



モバイル端末のための ユーザインタラクション技術(前編)

— 表示対話技術 —

旭 敏之
仙田 修司
磯谷 亮輔

日本電気(株) 共通基盤ソフトウェア研究所

ユーザインタフェース技術とその流れ

コンピュータにおけるユーザインタフェース(以下、UIと表記する)の概念を最初に提唱したのは、心理学者でもあったJ. C. R. Lickliderとされている。彼の代表的論文「Man-Computer Symbiosis」が発表されたのは、最初の大型科学技術計算用コンピュータIBM701が商用化されてから、ただだか8年後(1960年)のことである。

もともとヒトと機械の関係に関する研究は、目的とする作業を効果的に支援するため、道具や機械がいかに身体機能にフィットしているか、という課題に焦点が当てられてきた。これらの知見やノウハウは、やがてヒトと生産設備や産業機械との関係に引き継がれ、高生産性の達成を命題とした人間工学として体系化された。

他の機械や道具との比較において、コンピュータの際立った特徴は、使用者であるヒトとの間に高度で複雑な「対話」が存在することである。したがってUIの研究も、身体的特性とシステムとの適合性を探求する人間工学的なアプローチから、「ヒトとコンピュータとの対話をいかに円滑にかつ豊かなものにするか」といった心理的・感性的側面に関心が広がり、ソフトウェア(対話、応用技術)やハードウェア(入出力デバイス)はもちろん、設計や評価などの開発/運用技術、心理学や人文科学などヒトそのものの探求、アート技法など、幅広いスコープを有することとなった。図-1にこれらUI技術の研究領域を示す(本稿で解説するユーザインタラクション技術は、「対話技術」領域に位置する)。

UI研究の重要な成果の1つであり、ビジネス的にも

ユーザインタフェース(UI)の中心的パラダイムは、30年前に提唱されたグラフィカル・ユーザインタフェース(GUI)のそれがほぼそのまま踏襲されてきたが、近年、ユビキタス・コンピューティングの浸透とともにコンピュータの利用形態が多様化し、新たなアプローチが求められている。その中で、フルサイズ・キーボードやマウスを持たず、表示画面の小さいモバイル端末におけるユーザインタラクション技術の研究開発が喫緊の課題となっている。

本稿では、表示対話系の技術として、Webコンテンツの変換表示技術、マルチスケール表示技術を、また入力対話系の技術として、モバイル向けテキスト入力技術、携帯電話付属カメラを利用したテキスト入力技術、音声入力技術のモバイル応用、を例として採り上げ、最新の技術動向および今後の展望について2回に分けて解説する。

大きな成功を収めた技術がグラフィカル・ユーザインタフェース(以下、GUIと表記する)である。初めてGUIが搭載され、後のコンピュータ利用形態そのものに重大な影響を与えることになる端末が、1973年に発表されたXerox社のAltoワークステーションである。

Altoはビットマップディスプレイとマウスを備え、WYSIWYG(What You See Is What You Get)やWIMP(Window, Icon, Mouse, Pointing)と呼ばれる、今やおなじみになったデスクトップ環境を世に提案した。その

