweise Kurve des Ansprechverhaltens des menschlichen Sehens in Bezug auf Intensitätswechsel. Allgemein gilt, je größer der Bereich der Intensitätswerte, desto weniger ist das menschliche Sehen in der Lage, die Intensitätsänderungen zu erkennen. Wenn der Schwellenwert auf das Ansprechverhalten des menschlichen Sehvermögens angepasst ist, ist er deutlich präziser, und bei der Bildkantenverbesserung erscheint die optische Qualität eines derartigen verbesserten Bilds für den Menschen in besserer Qualität als das gleiche Bild ohne Verbesserung durch eine Schwellenwertbildung gemäß der "Wahrnehmung".

**[0027]** Die Wahrnehmungs-Schwellenwertbildung gemäß Schritt 130 umfasst ferner den Einsatz der Intensitätsmerkmale des Lokalisierungsbereichs, das heißt des Minimums, des Medians und des Maximums gemäß der Bestimmung in Schritt 120. Eine Möglichkeit zum Ausdrücken des relativen Intensitätsbereichs und der relativen Intensitätswerte in dem Lokalisierungsbereich ist die Definition eines prozentualen Anteils oder Quotienten, der die Gesamtverteilung der Intensitätswerte in dem Lokalisierungsbereich mit der Differenz zwischen Median und Minimum vergleicht. Somit kann eine "relative Lokalisierungsintensität" (RI) für die Schwellenwertbildung wie folgt definiert werden:

$$RI = (\text{Median} - \text{Min.}) / (\text{Max.} - \text{Min.})$$

**[0028]** Die RI kann für die Schwellenwertbildung verwendet werden, wobei der RI-Faktor jedoch dahingehend nicht optimal ist, dass er keine Korrelation zu dem menschlichen Sehvermögen vorsieht oder es wird eine Anwendung für die Kantendetektierung eingesetzt. Wie dies in der Abbildung aus [Fig. 2](#) dargestellt und auf dem Gebiet der Optik und Bilddarstellung allgemein bekannt ist, folgt das Ansprechverhalten des menschlichen Sehens auf Intensitätswechsel einem groben logarithmischen Muster. Zur Berücksichtigung der Wahrnehmungsfähigkeit des menschlichen Sehens sollte somit ein Leistungsfunktionsoperator auch einen Faktor für die Bestimmung des Schwellenwerts darstellen. Somit entspricht der Schwellenwert für die Kantendetektierung vorzugsweise  $RI^a$ , wobei "a" einen Leistungsfaktor darstellt. Simulationen und Testergebnisse haben gezeigt, dass ein Leistungsfaktor von 0,45 das menschliche Sehvermögen näherungsweise wiedergibt. Somit entspricht der Schwellenwert "T", der bei der Kantendetektierung für Zwecke des menschlichen Sehens eingesetzt wird:

$$T = \{(\text{Median} - \text{Min.}) / (\text{Max.} - \text{Min.})\}^{0,45}$$

**[0029]** Dieser Leistungsfunktionsoperator stellt lediglich ein Beispiel dar und kann in Übereinstimmung mit neuen Informationen modifiziert werden, die auf dem Gebiet des menschlichen Sehens bekannt werden, oder abhängig von der Anwendung, für welche die Kantendetektierung eingesetzt wird. Bei der Bildmustererkennung wie etwa OCR kann es zum Beispiel nützlich sein, von dem Ansprechverhalten des menschlichen Sehens abzuweichen und die Leistungsfunktion mit einem anderen Leistungsfaktor neu zu bilden oder den Operatortyp zu ändern, wie etwa auf einen linearen Operator.

**[0030]** Ein wichtiger Aspekt dieses Ausführungsbeispiels der vorliegenden Erfindung ist es, dass der Schwellenwert automatisch bestimmt und lokal dynamisch und ohne Benutzereingriffe angepasst wird. Die Schwellenwertbildung kann auch zusätzlich auf der Wahrnehmung des menschlichen Sehens basieren, wobei sie aber auch auf Wunsch für andere Sehsysteme oder Anwendungen modifiziert werden kann, welche die ausgeführte Kantendetektierung verwenden.

**[0031]** Nachdem ein Schwellenwert bestimmt worden ist, kann die Kantendetektierung unter Verwendung des lokalisierten Schwellenwerts ausgeführt werden (Schritt 140). Die Abbildung aus [Fig. 3](#) veranschaulicht einen Kantendetektor auf Gradientenbasis, der die Wahrnehmungs- und lokal angepasste Schwellenwertbildung nutzt. Die Kantendetektierung kann nach Bedarf Pixel für Pixel oder Abschnitt für Abschnitt ausgeführt werden. Anders ausgedrückt

können alle Pixel in dem Lokalisierungsbereich dem gleichen dynamisch ermittelten Schwellenwert unterliegen oder nur das zentrale Pixel, um das der Lokalisierungsbereich definiert ist. Demgemäß wird in dem Schritt **150** geprüft, ob Bildabschnitte zur Kantendetektierung verbleiben. Ein Bildabschnitt kann so klein wie ein einzelnes Pixel definiert sein, so dass der dynamisch ermittelte Schwellenwert für jedes einzelne Pixel erneut bestimmt wird, und wobei der Lokalisierungsbereich für jedes einzelne Pixel verschoben wird (und somit neu bestimmt wird). Wenn etwaige Bildabschnitte verbleiben, so wird der Lokalisierungsbereich erneut definiert (Schritt **110**), und die Schritte **120–150** werden wiederholt. Die Prozedur ist abgeschlossen, wenn keine Bildabschnitte für eine Kantendetektierung mehr verbleiben.

**[0032]** Die Abbildung aus [Fig. 2](#) zeigt eine Annäherung an das Ansprechverhalten des menschlichen Sehens auf Intensitätswechsel.

**[0033]** Der Graph aus [Fig. 2](#) weist eine horizontale Achse auf, die mögliche Intensitätswerte für ein Pixel anzeigt. Bei einem Bild mit einer Auflösung von 8 Bit je Pixel liegt der mögliche Bereich der Intensitätswerte zwischen 0 und 255. Die vertikale Achse stellt die wahrgenommene Intensitätsänderung durch das menschliche Sehen dar, wobei ein großteils nichtlineares Ansprechverhalten mit offenbar logarithmischer Funktion dargestellt ist. Im Wesentlichen gilt, je höher der Intensitätswert, dessen Änderung gemessen wird, desto weniger ist das menschliche Sehen in der Lage, die Unterschiede zu unterscheiden. Zum Beispiel wird eine Veränderung der Intensität von einem Wert von 25 auf einen Wert von 26 als eine Intensitätsveränderung von 4% durch das menschliche Sehen wahrgenommen. Im Gegensatz dazu zeigt die gleiche Änderung eines Intensitätswerts von 100 auf 101 lediglich eine Wahrnehmungsdifferenz von 1% des menschlichen Sehens. Die Wahrnehmungsfähigkeit nimmt weiter ab, so dass eine Veränderung der Werte von 200 auf 201 lediglich als eine Intensitätsänderung von 0,5% wahrgenommen wird. Der Luminanzwert einer Änderung von 201 auf 200 ist nicht zu groß wie bei einer Änderung von 25 auf 26.

**[0034]** Angenommen wird somit ein Abschnitt eines Bilds, der durch einen Schatten alle Werte im unteren Bereich aufweist, wie zum Beispiel zwischen 0 und 50, während der Rest des Bilds Werte in hohen Bereichen aufweist. Wenn auf das Bild ein einheitlicher Schwellenwert angewendet wird, wie etwa mit  $T = 0,6$ , der häufig im Fach verwendet wird, so könnten in dem Bildabschnitt keine Kantenmerkmale detektiert werden, da die Änderung nie größer wäre als 60% des verhältnismäßig hohen Maximums. Alle Kantenmerkmale in dem Bildabschnitt würden falsch gekennzeichnet werden. Im Gegensatz dazu würde eine lokalisierte Wahrnehmungs-Schwellenwertbildung den Schwellenwert lokal gemäß der maxima-

len, minimalen und mittleren Intensität in dem Bildabschnitt oder dem Bereich des Bildabschnitts, der als Lokalisierungsbereich definiert ist, festlegen. Bei einem Maximum von 50, einem Minimum von 0 und einem Median von 20 wäre der angewandte Schwellenwert gleich  $20/50$  bzw.  $0,40^{45}$ . Wenn die Intensitätsveränderung bzw. der Gradient ferner zu dem normalisiert oder auf den lokalen Bereich abgestimmt wird, ist der Vergleich für die Kantendetektierung aussagekräftiger, wie dies dargestellt und nachstehend im Text beschrieben ist.

