



(51) International Patent Classification:
G06K 9/20 (2006.01)

(21) International Application Number:
PCT/GB2010/001661

(22) International Filing Date:
1 September 2010 (01.09.2010)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
0915375.0 3 September 2009 (03.09.2009) GB

(71) Applicant (for all designated States except US): **DRS DATA SERVICES LIMITED** [GB/GB]; 1 Danbury Court, Linford Wood, Milton, Keynes MK14 6LR (GB).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **PU, Zhang** [GB/GB]; 31 Wardle Place, Oldbrook, Milton, Keynes MK6 2XS (GB). **CHASE, Julian**; 29 Sheppars Close, Newport Pagnell, Milton, Keynes MK16 0QN (GB). **BEATS, Stephen, Mark** [GB/GB]; 4 Redland Close, Steeple Claydon, Buckinghamshire MK18 2LH (GB). **HUDSON, Graham, Charles** [GB/GB]; "Monticello", 60 Cranmore Gardens, Aldeshot, Hampshire GU11 3BG (GB). **WHITELEY, Sean, Edward** [GB/GB]; 181 Grangeway, Rushden, Northamptonshire NN10 3GB (GB).

(74) Agents: **LEEMING, John, Gerard** et al.; J.A. KEMP & CO., 14 South Square, Gray's Inn, London WC1R 5JJ (GB).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

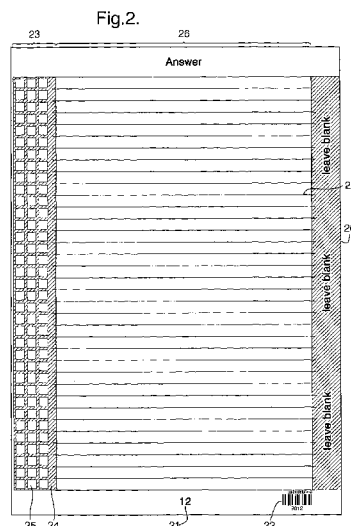
(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— with international search report (Art. 21(3))

[Continued on next page]

(54) Title: METHOD AND APPARATUS FOR SEGMENTING IMAGES



(57) Abstract: A method and an apparatus for obtaining segmentation information of a document image, wherein the document has a first and second section (23, 26), is disclosed. The method comprises the steps of : searching the first section (23) for a first indicating text and determining the location thereof; in the second section (26), setting a first position as the start of a segment based on the location of the first indicating text; searching the first section, downstream from the first indicating text, for a second indicating text and determining the location thereof; and in the second section, setting a second position as the end of the segment based on the location of the second indicating text. Optionally, the second section adjacent to the first and second indicating texts is searched for first and second blank areas respectively. Appropriate first and second blank areas are set as the start and end of the segment respectively.



— *before the expiration of the time limit for amending the
claims and to be republished in the event of receipt of*

amendments (Rule 48.2(h))

METHOD AND APPARATUS FOR SEGMENTING IMAGES

The present application claims priority from British Patent Application number 0915375.0 filed 3 September 2009, the entire contents of which are
5 incorporated herein by reference. This invention relates to a method and an apparatus for obtaining segment information of an image.

In the examination marking industry, it is commonplace for multiple examiners to be used to mark an exam answer document from a single exam candidate. Utilising multiple examiners per document can be more efficient than
10 utilising a single examiner per document as an examiner can focus on marking answers to only one or a few questions from a large number of candidates rather than answers for all questions in a document for a comparatively small number of candidates. This allows an examiner to specialise and be consistent in marking an answer over the large number of documents. This is especially true in circumstances
15 where optional questions are included as part of the examination and therefore the document could consist of a wide variety of answers to questions. It also means that parts of individual documents are marked by a number of examiners, which lessens marking bias across the whole document.

Currently marking using multiple examiners is achieved by sending each of
20 the examiners a copy of the entire answer document. In an alternative method currently utilised, the exam candidates are required to provide their answers in a fixed area within a page of a document so that each answer can be separated in the scanned image and sent to the relevant examiner. This has the disadvantage of constraining the exam candidate to provide an answer which must fit within a fixed area and so is
25 only suitable for certain uses and certain types of exam questions. This method of distributing answers to examiners is not possible with free form answers, where the length of an answer is unpredictable, the questions can be answered in a random order and the answers can be broken into separate sections.

A further problem faced in the examination marking industry when imaging
30 the answers is the difficulty in storing, organizing and tracking the images of individual answers simply and inexpensively.

Accuracy and reliability of the examiners are also of concern in the industry. Marking an examination script from an image stored in a server or online database can be used

as part of a quality control system, particularly if the individual answers can be identified from within the individual examination document.

Thus it is desirable to provide a scanned image of a free form answer booklet and separate the image so that only the relevant parts of an answer is sent to the relevant
5 examiner.

Thus it is desirable to provide an inexpensive, efficient system for storing, organizing and tracking the exam question response images.

The present invention provides a method of obtaining segment information of a document image, wherein the document has a first section and a second section,

10 comprising the steps of: searching the first section for a first indicating text and determining the location thereof; in the second section, setting a first position as the start of a segment based on the location of the first indicating text; searching the first section, downstream from the first indicating text, for a second indicating text and determining the location thereof; and in the second section, setting a second position
15 as the end of the segment based on the location of the second indicating text.

In an embodiment of the invention, the method described above further comprises the step of searching the second section adjacent to or in the vicinity of the location of the first indicating text for a first blank area having a predetermined height and width and not containing any free text, wherein the first blank area is set as the start of the
20 segment.

In another embodiment of the invention, the method described above further comprises, if a first blank area adjacent to or in the vicinity of the first indicating text is not found, searching the second section in an upstream direction for a blank area and setting the first found blank area as the start of the segment.

25 In this embodiment, the start of a free form text is determined by searching for a blank area across the second section and close to or above the location of the first indicating text (e.g. a question number). This has the advantage that any free form text written above the horizontal location of the first indicating text is not lost when setting the start of the segment. Thus, for example, if a writer has sloped writing, where the start
30 of a line of text may have started below the horizontal location of the question number but the end of the line may have ended above the horizontal location of the first indicating text, by searching for a blank area across the second section, the start location of the segment will account for sloped writing.

In another embodiment of the invention, the methods described above further comprise the step of searching the second section adjacent to or in the vicinity of the location of the second indicating text for a second blank area having a predetermined height and width and not containing any free text, wherein the second blank area is set
5 as the end of the segment.

In another embodiment of the invention, the method described above further comprises the step of searching in an upstream direction for further blank areas above the second blank area and setting the end of the segment as the location of a blank area that is adjacent to the free text or the start of the segment.

10 In another embodiment of the invention, in the method described above, if a second blank area adjacent to or in the vicinity of the second indicating text is not found, searching the second section in an upstream direction for a blank area and setting the first found blank area as the end of the segment.

The end of a free form answer is determined by searching for a blank area
15 across the second section and close to or above the location of the second indicating text (e.g. another, subsequently answered question number). The writer may leave spaces within a segment (e.g. when answering a question), thus a blank area in the second section may not indicate the end of a segment. The location of the start of a subsequent segment (e.g. an answer to another question) is likely to be close to the
20 location of the end of the previous segment. Thus searching for a blank area close to or above the location of the second indicating text in the first section provides greater certainty that such a blank area is the end of the segment.

In another embodiment of the invention, in the methods described above if a second indicating text is not found, searching the second section for a blank area that persists
25 to the end of the document and setting the start location of the blank area as the end of the segment.

In another embodiment of the invention, in the methods described above the first section has a predetermined format, and the step of searching the first section for a first indicating text comprises the steps of: searching for any mark in the first section
30 that is different to the predetermined format; applying character recognition to said mark to derive a character; comparing the character obtained from the mark to a list of reference characters; and if the derived character relates to a reference character, setting that character as the first indicating text.

In another embodiment of the invention, in the methods described above the characters that are horizontally adjacent to each other form a character string and if the character string relates to reference characters, setting that character string as the first indicating text.

5 In another embodiment of the invention, in the methods described above the first section has a predetermined format, and the step of searching the first section for a second indicating text comprises the steps of: searching for any mark in the first section that is different to the predetermined format; applying character recognition to said mark to derive a character; comparing a character obtained from the mark to a list
10 of reference characters; and if the character relates to a reference character, setting the character as the second indicating text.

In another embodiment of the invention, in the method described above, the derived characters that are horizontally adjacent to each other form a character string and if the character string relates to reference characters, setting that character string
15 as the first or second indicating text.

The predetermined format is known. Thus any marks made by the writer, which are not part of the predetermined format, are easily found. Comparing a recognised character against a list of reference characters (e.g question numbers) allows greater reliably and certainty that the mark and recognised character is a reference character
20 and not an erroneous mark made by the exam candidate.

In another embodiment of the invention, in the methods described above more than one character is recognised from the mark.

In another embodiment of the invention, in the methods described above when applying character recognition, a confidence level that a character is recognised from
25 the mark is higher than a confidence level threshold.

