

キーワード集合の可視化における視線追跡を用いた評価手法

Using eye tracking to evaluate visualization for keywords

東京工業大学大学院情報理工学研究科准教授 **藤井 敦**

PROFILE

1998年東京工業大学大学院博士課程修了。博士(工学)。筑波大学大学院准教授等を経て、2009年より現職。自然言語処理、情報検索、Webマイニング、特許情報処理の研究に従事。

1 はじめに

近年、情報爆発やビッグデータという用語に象徴されるように、個人が生涯をかけても読み尽くすことができないほど大量のテキスト情報が存在する。しかし、テキスト中のキーワードを抽出して可視化すれば、本文を読まなくても概要の効率的な把握に役立つことがある。このような用途でキーワードの集合を可視化した図を本稿ではキーワードマップと総称する。具体的には、大きさ、色、形状、位置、連結などの属性に基づいてキーワード間の関係を表現した図である。

図1～3にキーワードマップの例を示す。図1は、筆者が発表した英文論文の題目に含まれるキーワードの集合をタグクラウドの形式で可視化した図である。

キーワードの大きさが出現頻度を表しており、位置など他の属性に意味はない。図1より筆者の研究発表が information、patent、retrieval に関連していることが分かる。

図2は、筆者がウェブで公開している事典検索システム Cyclone (<http://cyclone.cl.cs.titech.ac.jp/>) が特許情報から自動生成した関連語マップであり、「トラッキング誤差」という用語を中心として関連語を連結し、さらに「光ディスク」のように関連度が大きい用語を近接させている。すなわち、図2ではキーワードの連結と位置の属性が重要である。

図3は、同じく筆者が開発した意見分析システム OpinionReader (オピニオンリーダー) が自動生成した論点キーワードのマップであり、「赤ちゃんポスト」に関する賛成と反対の意見テキストから抽出されたキー



