22nd Fuzzy System Symposium (Sapporo, Sept. 6-8, 2006)
6E3-2

サムネイル表示によるブラウザ履歴情報の可視化
Web Browser with Thumbnail View Interface for Web Browsing History

新美　礼彦 [1)]　片山　悠樹 [2)]　小西　修 [1)]
Ayahiko Niimi　Yuki Katayama　Osamu Konishi

1) 公立はこだて未来大学 システム情報科学部
School of Systems Information Science, Future University-Hakodate
2) 味の素システムテクノ株式会社
Ajinomoto System Techno Corporation

Abstract: The recent years, the Internet is rapidly spread, information gathering on various fields become easily. A lot of people use web browser to gather information, they want to visit again the pages that they visited in the past. Because the title and URL on the page which it looks for are not necessarily memorized, a target page may be not searchable well only by the browsing history of a browser. In such case, it is helpful to remember the page's visual image that the screen on each page can be thumbnail displayed. In this paper, we proposed that the web browsing history is made visible by a thumbnail display on each web page. Using our system, user can do easily in the step where the retrieval of the web page visited once is shorter. We constructed it as an enhanced feature of Mozilla Firefox. When the web page is read by a browser, the flow for which the history of making to visible is made with this system preserves the title, URL, and the visit date in the XML file, and makes the thumbnail. The history with the thumbnail can be displayed by opening the XML file by a browser.

1　はじめに

1993 年に NCSA によって発表されたブラウザ・Mosaic が登場した。これ以降、インターネットという言葉は、広く世間に認知され、研究的・学術的なものから、一般のユーザに利用されるようになった。近年はパソコンの性能が向上し、安価な値段で購入できるようになった上に、ネットワークインフラが整備されて、ブロードバンド環境が普及してきたために、インターネットはもはや生活する上で必需品になってきたといっても過言ではない。

多くのユーザは Web ブラウザ (以下ブラウザ) を用いて必要な情報を収集しているが、その時に一度訪れたことのある Web ページを再度訪問したいと思うときがある。その際によく使われる手段として、頻繁に訪れる Web ページや興味がある Web ページを手動で登録する “ブックマーク” 機能が挙げられる。しかし、ブックマークは全ての Web ページに対して行うのではなく、訪問した時に登録したいと思った Web ページに対してのみ行うので、最初は重要視せずにブックマークしなかったが後日再び訪問したいと思う Web ページを探したいという要望には応えられない。

次に目的の Web ページを探すための方法として考えられるのは、最初に訪問したときの検索方法を再現することである。検索エンジンから行ったなら、そのキーワードを思い出して入力する、別の Web ページからのリンクで行った場合は、リンク元の Web ページに行くなどがある。しかし、これらの手段は全て、外部のシステムに依存する方法なので、検索エンジンの出力結果が更新されたり、リンク元の Web ページからのリンクが切れたりといった、外部システムが変更したために、目的の Web ページに到達できないという場合がでてくる。

これらを解消する手段として、履歴機能が挙げられる。ほとんどのブラウザは Web ページの訪問履歴を一定日数保存しているので、これを参照することにより、外部に依存せず、確実に目的の Web ページに到達することができる。しかし、ブラウザ標準の履歴機能は、URL やタイトルを時系列順に羅列させているだけである。一日に数十〜数百の Web ページを訪れている

中で、過去に訪れた Web ページの URL やタイトルを覚えている可能性は低く、せっかく履歴情報を保存していても有効に活用するのは難しい。

本研究では、ブラウザの訪問履歴のデータを可視化するシステムを提案する。本研究では、履歴を用いて Web ページを探す際に、タイトルや URL ではなくレイアウトなどの見た目に着目した。上記で述べたように、数ある Web ページのタイトルや URL を記憶することは困難である。しかし、視覚的にどんなページであったかをなんとなくでも覚えていれば、再び画面を見た時に思い出すことは可能かもしれない。そこで、ブラウザの訪問履歴を各 Web ページのサムネイル表示による可視化を行う。それにより一度訪れた Web ページの検索に履歴情報から簡単に行うことができるようになる。

本研究でのシステムは拡張性の高さや、ブラウザのセキュリティ対応の状況などを考慮し、Mozilla Firefox の拡張機能として構築を行った。本システムで可視化履歴を作成する流れは、ブラウザで Web ページを読み込んだ際に、タイトルや URL、訪問日時を XML ファイルに保存し、同時にサムネイル画像を作成する。作成した XML ファイルをブラウザで開くことで、サムネイル付きの履歴を表示する。

