

検索行動の再利用のための検索クエリのトピック変換

湯本 高行[†] 森 悠太^{††} 角谷 和俊^{††}

本稿では、特定の目的に沿って検索を行った検索クエリ列を与えられた他のトピックに関する検索クエリ列に変換する手法を提案する。検索クエリの変換には、検索スケルトンという検索クエリに含まれる語の間の関係を表現するための概念を定義し、これを用いる。検索クエリの変換は、各検索クエリを検索スケルトンに変換し、検索スケルトン列のパターンに応じて、さらに別のトピックについての検索クエリに変換することによって行う。その際に、検索エンジンを用いて変換後の検索クエリに用いる語を発見する。提案手法の評価として、検索クエリの変換について、語の種類判定、関係性キーワードの発見、変換候補語の発見のそれぞれの精度を求めたところ、それぞれ、約 78.1%、約 50.0%、約 40.0%の結果が得られた。

Topic Conversion of Search Queries for Reuse of Search Activities

TAKAYUKI YUMOTO [†] YUTA MORI ^{††} and KAZUTOSHI SUMIYA^{††}

In this paper, we propose a method to convert a given sequence of search queries based on a certain topic into a sequence of search queries based on a given different topic. We define a concept of search skeleton for topic conversion. A search skeleton represents relationships between keywords in a query. A given sequence of search queries are converted into a sequence of search skeletons, and search skeletons are converted into a sequence of search queries for a target topic. We evaluated our method of search query conversion. In our evaluation, the precision to judge kinds of keywords in search queries was 78.1%, the precision to find relational keywords was 50.0%, and the precision to convert dynamic relational keywords was 40.0%.

1. はじめに

近年の Web の普及に伴い、Web 検索により情報を得ることが一般のユーザにとって日常的になっている。その一方で、検索に慣れたユーザは自由に自分の欲する情報を得ることが容易であるが、検索に慣れていないユーザにはこれは難しい。それを支援する手法として、クエリ拡張の手法^{1)~4)}が提案されており、また Google サジェスト⁵⁾のように関連するキーワードを提示するサービスが実現されている。これは検索キーワードが思いつかないという状態の 1 つの解決策にはなるかも知れないが、複数の検索を繰り返して知識を蓄積し、それを元に必要な情報に到達するというような高度な検索行動を行うには不十分である。我々は検索クエリ列から検索ノウハウを抽出し、それを他の目的の検索行動に適用することで、このような行動を

自動的に行う手法について研究を行っている(図 1 参照)。例えば、「デジタルカメラ」の「流行」について検索を行っている検索クエリ列を「テレビ」というトピックに変換することによって、「テレビ」の「流行」についての検索を行う検索クエリ列が得られる。我々はその中心的な概念として検索スケルトンを定義し、それを集団的 Web 探索に適用する手法について提案を行った⁶⁾。本稿では、その検索スケルトンの概念のモデルを改良して、より一般的な検索クエリのトピック変換に利用する手法について提案する。本手法では、検索スケルトンを仲介させて、検索クエリ列の変換を行うが、その際に検索エンジンを用いることで変換の候補の語を発見する。

本稿の構成は以下のとおりである。第 2 節では関連研究について述べる。第 3 節では改良した検索スケルトンのモデルについて、第 4 節では検索スケルトンを用いたクエリ変換の手法について述べる。第 5 節では提案手法の評価を行い、第 6 節では本稿のまとめと今後の課題について述べる。

