

情報探索におけるブラウジング作業を支援する WWW ブラウザの提案

武藤哲幸 新井克也
NTT 情報流通プラットフォーム研究所

概要

WWW を中心に爆発的な成長を続けるインターネットにおいて、全世界に散在する情報をいかに効率的に探索するかが問題となっている。一般的に、情報探索には『どこに何があるのか=探索範囲の限定』と『どれが自分に必要なのか=探索範囲のブラウジング』の二つの異なる課題が含まれている。現在、検索エンジンをはじめとするさまざまな情報フィルタリング方式が提案されているが、これらは『探索範囲の限定』を支援するものであり、次の段階である『探索範囲のブラウジング』における問題を解決するものではない。たとえば、探索範囲が数千の URL に及ぶ場合、現在広く使われている WWW ブラウザでは、これらすべてをブラウジングすることは現実的ではない。

NTT 情報流通プラットフォーム研究所で開発中の InfoLead®ブラウザは、同時に複数の Web ページを高速に表示するための機能を持っており、情報探索における『探索範囲のブラウジング』を積極的に支援する WWW ブラウザである。各 Web ページは3次元的に表示位置や角度、サイズ、透明度などを自由に指定して表示することができる。また、Web ページの配置自体に意味を持たせることによって、3つの評価軸で位置づけられた Web ページの集合を総論的にブラウズすることが可能となる。本稿では以上の機能の特徴とする InfoLead ブラウザを提案し、その実現方式について述べる。

A proposal of WWW browser which supports the browsing task of information search

Tetsuyuki Mutou Katsuya Arai
NTT Information Sharing Platform Laboratories

Abstract

In the rapid growth of the Internet in recent years, the main problem in the process of searching information that lie scattered in the huge area of the world, is the efficiency. Generally, the information searching is made of two part, the one is restricting the area of information searching (Where is the information located?), the other is browsing the information (Which is the best for me?). Information filtering such as search engine, supports the restriction task, but it doesn't supports the browsing task. In the case of browsing the list of huge number of URLs, it is not practical to browse all of the list.

InfoLead® browser which is under development by NTT platform labs. has a feature of fast display of a lot of Web pages at the same time to supports the browsing task of the information exploring. Each Web page can be transformed freely in the 3D coordinate space, be set the transparency. and categorized by the geometry. So user can make a comprehensive survey of Web pages which positioned by three degrees of estimation axis. In this paper, WWW browser which supports the browsing task of information exploring is proposed.

1. はじめに

近年、WWW を中心としたインターネットにおける情報量は爆発的な増加を続けており、すでに3億以上の Web ページが存在するといわれている。

この膨大な Web ページを効率的に探索するため、検索エンジンやパーソナライズサービスをはじめとする、さまざまな情報フィルタリングサービスが提供されている。

一般的に、情報探索には[Fig.1-1]のように二つの性質が異なる課題が含まれている。前者は、『どこに何があるのか = 探索範囲の限定』である。検索エンジンや、パーソナライズサービスなど、情報フィルタリングに関する研究は主にこの作業を支援するためのものである。また、情報空間の可視化は古くから研究されているテーマであり、ハイパーテキストや WWW の分野では、HTML 文書や Web サイト間のリンクを可視化したり、検索キーワードの動的変更を視覚化したものなどがある[1] [2]。これらは効率的に探索範囲を限定するための方式とみなすことができる。

そして後者が『どれが自分に必要なものか = 探索範囲のブラウジング』という作業である。ブラウジング作業を支援するための方式として、Web ページの切替えを、シナリオに従って自動的に行う Web 紙芝居[3]がある。

NTT 情報流通プラットフォーム研究所で開発中の InfoLead ブラウザは、このブラウジング作業の効率化をはかり、従来の WWW ブラウザに比べて、より広い探索範囲をブラウズ可能にすることを目標としている。

Yahoo!や NTT Directory などの検索エンジンは、WWW において初期の頃から提供されているサービスであり、現在の主要なポータルサイトのひとつである。検索エンジンをはじめとする情報フィルタリングサービスの出力結果は、URL とサマリのリストという形でユー