このようなシステムを使うことにより、ユーザは自分が今までどのような Web ページを訪問していたかがわかりやすくなり、探索している Web ページを一目でわかるようになると考えられる。我々は本システムをユーザが Web ページを閲覧する際の新たな手段の一つとして提案する。

2 Web 閲覧履歴の可視化

可視化とは、普段は見えない、見えにくいもの (情報) を見えやすく表示することであり、それにより、今まで理解できなかったものが、直感的に理解することができるようになる。

コンピュータの世界においての可視化は、シミュレーションによる科学技術計算の結果を、コンピューターグラフィックスを用いて画面に表示する科学可視化 (Scientific Visualization) が一般的であった。科学可視化は学術的な傾向が強く、可視化の対象が物理現象であることが多いので、物理的制約により可視化手法がある程度決まってしまう場合が多い。それに対し、抽象的な情報を効果的に表示することにより、ユーザが情報を理解することの助けとなる情報可視化 (Information Visualization) が近年注目されている。

情報可視化の例として、本研究で扱う Web データにも関わってくるネットワーク情報の可視化について述べる。ネットワークの骨組みがツリー構造をなしている場合、可視化が非常に有効である。ツリー構造の可視化手法は今までにいくつも提案されており、グラフ解析部分は Google API を用いて実装を行っている例もある。 [1, 2] Web データの可視化はサイト構成や Web サイト間の関係の可視化を対象にする研究と、ユーザの閲覧履歴の可視化を対象にする研究とがある。 [3] 前者は、サイト構成を検討したり、Web ページからコミュニティを発見するなどに使われる。後者はユーザの閲覧支援のために使われている。

閲覧履歴の可視化では、閲覧順番をグラフ表示したり、ページのサムネイルを表示して履歴の可視化を行っているものが多い。

iRider は、Wymea Bay 社が開発した IE のエンジンをベースに開発されたサムネイル表示を行う Web ブラウザである。 [4] 図 1 で表しているように、訪問履歴を更新日時順にサムネイル付で表示する。基本的にはブラウザを閉じると、表示されていた全ての履歴が消去されるが、杭のようなアイコンを打つことで、次回立ち上げたときにもその Web ページの履歴が残るという特長を持つ。しかし、サムネイルはブラウザを立ち上げる度に Web ページを読み込んで作成するため、残している履歴の数が増えると、その分だけブラウザを立ち上げたときの動作が重くなる。iRider では、ブラウザを開いてから閉じるまでの短時間の履歴の可視化を目的としていると考えられる。本研究では、数日から数週間単位での履歴の可視化を目的とした。

browseback は、Web ブラウザの履歴情報を取り込んで、Web ページのサムネイルを作成して履歴の可視化を行うソフトである。 [5] browseback は履歴の閲覧や検索に重点を置いたソフトであり、Web ページの閲覧はほかの Web ブラウザを使うようになっている。図 2 のようにサムネイルが積み重なって画面上に表示され、マウスを置いたサムネイルが選択され、履歴の詳細情報が表示される直感的なユーザインタフェースをもつ。優れたユーザインタフェースを持つが、Web ブラウザと browseback の 2 つの操作を行わなければならない。本研究では、普段使うブラウザの機能拡張としての実装を目的とした。

Mozilla Firefox は Mozilla Fundation が 2004 年に発表したブラウザで、発表以来、急激にシェアを伸ばしているブラウザである。 [6, 7] このブラウザの特徴とし