[†] 兵庫県立大学大学院工学研究科

Graduate School of Engineering, University of Hyogo

^{††} 兵庫県立大学環境人間学部

School of Human Science and Environment, University of Hyogo

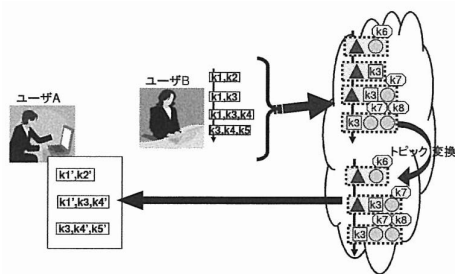


図1 検索ノウハウの適用手法

2. 関連研究

2つの事柄の対応関係を利用した情報収集手法としては、Nadamoto らの Comparative Web Browser⁷⁾ や中島らの相対的検索⁸⁾ がある。Comparative Web Browser は、同じ事柄について記述された2つの異なるニュースサイトの記事を自動的に検索して提示すると共に一方のスクロールやリンクに合わせてもう一方のページの操作を自動的に検索する機能を有する。Comparative Web Browser は2つのページの内容的な対応関係をパラグラフの類似度から求め、それを用いている。それに対して、本研究では変換する語の直接的な関連性を利用することができないため、他の語と共通した関連性キーワードを持つ語を発見する。また、中島らの相対的検索では各検索対象の重心からの距離を用いて相対的に位置付けを求め、その類似性を用いている。この手法では、検索対象の全体像がわかっていることが前提となっているが、それに対して、本手法では語の関係の保持のみを考慮しているため、ドメインの全体像を知る必要がなく、適用範囲が広いと言える。

また、変換した検索クエリに含まれるキーワードの候補の発見手法として、本研究では検索エンジンを用いてフレーズを検索する手法を用いているが、協調フィルタリング⁹⁾ を応用した手法なども考えられる。しかし、協調フィルタリングを応用した手法では大規模なクエリログデータが必要となるのに対し、本手法ではそれを必要としないため有用性が高いと考えている。検索エンジンを用いてフレーズを検索する手法を使う研究として、WebQA がある^{10)~12)}。WebQA では、具体的な質問に対して、フレーズを用いて指定された質問の答えを探す手法が主流となっており、手法には類似性がある。提案手法の改良にあたっては、WebQA の研究で得られた知見を活用できると考えている。

3. 検索スケルトン

検索スケルトンとは、検索クエリを抽象化した検索の骨格であり、検索クエリに対して、それを構成するキーワードの役割とその関係性を表現するものである。検索クエリ q は以下のようなキーワード集合によって表現される。

$$q = \{k_1, k_2, \dots, k_n\} \quad (1)$$

検索行動は以下のような検索クエリ列で表現する。

$$\langle q_1, q_2, \dots, q_n \rangle \quad (2)$$

ここで検索スケルトン s_1, s_2, \dots, s_n の検索スケルトン列は $\langle s_1, s_2, \dots, s_n \rangle$ で表す。変換は以下のように行う。

$$\langle q_1, q_2, \dots, q_n \rangle \rightarrow \langle s_1, s_2, \dots, s_n \rangle \quad (3)$$

$$\rightarrow \langle s'_1, s'_2, \dots, s'_m \rangle \quad (4)$$

$$\rightarrow \langle q'_1, q'_2, \dots, q'_m \rangle \quad (5)$$

式(3)の変換は検索クエリから検索スケルトンへの変換、式(4)の変換は検索スケルトンを目的のトピックに変換、式(5)は変換した検索スケルトンを検索クエリに変換することをそれぞれ表す。検索スケルトンは変換前の検索クエリ列のみでは決定されず、変換対象のトピックにも依存している。

3.1 キーワードの分類

検索クエリを構成する検索キーワードを検索行動における役割やそれらの間の関係性によって、以下の3つに分類する。

- 中心キーワード: 検索行動の中心となるキーワードで、対象となる検索行動中のすべての検索クエリに含まれるキーワード。
- 補助キーワード: 中心キーワードと共に使われるキーワード。
- 関係性キーワード: 中心キーワードと補助キーワードの関係性を示すキーワードで、検索クエリ中には直接現れない。

さらに、補助キーワードは静的補助キーワードと動的補助キーワードの2つに分類できる。静的補助キーワードとは、検索クエリを変換する際に変換しなくてもそのまま使用することができるようなキーワードである。例えば、「デジカメ」を中心キーワードとする検索クエリ「デジカメ 新製品」の中心キーワードを「テレビ」に変換する場合、キーワード「新製品」は意味が変わることなく使用できる静的補助キーワードである。また、動的補助キーワードとは、検索クエリを変換する際に他のキーワードに変換しなければ、元々の検索行動の意味と異なってしまうようなキーワードである。例えば、「デジカメ」を中心キーワードとするク

