

携帯端末での閲覧に向けた Web コンテンツ自動変換

吉川 裕章[†] 内田 理^{††} 中西 祥八郎^{††}

インターネットにアクセス可能な携帯端末の普及に伴い、携帯端末によるWebページ閲覧の頻度が増加している。しかし一般のPC向けに作成された通常のコンテンツは、携帯端末での閲覧には適しておらず、閲覧が不可能な場合もある。このような問題を解決するため、従来は人手によってコンテンツを携帯端末向けに作り直す作業が行われていたが、これはコンテンツ管理者の負担が大きく、効率的ではない。本研究では、このような問題を解決するため、PC向けに作成したWeb ページをリアルタイムで携帯端末に適した形式に変換し、表示するシステムを構築した。

Automatic Conversion of Web contents for Cellular Phone

Hiroaki Yoshikawa[†] Osamu Uchida^{††} Shohachiro Nakanishi^{††}

By reason of spread of cellular phones that can access the Internet, the opportunity of web browsing by a cellular phone has been increased. However, the web pages that was made for the PC are not suited for browsing by cellular phones. Moreover, we can not browse the web site by cellular phones, in some cases. Ordinarily, website developers have to make over the web pages for cellular phone. But, this task is inefficient for administration of website. In this paper, to tackle this problem, we implement an automatic conversion system of the web pages for cellular phone.

1. はじめに

近年、携帯電話やPDA等の携帯端末の普及は目覚ましく、携帯電話としては、平成15年12月末において、およそ8000万件の加入が確認されている[1]。これは、日本国内では約2人に1人が携帯電話を所有していることを意味しており、携帯電話は現代を代表する情報端末の一つであるといえる。

また、携帯電話の普及に伴い、携帯電話から

のWeb利用の頻度が増加している。しかし、携帯電話に搭載されたブラウザを用いて、一般のPCでの閲覧を想定して作成されたWebコンテンツにアクセスした場合、様々な問題が生じる。

まずディスプレイの解像度の低さが挙げられる。携帯端末の性能は年々向上しているが、PCと比較すると依然として携帯端末のディスプレイ解像度は低いため、1行あたりに表示できる文字数が少ない。そのため、PC向けに作成されたWebページを表示した場合、文字列の折り返し(改行)が頻繁に起こり、文章が読み難くなる原因となる。

また、端末ごとにディスプレイの解像度やキャッシュ容量、使用できる画像のフォーマット等が異なるため、特に画像を用いたコンテンツを作成する場合はページのレイアウトや画像

[†]東海大学大学院工学研究科
〒259-1292 神奈川県平塚市北金目 1117
Faculty of Engineering, Tokai University
1117 Kitakaname, Hiratsuka, Kanagawa, 259-1292
Japan

^{††}東海大学電子情報学部情報科学科
School of Information Technology and Electronics,
Tokai University

のサイズが制限されてしまう。

更に、端末によって対応するマークアップ言語が異なる問題が挙げられる。これにより、HTMLで記述された一般的なWebコンテンツを携帯電話で閲覧した場合、端末側で表示可能な形式に変換されて表示が行われる。この際、非対応のタグが原因となり、レイアウトが乱れることや、場合によっては閲覧不可能になることもある。従来これらの問題を解決するために、コンテンツ管理者は、手動でWebコンテンツを各端末で表示可能な形式に書き換え、更にアクセス解析などにより端末の機種を判定し、アクセス先を振り分けるなどの作業が必要であった[2][3]。しかし、これでは管理者の負担が大きく、効率的ではない。そのため、コンテンツを端末に適した表示形式に変換するための様々な手法が研究されている[4,5,6]。

本研究では、管理者がPC向けに作成したコンテンツを、リアルタイムでアクセス元の端末に適した形式に自動で変換し、表示を行うシステムを作成した。

2. 目標とするシステムの概要

図1に示すように、携帯電話のディスプレイ解像度は機種によって様々である[7,8,9,10]。そのため、画像を用いたコンテンツを端末で表示する場合、使用可能な画像のサイズやフォーマットが限られてしまう。

