

| | |
|---------|------------------------|
| 氏名 | なか じま しん すけ 中 島 伸 介 |
| 学位の種類 | 博 士 (情報学) |
| 学位記番号 | 情 博 第 98 号 |
| 学位授与の日付 | 平成 16 年 3 月 23 日 |
| 学位授与の要件 | 学位規則第 4 条第 1 項該当 |
| 研究科・専攻 | 情報学研究科社会情報学専攻 |
| 学位論文題目 | 相対性に基づく情報探索に関する研究 |

論文調査委員 (主査) 教授 田中克己 教授 上林彌彦 教授 石田 亨

論 文 内 容 の 要 旨

近年, WWW (World Wide Web, 以下 Web と略す) の普及により, 容易に情報を得るための環境が整いつつあり, また, 取得できる情報量も大幅に増加している。一方, 低品質な情報も Web 上に大量に散乱している状況であり, 膨大な Web 上の情報の中から適切な情報を効率的に検索することが困難になっていることも事実である。したがって, Web 情報検索を効率的かつ効果的に行うことが可能な情報検索手法の開発を行うことに対する社会的要請は極めて高い。

情報検索手法の研究開発は, 主にこれまで文献情報検索に関して長い歴史があり, キーワード語によるブーリアン検索, ベクトル空間モデルに基づく類似検索, システムから返される検索結果に対する利用者のフィードバックに基づき質問を段階的に修正する適合フィードバックなどの技術が開発され実用に供されている。また, Web 情報検索の分野においては, これらの伝統的な情報検索手法が適用されるとともに, Web 情報の特性を考慮して, ブックマーキングや, ハイパーリンク構造に基づく検索精度向上のための技術が開発されている。しかし, 依然として, 情報検索を行う利用者自身にとって, 検索のためのキーワード選定などの検索質問の形成が困難であることに関連する多くの課題が残されている。

本論文は, 従来の情報検索手法においては, 利用者が入力する検索質問で記述すべき検索条件が絶対的な条件として記述されなければならないが, 元来, 人間の情報選択のための条件記述が相対的な形で行っているという点に着目し, 検索条件指定に相対性という概念を導入することにより新たな検索手法を提案し, その有効性について検証を行ったものである。具体的には, 相対性の概念をブックマーキングに導入したコンテキスト依存型 Web ブックマーク, 適合フィードバックに相対性を導入した差異増幅型適合フィードバック, および, 利用者の嗜好や検索意図をサンプルデータ集合上でのデータ選択という形式で表す相対的マッピング処理という 3 つの手法を提案し, その有効性の検証を行っている。

本論文は以下の 6 章から構成されている。

第 2 章では, Web ブックマークや適合フィードバック画像検索における従来の関連研究を整理し, 本研究の位置づけを明確にしている。また, 本研究にて利用した関連技術について詳細に記述している。

第 3 章では, 本研究で新たに導入したコンテキスト依存型 Web ブックマークについて述べている。従来の Web ブックマークが最終的にブックマークしたページのタイトルと URL のみを保存しているのに対し, コンテキスト Web 依存型ブックマークでは, ブックマークを閲覧履歴の中での相対的評価として扱う。ブックマークの価値やブックマークに対するユーザの意図は, ブックマークに至るまでの閲覧履歴情報 (コンテキスト) に依存すると考え, ブックマークに至るまでのコンテキストを解析することで, ブックマークランキング値および各 Web ページの特徴キーワードを算出する方法を提案している。さらに, コンテキスト依存型 Web ブックマーク自身を, 新たな検索質問として利用する方法について提案し, その有効性の検証を行っている。

第 4 章では, 本研究で新たに提案した差異増幅型適合フィードバックに基づく画像データベース検索方式について述べている。従来の適合フィードバック検索が, 利用者が指定した正事例データ・負事例データの特徴量を直接反映する形で検索

