



(10) **DE 10 2012 102 797 B4** 2017.08.10

(12) **Patentschrift**

(21) Aktenzeichen: **10 2012 102 797.8**  
 (22) Anmeldetag: **30.03.2012**  
 (43) Offenlegungstag: **02.10.2013**  
 (45) Veröffentlichungstag  
 der Patenterteilung: **10.08.2017**

(51) Int Cl.: **G09G 5/22 (2006.01)**  
**G06T 11/60 (2006.01)**  
**G06T 7/40 (2017.01)**  
**G06K 9/00 (2006.01)**  
**H04W 88/02 (2009.01)**

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

(73) Patentinhaber:  
**beyo GmbH, 14467 Potsdam, DE**

(72) Erfinder:  
**Göktekin, Cüneyt, 14467, Potsdam, DE**

(74) Vertreter:  
**Farago, Peter, Dipl.-Ing.Univ., 80538 München, DE**

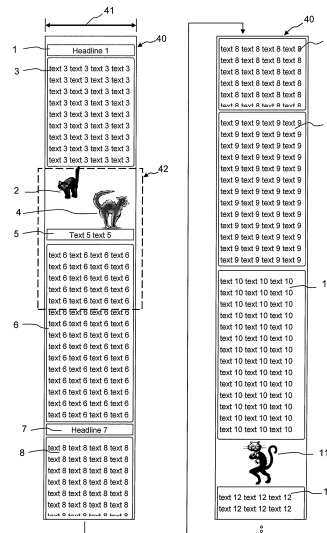
(56) Ermittelte Stand der Technik:  
**siehe Folgeseiten**

(54) Bezeichnung: **Kamerabasiertes Mobilfunkgerät zur Konvertierung eines Dokuments anhand von aufgenommenen Bildern in ein Format zur optimierten Anzeige auf dem kamerabasierten Mobilfunkgerät**

(57) Hauptanspruch: Verfahren zur Aufnahme eines Dokuments mit einem kamerabasierten Mobilfunkgerät und zur Konversion von Textinformation des Dokuments in ein Format zu einer geeigneten Darstellung auf dem Mobilfunkgerät, die folgenden auf dem Mobilfunkgerät ausgeführten Schritte umfassend:

- a) Aufnehmen des Dokuments mit dem Mobilfunkgerät durch mindestens zwei Bilder, von denen benachbarte Bilder jeweils einen gemeinsamen überlappenden Bereich des Dokuments aufweisen, und Abspeichern der mindestens zwei Bilder;
- b) Erkennen der benachbarten Bilder und Zusammen setzen zu einem jeweiligen zusammengesetzten Bild anhand der gemeinsamen überlappenden Bereiche des Dokuments (Stitching);
- c) Erkennen einer jeweiligen Layout-Struktur mit mindestens einem Textblock in dem jeweiligen zusammengesetzten Bild;
- d) Erkennen von Buchstabentext in dem jeweiligen Textblock durch Anwenden eines Optical Character Recognition Algorithmus (OCR) und Abspeichern des jeweiligen Buchstabentexts zu dem jeweiligen Textblock;
- e) Bestimmen und Abspeichern einer Reihenfolge der jeweiligen Textblöcke in dem Dokument unter Berücksichtigung der jeweiligen Layout-Struktur;
- f) Bestimmen einer Breite eines Displays des Mobilfunkgeräts durch automatisches Abrufen von entsprechenden Parametern vom Mobilfunkgerät und Bestimmen des Formats für die Darstellung der Buchstabentexte, das an die Breite des Displays angepasst ist, um beim Lesen der Buchstabentexte auf dem Display nur vertikal scrollen zu brauchen;
- g) Erzeugen einer Datei (40) in dem Format mit den jeweiligen Buchstabentexten in der bestimmten Reihenfolge der jeweiligen Textblöcke;

h) Bereitstellen der Datei (40) für das Mobilfunkgerät zur Darstellung auf dessen Display.



(56) Ermittelter Stand der Technik:

|    |                   |    |
|----|-------------------|----|
| US | 2004 / 0 255 244  | A1 |
| US | 2006 / 0 136 491  | A1 |
| US | 2008 / 0 235 564  | A1 |
| US | 2009 / 0 244 293  | A1 |
| US | 2010 / 0 128 131  | A1 |
| US | 2010 / 0 199 197  | A1 |
| US | 2011 / 0 280 450  | A1 |
| WO | 2006/ 124 473     | A2 |
| WO | 2009/ 104 193     | A1 |
| KR | 10 2008 0 050 272 | A  |

**COLLINS-THOMPSON K, NICKOLOV R.:**  
A Clustering-Based Algorithm for Automatic Document Separation, 2002, [[http://research.microsoft.com/en-us/um/people/kevynct/pubs/sigir02\\_docsep.pdf](http://research.microsoft.com/en-us/um/people/kevynct/pubs/sigir02_docsep.pdf)]

**HU J, KASHI R, WILFON G.:** Document Image Layout Comparison and Classification, In Proc. of the Conf. on Document Analysis and Recognition, 1999, S. 285-288

**MARINAI S.:** Introduction to Document Analysis and Recognition; In: Machine Learning in Document Analysis and Recognition; 2008; S. 1-20

**Beschreibung**

## Erfindungsgebiet

**[0001]** Die vorliegende Erfindung betrifft ein Verfahren und ein kamerabasiertes Mobilfunkgerät zur Aufnahme eines Dokuments mit dem kamerabasierten Mobilfunkgerät und Konvertierung des Dokuments mit darin enthaltenen Text- und Bildinformationen in ein Format, das vor allem in der Breite einem Display des kamerabasierten Mobilfunkgeräts angepasst ist.

## Hintergrund der Erfindung

**[0002]** Mobilfunkgeräte, wie beispielsweise ein Mobiltelefon, ein Smartphone, ein iPad® oder ähnliches, weisen aktuell zum größten Teil schon integrierte Kameras mit einer Auflösung von 5–12 Megapixel auf. Die Mobilfunkgeräte gelten als ständige Begleiter der jeweiligen Benutzer, und vor allem Geschäftsleute wünschen sich daher auch immer weitreichendere Einsatzmöglichkeiten. Oft werden die kamerabasierten Mobilfunkgeräte auch dafür genutzt, sich durch Kameraaufnahmen Notizen über einen Zeitungsartikel oder ein Dokument zu machen, oder sie für Übersetzungszwecke zu nutzen.

**[0003]** Aus US2011/0280,450A1 ist ein System bekannt, das die Qualität von mit dem Mobilfunkgerät aufgenommenen Dokumenten bestimmt und so das aufgenommene Bild akzeptieren oder zurückweisen kann.