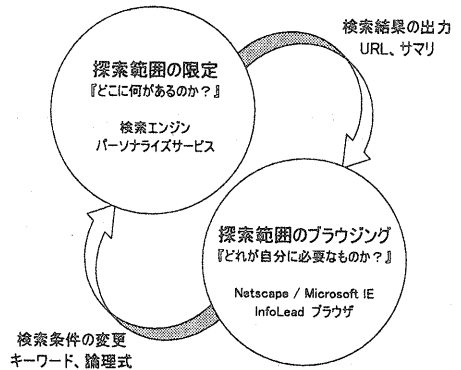


Fig.1-1 情報探索における2つの課題

ザに与えられる。通常、このリストは数千件以上におよぶことがあるが、これらのすべてがユーザの情報探索に利用されているわけではなく、実質的には先頭の十数件程度しか参照されないことが多い。このため、情報フィルタリングサービスに関する研究は、いかんして最適な十数件をユーザに提供できるか、ということに重点がおかれていることが多い。探索範囲の先頭十数件しか利用されない理由にはさまざまなものが考えられるが、主要要因として以下のものが挙げられる。

(1) 情報探索における試行作業の存在

実際に情報探索を行う場合、初期の段階では検索キーワードや論理式のアタリをつけるための試行錯誤が行われることが多い。この段階では検索結果に含まれる URL 数や、サマリの内容が主要な判断基準となるため、実際に参照される Web ページの数は少ない。

(2) ブラウジング操作量の問題

現在一般的に広く使われている WWW ブラウザでは、網羅的に数百件以上にのぼる URL リストのすべてをブラウズするためには、膨大な操作量が必要となる。また、現実のインターネットでは回線速度やサーバ応答時間によるオーバーヘッドが無視できないほど大きいため、一定時間内に参照可能なページ数に限りがある。

InfoLead®ブラウザは Web ページの表示に関して高い自由度を設け、一度にブラウズ可能な Web ページの数を増やすことによって上記(2)におけるブラウジング操作量の改善を行う。

2. ブラウジング作業の効率化

同時に数多くのオブジェクトを人間が直接取り扱う場合、たとえばコンピュータ上のファイル管理などには、アイコンやサムネイルによる視覚的な抽象表現がユーザインターフェースのひとつとして効果的に利用されている。InfoLead ブラウザでは、情報探索におけるブラウジング作業の効率化のため、Web ページの縮小画像による視覚的効果を利用している。

一般的に、電子ドキュメントを縮小表示することによって、同時に表示可能なページの数が増加するが、細かい文字などの可読性は低下するため、両者はトレードオフの関係にある。しかし、Yahoo!や NTT Directory などのディレクトリ型検索エンジンの出力結果の場合、大きなタイトル文字列や画像が使われているインデックスページがヒットする確率が高いため、サムネイルからでも人間がおよその内容を識別することが可能である。

現在広く利用されている WWW ブラウザは、Web ページをサムネイル化して一度に多数のページをブラウズするための機能は用意されていない。このため、たとえば検索エンジンの出力結果を網羅的にブラウズするためには、1ページずつ手動で辿る必要があり、探索範囲が広いほど、多くの操作量を必要とする。

InfoLead ブラウザではこの問題を解決するため、Web ページの高度なサムネイル表示機能を用意している。InfoLead ブラウザにおけるサムネイルは単純に固定倍率で縮小されたページイメージではなく、3次元空間内に配置された Web ページオブジェクトである。このため、任意の3次元的な変形、配置、透明度

の指定が可能である。Web ページオブジェクトは、一般のブラウザで HTML 文書を表示したものを原寸として生成される。このオブジェクトに縮小変換を適用、または視点位置から遠方に配置することによってサムネイルと同等の効果を出すことができる。従来の WWW ブラウザと同じ可読性を実現するためには、単に原寸で表示すればよい。

同時に表示する Web ページオブジェクトの数が増えると、互いに重なり合う部分が生じることがある。InfoLead ブラウザでは、各 Web ページオブジェクトに任意の透明度を設定できるため、重なり合う部分の可読性を高めることができる。また、カラーリングを施したり、任意の形状のウィンドウフレームを指定することによって選択状態を表現し、対話処理を実現することができる。