質問の修正を行うのに対し、差異増幅型適合フィードバックは、正事例として選択したデータだけでなく、選択データと非選択データの相対的な差異情報をもとに質問修正を行う。この提案手法が、正解に到達できるまでに必要となるフィードバック回数の低減化に貢献できることを実証するための評価実験結果を提示している。さらに、直前の検索質問を修正する差異増幅型適合フィードバックと、直前の検索質問とは独立に新たな質問を生成する差異増幅型適合フィードバックを混在させる場合の評価実験も行っている。

第5章では、本研究で新たに提案した、相対的マッピング処理に基づく相対的情報検索手法について述べている。本手法は、利用者が選択したデータのサンプルデータ集合内での相対的な位置を差分ベクトルで表現し、この差分ベクトルをターゲットとなるデータ集合にあてはめて検索解を得るものである。本手法は指数関数的な計算時間を要するため、データ集合の重心を用いた近似手法も提案し、近似手法の検索精度に関する評価実験結果も示している。

最後に第6章では、本研究で得られた研究成果をまとめ、さらに、今後の課題や展開について述べている。

論文審査の結果の要旨

本論文は、情報検索分野において、検索精度の向上を図るための新たな情報検索手法について検討したものである。

本論文で提案しているいくつかの情報検索手法は、学位申請者が新たに着目した利用者のデータ選択に関する相対性という概念に基づいている。従来の情報検索手法においては、利用者が入力する検索質問で記述すべき検索条件が絶対的な条件として記述されなければならない。一方、人間は、日常的には、情報選択のための条件記述を相対的な形で行っていることが多い。学位申請者はこの点に着目し、種々の検索条件指定に相対性という概念を導入することにより極めて新規性の高い検索手法を提案している。さらに、提案手法それぞれについて、それらの有効性検証のために評価実験を行い興味ある結果を提示している。

具体的には、本論文の研究によって得られた成果は以下のように要約される：

1. 相対性の概念をブックマーキングに導入したコンテキスト依存型 Web ブックマークという概念を導入すると共に、コンテキスト依存型 Web ブックマークを検索質問として活用する新たな方式を提案した。ブックマーク設定にいたるまでの閲覧履歴情報から、ブックマークされた Web ページの相対的な価値をランク付けする手法とその有効性を示した。さらに、コンテキスト依存型 Web ブックマーク自身を、新たな検索質問として利用することで適合率および再現率を向上させることができることを評価実験で示した。
2. 適合フィードバックに相対性を導入した差異増幅型適合フィードバック方式を提案した。従来の適合フィードバック検索が、利用者が指定した正事例データ・負事例データの特徴量を直接反映する形で検索質問の修正を行うのに対し、差異増幅型適合フィードバックは、正事例として選択したデータだけでなく、選択データと非選択データの相対的な差異情報をもとに質問修正を行うものである。この提案手法が、正解に到達できるまでに必要となるフィードバック回数の低減化に貢献できることを評価実験を通じて示した。さらに、直前の検索質問を修正する差異増幅型適合フィードバックと、直前の検索質問とは独立に新たな質問を生成する差異増幅型適合フィードバックの混在方法に関しても興味深い評価結果を得ている。
3. 利用者の嗜好や検索意図をサンプルデータ集合上でのデータ選択という形式で表す相対的質問という新たな概念を提案し、その有効性を示した。具体的には、利用者が選択したデータのサンプルデータ集合内での相対的な位置を差分ベクトルで表現し、この差分ベクトルをターゲットとなるデータ集合にあてはめて検索解を得るという新規性の高い手法を提案した。本手法は指数関数的な計算時間を要するため、データ集合の重心を用いた近似手法も提案し、この近似手法の検索精度も良好であることも示した。

以上、本論文は、相対性という新たな概念を従来の情報検索手法に導入し、新規性・有効性の高い情報検索手法を提案しており、学術上、實際上寄与するところが少なくない。よって、本論文は博士（情報学）の学位論文として価値あるものと認める。また、平成16年2月3日に論文内容とそれに関連した事項について試問を行った結果、合格と認めた。