図 1: iRider の履歴機能

図 2: browseback の履歴機能

てオープンソースである、クロスプラットフォームの実現している、非常に優れた拡張性を持つ、がある。ブラウザの仕組みは以下の図 3 のような構成をしている。

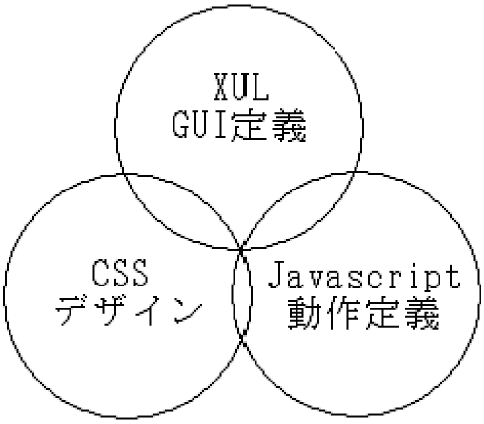

図 3: Firefox の GUI の仕組み

デザインと GUI 定義が分かれており、動作は Javascript によって拡張可能であることから、ユーザごとのインタフェースやデザインを簡単に実現することが可能である。Firefox では、Extension としてオリジナルにない機能を追加することが可能となっており、様々な拡張機能が公開されている。 [8, 9] 履歴情報は、一般的なブラウザと同様に、ブラウザは Web ページを読み込む度に、URL、タイトルなどの情報を一定日数分、記録している。履歴情報は、履歴インタフェースにより、日時ごと (今日、昨日、2 日前など) にフォルダ分けされ、閲覧した順番にソートされて表示される。Firefox の履歴情報は history.dat に、URL、タイトル、訪問回数、ホストネーム、最初に訪れた日時、最後に訪れた日時などが RDF 形式で記録されている。Firefox の履歴情報を拡張する Enhanced History Manager が公開されている。 [10] (図 4 参照) この Extension では、標準の履歴機能に比べ、最初に訪問した日時やホスト名など、多くの履歴情報を表示できる。また、より細かい条件を指定しての検索が可能であるが、視覚的に表示する機能は搭載されていない。

履歴情報ではないが、Google の検索結果のページをサムネイルとして表示する GooglePreview も公開されている。 [11] この Extension では、thumbshots.org や alexa.com によって公開されているサムネイル情報を用いて実装されている。そのため、イントラネット上の Web ページなどはサムネイル情報が表示されないという問題がある。

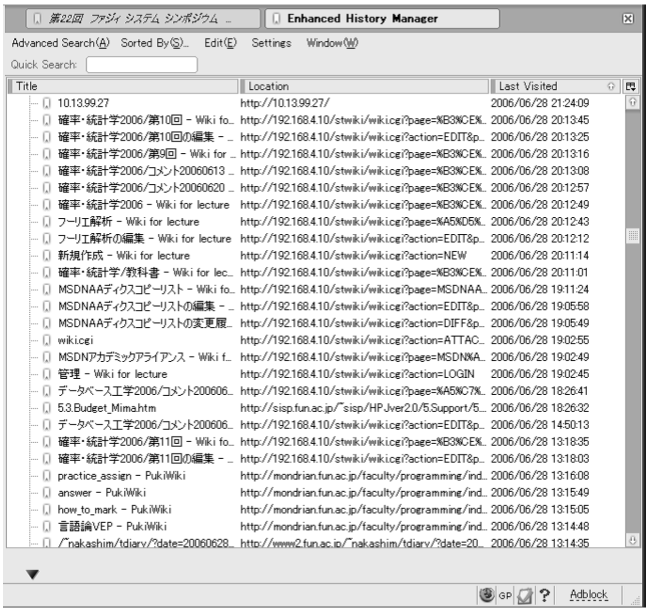

図 4: 既存の履歴機能

3 提案するシステム

iRider 等の履歴を可視化するブラウザはセッション情報の可視化しか行わない。しかしながら、履歴を可視化を最も必要とする場面は、辿り着いた経路を思い出せない Web ページを探索する時であると考えられるため、本研究では蓄積している履歴情報を可視化する。本システムにおいての動作は以下のようになる。

1. 閲覧履歴を XML に保存する。
2. 閲覧履歴を使い、Web ページが読み込まれる度にサムネイルを作成する。
3. 閲覧履歴の XML ファイルを開くと、サムネイル付きの履歴を表示する。

本システムはサーバを介さずに、クライアント側のみで実行・処理を行う。ブックマークの場合は、ソーシャルブックマークのようにサーバで保管して、インターネット上で複数の人と共有するシステムが幾つも公開されている。しかし、履歴はブックマークに比べ遥かに個人情報的要素が強く、基本的に他人が自由に見ることができるのは好ましくない。そのため、クライアント側で保存することにより、外部へ流出する危険性を減少させることができる。また、イントラネット上の Web ページなどの外部からアクセスできないページのサムネイルも作成することが可能となる。

本システムは、Firefox の拡張機能 (Extension) として構築した。これは、Microsoft 社の Internet Explorer(以下 IE) に比べてシェア率は落ちるが、Firefox は拡張の自由度が高いことと、IE に比べてセキュリティの安全面での不安が少ないと言われている点を考慮したからである。

閲覧履歴は Firefox の history.dat を用いずに、Firefox で履歴情報を取るための Extension の Slogger を利用した。 [12]