エリ「デジカメ 顔認識」の中心キーワードを「テレビ」に変換する場合、クエリ「テレビ 顔認識」で検索しても望む解は得られないと考え、「顔認識」は動的補助キーワードである。このとき、「デジカメ」における「顔認識」にあたる語を「テレビ」において探して置換することが妥当である。動的補助キーワードの置換にあたって重要になるのが、関係性キーワードである。関係性キーワードは、検索クエリに含まれるキーワードではなく、中心キーワードと補助キーワードの関係性を示すキーワードである。例えば、前述の「デジカメ」と「顔認識」の例であれば、関係性キーワードは「新技術」や「トレンド」などが考えられる。動的補助キーワードの変換は共通する関係性キーワードを持つキーワードによって行う。

3.2 検索スケルトンのモデル

検索クエリ q に対する検索スケルトン s は以下のよう
に定義する。

$$s = (t, R)$$

$$\text{ただし, } R = \{(k, r) | k \in q, r \notin q\} \quad (6)$$

t は中心キーワード、 k は補助キーワード、 r は中心キーワード t と補助キーワード k の関係を表す関係性キーワードである。ただし、 k が静的補助キーワードの場合は $r = *$ として表記する。

本研究は、単一の検索クエリの変換ではなく、検索クエリ列で表される検索行動の変換を目的とする。ここでは、1つの検索行動を共通したキーワード含む連続した最長の検索クエリ列とし、共通するキーワードを中心キーワード、それ以外を補助キーワードとする。

3.3 補助キーワードの分類の決定

静的補助キーワード、動的補助キーワードの判定は、検索エンジンを用い、判定対象となるキーワードと2つの中心キーワードのそれぞれとの関係を比較することによって行う。具体的には、判定対象キーワードと、それぞれの中心キーワードを合わせてひとつのクエリとして検索エンジンに問い合わせを行った際の検索結果を元に判定を行う。キーワード t と k の近傍共起度 Noc を以下のように定義する。

$$Noc(t, k) = P(t, k)/n \quad (7)$$

ただし、 $P(t, k)$ は検索キーワード「 $t k$ 」で検索したときの上位 n 件のスニペットにおいて、キーワード t と k が同一文中に出現した文書数を示す。これは、スニペットが検索キーワードの近傍の文字列からなるという性質を用いることで文書全体を分析することを避け、処理を高速化することが目的である。また、 Noc の値域は $[0, 1]$ である。 $Noc(t, k)$ が高いほど近傍で共起しやすい、すなわち、より関連が強いと見なすこと

ができる。そこで、検索行動を行っているユーザの中心キーワードを t_1 、検索行動を利用するユーザの中心キーワードを t_2 とし、判定を行う判定対象キーワードを k として、 Noc の値のよって補助キーワードを分類する。以下の式を満たす場合、すなわち k は t_1 とは関連が強いが、 t_2 とは関連が弱い場合は、 k は動的補助キーワードとし、それ以外の場合は k は静的補助キーワードとする。

$$Noc(t_1, k) \geq \theta \wedge Noc(t_2, k) < \theta \quad (8)$$

ただし、 θ は 0 以上 1 以下の閾値で本稿では $\theta = 0.5$ としている。表 1 に補助キーワードの分類条件をまとめる。

表 1 補助キーワードの分類
Table 1 Classification of sub-topic keywords

	t_1 と関連あり	t_1 と関連なし
t_2 と関連あり	静的補助キーワード	静的補助キーワード
t_2 と関連なし	動的補助キーワード	静的補助キーワード

3.4 関係性キーワードの抽出

中心キーワード t と動的補助キーワード d を用いて、関係性キーワード r を抽出する。検索エンジンに対して「 t の*は d 」というフレーズを用いて検索質問を生成し、検索エンジンに対して問い合わせを行い、検索結果の上位 n 件のスニペットから「 t の~は d 」というような文字列を検出し、「 t の」と「は d 」の間の文字数が 2 以上閾値 l 以下の場合、その間の文字列を関係性キーワードとして取得する。例えば、中心キーワードが「デジカメ」で、この検索行動中でのキーワード「顔認識」が動的補助キーワードの場合、生成される検索質問は「“デジカメの*は顔認識”」であり、「トレンド」などのキーワードが関係性キーワードとして抽出できる。