**[0035]** Die Kurve aus [Fig. 2](#) ist im Fach allgemein bekannt, wobei der Fachmann Werte zwischen 0,4 und 0,5 unterschiedlich zur Darstellung der die Kurve beschreibenden Leistungsfunktion verwendet hat. Vorzugsweise verwendet ein Ausführungsbeispiel der vorliegenden Erfindung einen Wert von 0,45 für den Leistungsfaktor für den Leistungsfunktionsoperator, wobei jedoch jeder Leistungsfaktor oder jeder andere geeignete Operator für die Anwendung oder Implementierung zur Anwendung auf den RI-Faktor ausgewählt werden kann.

**[0036]** Die Abbildung aus [Fig. 3](#) zeigt ein Flussdiagramm einer Gradienten-basierenden Kantendetektierung gemäß einem Ausführungsbeispiel der vorliegenden Erfindung.

**[0037]** Zuerst wird ein anfänglicher Lokalisierungsbereich in dem Bild um ein Pixel P definiert (Schritt **310**). Bei dem anfänglichen Lokalisierungsbereich kann es sich um einen rechteckigen Bereich von  $M \times N$  des Bilds handeln, wobei M und N auf Wunsch identisch sein können, um dadurch einen quadratischen Bereich zu definieren. Wenn das Bild über einen größeren Bereich einheitlichere Eigenschaften aufweist, kann der Lokalisierungsbereich als groß definiert werden. Wenn hingegen eine feine Detailunterscheidung wichtig ist oder das Bild in kleinen Bereichen große Eigenschaftsschwankungen aufweist, kann der Lokalisierungsbereich kleiner definiert werden.

**[0038]** Sobald der Lokalisierungsbereich definiert ist, wird in dem nächsten Schritt der Gradientenwert bestimmt, der jedem einzelnen Pixel in dem Lokalisierungsbereich zugeordnet ist, indem eine Maske oder ein Operator (Schritt **320**) angewandt wird. Die Maske oder der Gradienten-Operator wird auf eine kleine Nachbarschaft bzw. Umgebung um jedes Pixel angewandt, wobei diese Umgebung normalerweise kleiner ist als der Lokalisierungsbereich. Die Abbildungen der [Fig. 4\(a\)](#) bis [4\(b\)](#) und die zugehörige Beschreibung veranschaulichen die Parameter und die Konfiguration einer solchen Maske mit einer Umgebung von  $3 \times 3$  um und einschließlich des Pixels, dessen Gradient bestimmt wird. Durch wiederholte Anwendung des Operators oder der Maske auf jedes Pixel in dem Lokalisierungsbereich und die Bestim-

mung des Gradienten dadurch können die Gradienten-Informationen für den ganzen Lokalisierungsbereich abgeschlossen werden. Der Gradient bzw. der Verlauf misst die relative Intensitätsveränderung in verschiedene Richtungen von einem bestimmten Pixel in einer kleinen Umgebung (der Maskengröße) um dieses Pixel.

**[0039]** Bei dem Großteil der Kantendetektierung auf Gradientenbasis wird der Gradientenwert über das ganze Bild normalisiert, indem der Gradient durch den maximalen Gradient in dem ganzen Bild geteilt wird. Wenn eine derartige Normalisierung jedoch in Hardware implementiert wird, muss ein Ansatz mit zwei Durchläufen bestimmt werden, bei dem die Gradienteninformationen für alle Pixel in dem ganzen Bild bestimmt werden müssen, und wobei danach jeder Gradient durch den in dem Bild festgestellten maximalen Gradientenwert dividiert wird. Da ein Lokalisierungsbereich für den Zweck der Bestimmung eines lokalen Schwellenwerts definiert wird, sollte der gleiche Lokalisierungsbereich zur Normalisierung des Gradienten eines Pixels P gemäß dem maximalen Gradient in dem Lokalisierungsbereich verwendet werden (Schritt 325). Dies gewährleistet, dass der Vergleich zwischen Schwellenwert und Gradient die lokalisierten Bildtendenzen besser reflektiert.

**[0040]** Wenn der Gradient gemäß dem maximalen Gradient in dem ganzen Bild normalisiert würde, würde der lokalisierte Schwellenwert mit einem Gradienten im Verhältnis zu dem ganzen Bild und nicht dem betreffenden Lokalisierungsbereich verglichen.

**[0041]** Nach der Bestimmung der Gradienteninformationen kann der Schwellenwert automatisch auf der Basis lokaler Intensitätsinformationen bestimmt werden (Schritt 330). Die automatische Bestimmung des Schwellenwerts "T" wurde vorstehend in Bezug auf den Schritt 130 aus [Fig. 1](#) detailliert beschrieben, wobei er zur Wiederholung wie folgt ausgedrückt werden kann:  $T = \{(Median - Min.)/(Max. - Min.)\}^a$ , wobei "Median" den Mittelwert aller Intensitätswerte in dem Lokalisierungsbereich bezeichnet, wobei "Min." dem Minimum aller Intensitätswerte entspricht, wobei "Max." dem Maximum der Intensitätswerte in dem Bereich entspricht, und wobei "a" einen Leistungsfaktor auf der Basis der Anwendung oder des Sehsystems bzw. optischen Systems darstellt, auf die bzw. das die Kantendetektierung angewandt wird. Wenn die Anwendung oder das Sehsystem ein flaches lineares Ansprechverhalten auf Intensitätsveränderungen aufweist, ist der Leistungsfaktor gleich "1", wobei zumeist, bei der Charakterisierung des menschlichen Sehens, im Fach ein Leistungsfaktor von ungefähr "0,45" offenbart wurde.

**[0042]** Zur Bestimmung, ob das Pixel P ein Pixel darstellt, das einem Kantenmerkmal des Bilds zugehörig ist oder nicht, wird der normalisierte Gradient

mit dem dynamisch bestimmten lokalen Schwellenwert verglichen (Schritt 335). Wenn der normalisierte Gradient für das Pixel P den Schwellenwert überschreitet, kann das Pixel P als ein "Kantenpixel" klassifiziert werden (Schritt 340), bei dem es sich um ein Pixel handelt, das einem Kantenmerkmal des Bilds zugehörig ist, wie etwa einer Zeile bzw. einer Linie. Wenn dies nicht der Fall ist, ist das Pixel P als ein Nicht-Kantenpixel klassifiziert (Schritt 350). Zwar zeigt das Flussdiagramm aus [Fig. 3](#) ein Verfahren der Kantendetektierung, das heißt die Gradienten-basierte Kantendetektierung, wobei jedoch jeder geeignete Kantendetektor eingesetzt werden kann, der eine bestimmte Art der Schwellenwertbildung verwendet, wobei das lokalisierte Wahrnehmungs-Schwellenwertbildungskonzept verschiedener Ausführungsbeispiele der Erfindung zur Verbesserung der Fähigkeit des Kantendetektors verwendet werden kann. Zum Beispiel kann der allgemein bekannte Canny-Kantendetektor, der gefaltete Gaußsche Masken mit dem Bildraum bei der Klassifizierung von Kanten- und Nicht-Kantenpixeln verwendet, in Verbindung mit der Technik der lokalisierten und dynamischen Wahrnehmungs-Schwellenwertbildung zur Einstellung für die Hysterese verwendet werden, sowohl eines unteren als auch eines oberen Schwellenwerts auf der Basis lokaler Intensitätsinformationen. Für eine digitale Bilddarstellungsvorrichtung, wie etwa eine tragbare digitale Fotokamera, ist ein Kantendetektor wie etwa gemäß der Darstellung in den Abbildungen der [Fig. 3](#), 4(a) und 4(b) aufgrund der Rechengeschwindigkeit und dem Einsatz von weniger Siliziumfläche in der Vorrichtung effektiv.

**[0043]** Wenn in dem Lokalisierungsbereich (geprüft in dem Schritt 370) zu behandelnde Pixel verbleiben, werden als nächstes die Schritte 320–375 wiederholt. Wenn als Kanten- oder Nicht-Kantenpixel zu klassifizierende Pixel verbleiben, kann ein neuer Lokalisierungsbereich um das nächste Pixel P definiert werden (Schritt 380). Der Lokalisierungsbereich unterscheidet sich für jedes Pixel in dem Bild, und somit weist jedes Pixel einen eigenen angepassten lokalisierten Schwellenwert auf und ein lokales Gradienten-Maximum, um das normalisiert wird. Das Verfahren zur Bestimmung von Gradienteninformationen für ein bestimmtes Pixel ist in den Abbildungen der [Fig. 4\(a\)](#) und [4\(b\)](#) veranschaulicht.

**[0044]** Die Abbildung aus [Fig. 4\(a\)](#) veranschaulicht eine Maske zur Ermittlung eines ersten Richtungsdifferentials.