**[0004]** Aus US2009/244,293A1 ist ein System bekannt, mit dem die Polizei ein Bild einer Person und eines Nummernschildes aufnehmen kann und das Bild zusammen mit weiteren Daten, die im Mobilfunkgerät eingegeben werden, an einen Server senden kann. Dort werden die Daten beispielsweise geprüft, abgespeichert und weiterverarbeitet.

**[0005]** Aus WO2009/104193A1 ist ein System bekannt, mit dem ein bestimmter Teil eines Zeitungsartikels mit dem Mobilfunkgerät per Kamera aufgenommen und das aufgenommene Bild an einen Server gesendet wird. Der Server identifiziert den entsprechenden Zeitungsartikel und sendet genau diesen Inhalt dann an das Mobilfunkgerät zurück, um den Inhalt dann später lesen zu können.

**[0006]** KR102008/0050272A beschreibt ein Verfahren zu einer verbesserten Schrift- bzw. Buchstaben-erkennung in Dokumenten, die mit dem Mobilfunkgerät aufgenommenen worden sind, wobei insbesondere verschiedene Bildvorverarbeitungsmethoden zum Einsatz kommen.

**[0007]** WO2006/124473A2 beschreibt ein Verfahren und ein System, mit dem basierend auf dem Mobilfunkgerät Dokumente, wie beispielsweise Visitenkar-

ten, Quittungen, Formulare und sonstige Dokumente, aufgenommen und drahtlos zu einem Server gesendet werden. Im Server werden die Daten in eine Form gebracht, die beispielsweise mit einem PC weiterverarbeitet werden können. Auch ist eine Handschrifterkennung berücksichtigt.

**[0008]** US2010/0128131A1 beschreibt ein Verfahren mit dem das Dokument beispielsweise durch mehrere Bilder aufgenommen wird. Danach werden die aufgenommenen Bilder bearbeitet, indem beispielsweise ein Zusammenfügen der Bilder (Stitching), eine Blockerkennung und eine OCR Konversion in Buchstabentext ausgeführt wird. Das so aufbereitete Dokument kann archiviert und anderen Personen zur Verfügung gestellt werden. Ein Erkennen von Schlüsselworten und eine Abspeicherung bei der Archivierung können ein Wiederauffinden erleichtern.

**[0009]** Die Druckschrift von Simone Marinai mit dem Titel „Introduction to Document Analysis and Recognition“ der Zeitschrift „Machine Learning in Document Analysis and Recognition“ von 2008 (S. 1–20) offenbart ein Verfahren zur Dokumentenanalyse aus aufgenommenen Bildern. Darin wird eine Vor-Bildverarbeitung mit anschließender Layoutanalyse vorgestellt, in der Textzeilen und Textblöcke durch unterschiedliche Abstände von Textzeilen erkannt werden.

**[0010]** US2008/0235,564A1 offenbart ein Verfahren zur Darstellung von audiovisuellen Inhalten, wobei unterschiedliche Inhalte, wie beispielsweise Audio-Information, Bildinformation und Textinformation, erkannt, unterschiedlich aufgeteilt und dargestellt werden. Dabei finden auch semantische Analysen Anwendung. Verschiedenartige Mobiltelefone zur Darstellung der Informationen werden dabei nicht berücksichtigt.

**[0011]** US2006/0136,491A1 offenbart ein Verfahren zur Darstellung von Dokumenteninhalten, wobei unterschiedliche Inhalte, wie die Bildinformationen und Textinformationen, erkannt, unterschiedlich aufgeteilt und dargestellt werden. Dabei finden auch semantische Analysen Anwendung. Verschiedenartige Mobiltelefone zur Darstellung der Informationen werden dabei nicht berücksichtigt.

**[0012]** US2004/0255,244A1 offenbart ein Verfahren zur Darstellung von Webseiteninhalten auf kleinen Anzeigegeräten, wie beispielsweise auf Mobiltelefonen. Dabei werden HTML- oder XML-Webseiteninhalte in CSS-Daten (Cascading Style Sheets Data) umgewandelt, wobei bevorzugt von einem Server eine Reihe aufeinanderfolgender Inhalte analysiert, zerlegt und als die aufeinanderfolgenden Inhalte zum Anzeigegerät übertragen und dort spaltenförmig anzeigen werden.

**[0013]** Die oben beschriebenen Verfahren oder Systeme sind nicht jedoch darauf ausgerichtet, das jeweilige aufgenommene Dokument, das beispielsweise durch mehrere aufeinanderfolgende Bilder aufgenommen worden ist, geeignet auf einem kleinen Display eines Mobilfunkgerät adäquat darzustellen. Bei einem Lesen des Dokumentes ist meist ein Scrollen auch zur Seite nötig, wodurch ein Lesen einer breiten Zeitschrift für den Nutzer sehr umständlich und beschwerlich ist.

**[0014]** US2010/0199197A1 beschreibt ein Verfahren, welches Server-basiert eine gewünschte Web-Seite aufruft und in ein für ein Mobilfunkgerät besser lesbares Format umwandelt und dem Mobilfunkgerät über einen Internet-Link zusendet. Hier findet eine Umwandlung von dargestellten Informationen, die auf einen breiten PC-Bildschirm ausgerichtet sind, in ein Format statt, das auf einen schmalen Bildschirm des Mobilfunkgeräts angepasst ist. Dabei können beispielsweise auch unnötige Werbebilder an der Seite einer Web-Seite unterdrückt werden.

**[0015]** Eine befriedigende Darstellung von mit dem Mobilfunkgerät aufgenommenen Textdokumenten auf dem Mobilfunkgerät ist nicht verfügbar, ohne dass beispielsweise seitlich auch gescrollt werden müsste.

#### Zusammenfassung der Erfindung

**[0016]** Daher besteht die Aufgabe der Erfindung, in der Bereitstellung eines Verfahrens und einer Vorrichtung zum Aufnehmen eines Dokuments mit einem Mobilfunkgerät mit integrierter Kamera und zum Konvertieren von Textinformationen der aufgenommenen Bilder des Dokuments in ein Format, das möglichst gut auf dem Mobilfunkgerät angezeigt werden kann, um dabei einem Anwender beim Lesen ein seitliches Scrollen zu ersparen. Dabei soll die richtige Reihenfolge von Textstellen möglichst gut erkannt und beibehalten werden. Dadurch sollen auch die Nachteile aus dem Stand der Technik beseitigt werden.

