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

西原 陽子¹⁾, 田中 大智²⁾, 砂山 渡²⁾

Visualization Method for Transition Recognition between Keyword Relationships Varying with Multiple Viewpoints

Yoko NISHIHARA, Taichi TANAKA and Wataru SUNAYAMA

ABSTRACT

Many researches have been conducted extracting keyword relationships using keyword hit counts of a search engine. Since the relationships vary with multiple viewpoints, it is important to recognize transitions between them. However, previous methods are not useful for variance recognition. This paper proposes a visualization method for transition recognition between keyword relationships. The interface shows animations where the differences move between two relationships in order to help users recognizing the transitions. The proposed method shows the relationships as maps including islands, houses, bridges, rail tracks, and roads in order to support users in recognizing the whole relationships. Experimental results showed that the interface supported users in understanding the whole transitions between maps.

Keywords : Transition of Keyword Relationships, Map Metaphor Animation, Global/Local Relationships

1. はじめに

Web での検索ヒット数を用いてキーワード間の関係を抽出する研究が盛んに行われている^{6,7,9)}。これらの研究では、Web 上で同じページに出現しやすいキーワード間に、強い関係があると推定しているが、それらのキーワードが、どのような文脈において同時に用いられたかまでは、十分に考慮されていない。

すなわちキーワード間の関係には、文脈に依存しない普遍的な関係と、文脈に依存して、ある文脈のもとでは関係が強く、また別の文脈のもとでは関係が弱くなる特殊な関係とが存在する。本稿では、この関係の強さを定める文脈のことを「観点」と呼び、観点の違いによって変化する特殊な関係に着目する。

例えば、企業名をキーワードとしたときには、「訴訟」という観点から得られるキーワード間の関係と、「提携」という観点から得られるキーワード間の関係は異なってくると考えられる。そのため、訴訟のみに現れる関係と提携のみに現れる関係を比較することができれば、提携企業間での訴訟の有無がわかり、株式投資の一指針にすることが可能になる。また、サッカーJリーグのチーム名

というキーワードに対しては、「攻撃」という観点から得られるキーワード間の関係と、「守備」という観点から得られるキーワード間の関係の違いをもとに、攻撃に特化した試合をするチーム同士や、守備に特化した試合をするチーム同士を把握することも可能になる。その他、「テキスト」「データ」「マイニング」など、研究のキーワード集合を入力し、「システム」「可視化」などの観点でキーワードの関係を比較することができれば、システムとしては存在するが、可視化が行われていないキーワードの組み合わせを知り、新たに狙うべき研究テーマの所在を知ることなども可能になる。

そこで本稿では、観点に基づくキーワード間の関係を提示し、異なる観点間の変化を捉えるための可視化手法を提案する。ただし提案手法は、調査や概念把握の初期段階での利用を想定しており、提案システムを利用するユーザの調査や活動において、対象領域の全体像を把握したり、幅広い選択肢の中から可能性を絞り込んでいくことに役立てることを想定している。

本手法では、関係の変化を捉えやすくするために、2つの観点間で関係の異なる部分を動かすアニメーションを表示する。また、アニメーションによる変化を印象づけるために、ノードや線によるネットワーク表現ではなく、キーワードの集合を島、キーワード間の関係を橋、線路、道

* 原稿受付 2009 年 2 月 19 日

1) 会員 東京大学大学院工学系研究科 (〒113-8656 東京都文京区本郷 7-3-1, E-mail: nishihara@sys.t.u-tokyo.ac.jp)

2) 非会員 広島市立大学大学院情報科学研究科

Fig. 1 Outline of proposed interface

路といった地図形式で表現する。

従来手法においても、アニメーションや身近なメタファを用いた可視化手法は提案されているが、地図とアニメーションを組み合わせた可視化手法の提案および効果は検証されていない。そこで、本稿では両者を組み合わせた地図型アニメーションを提案し、関係の変化の理解支援に役立てられることを検証する。

2. 地図型アニメーション可視化手法

Figure 1 を用いて提案手法の構成を説明する。ユーザが、1) 関係を把握したいキーワード集合、2) 関係を定める観点となるキーワード（以下、観点キーワード）、3) キーワード間の関係値、4) キーワードのクラスタリング手法の4つを入力として与えると、キーワード間の関係を表す地図が作成される。地図は観点キーワードごとに作成され、各地図と観点間の変化を表すアニメーションが表示可能になる。以下で、本手法のアルゴリズムと機能について説明する。