**[0045]** Eine Möglichkeit zur Bestimmung des Gradienten für ein Pixel ist die Anwendung einer Maske A, die im Wesentlichen die "Stärke" des Pixels um eine benachbarte Begrenzung ermittelt, welche die Maske abdeckt. Diese Stärke ist im Wesentlichen ein Differential in eine bestimmte Richtung entlang zweier Grenzen eines Pixels. Die erste Richtung liefert in

Kombination mit einer ähnlichen Berechnung der Stärke entlang einer zweiten Richtungsbegrenzung eine gute Annäherung an die relative Stärke des zentralen Pixels in Bezug auf dessen Nachbarn. Die Abbildung aus **Fig. 4(a)** zeigt eine  $3 \times 3$ -Maske, die in den ersten und dritten Spalten Koeffizienten von ungleich Null aufweist und Nullkoeffizienten in der mittleren bzw. der zweiten Spalte. Die Maske A ist mit sechs Koeffizienten von ungleich Null A<sub>1</sub>, A<sub>2</sub>, A<sub>3</sub>, A<sub>4</sub>, A<sub>5</sub> und A<sub>6</sub> dargestellt. Wenn die Maske auf einen Bildbereich I mit Pixeln mit den Intensitätswerten I<sub>0</sub>, I<sub>1</sub>, I<sub>2</sub>, I<sub>3</sub>, I<sub>4</sub>, I<sub>5</sub>, I<sub>6</sub>, I<sub>7</sub> und I<sub>8</sub> angewandt wird, ist das resultierende Differential  $\Delta x$  wie folgt ausgedrückt:

$$\Delta x = A_1 \cdot I_0 + A_2 \cdot I_3 + A_3 \cdot I_6 + A_4 \cdot I_1 + A_5 \cdot I_4 + A_6 \cdot I_7.$$

**[0046]** Das Differential  $\Delta x$  kann bei der Berechnung des Gradienten des Pixels mit dem Intensitätswert I<sub>4</sub> verwendet werden, der sich in der Abbildung in der Mitte des Bildbereichs I befindet. Normalerweise weist die Maske A Koeffizienten auf, die summiert Null ergeben. Bei der allgemein bekannten Sobel-Maske weisen die Koeffizienten A<sub>1</sub>, A<sub>2</sub> und A<sub>3</sub> alle Werte von -1 auf, während die Koeffizienten A<sub>4</sub>, A<sub>5</sub> und A<sub>6</sub> Werte von +1 aufweisen. Bei Koeffizienten gemäß der Sobel-Maske entspricht das Differential  $\Delta x = I_2 + I_5 + I_8 - I_0 - I_3 - I_6$ . Die Auswahl der Maskenkoeffizienten, die, wie dies im Fach bekannt ist, einen aussagekräftigen Gradientenwert erreicht, vereinfacht somit die differentielle Berechnung auf eine reine Addition und Subtraktion von Pixelintensitätswerten, die leicht durch Addierer in Hardware implementiert werden können. Das Differential lässt sich somit leicht und schnell implementieren. Eine andere angewandte Maske, wie etwa die Prewitt-Maske, lässt sich leicht implementieren. Die Prewitt-Maske weist die Koeffizienten A<sub>1</sub> und A<sub>3</sub> mit -1 und einen Koeffizienten A<sub>2</sub> von -2 auf. Zur Spiegelung weist die Prewitt-Maske Koeffizienten an der anderen direktionalen Begrenzung von A<sub>4</sub> und A<sub>6</sub> gleich +1 und einen Koeffizienten A<sub>5</sub> von +2 auf.

**[0047]** Unter Verwendung der Prewitt-Maske wird  $\Delta x$  oder der Differentialwert zu  $\Delta x = I_2 + 2 \cdot I_5 + I_8 - I_0 - 2 \cdot I_3 - I_6$ . Zwar wird eine Multiplikation von 2 eingeführt, wobei diese jedoch leicht durch eine Schiebeoperation von I<sub>5</sub> und I<sub>3</sub> der durch Aufteilung der Berechnung von  $\Delta x$  zu  $\Delta x = I_2 + I_5 + I_5 + I_8 - I_0 - I_3 - I_3 - I_6$  implementiert werden kann, wenn nur Additions- und Subtraktionsoperationen in der Hardware implementiert sind.

**[0048]** Die Gradientenoperation wird zur Kantendetektierung verwendet, da angenommen wird, dass eine Kante eine bestimmte Richtung in dem Bildraum aufweist. Die Pixel entlang der Seitenbegrenzung (eine Seite I<sub>0</sub>, I<sub>3</sub>, I<sub>6</sub> und die andere Seite I<sub>2</sub>, I<sub>5</sub>, I<sub>8</sub>) des Pixels sollten eine Stärke oder einen relativen kombinierten Intensitätswert von weniger als das Pixel I<sub>4</sub> selbst aufweisen, wenn I<sub>4</sub> als Kantenpixel gilt. Diesbezüglich kann ein erstes Richtungsdifferential im

Wesentlichen in die horizontale Richtung (die Maske weist Werte auf, welche die Stärke auf den linken und rechten Seiten des Bilds bestimmen) gemäß den vorstehenden Ausführungen berechnet werden. Da der Bildraum in zwei Dimensionen verläuft, vertikal und horizontal, muss nur ein weiteres Richtungsdifferential um das Pixel I<sub>4</sub> berechnet werden.

**[0049]** Die Abbildung aus **Fig. 4(b)** veranschaulicht eine Maske zur Ermittlung eines zweiten Richtungsdifferentials. Die Abbildung aus **Fig. 4(b)** zeigt eine zweite  $3 \times 3$ -Maske, die in den ersten und dritten Spalten Koeffizienten von ungleich Null aufweist und Nullkoeffizienten in der mittleren bzw. der zweiten Spalte. Die Maske B ist mit sechs Koeffizienten von ungleich Null B<sub>1</sub>, B<sub>2</sub>, B<sub>3</sub>, B<sub>4</sub>, B<sub>5</sub> und B<sub>6</sub> dargestellt. Wenn die Maske auf einen Bildbereich I mit Pixeln mit den Intensitätswerten I<sub>0</sub>, I<sub>1</sub>, I<sub>2</sub>, I<sub>3</sub>, I<sub>4</sub>, I<sub>5</sub>, I<sub>6</sub>, I<sub>7</sub> und I<sub>8</sub> angewandt wird, ist das resultierende Differential  $\Delta y$  wie folgt ausgedrückt:

$$\Delta y = B_1 \cdot I_0 + B_2 \cdot I_3 + B_3 \cdot I_2 + B_4 \cdot I_1 + B_5 \cdot I_4 + B_6 \cdot I_7.$$

Das Differential  $\Delta y$  kann bei der Berechnung des Gradienten des Pixels mit dem Intensitätswert I<sub>4</sub> verwendet werden, der sich in der Abbildung in der Mitte des Bildbereichs I befindet. Normalerweise weist die Maske B Koeffizienten auf, die summiert Null ergeben. Bei der allgemein bekannten Sobel-Maske weisen die Koeffizienten B<sub>1</sub>, B<sub>2</sub> und B<sub>3</sub> alle Werte von -1 auf, während die Koeffizienten B<sub>4</sub>, B<sub>5</sub> und B<sub>6</sub> Werte von +1 aufweisen. Bei Koeffizienten gemäß der Sobel-Maske entspricht das Differential  $\Delta y = I_6 + I_7 + I_8 - I_0 - I_1 - I_2$ . Die Auswahl der Maskenkoeffizienten, die, wie dies im Fach bekannt ist, einen aussagekräftigen Gradientenwert erreicht, vereinfacht somit die differentielle Berechnung auf eine reine Addition und Subtraktion von Pixelintensitätswerten, die leicht durch Addierer in Hardware implementiert werden können. Das Differential lässt sich somit leicht und schnell implementieren. Eine andere angewandte Maske, wie etwa die Prewitt-Maske, lässt sich leicht implementieren. Die Prewitt-Maske weist die Koeffizienten B<sub>1</sub> und B<sub>3</sub> mit -1 und einen Koeffizienten B<sub>2</sub> von -2 auf. Zur Spiegelung weist die Prewitt-Maske Koeffizienten an der anderen direktionalen Begrenzung von B<sub>4</sub> und B<sub>6</sub> gleich +1 und einen Koeffizienten B<sub>5</sub> von +2 auf. Unter Verwendung der Prewitt-Maske wird  $\Delta y$  oder der Differentialwert zu  $\Delta y = I_6 + 2 \cdot I_7 + I_8 - I_0 - 2 \cdot I_1 - I_2$ . Zwar wird eine Multiplikation von 2 eingeführt, wobei diese jedoch leicht durch eine Schiebeoperation von I<sub>7</sub> und I<sub>1</sub> der durch Aufteilung der Berechnung von  $\Delta y$  zu  $\Delta y = I_6 + I_7 + I_7 + I_8 - I_0 - I_1 - I_1 - I_2$  implementiert werden kann, wenn nur Additions- und Subtraktionsoperationen in der Hardware implementiert sind. Vorzugsweise sollte die Maske A, die zur Ermittlung von  $\Delta x$  oder dem ersten Richtungsdifferential verwendet wird, den gleichen Typ wie die Maske B aufweisen, die zur Ermittlung des zweiten Richtungsdifferentials verwendet wird. Die Verwendung einer Sobel-Maske in eine Richtung und einer Prewitt-Maske in eine andere Richtung kann zur Kompensation wünschens-

wert sein, da die bilddarstellende Vorrichtung unter Umständen keine quadratischen Pixel sondern rechteckige Pixel oder Pixel in einer anderen Form ergibt. In diesem Fall kann die Geometrie eines Pixels bei der Auswahl der Maskenkoeffizienten berücksichtigt werden.