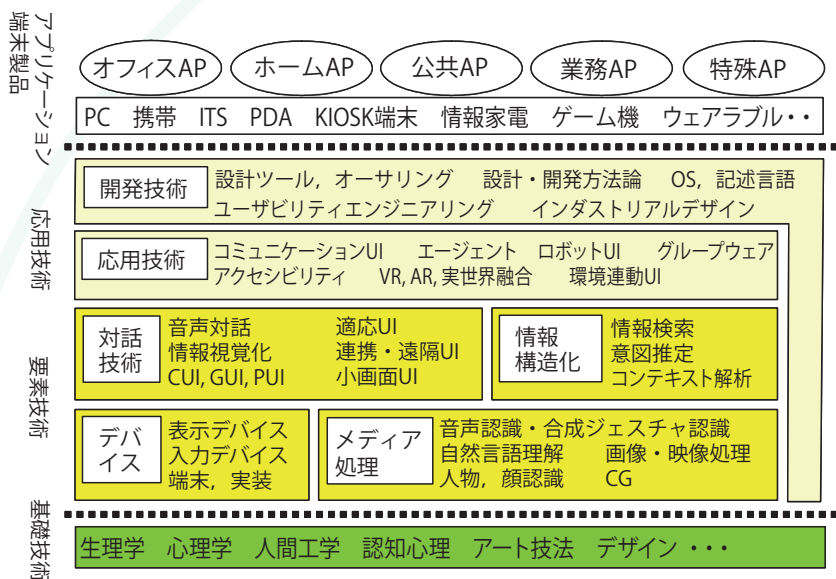


図-1 ユーザインタフェース技術の研究領域

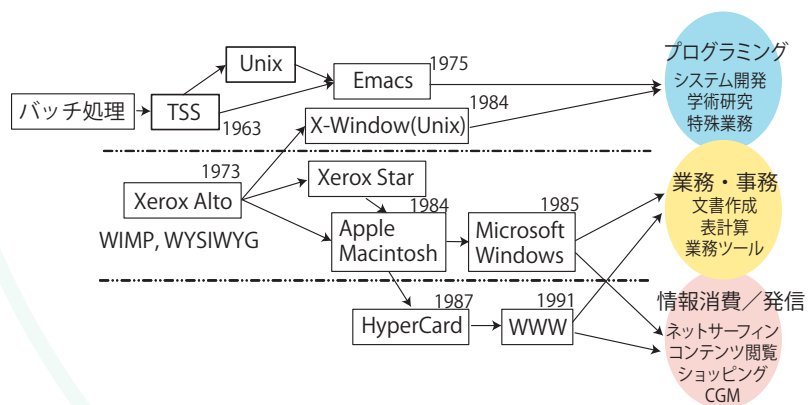


図-2 GUIを中心としたユーザインタラクション技術の変遷

後、パソコンの高性能化（速度や容量などの性能はここ20年で300～1,000倍にもなった）とともに普及が進み、また、グループウェア、ハイパーメディア、バーチャルリアリティ、エージェント、インターネット&Webコンピューティングなど、ヒトとコンピュータの関係性にかかわる新たな応用技術が次々と提案されてきたが、対話技法そのものの基本思想は、（特に実用・商用化実績を重要視すると）30年以上前のGUIのそれがそのまま継承されてきたと言える（図-2）。

ユビキタス時代における課題

進化の激しいコンピュータ技術の中で、長らくGUIが主流であった理由としては、利用場面がオフィス／工場や家庭内に、また端末形態がワークステーションやPCに限定されてきたことが大きい。しかし近年、IT技術とともに社会のユビキタス化が進展し、場所や時間の制約を超えてITサービスを活用したいというニーズが

顕著になってきた。これに応えるかたちで、通信やコンテンツ／サービスとともに、端末の多様化が急速に進んでいる。

ユビキタスは、Xerox EuroPARCの研究者であったM. Weiserが1991年に提唱したコンピュータの利用概念であるが、そこで描かれたのは極小化したコンピュータが環境の中に埋め込まれ、ヒトがその存在を意識せずにサービスが享受できる世界であった。現実のユビキタス化は、携帯電話などモバイル端末の普及とRFIDなど無線タグの実用化が中心になって進展している。