ディスプレイ解像度の横幅(ピクセル)	機種数	ディスプレイ解像度の横幅(ピクセル)	機種数
100未満	39	140~144	14
100~104	1	145~149	1
105~109	2	150~154	0
110~114	3	155~159	0
115~119	15	160~164	8
120~124	83	165~169	0
125~129	14	170~174	0
130~134	25	175~179	9
135~139	0	180以上	29

図1. 携帯電話のディスプレイ解像度

また、端末が対応するマークアップ言語や画像フォーマットもキャリアによって異なり、統一されていない(表1)。これは同一ページを複数の端末で閲覧した際、表示結果が異なる原因となる。

表1. 携帯電話の対応言語, 画像フォーマット

キャリア名	端末名	対応言語	対応画像フォーマット
NTT DoCoMo	3G11	HTML 1.0	GIF
	2000/2000/2000	HTML 2.0	GIF
	210/200	HTML 2.0	GIF
	210/210/200	HTML 4.0	GIF, JPEG
	200	HTML 3.0	GIF, JPEG
	200/200/200V	HTML 3.0+HTML	GIF, JPEG
Vodafone (J-PHONE)	スペース/非対応機種(02型)	J-800/HTML	PNG
	スペース/対応機種(02型以外)	J-800/HTML (カラー/255色) AMR	PNG, JPEG
	117対応対応機種(02型)	J-800/HTML (117対応) AMR	PNG, JPEG
au	0200/0200/400/1800/2100	HTML	PNG, JPEG, GIF
	0200/0200/400/1800/2100/3000/4000/5000	HTML, CSS	GIF, JPEG, PNG, BMP
T-Mobile	1100/1100/1100/1100/1100	HTML 4.0	BMP
	1100/1100/1100/1100/1100/1100	HTML 4.0	BMP, JPEG

そこで、各端末に対して最適な状態にWebコンテンツを変換、配信するため、CGIプログラムによるアクセス解析によって機種を特定し、機種ごとに適したコンテンツ変換を行うことにした。

コンテンツ変換の主な手順を図2に示す。

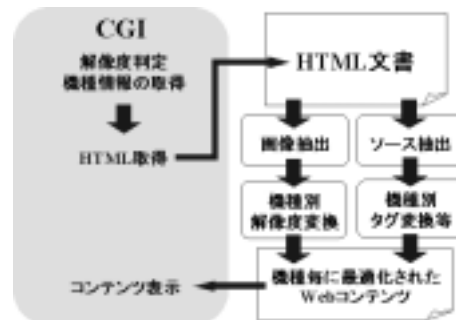


図2. コンテンツ変換の概要

3. 画像の解像度変換

先に述べた通り、携帯電話のディスプレイ解像度は機種によって異なっており、また1ページあたりに表示可能なデータ容量も異なるため、サイズの大きい画像を表示させようとする表示できないことも多い。そのため、従来はコンテンツに用いられている全ての画像に関して、機種ごとに適切なサイズの画像を用意しなければならなかった。

このような管理者の負担を軽減するため、機種ごとに画像を用意しておくのではなく、携帯電話からのアクセスの度に最適な画像が自動的に生成されることが望ましい。そこで、CGIプログラムにより端末のディスプレイ解像度を判定した後、表示対象となるコンテンツに含まれる画像ファイルを抽出し、端末に合わせて適切なサイズに解像度変換を行い、表示する。

ここで問題となるのが、コンテンツ内の画像の元の解像度と、端末側で表示する変換後の解像度である。図 3(a)のような画像（以下、画像タイプ A）は、解像度変換後もなるべく大きく表示される必要がある。逆に、図 3(b)のような画像（画像タイプ B）は、ページ全体の容量削減のためにも、大幅に縮小してしまっても構わないと考えられる。図 3(c)のような画像（画像タイプ C）は、その中間にあたる。