**[0050]** Bei einer Ermittlung des zweiten Richtungs-differentials  $\Delta y$  gemäß der vorstehenden Beschreibung werden beide direktonalen Begrenzungspaire horizontal und vertikal berücksichtigt. Danach entspricht der Gradient von I4, das zentrale Pixel in dem Bildraum I der Größe des durch einen Vektor erhaltenen Vektors, der die Summe der in jede Richtung erhaltenen Vektoren ist. Der Gradient G entspricht somit einfach  $\sqrt{\Delta x^2 + \Delta y^2}$  oder der Größe des Vektors, der durch Kombination von  $\Delta x$  und  $\Delta y$  gebildet wird. Der Gradient jedes Pixels kann somit durch Berücksichtigung einer Umgebung, vorzugsweise einer kleinen Umgebung von  $3 \times 3$ , welche jedes Pixel umgibt, und folgendes Anwenden einer Maske darauf in jede von zwei Richtungen und der entsprechenden Grenzpaare erhalten werden, wodurch im Wesentlichen jedes unmittelbar benachbarte bzw. angrenzende Pixel bei der Gradientenbestimmung einbezogen wird. Wie dies vorstehend im Text erwähnt worden ist, kann der Gradientenwert für den maximalen vorhandenen Gradient n dem Lokalisierungsbereich normalisiert werden, so dass relative Gewichtungen für das Gradientenmaß vorgesehen werden. In einem Bildabschnitt mit mehr Textur an Stelle von harten Kanten wie etwa Linien können die Gradientenwerte zum Beispiel in einem schmalen Bereich liegen. Wenn diese auf den maximalen Gradient in dem lokalen Bereich normalisiert werden, können die erhaltenen Gradientenwerte effektiver analysiert werden. Der normalisierte Gradient  $G = \sqrt{\Delta x^2 + \Delta y^2}/G_{LMAX}$  mit  $G_{LMAX}$  ist der maximale Gradient aller Pixel in dem Lokalisierungsbereich.

**[0051]** Die Abbildung aus [Fig. 5](#) zeigt ein Systemdiagramm eines Ausführungsbeispiels der vorliegenden Erfindung.

**[0052]** Die Abbildung aus [Fig. 5](#) veranschaulicht ein Computersystem **510**, bei dem es sich um eine allgemeine oder spezielle Bestimmungs- oder Datenverarbeitungsvorrichtung wie etwa einen PC (Personalcomputer) handeln kann, die bzw. der mit einer Kamera **530** verbunden werden kann. Bei der Kamera **530** kann es sich um eine Digitalkamera, eine digitale Videokamera oder eine andere Bilderfassungsvorrichtung oder ein entsprechendes Bilddarstellungs-system handeln, und wobei sie zur Erfassung eines Sensorbilds eines Objekts oder einer Szene **540** verwendet wird. Die erfassten Bilder werden im Wesentlichen durch eine Bildverarbeitungsschaltung **532** komprimiert und verarbeitet, so dass sie effizient in einer Bildspeichereinheit **534** gespeichert werden können, bei der es sich um einen RAM oder eine an-

dere Speichervorrichtung wie etwa eine Festplatte, eine Speicherplatte, etc. handeln kann. In den meisten Kameras werden die von dem Bilderfassungsmechanismus wie etwa einem Sensor erhaltenen unbehandelten Bilder zuerst in der Kamera **530** gespeichert und danach zur Ausgabe, Anzeige oder Bearbeitung heruntergeladen. Dies ermöglicht der Kamera **530** das schnelle Erfassen des nächsten Objekts ohne Verzögerung. Eine praktische Darstellung für unbehandelte Bilder ist das vorstehend beschriebene 8-Bit-Bayer-Muster CFA, das jedem Graustufenintensitätswert eines Pixels eine Farbe (rot, grün oder blau) zuordnet.

**[0053]** In einem Ausführungsbeispiel wird ein Bild durch die Kamera **530** als unbehandeltes 8-Bit-Bayer-Muster CFA erfasst und danach in ein anderes Daten- oder codiertes Format komprimiert. Die lokalisierte Schwellenwertbildung für die Kantendetektierung gemäß einem Ausführungsbeispiel der vorliegenden Erfindung funktioniert wie folgt: Zuerst wird das komprimierte Bild aus der Kamera **530** über eine E/A-Schnittstelle **517** auf einen Plattspeicher **518** oder eine andere Vorrichtung in dem Computersystem **510** heruntergeladen. Das komprimierte Bild wird danach in dem Computersystem **510** dekomprimiert. Danach führt das Computersystem die lokalisierte Schwellenwertbildung für die Kantendetektierung an dem wiederhergestellten CFA-Bild durch Ausführung von Anweisungen durch, welche den Prozessor **512** diesbezüglich konfigurieren, wie dies in verschiedenen Ausführungsbeispielen der Erfindung beschrieben ist.

**[0054]** Die in den verschiedenen Ausführungsbeispielen der Erfindung beschriebene Methodik kann unter Verwendung eines Prozessors **512** wie etwa eines Prozessors Pentium<sup>TM</sup> (ein Produkt der Intel Corporation) und eines Speichers **511** wie etwa eines RAM-Speichers ausgeführt werden, der zum Speichern/Laden von Anweisungen, Adressen und resultierenden Daten verwendet wird. Die verwendeten Anwendungen zur Ausführung der Wahrnehmungs-Schwellenwertbildung und der Kantendetektierung können ein aus einer Quelle kompliertes ausführbares Modul darstellen, geschrieben in einer Sprache wie etwa C++. Die Anweisungen bzw. Befehle des ausführbaren Moduls, die Anweisungen entsprechen, welche die Bestimmung und Normalisierung von Gradientenwerten, das Ermitteln lokaler Intensitätsmerkmale und die automatische Bestimmung eines lokalisierter Schwellenwerts unterstützen, können auf einem Plattspeicher **518** oder in einem Speicher **511** gespeichert werden, und wobei sie somit über ein bestimmtes computerlesbares Medium geladen werden können. Für den Durchschnittsfachmann auf dem Gebiet der Informatik ist die Programmierung einer Bestimmungsvorrichtung zur Ausführung der Interpolationsmethodologie einfach, wie dies in verschiedenen Ausführungsbeispielen

len der Erfindung beschrieben ist.

**[0055]** Das Computersystem **510** weist einen Systembus **513** auf, der die Informationsübertragung zu/von dem Prozessor **512** und dem Speicher **511** erleichtert, und ferner mit einer Brücke **514**, die den Systembus **513** mit einem E/A-Bus **515** koppelt. Der E/A-Bus **515** verbindet verschiedene E/A-Vorrichtungen, wie etwa einen Display-Adapter **516**, eine Disk **518** und einen E/A-Port, wie etwa einen seriellen Anschluss. Viele derartige Kombinationen aus E/A-Vorrichtungen, Bussen und Brücken können gemäß der vorliegenden Erfindung verwendet werden, wobei die dargestellte Kombination lediglich ein Beispiel für eine mögliche Kombination veranschaulicht.

**[0056]** Wenn ein Bild wie etwa ein Bild eines Objekts/einer Szene **540** erfasst wird, wird das Bild als R-, G- und B-Pixelintensitätswerte in einem CFA wie etwa dem allgemeinen Bayer-Muster erfasst. Die Pixelintensitätswerte werden an die Bildverarbeitungsschaltung **532** übermittelt. Die Bildverarbeitungsschaltung **532** besteht aus integrierten Schaltungen (ICs) und anderen Bauteilen, die neben anderen Funktionen eine Bildkomprimierung ausführen, um die Größe der Übertragung zwischen der Kamera **530** und dem Computersystem **510** zu reduzieren. Die Notwendigkeit für eine Farbinterpolation ist am signifikantesten, wenn das Bild angezeigt oder ausgegeben werden soll, und somit muss sie nicht in der Kamera **530** ausgeführt werden. Wenn der Benutzer oder die Anwendung das Herunterladen von Bildern wünscht/verlangt, werden die in der Bildspeichereinheit gespeicherten komprimierten Bilder aus der Bildspeichereinheit **534** zu dem E/A-Port **517** übertragen. Der E/A-Port **517** verwendet die abgebildete Bus-Brücken-Hierarchie (E/A-Bus **515** an Brücke **514** an Systembus **513**) zum temporären Speichern der Bilddaten in dem Speicher **511** oder optional auf der Platte **518**. Die komprimierten Bilder werden durch eine geeignete Anwendungssoftware (oder dedizierte Hardware) dekomprimiert und können dafür den Prozessor **512** verwenden. Die dekomprimierten Bilddaten werden als 8-Bit je Pixel Anordnung von Intensitätswerten wiederhergestellt, ähnlich oder identisch zu den Bilddaten gemäß der Erfassung durch den Sensor der Kamera **530**.

