

アノテーションを用いた Web 情報リファインディング支援

飯野亜耶* 奥野拓
(はこだて未来大)[†]

1 はじめに

近年, Web 上に存在する情報は爆発的に増えている. そのため情報収集は書籍など紙媒体での収集と共に, Web 上での収集も重要になった. また個人の情報発見能力, 管理能力も問われる時代になっている.

人が Web 情報を閲覧する際, 閲覧した情報の保存, 既読情報の再閲覧といった作業を行っている. 従来の Web における既読情報の保存は, ブックマーク登録が主な手段である. しかし, これには URL やページ名といったごく一部の情報しか保存されない. そのため時間の経過や登録ページの増大などにより, 再閲覧したいページの検索 (リファインディング) が難しくなる.

2 目的

本研究は, Web ブラウザ上で閲覧した情報の保存・再閲覧を効果的に支援する手法の構築を目的とする. ブックマークなど, 限られた情報をテキストで保存・表示するだけの機能ではリファインディング支援に限界があると考えられる. そこで, 人が紙媒体の情報を再閲覧するプロセスを参考に, 新たな手法の構築を行う.

紙媒体では, Fig. 1 のように情報を閲覧する際には本の特定, ページの特定, 記述の特定と, 徐々に情報を絞り込みながら検索を行っている. その際には付箋やマーカー等, ユーザが付加させたメタ情報を再び閲覧するための補助として利用している. また, これらのメタ情報に加え, 紙面のレイアウトや読んだ時期など, さまざまな記憶も参照して目的の情報に辿り着く.

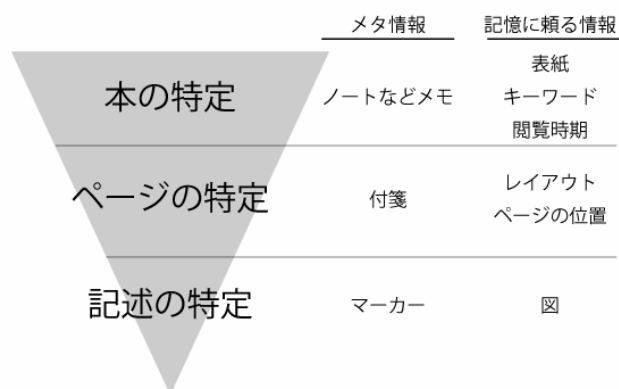


Fig.1 紙媒体におけるリファインディングのプロセス

Web ページは検索からダイレクトに記述にたどり着くことが大半である. そのため, 紙媒体のような本 (Web にたとえると 1 つのサイト) の特定というプロセスは Web には不向きであると考えられる. しかし, メタ情報の付加 (アノテーション) やレイアウトの提示によるページ・記述の特定は Web にも有効であると考えられる. そこで, 本研究ではアノテーションとレイアウトの提示に着目し, これを用いた Web 情報リファインディング支援手法を構築する.

3 先行研究

リファインディングに関する研究が Capra らや西本によって行われている.

初回の情報検索とリファインディングは, 全く異なる性質のプロセスであることを, Capra らが示している[1]. 初回検索時, ユーザは目的の情報が存在するのか, また閲覧している情報が本当にユーザの欲しい情報だったのか確信を持っていない. これに対し, リファインディング時は, 欲しい情報が明確であり, 目的の情報が確かに存在したと確信している.

Web ページの閲覧履歴をリファインディング支援に利用した手法がある[2]. この手法では, リファインディングの際に「以前確かにここを通過した」と自覚できる特徴的なページを Waypoint とし, ブックマークやアクセス履歴などからそれを推測し提示する. Waypoint の提示によって, ユーザに閲覧記憶を想起させるリファインディング支援手法を構築した.

また, アノテーションに関する研究も, いくつか行われている.

松岡らの研究[3]では, ユーザが論文概要に引いたアンダーラインが重要語を多く含むことが実証されている. また, ユーザによってアンダーラインの引き方に個性があるため, パーソナライゼーションへ応用できると述べている. 岩沢らの研究[4]では, 電子媒体ドキュメントへアノテーションを付加し, これをデジタルノートや情報共有手段として導入している. これにより, アノテーションによる情報の切り出し・加工が学習支援に有効であることが示された. また, アノテーションがリファイ

*b1006003@fun.ac.jp

[†]北海道函館市亀田中野町 116 番地 2 公立はこだて未来大学

ンディングに有効であることが Arikawa らによって示されている[5].

これらの先行研究より、本研究では、アンダーラインに重要語が含まれているならば、ラインを引いた文字列を強調して見せることでユーザの記憶を想起させリファインディングに活用できるのではないかと考察する。また、Waypoint とその前後に閲覧したページの関連性を、アノテーションや閲覧時間から導き出しユーザに提示することで、より効果的な支援ができるのではないかと考える。

4 リファインディング支援の手順

想定している支援の手順は、「初回閲覧・記録フェーズ」と「リファインディングフェーズ」という2つの大きなフェーズに分けることができる。初回閲覧・記録フェーズはユーザが初回閲覧時にアノテーションを付加し、重要箇所の記録を行うフェーズである。リファインディングフェーズは過去に見た情報をユーザが検索するフェーズである。

4.1 初回閲覧・記録フェーズ

Web ページを閲覧しながら、ユーザが大切だと判断した箇所にアノテーションを付加するフェーズである。アノテーションには、しおりの挟み込み・付箋の貼り付け・メモの書き込み・テキストへのライン引き（マーカー）といった手段が考えられるが、本研究ではマーカーを採用した。その理由としては、メモの書き込みではテキストを打つコストがかかり、しおりや付箋はコストが低すぎるためにユーザの記憶に残りにくい。その一方で、マーカーはドラッグで手軽にラインを引きながらも、意識的にテキストを読むために記憶に残りやすいといった利点がある。