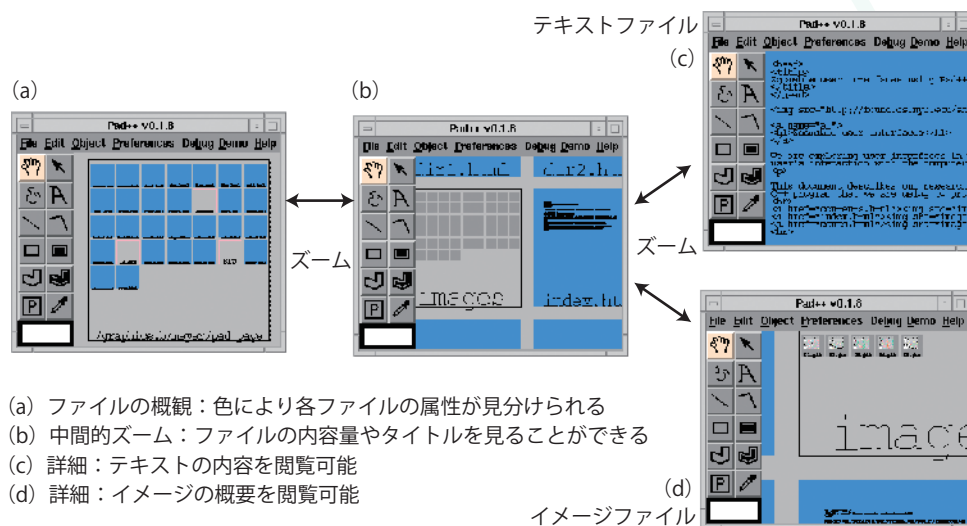
モバイル端末の中でも携帯電話の普及は目覚しく、幼児を除くあらゆる層の利用者が日常的に所持する端末となった。当初はCPU性能、メモリ容量、通信容量といったリソース面の制約による技術的ハードルがあったが、（日本では）1999年のi-mode™のサービス開始以来、メールや専用サイトやJavaアプリの利用など、携帯電話によるインターネット利用が急速に進んだ。また、GPSやRFIDなどのセンサを組み合わせ、環境と連動したモバイル端末ならではのアプリケーションが開発／実用化されてきた。これに伴い、さまざまなサービス／コンテンツを携帯電話で利用したいという

要求が、ビジネス、ホーム、パーソナルにかかわらずあらゆる生活場面で高まってきている。

ユーザインタラクションの視点から見たモバイル端末の特徴（制約）は、1）表示画面が小さく、2）入力・操作の制約がフルサイズ・キーボードやマウスなどPC標準の入力デバイスに比べて大きい点である。このため、WIMPで構成されるGUI技法をそのまま適用することは困難である。これら制約を克服し、使いやすく快適な環境を利用者に提供することが求められている。

このような動向を背景にして、「モバイル端末やデバイスの制約を克服し、使いやすく快適な対話技法を確立する」ことは、ユーザインタラクション研究の重要な課題となっている。本稿では、表示系の対話技術の例として、マルチスケール表示、Webページ変換表示について、また入力系の対話技術として、携帯電話のキーやカメラを利用した文字入力や音声による入出力について解説する。

なおモバイル端末としては、本稿で念頭に置いた携帯



(画像は Maryland 大学の Pad++ 説明サイト "<http://www.cs.umd.edu/hcil/pad++/>" から著者許諾を得て転載)

図-3 Pad++ によるファイルブラウジング

電話や PDA (Personal Digital Assistant) のほかに、ゲーム端末や音楽端末、マイクロ PC など普及／実用化が進んでおり、それぞれ特色ある UI が搭載されているが、紙面の制約上から解説を省いた。また、TV はモバイル端末ではないが、小画面端末と同様の課題があるために、本文において一部応用例として触れた。

表示対話技術：マルチスケール表示

「画面が小さい」「解像度が低い」など、一般に表示画面の制約が強いと一度に提示できる情報量が抑えられ、大局的な情報(全体像や構造など)と局所的な情報(情報の詳細)の両方を把握することが困難になる(この問題は "Small Screen Problem" と総称される。ちなみに、大画面／高解像度ディスプレイでも、TV 画面や公共パネルのように視距離が長い場合には同じ問題が発生する)。この課題に対して、これまで、以下のような手法が試みられてきた。

[3 次元表現]：奥行きをグラフィカルに表現することで、見かけの表示量を増加させたり、情報構造を分かりやすく提示するもの。代表的な例として Cone Tree や Perspective Wall (共に Xerox PARC) などがある。

[半透明表現]：詳細な情報の上に大局的な情報を重ね合わせて表示する手法。地図に適用した例として Macroscopic (MIT) がある。

[歪み表現]：着目個所を詳細に、それ以外の部分を小さく表示する手法。Fisheye View (ATT Bell 研)、DUALQUEST (NEC) などが知られている。