#### Offenbarung der Erfindung

**[0017]** Die vorstehenden Aufgaben sowie weitere der Beschreibung zu entnehmende Aufgaben werden von einem Verfahren und von einer Vorrichtung zum Aufnehmen und zum Konvertieren eines Dokuments mit einem kamerabasierten Mobilfunkgerät gemäß dem unabhängigen Anspruch 1 bzw. 13 gelöst. Weitere vorteilhafte Ausbildungen der Erfindung sind in den abhängigen Ansprüchen angegeben.

**[0018]** Die durch die Erfindung erreichten Vorteile bestehen darin, dass Dokumente mit einem kamerabasierten Mobilfunkgerät einfach der Reihe nach, seitenweise oder Textbaustein für Textbaustein aufgenommen werden können und danach auf dem kamerabasierten Mobilfunkgerät eine automatische Kon-

vertierung der Texte so erfolgt, dass aufeinanderfolgende Textbausteine in der richtigen Reihenfolge untereinander gesetzt werden und so auf dem Mobilfunkgerät dargestellt werden. Demgemäß müssen bei im Originaldokument nebeneinander angeordneten, aufeinanderfolgenden Textblöcken nicht seitlich, sondern nur vertikal gescrollt werden, was auf Mobilfunkgeräten für einen Nutzer sehr umständlich wäre. Aufeinanderfolgende Textbausteine eines Dokuments werden in der Reihenfolge erkannt und neu angeordnet, indem sie in ein geeignetes Layout bzw. Format gebracht werden, das gerade so breit ist, dass es einem Display des Mobilfunkgeräts entspricht. Die Textblöcke werden entsprechend dem geeigneten Format als Datei in dem kamerabasierten Mobilfunkgerät abgespeichert und stehen so bequem lesbar als ein Textdokument auf dem Mobilfunkgerät zur Verfügung. Insbesondere bei breiten Dokumenten, wie beispielsweise bei Zeitungen oder bei Dokumenten im Querformat ist diese Art der Konvertierung sehr angenehm, einen Fließtext auf dem Display des Mobilfunkgeräts dargestellt zu bekommen, der automatisch der Breite des Displays angepasst ist, und bei dem nur vertikal, entlang dem Text und nicht seitlich gescrollt zu werden braucht.

**[0019]** Bei dem erfindungsgemäßen Verfahren wird die Konvertierung der Bilder ganz auf dem Mobilfunkgerät bearbeitet. Bei einem anderen alternativen Verfahren wird die Konvertierung der Bilder bevorzugt zum überwiegenden Teil auf einem Server bearbeitet, um Rechenleistung zu sparen und zugleich eine Kopie in einem Dokumentenarchiv abzulegen.

**[0020]** Auch können zusätzlich Dokumente in einer zusätzlichen zweiten Form abgespeichert werden, die beispielsweise auf PC Monitore angepasst sind. So kann auch der Textinhalt von Zeitungen, die ein Vielfaches breiter als DIN A4 sind, auf eine DIN A4 Breite reduziert und abgespeichert werden. Eine solche Optimierung der Darstellung macht es dem Nutzer angenehm, einen abfotografierten Dokumententext ohne Klimmzüge und Suchen, wo eine aktuelle Textpassage fortgesetzt wird, lesen zu können.

**[0021]** Eine bevorzugte Ausführungsform gemäß der vorliegenden Erfindung ist in nachfolgenden Zeichnungen und in einer detaillierten Beschreibung dargestellt, soll aber die vorliegende Erfindung nicht darauf begrenzen.

#### Kurzbeschreibung der Zeichnungen

**[0022]** Fig. 1a zeigt ein Bild eines Dokuments mit verschiedenen Textblöcken, Artikelüberschriften und zwei Abbildungen.

**[0023]** Fig. 1b zeigt ein weiteres Bild eines Dokuments mit verschiedenen Textblöcken, Artikelüberschriften und drei Abbildungen.

**[0024]** Fig. 2 zeigt links im Bild eine Darstellung eines ersten Teils einer Datei in einem geeigneten Format für eine Darstellung auf einem Mobilfunkgerät mit den Textblöcken aus Fig. 1a, wobei rechts im Bild ein zweiter Teil der Datei dargestellt ist, der den ersten Teil der Datei fortsetzt.

**[0025]** Fig. 3 zeigt das gleiche Dokument mit den verschiedenen Textblöcken aus Fig. 1a, wobei die verschiedenen Textblöcke durch drei Bilder von einem kamerabasierten Mobilfunkgerät aufgenommen werden.

Detaillierte Beschreibung  
von Ausführungsbeispielen

**[0026]** Fig. 1a und Fig. 1b repräsentieren jeweils eine Seite eines beispielhaften Dokuments, wobei Fig. 1a eine erste Seite und Fig. 1b eine zweite Seite des Dokuments zeigen. Die erste Seite des Dokuments zeigt beispielsweise eine erste Überschrift an einer Stelle **1** eines ersten Artikels mit zugehörigen Textblöcken an Stellen **3, 5, 6** und Abbildungen an Stellen **2, 4** und eine zweite Überschrift an einer Stelle **7** eines zweiten Artikels mit zugehörigen Textblöcken an Stellen **8, 9**. Auf der zweiten Seite in Fig. 1b wird der zweite Artikel fortgesetzt durch entsprechende weitere Textblöcke an Stellen **10, 12, 13** und durch eine Abbildung an einer Stelle **11**. Auf der zweiten Seite folgt dann noch ein dritter Artikel mit einer dritten Überschrift an einer Stelle **14**, mit zugehörigen Textblöcken an Stellen **15, 17, 19** und zwei Abbildungen an Stellen **16, 18**.

**[0027]** Mit dem Ziel das in Fig. 1a und Fig. 1b dargestellte Dokument zu registrieren bzw. zu archivieren und auf einem Mobilfunkgerät lesbar zu machen, werden die erste und die zweite Seite vorzugsweise mit dem Mobilfunkgerät abfotografiert, wobei in diesem Beispiel von der Kamera des Mobilfunkgeräts ein erstes Bild **30** und ein zweites Bild **31** erhalten abgespeichert werden. Im dargestellten Beispiel umfasst das erste Bild **30** einen ersten Bildbereich und das zweite Bild **31** einen zweiten Bildbereich. Bei Dokumenten mit mehreren Seiten werden entsprechend alle Seiten, die archiviert werden sollen, abfotografiert. Denkbar ist bei der Aufnahme des Dokuments durch das Mobilfunkgerät auch eine Erkennung einer genügend guten Ausrichtung der Kamera zu dem aufzunehmenden Text. Dabei können beispielsweise auch akustische Feed-back Verfahren zur genügend guten Ausrichtung eingesetzt werden.

