

## 言語横断情報検索のクエリ翻訳曖昧性解消のための画像情報の利用

林 良彦<sup>1</sup>, Bora Savas<sup>1</sup>, 永田昌明<sup>2</sup>

<sup>1</sup> 大阪大学大学院言語文化研究科

<sup>2</sup> NTT コミュニケーション科学基礎研究所

### 1. はじめに

クエリ翻訳型の言語横断情報検索においては, ユーザによる原言語のクエリを適切に目的言語に翻訳することが必要である. そこで, 人手の介入を許すインタラクティブな検索インタフェースにおいて, クエリの翻訳曖昧性を解消する手がかりとして画像情報をユーザに提示することの有効性について検討する. 実現したインタラクティブ検索インタフェースは, クエリ中の各語に訳語候補に対する画像検索を実行し, その結果のサムネール一覧をユーザに提示する. ユーザはこの画像情報を手掛かりとしてクエリ語の翻訳を決定する. 本稿では, 画像情報の提示の有効性に関する評価実験の結果について報告する.

### 2. 言語横断情報検索における課題

言語横断情報検索システムを構成する方法論としてはいくつかのものがあるが, ユーザによる原言語のクエリを目的言語へと変換し, 目的言語における情報検索を行うシステムが主流である. このようなクエリ翻訳型の言語横断情報検索における課題は以下のようにまとめられる.

1. ユーザのクエリを適切に目的言語に変換する
2. 目的言語における情報検索を適切に行う
3. 目的言語における検索結果を適切に提示する

これらの中でとくに重要なのは 1 の課題である. 一般に情報検索のクエリは数語の単語からなるため, 目的言語への変換を適切に行うための情報が不足していることが多く, クエリ翻訳の不適切さが検索の有効性に影響を与えることがある. この問題を解決するため, 何らかのユーザの介入を許すインタラクティブな検索システムも多く提案されている(Oard, 2001). ユーザが十分な目的言語の知識を持つことが想定できない場合は, これを考慮した支援が必要となる. 2 の課題は通常の情報検索における課題でもあるが, 特に 1 との関連でいえば, クエリの適切な変換がなされたとしても, その結果が目的言語で有する多義性の対処が必要となる. さらに, 3 に関しては, 実際に

アクセスすべき文書をユーザが選択する過程や文書の読解, また, 次の検索クエリの選択などへの支援が必要となる. この課題に関してどのような支援が可能であるかは, 最終的な情報検索のターゲットがどのようなものか, また, 検索対象のドキュメントに関して利用できる情報に依存する. 以上の課題は, ある程度相互に関連しており, 総合的な対処が必要であるが, 本稿では特に 1 の課題をとりあげ, クエリ語を翻訳する際に画像情報を利用するインタラクティブなインタフェースにおける画像情報の利用の有効性を実験的に示す.

### 3. インタラクティブな言語横断情報検索インタフェース

今回実現した言語横断情報検索インタフェースの構成を図 1 に示す. 人手の介入により原言語におけるクエリの翻訳曖昧性を解消するためにクエリ中の各語に対する訳語候補に対応する画像情報を提示する. このために, 外部の画像サーチエンジンを利用する. 翻訳の曖昧性解消後は, ターゲットとなるサーチエンジンにアクセスして目的言語における情報検索を行う. 本インタフェースは, 特定の外部エンジンに固有の処理は行わないため, 画像サーチエンジン, ターゲットのサーチエンジンともに Web 上で利用可能な適当なものを利用することが可能である.

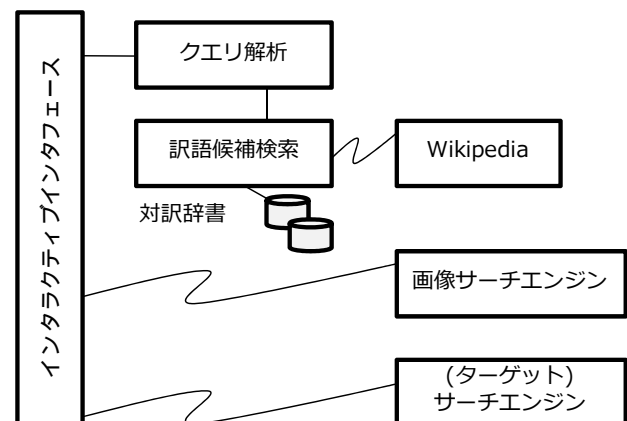


図 1: インタラクティブな検索インタフェースの構成

現時点では以下の理由により、画像サーチエンジンとして、Google イメージ検索<sup>1</sup>を利用している。

- 蓄積している画像情報の量が多く、様々な言語の語義に対応した画像が得られることが期待できる
- 蓄積している画像のタイプが写真に限定されていない
- Ajax API<sup>2</sup>によりサムネイル画像が取得可能である

