

Development of the Document Image Layout Reconstruction Technology “GT-Layout” for Mobile Devices

Hironori YANO*, Arito ASAI*, Tetsuya SAWANO*, and Norihisa HANEDA**

Abstract

Document browsing on the small displays of mobile devices impairs the readability because the typical page layout of document files is designed for browsing their contents on a PC display or printed paper. To improve the readability of mobile devices, we have developed the document layout reconstruction technology “GT-Layout”. In this report, we introduce the technologies used in GT-Layout.

1. Introduction

As the processing power of mobile devices, such as cell phones and smartphones, is increasing and their communications band is expanding, solutions to use mobile devices as a business tool are developing.

Under this situation, we have developed “GT-Document” (GT-Doc), document viewing technology for mobile devices, making the most of the image technology, our core technology¹⁾. GT-Doc converts a document file created with Microsoft Office, Adobe Acrobat or other document software to an image format file, such as JPEG. This technology is mainly used to view a document file in an office’s file server or a document attached to email on a mobile device. To view a document file, the corresponding viewer application for the file is usually required. This technology, if a mobile device has JPEG display function, enables easier creation of a viewer application to view the document on the device. Some mobile devices have viewer applications for document files. However, viewable layout and design may vary depending on type of mobile device. GT-Doc, as it transmits data as an image, reproduces the original document file exactly regardless of type of mobile device. It is only the image file that is sent to the mobile device and the original document file is not transmitted. That reduces the risk of information leakage due to loss of the mobile phone. There are several solutions for businesses to enable viewing of document files on mobile devices. GT-Doc is given a competitive edge by the perfect reproducibility and security, which are essential for use in business.

Although GT-Doc enables viewing of a document file on a mobile device without the viewer application, the document is not always easy to read. A document file is designed for viewing on a PC monitor, a printout or any other medium

with a large display area. When such a document is viewed on the small screen of a mobile device, it is displayed only partly (Fig. 1). The entire document may be displayed on a mobile device if it is reduced in size. But then the letters may be too small to read.

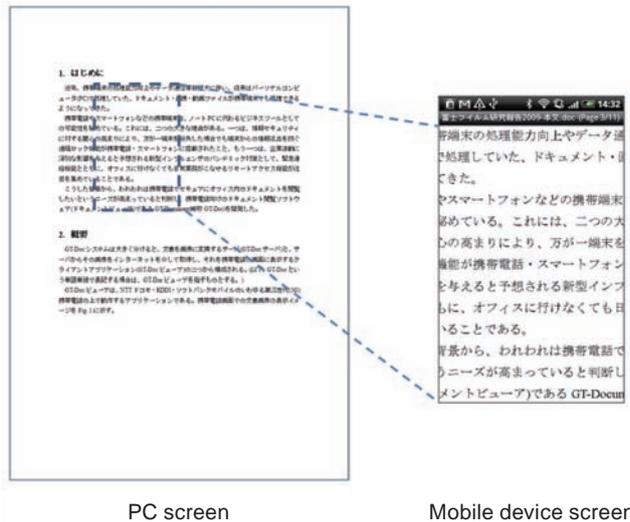


Fig. 1 Document view on a mobile device.

When the document is enlarged until the letters are large enough to read, each line is partly missing and thus the sentences do not make sense. To read the whole text, you have to scroll the screen every time you reach the end of a line on screen and then scroll back at the end of each line of the original text to return to the head of the next line. In other words, the user is disturbed by those operations while reading a document. The problems arising from the layout of a document file and the size of a mobile device screen are roughly divided into two.

Original paper (Received December 14, 2011)

* Internet Business Development Division
FUJIFILM Corporation
Nishiazabu, Minato-ku, Tokyo 106-8620, Japan

** Internet Business Development Division
FUJIFILM Corporation
Akasaka, Minato-ku, Tokyo 107-0052, Japan

- (1) Complete sentences cannot be displayed on one screen.
(Problem of screen size)
- (2) That causes extra steps and makes the operation annoying.
(Problem of ease of operation)

To solve the difficulties in viewing document files on mobile devices, we have developed “GT-Layout”, document layout reconstruction technology. This report provides an overview of the GT-Layout technology and explanations of the features.