Setting a confidence level threshold allows a user to specify the reliability and certainty of the Intelligent Character recognition (ICR). The user may accept more exceptions in order to reduce the erroneous recognition of characters, or vice versa.

In another embodiment of the invention, in the methods described above the
30 predetermined format comprises a plurality of regularly spaced response areas that are distinguished from the surroundings.

In another embodiment of the invention, in the methods described above the surroundings are shaded and the response areas are unshaded.

The response areas indicate to the writer where the reference character (e.g. exam question number) should be written. The surroundings may be shaded in a certain colour that is invisible to the scanner. The searching means may remove the shading colour from the image, leaving a clear background for character recognition.

- 5 This clearer image allows easier searching and recognition of the marks.

In another embodiment of the invention, in the methods described above only the response areas are searched for marks.

Searching the response areas only and not the whole of the first section can increase the speed and reliability of obtaining segmentation information.

- 10 In another embodiment of the invention, in the methods described above the document has one or more pages, each page having a sequence indicator which indicates the downstream direction.

- When the writer reaches the end of a page, the writer can continue to write on the following page without the segment being ended. A barcode, for example, can
15 indicate the sequence of pages in a document. Thus the writer is not restricted in the length of text (e.g. an answer to a question) he or she may write. The writer may use more than one page for a segment of text.

In another embodiment of the invention, in the methods described above the first section is adjacent to and left of the second section.

- 20 In another embodiment of the invention, in the methods described above the first section is adjacent to and right of the second section.

In another embodiment of the invention, in the methods described above the first and second indicating text are handwritten. In another embodiment of the invention, in the methods described above the free text is handwritten.

- 25 In another embodiment of the invention, in the methods described above the segment information comprises information related to the first indicating text, the location of the start of the segment and the location of the end of the segment.

- The invention also provides a second method comprising the steps of: scanning a document; while scanning the document, obtaining segment information of the
30 document; compressing an image of the document; and storing the segment information and the compressed image related to the document.

Effective character recognition such as ICR requires high resolution images. Thus, segment information is obtained while scanning the document thereby allowing character recognition to be carried out prior to compressing the image. Storing a

compressed image and associated segment information data requires less storage space than storing a high resolution image.

In another embodiment of the invention, in the second method described above the step of obtaining segment information is done in a buffer memory of a scanner.

5 A temporary high resolution image of the scanned document is stored in the buffer memory of the scanner. Character recognition can be carried out with greater accuracy on the high resolution image compared with a compressed, low resolution image that is used for storage and transmission purposes. Thus segment information can be obtained with greater accuracy from the high resolution image. Furthermore,
10 as the processing for obtaining segment information can be carried out by the scanner, there is no requirement to carry out additional processing of the image by a computer or server connected to the scanner, thus saving time and processing resources.

In another embodiment of the invention, in the second method described above the step of obtaining segment information comprises a method according to the method
15 described above.

In another embodiment of the invention, the second method described above further comprises a method of displaying a segment, the method of displaying a segment comprising the steps of: requesting a segment; obtaining only parts of the compressed image which correspond to the requested segment; and displaying said
20 parts.

Sending and displaying a compressed image of a segment of text only (e.g. an answer), rather than sending and displaying an uncompressed image or an image of the whole document, requires less data to be sent over a network and thus saving on network resources and enabling faster transfer of the image.

25 In another embodiment of the invention, the second method described above further comprises the step of assigning and submitting, to a database, data related to the displayed parts.

Assigning, submitting and storing data for each segment (e.g. marks for each individual answer), rather than a whole document, allows efficient storage and
30 retrieval such data. For example, the marks assigned to an answer can be easily retrieved and compared to marks assigned by other examiners for the same question for quality control purposes.

In another embodiment of the invention, in the second method described above said parts are displayed in a sequence according to the segment information and the downstream direction.

5 In another embodiment of the invention, in the second method described above the sequence of parts are concatenated and displayed as a single image.

The writer is free to write text associated with an indicating text in different areas of the document (e.g. an answer to a question can be written in different areas of the document). However, viewing text with parts dispersed in different areas is inefficient and difficult (e.g. examining an answer is difficult if it is dispersed in
10 different areas). Thus, from the segment information data, the present invention is able to collate the different parts and display them in sequence or as a single image. This allows, for example, easier and more efficient marking of an answer.

In another embodiment of the invention, in the second method described above the step of requesting and displaying can occur at a location remote from a location of the
15 stored segment information and the stored compressed image.

Storing the compressed images and segment information data in a server allows easy access over a network such as the internet. Thus, for example, multiple examiners can be utilised.

20 In another embodiment of the invention, in the methods described above the document is an examination script.

The invention also provides an apparatus for obtaining segment information of a document image, wherein the document has a first section and a second section, the apparatus comprising: a means for searching the first section; and a means for setting segment information, wherein: the means for searching the first section searches for a
25 first indicating text and determines the location thereof; the means for setting segment information sets, based on the location of the first indicating text, a first position in the second section as the start of a segment; the means for searching the first section searches, downstream from the first indicating text, for a second indicating text and determines the location thereof; and the means for setting segment information sets,
30 based on the location of the second indicating text, a second position in the second section as the end of the segment.

In another embodiment of the invention, the apparatus described above further comprising a means for searching the second section, wherein: the means for searching the second section searches adjacent to or in the vicinity of the location of

the first indicating text for a first blank area having a predetermined height and width and not containing any free text; and the means for setting the segment information sets the first blank area as the start of the segment.

5 In another embodiment of the invention, in the apparatus described above if the means for searching the second section can not find a first blank area adjacent to or in the vicinity of the first indicating text, the means for searching the second section searches in an upstream direction for a blank area and the means for setting the segment information sets the first found blank area as the start of the segment.

10 In another embodiment of the invention, in the apparatus described above the means for searching the second section searches adjacent to or in the vicinity of the location of the second indicating text for a second blank area having a predetermined height and width and not containing any free text; and the means for setting the segment information sets the second blank area as the end of the segment.

15 In another embodiment of the invention, in the apparatus described above, the means for searching the second section searches in an upstream direction for further blank areas above the second blank area and the means for setting the segment information sets the end of the segment as the location of a blank area that is adjacent to the free text or the start of the segment.

20 In another embodiment of the invention, in the apparatus described above if the means for searching the second section can not find a second blank area adjacent to or in the vicinity of the second indicating text, the means for searching the second section searches in an upstream direction for a blank area and the means for setting the segment information sets the first found blank area as the end of the segment.

25 In another embodiment of the invention, in the apparatus described above if the means for searching the second section can not find a blank area below the horizontal location of the first indicating text, then the means for searching the second section searches, in a downstream direction from the second indicating text for a blank area and the means for setting the segment information sets the first found blank area as the end of the segment.

30 In another embodiment of the invention, in the apparatus described above if the means for searching the first section does not find a second indicating text, the means for searching the second section searches for a blank area that persists to the end of the document and the means for setting segment information sets the start location of the blank area as the end of the segment.

In another embodiment of the invention, in the apparatus described above the first section has a predetermined format, and the means for searching the first section comprises: a means for searching for any mark that is different to the predetermined format; a means for character recognition of said mark; a means for comparing a
5 character obtained from the character recognition means to a list of reference characters; and a means for setting the character as the first or second indicating text if the character relates to a reference character.

In another embodiment of the invention, in the apparatus described above, the derived characters that are horizontally adjacent to each other form a character string
10 and if the character string relates to reference characters, the means for setting the character sets that character string as the first or second indicating text.

In another embodiment of the invention, in the apparatus described above the means for character recognition recognises more than one character from the mark.

In another embodiment of the invention, in the apparatus described above the means
15 for character recognition comprises a confidence level threshold, wherein a confidence level that a character is recognised from the mark is higher than the confidence level threshold.

In another embodiment of the invention, in the apparatus described above the predetermined format comprises a plurality of regularly spaced response areas that are
20 distinguished from the surroundings.

In another embodiment of the invention, in the apparatus described above wherein surroundings are shaded and the response areas are unshaded.

In another embodiment of the invention, in the apparatus described above the means for searching the first section only searches the response areas for marks.

25 In another embodiment of the invention, in the apparatus described above the document has one or more pages and each page has a sequence indicator which indicates the downstream direction.

In another embodiment of the invention, in the apparatus described above the first section is adjacent to and left of the second section.

30 In another embodiment of the invention, in the apparatus described above the first section is adjacent to and right of the second section

In another embodiment of the invention, in the apparatus described above the first and second indicating text are handwritten.

In another embodiment of the invention, in the apparatus described above the free text is handwritten.

In another embodiment of the invention, in the apparatus described above the segment information comprises information related to the first indicating text, the

5 location of the start of the segment and the location of the end of the segment.

The invention also provides for a second apparatus comprising: a means for scanning a document; a means for obtaining segment information of the scanned document; a means for compressing an image of the scanned document; and a means for storing the segment information and the compressed image related to the scanned document, 10 wherein the segment information is obtained while scanning the document.