[ズーム／パン]：情報の表示粒度をユーザ操作でコントロールする手法。表示の変化をアニメーションで見

せることで、階層的な構造が把握しやすくなる効果がある。

この中で、ズーム／パン効果を利用したものはマルチスケール表示技術と呼ばれる。大量情報を表示する際にも、歪みや隠蔽 (Occlusion) を伴わない点が長所とされる。Pad や Pad++¹⁾ はその先駆的な例であるが、無限の広がり可能な 2 次元空間に配置されたオブジェクトに、利用者はパン／ズーム操作を駆使してアクセスする。図形描画や情報の検索・閲覧だけでなく、ファイルやドキュメントのブラウジングなど、従来のデスクトップ・インタフェースに替わる環境としても考案された(図-3)。

Pad や Pad++ が持つ連続的なズームイン／アウト機能により、利用者は自分が意図する特定の詳細情報に辿り着くことができる。ズーム機能を簡単かつ直感的な操作で提供することで、さまざまな詳細度や階層のビューを迅速に行き来することが可能になり、表示されているデータが全体のどの部分にあるかが把握しやすくなる効果が生じる。また、単に表示内容全体を一律に拡大・縮小するのではなく、データの属性を考慮しながらオブジェクトごとに拡大・縮小率をコントロールする技術 (セマンティック・ズーム手法) も提案されている。

ズーム型のユーザインタフェースが商用化された例としては、GeoPhoenix 社の ZoominatorTM がある。ZoomspaceTM と呼ばれる表示 UI では、画面上に配置されたアイコン型のメニューを選択することで、より詳細なメニューや内容が提示され(ズームイン)、簡単な操作で上位画面に戻ることができる(ズームアウト)。擬似アニメーションによって階層間の移動を連続的に表現する

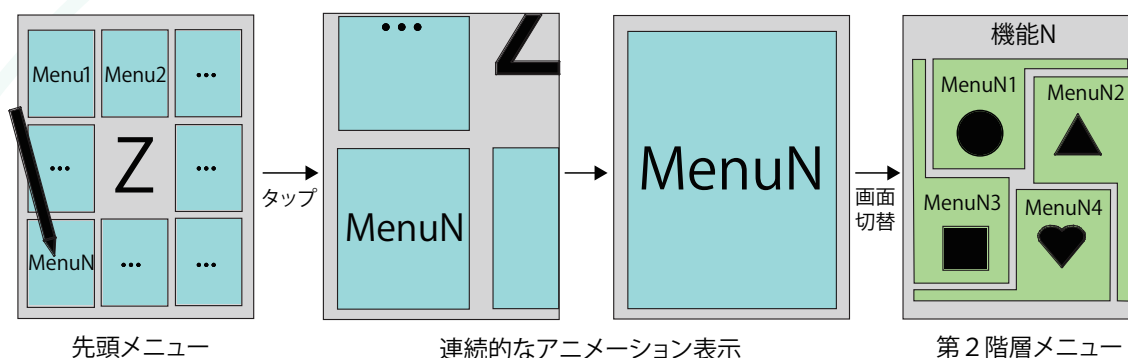


図-4 Zoomspace によるメニュー選択 UI の例

ことで、ズーム感覚が強調されている (図-4)。本方式の特長は画面サイズにかかわらず一貫した操作感を提供できることであるが、特に PDA や携帯などの小画面で大量のデータや機能メニューを検索／閲覧／選択する際に効果が大きいと考えられる。

以上、表示画面の制約を克服する手段として、ズーム／パン効果を利用したマルチスケール表示技術を紹介した。これらの技術により、大局的な情報と局所的な情報の両者を把握しつつ、メニューやディレクトリなど階層構造を持つ情報や 2 次元的な広がりを持ったデータに迅速にアクセスすることが可能になる。

今後は一層の実用化・普及に向けて、このような UI を短期間・低コストで開発するための構築環境の開発や、既存アプリケーションへの適用を容易にする工夫が求められる。また、より効果的で汎用的な技術を目指して、他の手法 (3 次元表示や半透明、歪み表現など) との融合や、階層化されていない情報などへの展開が課題である。