2.1 提案手法への入力

入力は、関係を調べたいキーワード集合、関係を定める観点キーワード、観点ごとのキーワード間の関係値、クラスタリング手法の4つとする。キーワード集合は関係を調べたい単語の集合として与える。表示領域の制限と一度に認識できる数を考慮して、キーワード数は50個以下とする。観点キーワードには、キーワード間の関係を定める観点になり得る単語の集合を与える。キーワード間の関係値は、各観点語が定める文脈において、「広く関係が認められる大局的な関係値」と、「与えられたキーワード集合の中では関係が認められる局所的な関係値」の2種類を与える¹。クラスタリング手法は、与えられた関係値からキーワード集合をクラスタに分類し、クラスタを明示して地図を描画するために用いる。

入力となる関係値の算出方法や、クラスタリング手法の

具体的なアルゴリズムは、本稿の主題ではなく、ユーザが独自に定めたものを用いることができる。3章の評価実験で用いた、関係値の算出方法およびクラスタリング手法は、提案手法の一つの形を実現するためのものであり、本稿で用いた手法が最適と主張するものではない。

2.2 地図の描画

地図上には、ノード、リンク、クラスタの3つを、地図形式で描画する。

まず、各キーワードをノードとして、閾値以上の局所的な関係値を持つノード間にリンクを与える。また、ノード集合とリンク集合にクラスタリング手法を適用することでキーワードのクラスタを得る。

次に各ノードの座標を定める。座標は、表示画面の左上を原点、横軸を x 軸、縦軸を y 軸とし、キーワードの50音順に $y=x$ を満たす等間隔の値を初期座標として与えた後、バネモデルのアルゴリズム¹⁰を用いて決定する。バネの力には、キーワード間の大局的な関係値を用い、視認性を高めるため、どのクラスタにも属さない単一のノードは描画しない。

最後に、ノード、リンク、クラスタを、地図形式の表現で描画する。各ノードは「家」として、家の画像を貼りつけて描画する。ノード集合のクラスタは「島」として、以下の凸包アルゴリズム¹¹によって輪郭を作成し、輪郭内部を塗りつぶして描画する。

- 1) 1つのクラスタの中で x 座標値が最小になるノードを選択し、出発点とする
- 2) 出発点から x 軸に平行に伸ばした直線を反時計回りに回転させ、最初に当たるノードを探索する
- 3) 当たったノードが1)の出発点ならば4)へ。それ以外は当たったノードを新しい出発点として、2)へ戻る
- 4) 直線が当たったノードを出発点から順に線で接続する。

リンクは「橋」「線路」「道路」のいずれかとして描画する。橋は、島間をつなぐリンク（格子状の赤線）として描画し、線路は、島内部での大局的な関係値によるリンク（黒と白のまだら線）として、道路は、島内部での局所的な関係値によるリンク（帯状の灰色線）として描画する²。

2.3 アニメーションの描画

アニメーションでは、始点および終点となる2つの地図を指定し、それらの地図間の連続的な変化を描画する。

各ノード（家）の座標値は、始点の地図中の各ノードの間に、終点の地図における大局的な関係値を与え、与えた関係値をバネの強さとしてバネモデルのアルゴリズムを

¹ 提案手法は2種類の関係値を扱う仕様だが、1種類のみの関係値の入力でも、関係の比較を行うことは可能である。

² 1組のノードの間に道路と線路の両方がある時は、正規化された関係値が高い方のみを描画する。

Fig. 2 Output of proposed system.