In another embodiment of the invention, in the second apparatus described above the segment information is obtained from an image in a buffer memory of a scanner.

In another embodiment of the invention, in the second apparatus described above the means for obtaining segment information comprises an apparatus according to any 15 one of the apparatuses described above.

In another embodiment of the invention, the second apparatus described above further comprises: a means for requesting a segment; a means for obtaining parts of the compressed image which correspond to the requested segment; and a means for displaying said parts.

20 In another embodiment of the invention, the second apparatus further comprising a means for assigning and submitting, to a database, data related to the displayed parts.

In another embodiment of the invention, in the second apparatus described above said parts are displayed in a sequence according to the segment information and 25 the downstream direction.

In another embodiment of the invention, in the second apparatus described above the sequence of parts are concatenated and displayed as a single image.

In another embodiment of the invention, in the second apparatus described above means for requesting a segment and the means for displaying the parts are at a 30 location remote from a location of the means for storing the segment information and the compressed image related to the scanned document.

In another embodiment of the invention, in the apparatuses described above the document is an examination script.

The invention also provides for a system comprising: an apparatus, connected to a network, comprising: a means for scanning a document; a means for obtaining segment information of the scanned document while scanning the document according to the apparatus described above; and a means for compressing an image of the scanned document; a network device, connected to the network, comprising: a means for storing the segment information and the compressed image related to the scanned document; and a client computer, connected to the network, comprising: a means for requesting a segment; a means for obtaining parts of the compressed image which correspond to the requested segment from the apparatus; and a means for displaying said parts.

In another embodiment of the invention, in the system described above the client computer further comprises a means for assigning and submitting, to the apparatus, data related to the displayed parts.

The invention also provides for a computer program comprising instructions that when executed by a computer system, cause the computer system to perform a method according to any one of the above methods.

The present invention will be described further below with reference to exemplary embodiments and the accompanying schematic drawings, in which:

Figure 1 depicts an example of a front page of an exam document;
Figure 2 depicts an example of an answer page of an exam document;
Figure 3 depicts a system according to an embodiment of the present invention;

Figure 4 depicts the architecture of a scanner according to an embodiment of the present invention; and
Figures 5A to 5D depict a process of segmenting an image according to an embodiment of the present invention.

In the drawings, like reference numerals indicate like parts.

Figure 1 depicts an example of the front page 10 of a type of exam document used in the invention. The document has been designed to maximise the ability to read characters using automated data capture techniques. The document includes features that permit the document to be analysed when it is passed through an imaging scanner. The front page comprises predetermined formatted areas 11 (which can, for example, be shaded e.g. coloured), each with a number of response areas 12 (which

can, for example, be unshaded) where the exam candidate can write the details of the examination being attempted. The details on the front page can include the exam paper number, the candidate number, the examination date, the location number and any other details. These written details can be read by a variety of means, which may

5 be ICR software, and inputted into a central database. The front page also comprises an indicator 13. This indicator can be a barcode or any other machine-readable image. The indicator can be used to identify the orientation, the type of document, the number of pages in the document and any other information. The front page also comprises the document version indicator 14, which can be a barcode and identifies

10 the specific document type. These indicators are linked to the written details, such as the candidate number and the exam paper number, stored in the central database. Figure 2 depicts an example of an answer page 20 that is part of the exam document. Each page can comprise a page number 21 to indicate the page number and order of the pages in the document. The pages can comprise a barcode 22, or any other type of

15 machine-readable image, to indicate the page number, the document the page belongs to and any other information.

The page comprises a first section 23 which is used to enter the characters which indicates the exam question being answered. The characters can be numbers, symbols or text in any language (for example, it is possible to use Latin, Greek,

20 Cyrillic, Arabic or Asian characters). The first section 23 can comprise a predetermined formatted area 24 (which can, for example, be shaded) and a number of response areas 25 (which can, for example, be unshaded) where the characters which indicates the exam question can be entered. The response areas are shown as boxes in Figure 2, however the response areas can be of any shape. The exam candidate enters

25 the exam question details in the response areas. Figure 2 shows three boxes per line where the exam question details are entered, however one or more response areas per line can be utilised. Some variant designs involve all or some of the response areas of some lines being removed.

Adjacent to the first section 23 is a second section 26. The second section 26

30 can have a larger area than the first section. The second section can comprise horizontal lines 27 that are vertically spaced to correspond to the top and bottom of the response areas in the first section. The second section can also comprise horizontal lines in the form of a musical stave. The second section can also be a grid design (for example, graph paper). The second section can also be blank with no

lines. Free form text (e.g. an answer to an exam question that can be hand written or typed or a diagram) is entered into the second section 26.

The exam candidate is instructed to write a question number in the response areas 25 in the first section 23 and begin writing an answer to the question on a line in the second section 26 adjacent to the question number in the first section 23. The exam candidate writes the numbers in response areas 25 that are created within the predetermined format area 24 to give structure to the blank document. The predetermined format area 24 can be removed from the image when scanned to leave a clear image of the character that needs to be recognised. As the response areas 25 are regularly repeated down the answer page and the exam candidate writes which answer is being attempted adjacent to the beginning of the answer, the beginning of each answer can be electronically deduced. There is some tolerance in the system in case the exam candidate makes a mistake in writing the number and uses nearby space up or down the page. The system can be configured to include all of the answer text below detection of a clear line in the second section 26. The candidates are instructed to leave a clear line between each question attempted to assist facilitation this procedure.

The document can be designed in a variety of formats depending upon the questions being used and the language in which the examination is being completed. For example, the first section can be placed on the right hand side of the second section. This is advantageous for languages where text is read and written from right to left.

The document can be scanned to create an image and associated metadata, which can be used to send the image of that answer to an appropriate qualified examiner (note that it is only that answer, not the entire document). The examiner is given tools to assist in marking the candidate's work from the image and to pass those marks and comments back to a central database, where they can be collated and ultimately exported as the candidate's mark.

Figure 3 depicts a system comprising a scanning device 30, a central database 31 connected to the scanning device 30, a network administering device 32 connected to the central database 31 via a network 33 and a client computer 34 connected to the network administering device 32 via a network 33. The scanning device 30, central database 31 and network administering device 32 is known as the technical infrastructure. The technical infrastructure can be a single, combined device. The

central database 31 holds a set of operational instructions (a computer program) to facilitate the purposeful running of the technical infrastructure. More than one scanning device 30 and more than one client computer can be connected to the central database 31 or network administration device 32 at any one time.

5 The scanning device 30 scans the front page and the answer pages of the exam document. While scanning the answer pages in the document, the scanning device 30 performs a segmentation routine, where the answer pages are segmented according to the exam questions in the first section 23 and the written answers in the second section 26. Segmentation information data and the image data of the scanned pages
10 are sent to the central database 31. The scanning device 30 or the central database 31 can compress the image data to reduce the storage space required to store the image data in the database. The segmentation information related to the image data can be correspondingly related to the compressed image data.

 The segment information data can comprise data that indicates the document
15 scanned, the pages scanned, the location of the start and end of each segment, the question number associated with each segment, the exam candidate, the examination paper, the date and time of scanning and the filing reference of the original physical document.

 A client computer 34 can be connected to the technical infrastructure via a
20 local network or the internet. An examiner, using the client computer 34, can request an answer to a specified question to mark. The technical infrastructure identifies, from the segmentation information data, which parts of the images or compressed images correspond to the requested answers. The identified parts of the images or compressed images are sent to the client computer 34. The client computer 34 then
25 displays the answer to the question. The examiner can assign a mark, comment or special instruction to the answer and send this to the technical infrastructure.
 If an exam candidate has answered a question in two or more locations (e.g. different parts of a page or different pages in the document) the nonadjacent parts will be joined together to display a single image. The parts are displayed in an order
30 corresponding to the downstream direction. The downstream direction is the direction from the top of a page to the bottom of a page and then continuing to the top of the next page in the sequence of pages in the document.

The displayed image of a segment can comprise the question number from the first section and the written answer given in the second section. Alternatively, the image of the segment comprises the answer from the second section only.

ICR requires high resolution images. Thus performing segmentation in real time has the benefit of not having to transfer and store large volumes of high resolution image data. A high resolution image is captured into buffer memory of the scanner and is segmented to obtain segmentation information data. The high resolution image data can then be compressed into low resolution image data and stored in the central database.

Figure 4 depicts the architecture of the scanner. The scanner comprises an optical scanning means 40 to convert the pages of the document into image data, a buffer memory 41 which temporarily stores the image data, a processor 42 and a network connection terminal 43. The image data of the scanned answer pages stored in the buffer memory is segmented by the processor 42. The resulting segment information data associated with the answer pages and the image data is sent to the technical infrastructure via the connection terminal 43.