表示対話技術：Web ページ変換表示

Web が情報インフラとして生活やビジネスに浸透した現在、モバイル端末や情報家電などの非 PC 端末を利用して Web サイトにアクセスしたいという要望は強い。世界で 1 億を超えたとされる Web サイトであるが (Netcraft 社 2006 年 11 月調査)、そのほとんどが PC 端末での利用を想定して設計されたものである。すなわち、その UI は画面上に表示されたリンクや操作オブジェクトをマウスカースルでポインティングするものであり、基本的には GUI コンセプトが継承されていると言える。これを非 PC 端末で利用する際には、端末のリソースや記述言語など実装上の問題だけでなく、Small Screen Problem や、入力デバイスの操作性の低さに起因するユーザビリティ上の問題が発生する。これを解決するためには、大まかに、①非 PC 端末向けの専用サイトを制作

する、② PC 向けサイトを非 PC 端末に適した形式に変換する、の 2 つの方策があるが、ここでは②を自動的に行う Web ページ変換表示技術とその応用について解説する。

変換例

もともと PC 向けに設計された Web ページを、非 PC 端末に適した形で表示する技術が Web ページ変換表示技術である。筆者らが開発した「セマンティックズーム」はその 1 つであり、以下に示すように携帯電話やテレビ画面に適した形で Web ページを表示することが可能になる。

・携帯電話向け(テキスト主体)

一般の Web ページを i-mode や EZWeb ™ などテキスト表示主体の携帯用ブラウザで閲覧可能にする変換技術である。通常は文字情報を抽出して表示することが基本であるが、情報量の多いページの場合は、利用者に過剰なスクロール操作を強いることになりかねない。そこで本手法では、1 ページを適当な矩形領域 (セクション) に区切り、各セクションごとに“タイトル”と“内容”テキストを分けて抽出する。利用者はまずタイトルだけを集めた“目次ページ”を閲覧し、興味のあるタイトルを選択することで、所望の内容を閲覧することができる (図-5)。

・携帯電話向け(グラフィック主体)²⁾

フルブラウザを装備したモバイル端末向けの変換表示技術である。現行のフルブラウザでは、全体を縮小して表示するオーバービュー表示機能が提供されているが、通常、この状態では文字が読めず内容を把握することは難しい。

本方式では、オーバービュー表示において適当なセクションごとにカーソルが移動し、そのセクション内の“タイトル”を拡大表示する方式が実現されている。カーソル移動が迅速になるだけでなく、タイトルを読むことでそのセクションの内容を把握することができる (図-6)。

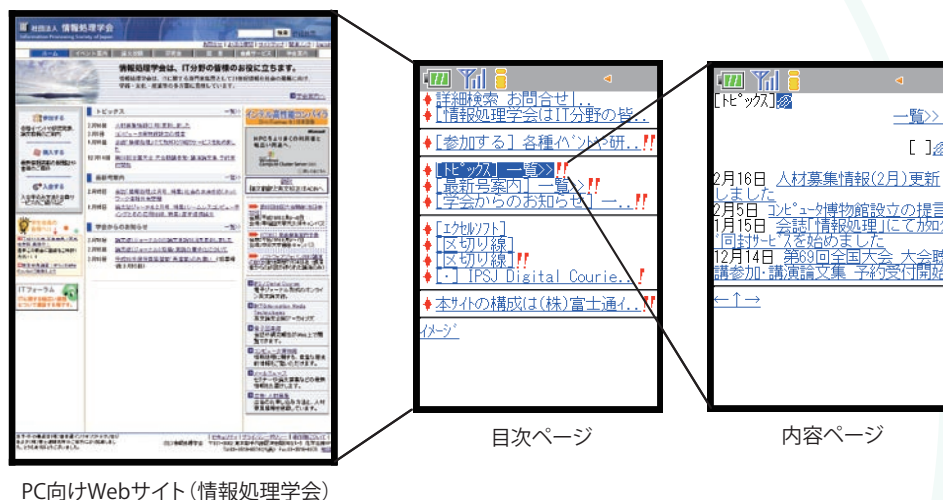


図-5 Web ページ変換表示例(1) 携帯電話向け(テキスト主体)

