

携帯端末向け文書画像レイアウト再構成技術「GT-Layout」の開発

矢野 浩教*, 浅井 有人*, 沢野 哲也*, 羽田 典久**

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. はじめに

近年の携帯電話やスマートフォンなどの携帯端末の処理能力向上や通信帯域の拡大に伴い、携帯端末をビジネスツールとして利用する法人向けソリューションが発展している。

そのような背景のなか、われわれは当社のコア技術であり画像技術を活用した携帯端末における文書閲覧技術「GT-Document」(以下GT-Doc)を開発した¹⁾。GT-DocとはMicrosoft OfficeやAdobe Acrobatなどで作成された文書ファイルをJPEGなどの画像形式に変換し、変換したデータを配信する技術である。本技術は主に社内ファイルサーバ内の文書ファイルやメールの添付ファイルを、携帯端末で閲覧する用途で利用される。通常文書ファイルの閲覧には閲覧対象となる文書ファイルの専用ビューアアプリケーションが必要であるが、本技術により携帯端末がJPEG表示機能を持っていれば、目的文書ファイルを閲覧可能なビューアアプリケーションを容易に作成可能となる。仮に携帯端末に専用ビューアアプリケーションが存在する場合でも、携帯端末の種別によって表示可能な文書ファイルのレイアウトやデザインが異なることが多いが、GT-Docでは画像化したデータを配信するため、端末種別に依存せず完璧な再現性で元文書ファイルと同一の表示が可能となる。また携帯端末に送信するのは画像化された文書ファイルのみであり、文書ファイル本体は送信しないため、携帯端末紛失などによる情報流出の危険性が低減される。携帯端末上で文書ファイルを閲覧する法人向けソリューションは幾つか存在するが、GT-Docはビジネス用途で利用するために必

要な要件である「完璧な再現性」と「安全性」を持つことで優位性を高めている。

しかしながらGT-Docにより専用ビューアアプリケーションが存在しない携帯端末でも文書ファイルを閲覧することが可能となったとはいえ、「快適な」閲覧が可能になったとは言い難い。本来文書ファイルはPCディスプレイ上や印刷紙のように表示領域が広い媒体で閲覧することを前提としたレイアウトであることが多く、携帯端末上で文書ファイルを閲覧した場合、携帯端末ディスプレイのように小さい表示領域では文書ファイルの一部しか表示することができない (Fig. 1)。一方携帯端末上で全体を表示可能なように、文書画像を縮小した場合、

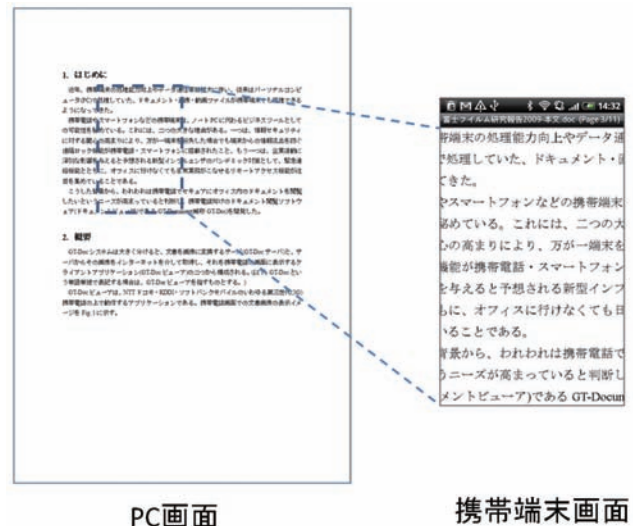


Fig. 1 Document view on a mobile device.

本誌投稿論文 (受理 2011 年 12 月 14 日)

*富士フィルム (株) ネット応用ビジネス推進部
〒106-8620 東京都港区西麻布 2-26-30

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

**富士フィルム (株) ネット応用ビジネス推進部
〒107-0052 東京都港区赤坂 9-7-3

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

各文字が判読不可能になるほど小さくなるため、そもそも閲覧には適さない。

文字が閲覧可能なサイズまで拡大した場合には画面に表示される文字列は各行間で文脈が繋がっていない。文脈を繋げたままで閲覧するためには画面上閲覧中の行が端に達した場合、画面外の領域を表示するために、スクロール操作を行なう必要があり、また文書ファイル上で行末にきた場合、行頭に戻るスクロール操作が必要となる。このように携帯端末における文書閲覧では、閲覧行為と操作行為を交互に意識する必要があり、持続した閲覧行為を行なうことが不可能となる。この文書ファイルのレイアウトと携帯端末表示領域の狭さから生じる閲覧性の問題は大きく以下の2つに大別される。