また、ターゲットとなるサーチエンジンとして、今回は一般の Web 検索ではなく、主に写真画像の共有を目的としたサイトである flickr<sup>3</sup>を指定している。これは言語横断情報検索のターゲットとして非テキストメディアの情報検索が有用であり(Frank, 2000)、とくに画像を検索対象とする言語横断情報検索 (Cross Language Image Retrieval) の研究が盛んに行われている<sup>4</sup>ということが理由である。

現時点のシステムは、日本語クエリに対して、英語、フランス語、ドイツ語、韓国語、中国語などの約 20 言語の多言語同時並行検索を行うことができる。クエリ翻訳においては、EDICT/JMDict<sup>5</sup>、NTT 研究所による日・中・韓対訳辞書などの言語資源を用いるほか、Wikipedia を on-the-fly で検索して各国語の訳語を得ている<sup>6</sup>。

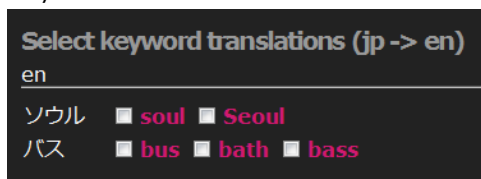


図 2(a):訳語候補の提示



図 2(b):“bus”に対する画像・逆翻訳提示

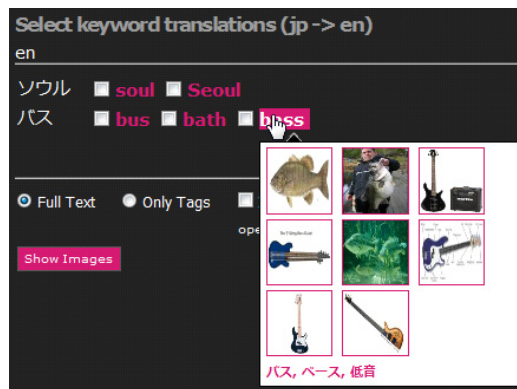


図 2(c):“bass”に対する画像・逆翻訳提示

図 2 に“ソウルのバス”という日本語クエリに対する検索インタフェースの情報提示の例(英語)を示す。(a)は、クエリ中の各単語に対する英訳語候補の提示を示す。“ソウル”に関して 2 つの訳語候補，“バス”に関して 3 つの訳語候補が得られている。ユーザは、(b), (c) に示すような画像サーチエンジンから得られたサムネイル画像と、訳語候補を英日対訳辞書により日本語へ逆翻訳した結果を手掛かりとして、実際の検索に用いる訳語を決定する。(b)は、「乗合自動車」の意味の“bus”に関する画像検索の結果、(c)は「楽器」または「魚」の意味の“bass”に関する画像検索の結果を示している。(c)の例は、目的言語において多義性があることを視覚的に示すことに成功しているが、一般に Web 上の情報には一定のバイアスがあるため、語義に対応する画像が適切に得られる保証はない。このことは、語義の数と提示するサムネイルの枚数にも影響を受ける。

## 4. 画像情報利用の有効性評価

### 4.1 評価実験の設定

クエリ翻訳の曖昧性解消のために画像情報を利用することの基本的な有効性を確認するため、クエリ翻訳の曖昧性解消の手がかりとして、クエリ語の逆翻訳を提示する場合と Google イメージ検索を利用したサムネイルを提示する場合を比較した。また、目的言語の知識の影響をしらべるため、原言語のクエリ語の翻訳結果をユーザに同時に提示する場合としない場合の比較も行った。被験者は、本論文の著者 3 名のほか 1 名による計 4 名であり、いずれも一定の英語の知識・能力を有している。