適用する。バネモデルが収束する座標が、終点の地図における座標となる。

クラスタ（島）とリンク（橋，線路，道路）については，始点から中間点までは，島の透明度を増して透過させていき（Figure 3 中列），中間点（Figure 3 右）で，各ノードが属するクラスタおよびリンク情報を切り替え，中間点から終点の地図に向けては，再び島が出現させていく．新しく出現するノードや消えるノードも，2つの地図が切り替わる中間点で表示変更を行う．

2.4 ユーザに提供する機能

ユーザが操作可能な3つの機能を以下で，実際の出力画面となる Figure 2 を用いて説明する．

2.4.1 表示地図の切替

ユーザは Figure 2 左上の「関係読み込み」ボタンを押すことにより，データベースに保存されている地図を選択し，表示する地図を切り替えることができる．

2.4.2 描画情報の変更

ユーザはマウスを用いてノードを移動することができ，ユーザのノード移動に合わせて島の形が変化する．Figure 2 左下のスライダーにより表示する関係値の閾値を変更でき，左へ動かすと閾値を下げてリンクの数を増やすことができ，右へ動かすとリンクの数を減らすことができる．Figure 2 左の「地図の再描画」ボタンを押すことにより，移動したノードをバネモデルのアルゴリズムを用いて再配置することができる．

2.4.3 地図間のアニメーション表示

ユーザは Figure 2 左，中央部分の「>」，「<」ボタンを押すことにより，選択した2つの地図間で関係が異なる部分を動かすアニメーションを表示できる．表示ボタンを押してから，アニメーションが終了するまでのアニメーションの表示時間はおおよそ3秒になる．「表示」ボタンを押すことにより選択した地図を表示でき，「観点なし保存」，「観点あり保存」ボタンを押すことにより，ノードを移動して作成した任意の地図を保存することができる．また，その

地図間のアニメーションも表示できる．

2.5 提案手法の使用手法

上京したばかりの学生であるユーザが，山手線の沿線で，乗り換え駅の関係とショッピングに便利な駅の関係をもとに，週末のショッピングプランを立てたいと考えた．そこでユーザは，山手線の駅名をキーワード，観点キーワードを「乗り換え」「ショッピング」として入力を与えた．すると，Figure 3 左列の地図が出力された³．

まず，ユーザは出力されたそれぞれの観点による地図を見て，観点ごとの駅間の関係を解釈する．Figure 3 左上の観点「乗り換え」の地図は，5つのエリア（東京駅周辺，浜松町駅周辺，大崎駅周辺，新宿駅周辺，池袋駅周辺）に分類され，左下の観点「ショッピング」の地図は，3つのエリア（品川駅周辺，池袋駅周辺，上野駅周辺）に分類されていることがわかる．

次に，ユーザは2つの地図間のアニメーションを表示させ，アニメーション中の変化に着目して，どのような関係の違いがあるかを確認する．

観点「乗り換え」の地図から，観点「ショッピング」の地図へのアニメーションを見ると，観点「乗り換え」の地図にあった，「浜松町周辺」の島が消えたことから，この周辺は乗り換えには重要なポイントだが，ショッピング向けの店はあまり多くないということが解釈できる．また，「上野駅周辺」の島が現れたことから，この辺りを乗り換え目的で使う人は少ないが，ショッピング向けの店がわりとあると解釈することができ，東京在住の人以外にはわかりにくかった事実を確認することができる．

また，「池袋駅周辺」や「東京・品川駅周辺」は両方の地図に共通に現れているため，乗り換えも便利で，ショッピングもできる場所として広く認知されている場所だと確認することができる．

このように提案手法を用いることにより，ユーザは表示された地図からキーワード間の関係について既知の情報を確認し，ユーザにとって曖昧であったり，未知だった情報を補完することが可能になる．また，書き手の主観が入ったガイドブックの情報ではなく，Web ページの数を基にして得られる客観的な情報に基づいて，自分自身で関係の解釈を行いたい場面で提案手法は有効になる．

3. 評価実験

本実験では，観点の違いによる関係の変化の理解と解釈を助ける，地図型アニメーション表示の効果を確認する．

3.1 使用データ

Table 1 に実験に用いたテーマと観点キーワード，用意し

³ 関係値およびクラスタリング手法は，後述の3章に示す方法を用いた．

Fig.3 Animation between two maps. Keywords are the name of stations in Yamanote line. Viewpoint keywords are transfer and shopping.

Table 1 Data sets for experiment.

テーマ	観点キーワード	キーワード数
食べ物	おいしい, 健康	40
生き物	気持ち悪い, 怖い	40
偉人	芸術, 科学	30
スポーツ	球技, 格闘技	30

Table 2 Example of keyword set. Theme is food.