本システムを用いて、サムネイル付履歴表示を行っている様子を図 5 に示す。

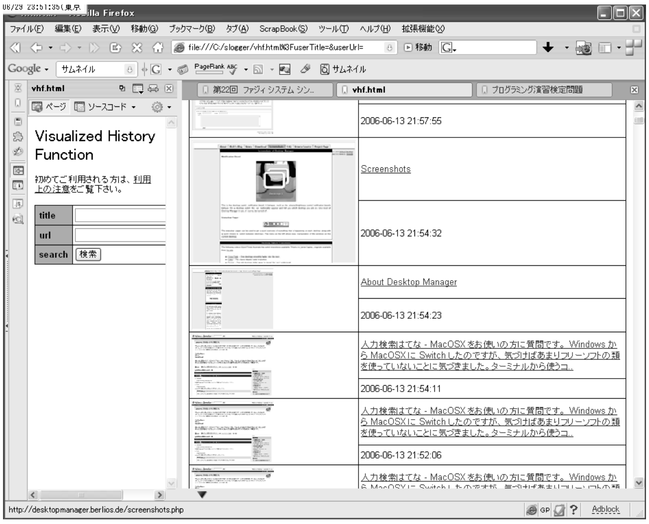

図 5: サムネイル付の履歴機能

閲覧履歴の XML ファイルを XSL によって変換し、CSS ファイルを用いて出力する項目とレイアウトを決めている。画面上の URL またはサムネイルをクリックすることにより、該当のページを開くことができる。

図 5 で 1 番目の画像が大きく表示されている。本システムでは、MSXML の機能によりサムネイルの拡大とタイトル・URL 検索が行えるようになっている。Firefox でも “IE View” や “IE Tab” [13] といった、タブを IE のエンジンを使って表示できる Extention を用いて表示すると、次の機能を利用することができる。

1. サムネイルにマウスを合わせると、サムネイルを拡大して表示する。
2. タイトルと URL をキーワードとして、履歴の検索表示ができる。

入力フォームの様子を図 5 の左側に示す。インタフェースはもうけていないが、XML ファイルには訪問日

時やホスト名も記録しているため、これらを検索することも可能である。

当初、サムネイルの作成には、url2bmp や Webnail2 を用いた。[14, 15] Web ページを読み込むたびに、Extension から呼び出し、ソフトに URL とファイル名を指定しサムネイルを作成した。しかし、ポップアップや JAVASCRIPT エラー等のダイアログボックス表示を抑制することが出来なかったため、ページによってはサムネイルを作るための動作で止まってしまうことがあった。そこで、サムネイル作成を CrenaHtml2jpg を使うように変更した。[16] このソフトは上記の表示を抑制することが出来るようになっている。これらのソフトは IE の機能を使っている。

4 実験と評価

本システムを使用することにより、一度訪問した Web ページを探索する際に、従来の探索方法に比べ素早く手軽に行うことができる、という本研究の目標が達成されているかについて評価した。

本学の学生数名に対して、本システムの趣旨を説明した上で、実験期間中は今までと変わらない方法でインターネットを利用してもらい、Web ページ探索の際の一つの手段として、本システムを使用してもらう。実験期間終了後、本システムについてのアンケート調査を行い、結果を評価した。

全体的には回答傾向としては、主に以下の 6 点が挙げられる。

- 過去に見た Web ページがどのようなものだったか、一目でわかる
- どんなページを見たか思い出しやすい
- 画面を保存するウィンドウが表示されるのが気になる
- 同じ Web ページをまとめることができると良い
- 画像データがたまると HDD の容量が気になる
- あとでサムネイルで表示したくないページまで表示してしまう

全体を通して、技術的な問題点についていくつか指摘があったが、概ね本研究の目的を達成する結果がでたと考えられる。

データ量に関して、無作為に 100 の Web サイトを閲覧し、本システムの履歴 XML ファイルと生成したサムネイルのデータ量を既存の履歴ファイルである history.dat と比較した。結果を表 1 に示す。

表 1: 履歴データのサイズ

種類	使用サイズ
history.dat	41.3KB
XML	30.8KB
サムネイル	633KB

XML ファイルに関しては、history.dat よりも小さな容量で保存されることがわかった。これは、history.dat に比べて XML ファイルでは少ない項目のみを保存しているためだと考えられる。サムネイルに関しては、history.dat に比べて約 15 倍の容量になった。しかし、一日 100 サイトを閲覧したとすると、1ヶ月で history.dat が約 1.2 メガバイト、サムネイルが約 20 メガバイトと予想される。ここ数年、PC の HDD 容量が増大していることを考えると、許容範囲であると考えられる。それよりも、多量の画像ファイルをどのように保存するか、どのように表示するかが問題だと考える。