図 3. 画像ファイルの分類

そこで、Web 上のサイトから 500 件の画像をランダムに選出し、解像度に関する調査を行った。調査結果を図 4 に示す。

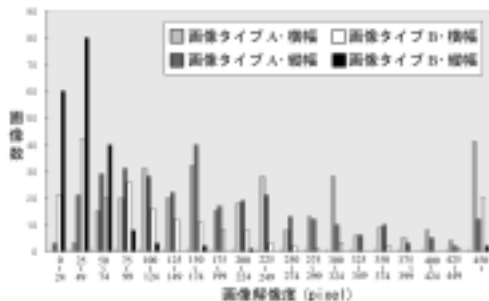


図 4. Web 上の画像解像度

この結果を踏まえ、元の画像サイズに対する変換後の表示サイズを表 2 に示すように 3 段階に分けて設定した。

表 2. 携帯電話上での画像の表示サイズ

元の画像サイズ (変換前)	変換後の表示サイズ (画面解像度に対する割合)
横130ピクセル以上 または 縦50ピクセル以上	画面の横幅または縦幅の 60%表示(タイプA)
横100ピクセル以下 または 縦60ピクセル以下	画面の横幅または縦幅の 5%表示(タイプB)
どちらにも含まれない場合	画面の横幅または縦幅の 20%表示(タイプC)

元の画像の横幅が 130 ピクセル以上または縦幅が 150 ピクセル以上の場合をタイプ A、横幅が 100 ピクセル以下または縦幅が 100 ピクセル以下の場合をタイプ B、どちらにも当てはまらなかった場合をタイプ C の画像として扱う。これらを表示する際、携帯電話のディスプレイ表示領域に対し、タイプ A を 60%、タイプ B を 20%、タイプ C を 5%の大きさに解像度変換して表示するものとした。

4. 表形式データ(table)の変換・表示

HTML で表形式データを表現するためには、通常<table>タグが用いられる。しかし<table>タグで記述された表形式データを携帯端末で閲覧した場合、ブラウザの表示性能により表示結果が異なる。

ブラウザが<table>タグに対応していない場合、タグ情報は単純に無視され、表の中に記述された文書のみがそのまま表示される。そのため、表として配置された文書の位置関係にずれが生じ、読み取ることが困難である(図 6(a))。また、<table>タグには対応していても、セルの連結に対応していなければ、セルの表示位置が狂う危険性があり、表自体が誤って表示されてしまうことも考えられる(図 6(b))。

<table>タグに対応した端末によって問題なく表形式データが表示できた場合でも、文字列の折り返しや、横方向スクロールの手間といった問題がある。端末によってはブラウザ表示時の横方向スクロールが不可能であるため、表中に書かれた文字列は、ブラウザの横幅内に収まるように強制的に改行されて表示される。このように頻繁に改行された文字列は非常に読み取りにくく、解読不可能な場合もある。(図 6(c))



図 6. 表形式データの表示の違い

以上のように、携帯端末で表形式データを正確に表示し、閲覧することは難しい。このような問題を解決するため、表形式データ内の境界線と項目名、要素等を自動判別することにより、データの構造を認識し、内容をテキストとして書き出す手法が提案されている [5]。しかし、表の内部に画像が含まれる場合や特殊な表形式データなど、[5]の方法が適用できない場合も多い。

そこで本システムでは、表形式データに含まれる要素を表から分離し、テキストとして表現するアルゴリズムを数パターン用意することにした。これにより、各種の表形式データは、必要に応じてユーザーがモードを選択することにより、テキストデータとして書き出され、内容の理解が可能となる。

具体的な手法を以下に示す。

4.1 表形式データの整形表示

図 7(a)のように、表形式データには rowspan や colspan 等[11]を用いることにより、2つ以上のセルを1つに結合しているものがある。この場合、結合されたセルを分割し、図 7(b)のように整形する。