図1 タグクラウドの例

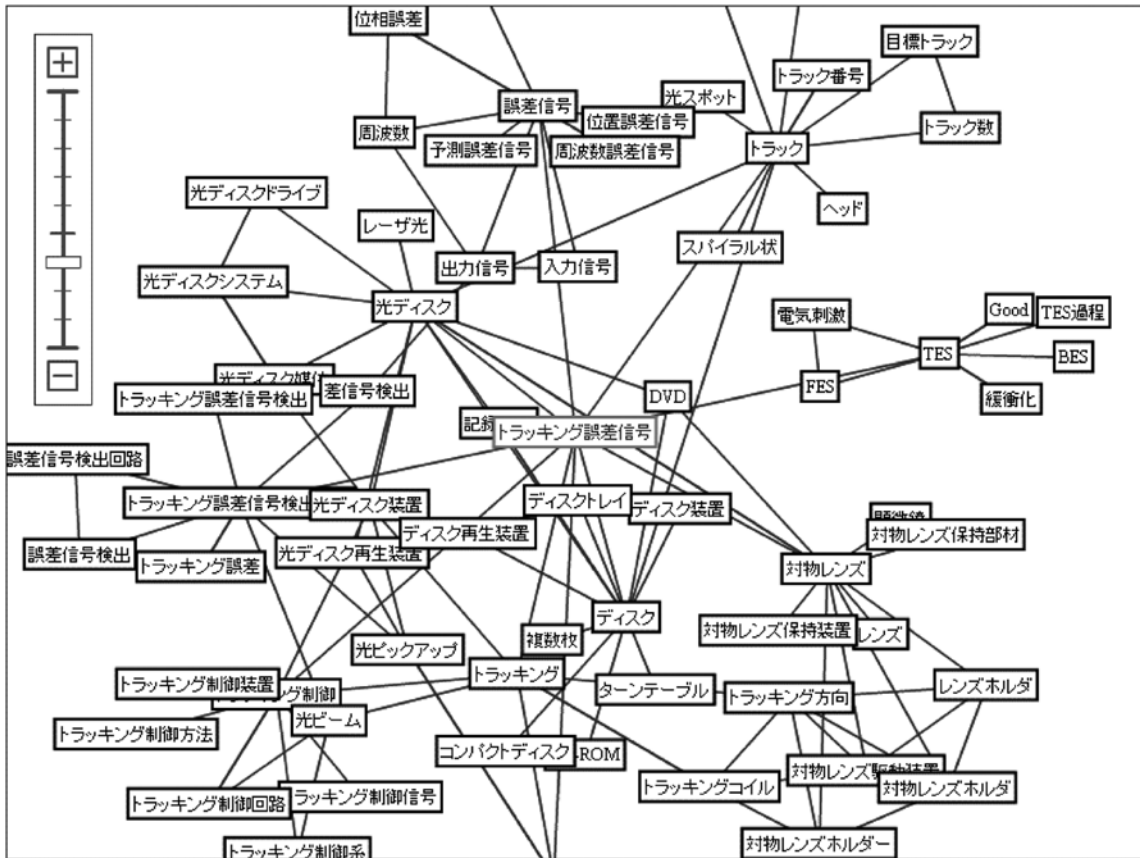


図2 「トラッキング誤差」に関する関連語マップの例

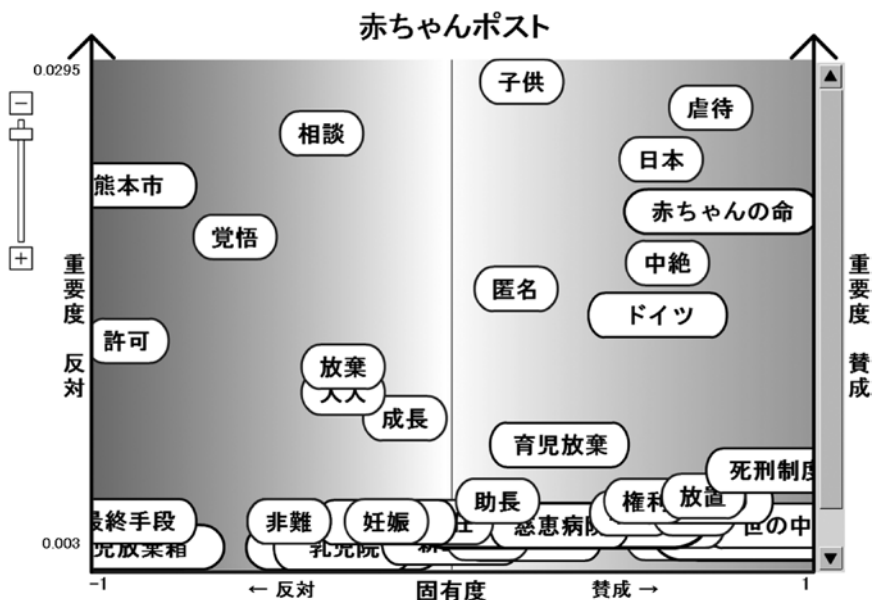


図3 「赤ちゃんポスト」に関する論点キーワードマップの例

ワードが二次元平面に配置されている。横軸と縦軸は、各立場における関連度および出現頻度を表している。すなわち、図3ではキーワードの位置属性が重要である。図3より賛成派と反対派がそれぞれ「虐待」や「相談」といったキーワードを頻繁に使用することが分かる。

図1～3のいずれにおいてもキーワードは抽出元テキストと対応付けられており、キーワードが使用される文脈を確認することができる。すなわち、読むべきテキストを効率的に特定することが可能になる。このように、キーワードマップには一定の有用性がある。

しかし、キーワードマップの品質を定量的に評価する手法が確立されていない。現状では利用者による判定に頼らざるを得ないため、労力や恣意性の点で問題がある。

折しも視線追跡技術が発展したことで、利用者の視線情報に基づく情報処理技術の研究が盛んになりつつある。筆者は、キーワードマップの品質によって利用者の視線に違いが出ると考え、視線情報を解析してキーワードマップの品質を自動的に評価する手法について探求している。現在は、キーワードの位置属性に焦点を当てている。位置以外の属性については今後検討する。

自動評価の方法論が確立されると、複数のキーワードマップから有用な物だけを選別して情報分析の効率を高めることが可能になる。また、キーワードの上手い並べ方に関する理論が発展し、コンピュータを用いて高品質のキーワードマップを自動的に作成することが可能になる。本研究は、高度情報化社会における情報分析の高度化および効率化に貢献することを目指している。