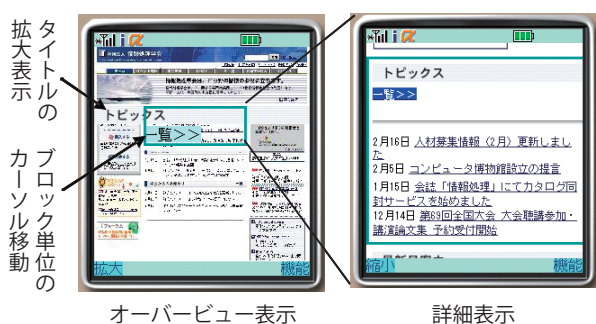


図-6 Web ページ変換表示例(2) 携帯電話向け(グラフィック主体)

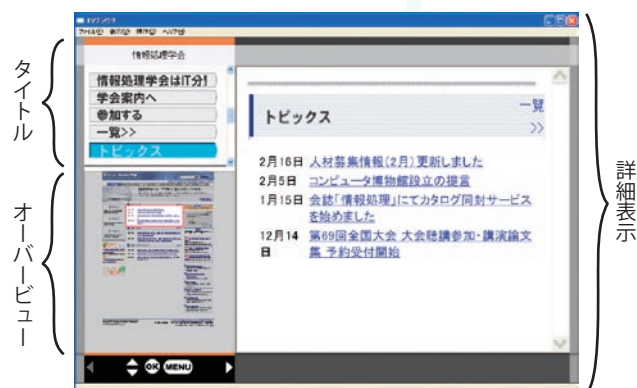


図-7 Web ページ変換表示例(3) TV 画面向け

• TV 画面向け

放送／インターネット融合の進展により、TV 画面で Web を閲覧したいという要求が高まっている。TV の場合、ハイビジョン対応の大画面・高精細端末が普及してきているが、利用者の視距離が長いため、小画面端末と同様の問題（拡大表示しないと文字等の詳細情報が読み取れない）がある。本方式では、画面の左にオーバービュー画面とカーソルを提示し、右の画面にカーソル位置に相当する部分を拡大表示することで、全体的な位置情報を把握しながら詳細な情報が参照できるように工夫されている(図-7)。

レイアウト解析

以上、説明したような変換表示を実現するためには、Web ページを適当な小領域（セクション）に分割し、かつ各セクションを代表するタイトルを抽出する必要がある。以下では、これを実現するためのレイアウト解析技術を説明する³⁾。

本来、Web ページの構造はタグによって論理的に与えられているが、単純にこれをセクションとすると、往々にして、ページを見た人間の印象とは隔たったものになる。Web ページの記述スクリプトである HTML では、多彩で柔軟な視覚表現が可能であり、実際にはこれらの組合せでセクションの区切りやタイトルが表現されるからである。したがって、レイアウト解析では HTML 文書が持つ論理情報(名前、属性、構造など)、表現情報(位置や色、テキストサイズ、フォントウェイトなど)、内容情報(内部テキストやデータに関する情報)を併せて解析することが必要になる。

ここでは、できるだけ見た目の印象に近い形でセクションとタイトルを抽出するために考案された、以下の3段階処理によるレイアウト解析技術を紹介する。

- (1) ブロック抽出：画面を矩形領域(ブロック)に分割。
- (2) タイトル抽出：画面内のタイトルを判定。
- (3) セクション抽出：ブロックをセクションにまとめあげる。