履歴情報を保存したくない場合は、Slogger の機能によりブラックリスト登録をすることができる。

今回の実装では、ブラウザの裏でサムネイル作成用にソフトを動かした。そのため、ユーザ認証やエラーダイヤログが複数回表示されてしまい、表示がうるさく感じられるという問題があった。また、サムネイル作成用のソフトが IE の機能を使っているため、Firefox のクロスプラットフォームという特徴を生かすことができなかった。

5 考察

これまでの結果を踏まえて、本研究で構築したシステムが果たした役割は主に次のようなものになると考えられる。

- 既存の履歴機能に比べ、情報を見つけやすい。
- どのような Web ページであったかを瞬時に把握できる。
- 検索エンジンで検索しても現れない Web ページを探し出す手段として有効である。

これらの特徴により、一度訪れた Web ページの検索がより短いステップで簡単に行うことができるという本

研究の目的を達成することはできたと考えられる。しかしながら、いくつかの問題点も挙げられる。今回の実装ではページをリロードしただけで、サムネイルが作成されてしまうため、あとから履歴を閲覧するときに同じサムネイルが並んで表示されてしまうことがある。また、サムネイルだけでは細かい本文は読み取れないので、文字列を探すときは、結局 1 ページずつ表示させて自分で調べなければならない。しかし、履歴本文を全文検索するシステムと組み合わせることで、文字列検索への要望へ対応することが出来ると考えられる。また、履歴の保存条件などをもう少し検討する必要がある。履歴を保存する際に、検索エンジンで使用した検索キーワードを一緒に保存することができると、検索キーワード＋サムネイルで、より強力な検索機能となると考えられる。

本システムでは、作成したサムネイルは画面上に表示する以外に使っていないが、サムネイルを画像解析することで、似たようなページを集めたりすることによりユーザインタフェースを改良することが可能であると考えられる。

クラスプラットフォームによるサムネイル作成であるが、Firefox の Ver.1.5 より canvas 機能が実装され、Web ページのサムネイルを作成することが可能なった。この機能を使うことにより、本システムのクロスプラットフォーム対応が可能になると考えられる。

6 結言

本論文では、再訪問する Web ページを検索する際の手助けとなるための手段として、ブラウザが持つ履歴情報を可視化するシステムの提案を行い、システムを Firefox の Extension として実装し、ユーザによる評価、保存データサイズの評価を行った。実験・評価の結果、提案システムを用いることにより一度訪れた Web ページの検索がより短いステップで簡単に行うことができるという本研究の目的を達成することはできたと考えられる。

今後、canvas によるサムネイル作成機能や、似た色をまとめる機能などが実現すると、よりスムーズに、直感的に Web ページの再検索ができるようになると考えている。

参考文献

[1] 村田剛志. Google API を用いたグラフ構造の生成: 第 18 回人工知能学会全国大会, 3pages in CD-ROM (2004).

[2] Tera Calishain, Rael Dornfest, 田中裕子 (訳). Google Hacks: ORAILLY JAPAN (2003).

[3] 柴山悦也. ブラウザのための視覚化とナビゲーション支援: 人口知能学会誌, Vol.16, No.4, pp.509–P514 (2001).

[4] iRider Web Browser http://www.irider.com/

[5] browseback: http://www.bridge1.com/browseback.html

[6] Mozilla project http://www.mozilla.org/

[7] もじら組 http://www.mozilla.gr.jp/

[8] 松澤太郎, 下田洋志. Firefox の全て: C MAGAZINE, 10 月号特集, pp.36–71 (2005).

[9] XULPlanet http://www.xulplanet.com/

[10] Enhanced History Manager http://home.ripway.com/2005-6/326315/FireFox/

[11] GooglePreview http://ackroyd.de/googlepreview/

[12] Slogger http://www.kenschutte.com/firefoxext/

[13] IE Tab http://ietab.mozdev.org/

[14] url2bmp http://www.pixel-technology.com/freeware/url2bmp/english/index.html

[15] Webnail2 http://www.hirax.net/

[16] CrenaHtml2jpg http://dip.picolix.jp/disp5.html

[問い合わせ先]
新美　礼彦
公立はこだて未来大学 システム情報科学部
情報アーキテクチャ学科
〒 041–8655 北海道函館市亀田中野町 116–2
Phone:0138–34–6222　FAX:0138–34–6301
E-mail:niimi@fun.ac.jp