ジュース, 牛乳, 紅茶, コーヒー, 栄養ドリンク, チョコ, プリン, ケーキ, パフエ, お煎餅, クッキー, アイス, お饅頭, ゼリー, ハンバーガー, サンドイッチ, あんぱん, 食パン, メロンパン, クリームパン, フランスパン, にんじん, キャベツ, トマト, かぼちゃ, 桃, なすび, リンゴ, ブドウ, バナナ, みかん, カレー, チャーハン, 炊き込みご飯, おむすび, オムライス, ピラフ, ドリア, 寿司, カツ丼

たキーワード数を示す. 例として Table 2 にテーマ「食べ物」で用意したキーワード集合を示す.

3.2 関係値算出手法

本実験では大局的な関係値として相互情報量を用い, 局所的な関係値としてカイ 2 乗値を用いた. 相互情報量 pmi はキーワードの検索ヒット数から式(1)により算出した.

$$pmi(w_a, w_b) = \log \frac{hit(w_a \wedge w_b)}{hit(w_a) \times hit(w_b)} \quad (1)$$

式(1)で w_a, w_b はキーワード, $hit()$ は検索エンジン (Google を使用) でのキーワードのヒット数を表す. 観点がある時の評価値 pmi_v は, 観点キーワード v を式(1)に加え, $hit(w)$ を $hit(w \wedge v)$ として同様に求めた.

キーワード集合 W でのキーワード間のカイ 2 乗値 χ^2 は式(2)により算出された⁹⁾.

$$\chi^2(w_a, w_b) = \frac{N' \times (a \times d' - b' \times c')^2}{(a + b') \times (a + c') \times (b' + d') \times (c' + d')} \quad (2)$$

$$N' = \sum_{w_a, w_b \in W} hit(w_a \wedge w_b) u$$

$$a = hit(w_a \wedge w_b)$$

$$b' = \sum_{w_i \in W; w_i \neq w_a, w_b} hit(w_a \wedge w_i)$$

$$c' = \sum_{w_i \in W; w_i \neq w_a, w_b} hit(w_b \wedge w_i)$$

$$d' = \sum_{w_i, w_j \in W; w_i, w_j \neq w_a, w_b} \text{hit}(w_i \wedge w_j)$$

相互情報量は入力されたキーワード集合 W に依存しない絶対的な値を表すのに対して、カイ 2 乗値は入力されたキーワード集合 W に依存する相対的な値を表す。このことから相互情報量をより大局的な関係を表す値、カイ 2 乗値をより局所的な関係を表す値として用いた。観点がある時の評価値 χ^2_v は式(2) の $N \sim d'$ までの $\text{hit}(w_a \wedge w_b)$ を $\text{hit}(w_a \wedge w_b \wedge v)$ とし同様に求めた。全ての評価値は最小値を 0、最大値を 1 として正規化した。

3.3 クラスタリング手法

クラスタリング手法には以下に示す Newman のアルゴリズム¹¹⁾を用いた。

- 1) 全てのノードを大きさ 1 のクラスタとして定義
 - 2) 評価値 Q の増減 ΔQ が最大になるノードを結合
 - 3) 2) を繰り返し、 Q が最大になれば終了
- 評価値 Q と ΔQ は以下の式により定義される。

$$Q = \sum_l (u_{ll} - v_l^2)$$

$$\Delta Q = u_{lm} + u_{ml} - 2v_l v_m = 2(u_{lm} - v_l v_m)$$

ただし、 u_{lm} はクラスタ l とクラスタ m の間のリンク数、

$$v_l \equiv \sum_m v_{lm} \text{ はクラスタ } l \text{ 内への総リンク数とする。}$$

また、クラスタリングのためのノード間のリンクの有無は、ノード間のカイ 2 乗値を用いて、値が閾値 T 以上の場合にリンクがあるとした。

3.4 被験者の作業手順

被験者は情報科学を専攻する大学生・大学院生 12 名で全て男性であり、以下の作業を行った。

- 1) 観点がない時のキーワード間の関係と解釈⁵⁾を記述
- 2) 観点がある時のキーワード間の関係と解釈を記述
- 3) 観点がある 2 つの関係の変化と解釈を記述

実験結果を比較するため、キーワード間の関係をグラフで表す手法を用意した (Figure 4. 以下、比較手法)。比較手法は島を表示せず、家をノード、橋、線路、道路をリンクにより表す (橋: 赤色、線路: 黒色、道路: 灰色)。アニメーション機能は設けず、関係の変化を理解し解釈を与える際は 2 つのグラフを並べて表示した。島は比較手法上でリンクにより接続されるキーワード集合を表すものであり、比較手法上でも島に相当する単語グループの認識

Fig. 4 Output of compared interface.