2 予備実験

複数の利用者に様々な品質のキーワードマップを見せよう実験を行い、その視線データを解析すればキーワードマップの品質に関係する特徴量を特定することが可能になると考えた。キーワードマップを作成するために、スポーツや地名のようなカテゴリごとにキーワードを種々の辞書から収集した。品質の異なるキーワードマップを作成するために、キーワードを無作為に配置し

た「低品質」のマップと、同じカテゴリに属するキーワードが近接するように配置した「高品質」のマップをそれぞれ複数作成した。キーワードの配置だけに焦点を当てるため、文字の大きさや色は全てのキーワードで共通とし、キーワードどうしは連結しなかった。

利用者がキーワードマップを見る目的を明確にするために、画面に表示されたキーワードマップを見ながら簡単な質問に口頭で答えてもらった。具体的には、画面上のキーワード集合を同類どうしでまとめた場合に、(1)いくつのグループができますか、(2)各グループに名前を付けて下さい、という二つの質問に答えてもらった。例えば、「サッカー、野球、水泳」で構成されるグループには「スポーツ」や「競技」などの名称が付く。回答が終了したらキー入力によって次のキーワードマップを表示して作業を繰り返した。視線追跡装置を内蔵したディスプレイを使用して、利用者の視線データ（ディスプレイ上のどこをどのような順番で見たのか）を収集した。

収集した視線データを様々な観点から分析した結果として、一つの知見が得られた。具体的には、複数の利用者が同一のキーワードマップを見た時の視線を比較すると、低品質のキーワードマップでは利用者による視線の差異が比較的小さいことが分かった。図4と5は、それぞれ低品質および高品質のキーワードマップにおける二人の視線を並べている。