これらの機能を用いることによって、従来のページ単位のブラウジングに加えて、複数ページを同時に、総覧的にブラウジングすることが可能となる。

3. 総覧型情報探索

現在の検索エンジンと、WWW ブラウザによる情報探索は、1次元上に配列された探索範囲の、ごく一部のみに着目してブラウジングを行っていると考えられる [Fig. 3-1]。本稿では、この従来方式による情報探索を逐次型情報探索と呼ぶ。

これに対して、我々の提案する InfoLead ブラウザによる総覧型情報探索では、1~3次元上に配列された探索範囲の、ある『連続した部分空間』に着目して総覧的にブラウジングを行う [Fig. 3-2]。

逐次型情報探索では、一度にブラウジングできる領域が狭いため、数百件以上の広い探索範囲を網羅的にブラウズすることは困難である。近い将来、光ファイバの普及によってネットワークの帯域幅が十分に広くなり、ページを呼び出す時間が大幅に短縮されたら

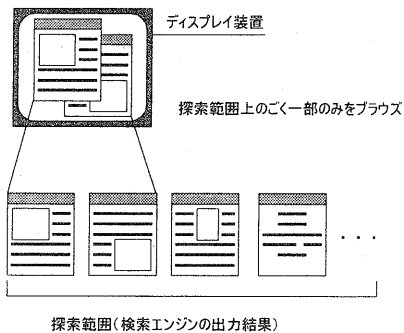


Fig. 3-1 逐次型情報探索 (従来方式)

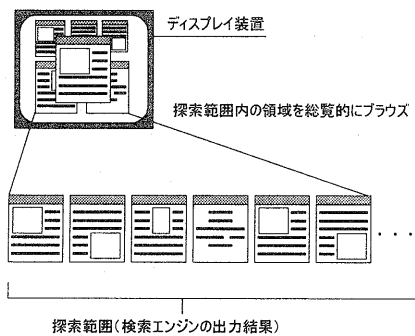


Fig. 3-2 総覧型情報探索 (1)

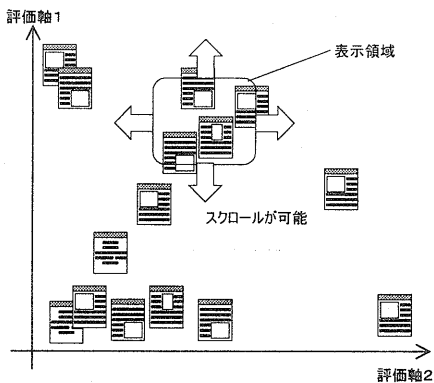


Fig. 3-3 総覧型情報探索 (2)

しても、広い探索範囲における操作量の問題は依然として解決しない。また、探索範囲が何らかの評価軸によって順序付けされていたとしても、ユーザは離散的にしかその傾向を理解することができない。

InfoLeadブラウザを利用した総覧型情報探索では、探索範囲内の連続した領域を自由自在にスクロール

させることができるため、探索範囲の網羅的なブラウジングが可能である。さらに、評価軸に沿って表示領域をスクロールさせることによって、順位付けの意味と傾向を直感的に理解することができる[Fig. 3-3]。

4. InfoLead ブラウザの実装

以上のような特徴を持つ新しいWWWブラウザは、[Fig. 4-1]のように3つの機能ブロックから構成することができる。

・GUI 処理モジュール

ユーザの対話処理と、Web ページオブジェクトの表示を行う。

・HTML レンダリングモジュール

Web ページオブジェクトの表示に必要なビットイメージをHTML 文書から生成する。

・Web ページ配置制御モジュール

情報フィルタリングサービスによって限定された探索範囲に含まれる Web ページに対して、ジオメトリを定義する。

今回、図中の GUI 処理モジュール、および HTML レンダリングモジュールについてプロトタイプの実装を行った。Web ページ配置制御モジュールについては、現在開発中である。

