

研究種目：若手研究(B)
 研究期間：2007～2009
 課題番号：19700151
 研究課題名（和文） 大規模データの意味理解と解釈を支援する思考活性化インタフェース
 研究課題名（英文） Thinking activation interface for understanding and interpretation of massive data
 研究代表者
 砂山 渡 (SUNAYAMA WATARU)
 広島市立大学・情報科学研究科・准教授
 研究者番号：40314398

研究成果の概要（和文）：

データの意味理解と解釈を支援するための、データを様々な観点により可視化するインタフェースを作成した。観点間の比較を行えるインタフェースを作成するとともに、その積極的な解釈を促すための特徴点を提示するインタフェースを構築した。

研究成果の概要（英文）：

Interface that supports users' understanding and interpretation of data by visualizing methods was constructed. The interface enables us to compare data by some viewpoints, and supplies difference points for activating interpretation.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2007 年度	1,300,000	0	1,300,000
2008 年度	1,300,000	390,000	1,690,000
2009 年度	700,000	210,000	910,000
年度			
年度			
総計	3,300,000	600,000	3,900,000

研究分野：総合領域

科研費の分科・細目：情報学・知能情報学

キーワード：知識発見とデータマイニング

1. 研究開始当初の背景

(1) コンピュータの普及とインターネットの発展を背景に、多量の電子データを容易に獲得して解析できる環境が整ってきた。「データマイニング」という言葉がよく使われ、大量のデータの中から有効なデータを抽出する技術の総称として分野として確立される中、研究も競って行われてきた。

(2) どんなマイニングの技術やツールが開発されたとしても、それらを用いて最終的な意思決定を行うのは人間である。ツールの出

力を有効に活用できるか否かは、出力の制御と出力結果の理解を行う人間の能力に依存する部分が多い。

(3) データや解析結果をコンピュータ上に視覚化して表示し、人間が直感的にデータを理解できる環境の構築を目指す「情報視覚化」の研究も多く行われるようになってきているが、ただデータを視覚化すれば、それでデータの解釈が即座に可能になるわけではない。

(4) 視覚化研究の多くは、生の大量の数値データをそのまま見るよりは、十分に見やすい環境を提供しているが、具体的に視覚化されたデータのどこ、および何に注目して、どのように見れば解釈できるか、という視覚化されたデータの見方までを考慮したものは少ない。

2. 研究の目的

(1) 視覚化されたデータの中で着目すべき点を明示し、出力データの意味理解を促す。

(2) 意味理解の結果をまとめ、データ全体を人間が総合的に解釈することの支援を行う。

3. 研究の方法

(1) データの意味理解の支援として、着目して欲しいデータに対して「気づきのもと」を与える（たとえば具体的な印をつける）ことで、データ解析者にそのデータへの意味の付与を促す。意味理解への第一段階は、無意識の状態から意識を促す「気づき」にあるため、この「気づきのもと」の有無によって意味理解の絶対量は増すと考えられる。

(2) データの総合的解釈の支援として、データの意味理解で集められた意味の集合を元に、データ全体としての総合的解釈、知識の抽象化を支援し、次の行動や意思決定に結びつく結論の獲得を支援する。データの解釈を行うデータ解析者には、データを扱う理由（観点）が必ず存在する。たとえば、データ解析者は、「自身の興味」「データの解析結果に対する予想」「理想的なデータが持つ特性」などの観点をもってデータの確認作業を行い、自身がもつ知識の範囲内で、データの総合的な解釈を行う。

4. 研究成果

(1) 多くのWebページ集合の中から、自分が着目しているページに比べて、より抽象度が高いページや、具体性の高いページを探し出すシステム(図1)の構築と評価実験とを行った。実験により具体抽象の判定を行なうために妥当なしきい値を得ることができ、特に具体的なページを探すときに本システムが役に立つことを確認した。このデータの分布と、各データのキーワード表示によりデータ全体の傾向を把握することができる。



図1 具体抽象可視化インタフェース

(2) また、単一のテキストにおいて、テキストの主題と関わりのある単語、および明確な関わりが記述されていない単語を区別したうえで、単語出現の時系列で表示する川下りシステム(図2)を構築した。このシステムを用いることで、各テキストが、その主題についてどれくらい一貫性があるかを、テキストの本文を読むことなく、直感的に把握できることを確認した。