図4と5のような違いが生じる要因として、低品質のキーワードマップではキーワードの配置に意味がないため誰が見ても端から順番に走査するような視線に近くな

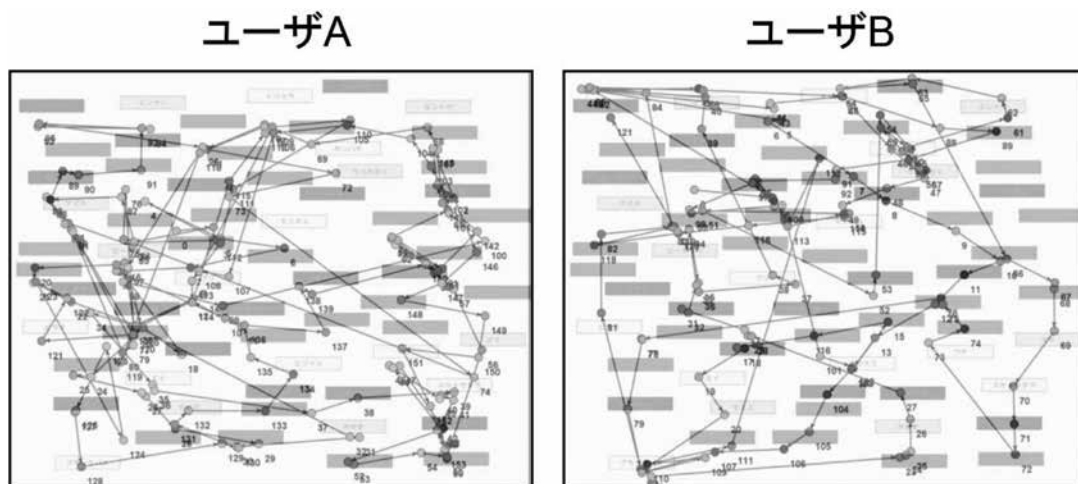


図4 キーワードが無作為に配置されたキーワードマップにおける視線情報

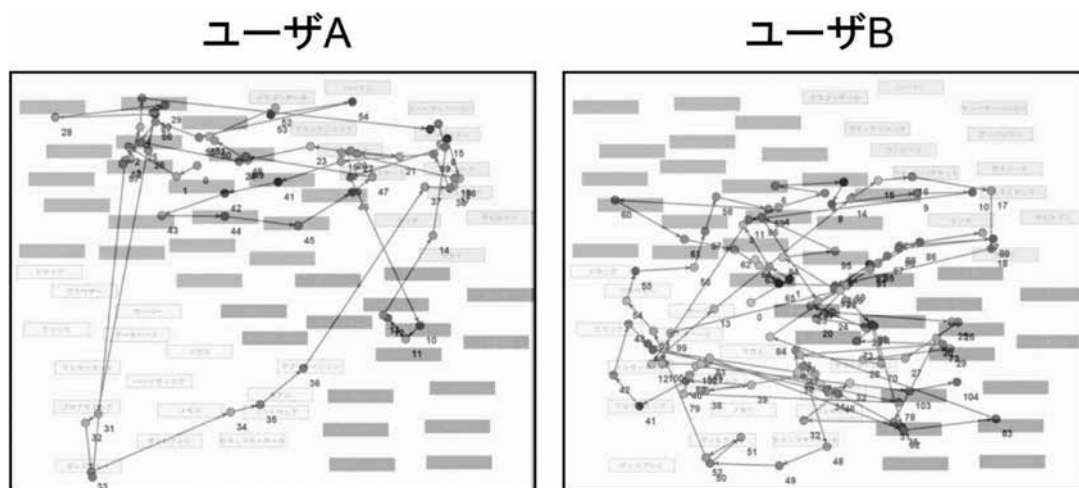


図5 同類のキーワードが近接するキーワードマップにおける視線情報

り、個人差が小さくなる。それに対して、高品質のキーワードマップでは見方に関する個人差が大きくなるのではないかと予想している。

3 キーワードマップの評価手法

キーワードマップの見方に関する個人差を数値化することができれば、その値が大きいほど当該キーワードマップの品質が高いと判定することができる。言い換えれば、視線情報を意味のある特徴量に分解することが重要である。現在、以下に示す三種類の特徴量を用いている。

・注視点間の距離

近接するキーワードを慎重に読み進めるほど値が小さくなる。図5ではAの方がBに比べて値の大きい部分が多い。

・三注視点の角度

視線が同じ場所を何度も往復するほど視線が鋭角を描き、値が小さくなる。

・注視点の密度

キーワードマップ中の各領域における注視の度合いを表す値である。図5ではAとBで注視している領域が顕著に異なる。

図5と比べると、図4ではいずれの特徴量もAとBによる値の差があまりない。

以上の考えに基づくキーワードマップ評価手法の概要を図6に示す。キーワードマップの評価手法自体が存在

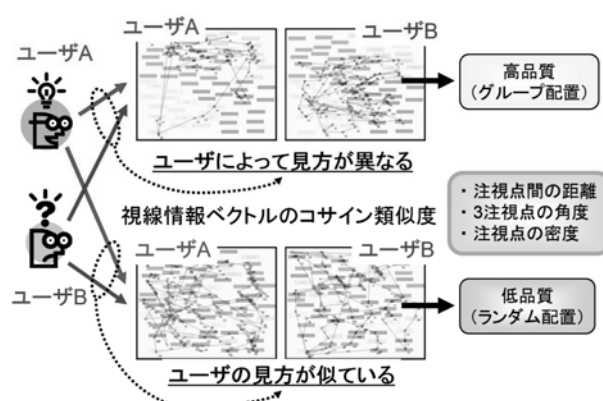


図6 キーワードマップ評価手法の概要

しないため、本研究のような着想に基づく評価手法も存在しない。

本評価手法の有効性を評価するために、キーワードマップ二枚の相対評価を行った。具体的には、高品質と低品質のキーワードマップ各一枚を対にして入力し、プログラムにどちらが高品質のキーワードマップであるかを予測させた。キーワードマップの対を複数用意して、さらに条件を変更しながら実験を繰り返したところ、品質予測の正解率は70~80%台であった。実際のキーワードマップは高品質と低品質の両極だけではないため、より細かな粒度に基づく品質評価が今後の課題である。

4 おわりに

キーワードマップの評価に関する研究には未解決の課題が山積している。地道に研究を進めながら、実用的な技術を実現することを目指している。