InfoLead ブラウザを実行するクライアントは、高速な3Dグラフィックス機構が必要となるだけでなく、複数のHTML 文書を高速に処理しなければならない。

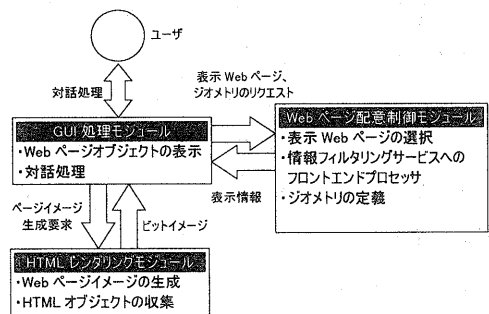


Fig. 4-1 InfoLead ブラウザの機能ブロック

このため、以下の要求を満たすような実装方式が必要となる。

(1) 高速な3Dグラフィックス処理能力

同時に多数の Web ページオブジェクトを表示し、円滑な対話処理に必要なフレームレートを實現するためには、大容量のテキストメモリを持った高速な3Dグラフィックス機構が必要である。

(2) 高度な HTMLドキュメントの再現能力

InfoLead ブラウザは、Web ページの表示イメージを利用してブラウジング作業を効率化する。よって、ページのレイアウトを忠実に再現し、製作者の意図に近い Web ページのイメージを再現する必要がある。

(3) 高速な HTMLレンダリング処理

同時に複数の Web ページを短時間で総覧表示するためには、素材となる Web ページの表示イメージを高速に生成する必要がある。

これらの要求条件を満たすため、現バージョンの InfoLead ブラウザは、機能分散と負荷分散を組み合わせた分散システムとして実装されている。HTML レンダラは、並列処理による性能向上を図るため、GUI クライアントとは別の計算機上に実装した。実際のシステム構成を [Fig. 4-2] に示す。

GUI クライアントは、大容量のテキストメモリを持つ米 SGI 社のグラフィックスワークステーション Onyx2 Infinite Reality 上で実装を行った。グラフィックス API では、OpenGL と IRIS Performer を利用している。HTML レンダラで生成されたビットイメージは GUI クライアントに転送され、Web ページオブジェクト上のテキストチャとして処理される。

HTML レンダラには Microsoft Internet Explorer 4.0 をレンダリングエンジンとして利用した。これを WindowsNT 4.0 Server 上の Internet Information Server 経由で制御するためのインターフェースを実

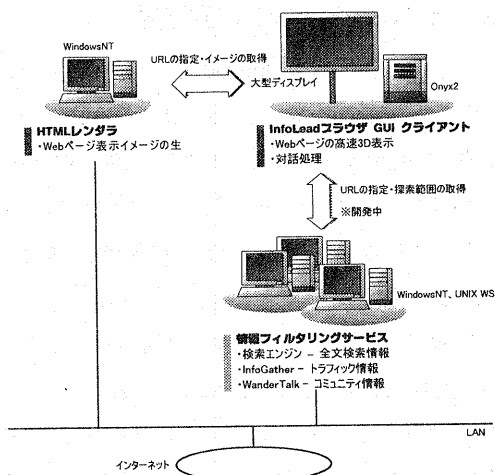


Fig. 4-2 InfoLead ブラウザの構成

HTML レンダラ

HTML 文書の平均容量	約 50 Kbytes
Web ページの表示イメージ容量 (512 x 512 ピクセル・非圧縮)	512 Kbytes
レンダリング速度 (ネットワークの待ち時間含まず)	約 4pages/sec
ハードウェア構成	Intel PentiumII 450MHz x 1 / 128Mbytes RAM

GUI クライアント

テキストメモリサイズ	64 Mbytes
同時表示可能な Web ページオブジェクト数	120 pages
同時表示可能なパーティクルオブジェクト数	1000 pieces
総表示ポリゴン数	約 8000 polygons
実効フレームレート	10 ~ 18 fps
ハードウェア構成	MIPS R10000 195MHz x 2 / 256Mbytes RAM Infinite Reality Graphics (1RM)