これらの処理に込められた意図は、(1) 画面を見た目の

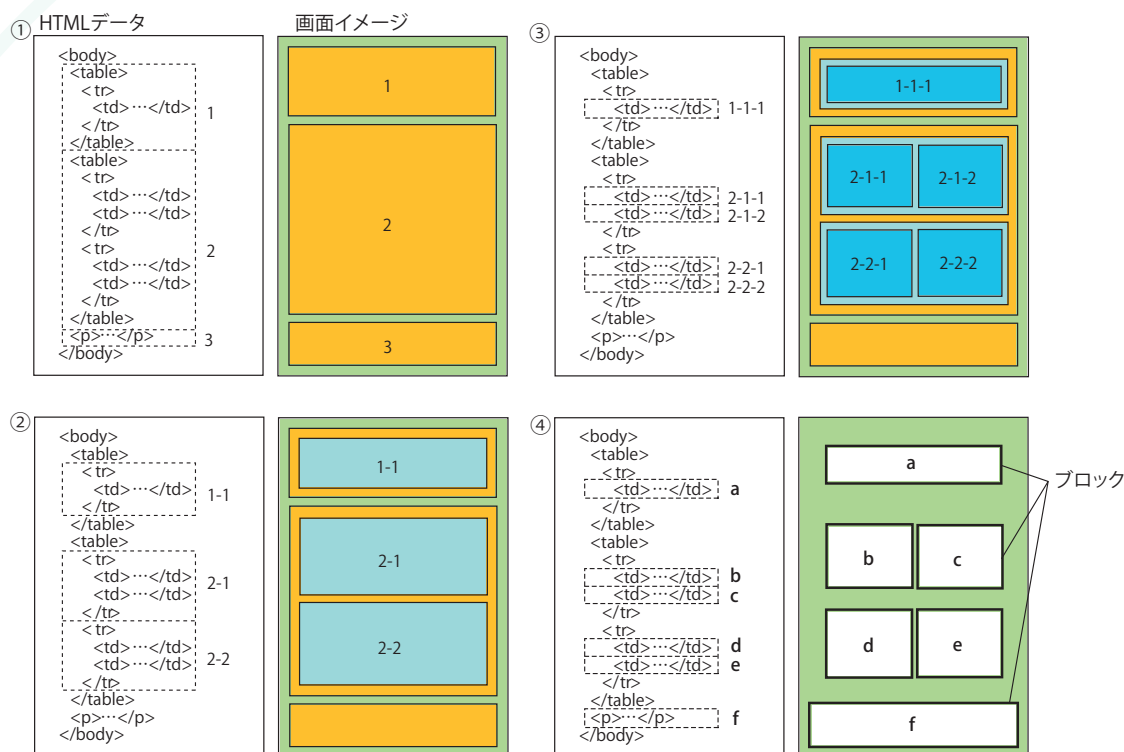


図-8 ブロック抽出の処理イメージ(第1階層)
 ①② 画面全幅を水平に分割する区切りを抽出
 ③ ブロック内を垂直に分割する区切りを抽出
 ④ 1階層目のブロックとして確定

区切りで小領域に分割し、(2) タイトルを抽出し、(3) このタイトルを手がかりにして (1) の小領域を意味のあるセクションにまとめ上げる、というものである。以下、各処理について簡単に説明する。

ブロック抽出

HTML タグの中で、表示上の区切りを生じるもの（以下、レイアウト形成タグと呼ぶ。table, tr, td, p, div, h1 ~ h6 ... などが相当する）に注目し、HTML 文書がレンダリングされたときのブロック構造を求める。まず HTML の <body> タグ内のレイアウト形成タグを検出し、画面全体を水平に分割する区切りを探す。次に分割された各領域について垂直に分割できる区切りを検出する。水平、垂直の2つの分割が終了した時点で、第1階層目のブロックが形成される（図-8）。この処理を各ブロックに対して再帰的に繰り返していくことで、階層ごとのブロックを抽出していく。

タイトル抽出

HTML 文書ではタイトルを h1 ~ h6 タグで指定することになっているが、実際の Web ページにはそれ以外にも見かけ上の特徴（たとえば、文字サイズ、色、他と異なる背景など）でタイトルに見える文字列も多い。これらをできるだけ漏れなく抽出するため、たとえば図-9

タグ情報	h1-h6 要素の場合は無条件に見出しと判定 img 要素の場合はサイズや内容を考慮
表現情報	目立つ背景色やイメージを持つ テキストサイズが閾値以上である フォントウェイトが閾値以上である サイズ（幅や高さ、アスペクト比）が適当
内容情報	img 要素であり、alt 属性を持ち、使用回数が適当 img 要素ではないが、内部テキスト数が適当