は可能である。本実験では可視化で最も一般的に使われるグラフ表示を採用し、地図表示の効果を検証した。

各被験者に異なるテーマと可視化手法を 2 つずつ与えた。作業手順の 1) から 3) は自由記述による回答を指示し、提案手法を用いたときは「島、橋、線路、道路」を、比較手法を用いたときは「単語グループ、赤線、黒線、灰線」の単語を用いて回答する旨を指示した。

4. 実験結果

4.1 被験者の回答例

テーマ「食べ物」で、観点がないときの二人の被験者の回答例を Table 3 に、その解釈の元になった図を Figure 5 に示す。提案手法では Figure 5 左の地図を見ると大きな 3 つの島が目につき、例えば右上の島に着目して、そこには米料理が集まっていたため、被験者はキーワード間の関係を米料理と回答したと考えられる。比較手法では、別の被験者が Figure 5 右のグラフを眺め、例えば左下にあるキーワードの関係に着目し、被験者はキーワード間の関係をパンと回答したと考えられる。これらを含め、全ての回答は著者らに確認されており、不適切な記述が含まれるものはなかった。また、本手法は既存の関係を解釈することを前提にしているため、不適切でない回答の質は問わないものとしている。

提案手法を用いた被験者は地図メタファを自然に理解していたと考えられる。Figure 5 左に類似した図を見た被験者は、パンの島に着目し、道路により接続された食パンとフランスパンの関係を「食パンとフランスパンは味が合わないパン」と記述し、線路により接続されたフランスパン、メロンパンの関係を「フランスパンとメロンパンは関係があるが、メロンパンには味がある」記述していた。これは、フランスパンとメロンパンは、より大きなパンというカテゴリに属している大局的な関係にあるということ、また、食パンとフランスパンは、これらのパンの中に限っては味が合わない二つのパンとして局所的な関係にあり、それらが自

⁴ 0.8 とした。この値は経験的に定めた。

⁵ キーワード間の関係を表す説明文を記述する旨を教示した。

Fig.5 Examples of proposed interface and compared interface. Theme keyword was food. No viewpoint keyword was used. Two figures are not the same, because each participant used different thresholds for showing links. This figure is corresponded with Table 3.

Table 3 Example of keyword relationships and interruptions for the relationships. Theme keyword was food. No viewpoint keyword was used. This table is corresponded with Figure 5.

提案手法 (被験者1)
米料理(カレー, チャーハン, 寿司が道路で接続されていて, ピラフ, オムライス, チャーハンが線路で接続)
甘いもの(ケーキ, チョコが同じ島にある)
飲み物(紅茶, コーヒーが道路で接続)
果物(リンゴ, ブドウが道路で接続)
野菜(ニンジン, キャベツ, トマト, カボチャが道路で接続)
パン(食パン, クリームパン, メロンパン, あんぱん, フランスパン, サンドイッチが線路で接続)
比較手法 (被験者2)
飲み物(コーヒー, 紅茶が灰線で連結)
パン(フランスパン, 食パン, メロンパン, クリームパンが灰線で接続)
野菜(トマト, キャベツ, かぼちゃ, にんじんが灰線で接続)
飯物(寿司, カレーが黒線で接続)

然に理解されている。同様にメタファの違いを理解した上で、関係を記述していた回答は多数あったことから、地図メタファは被験者にとって容易に理解できるものであったと考えられる。

被験者ごとに、各自が異なるしきい値を設定したため、画面に表示されるリンクとノードの数は、各々で異なっていたが、それらの平均値は提案手法と比較手法で差がなかった。そのため以降の節では、回答数が多いほど手法が有効に働いたとみなして評価を与える。

4.2 アニメーションの効果

関係の変化が記述された回答数の平均 (1つのテーマと、1つの手法に割り当てられた3名の平均) を Table 4 に示す。回答数の平均は提案手法の方が多かった ($t=2.3$, $P<.05$)。テーマが「生き物」のときの回答数の平均は比較手法を用いたときが多かったが、有意差はなかった。