Fig. 4-3 InfoLead ブラウザの諸元

装し、http によって GUI クライアントから呼び出しを行う。IE を用いることで Java や、さまざまなプラグインの実行イメージを得ることができる。

Fig. 4-3に、今回試作した InfoLead ブラウザの諸元を示す。現在はまだプロトタイプを開発中であるが、120 枚までの Web ページと、1000 個のパーティクルを同時に表示することが可能である。

5. InfoLead ブラウザの実行例

[Fig.6-1]~[Fig.6-4]に InfoLead ブラウザの実行例を示す。Yahoo! JAPAN にて『雑誌』というキーワードで検索を実行し、検索結果の先頭から 120 件を同時に表示している。各 Web ページオブジェクトのジオメトリについては該当モジュールがまだ開発中のため、単純に画面手前から機械的に配列している。

[Fig.6-4]では、各ページを訪れているユーザの状態や、遷移状況をパーティクルアニメーションによって表現した例である。ユーザのクライアントサイドにおける挙動は、NTT 情報流通プラットフォーム研究所が開発した InfoGather@[4]や WanderTalk を利用して収集することができる。

6. まとめ

本稿では、情報探索におけるブラウジング作業を支援するための新しい WWW ブラウザを提案し、実現方式について述べた。また、このブラウザを利用することによって総覧型情報探索が可能であることを示した。

今後の課題として、

- ・総覧型情報探索による認知モデルの評価
 - ・並列 HTML レンダラの実装
 - ・Web ページ配置制御モジュールの実装
- を行っていく予定である。

参考文献

- [1] 館村: DocSpace: 文献空間のインタラクティブ視覚化、インタラクティブシステムとソフトウェア IV、近代科学社、1996
- [2] 早川、大久保、田中: “WWW 検索サービスにおける検索結果絞込み用インタフェースの開発”、情処研報 98-HI-76, pp.25-30
- [3] 林、新井、西田: “自動ブラウジングシステムの検索サービスへの適用”、第 55 回情処全国大会、1997
- [4] <http://www.infogather.com/>

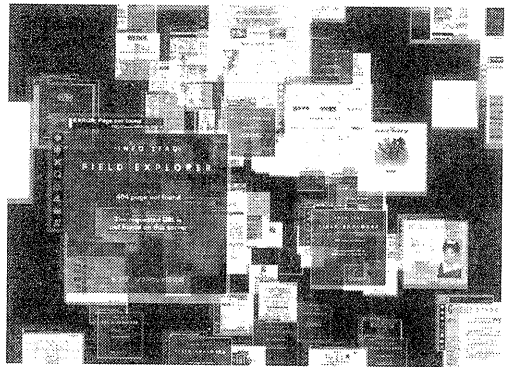


Fig. 6-1



Fig. 6-2

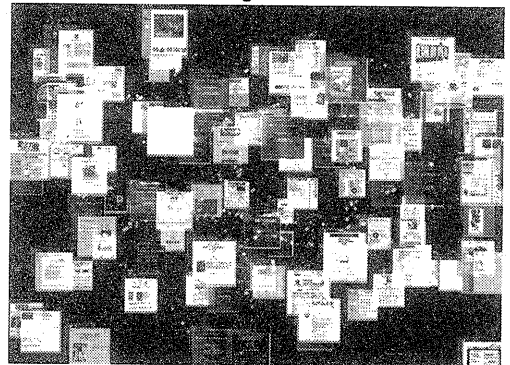


Fig. 6-3

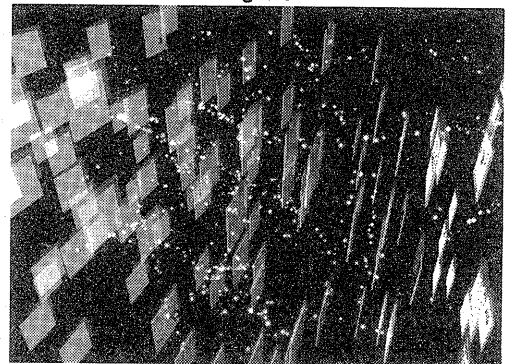


Fig. 6-4