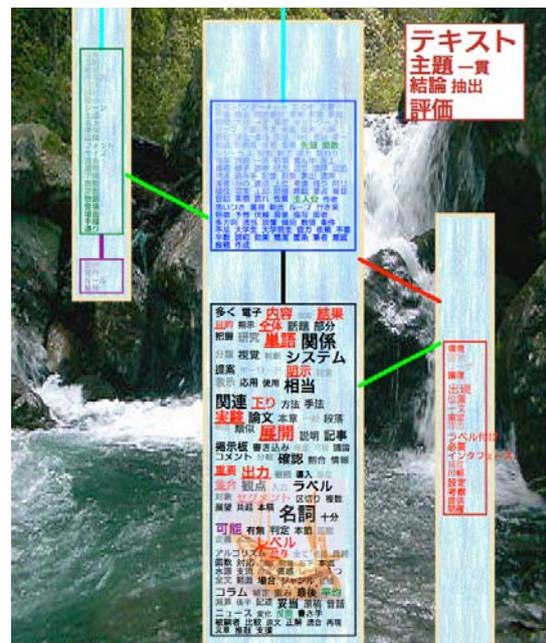


図2 川下りシステムインタフェース

(3) ある分野のキーワード集合を、検索エンジンによるWeb上でのヒット件数をもとにクラスタリングして表示するシステムを基盤として、そのキーワード集合に「観点」となる単語を与え、異なる観点を与えたときに、キーワード間の関係がどのように変わるかを把握することの支援を行うシステム(図3)を構築した。これにより、観点間の関係の違いを比較して、違いの解釈を有効に行えることを確認した。

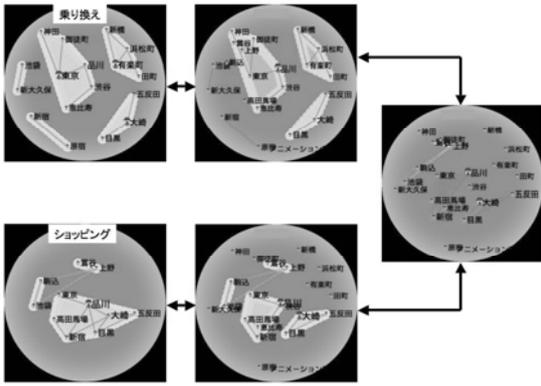


図3 観点間の関係比較インタフェース

(4) 複数の時系列データを連続的に可視化
 する中で、データの比較が行なえるインタ
 フェース(図4)を構築し、その中で着目す
 べき点を明示するシステムを作成した。

データ全体の傾向を捉えるためには、さ
 まざまな観点(基準)に基づく可視化結果を
 比較する必要がある。2つのデータを比較の
 助けとなる着目点を明示して、違いの解釈
 を促すシステムを構築し、実験により、ラン
 ダムに提示する着目点に比べ、解釈の量と
 質の違いについて検証し、提案インタフェ
 ースの有効性を確認した。

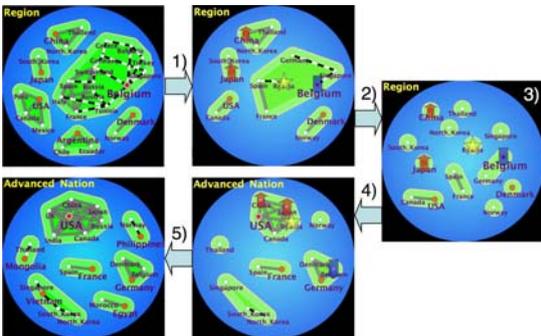


図4 価値変化を表す着目点を可視化
 するアニメーションインタフェース

(5) (4)の2つのデータを比較できるシ
 ステムを改良し、3つ以上のデータの変
 化を可能な限り滑らかにつなぎ、かつ全
 体の中で注目すべき点を示唆するインタ
 フェース(図5)を作成した。これにより、
 時系列データを有効に解釈できることを
 確認した。