**[0028]** Bevorzugt wird das erste Bild **30** durch einen Layout Erkennungsalgorithmus verarbeitet, so dass in dem ersten Bild **30** die Textblöcke an den Stellen **1, 3, 5, 6, 7, 8, 9** und bevorzugt die Abbildungen an den Stellen **2, 4** erkannt werden. Dabei wird auch eine Layout-Struktur, bzw. eine Verteilung der Textblöcke und bevorzugt der Abbildungen in dem ersten Bild

**30** erkannt und abgespeichert. Die Layout-Struktur wird danach ausgewertet, um festzustellen, welche jeweils benachbarten Textblöcke und Abbildungen zu jeweils einem Artikel gehören und welche zu einem weiteren Artikel. Der Layout Erkennungsalgorithmus basiert auf bekannten Algorithmen zur digitalen Bildverarbeitung, um bevorzugt Kanten zu erkennen und Textblöcke, die Bereiche mit Text sind. Auch können dabei bevorzugt Abbildungen erkannt werden. Der Layout Erkennungsalgorithmus erkennt dabei bevorzugt auch Abstände zwischen den Textblöcken und Abbildungen. Desweiteren erkennt der Layout Erkennungsalgorithmus auch bevorzugt Überschriften in einem Textblock, das heißt, ob ein bestimmter Textblock eine Überschrift ist, wie beispielsweise an den Stellen **1** und **7** in Fig. 1a. Zur Erkennung einer Überschrift kann der Layout Erkennungsalgorithmus eine Erkennung entweder durch eine Schriftgrößenbestimmung bei einem Vergleich mit der Schriftgröße benachbarter Textblöcke vornehmen oder auch beispielsweise syntaktische Charakteristika berücksichtigen. Es können dazu auch Ausgabeparameter aus einer nachfolgenden OCR Analyse mit verwendet werden. Mit dem zweiten Bild **31** und mit weiteren Bildern, falls vorhanden, wird wie mit dem ersten Bild **30** verfahren.

**[0029]** Die erkannten Textblöcke werden jeweils einem Optical Character Recognition (OCR) Algorithmus zugeführt, der darin Buchstabentext erkennt und ausgibt. Alternativ kann auch das gesamte erste Bild **30** dem OCR Algorithmus zugeführt werden. Der Buchstabentext oder die Buchstabentexte werden bevorzugt in einem Standard-Text-Format oder als Fließtext abgespeichert und mit dem entsprechenden Textblock verknüpft. Trennstriche zu einer Zeilenumbruchtrennung werden bevorzugt entfernt. Dabei ist es auch denkbar, dass noch weitere digitale Vorverarbeitungsalgorithmen zur Bildverbesserung angewendet werden, um dadurch eine bessere Erkennung des Buchstabentextes zu ermöglichen. Mit dem zweiten Bild **31** und mit weiteren Bildern, falls vorhanden, wird wie mit dem ersten Bild **30** verfahren.

**[0030]** In einem folgenden Schritt wird durch einen Reihenfolgeerkennungsalgorithmus eine Reihenfolge der zuvor bestimmten Textblöcke und bevorzugt auch der Abbildungen bestimmt. Der Reihenfolgeerkennungsalgorithmus berücksichtigt dabei bevorzugt folgende Parameter und Sub-Funktionen bzw. Sub-Algorithmen:

- ein Abstand der Textblöcke zueinander;
- eine syntaktische Verknüpfung, wenn beispielsweise ein letzter Satz in einem Textblock in einem nächsten Textblock weitergeführt und dort beendet wird;
- eine Anordnungsregel, die erkannt wird, wie beispielsweise links oben – links unten, rechts oben – rechts unten;
- Trennstriche und/oder Rahmen um Textblöcke;

- Weiterführung von Trennstrichen und/oder Rahmen um Textblöcke in benachbarten Bildern;
- Erkennung und Analyse von mindestens einem Schlüsselwort in benachbarten Textblöcken;
- Erkennung von aufeinanderfolgenden Texten mit neuronalen Netzen.

**[0031]** Bei der Reihenfolgeerkennung werden bevorzugt zahlreiche parallel arbeitende Algorithmen eingesetzt, die die Reihenfolge der Textblöcke und bevorzugt der Abbildungen erkennen. Dabei können auch semantische Erkennungs-Algorithmen eingesetzt werden. Auch wird bevorzugt eine Wort-Histogramm-Analyse zu den jeweiligen Buchstabentexten der Textblöcke durchgeführt, die eine Zuordnung zu einem bestimmten Artikel in dem Dokument ermöglicht. Ebenso denkbar ist zusätzlich eine Anwendung eines Kantenfilters mit anschließender Cluster-Bestimmung, um dadurch die Textblöcke zu erkennen. Ebenso werden bevorzugt aufeinanderfolgende Bilder **30**, **31** auf eine Zusammengehörigkeit von Textblöcken bzw. hinsichtlich deren Reihenfolge untersucht. Die entsprechende Reihenfolge der Textblöcke wird abgespeichert und lässt sich so ebenfalls auf die mit den Textblöcken verknüpften Buchstabentexte anwenden.

**[0032]** Anschließend wird ein geeignetes Format bestimmt, in dem die Buchstabentexte und bevorzugt die Abbildungen abgespeichert werden. Das geeignete Format wird so bestimmt, dass eine Breite der darin eingefüllten Buchstabentexte, also die Zeilenbreite, und bevorzugt eine weitere Breite der Abbildungen eine Breite eines Displays des Mobilfunkgeräts nicht übersteigt, oder in anderen Worten wird die Breite des geeigneten Formats so bestimmt, dass die Zeilenbreite gerade der Breite des Displays entspricht. Das geeignete Format weist ferner eine Schriftart und eine Schriftgröße auf, mit denen die Buchstabentexte auf dem Display dargestellt werden und die vom Nutzer gut lesbar sein sollen. Vorzugsweise kann der Nutzer die gewünschte Schriftart und Schriftgröße einstellen, wie beispielsweise Arial 10, Times 11, oder dergleichen. Nun werden die Buchstabentexte in der vorher bestimmten Reihenfolge in dem geeigneten Format abgespeichert, wobei entsprechende Zeilenumbrüche und bevorzugt Worttrennungen eingefügt werden. Sind Abbildungen in der Layout-Struktur erkannt und abgespeichert worden, werden die Abbildungen bevorzugt auch an den entsprechenden Stellen zwischen den Buchstabentexten in der entsprechenden Reihenfolge mit abgespeichert. Das geeignete Format ist entweder ein festes Format, das in einem Speicher abgespeichert ist, oder es ist ein variables Format, das Parameter berücksichtigt, die entweder vom Nutzer eingegeben werden oder vom Mobilfunkgerät abgerufen werden können, wie beispielsweise die gewünschte Schriftgröße und Schriftart im Display.