The pages of the document may be fed through the scanner using a high speed feeder. This allows a large number of pages to be scanned quickly. However, a problem that can arise with high speed scanning is the stretching of areas of a scanned image along the scanning direction. This problem can be solved by providing regularly spaced markers of a predetermined distance along the scanning direction of a document page. The predetermined distance between the markers is known by the processor. For example, the document version indicator 14 on the front page of the document can be used to indicate to the processor 42 the predetermined distance between the markers. The processor 42 of the scanner measures the distance between the markers and if the distance between the markers is larger than the predetermined distance, then stretching of the image in that area has occurred. Image processing methods can be used to correct the stretching of the image. For example, the distance between the markers is related to a number of pixels for a given resolution. Thus a greater distance between the markers due to stretching can indicate the number of excess pixels that have arisen along the scanning direction due to stretching. The processor can remove the excess pixels along the scanning direction to correct the stretching of the image. For example, where two adjacent rows of pixels are identified, one is deleted. The image correction can also be carried out on a processor

on computer, for example by a processor on the central database 31. This technique can be used in other methods involving scanning of documents.

Figures 5A-D depict a process of segmentation of an image. This process can be carried out on an image stored in the buffer memory of a scanner or any other
5 electronic storage means.

In step S1, the image is checked for certain indicators (such as rule lines, margins, registration marks, page numbers etc), which indicate that the image is of a known exam page. In step S2, it is determined whether or not the correct indications are found. If the indicators are not found, then the scanner rejects the page and alerts
10 a user of the scanner, as shown by in step S3. In step S4, the image of the page in the buffer memory removed and the process is ended. The segmentation process may then start again for another image.

If the indicators are found, then the process proceeds to step S5. In step S5 the processor begins searching the image of the first answer page of the document. The
15 first section is searched for any marks that were not part of the predetermined format, i.e. any marks made by the exam candidate. The search begins from the top of the page towards the bottom of the page and then on to the next page in the sequence (the downstream direction). The indicators mentioned above can be used to indicate the top and bottom of each page. An embodiment can be configured to search the whole
20 of the first area or just in the response areas.

A segment can span more than one page. From the page indicators the downstream direction can be deduced and the segment continues from the bottom of a page to the top of the next page in the page sequence of the document.

In step S6, a mark is found. In step S7, the location of the mark is recorded.
25 Intelligent Character Recognition is applied to the mark in step S8. To leave a clear image of the mark to be recognised, the processor can remove the imaged artefacts of the predetermined format area (for example, the shading). This improves the accuracy of the character recognition. Alternatively, the colour of the predetermined formatted areas (for example, the shading colour) can be removed optically by using a
30 certain wavelength of light while scanning or by using optical filters. Character recognition can be carried out on all marks found in the first section, except where these marks may be undesired artefacts of the imaging process which can be electronically removed.

When performing character recognition, it is possible to set acceptable confidence limits. This allows the character recognition software to determine if a character is deemed to be recognised from a mark and thus passed on to the next stage. If the confidence level of a character recognised from a mark falls below the acceptable configurable confidence limit as determined in Step S9, then the process
5 moves on to step S10. In another embodiment, the character is flagged for intervention via a keying process which presents the character to a human operator for a decision on what the character represents.

In step S10, the search for marks continues in the downstream direction from the location of the last mark found and then continues to step S6.
10 If a character is recognised in step S9, its location, character and confidence rating is recorded. Where the predetermined formatted design includes other response areas in the immediate vicinity of each other, for instance on the same row, any characters that can be recognised from marks in such areas are concatenated and then the process
15 proceeds to step S11.

In step S11, the recognised characters are compared with a reference list of characters, which correspond to the nomenclature of the exam questions. If the recognised characters match with characters in the reference list, then the process moves on to step S13. In step S13, the recognised characters are recorded as the
20 question number.

If the recognised characters are not in the reference list, then the process proceeds to step S10. Alternatively, the character is flagged and checked by a user. Using the example of three response boxes in the answer page shown in Figure 2, steps S8 to S12 are carried out on any marks on all three boxes. Thus, all three marks
25 must be recognised and compared to the list of question numbers. The three recognised characters must match with a three-digit question number in the reference list to proceed on to step S13. If the number derived from all three recognised characters is not in the reference list, then the process proceeds to step S10. Alternatively, the characters are flagged and checked by a user. It is also possible that
30 more than one character can be recognised in a single response area.

In step S14, the second section is searched for a blank area. The search is carried out in a location horizontally adjacent to the location of the question number or alternatively a location where there is a mark that provides suspicion of a question number. The parameters of the blank area search can be configured so that the search

can be carried out in the vicinity of the location of the question number. For example, the search parameters can be configured such that the blank area search is carried out two lines above and two lines below the horizontal location of the question number.

The blank area can have a predetermined height and width, which can be a blank line
5 across the page, and does not have any marks or written text within it. The processor allows for sloped writing when searching for a blank area and does not assume a level/horizontal division between answers.

If it is determined in Step S15 that a blank area is found, the process proceeds to step S18. In step S18, the processor sets the location of the blank space as the start
10 of the segment for the question number or the suspected question number.

It is possible to configure the processor to set the start of the segment directly from the location of the question number. For example, it is possible to configure the processor to set the start of the segment at a location horizontally adjacent to the question number or at a location horizontally adjacent to a predetermined vertical
15 distance from the location of the question number. It is also possible to set the start of the segment at a horizontal line that is adjacent to or in the vicinity of the question number and is drawn across the second section by the exam candidate. Thus, steps S14 to S18 would not be required.

If the blank area is not found, i.e. there is writing in the areas adjacent to and
20 in the vicinity of the question number, then the process moves on to step S16. In step S16, the scanner processor searches a configurable distance in the upstream direction for a blank area. If one is not found, it searches downwards a configurable distance. A blank space that is located below the question number can be set as the start of the question number if there is no other question number adjacent to or in the vicinity of it
25 in the first section. The location of the first blank area that is found is recorded and set as the start of the segment of the question number in step S18.

The start of the segment can be set from the location of the question number in the manner described above in circumstances where a blank space cannot be located in step S16.

30 In step S19, the scanner processor then searches the first section, downstream from the location of the question number, for marks which correspond to another question number in the reference list or alternatively another mark. Steps S20 to S26 correspond to earlier steps S6 to S12 for finding and recognising a mark.

The second recognisable character can indicate the location around which the end of the answer of the question is located and the beginning of the answer to another question is located. In step S27, the second recognisable character is recorded as the subsequent question number attempted by the exam candidate. The subsequent question number can be any question number in the reference list and does not have to be the next question in the sequence of question numbers in the exam paper.

Similarly to step S14 above, in step S28 the second section is searched for a blank area horizontally adjacent to or in the vicinity of the location of the subsequent question number.

10 If a blank area is found (S29), the processor can check for further blank areas above the blank area adjacent to the location of the subsequent question number. If further blank areas are found, the end of the segment will be set as the location of a blank area that is adjacent to the last detected marks in the second section. If further blank areas adjacent to the location of the subsequent question number are not found, 15 the location of the end of segment will be set as the blank area adjacent to the location of the subsequent question number (S32).

 If the blank area is not found, then the processor moves on to step S30. In step S30, the scanner processor searches the document image in the upstream direction for a blank area. The processor then proceeds to step S32 where the location 20 of the first blank area that is found is set as the end of the segment for the question number. This will also set the start of the segment for the subsequent question number. It is also possible to allow segments to overlap.

It is possible to set the end of the segment for the question number and the start of the segment for the subsequent question number directly from the location of the 25 subsequent question number. The location of the subsequent question number can be used to determine the end of the segment for the question number and the start of the segment for the subsequent question number by plotting the segment coordinates to begin on the upper boundary of response area in the first section and extending through the horizontal plane into the second section. It is possible to configure the 30 processor to set the end of the segment at a location horizontally adjacent to the subsequent question number or at a location horizontally adjacent to a predetermined vertical distance from the location of the subsequent question number. It is also possible to set the end of the segment at a horizontal line that is adjacent to or in the

vicinity of the subsequent question number and is drawn across the second section by the exam candidate. Thus, steps S28 to S32 would not be required.

If a blank area is not found in step S30, the location of the subsequent question number can be used to determine the end of a segment in the manner described above.

- 5 The segmentation process is then repeated for the remaining images of the pages in the document.

Problems with recognition of segments are resolved through “clip-fixing” software. Clip fixing is an exception process where characters that could not be recognised by the standard character recognition approaches and according to the
10 configurable confidence levels in place are escalated to a suitably qualified user in order that a decision on what the mark represents can be made. Characters that have not corresponding segmentation areas, or those that are too thin to practically represent content, or segmentations that overlap others are escalated to a suitably qualified user in order that a decision on what the segment should represent.

- 15 Marking from an image using the approach described herewith can be used as part of a quality control system. Such a system can actively intervene if marking quality is sub-optimal. This reduces the cost in the processing of examination documents because examiners whose work that is not of sufficient standard would not be permitted to continue.

- 20 Also, marking an exam at the level of individual questions brings quality control benefits as it reduces examiner bias that would often be present when the whole document paper is sent to be marked by an individual examiner. It also permits examiners to specialise in certain question types, which also increases the consistency of marking standards.