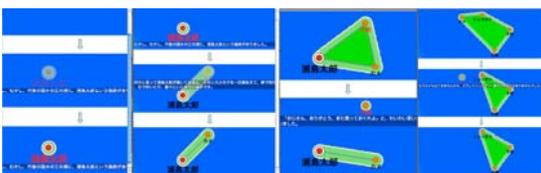


図5 時系列変化を可視化するアニメ
 ーションインタフェース

(6) 電子掲示板において有効なコミュニ
 ケーションが活発に行われているスレ
 ッドを評価するシステムを構築した。こ
 のシステムにより、有効なスレッドを
 探してその内容を確認できるだけで
 はなく、掲示板全体でどのような
 話題のコミュニケーションが活
 発に行われているかを知り、世
 の中全体の傾向を把握する事にも
 役立てられる。

(7) 電子的に提出されたレポート集
 合において、それらの類似性をもと
 に、各レポートの独自性をインタ
 フェース上に可視化するシステ
 ムを構築した。また、各レポート
 中で用いられる単語情報をもと
 に、レポートのテーマに関連する
 キーワードと、独創的なキー
 ワードを評価し、各レポートで
 述べられる事実と意見をそれぞ
 れ評価するシステムを構築した。
 これらのキーワードは、テキ
 スト集合全体を解釈する際の
 ヒントとして用いることができる。

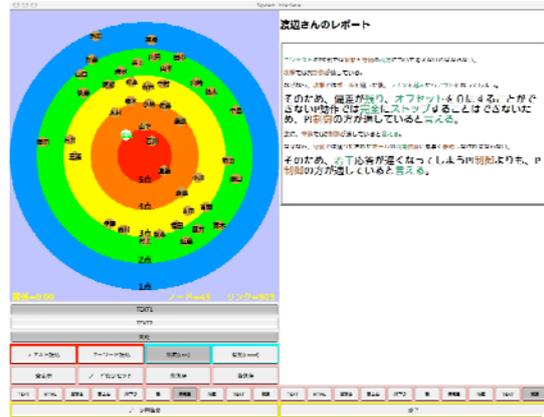


図6 レポートの独自性可視化インタ
 フェース

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者に
 は下線)

〔雑誌論文〕(計7件)

- ① 西原陽子、田中大智、砂山渡：観点の
 違いによるキーワード間の関係の
 変化を捉えるための可視化手法、
 可視化情報学会論文集、査読有、
 Vol. 29, No. 6, pp. 9-16, (2009).
- ② 砂山渡、錦戸拓也、西原陽子：
 地図型アニメーションインタフェ
 ースにおけるキーワードの価値
 変化の検出、日本知能情報ファ
 ジィ学会誌、査読有、Vol. 21, No. 3,
 pp. 304-315, (2009).
- ③ 砂山渡、鮫島聡志、西原陽子：
 Web ページ間の相対的な具体
 抽象関係の視覚化による情報
 収集支援、電子情報通信学会
 論文誌、査読有、Vol. J92-D, No. 3,
 pp. 271-280, (2008).

〔学会発表〕(計 10 件)

① 清水次朗, 砂山渡: 電子掲示板の勢いを評価したスレッド選択支援, 第 17 回 Web インテリジェンスとインタラクション研究会資料, pp. 39-44, (2010).

② 川口俊明, 砂山渡: レポートのテーマ関連度と意見文抽出による情報量評価, 第 132 回情報処理学会ヒューマンコンピュータインタラクション研究会資料, pp. 77-84, (2009).

③ 砂山渡: テキストの話の流れを視覚化するインタフェース, 第 22 回人工知能学会全国大会, 1B1-1, (2008).

④ 川口俊明, 砂山渡: 内容の独自性を視覚化するレポート評価支援システム, 第 21 回人工知能学会全国大会, 2H4-2, (2007).

〔図書〕(計 0 件)

〔産業財産権〕

○出願状況 (計 0 件)

名称:

発明者:

権利者:

種類:

番号:

出願年月日:

国内外の別:

○取得状況 (計 0 件)

名称:

発明者:

権利者:

種類:

番号:

取得年月日:

国内外の別:

〔その他〕

ホームページ等

6. 研究組織

(1) 研究代表者

砂山 渡 (SUNAYAMA WATARU)

広島市立大学・情報科学研究科・准教授

研究者番号: 40314398

(2) 研究分担者

なし

(3) 連携研究者

なし