各被験者が記述した関係の変化を表す単語が含まれる回答の数を Table 5 に示す。ただし、関係の変化を表す単語は「消えた, 出てきた, くつついた」と「有り, 無し」などとした。Table 5 から、提案手法では関係の変化を表す単語が比較手法よりも多かった。すなわち、提案手法を用いた被験者は、関係の変化を単純に何かの有無として捉えるのではなく、関係の変化をイメージとして理解しようと努めたことにより、回答数が多くなったと考えられる。

実際の被験者の回答例からも、提案手法を用いた被験者は「テニスの居た島において、他の全ては消えたが、テニスだけが残った」のように、2つの地図で関係が変化した様子を頭の中にイメージとして残し、記憶が強化されていたことが伺える。対して比較手法では「プリン, チョコ, ケーキは『おいしい』の地図にはあるが、『健康』の地図ではない」など、キーワードの有無に関する回答が多く、関係が変化した様子をイメージとして残すことは難しかったと考えられる。

すなわち、提案手法における回答数が多く、回答文からも関係の変化をイメージとしてとらえた旨が伺えることから、アニメーションを表示することにより関係の変化の理解がしやすくなることが分かった。

4.3 地図とアニメーションの組み合わせの効果

アニメーションと地図表示を組み合わせたことによる効果について考察する。

Table 6 に「島 (単語グループ)」と「橋, 線路, 道路 (赤線, 黒線, 灰線)」が含まれた関係の変化に対する回答数を示す。Table 6 の全テーマの平均において提案手法は「島」を含む回答が多く ($t=3.4$, $P<.05$)、比較手法は「赤線, 黒線, 灰線」を含む回答が多かった ($t=4.8$, $P<.05$)。

提案手法を用いた被験者には「昆虫の島, 鳥の島, は虫類の島, 肉食獣の島がある。同じ島にカマキリ, カブトムシ, トンボ, チョウがあり, カマキリを中心に道路で接続」と初めに地図上の「島」を見て、続いて「島」内のキーワ

Table 4 Averages of keyword relationship varying with multiple viewpoints.

テーマ	提案手法	比較手法
食べ物	3.33	2.33
生き物	1.33	1.67
偉人	3.00	2.00
スポーツ	3.67	3.50
平均	2.83	2.37

Table 5 Number of keywords representing relationships varying with multiple viewpoints.

	「消えた」など	「有り, 無し」など
提案	17	6
比較	2	7

ード間の関係を記述した人が多かった。このことから、提案手法を用いた被験者は、まず地図の全体的な関係を理解した後に、部分的な関係を理解しようとしていたことがわかる。

逆に比較手法を用いた被験者は、「コーヒー、紅茶が灰色の線で接続」のように「線」で結ばれた関係を記述した人が多く、単語グループに関する関係を記述した人は少なかった。すなわち、線を主体としたグラフ表現では、まず線の有無に注目が集まって、キーワード間の接続関係を見た後に、その線が属する「単語グループ」の意味を考えていったと推定される。

すなわち、提案手法においては、全体を見てから部分を見るという順番ができ、比較手法においては、部分を見てから全体を見るという順番であったと考えられる。提案手法を用いた被験者による関係の変化と解釈に関する回答例を Table 7 に示す。被験者の回答からはキーワードや島の変化を「くつついた、消えた」と関係の変化を表す単語を用いて記述しており、適切な解釈が行えていることが読み取れる。

提案手法を用いた被験者の回答を見ると、どの回答でもキーワードや島の変化に関する記述はあるが、「橋、線路、道路」の消滅と出現、線路から道路へとメタファの変化、すなわち部分的な関係の変化に関する記述はほとんど見られなかった。これは島に注目してその変化を記述した被験者が、大きな変化を記述したことによって満足し、それ以上細かい変化に着目しようと思わなかったためではないかと考えられる。そのため、比較手法での部分的な理解の回答数の多さにも着目して、全体像の解釈が進んだ後の段階においては、部分的な解釈の支援を地図アニメーションなしで行う方法も検討の余地があると考えている。