**[0033]** Fig. 2 zeigt eine Darstellung einer Datei **40**, die Buchstabentexte und Abbildungen umfasst, die in dem geeigneten Format in der bestimmten Reihenfolge abgespeichert worden sind. Dabei wird die erkannte erste Überschrift an Stelle **1** des ersten Bilds **30** gemäß Fig. 1a ganz oben in der Reihenfolge an Stelle **1** in der Datei **40** abgespeichert. Auf die Überschrift an der Stelle **1** folgt der Buchstabentext an Stelle **3** in der Datei **40**, der vom Textblock an der Stelle **3** aus dem ersten Bild **30** stammt. Als nächstes in der Reihenfolge wird die Abbildung **2** und dann Abbildung **4** in der Datei **40** abgespeichert, genauso, wie sie in der Reihenfolge im ersten Bild **30** erkannt worden sind. Nach dem Buchstabentext an Stelle **5** folgt der Buchstabentext an Stelle **6** in der Datei **40**. Eine zweite Länge des Buchstabentexts an Stelle **6** der Datei **40** ist länger als eine erste Länge des gleichen Buchstabentexts des Textblocks an der Stelle **6** im ersten Bild **30**, was darauf zurückzuführen ist, dass die Schriftgröße in der Datei **40** größer ist als im entsprechenden Textblock an der Stelle **6** im ersten Bild **30**.

**[0034]** An Stelle **7** in der Datei **40** folgt eine neue Überschrift, die im Textblock an Stelle **7** im ersten Bild **30** erkannt worden ist. Anschließend folgen die Buchstabentexte an den Stellen **8**, **9** und **10**, dann die Abbildung an Stelle **11** in der Datei **40** und weitere Buchstabentexte und Abbildungen, die in Fig. 2 nicht mehr dargestellt worden sind. Die Datei **40** kann beliebig kurz oder lang sein und umfasst eben die Buchstabentexte und bevorzugt die Abbildungen, die in den aufgenommenen Bildern **30**, **31** des Dokuments erkannt worden sind.

**[0035]** Die Datei **40** zeigt eine Breite **41** des Formats, und es ist dabei auch strichliert ein Ausschnitt **42** zu sehen, der jeweils am Display des Mobilfunkgeräts dargestellt wird. Wenn der Fachmann Bild **2** betrachtet, ist es für ihn leicht vorstellbar, dass bei einem Lesen des Dokuments in der oben beschriebenen konvertierten Form in dem geeigneten Format kein seitliches Scrollen mehr notwendig ist, sondern dass der Nutzer einfach auf- und abwärts zu scrollen braucht, um den Text fortlaufend zu lesen.

**[0036]** Die Datei **40** kann dabei auf dem Mobilfunkgerät abgespeichert werden, wobei die Datei **40** bevorzugt eine Standard-Text-Datei ohne Abbildungen ist, wie beispielsweise eine ASCII-Text-Datei, oder ein PDF-File bevorzugt mit Abbildungen ist, eine Microsoft-Word-Datei ist oder eine Datei in einem anderen Standard Format ist. In einem alternativen Verfahren kann die Datei **40** ebenso auf einem Server abgespeichert werden, wobei die vorangegangenen Verarbeitungsschritte ebenfalls auf einem Server stattgefunden haben.

**[0037]** In Fig. 3 ist dargestellt, dass die erste Seite des Dokuments, die in Fig. 1a durch das erste Bild **30**

abfotografiert worden ist, auch durch drei Bilder, bzw. durch ein drittes Bild **32**, ein viertes Bild **33** und ein fünftes Bild **34** abfotografiert werden kann. Durch ein Abfotografieren einer Dokumentenseite durch mehrere Bilder, kann für jedes Bild eine höhere Auflösung der jeweils darin enthaltenen Textblöcke erreicht werden. Durch die höhere Auflösung wird in der Regel eine bessere OCR-Erkennungsrate ermöglicht, was den höheren Aufwand je nach Dokumentenmaterial und Qualität der Kamera des Mobilfunkgeräts rechtfertigen kann. Im Beispiel von **Fig. 3** werden im dritten Bild **32** zwei Textblöcke und eine halbe Abbildung an Stelle **4** erkannt. Im folgenden vierten Bild **33** werden zwei Textblöcke und die Abbildung an der Stelle **4** erkannt.

**[0038]** Dies macht deutlich, dass der Reihenfolgeerkennungsalgorithmus zweckmäßigerweise auch überlappende gemeinsame Bereiche zwischen Bildern erkennt. Auf diese Weise kann in dem dritten Bild **32** und in dem vierten Bild **33** beispielsweise erkannt werden, dass nach dem Textblock an Stelle **3** aus dem dritten Bild **32** die Abbildung an Stelle **4** aus dem vierten Bild **33** folgt und nach der Abbildung an Stelle **4** der Textblock an Stelle **5** aus dem vierten Bild **33** folgt.

**[0039]** Die Reihenfolgeerkennung umfasst einen Algorithmus zum Erkennen bzw. Zusammensetzen benachbarter Bilder **32–34** (Stitching), um dadurch besser benachbarte und aufeinander folgende Textblöcke erkennen bzw. bestimmen zu können.

**[0040]** Erfindungsgemäß werden alle Verarbeitungsprozesse auf dem Mobilfunkgerät ausgeführt. Alternativ ist es aber auch denkbar, dass alle Verarbeitungsprozesse auf dem Mobilfunkgerät ausgeführt werden, ein Teil der Verarbeitungsprozesse auf einem externen Server ausgeführt werden. So ist es gut denkbar, dass die aufgenommenen Bilder **30–31** bzw. **32–34**, wie beispielsweise das erste Bild **30** und das zweite Bild **31**, sowie weitere Bilder **32–34** des Dokuments durch das Mobilfunkgerät an einen Server übertragen werden. Der Server verarbeitet die Bilder **30–31** bzw. **32–34** durch Erkennung der entsprechenden Layout-Strukturen, führt eine OCR Erkennung und eine Reihenfolgeerkennung durch und bildet die Datei **40**, wie oben beschrieben. Danach sendet der Server die Datei **40** an das Mobilfunkgerät zurück, wo die Datei **40** bzw. das Dokument dann auf dem Mobilfunkgerät in dem geeigneten Format angesehen werden kann. Dabei sind auch andere Teilprozesse als Auslagerung auf den Server denkbar, beispielsweise nur die OCR-Konvertierung oder die Reihenfolgeerkennung und dergleichen.