場所	リフト券	ナイター	レンタル
白馬	¥3600		¥1500
苗場	¥3600	¥4000	¥2000

(a) 整形前

場所	リフト券	ナイター	レンタル
白馬	¥3600	¥3600	¥1500
苗場	¥3600	¥4000	¥2000

(b) 整形後

図 7. 表形式データの整形

4.2 縦に記述された表形式データの分離

続いて、整形された表から要素を分離し、テキストとして書き出す。ここで、表に項目名が含まれる場合、表から分離した要素をただ並べただけでは、スクロールや改行によって、表に含まれていた項目名と各要素の位置関係がわからなくなってしまう危険性がある。そのため、全要素を書き出す際に、もとの項目名を併記することにした。これにより、元の表の内容をテキストとして表現することが可能となった(図 8)。

場所	リフト券	ナイター	レンタル
白馬	¥3600	¥3600	¥1500
苗場	¥3600	¥4000	¥2000

→

場所 白馬
リフト券 ¥3600
ナイター ¥3600
レンタル ¥1500
場所 苗場
リフト券 ¥3600
ナイター ¥4000
レンタル ¥2000

図 8. 縦方向への分離

4.3 横に記述された表形式データの分離

4.2と同様に、整形された表形式データから、項目名にあたる部分と要素を、テキストとして書き出す(図 9)。

場所	白馬	苗場
リフト券	¥3600	¥3600
ナイター	¥3600	¥4000
レンタル	¥1500	¥2000

→

場所 白馬
リフト券 ¥3600
ナイター ¥3600
レンタル ¥1500
場所 苗場
リフト券 ¥3600
ナイター ¥4000
レンタル ¥2000

図 9. 横方向への分離

4.4 例外となる表について

表形式データの中には、先述の方法では読み取ることの出来ない特殊なものが存在する。例として、図 10 に示すように、表の 1 行目に表全体のタイトルが記述されている場合や、2 段組みの表などが挙げられる。

料 金 表			
場所	リフト券	ナイター	レンタル
白馬	¥3,600		¥1,500
苗場	¥3,600	¥4,000	¥2,000

(a) 表の 1 行目に表全体のタイトルが記述されている場合

場所	リフト券	ナイター	レンタル
白馬	¥3,600		¥1,500
場所	リフト券	ナイター	レンタル
苗場	¥3,600	¥4,000	¥2,000

(b) 2段組みの表

図 10. 例外となる特殊な表形式データ

これらについては、今後更に別の手法を検討する必要がある。

4.5 表形式データの変換・表示システム

以上を踏まえ、表形式データをテキストとして書き出すシステムを作成した。ここでは表の内容の自動判定は想定していないため、縦方向への分離、横方向への分離といった表示モードの選択は、ユーザーが必要に応じて行うものと

する。実際の変換例を図 11 に示す。



図 11. 表形式データの変換例

5. フレームを用いたページの変換

通常、携帯電話でフレームを用いたページを表示しようとした場合、図 12(a)のようにエラー処理が行われるか、機種によっては同図(b)のように、表示対象ページへのリンク一覧が表示される。このように、フレームが使用されているコンテンツを閲覧する際、携帯電話のようにブラウザがフレーム表示に対応していなければ、コンテンツを表示することはできない。



図 12. フレーム非対応機種における表示

そこで、フレーム表示に非対応のブラウザで、フレームを用いたページを閲覧するための手法について考える。

まず、表示対象ページを個別に閲覧する方法が挙げられる。これは図 12(b)のように、フレーム内の表示対象コンテンツのファイル名をリストとして表示し、このリストから各コンテンツへリンクすることによって実現できる。この際、各コンテンツはアクセス元の端末に適した形式に変換して表示する。

次に、本来フレーム表示によって別々に表示されている複数のページを、1枚のページとして連結して表示する方法が考えられる。このためには、システムから見たページのディレクト