異なる観点による関係の変化について、その結論や解釈を得るためには、最終的に各変化点を抽象化、一般化する必要がある。提示された関係の全体もしくは広い部分に関する変化を捉える必要がある。この意味で、実際に全体的な関係の変化を捉えた記述が多かった提案手法が、変化を捉えて解釈を行なうという目的に対して、有効に働くと考えられる。それに対して、比較手法では部分的な関係が主に着目されており、全体像の理解の伴わない部分的な理解のみでは、効果的な解釈を導くことは難しいと考えられる。以上のことから、提案手法は、アニメーションと地図の組み合わせによって、島の変化により全体的な関係の変化を捉え、既存の関係に対する解釈を与えるのに有効であることがわかった。

5. 関連研究

キーワード間の関係の可視化、およびキーワード間の関係の変化の理解を支援する関連研究をまとめる。

Table 6 Averages of islands, bridges, railways, and roads written by participants.

テーマ	提案手法		比較手法	
	島	橋など	単語グループ	赤線など
食べ物	1.67	1.67	1.00	1.33
スポーツ	1.00	0.33	0.33	1.33
偉人	2.33	0.67	0.33	1.67
生き物	2.00	0.67	0.67	1.67
平均	1.75	0.83	0.58	1.50

Table 7 Examples of keyword relationships and interruptions by participants.

関係の変化と解釈に関する回答例 (矢印の前が関係の変化、矢印の後が解釈)
「おいしい」から「健康」に変わると、野菜、果物、米料理、甘いものの島がくつついて1つになった → 不健康なものはあまり種類が分かれていない
「芸術」では島に含まれたマルクスと孔子が「科学」では消えた → マルクスと孔子は思想家としての関連が強い
「怖い」のときは島の中でサンマ、マグロとつながっていたサメが「気持ち悪い」のときは消えた → サメは怖いが気持ち悪くはないと思われる

5.1 キーワード間の関係の可視化

Web ページ集合間の関係を立体のグラフで表現する手法¹⁴⁾や、掲示板への書き込みに含まれるキーワード間の関係を平面のグラフで表示する手法がある¹⁵⁾。これらの手法は、特定の1つのノードに着目して、そのノードに関連するノード集合を把握することには便利であるが、複数のノード集合同士を比較して、それらキーワードの全体像を把握するのには不向きとなる。これらに対し、提案手法ではキーワード集合を島として表示して、集合の関係を理解することを支援する。

キーワード間の関係をグラフで表示し、ユーザが選択したキーワードにより新しくキーワード集合を定義して表示する手法がある⁴⁾。キーワード集合を定義できる部分は提案手法でキーワード集合を島とする部分に似ているが、提案手法では実際に島を描画するため、キーワード集合を見て言語によって理解すること、島を見てイメージによって理解することの両方が行え、記憶に残りやすくなる利点がある²⁾。

検索結果をクラスタリングし、クラスタを島のように表示する手法があり^{3),5)}、これらの手法は選択されたクラスタに属するクラスタを線で接続して表示する。これらに対し提案手法はキーワード集合同士の関係を「橋」、キーワード間の大局的な関係と局所的な関係を「線路、道路」を用いて区別して描画するため、キーワード間の関係を詳細に理解することができる。

可視化で用いられるメタファには身近なところで使われるものが多く、年輪¹²⁾や本¹⁸⁾を使う手法が存在する。

これらの中で、1つの図で複数の関係を表せるメタファは少なく、その中で地図は最も人間にとって身近なメタファだと考えられる。地図メタファを用いることによって、人間がキーワード間の関係を直感により理解することを支援することが可能となる。

5.2 キーワード間の関係の変化の可視化

情報の状態変化をアニメーションで表示する手法があり¹⁰⁾、この手法では変化の前後で共通に存在する部分をアニメーションで表示する。yFiles¹⁷⁾も事象の関係をグラフ表示し、グラフ上のノードの追加、削除やエッジラベルの移動にアニメーションを用いている。これらに対し、提案手法では存在する情報の一部分ではなく、全体の変化を示すアニメーションを用いているため、ユーザは全体の関係の変化をより理解しやすくなる。

他にアニメーションを用いた研究には人間の嗜好の違いの解釈を支援するものがある¹³⁾。この手法ではアニメーションの途中で情報が増減しないと仮定している。対して提案手法では情報の増減に対応しており、ユーザの情報理解を妨げることは少ないと考えられる。