- ① 一画面で文章の文脈を追うことができない。(表示画面領域の問題)
- ② そのため操作が煩雑となる。(操作性の問題)

そこでわれわれは携帯端末における文書ファイル閲覧性の問題を解消すべく、文書レイアウト再構築技術「GT-Layout」の開発を行なった。本報告では、GT-Layoutの技術概要と特長を解説する。

2. 技術コンセプト

携帯端末においてPC向けに作成された文書ファイルを表示する場合、(1) 表示画面領域の問題と(2) 操作性の問題が存在し、閲覧性を下げていることは前述のとおりである。GT-Layoutではこれらの問題に対し、文書ファイルのレイアウトを、携帯端末表示領域に合わせ再構築することで、快適な閲覧と操作を実現した (Fig. 2)。

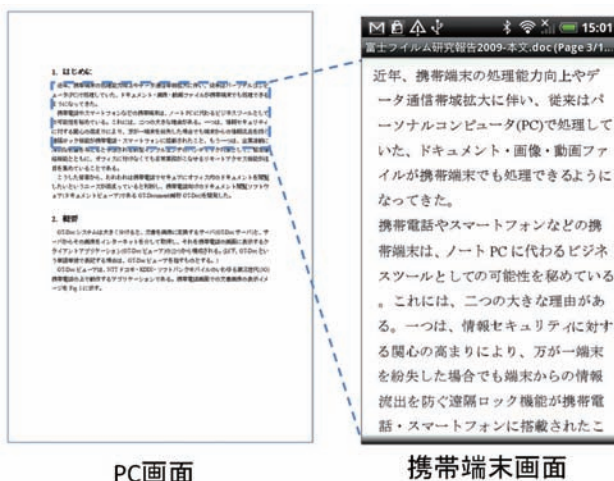


Fig. 2 Layout reconstruction by GT-Layout on a mobile device.

Fig. 2はGT-Layoutにより携帯端末向けにレイアウトを再構築した文書閲覧画面を示している。GT-Layoutでは段落のように文字が密集した領域を、文字の並びを保持した形式で携帯端末表示領域に合わせ再レイアウトする。これにより(1) 表示画面領域の問題を解決可能と

なる。また煩雑な操作は必要なくなり、縦スクロールのみで操作可能となり(2) 操作性の問題も解決される。次章では本技術の構成を説明する。

3. 技術

GT-Layoutは文書ファイルを画像化するGT-Docサーバ、文書画像のレイアウトを解析するGT-Layoutサーバと、解析したレイアウト情報を取得して携帯端末に適したレイアウトに画面を再構築するクライアントアプリケーションで構成される (Fig. 3)。



Fig. 3 System configuration of GT-Layout.

本章では、各構成要素で使用されている技術と動作フローについて説明する。

3.1 GT-Docサーバ

GT-Documentサーバでは文書ファイルをJPEGなどのフォーマットに画像化する。画像化したデータを携帯端末上のクライアントアプリケーションに送信することで、携帯端末において文書の閲覧が可能となる。またクライアントアプリケーションから送信されるレイアウト情報取得のリクエストを受け取り、GT-Layoutサーバにレイアウト取得対象となる文書画像と共に、レイアウト情報取得リクエストを送信する。その他のGT-Docの基本的な機能は参考文献¹⁾に記載されている。

3.2 GT-Layoutサーバ

GT-LayoutサーバではGT-Docサーバから送信されるレイアウト情報取得リクエストと文書画像を受信しレイアウト解析処理を行ない、文書画像のレイアウト情報を抽出する。通常文書ファイルのレイアウト解析には当該文書ファイルを再生可能なビューアアプリケーションが必要となるが、GT-Layoutでは、一度GT-Docサーバにより画像化した文書を扱うため、OCRなどの光学文字認識技術を利用する。このためGT-Docサーバが対応している文書ファイルの形式全てに対応することが可能となる。

レイアウト解析処理において抽出する情報は携帯端末側で文書画像のレイアウトを再構成するために必要な情報であり、文字の領域、文字が連なった行の領域、段落の領域から構成される。解析処理は以下の流れで行なわれる。