- 25 The present invention can be utilised in areas other than examination documents. The invention can be applied to any type of hand written or typed document. For example, the present invention is advantageously applied to laboratory books where the images of the laboratory book may be segmented according to a project code, associated data files, the date or any other information. Other areas of
30 application can include forms, notepads, project books, diaries, planners, address books, visitors books, time management, stock control, etc...

It will be appreciated that the above description is not intended to be limiting and that other modifications and variations fall within the scope of the present invention, which is defined by the appended claims.

CLAIMS

1. A method of obtaining segment information of a document image, wherein the document has a first section and a second section, comprising the steps of:

searching the first section for a first indicating text and determining the location thereof;

in the second section, setting a first position as the start of a segment based on the location of the first indicating text;

searching the first section, downstream from the first indicating text, for a second indicating text and determining the location thereof; and

in the second section, setting a second position as the end of the segment based on the location of the second indicating text.

2. A method according to claim 1, further comprising the step of searching the second section adjacent to or in the vicinity of the location of the first indicating text for a first blank area having a predetermined height and width and not containing any free text, wherein the first blank area is set as the start of the segment.

3. A method according to claim 2, wherein, if a first blank area adjacent to or in the vicinity of the first indicating text is not found, searching the second section in an upstream direction for a blank area and setting the first found blank area as the start of the segment.

4. A method according to any one of claim 1 to 3, further comprising the step of searching the second section adjacent to or in the vicinity of the location of the second indicating text for a second blank area having a predetermined height and width and not containing any free text, wherein the second blank area is set as the end of the segment.

5. A method according to claim 4, further comprising the step of searching in an upstream direction for further blank areas above the second blank area and setting the end of the segment as the location of a blank area that is adjacent to the free text or the start of the segment.

6. A method according to claim 4, wherein, if a second blank area adjacent to or in the vicinity of the second indicating text is not found, searching the second section in an upstream direction for a blank area and setting the first found blank area as the end of the segment.
7. A method according to claim 4, further comprising the step of searching in an upstream direction for further blank areas above the second blank area and setting the end of the segment as the location of a blank area that is adjacent to the free text or the start of the segment.
8. A method according to any one of claims 1 to 3, wherein, if a second indicating text is not found, searching the second section for a blank area that persists to the end of the document and setting the start location of the blank area as the end of the segment.
9. A method according to any one of the preceding claims, wherein the first section has a predetermined format, and wherein the step of searching the first section for a first indicating text comprises the steps of:
 - searching for any mark in the first section that is different to the predetermined format;
 - applying character recognition to said mark;
 - comparing a character obtained from the mark to a list of reference characters;
 - and
 - if the character relates to a reference character, setting the character as the first indicating text.
10. A method according to claim 9, wherein characters that are horizontally adjacent to each other form a character string and if the character string relates to reference characters, setting that character string as the first indicating text.
11. A method according to any one of the preceding claims, wherein the first section has a predetermined format, and wherein the step of searching the first section for a second indicating text comprises the steps of:

searching for any mark in the first section that is different to the predetermined format;

applying character recognition to said mark;

comparing a character obtained from the mark to a list of reference characters;

and

if the character relates to a reference character, setting the character as the second indicating text.

12. A method according to claim 11, wherein characters that are horizontally adjacent to each other form a character string and if the character string relates to reference characters, setting that character string as the second indicating text.

13. A method according to any one of claims 9 to 12, wherein more than one character is recognised from the mark.

14. A method according to any one of claims 9 to 13, wherein, when applying character recognition, a confidence level that a character is recognised from the mark is higher than a confidence level threshold.

15. A method according to any one of claims 8 to 14, wherein the predetermined format comprises a plurality of regularly spaced response areas that are distinguished from the surroundings.

16. A method according to claim 15, wherein the surroundings are shaded and the response area is unshaded.

17. A method according to claim 15 or 16, wherein only the response areas are searched for marks.

18. A method according to any one of the preceding claims, wherein the document has one or more pages, each page having a sequence indicator which indicates the downstream direction.

19. A method according to any one of the preceding claims, wherein the first section is adjacent to and left of the second section.
20. A method according to any one of the preceding claims, wherein the first section is adjacent to and right of the second section.
21. A method according to any one of the preceding claims, wherein the first and second indicating text are handwritten.
22. A method according to any one of claims 2 to 21, wherein the free text is handwritten.
23. A method according to any one of the preceding claims, wherein the segment information comprises information related to the first indicating text, the location of the start of the segment and the location of the end of the segment.
24. An apparatus for obtaining segment information of a document image, wherein the document has a first section and a second section, the apparatus comprising:
- a means for searching the first section; and
 - a means for setting segment information, wherein:
 - the means for searching the first section searches for a first indicating text and determines the location thereof;
 - the means for setting segment information sets, based on the location of the first indicating text, a first position in the second section as the start of a segment;
 - the means for searching the first section searches, downstream from the first indicating text, for a second indicating text and determines the location thereof; and
 - the means for setting segment information sets, based on the location of the second indicating text, a second position in the second section as the end of the segment.

25. An apparatus according to claim 24, further comprising a means for searching the second section, wherein:

the means for searching the second section searches adjacent to or in the vicinity of the location of the first indicating text for a first blank area having a predetermined height and width and not containing any free text; and

the means for setting the segment information sets the first blank area as the start of the segment.

26. An apparatus according to claim 25, wherein, if the means for searching the second section does not find a first blank area adjacent to or in the vicinity of the first indicating text, the means for searching the second section searches in an upstream direction for a blank area and the means for setting the segment information sets the first found blank area as the start of the segment.

27. An apparatus according to any one of claims 24 to 26, wherein:

the means for searching the second section searches adjacent to or in the vicinity of the location of the second indicating text for a second blank area having a predetermined height and width and not containing any free text; and

the means for setting the segment information sets the second blank area as the end of the segment.

28. An apparatus according to claim 27, wherein the means for searching the second section searches in an upstream direction for further blank areas above the second blank area and the means for setting the segment information sets the end of the segment as the location of a blank area that is adjacent to the free text or the start of the segment.

29. An apparatus according to claim 27, wherein, if the means for searching the second section can not find a second blank area adjacent to or in the vicinity of the second indicating text, the means for searching the second section searches in an upstream direction for a blank area and the means for setting the segment information sets the first found blank area as the end of the segment.

30. An apparatus according to any one of claims 24 to 26, wherein, if the means for searching the first section does not find a second indicating text, the means for searching the second section searches for a blank area that persists to the end of the document and the means for setting segment information sets the start location of the blank area as the end of the segment.

31. An apparatus according to any one claims 24 to 30, wherein the first section has a predetermined format, and the means for searching the first section comprises:

- a means for searching for any mark that is different to the predetermined format;
- a means for character recognition of said mark;
- a means for comparing a character obtained from the character recognition means to a list of reference characters; and
- a means for setting the character as the first or second indicating text if the character relates to a reference character.

32. An apparatus according to claim 31, wherein the derived characters that are horizontally adjacent to each other form a character string and if the character string relates to reference characters, the means for setting the character sets that character string as the first or second indicating text.

33. An apparatus according to claim 31 or 32, wherein the means for character recognition recognises more than one character from the mark.

34. An apparatus according to any one of claims 31 or 33, wherein the means for character recognition comprises a confidence level threshold, wherein a confidence level that a character is recognised from the mark is higher than the confidence level threshold.

35. An apparatus according to any one of claims 31 to 34, wherein the predetermined format comprises a formatted area with a plurality of response areas, and wherein the response areas are regularly spaced.

36. An apparatus according to claim 35, wherein the surroundings are shaded and the response areas are unshaded.
37. An apparatus according to claim 35 or 36, wherein the means for searching the first section only searches the response areas for marks.
38. An apparatus according to any one of claims 24 to 37, wherein the document has one or more pages and each page has a sequence indicator which indicates the downstream direction.
39. An apparatus according to any one of claims 24 to 38, wherein the first section is adjacent to and left of the second section.
40. An apparatus according to any one of claims 24 to 38, wherein the first section is adjacent to and right of the second section.
41. An apparatus according to any one of claims 24 to 40, wherein the first and second indicating texts are handwritten.
42. An apparatus according to claim 25 or any claim dependent thereon, wherein the free text is handwritten.
43. An apparatus according to any one of claims 24 to 42, wherein the segment information comprises information related to the first indicating text, the location of the start of the segment and the location of the end of the segment.
44. A method comprising the steps of:
 scanning a document;
 while scanning the document, obtaining segment information of the document;
 compressing an image of the document; and
 storing the segment information and the compressed image related to the document.