4. トピック変換

検索スケルトンを対象となるトピックに変換するためには、検索パターンによって検索スケルトンを変形した後、中心キーワードの置換、動的補助キーワードの変換を行う。具体的には以下の手順で変換を行う。

- (1) 検索クエリから検索スケルトンの導出: 3 に示した方法で、検索スケルトンを導出する。
- (2) 検索スケルトン列の変形: 検索スケルトン列をそのパターンに応じて変形する。詳しくは 4.1 にて述べる。
- (3) 検索スケルトンのトピックの変換: 検索スケルトンのトピックを与えられたトピックに変換する。詳しくは 4.2 にて述べる。

- (4) 検索スケルトンから検索クエリへの変換: 以下のようにして, 検索スケルトンから検索クエリを求める.

$$\begin{aligned} & (t, \{(k_1, r_1), \dots, (k_n, r_n)\}) \\ & \rightarrow \{t, k_1, \dots, k_n\} \end{aligned} \quad (9)$$

- (5) 重複した検索クエリの削除: 前方から走査し, 重複した検索クエリを削除する.

4.1 検索パターンによる検索スケルトンの変形

検索行動には試行錯誤状態が含まれていることがあるが, その部分まで変換して検索クエリとして提示することは有効ではない. そこで, いくつかの検索パターンを定義し, それに合致する検索スケルトン列は簡略化するものとする. ここで定義する変換は多対一の変換を行うものである. 検索パターンは以下の3種類を定義する.

4.1.1 比較検索パターン

連続する3つ以上の検索スケルトンにおいて, 動的補助キーワードのみが変化し, 関係性キーワードが同じである場合, 比較検索パターンと定義する. この場合, 共通の上位概念を持つ並列関係のキーワードに関する情報を調べようという意図を持っていると考え, 比較検索パターンが当てはまる部分の検索スケルトンのうち, 変更されている補助キーワードを共通する関係性キーワードに変更し, さらにそれを静的補助キーワードとして扱う1つの検索スケルトンへと集約する. すなわち, 以下のような変形を行う.

$$\langle s_1, s_2, \dots, s_n \rangle \rightarrow \langle (t, \{r, *\}) \rangle \quad (10)$$

ただし, $n \geq 3$, 任意の $i = 1, \dots, n$ に対して, s_i は以下を満たす.

$$s_i = (t, R_i) \quad (\text{ただし}, (k_i, r) \in R_i) \quad (11)$$

4.1.2 詳細化検索パターン

2つの連続する検索スケルトンにおいて, キーワードの追加が行われている場合, 詳細化検索パターンと定義する. このとき, ユーザは直前のキーワードにおける, 追加したキーワードについての情報をより詳しく知りたいという意図を持っていると考え, パターンに当てはまる部分の検索スケルトンで, 前半の部分に当たる検索スケルトンを削除する. すなわち, 以下のように変形を行う.

$$\langle (k, R_1), (k, R_2) \rangle \rightarrow \langle (k, R_2) \rangle \quad (12)$$

ただし, $R_1 \subset R_2$ である.

4.1.3 緩和検索パターン

2つの連続する検索スケルトンにおいて, キーワードの削除が行われている場合, 緩和検索パターンと定義する. このとき, ユーザは削除したキーワード以外の情報をより広範に渡って知りたいという意図を持つ

ていると考え, パターンに当てはまる部分の検索スケルトンで, 前半の部分に当たる検索スケルトンを削除する. すなわち, 以下のように変形を行う.

$$\langle (k, R_1), (k, R_2) \rangle \rightarrow \langle (k, R_2) \rangle \quad (13)$$

ただし, $R_1 \supset R_2$ である.