評価の指標は、あらかじめ準備したクエリセットにおいて、意図した目的言語のクエリ語句を選ぶことができたクエリの割合(正解訳語選択率)と、手がかり情報が表示されてからクエリ選択の決定を下すまで経過時間(選択時間)

<sup>1</sup> <http://images.google.co.jp/>

<sup>2</sup> <http://code.google.com/intl/ja/apis/ajaxsearch/>

<sup>3</sup> <http://www.flickr.com/>

<sup>4</sup> ImageCLEF <http://imageclef.org/>

<sup>5</sup> <http://www.csse.monash.edu.au/~jwb/jmdict.html>

<sup>6</sup> 日本語のクエリ語で Wikipedia(<http://ja.wikipedia.org/>)を検索し、「他の言語」のリンク集における URL 文字列から各国語の訳語を得る。

の実測値の二つとした。すなわち、より素早く所定の目的言語のクエリ語句を選択できるケースを確認し、また、これらの二つの指標値が損なわれるクエリがあれば、その特性を調べることを評価実験の目的とする。

#### 4.2 クエリセット

クエリセットは 39 の原言語(日本語)におけるクエリ語からなる。実際の言語横断検索において、“ソウルのバス”のようなクエリの場合、構成語の語義間の共起関係を制約として翻訳語を絞り込むことができるが、今回は純粋に画像情報の有用性を調べるために、すべて単一の語基からなる単語とした。表 1 にクエリセットの概要を示す。基本的にはすべてが日本語側で多義であるが、とくに 2 語(“高速”, “携帯”)は略語となっている。

表 1: クエリセットの概要

特性	数	例
カタカナ語	13	ミス, ドライバー, ソウル
うち正解が固有名詞	2	ソウル, ルーブル
漢字語	26	勉強, 計算機
うち一文字語	21	橋, 島, 車
物体画像が想起しやすい	34	髭, 犬, スタンド
想起しにくい	6	青, 空, 家庭
日本語での語義が英語側でマイナー	2	グローブ (glove), 表 (table)
正解と同義・類義の英単語がある	3	車 (car/vehicle), 髭 (beard/mustache)
systematic polysemy (Cruse, 2004, p.111)	2	鳥 (動物:食料), 木 (植物:材料)

各クエリ語に対しては、EDICT/JMDict から選択した最大 5 つの候補翻訳語句が与えられている。また、各クエリ語に対応する検索意図があらかじめ想定されており、これに対応する正解翻訳語句が一つ定められている。なお、ターゲットのサーチエンジンを flickr とし、正解翻訳語句、および、候補翻訳語句のすべてを OR 結合した場合のクエリに対する関連性の評定を各クエリの検索結果上位 30 件に対して行っている。評定は本論文の著者の 1 名が行った。その結果、正解翻訳語句に対する{P@30, S@30}<sup>7</sup>は{52.5%, 96.7%}, OR クエリに対しては{26.1%, 83.3%}であった。

<sup>7</sup> P@30:上位 30 件までの結果で relevant ドキュメントの割合。S@30:上位 30 件まで見た場合に relevant なドキュメントがひとつでも含まれるクエリの割合。

#### 4.3 実験結果

実験におけるケースは、逆翻訳結果/画像情報(サムネール)のどちらを表示するか、翻訳語句表示の有無の組み合わせから 4 通り存在する。表 2 に各ケースの 4 名の被験者の平均正解翻訳語句選択率を、表 3 に各ケースの 4 名の被験者の平均選択時間を示す。なお、実験において、各翻訳語句に対応する情報(逆翻訳, 画像)が固定位置に表示されるバイアスを避けるため、表示順は適宜ランダム化している。

表 2: 平均正解翻訳語句選択率 (%)

	翻訳語句表示:有	翻訳語句表示:無
逆翻訳結果	91.0	60.3
画像情報	91.0	89.1

表 2 の結果から以下のことが言える。

- 翻訳語句の表示のない場合、画像情報が非常に有用な手掛かりとなる。このケースは、目的言語の情報が表示されていても全く役に立たない状況を近似するものであるから、目的言語の知識・能力が不十分なユーザにおいても画像情報を手掛かりに適切なクエリ翻訳を行える可能性を示唆する。
- また、上記のケースにおいて、逆翻訳結果だけの提示では十分にクエリ翻訳の曖昧性が解消できない。これは、逆十分な手がかりとならない場合が多いことを示唆する。詳細については、次節で議論する。
- 一方で、翻訳語句の表示も正解選択に有効である。とくに、画像情報の表示がなくとも翻訳語句の表示があれば高い率で正解語句の選択が行えていることがわかる。ただしこれは、今回の実験の被験者が一定の英語の知識・能力を持つことによるためである。