6. むすび

本稿では、観点の違いによる、キーワード間の関係の変化を理解し、その解釈を与えることを支援する地図型アニメーションによる可視化手法を提案した。地図のアニメーション表示を行うと、島の出現や消滅をイメージとして印象づけることができ、地図の全体的な変化を理解しやすくなることを、実験により確認した。

今後は、線路や道路といった地図メタファの描画方法の改善により、大局的な関係や局所的な関係が直感的に理解可能な支援を行うこと、また、ユーザによって情報を理解したいテーマが与えられると、キーワード集合や観点キーワードを自動獲得し、ユーザの負担を軽減する方法を考えていきたい。

参 考 文 献

- 1) C. B. Barber, D. F. Dobkin, and H. Huhdanpaa, The Quickhull Algorithm for Convex Hulls, ACM Transactions on Mathematical Software, Vol. 22, No. 4, pp. 469–483 (1996).
- 2) J. M. Clark and A. Paivio, Dual coding theory and education, Educational Psychology Review, Vol. 3, No. 3, pp. 149–170 (1991).
- 3) <http://www.grokker.com/>
- 4) K. Horie and Y. Ohsawa, Product Designed on Scenario Maps Using Pictorial KeyGraph, WSEAS Transaction on Information Science and Application, Vol. 3, No. 7, pp. 1324–1331 (2006).
- 5) <http://www.kartoo.com/>
- 6) H. Kautz, B. Selman, and M. Shah, The Hidden Web, AI magazine, Vol. 18, No. 2, pp. 27–35 (1997).
- 7) 金英子, 松尾豊, 石塚満, Web 上の情報を用いた企業間関係の抽出, 人工知能誌, Vol. 22, No. 1, pp. 48–57 (2007).
- 8) Y. Matsuo, T. Sakaki, K. Uchiyama, and M. Ishizuka, Graph-based Word Clustering using Web Search Engine, in Proc. of 2006 Conference on Empirical Methods in Natural Language Processing (2006).
- 9) 松尾豊, 友部博教, 橋田浩一, 石塚満, Web 上の情報からの人間関係ネットワークの抽出, 人工知能誌, Vol. 20, No. 1, pp. 46–56 (2005).
- 10) 松下光範, 加藤恒昭, コンテキスト保持による探索的データ分析支援の枠組み, 知能と情報, Vol. 18, No. 2, pp. 251–164 (2006).
- 11) M. E. J. Newman, Finding and Evaluating Community Structure in Networks, Physical Review E69, 026113 (2004).
- 12) 野田尚志, 上窪真一, 旭敏之, 時空間の視覚化手法～年輪メタファを組み込んだ時空間ブラウジングコンテンツ～, インタラクシオン1998 論文集, pp. 135–136 (1998).
- 13) 西原陽子, 辻由紀子, 田中大智, 砂山渡, 嗜好の違いの解釈を支援するアニメーションインタフェース, 知能と情報, Vol. 19, No. 1, pp. 3–12 (2007).
- 14) 塩澤秀和, 西山晴彦, 松下温, 納豆ビューの対話的な情報視覚化における位置づけ, 情報学論, Vol. 38, No. 11, pp. 2331–2342 (1997).
- 15) Y. Seo, Y. Iwase, and Y. Takama, KeyGraph-based BBS for Supporting Online Chance Discovery Process, in Proc. of Joint 3rd International Conference on Soft Computing and Intelligent Systems and 7th International Symposium on advanced Intelligent Systems, pp. 1210–1214 (2006).
- 16) 杉山公造, グラフ自動描画法とその応用, 計測自動制御学会 (1993).
- 17) <http://www.hulinks.co.jp/software/yfiles/>
- 18) 米村俊一, 小川克彦, HI 設計ガイドラインデータベース: ブックメタファの可視化とその効果, 情報処理学会論文誌, Vol. 35, No. 10, pp. 2159–2169 (1994).

- 1) C. B. Barber, D. F. Dobkin, and H. Huhdanpaa, The Quickhull Algorithm for Convex Hulls, ACM Transactions on Mathematical Software, Vol. 22, No. 4, pp. 469–483