図-9 タイトルらしく見える条件

に示すような「タイトルらしく見える条件」をルール化して判定に用いている。

図-10 に、情報処理学会の Web ページに本手法を適用した例を示す（オレンジ色に着色された部分がタイトルと判定されている）。

セクション抽出

先に述べたようなブロック抽出処理を行うと、ブロックが必要以上に細分化されてしまうことが多く、またそれぞれのブロックが必ずしもタイトルを含むとは限らない。セクション抽出は、タイトル抽出で得られたタイトルを用いて、適当な数のブロックを、タイトルを持ったセクションとしてまとめ上げる処理である。

本処理は基本的に、①ブロックをタイトルを含むものと含まないものに分け、②タイトルを含まないブロック



図-10 タイトル抽出処理の結果例

を“近くの”タイトルを含むブロックに併合し、③それ以上併合するブロックが存在しないまとまりをセクションとする、ことにより実行される。ただし、この処理はブロックの階層ごとに施され、応用(変換表示)によって最適な階層のセクションを選択できるようになっている。

以上、Web ページ表示変換技術として、セクションとタイトルを用いた表示変換と、これを実現するためのレイアウト解析技術を紹介した。これらにより、PC 向けにデザインされた Web ページを、モバイル端末からでも迅速に閲覧することができる。

ただし、Web ページは表現の自由度が高く、また Java や Flash など新技術の開発・普及も速いため、汎用的に適用できる決定的なアルゴリズムはないと言える。それぞれの応用に応じて必要な機能・性能を設定し、アルゴリズムも含めてチューンアップする必要がある。ただし、本稿で述べたような解析技術には 100% の精度は望めないし、そもそも実際の Web ページの中には、人間でもタイトルやセクションの同定が難しいものがある。

解析誤差があったとしても、実用上支障がないような対話方式の開発も必要になるだろう。

なお、本稿では簡単のため HTML 文書を直接解析するイメージで説明したが、実際にはブラウザなどで生成される DOM (Document Object Model: Web ページのコンテンツやその配置、スタイルをオブジェクトとして扱い、スクリプト言語を使って制御するための規則) ツリーを用いて処理している。

次号に向けて

モバイル端末のためのユーザインタラクション技術の例として、マルチスケール表示技術と Web ページ変換表示技術を紹介した。次号では、主として入力系の対話技術に焦点を当て、テキスト入力や音声入出力技術について解説する。

参考文献

- 1) Bederson, B. and Meyer, J.: Implementing a Zooming User Interface: Experience Building Pad++, Software Practice & Experiences, Vol.28, No.10, pp.1101-1135 (1998).
- 2) 野田尚志, 池上輝哉, 辰巳勇臣, 福住伸一: 多様な端末に適応した Web 閲覧方式の提案 (1), ヒューマンインタフェースシンポジウム 2006 論文集, pp.955-958 (2006).
- 3) 辰巳勇臣, 森口昌和, 旭 敏之: DOM の画面レイアウトに基づく Web ページのセグメンテーション手法, 第 8 回 Web インテリジェンスとインタラクション研究会論文集, pp.13-18 (2007).

(平成 19 年 4 月 3 日受付)

旭 敏之 (正会員)

t-asahi@bx.jp.nec.com

1984 年大阪大学大学院基礎研究科修士課程修了。同年 NEC 入社。1997 年奈良先端科学技術大学院情報科学研究科博士後期課程修了。ユーザインタフェースの研究開発に従事。現在、NEC 共通基盤ソフトウェア研究所勤務。ヒューマンインタフェース学会会員。

仙田 修司 (正会員)

s-senda@ap.jp.nec.com

NEC 共通基盤ソフトウェア研究所勤務。1996 年京都大学大学院工学研究科情報工学専攻修士課程卒業。工学博士。同年 NEC 入社。手書き/印刷文字認識、ユーザインタフェース、画像圧縮応用の研究に従事。電子情報通信学会会員。

磯谷 亮輔

r-isotani@bp.jp.nec.com

1985 年東大・工・計数卒業。1987 年同大学院修士課程修了。同年 NEC 入社。以来、音声認識、自動通訳システムの研究開発に従事。現在、NEC 共通基盤ソフトウェア研究所主任研究員。電子情報通信学会、日本音響学会各会員。