(1) 文字領域の抽出

文書画像に含まれている文字領域の取得にはOCRを利用する (Fig. 4)。本来OCRは画像中に含まれている文字の座標と種別の認識に利用される。しかしGT-Layoutにおいては後述する手法により、文字座標の情報のみで文書レイアウトの再構築が可能となるため、文字種別認識の精度を下げることで、OCR処理の速度を高速化している。



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

(2) 行領域の抽出

抽出した文字領域から行領域を構成する (Fig. 5)。ある文字を基準とし、文字右端から文字幅程度の領域内に存在する文字を隣接文字と判定し、さらに隣接文字と判定された文字に対しても、同様の判定を繰り返し抽出された文字列を行領域として構成する。文字列内に句読点など領域の狭い文字が存在する場合、通常なら行と判定されるべき領域が句読点で分断されてしまう。この分断は行の上下位置がほぼ同じで間隔が平均文字幅内の行を同一の行と判定することで、回避可能である。

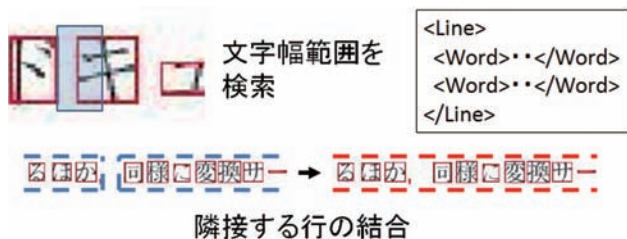


Fig. 5 Calculation of line regions.

(3) 段落領域の抽出

構成した行領域から段落領域を構成する (Fig. 6)。行下端から行高さ程度の領域内に存在する行を隣接行として判定する。この判定を繰り返すことにより、行の塊である段落領域が抽出される。

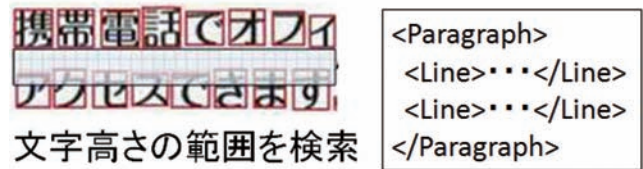


Fig. 6 Calculation of paragraph regions.

上記の (1) (2) (3) の処理により、文書画像のレイアウト情報が抽出される。抽出したレイアウト情報と文書画像から携帯端末に合わせたレイアウト画面を表示する方法を次節で紹介する。

3.3 クライアントアプリケーション

クライアントアプリケーションは、GT-Docサーバとやり取りを行わない文書画像の受信と、端末上画面への表示を行なう¹⁾。

表示中の文書画像のレイアウトを再構築する場合、GT-Docサーバにレイアウト情報取得要求を送信し、レイアウト情報を取得する。以下ではレイアウトを再構築する方法とアプリケーションUIを記載する。

3.3.1 レイアウト再構築

クライアントアプリケーションは表示中文書画像とレイアウト情報を元に、携帯端末表示領域に合わせたレイアウトを再構築する。

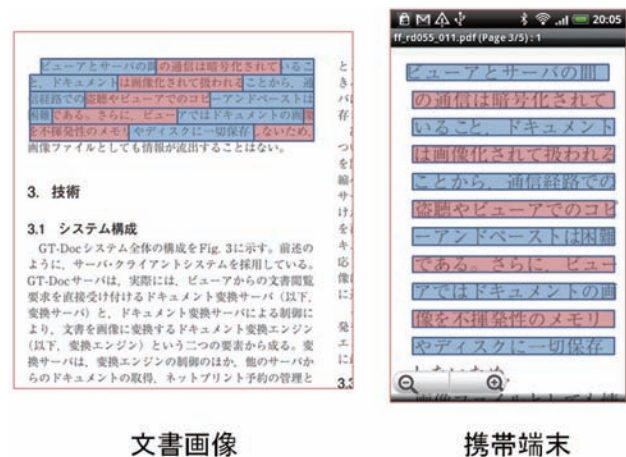


Fig. 7 Reconstruction of paragraph layout.