ユーザは、普段 Web ページを閲覧しているときと同じように情報の検索・閲覧を行うが、重要な記述を見つけた際には、マーカーを閲覧ページ内の文字列に引く。また、システムは常にユーザの閲覧しているページとその閲覧時間、同時に開いていたページ、マーカーを引いた位置を記録していく。

4.2 リファインディングフェーズ

過去に閲覧したページを検索し、該当するページの候補を縮小表示（サムネイル）から探し出すフェーズである。その手順を以下に示す。

ユーザは探し出したいページのキーワードを入力し、



Fig.2 検索結果のイメージ

検索を行う。システムは検索キーワードを含むページを閲覧履歴から探し出し、そのサムネイルをユーザに提示する (Fig.2)。この際、マーカーが引かれた文字列を吹き出しに表示する。検索結果のページ数に応じてサムネイルを複数表示し、比較しながら探すことができるようにする。サムネイル、閲覧時期の新しいものほど左上に表示し、古いものほど右下に表示する。ユーザはサムネイルからそのページに直接アクセスすることができる。

キーワードが分からない場合、閲覧時期を入力し検索を行う。指定された閲覧時期に閲覧されたページの中から、マーカーが引かれているページや閲覧時間の長いページを検索結果として表示する。サムネイル表示順は、キーワードで検索を行った際と同様とする。

検索結果に該当するページが無い場合、次のような手順でサムネイルを2次元平面に配置した俯瞰マップ (Fig.3) を作成し、その俯瞰マップから該当ページを探し出す。ユーザは Fig.2 の閲覧時期やマーカーの情報を参照し、探し出したいページと近い時期に閲覧した1つのサムネイルを「中心ページ」として検索結果から選択する。これは、ユーザが探し出したいページを想起する際、該当ページの前後や同時に閲覧していたページを記憶している傾向があるためである。システムは中心ページを中心とした俯瞰マップを描画する。図中のページ1

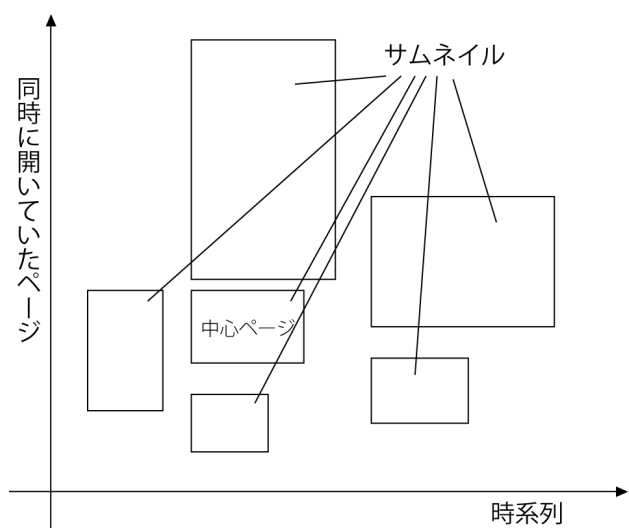


Fig. 3 俯瞰マップのイメージ

が中心ページであり、Waypoint[2]のような役割を持っている。このとき、すでに Fig.2 に表示されていたページは、俯瞰マップには表示しない。サムネイルは時系列を横軸、同時に表示していたページを縦軸とした2次元平面に配置し、キーワードと閲覧時間によって決定された重み付けによって大きさを変化させる。

キーワードによる重み付けは、中心ページに含まれるキーワードとの関連性の強さによって決定する。中心ページ内の文字列を形態素解析し、出現頻度の高い単語とマーカーが引かれた文字列内の単語を中心ページのキーワードとする。俯瞰マップに描画されたページのうち、キーワードが多く出現しているページほど重みを大きくする。閲覧時間による重み付けは、閲覧時間の長いページほど重みを大きくする。2つのプロセスによって設定された重みを足し合わせ、重みの大きいページほど大きく表示する。こうすることでユーザが着目すべきページの優先順を提示することができる。

5 実装・評価

4で述べた手法を既存のWebブラウザに導入し、有効性を確かめる。ブラウザは拡張機能の導入が容易なFirefoxを検討している。

評価では、マーカーを引いた文字列がリファインディングの際に重要な情報となりえるのか、またマーカーを引いたページと他のページの関連性がユーザの記憶想起に役立つのかを示すことが重要である。今後は構築した手法をツールとして実装し、アンケートによるユーザビリティ評価や、リファインディングにかかった時間な

どを従来ツールと比較評価し、構築手法の有効性を確かめる予定である。

6 まとめ

本研究では紙媒体で行われているメタ情報を付加したリファインディングに着目し、ユーザがWebページにマーカーを引き、それをリファインディングに活用する手法を構築した。今後は、マーカー以外のアノテーションの有効性についても検証していきたい。

参考文献

- [1] Robert Capra et al: Refinding Is Not Finding Again, Technical Report TR-05-10, Computer Science, Virginia Tech, pp. 1-10 (2005).
- [2] 西本一平: 履歴情報を利用した Web ブラウジング支援, 公立ほこだて未来大学大学院 修士論文 (2007).
- [3] 松岡有希ら: 論文概要に対する色つきアンダーライン付与システムの運用・分析, DEWS2006, (2006).
- [4] 岩沢和男ら: デジタルコンテンツへのアノテーション機能, 情報処理学会研究報告, CE-87, pp. 67-74 (2006).
- [5] Takeshi Arikawa et al: A Study on Web Page Search Support System Using Web Annotation, Proceedings of The 8th Asia Pacific Industrial Engineering & Management System 2007, T2-R06, pp. 127-131 (2007).