**[0057]** Das einmal dekomprimierte heruntergeladene Bild weist die Pixel somit in ihrem Bayer-Musterformat auf, das heißt mit einem 8-Bit-Intensitätswert für ein Pixel, jedoch nur einer von drei Farbbebenen Rot, Grün oder Blau zugeordnet. Wenn bei einer Anwendung eine Version des Bilds mit verbesserter Kantenwiedergabe auf einem Monitor **250** ausgegeben werden soll, so sollte die lokalisierte Schwellenwertbildung und Kantendetektierung gemäß den verschiedenen Ausführungsbeispielen der vorliegenden Erfindung eingesetzt werden. Das heruntergeladene und dekomprimierte Bild wird Pixel für Pixel und unter

Verwendung der vorstehend beschriebenen Techniken analysiert, wobei jedes Pixel entweder als Kanten- oder Nicht-Kantenpixel klassifiziert wird. Die detektierten Kanten können durch eine andere Anwendung oder ein Codemodul verbessert oder anderweitig verarbeitet werden. Während der Ausführung der lokalierten Schwellenwertbildung und der Kanten-detectierung können die Gradientenwerte und die normalisierten Gradientenwerte in dem Speicher **511** oder auf der Platte **518** in einer Array gespeichert werden. Wenn jedes Element der gespeicherten Array einen Datensatz an Stelle von Einzeldaten speichern kann, können der normalisierte Gradient, der Gradient, die ursprüngliche Pixelintensität, der lokalisierte Schwellenwert für das Pixel alle so gespeichert werden, dass sie für andere Anwendungen wieder verwendet werden können. Wenn der Zweck der Kantendetektierung eine Kantenverbesserung ist, so erscheint das gerenderte Bild **550** auf entsprechenden Anzeigegeräten (Monitor **520** und Adapter **516**) optisch mit höherer Detailgenauigkeit und schärferen Kanten als bei einem herkömmlichen Kantendetektionsverfahren. Das Bild mit verbesserten Kanten kann nach dem Rendering durch den Display-Adapter **516** auf der Platte **518** oder in dem Speicher **511** gespeichert und/oder direkt an den Monitor **520** ausgegeben werden. Unter Verwendung des menschlichen Sehens als Faktor bei der Schwellenwertbildung gemäß der vorstehenden Beschreibung in Bezug auf verschiedene Ausführungsbeispiele der Erfindung, wird das gerenderte Bild **550** Anwendern in einer höheren Bildqualität angezeigt. Eine umfassende Vielzahl von Bilddarstellungsanwendungen und Anwendungen für maschinelles Sehen, wie etwa OCR, die teilweise in dem Computersystem **500** implementiert sein können, können die lokalisierte Schwellenwertbildung und Kantendetektierung gemäß den verschiedenen Ausführungsbeispielen der Erfindung verwenden. Wenn ferner ein anderes Kantendetektierungssystem als die hierin beschriebenen bevorzugt wird, so kann die lokalisierte Schwellenwertbildungstechnik mit geringfügigen Modifikationen auch darauf angewandt werden, so dass das System effektiv arbeitet.

**[0058]** Der Computerprogrammcode zur lokalisier-ten Schwellenwertbildung und Kantendetektierung kann auf einem entfernbaren, computerlesbaren Me- dium wie etwa einer Floppy-Disk oder CD-ROM vor- gesehen werden und auch in Verbindung mit Softwa- re, welche die Dekomprimierung der von der Kamera **530** heruntergeladenen Bilder ausführt. Wie jede an- dere Software kann sie über ein Netzwerk an einem Kommu-nikationssystem heruntergeladen/verteilt werden, oder als Firmware verfügbar sein. Die herun- tergeladenen Bilder können gemäß ihrer Kanten- merkmale verbessert werden, so dass die Bilder dem Benutzer optisch präziser das Objekt/die Szene **540** anzeigen. Die Kantendetektierung und die lokalisier- te Schwellenwertbildung gemäß einem weiteren Aus-

führungsbeispiel der Erfindung können ebenfalls in Hardware an der Kamera **530** selbst implementiert werden.

**[0059]** In einem alternativen Ausführungsbeispiel der Erfindung kann die Bildverarbeitungsschaltung **532** so angepasst werden, dass sie eine Schaltkreisanordnung zur Ausführung der lokalisierten Schwellenwertbildung und der Kantendetektierung an der Kamera aufweist, so dass es sich bei dem in der Bildspeichereinheit **534** gespeicherten Bild um ein rauschfreies Bild handelt. Die Kantenmerkmale sind somit bereits klassifiziert, wenn das Bild auf ein Computersystem **510** oder eine Rendering-Vorrichtung heruntergeladen wird. In einem derartigen Ausführungsbeispiel können die Kantenmerkmale des Bilds vorab klassifiziert werden, so dass Techniken wie Rauschunterdrückung und Farbinterpolation präziser und besser ausgeführt werden. Die vorstehend beschriebenen Verfahren zur Ermittlung/Normalisierung von Pixelgradienten und lokalisierten Schwellenwerten sowie zum Detektieren von Kanten, werden einfach in VLSI- (Very Large Scale Integration) oder anderen integrierten Schaltkreisanordnungen implementiert.

**[0060]** Die Abbildung aus [Fig. 6](#) zeigt ein Flussdiagramm mit Modifikationen in Bezug auf die Gradienten-basierte Flankendetektierung aus [Fig. 3](#), anwendbar auf eine direkte Hardware-Implementierung.

**[0061]** Bei einer Implementierung in Hardware wie etwa der Kamera **530** aus [Fig. 5](#) können verschiedenen Modifikationen in Bezug auf das Verfahren aus [Fig. 3](#) eingesetzt werden. In der Annahme eines quadratischen Lokalisierungsbereichs mit  $M \times M$  müssen anfänglich die ersten  $M^2$  Pixelgradientenwerte bestimmt (Schritt **600**) und in einer Gradientenanordnung gespeichert werden. Diese Anzahl der  $M^2$  Berechnungen muss für ein bestimmtes Bild nur einmal ausgeführt werden, wie dies nachstehend im Text beschrieben ist. Als nächstes wird gemäß dem Schritt **620** der maximale Gradient in dem Lokalisierungsbereich bestimmt. In Hardware können ein Komparator und ein Register in Verbindung mit dem Schritt **600** eingesetzt werden, so dass während der Berechnung der Gradientenwerte für alle  $M^2$  Pixel auch die Bestimmung des Maximums vorgenommen wird. Die Normalisierung des Gradientenwerts bzw. der Gradientenwerte kann nur an dem Gradient des Pixels P ausgeführt werden, um das der Lokalisierungsbereich definiert ist (Schritt **630**). Als nächstes werden die minimalen und maximalen Intensitätswerte ermittelt (Schritt **640**). Bei Hardware kann dieser Schritt auch parallel zu der Bestimmung aller  $M^2$  Pixelgradientenwerte gemäß Schritt **600** durch ein Paar von Registern und Komparatoren ausgeführt werden. Als nächstes wird der Mittelwert der Intensitätswerte bestimmt (Schritt **650**). Schließlich kann die Kantende-

tekтирung durch die Bestimmung des lokalisierten Schwellenwerts und Vergleichen dieses Werts mit dem normalisierten Gradient für das Pixel P abgeschlossen werden (Schritt **660**).

**[0062]** Wenn das nächste Pixel für die Kantendetektierung erfasst wird (Schritt **670**), gibt es zwei Möglichkeiten. Entweder gehört das Pixel zu einer neuen Zeile von Bildern oder nicht (Prüfung in dem Schritt **678**). Wenn das Pixel keiner neuen Zeile angehört, ist es Bestandteil der gleichen Zeile, und in der Annahme einer Pixelabtastung von links nach rechts, wird die erste Spalte der  $M$  Pixelgradientenwerte nach außen verschoben oder aus der Gradientenanordnung gelöscht (Schritt **680**). Danach muss nur eine neue Spalte von  $M$  Pixelgradientenwerten für das nächste Pixel der gleichen Zeile, jedoch einer anderen Spalte berechnet werden (Schritt **685**). Bei einer Linksverschiebung (für eine Abtastung von links nach rechts) oder einer Rechtsverschiebung (für eine Abtastung von rechts nach links) der Spalten in der Gradientenanordnung in dem Schritt **680** muss nur die letzte Spalte in der Anordnung mit den neu berechneten Gradientenwerten aktualisiert werden. Dies macht die Hardware erheblich schneller.