Fig. 7は右側に元の文書画像、左側にレイアウトを再構築した携帯端末画面を示しており、元文書の文字領域の携帯端末上で配置を色分けして表示している。レイアウトを再構築する手順は以下のとおりである。

(1) 段落の指定

レイアウト情報からレイアウトを再構築する段落を指定する。

(2) 段落領域のレイアウト

指定した段落の行領域情報を取得し、文書画像段落領

域の先頭行から順に、携帯端末の画面に収まる行領域を切り出し、端末の画面に表示する。切り出しと携帯端末での表示を行末まで繰り返し、行末に達した場合、携帯端末画面上で表示領域に余白があれば、次行の先頭から余白分だけ切り出し表示する。この一連の操作を段落の最終行まで繰り返し携帯端末画面に適した段落領域のレイアウトを行なう。

OCR情報に含まれる文字種別により文書領域画面を再構築することも可能であるが、文書画像をそのまま利用することにより、元文書のフォントや文字間/行間などの見た目の印象を、レイアウト再構築後も保つことができるという点でGT-Layoutの方式が優れている。

3.3.2 アプリケーションUI

本節ではクライアントアプリケーションのUIについて記述する。以下はクライアントアプリケーションの画面はタッチスクリーンであることを前提としているが、端末の十字キーでも同様の操作が可能である。

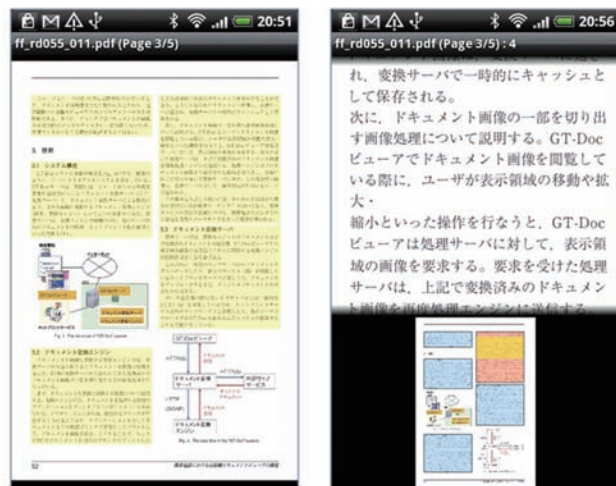


Fig. 8 Client application UI.

Fig. 8にはレイアウト情報を受信したクライアントアプリケーション画面が示されている。

レイアウト情報を取得するとFig. 8左側の段落選択画面が表示される。レイアウト情報内の段落領域情報から、レイアウト再構築可能な段落の領域を取得し、画面上の文書画像にハイライトする。これにより容易に再構築可能な段落の判定が可能となる。ハイライトされた領域をタッチすることで、段落のレイアウトを再構築する。

Fig. 8右側はレイアウト再構築した画面である。画面上部にはレイアウト再構築した段落領域が表示されており、上下スクロールのみで連続した文書閲覧が可能となっている。またスクリーンのピンチインアウトのより文字サイズの拡大・縮小などの操作を行なうことができる。画面下部にはページ全体のサムネイル画像が表示されている。現在画面上部に表示されている段落領域を赤でハイライトし、また再構築可能な段落領域を青でハイライトしている。サムネイルの青領域をタッチすること

で、画面上の再構築された段落の画面がタッチ領域の段落となり、直感的な操作で段落間遷移が可能となる上に、現在表示されている段落の文書画像内の位置がわかり、携帯端末の狭い表示領域においても、全体と部分の同時確認が可能となる。

4. まとめ

本報告では、画像技術を応用した、携帯端末における文書画像閲覧の閲覧性向上技術であるGT-Layoutについて解説した。本技術により携帯端末において文書閲覧行為が快適なものとなり、ビジネスツールとしての携帯端末の利用が促進され、業務の効率化を行なうことができると期待される。今後、本技術の機能・性能の向上に尽力するとともに、携帯端末以外での活用やさまざまな用途での利用も考慮しつつ、同技術を利用した市場開拓を進めていきたい。

参考文献

- 1) 大谷洋, 池谷直樹, 上杉英司, 浅井有人, 瀬川英生, 羽田典久. 携帯電話における高精度ドキュメントビューアの開発. 富士フィルム研究報告, No.55, 50-54 (2010).

(本報告中にある“Microsoft Office”はMicrosoft Corp.の, “Adobe”, “Acrobat”はAdobe Systems Inc.の登録商標です。“GT-Document”, “GT-Doc”は富士フィルム(株)の登録商標です。)