45. A method according to claim 44, wherein the step of obtaining segment information is done in a buffer memory of a scanner.
46. A method according to claim 44 or 45, wherein the step of obtaining segment information comprises a method according to any one of claims 1 to 16.
47. A method according to any one of claim 44 to 46, further comprising a method of displaying a segment, the method of displaying a segment comprising the steps of:
requesting a segment;
obtaining parts of the compressed image which correspond to the requested segment; and
displaying said parts.
48. A method according to claim 47, further comprising the step of assigning and submitting, to a database, data related to the displayed parts.
49. A method according to claim 47 or 48, wherein said parts are displayed in a sequence according to the segment information and the downstream direction.
50. A method according to claim 49, wherein the sequence of parts are concatenated and displayed as a single image.
51. A method according to any one of claims 47 to 50, wherein the step of requesting and displaying occurs at a location remote from a location of the stored segment information and the stored compressed image.
52. A method according to any one of claims 1 to 23 and 44 to 51, wherein the document is an examination script.
53. An apparatus comprising:
a means for scanning a document;
a means for obtaining segment information of the scanned document;
a means for compressing an image of the scanned document; and

a means for storing the segment information and the compressed image related to the scanned document, wherein
the segment information is obtained while scanning the document.

54. An apparatus according to claim 53, wherein the segment information is obtained from an image in a buffer memory of a scanner.

55. An apparatus according to claim 53 or 54, wherein the means for obtaining segment information comprises an apparatus according to any one of claims 33 to 52.

56. An apparatus according to any one of claims 53 to 55, further comprising:
a means for requesting a segment;
a means for obtaining parts of the compressed image which correspond to the requested segment; and
a means for displaying said parts.

57. An apparatus according to claim 56, further comprising a means for assigning and submitting, to a database, data related to the displayed parts.

58. An apparatus according to claim 56 or 57, wherein said parts are displayed in a sequence according to the segment information and the downstream direction.

59. An apparatus according to claim 58, wherein the sequence of parts are concatenated and displayed as a single image.

60. An apparatus according to any one of claims 56 to 59, wherein the means for requesting a segment and the means for displaying the parts are at a location remote from a location of the means for storing the segment information and the compressed image related to the scanned document.

61. An apparatus according to any one of claims 33 to 60, wherein the document is an examination script.

62. A system comprising:

an apparatus, connected to a network, comprising:
a means for scanning a document;
a means for obtaining segment information of the scanned document
while scanning the document according to any one of claims 26 to 39; and
a means for compressing an image of the scanned document;
a network device, connected to the network, comprising:
a means for storing the segment information and the compressed image
related to the scanned document; and
a client computer, connected to the network, comprising:
a means for requesting a segment;
a means for obtaining parts of the compressed image which correspond
to the requested segment from the apparatus; and
a means for displaying said parts.

63. A system according to claim 62, wherein the client computer further comprises a means for assigning and submitting, to the apparatus, data related to the displayed parts.

64. A computer program comprising instructions that when executed by a computer system, cause the computer system to perform a method according to any one of claims 1 to 23 and 44 to 51.

1/8

Fig.1.

12 PAGE ANSWER BOOK

EXAMINATION CODE: [] [] [] [] [] []

EXAMINATION DATE: [] [] [] [] [] []
D D / M M / Y Y

CENTRE NUMBER: [] [] [] [] []

CANDIDATE NUMBER: [] [] [] [] [] []

11

leave blank

leave blank

leave blank

14

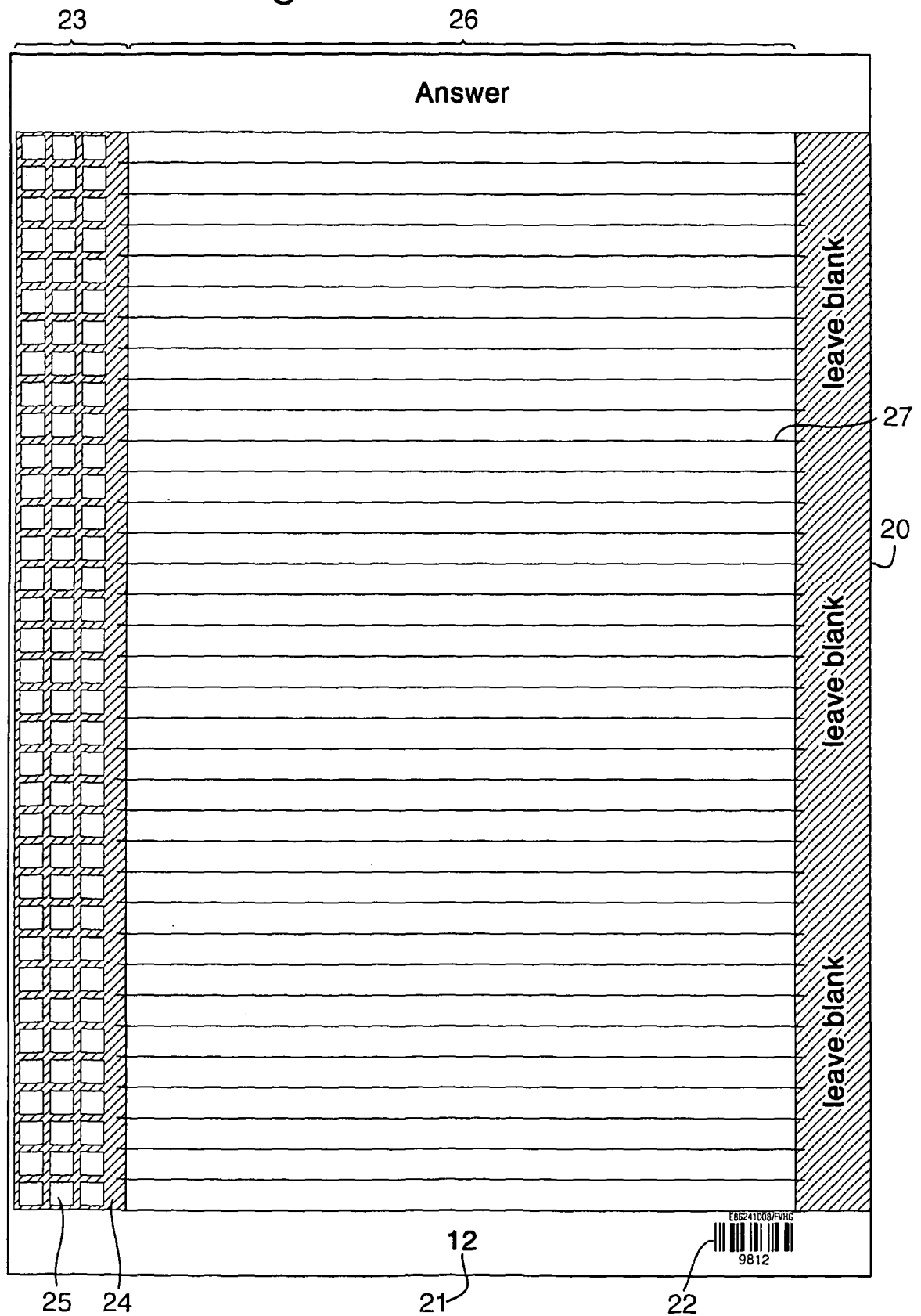
1

13

F86131008/ADW
9801

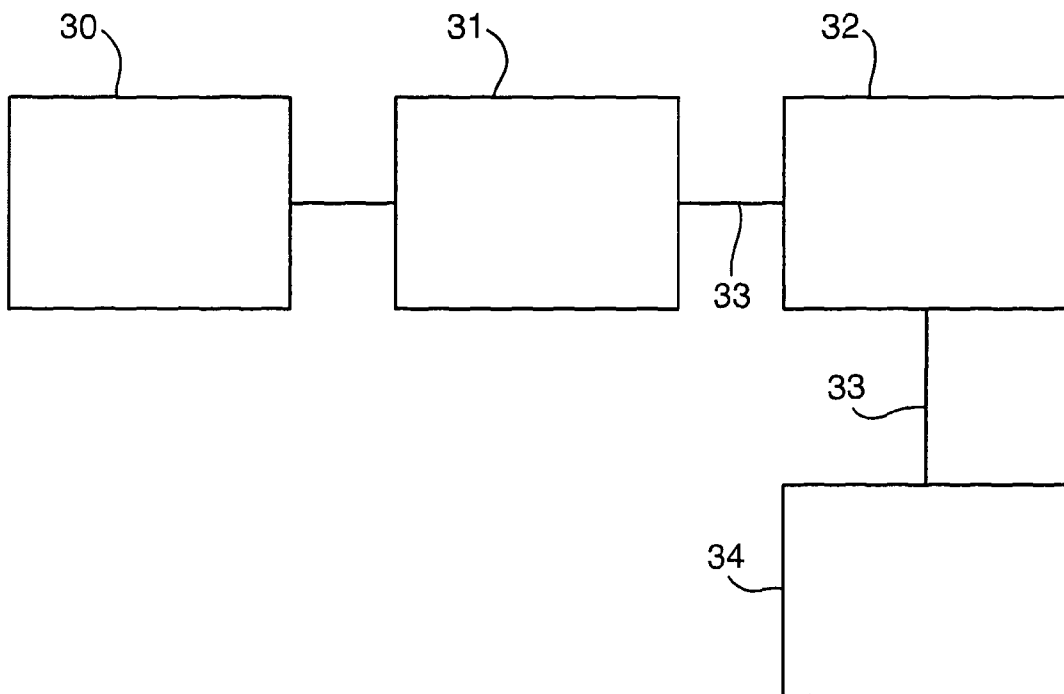
2/8

Fig.2.