4.2 検索スケルトンのトピックの変換

検索スケルトンのトピック変換では, 中心キーワードと動的補助キーワードの置換を行う. 置換先の中心キーワードは変換先のトピックとして与えられているが, 動的補助キーワードは, 対象となるの中心キーワードに適したキーワードを発見する必要がある. ユーザの検索行動中における中心キーワード t_1 と動的補助キーワード d_1 の関係を表すような関係性キーワード r が存在するとする. これに対して, 対象となる中心キーワードを t_2 とする場合, 関係性キーワードとして r を持つ t_2 の動的補助キーワード d_2 を求め, d_1 を d_2 に置換する. すなわち, 以下のようにして, 検索クエリ「 $t_1 d_1$ 」は「 $t_2 d_2$ 」に変換される.

$$\{t_1, d_1\} \rightarrow (t_1, (d_1, r)) \quad (14)$$

$$\rightarrow (t_2, (d_2, r)) \quad (15)$$

$$\rightarrow \{t_2, d_2\} \quad (16)$$

変換された動的補助キーワード d_2 は関係性キーワード r と, 対象となる中心キーワード t_2 を用いて発見する. まず, これら2つのキーワードを組み合わせて「 t_2 の r は」という検索質問を生成し, 検索エンジンに対して問い合わせ, この検索質問に対する検索結果 Web ページ上位 n 件のリストを取得する. そこから「 t_2 の r は」という文字列を検出し, その直後に出てくる名詞または複合名詞を, 検索トピック t_2 の範囲における変換後の動的キーワード d_2 とする. また, 動的補助キーワードと判断された場合においても変換キーワードが発見できない場合がある. そのような場合は, 静的補助キーワードと同等に扱い, 変換を行わない.

ここで例を示す. 検索行動中の中心キーワードが「デジカメ」で, この検索行動中のキーワード「顔認識」が動的補助キーワードとして判定され, 関係性キーワードとして「トレンド」が抽出された場合の例で説明する. この「顔認識」という動的キーワードを「テレビ」という検索トピックに適合したものに変換させるとき, 動的キーワード変換のために生成される検索質問は「テレビのトレンドは」である. この検索質問で検索エンジンに対して問い合わせを行った際に, 検索結果中より得られる「テレビのトレンドは」の直後に出てくる複合名詞として「有機 EL」などのキーワードを取得することができる.

5. 評価

5.1 検索クエリの変換

検索クエリの一対一の変換の精度を求め、評価を行う。検索クエリの一対一の変換は主に以下の3つのステップにて実行される。

- (1) 補助キーワードの種類の決定
- (2) 関係性キーワードの発見
- (3) 動的補助キーワードの変換

上記の3つのそれぞれに対して、評価を行った。

5.1.1 補助キーワードの種類の決定

提案したアルゴリズムによって補助キーワードを動的補助キーワードと静的補助キーワードに分類を行い、その分類結果をユーザにより判定させることにより、精度を求めた。用意した補助キーワードは表2に示すとおりである。このうち、括弧で囲まれたキーワードはユーザによって不正解とされたものである。動的・静的補助キーワードの判定の全体の精度は $25/32=78.1\%$ であったが、動的補助キーワードの判定の精度は46.2%であった。

5.1.2 関係性キーワードの発見

5.1.1の実験にてユーザによって動的補助キーワードと判定されたキーワードの関係性キーワードを提案したアルゴリズムによって求め、それをユーザに判定させた。結果を表3に示す。このうち、括弧で囲まれたキーワードはユーザによって不正解とされたキーワードである。表3より、精度は50.0%となった。

5.1.3 動的補助キーワードの変換

5.1.1の実験にてユーザによって動的補助キーワードと判断されたキーワードの変換を行い、その結果をユーザに判定させた。表4に実験に使用したキーワードとその変換結果を示す。括弧で囲まれたキーワードはユーザにより不正解とされたキーワードで、空欄は対応する種類のキーワードが発見できなかったことを示す。ユーザが正解と判定した動的補助キーワードを発見できた関係性キーワードの割合を精度とすると、精度は40.0%であった。

5.2 トピック変換の例

デジタルカメラについての検索行動をテレビに変換する場合の例を示す。検索クエリ列を表5に示す。このクエリから検索スケルトンを抽出し、検索行動パターンによって検索スケルトン変換すると、表6のようになる。ここでは、「デジカメ」が検索行動における中心キーワードと判定され、「新製品」「格安」というキーワードは静的補助キーワードとして判定された。「顔認識」「手ブレ」「高感度」「一眼レフ」「ミノ