リ階層および HTML ソースを取得し、ソースの共通部分、不要部分等を除いて連結することが必要となる。その際、ページ内に含まれるリンクを正しく動作させるため、表示対象の HTML ファイル名の他に、フレームを指定したページに記述されるフレーム名 (TARGET の指定) を随時保持していなければならない。実際の変換例を図 13 に示す。



図 13. フレームを用いたページの結合表示

また、本システムではフレームを個別に表示、または結合して表示といったモードは、ユーザーが任意に選択可能とした。表示モードの選択は、閲覧しようとするコンテンツにフレームが含まれていた場合に自動的に表示される。携帯電話に表示される。実際のモード選択画面を図 14 に示す。



図 14. フレーム表示モードの選択

6. システムの検証

本システムを、Red Hat Linux release 6.2 を搭載したサーバー上に実装し、検証実験を行った。今回の実験では、端末 1 (A5501T: 解像度 240×270)、端末 2 (N2051: 解像度 176×198)、端末 3 (J-SA05: 解像度 120×112) の 3 機種の携帯電話を用意し、画像や表形式デ

ータ, フレームを使用した Web コンテンツの携帯電話での表示を試みた. この結果, 全ての端末において変換されたコンテンツが問題なく表示されており(図 14), システムが良好に動作していることが確認できた. その他の端末においても, 一部機種を除いて良好に動作することを確認している.



図 15. 表示結果

7. まとめ

本研究で作成したシステムにより, PC 向けに作成した通常の Web コンテンツを携帯電話で閲覧可能な形式に変換し, 表示することができた. これにより Web コンテンツ管理者は, PC 向けのコンテンツと共に本システムを設置することにより, 携帯電話向けにコンテンツ配信が可能となるため, 非常に効率的である.

しかし, 現段階では画像の表示やフレームの結合などを行うことにより, 端末が 1 ページ内に表示可能なキャッシュ容量をオーバーしてしまう可能性がある. 今後の研究課題として, 出力されたコンテンツを一定量ごとに分割するなどして対応しなければならない. その際, 正しくページを分割するために, 各キャリアの対応したマークアップ言語に変換し, 正確なデ

ータ量を測る必要がある.

また, 本システムを使用するためには, CGI(Perl5.0)および JAVA(JDK1.4)が許可されている環境が必要であるが, なるべく環境に左右されず, 誰にでも手軽に設置できるシステムを構築することが目標である.

謝辞

本研究に対し有益な御助言をいただいた, 東海大学電子情報学部情報メディア学科, 菊池浩明助教授に深く感謝致します.

参考文献

- [1] 総務省 情報通信統計データベース
<http://www.johotsusintokei.soumu.go.jp/>, (2004.2)
- [2] 水野 貴明, “自分で出来る Web アクセス解析とデータ活用術”, ソーテック社, 2002.
- [3] インフォシエル, “ケータイサイト構築完全ガイド”, 毎日コミュニケーションズ, 2002.
- [4] 吉川 裕章, 高梨 崇広, 内田 理, 中里 純二, 中西 祥八郎, “携帯端末向け画像解像度変換システム”, 情報処理学会研究報告 2002-DP S-111, pp.59-64, 2003.
- [5] 塚本 修一, 増田 英孝, 中川 裕志, “HTML の表形式データの変換と携帯端末表示への応用”, 情報処理学会研究報告 2002-NL-151, pp.35-42, 2002.
- [6] 渡部 聡彦, 中川 裕志, “多種の表示デバイスへ適応可能なコンテンツ中間表現形式の提案”, 情報処理学会研究報告 2001-DD-32, pp.63-70, 2002.
- [7] NTT DoCoMo ホームページ
<http://www.nttdocomo.co.jp/>, (2004.2)
- [8] au ホームページ
<http://www.au.kddi.com/>, (2004.2)
- [9] Vodafone ホームページ
<http://www.vodafone.jp/>, (2004.2)
- [10] TU-KA ホームページ
<http://www.tu-ka.co.jp/>, (2004.2)
- [11] 越川 彰彦, “HTML 重要タグ・リファレンス”, ナツメ社, 2002.