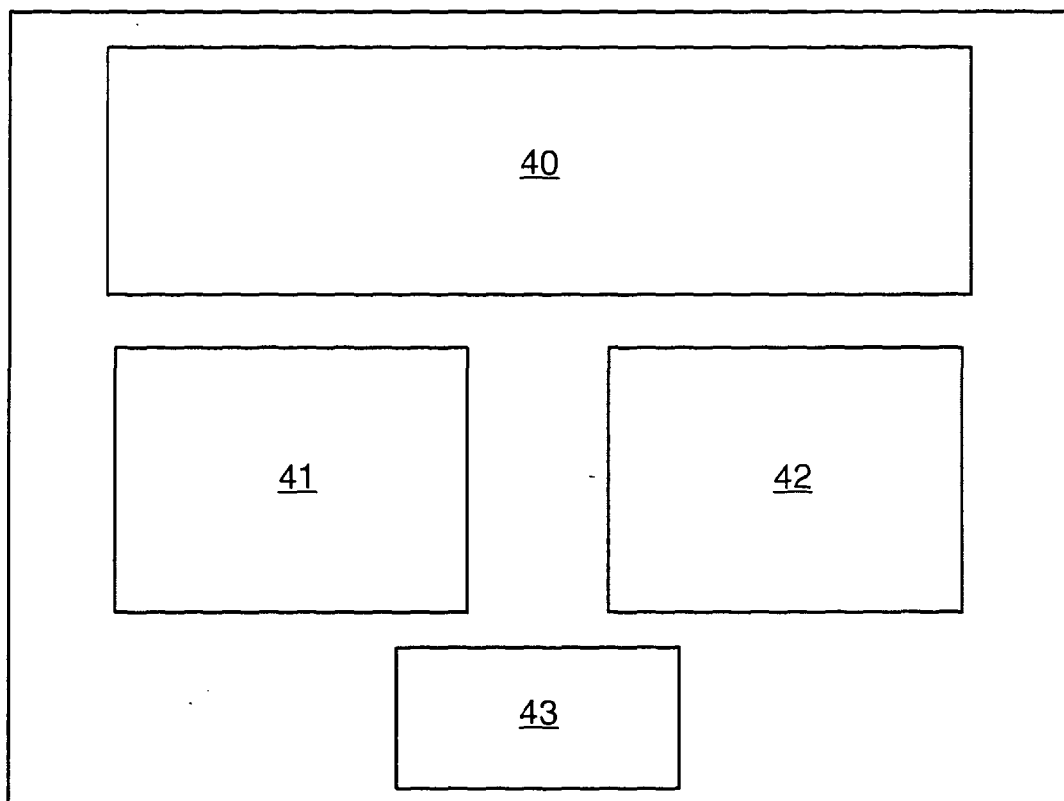
3/8

Fig.3.



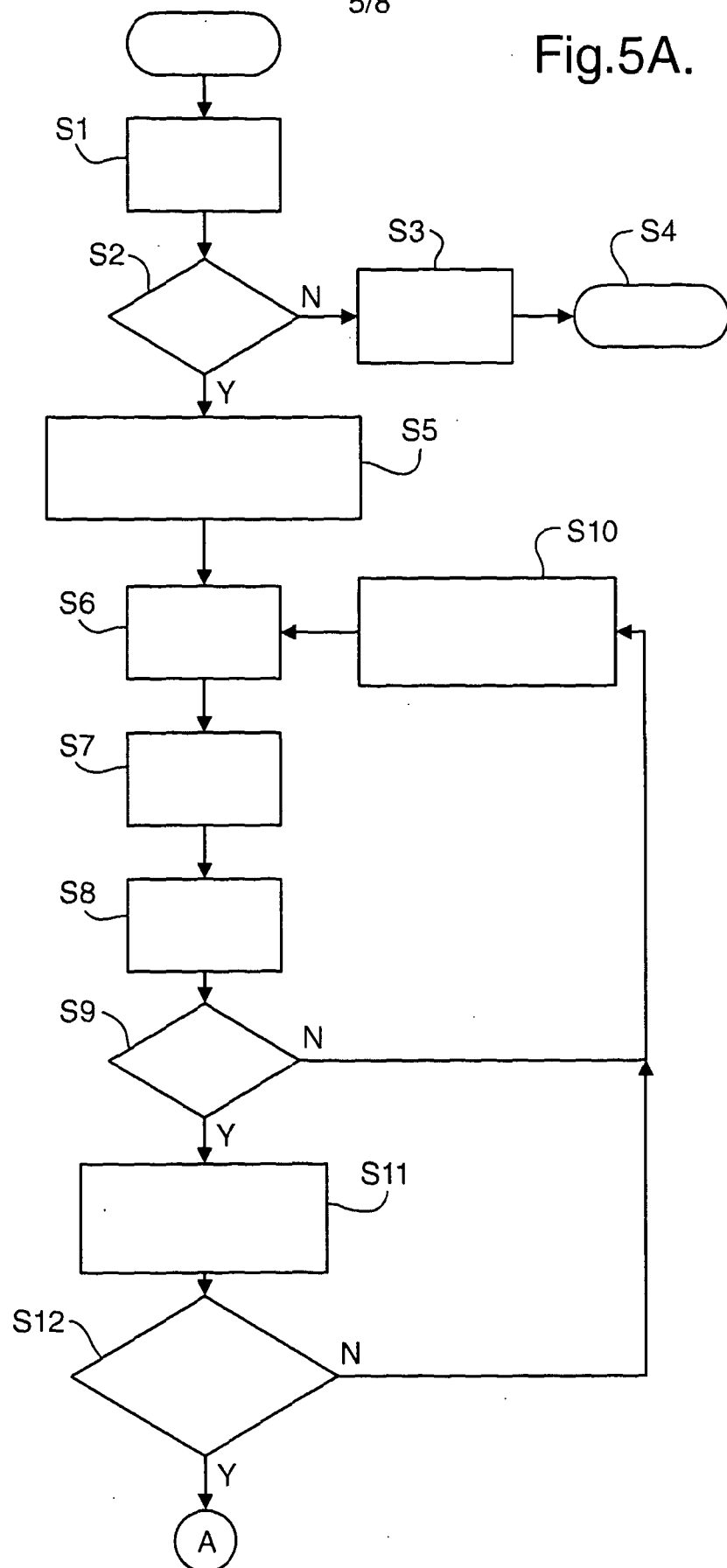
4/8

Fig.4.



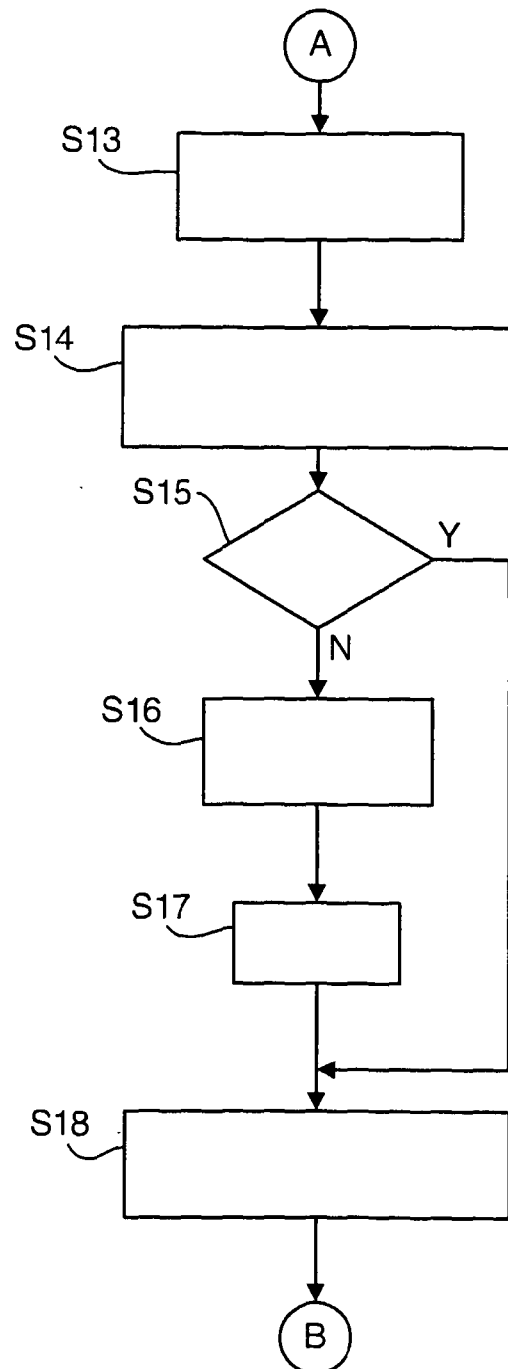
5/8

Fig.5A.