**[0063]** Wenn das neue Pixel nicht aus der gleichen Zeile stammt, so wurde an Stelle einer Verschiebung einer Spalte aus der Gradientenanordnung ein Zeilenwechsel bewirkt, wobei eine Zeile von  $M$  Werten nach außen verschoben wird (Schritt **690**). Bei Annahme einer Abtastung von oben nach unten wird die erste Zeile nach außen verschoben, wobei alle Zeilen nach oben verschoben und die letzte leere Zeile in der Gradientenanordnung durch die neu berechneten  $M$  Gradientenwerte ersetzt wird (Schritt **695**). Wenn eine Pixelabtastung Zeile für Zeile ausgeführt wird, so sollte die Abtastrichtung (links nach rechts und rechts nach links) mit jeder neuen Zeile wechseln. Auf diese Weise wird die geringste Anzahl der Gradientenberechnungen ausgeführt. Eine Anzahl von  $M^2$  Gradientenoperationen wird nur ausgeführt, wenn ein Zick-Zack-Muster der Abtastung und der Pixelverarbeitung angewandt wird. Eine Änderung der Abtastrichtung macht auch eine Änderung der Richtung der Verschiebung der Spalte/der Zeile der Gradientenanordnung erforderlich. Eine derartige Implementierung ist wünschenswert, wenn die Gradientenwerte für das ganze Bild nicht gespeichert werden können oder wenn ein Ansatz mit einem Durchlauf aufgrund der Geschwindigkeitsbeschränkungen bevorzugt wird.

**[0064]** Selbst bei einer Reduzierung der Gradientenberechnungen muss die Ermittlung der Minima und Maxima der Intensitäts- und Gradientenwerte für jedes einzelne Pixel wiederholt werden, da es nicht ausreicht, die  $M$  neuen Werte lediglich mit den vorher bestimmten Minima und Maxima zu vergleichen, da die nach außen verschobenen und gelöschten An-

ordnungswerte diese Minima und/oder Maxima aufweisen können und es somit zu fehlerhaften Ergebnissen kommen kann.

### **Patentansprüche**

1. Verfahren, das folgendes umfasst:

Definieren (**310**) eines Lokalisierungsbereichs eines Bilds, wobei der genannte Lokalisierungsbereich einen ausgewählten Abschnitt des genannten Bilds aufweist; und  
 automatisches Bestimmen (**330**) ohne Benutzereingriff eines zur Kantendetektierung innerhalb des genannten Lokalisierungsbereichs angewandten Schwellenwerts, wobei der genannte Schwellenwert durch einen Leistungsfunktionsoperator bestimmt wird, der die Wahrnehmungseigenschaften in Bezug auf die Intensität des menschlichen Sehvermögens reflektiert, wobei der genannte Leistungsfunktionsoperator ein Maß der relativen Lokalisierungsintensität erhöht auf einen Leistungsfaktor darstellt, wobei die genannte relative Lokalisierungsintensität der Quotient der Differenz zwischen dem Mittelwert und dem Minimum aller Intensitätswerte innerhalb des genannten Lokalisierungsbereichs ist sowie der Differenz zwischen dem Maximum und dem genannten Minimum aller Intensitätswerte in dem genannten Lokalisierungsbereich.

2. Verfahren nach Anspruch 1, wobei der genannte Leistungsfaktor einer Wahrnehmungsfähigkeit des menschlichen Sehvermögens auf Intensitätsveränderungen entspricht.

3. Verfahren nach Anspruch 2, wobei der genannte Leistungsfaktor gleich 0,45 ist.

4. Verfahren nach Anspruch 1, wobei dieses ferner folgendes umfasst:

Detektieren (**335**) von Kantenmerkmalen des genannten Bilds durch Vergleichen des genannten Schwellenwertes mit einem normalisierten Gradient für ein Pixel P, um das der genannte Lokalisierungsbereich definiert ist.

5. Verfahren nach Anspruch 4, wobei das genannte Pixel als ein Pixel klassifiziert ist, das zu einem Kantenmerkmal des genannten Bilds gehört, wenn der genannte normalisierte Gradient des genannten Pixels den genannten Schwellenwert überschreitet.

6. Computerprogramm, das eine Computerprogramm-Codeeinrichtung umfasst, die alle Schritte der Ansprüche 1 bis 5 ausführen kann, wenn das Programm auf einem Computer ausgeführt wird.