表5 変換前のクエリ

Table 5 Original queries

デジカメ	新製品	
デジカメ	格安	顔認識
デジカメ	格安	手ブレ
デジカメ	格安	高感度
デジカメ	格安	一眼レフ
デジカメ	顔認識	
デジカメ	顔認識	ミノルタ

ルタ」というキーワードは、動的補助キーワードとして判定され、関係性キーワードとして「トレンド」や「進歩」、「メーカー」といったキーワードが抽出された。また、比較検索パターンと緩和検索パターンが判定され、その範囲に該当するスケルトンの変換が行われた。その他の範囲に関しては検索行動パターンが検出されないため、そのままの構造となった。検索スケルトン内の動的補助キーワードは、変換先の中心キーワードに合わせて、それぞれ同じ関係性キーワードを持つキーワードに変換された。表6の検索スケルトンを検索クエリに変換すると、表7のような検索クエリ列が得られた。

表7 変換後のクエリ

Table 7 Converted queries

テレビ	新製品	
テレビ	格安	トレンド
テレビ	格安	薄型
テレビ	有機 EL	シャープ

6. おわりに

本稿では、検索行動を他のトピックに対して適用することによって、クエリの変換を行う手法について提案を行った。具体的には、検索スケルトンを提案し、それをを用いて検索クエリを与えられた検索キーワードに適用する手法を提案した。さらに、検索クエリ変換の評価を行うと共に、本手法の利用例を示した。検索クエリ変換の評価については、補助キーワードの種類の判定については78.1%、関係性キーワードの発見では50.0%、動的補助キーワードの変換については40.0%の精度が得られた。

今後は補助キーワードの種類の判定や関係性キーワードの発見に用いるフレーズの種類数を増やしたり、フレーズの自由度を高めるなどの手法を検討し、検索クエリの変換精度の向上を図る必要がある。また、現在は1単位分の検索行動を共通した1つのキーワードが含まれる連続した検索クエリ列として定義したが、2つの事柄を比較を行いながら検索を行う場合など必

表 2 補助キーワードの分類実験
Table 2 Experiment of classifying sub-topic keywords

変換前のトピック	変換後のトピック	動的補助キーワード	静的補助キーワード
デジカメ	テレビ	顔認識 手ブレ (光学ズーム) (ニコン) ミノルタ (高感度)	格安 人気 ソニー 薄型
液晶	プラズマ	(シャープ) アクオス	画質 消費電力 サイズ 寿命 価格 動画 パナソニック ピエラ
武田信玄	上杉謙信	(風林火山) (長篠の戦い)	川中島の戦い 肖像 国 時代
姫路	小樽	(城) 駅そば 菓子博	港 観光 食事
精度		6/13=46.2%	19/19=100.0%

表 3 関係性キーワードの発見実験
Table 3 Experiment of finding relational keywords

変換前のトピック	変換後のトピック	動的補助キーワード	関係性キーワード
デジカメ	テレビ	顔認識	トレンド
		手ブレ	はやり (映像) 進化 トレンド
		ミノルタ	メーカー
液晶	プラズマ	アクオス	(画質) (シャープ)
姫路	小樽	菓子博	(動物園)
		駅そば	(人間)
精度			5/10=50.0%

ずしも同じキーワードが連続しない場合も存在する。そのような状況における1単位分の検索行動についても考察していく必要があると考えている。さらに、現在は適用する検索行動を陽に与えて変換を行っているが、他のユーザと検索ノウハウを共有するなど検索行動が増えていった場合のユーザの利便性を考えると、ユーザが与えたトピックに関して適用可能な検索行動を動的に選択する手法についても検討が必要である。

謝辞 本研究の一部は、平成20年度文部科学省科学研究費補助金基盤研究(B)「ユーザの潜在的意図を用いたレス・コンシャス情報検索基盤の構築」(研究代表者:角谷和俊, 課題番号:20300039)によるものです。ここに記して謝意を表すものとします。

参考文献

- 1) Xu, J. and Croft, W. B.: Query expansion using local and global document analysis, *SIGIR '96: Proceedings of the 19th annual international ACM SIGIR conference on Research and development in information retrieval*, New York, NY, USA, ACM, pp. 4-11 (1996).
- 2) Cui, H., Wen, J.-R., Nie, J.-Y. and Ma, W.-Y.: Probabilistic query expansion using query logs, *WWW '02: Proceedings of the 11th international conference on World Wide Web*, New York, NY, USA, ACM, pp. 325-332 (2002).
- 3) Salton, G. and Buckley, C.: Improving retrieval performance by relevance feedback, pp. 355-364 (1997).