**[0041]** Ein Vorteil die Bilder **30–31** bzw. **32–34** zu einem Server zu senden besteht auch darin, dass auf dem Server dann zugleich auch eine Datei **40** in einem zweiten Format erstellt werden kann, wobei

das zweite Format im Wesentlichen für die Darstellung auf einem PC-Monitor oder für einen Ausdruck auf standardisiertem Druckerpapier geeignet ist. Das zweite Format weist dabei eine zweite Breite auf, die beispielsweise einer Breite des DIN A4 Formats oder des US Letter-Formats entspricht. Dabei kann beispielsweise auch eingestellt werden, ob das zweite Format an ein Längs- oder an ein Querformat angepasst sein soll oder eine andere Breite aufweisen soll. Bevorzugt kann dabei ebenso die Schriftart und/oder die Schriftgröße eingestellt werden.

**[0042]** Eine weitere Art der Konvertierung der aufgenommenen Bilder **30–31** bzw. **32–34** des Dokuments in ein anderes geeignetes Format ist zunächst wieder eine Erkennung der Layout-Struktur und der Reihenfolge der Textblöcke und bevorzugt der Abbildungen. Bei der weiteren Art werden die Textblöcke und bevorzugt die Abbildungen dann jedoch in dem anderen geeigneten Format als Bildteile zusammengesetzt. Dabei wird also keine OCR-Konvertierung der Textinformation durchgeführt, sondern es werden lediglich die Bildteile der Textblöcke und der Abbildungen, die aus dem jeweiligen Bild automatisch digital ausgeschnitten werden, in der bestimmten Reihenfolge angeordnet und zu einer Datei **40** zusammengefügt.

**[0043]** Weitere mögliche Ausbildungsformen sind in den folgenden Ansprüchen beschrieben.

**[0044]** Die In den Ansprüchen genannten Bezugszeichen dienen der besseren Verständlichkeit, beschränken aber die Ansprüche nicht auf die in den Figuren dargestellten Formen.

#### Bezugszeichenliste

|             |                    |
|-------------|--------------------|
| <b>1–19</b> | Stelle             |
| <b>30</b>   | erstes Bild        |
| <b>31</b>   | zweites Bild       |
| <b>32</b>   | drittes Bild       |
| <b>33</b>   | viertes Bild       |
| <b>34</b>   | fünftes Bild       |
| <b>40</b>   | Datei              |
| <b>41</b>   | Breite des Formats |
| <b>42</b>   | Ausschnitt         |

#### Patentansprüche

1. Verfahren zur Aufnahme eines Dokuments mit einem kamerabasierten Mobilfunkgerät und zur Konversion von Textinformation des Dokuments in ein Format zu einer geeigneten Darstellung auf dem Mobilfunkgerät, die folgenden auf dem Mobilfunkgerät ausgeführten Schritte umfassend:

a) Aufnehmen des Dokuments mit dem Mobilfunkgerät durch mindestens zwei Bilder, von denen benachbarte Bilder jeweils einen gemeinsamen überlappen-

den Bereich des Dokuments aufweisen, und Abspeichern der mindestens zwei Bilder;

b) Erkennen der benachbarten Bilder und Zusammensetzen zu einem jeweiligen zusammengesetzten Bild anhand der gemeinsamen überlappenden Bereiche des Dokuments (Stitching);

c) Erkennen einer jeweiligen Layout-Struktur mit mindestens einem Textblock in dem jeweiligen zusammengesetzten Bild;

d) Erkennen von Buchstabentext in dem jeweiligen Textblock durch Anwenden eines Optical Character Recognition Algorithmus (OCR) und Abspeichern des jeweiligen Buchstabentexts zu dem jeweiligen Textblock;

e) Bestimmen und Abspeichern einer Reihenfolge der jeweiligen Textblöcke in dem Dokument unter Berücksichtigung der jeweiligen Layout-Struktur;

f) Bestimmen einer Breite eines Displays des Mobilfunkgeräts durch automatisches Abrufen von entsprechenden Parametern vom Mobilfunkgerät und Bestimmen des Formats für die Darstellung der Buchstabentexte, das an die Breite des Displays angepasst ist, um beim Lesen der Buchstabentexte auf dem Display nur vertikal scrollen zu brauchen;

g) Erzeugen einer Datei (**40**) in dem Format mit den jeweiligen Buchstabentexten in der bestimmten Reihenfolge der jeweiligen Textblöcke;

h) Bereitstellen der Datei (**40**) für das Mobilfunkgerät zur Darstellung auf dessen Display.

2. Verfahren gemäß Anspruch 1, wobei in Schritt c) in der Layout-Struktur auch Abbildungen erkannt werden, wobei in Schritt e) die Abbildungen in der Reihenfolge der Textblöcke mit einbezogen werden, so dass eine Reihenfolge aus den Textblöcken und den Abbildungen erzeugt wird, und wobei in Schritt g) die Abbildungen der Breite des Displays angepasst werden, so dass bei der Darstellung auf dem Mobilfunkgerät nur vertikal gescrollt werden muss.

3. Verfahren gemäß Anspruch 1 oder 2, wobei die Erkennung der Layout-Struktur auf Basis einer Histogramm-Analyse der Schwärzungsverteilung in dem jeweiligen Bild erfolgt.

4. Verfahren gemäß einem der vorstehenden Ansprüche 1–3, wobei in Schritt e) bei der Bestimmung der Reihenfolge der jeweiligen Textblöcke ferner – die Reihenfolge untereinander stehender Textblöcke berücksichtigt wird;

– eine syntaktische Verbindung zwischen nebeneinanderstehenden Textblöcken analysiert wird, wobei sich beispielsweise nebeneinanderstehende Textblöcke syntaktisch ergänzen und so mit hoher Wahrscheinlichkeit auf einander folgen.

5. Verfahren gemäß einem der vorstehenden Ansprüche 1–4, wobei in Schritt f) das Format zudem so bestimmt wird, indem von einem Nutzer eine gewünschte Schriftart und Schriftgröße für die Darstel-

lung der Buchstabentexte auf dem Display einstellbar ist.

6. Verfahren gemäß einem der vorstehenden Ansprüche 1–5, wobei in Schritt e) bei der Bestimmung der Reihenfolge der jeweiligen Textblöcke ferner – bei benachbarten Textblöcken eine Histogramm-Wortanalyse oder eine Kantenfilterung mit anschließender Cluster-Bestimmung vorgenommen und berücksichtigt wird, um festzustellen, ob diese dem gleichen oder einem nächsten Themengebiet zugehören.