2. Technological Concept

When a document file created for a PC is viewed on a mobile device, there arise (1) problem of screen size and (2) problem of ease of operation. These problems reduce the readability. GT-Layout has achieved user-friendly viewing and operation by reconstructing the layout of a document file to suit the display screen of a mobile device (Fig. 2).

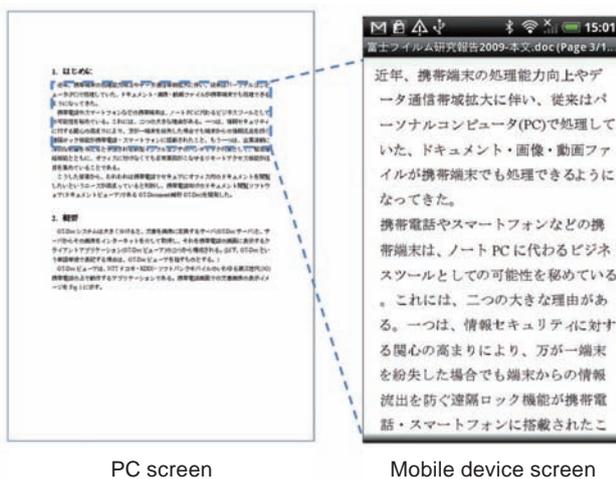


Fig. 2 Layout reconstruction by GT-Layout on a mobile device.

Fig. 2 shows a document reconstructed to suit a mobile device by GT-Layout. GT-Layout reconstructs a group of letters, such as a paragraph, to fit into the screen of a mobile device while keeping the sequence of the letters. That solves the (1) problem of screen size. Extra operations are not necessary. The user can read the text just scrolling down. Therefore, (2) problem of ease of operation is also solved. Next section will explain the system configuration of this technology.

3. Technology

A GT-Layout system consists of the GT-Doc server for converting a document file to an image file, the GT-Layout server for analyzing the layout of the document image, and a client application for acquiring the analyzed layout information and reconstructing the layout to suit a mobile device (Fig. 3).

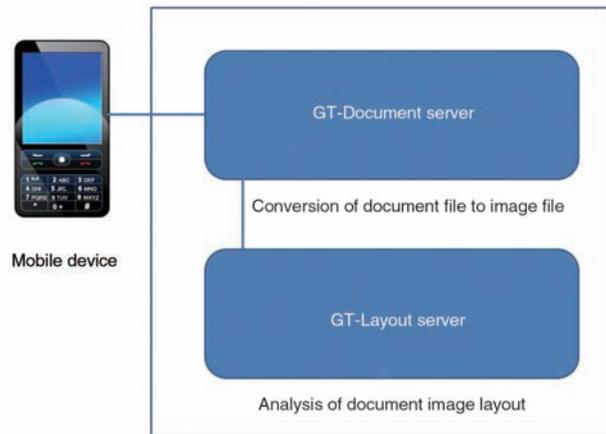


Fig. 3 System configuration of GT-Layout.

This section explains the technologies used for individual components of the system and the operation flow.

3.1 GT-Doc Server

The GT-Doc server converts a document file to an image format file, such as JPEG, and transmits the converted image data to a mobile device so that the user can view the document on the mobile device. The server also receives a request for acquisition of layout information from the client application. The server sends the GT-Layout server the request together with the document image from which the layout is to be acquired. Other basic functions of GT-Doc are provided in the reference literature¹⁾.

3.2 GT-Layout Server

The GT-Layout server receives a request for acquisition of layout information and a document image from the GT-Doc server. The server performs layout analysis and acquires layout information. Layout analysis of a document file normally requires a viewer application that can reproduce the document file. But GT-Layout does not need the application. Instead, GT-Layout applies the optical character recognition (OCR) technology to the document image file converted by the GT-Doc server. That makes it possible to support any document file format the GT-Doc server supports.