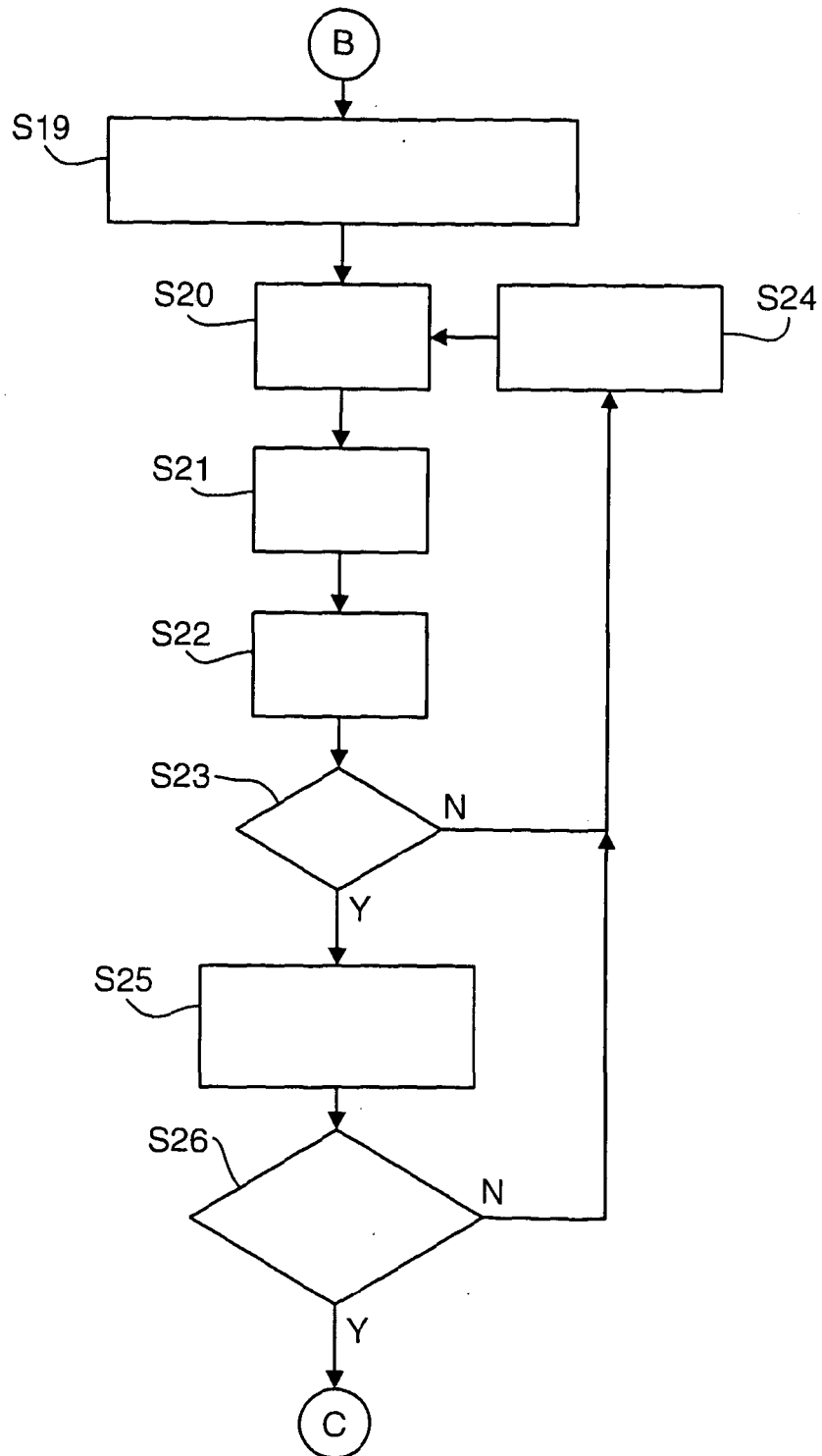
6/8

Fig.5B.



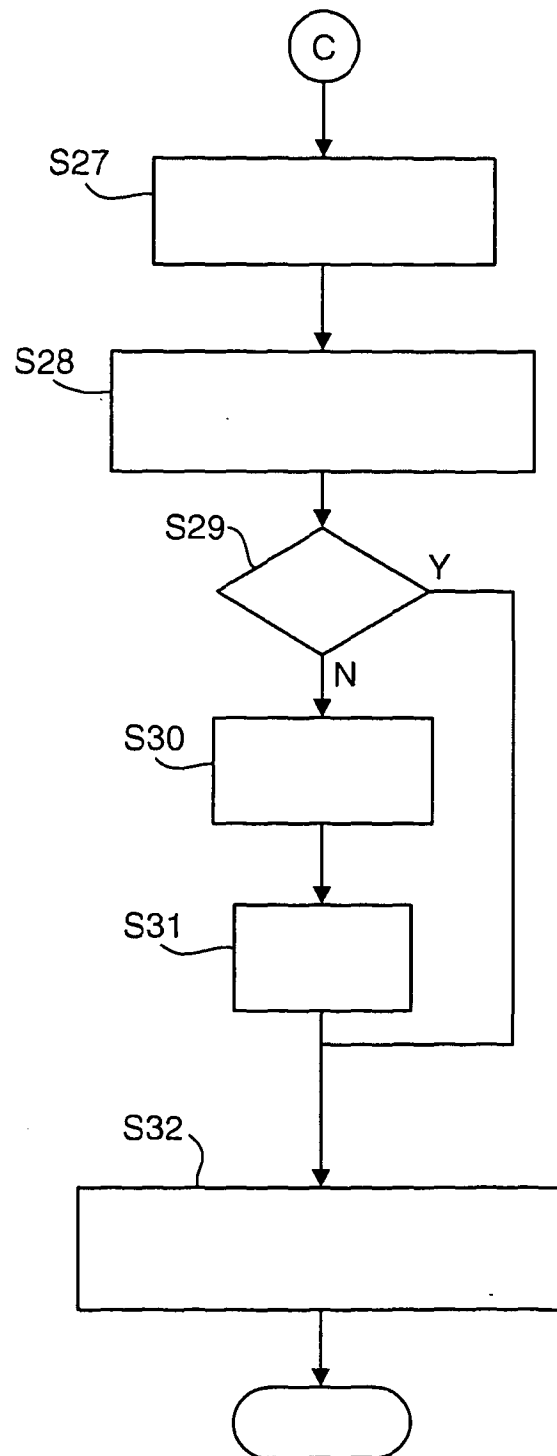
7/8

Fig.5C.



8/8

Fig.5D.



INTERNATIONAL SEARCH REPORT

International application No

PCT/GB2010/001661

A. CLASSIFICATION OF SUBJECT MATTER

INV. G06K9/20

ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X A	US 2003/142358 A1 (BEAN HEATHER N [US] ET AL) 31 July 2003 (2003-07-31) * abstract; claim 1; figures paragraph [0015] - paragraph [0016]	1, 24, 44, 53 2-23, 25-43, 45-52, 54-64
X A	EP 1 833 022 A1 (FUJITSU LTD [JP]) 12 September 2007 (2007-09-12) * abstract; figure 1 paragraph [0017] - paragraph [0021]	1, 24 2-23, 25-64
A	EP 0 081 767 A2 (TOKYO SHIBAURA ELECTRIC CO [JP]) 22 June 1983 (1983-06-22) the whole document	1-64
	----- -/--	



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents :

A document defining the general state of the art which is not considered to be of particular relevance

E earlier document but published on or after the international filing date

L document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

O document referring to an oral disclosure, use, exhibition or other means

P document published prior to the international filing date but later than the priority date claimed

T later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

X document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

Y document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

& document member of the same patent family

Date of the actual completion of the international search

20 December 2010

Date of mailing of the international search report

28/12/2010

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Authorized officer

Isa, Sabine

INTERNATIONAL SEARCH REPORT

International application No

PCT/GB2010/001661

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 11 085895 A (OLYMPUS OPTICAL CO) 30 March 1999 (1999-03-30) the whole document -----	1-64
A	US 6 023 342 A (YANAGIDA MASATO [JP]) 8 February 2000 (2000-02-08) the whole document -----	1-64
A	US 6 616 038 B1 (OLSCHAFSKIE FRANCIS [US] ET AL) 9 September 2003 (2003-09-09) the whole document -----	1-64

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/GB2010/001661

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
US 2003142358	A1	31-07-2003	NONE	
EP 1833022	A1	12-09-2007	WO 2006070476 A1 JP 4398474 B2 US 2007242882 A1	06-07-2006 13-01-2010 18-10-2007
EP 0081767	A2	22-06-1983	DE 3278802 D1 JP 58103266 A US 4589144 A	25-08-1988 20-06-1983 13-05-1986
JP 11085895	A	30-03-1999	NONE	
US 6023342	A	08-02-2000	NONE	
US 6616038	B1	09-09-2003	NONE	