7. Verfahren gemäß Anspruch 6, wobei bei der Histogramm-Wortanalyse Schlüsselworte wie beispielsweise die Schlüsselworte einer Überschrift eines jeweiligen Artikels des Dokuments verwendet werden.

8. Verfahren gemäß einem der vorstehenden Ansprüche 1–7, wobei in Schritt e) bei der Bestimmung der Reihenfolge der jeweiligen Textblöcke ferner – auch die Abfolge der aufeinander folgenden Bilder berücksichtigt wird.

9. Verfahren gemäß einem der vorstehenden Ansprüche 1–8, wobei in Schritt e) bei der Bestimmung der Reihenfolge der jeweiligen Textblöcke ferner – auch semantisch zusammengehörende Textblöcke aufeinander folgender Bilder untersucht und berücksichtigt werden.

10. Verfahren gemäß einem der vorstehenden Ansprüche 1–9, wobei in Schritt f) das bestimmte Format ein PDF Format mit einer Breite ist, die der Breite des Displays des Mobilfunkgeräts entspricht.

11. Verfahren gemäß einem der Ansprüche 1–9, wobei in Schritt f) das bestimmte Format ein Standard-Text-Format mit einer Breite ist, die der Breite des Displays des Mobilfunkgeräts entspricht.

12. Verfahren gemäß einem der vorstehenden Ansprüche 1–11, wobei die jeweiligen Buchstabentexte in der bestimmten Reihenfolge der jeweiligen Textblöcke zudem in einem zweiten Format abgespeichert werden, wobei das zweite Format ein gut ausdruckbares Standardpapierformat ist und gut auf PC-Monitoren zum Lesen von Text verwendbar ist, wie beispielsweise das DIN A4 Format oder das US Letter-Format.

13. Kamera basiertes Mobilfunkgerät zur Konvertierung eines Textinhalts aus Bildern, die vom Mobilfunkgerät von einem Dokument aufgenommen worden sind, wobei der Textinhalt in ein Format konvertiert wird, das für eine Darstellung auf einem Display des Mobilfunkgeräts geeignet ist, umfassend:

a) Erste Mittel zum Aufnehmen der Bilder mit dem Mobilfunkgerät, wobei die Bilder jeweils mindestens einen Textblock enthalten und zum Detektieren von

- entsprechenden Parametern zu einer Breite des Displays des Mobilfunkgeräts;
- b) Zweite Mittel zu einem Erkennen von Überlappungsbereichen in aufeinanderfolgenden benachbarten Bildern des Dokuments und zu einem entsprechenden Stitching der benachbarten Bilder zu einem zusammengesetzten Bild, basierend auf den Überlappungsbereichen;
- c) Dritte Mittel zum Erkennen einer jeweiligen Layout-Struktur mit dem mindestens einen Textblock in dem jeweiligen zusammengesetzten Bild;
- d) Vierte Mittel zum Erkennen von Buchstabentext in dem mindestens einen Textblock durch Anwenden eines Optical Character Recognition Algorithmus (OCR) und zum Abspeichern des jeweiligen Buchstabentexts des jeweiligen Textblocks;
- e) Fünfte Mittel zum Bestimmen und Abspeichern einer Reihenfolge der jeweiligen Textblöcke des Dokuments unter Berücksichtigung der jeweiligen Layout-Struktur;
- f) Sechste Mittel zur Bestimmung der Breite des Displays des Mobilfunkgeräts anhand der detektierten Parameter zur Breite des Displays, und Auswahl und automatische Anpassung des Formats für die Darstellung der Buchstabentexte an die Breite des Displays, um beim Lesen der Buchstabentexte auf dem Display nur vertikal scrollen zu brauchen;
- g) Siebte Mittel zur Erzeugung einer Datei (**40**) in dem Format mit den jeweiligen Buchstabentexten in der bestimmten Reihenfolge der jeweiligen Textblöcke;
- h) Achte Mittel zum Abspeichern der erzeugten Datei (**40**) im Mobilfunkgerät.