表 4 動的補助キーワードの変換実験

Table 4 Experiment of converting dynamic sub-topic keywords

変換前の動的補助キーワード	関係性キーワード	変換後の動的補助キーワード
顔認識	トレンド	有機 EL, デザイン, (リンク録画), (動画ボケ), (画素), (国)
	はやり	
手ブレ	(映像)	(多数), (スムーズ), (特別), (以下), (強烈), (イマイチ), (横)
	進化	画質, (劇的)
	トレンド	有機 EL, デザイン, (リンク録画), (動画ボケ), (画素), (国)
ミノルタ	メーカー	(goo), (モニター), シャープ, (HDD), (設定), SAMSUNG
(アコス)	(画質)	(数値), (ブラウン管), (机上), (今一つ), (レベル), (圧巻)
	(シャープ)	
葉子博	(動物園)	
駅そば	(人間)	(鉄道)
精度		4/10=40.0%

表 6 検索スケルトンの変形

Table 6 Transformation of search skeleton

変形前の検索スケルトン	変形後の検索スケルトン
(デジカメ, {(新製品, *)})	(テレビ, {(新製品, *)})
(デジカメ, {(格安, *), (顔認識, トレンド)})	(テレビ, {(格安, *), (トレンド, *)})
(デジカメ, {(格安, *), (手ブレ, トレンド)})	
(デジカメ, {(格安, *), (高感度, トレンド)})	(テレビ, {(格安, *), (魅力, *)})
(デジカメ, {(格安, *), (一眼レフ, 魅力)})	
(デジカメ, {(顔認識, はやり)})	(テレビ, {(有機 EL, はやり), (シャープ, メーカー)})
(デジカメ, {(顔認識, はやり), (ミノルタ, メーカー)})	

- 4) 村田真哉, 戸田浩之, 松浦由美子, 片岡良治: 検索結果中のアクセス集中サイトを利用したクエリ拡張法の提案, 日本データベース学会 Letters, Vol. 6, No. 4, pp. 45-48 (2008).
- 5) Google: Google サジェスト.
- 6) 森悠太, 湯本高行, 角谷和俊: 集团的 Web 情報探索のための検索行動の抽出・適用方式, 電子情報通信学会 第 19 回データ工学ワークショップ (2008).
- 7) Nadamoto, A. and Tanaka, K.: A comparative web browser (CWB) for browsing and comparing web pages, *WWW '03: Proceedings of the 12th international conference on World Wide Web*, New York, NY, USA, ACM, pp. 727-735 (2003).
- 8) 中島伸介, 田中克己: 相対的マッピング処理に基づく相対的情報検索手法, 情報処理学会論文誌: データベース, Vol. 45, No. SIG 4(TOD 21), pp. 63-75 (2004).
- 9) Resnick, P., Iacovou, N., Suchak, M., Bergstorm, P. and Riedl, J.: GroupLens: An Open Architecture for Collaborative Filtering of Netnews, *Proceedings of ACM 1994 Conference on Computer Supported Cooperative Work*, Chapel Hill, North Carolina, ACM, pp. 175-186 (1994).
- 10) Dumais, S., Banko, M., Brill, E., Lin, J. and Ng, A.: Web question answering: is more always better?, *SIGIR '02: Proceedings of the 25th annual international ACM SIGIR conference on*

Research and development in information retrieval, New York, NY, USA, ACM, pp. 291-298 (2002).

- 11) Kwok, C., Etzioni, O. and Weld, D. S.: Scaling question answering to the web, *ACM Trans. Inf. Syst.*, Vol. 19, No. 3, pp. 242-262 (2001).
- 12) Jijkoun, V. and de Rijke, M.: Retrieving answers from frequently asked questions pages on the web, *CIKM '05: Proceedings of the 14th ACM international conference on Information and knowledge management*, New York, NY, USA, ACM, pp. 76-83 (2005).