The information acquired by layout analysis is needed for reconstruction of the layout of the document image on a mobile device. It consists of character areas, line areas and paragraph areas. Analysis is performed in the flow below.

(1) Acquisition of Character Area

Character areas are acquired from a document image with OCR (Fig. 4). OCR is designed to locate the coordinates of characters and identify their types. But, GT-Layout uses OCR only for locating the coordinates as reconstruction of the layout of a document requires only the coordinates of characters. We have decreased the precision in the character type recognition to speed the OCR processing.



Fig. 4 Recognition of character positions in a document image.

(2) Acquisition of Line Area

A line area is constructed from character areas (Fig. 5). Using a character as a reference, when another character exists at a distance of a character width from the right edge of the reference character, the character is judged to be an adjacent character. The same process is performed using the adjacent character as a reference. Repeating this process, the server makes character strings into a line. When there is a character with a small area, such as punctuation, the line may be broken at the punctuation. To avoid that, when characters are about the same in top and bottom positions and when each space between them is within an average character width, those characters are judged as a line.

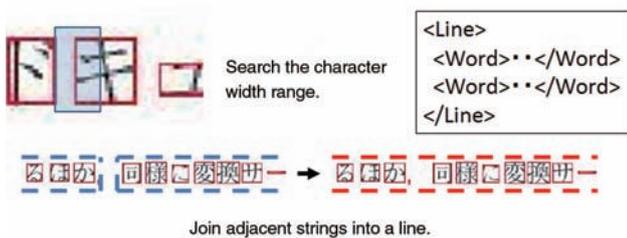


Fig. 5 Calculation of line regions.

(3) Acquisition of Paragraph Area

A paragraph area is constructed from line areas (Fig. 6). When another line exists at a distance of a line height from the bottom edge of a line, the line is judged to be an adjacent line. Repeating this process, the server acquires a paragraph, a set of lines.

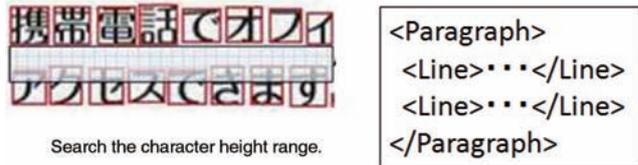


Fig. 6 Calculation of paragraph regions.

The layout information of a document image is acquired by the processes (1), (2) and (3) above. Next section explains how layout suitable for a mobile device is created from the acquired layout information and the document image.

3.3 Client Application

The client application, communicating with the GT-Doc server, receives a document image and displays it on the screen of a mobile device¹⁾.

To reconstruct the layout of a document image on screen, the application sends a request for acquisition of layout information to the GT-Doc server and obtains the layout information. Next section explains the method of layout reconstruction and describes UI of the application.

3.3.1 Layout Reconstruction

The client application reconstructs the layout to fit the screen of the mobile device using the document image and the layout information.

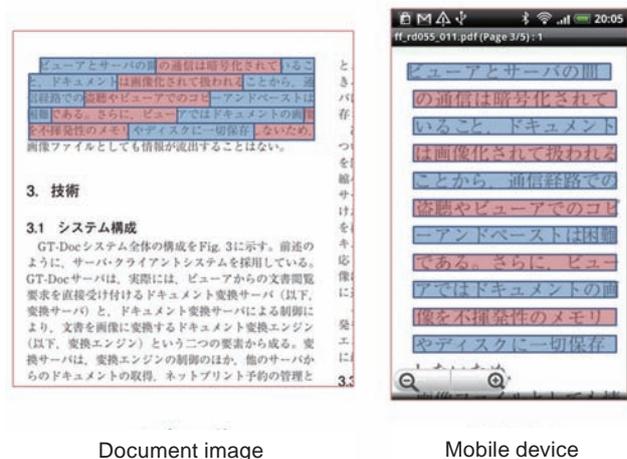


Fig. 7 Reconstruction of paragraph layout.

Fig. 7 shows an original document image on the left and a mobile device screen with reconstructed layout on the right. The text on the mobile screen and its corresponding character areas of the original document are shown in color. The layout is reconstructed in the following procedure.