Es folgen 3 Seiten Zeichnungen

Anhängende Zeichnungen

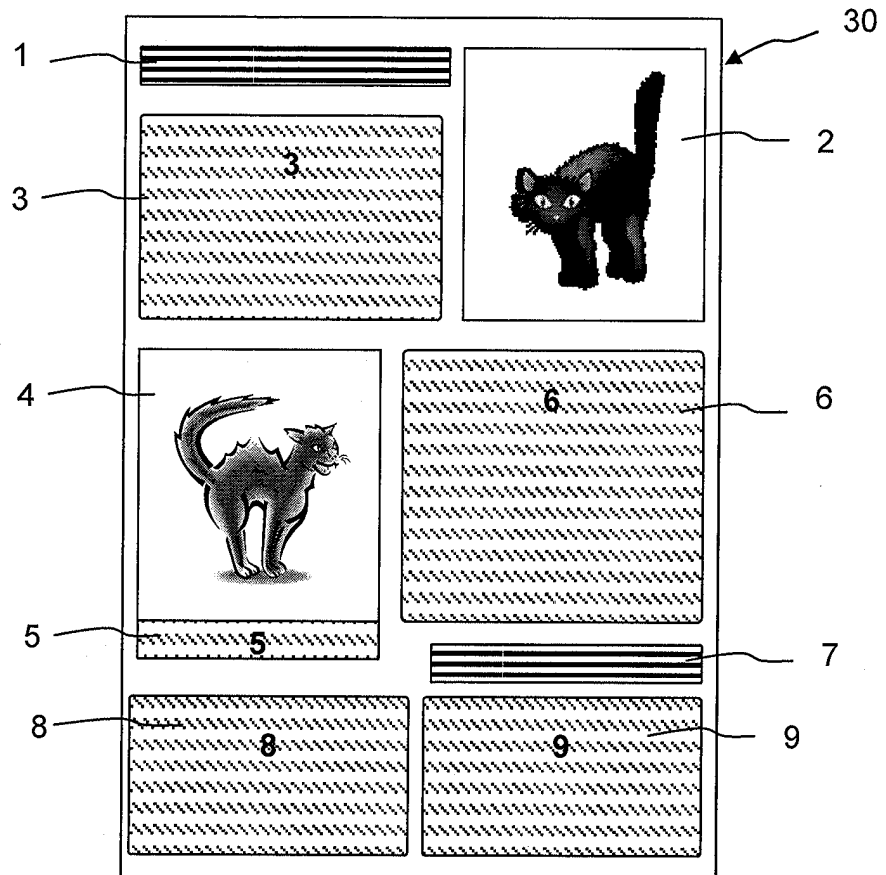


Fig. 1a

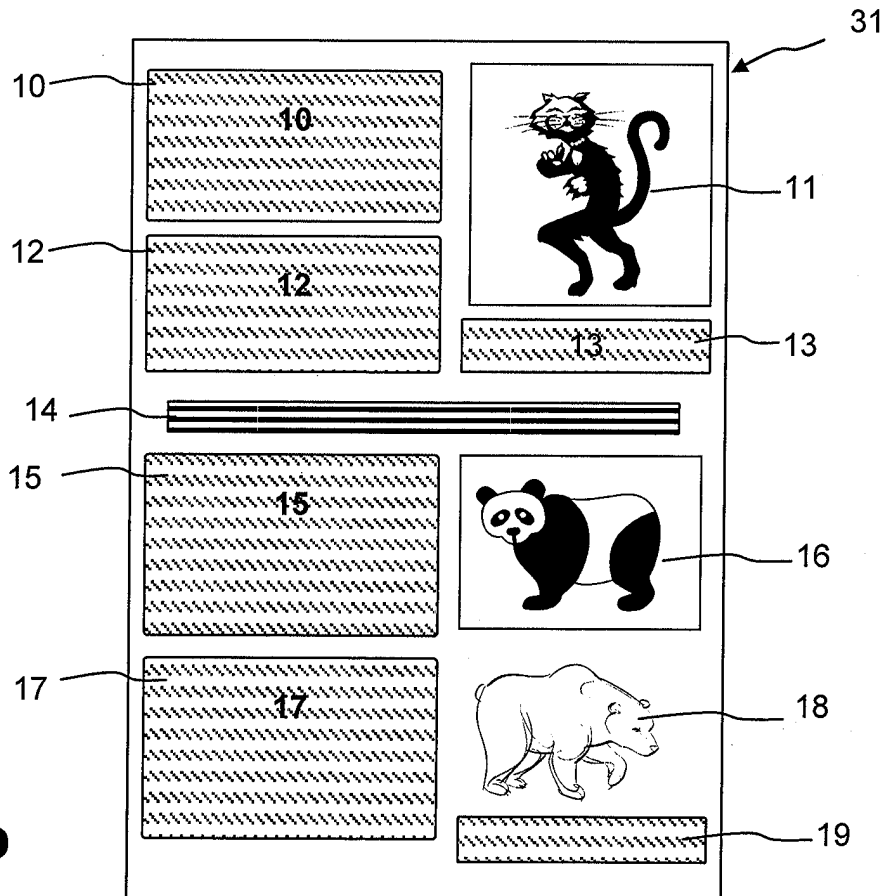


Fig. 1b

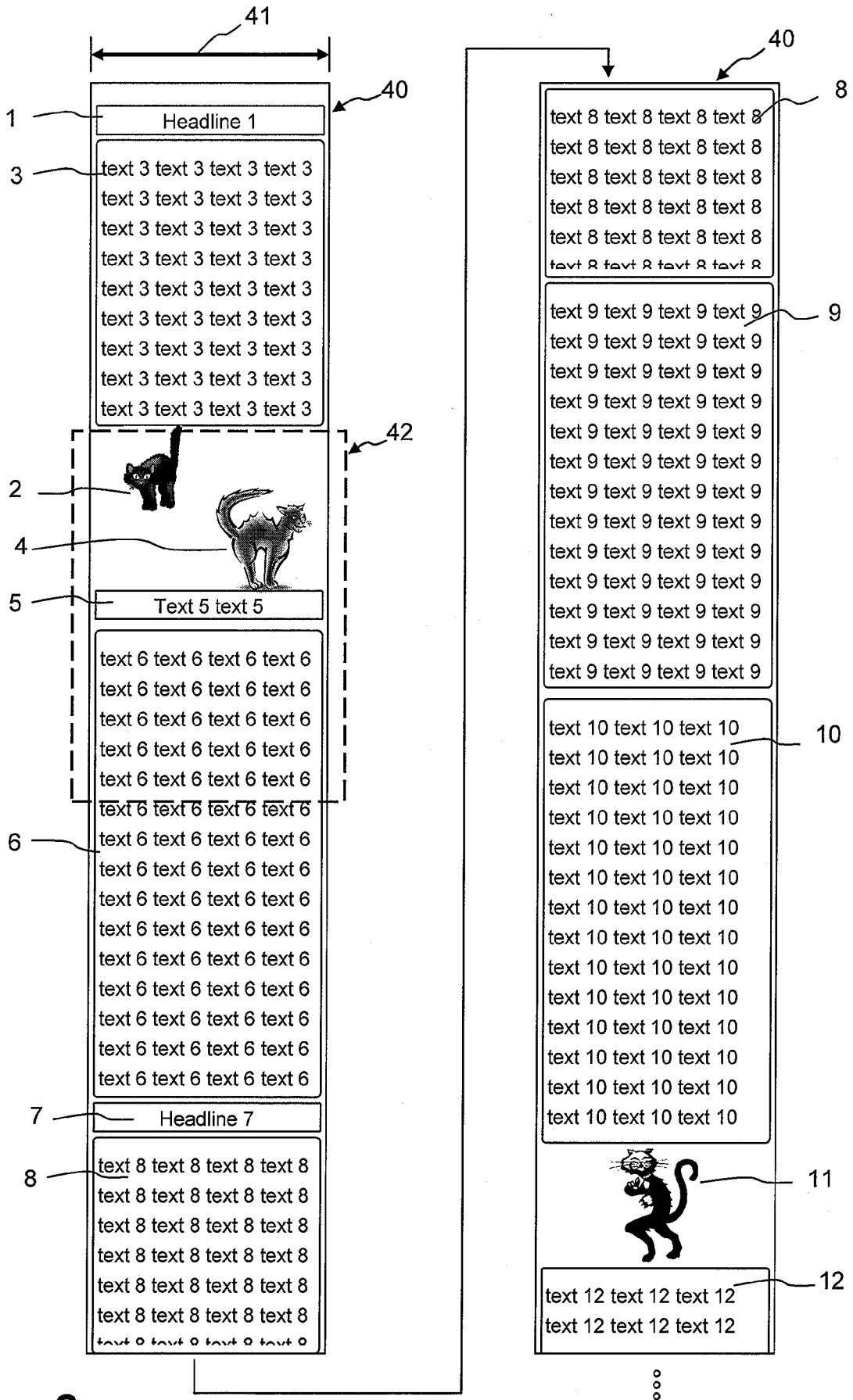


Fig. 2