7. Computerprogramm nach Anspruch 6, ausgeführt auf einem computerlesbaren Medium.

Es folgen 6 Blatt Zeichnungen

## Anhängende Zeichnungen

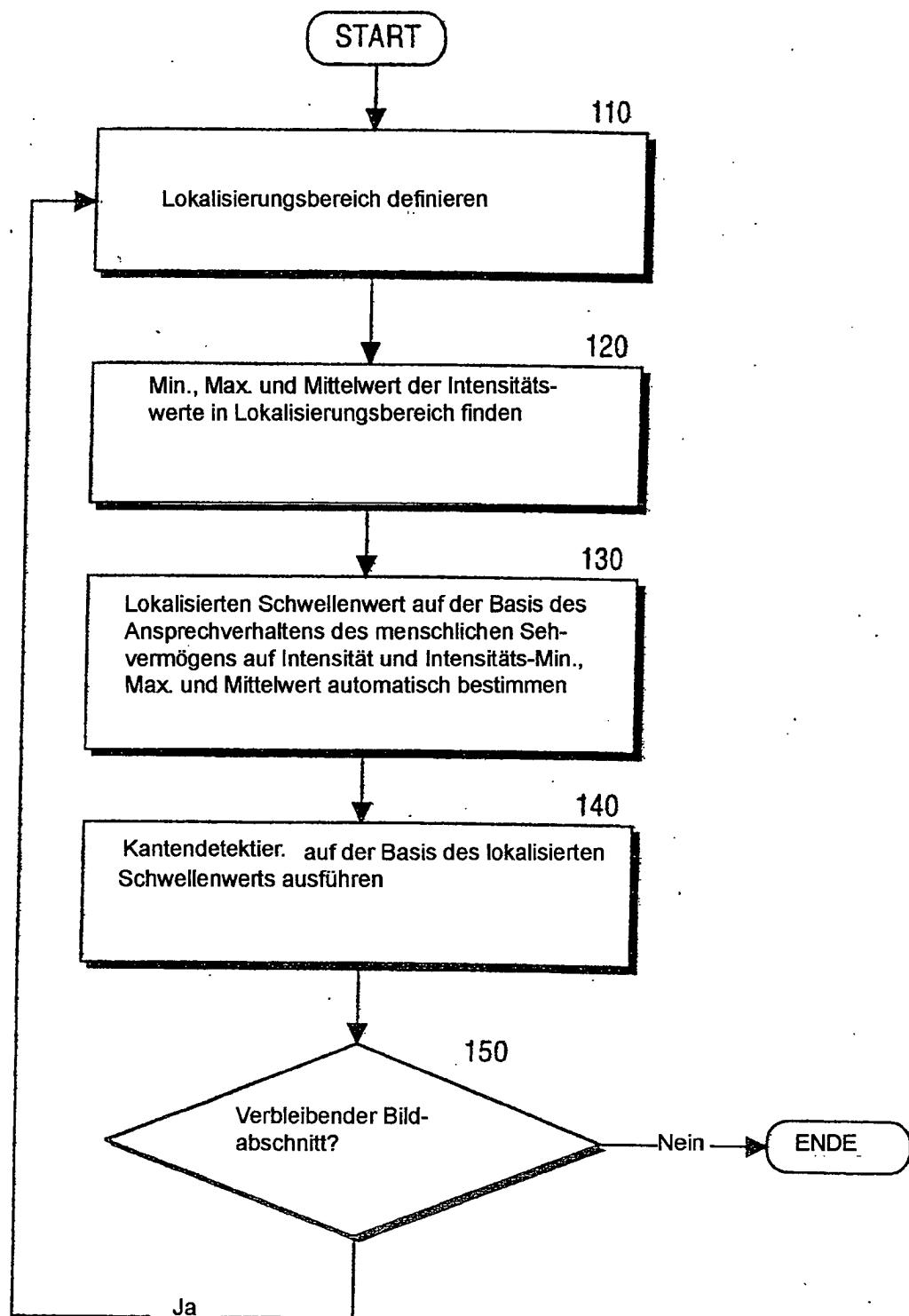


Fig. 1

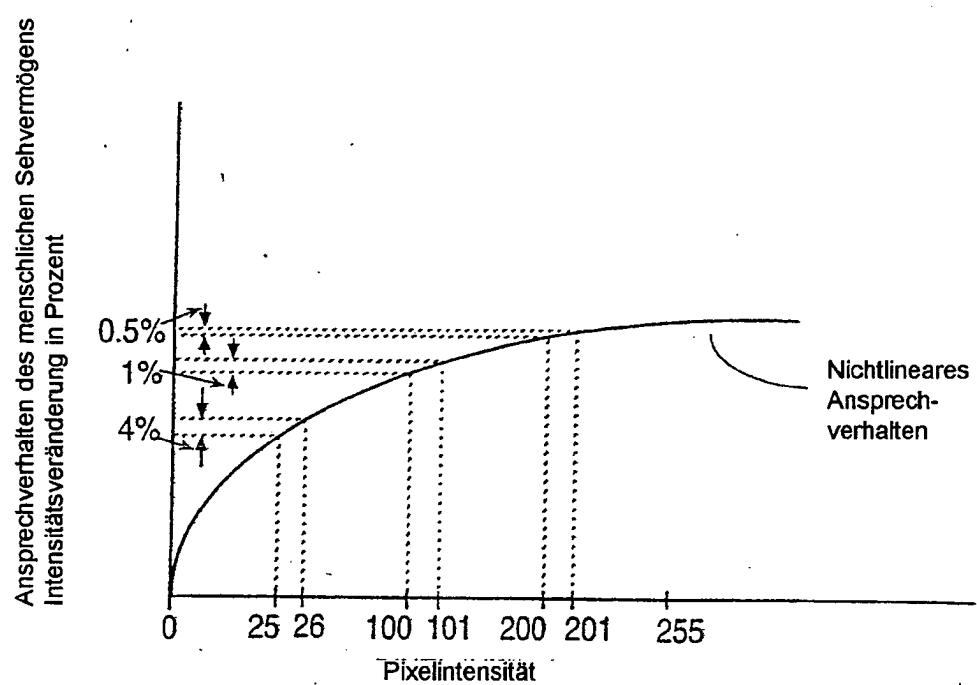


Fig. 2

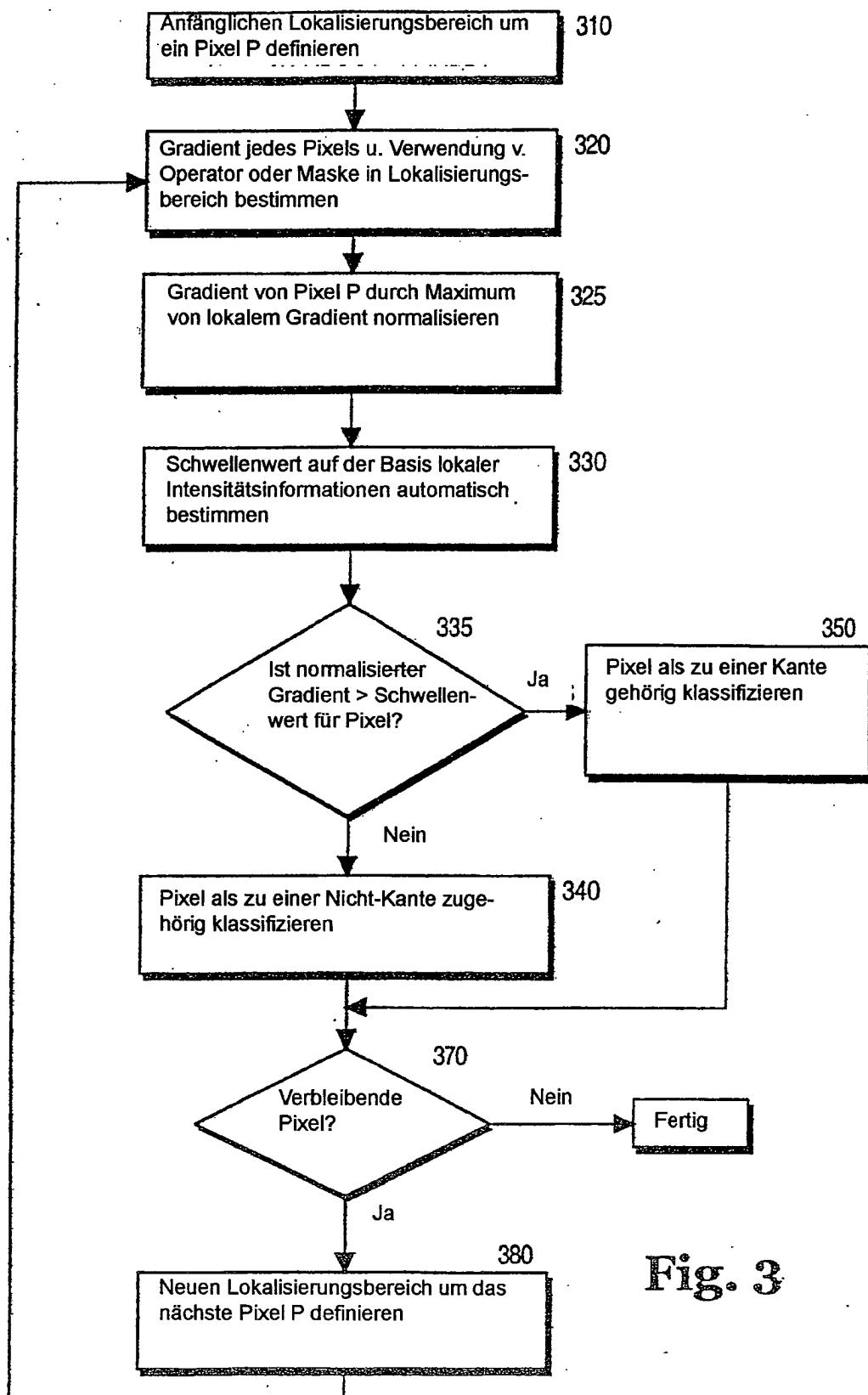
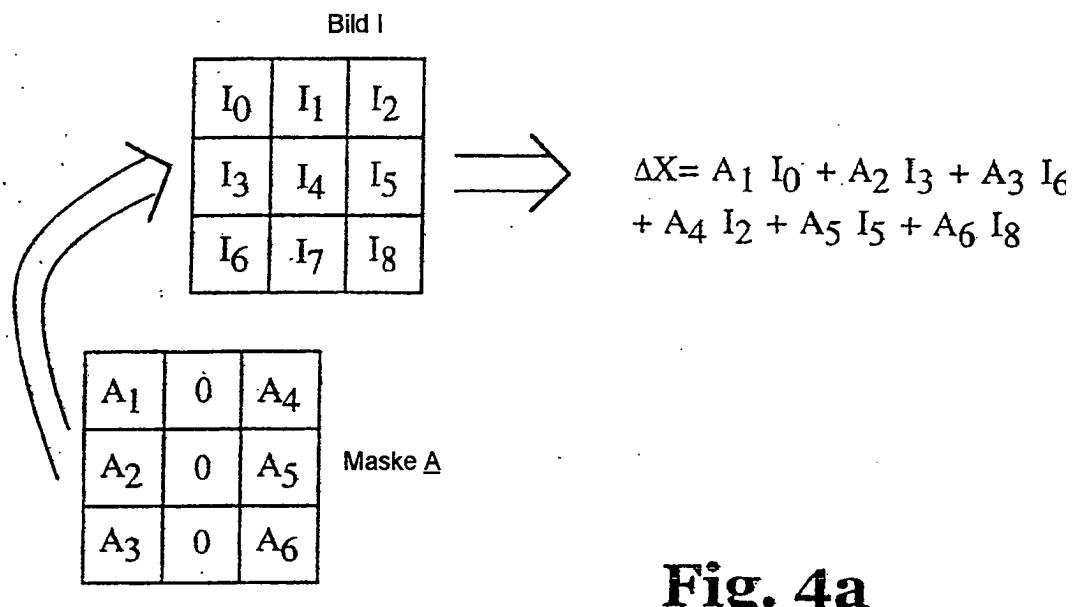
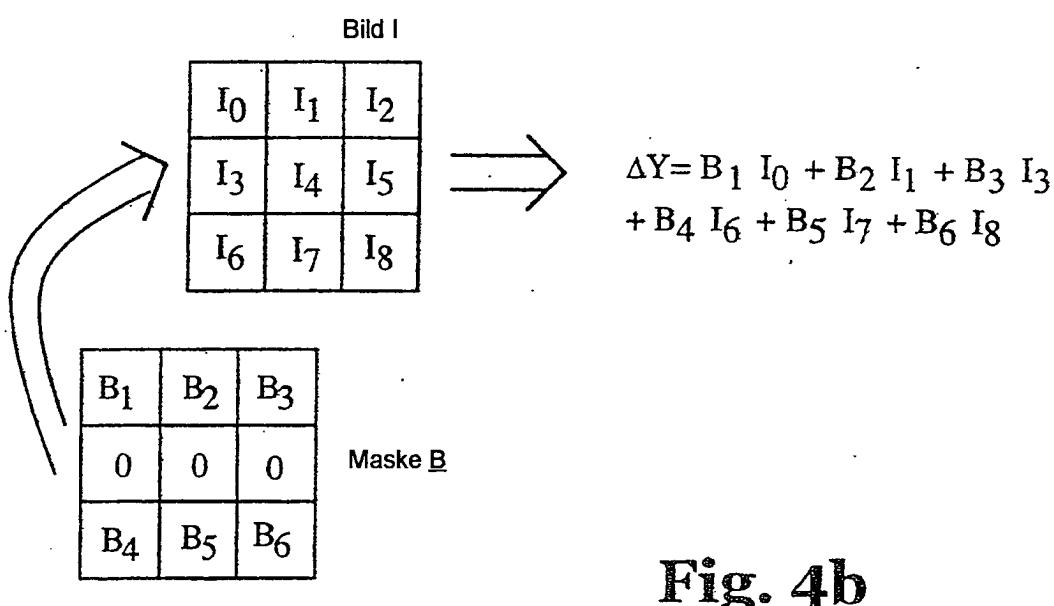