(1) Specifying a Paragraph

A paragraph of which the layout is to be reconstructed is specified in the layout information.

(2) Laying out a Paragraph Area

The line area information of the specified paragraph is

acquired. A part of the first line of the paragraph to fit into the screen of the mobile device is cut out from the document image in order from the head of the line and displayed on the mobile screen. This process is repeated until the last part of the first line is cut out and displayed. If there is space left next to the end of the first line on the screen, a part of the next line to fit into the space is cut out and displayed. The process is repeated until the last line of the paragraph is cut out and displayed.

It is possible to reconstruct a document image using the character types OCR information contains. But, GT-Layout is better in that the document image is used as it is and that the appearance, for instance, font, character and line spaces, is maintained.

3.3.2 Application UI

This section describes the UI of the client application, assuming that the touch screen is used to operate the mobile device. The arrow key can be also used to perform the same operations.

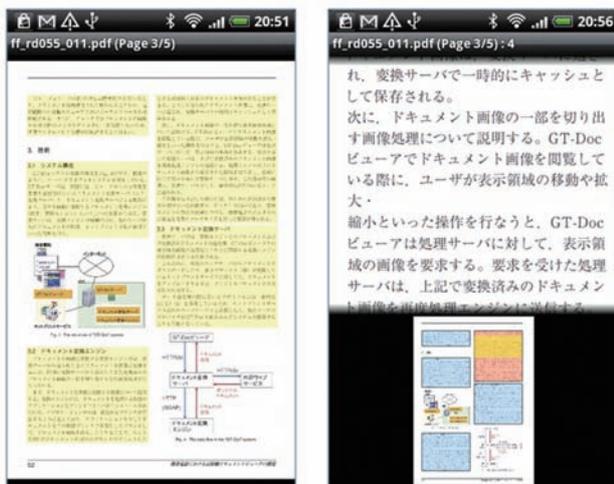


Fig. 8 Client application UI.

Fig. 8 shows screens of the client application after it has received layout information.

When the client application has acquired layout information, the paragraph selection screen shown on the left of Fig. 8 is displayed. The application acquires paragraph areas from the paragraph area information in the layout information and highlights them in the document image on screen. This helps to identify the paragraph to be reconstructed. By touching the highlighted area, the client application reconstructs the paragraph.

The reconstructed layout is shown on the right of Fig. 8. The upper half of the screen shows a paragraph with the reconstructed layout and the user can read it simply by scrolling up and down. The font size can be increased and decreased by pinch-in and pinch-out gestures on the screen. The lower half shows a thumbnail of the whole page. The

paragraph currently displayed in the upper half of the screen is highlighted in red in the thumbnail. All the paragraphs that can be reconstructed are highlighted in blue. When the user touches one of the blue areas, the touched paragraph will replace the currently shown paragraph. The user can move from one paragraph to another by intuitive operations. The user can also see which part of the document he or she is viewing. In other words, it is possible to check both the whole page and a part of it simultaneously even in a small screen of a mobile device.

4. Conclusion

This report has explained GT-Layout, technology for improving readability of document images on mobile devices using our image technology. This technology makes it easy to read documents on mobile devices. That will increase the use of mobile devices as a business tool and enhance business efficiency. As well as continuing our efforts to further improve the functions and performance of this technology, we will work to expand the market for the technology, exploring applications other than mobile devices and various other purposes.

Reference

- 1) Ohtani, Hiroshi; Ikeya, Naoki; Uesugi, Eiji; Asai, Arito; Segawa, Hideo; Haneda, Norihisa. "High Quality Document Viewer on Mobile Phones". FUJIFILM RESEARCH & DEVELOPMENT. No.55, 50-54 (2010).

(In this report, "Microsoft Office" is a registered trademark of Microsoft Corporation. "Adobe" and "Acrobat" are the registered trademarks of Adobe Systems Inc. "GT-Document" and "GT-Doc" are the registered trademarks of FUJIFILM Corporation.)