Fig. 3

**Fig. 4a****Fig. 4b**

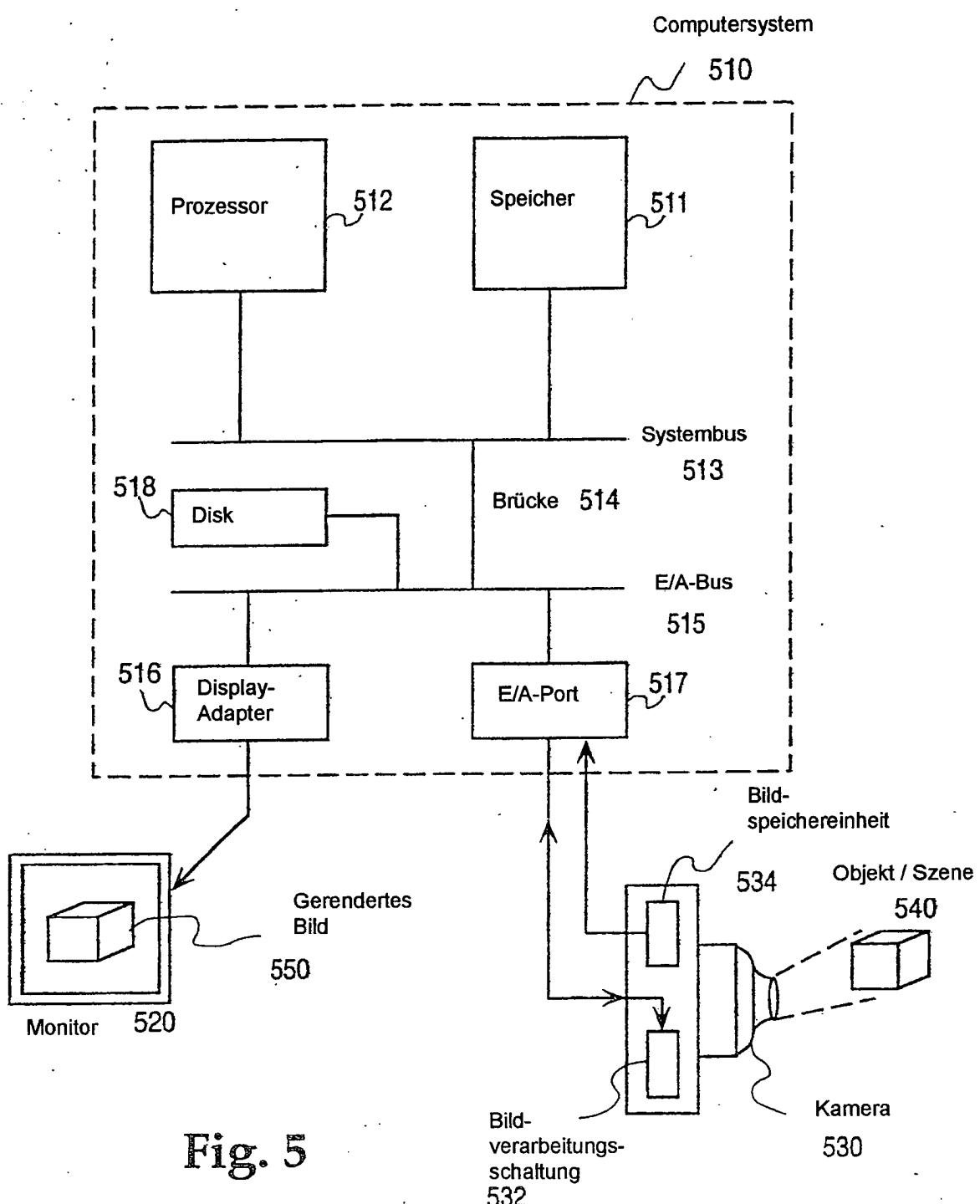


Fig. 5

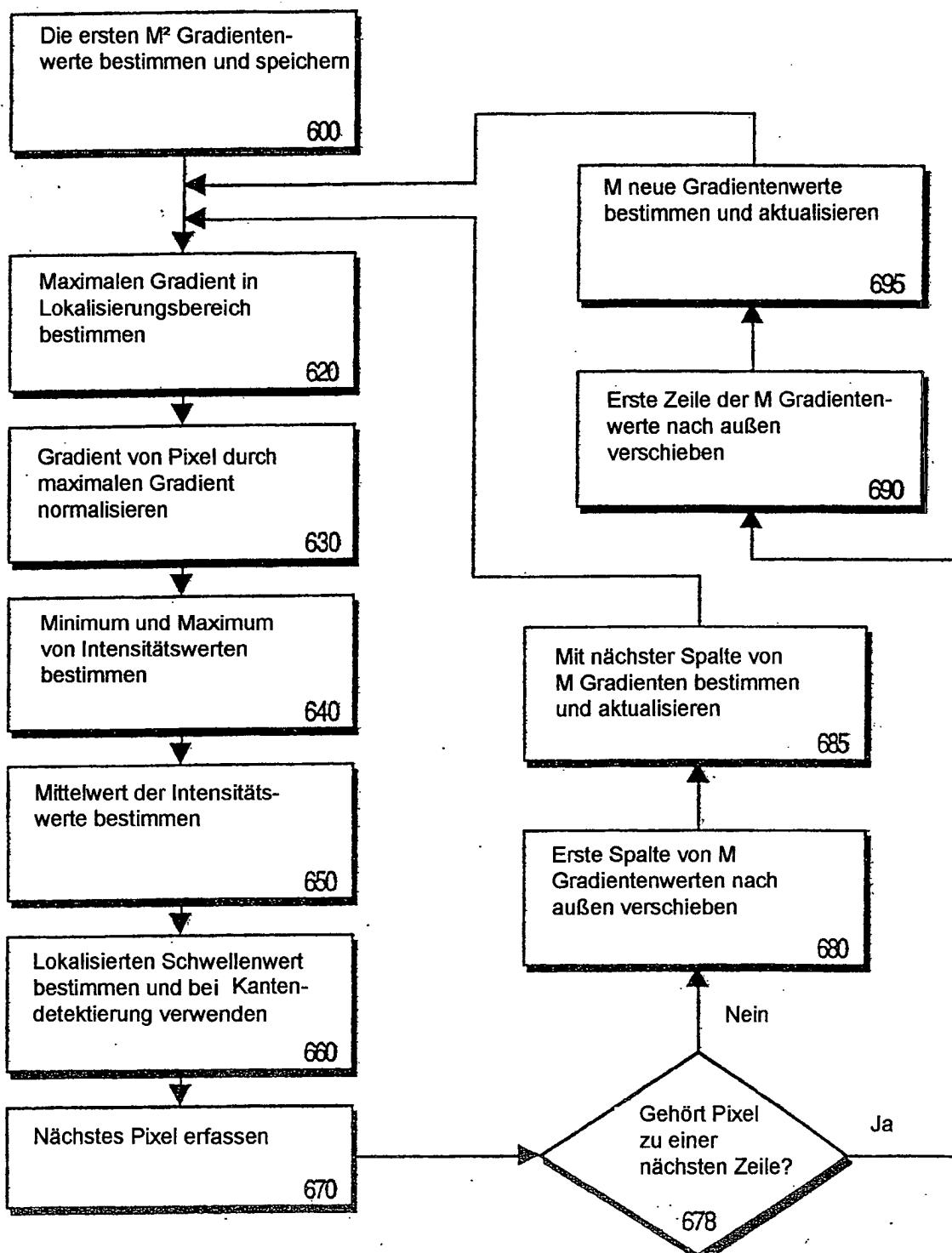


Fig. 6