表 3: 平均選択時間 (ms)

	翻訳語句表示:有	翻訳語句表示:無
逆翻訳結果	3718.0	5166.1
画像情報	3481.4	4305.8
(時間比 %)	(93.6%)	(83.3%)

表 3 の結果から以下のことが言える。

- 画像情報の提示は選択時間の短縮に有効である。画像情報を提示することにより、被験者に示される情報量は増加しているが選択時間は短縮しており、これは画像情報の持つ高い一貫性によるものと考えられる。

- とくに翻訳語句表示がない場合に選択時間の短縮に寄与している。これは前述のように逆翻訳が十分な手がかりとならない場合が多いことを示唆する。

#### 4.4 考察：画像情報のみを表示するケース

本研究の目的に近いケースとして、画像表示有、かつ、翻訳語句表示なしの場合をとりあげる。4名の被験者中、2名以上が正解の翻訳語句を選択できなかったクエリを「難しいクエリ」とすると、このようなクエリは5件あり、その内訳は、物体画像が想起できにくい抽象概念:2件(“青”, “勉強”), 英語側の類似語義が画像上からは判定しづらいもの:1件(“髭”/beard), 英語側の優勢な意味(“表”/table における「家具:テーブル」)により適切な画像が表示できなかったもの:1件, その他:1件(検索要求の意図の被験者への指示が適切でなかった)であった。

これらの難しいクエリに限った平均選択時間をみると6494.0(ms)となっており、難しいクエリの場合は、画像情報の手がかりがあっても選択に時間を要していることが分かる。なお、上記の難しいクエリを除いた34件のクエリにおける平均選択時間をみると、4名の被験者中3名において、クエリ全体の場合よりも平均選択時間が短かった(4305.8(ms) / 3984.0(ms))。

#### 5. 関連研究

画像を単語の語義解消(WSD: Word Sense Disambiguation)のための手がかりとして利用する研究の代表的なものに(Bernard and Johnson, 2005)がある。この手法は、画像処理によって抽出された画像の代表的な部分領域に対して、限定された語彙集合から選ばれた単語が注釈として付与されている画像データベースを用いて、これらの単語の語義解消を行う。一方、(藤井, 2005)や(Alm et al., 2006)は、画像に適切な意味を対応付ける問題を「画像の多義性解消」(ISD: Image Sense Disambiguation)と呼んでいる。(藤井, 2005)は、語義説明文と画像のキャプションのマッチングによる対応付け手法を提案している。(Alm et al., 2006)は、ISDにおける問題点を分類している。提示されている問題点は語義と図像の曖昧性の差異に起因するものであり、ターゲットの情報検索対象を画像とした場合のとくにクエリに対する画像の関連性(relevancy)の問題に関連する。

なお、画像を検索対象とする言語横断情報検索の評価に関しては、(Sanderson et al., 2004)がある。この論文では、システムの usability の観点から、MAP などの通

常の情報検索の指標ではなくクエリ成功率を用いることを提案している。

#### 6. おわりに

インタラクティブな言語横断情報検索において、クエリ翻訳の曖昧性の解消のために画像情報の提示が有用である可能性を実験的に示すとともに、画像情報の提示が必ずしも有効でないクエリの考察を行った。今後は以下のような課題に取り組む予定である。

- 評価実験の精緻化、規模の拡大を行う。
  - 選択時間の計測をより精緻化し信頼性を高める
  - 英語以外を目的言語とした場合の評価を行う
  - 表1に示したようなクエリのカテゴリを詳細化し、適切にバランスさせたクエリセットを定める
- 検索インタフェースに関しては、ターゲットのサーチエンジン側の情報(例:flickr における tag 集合など)を利用して、検索過程のナビゲーションを高度化する。
- 今回は Google イメージ検索の利用により比較的良好な画像情報を提示できたが、語義の曖昧性を解消するための適切な画像リソースの体系を探究する必要がある。このため、語彙的なリソースと画像リソースの対応付けの研究を進めるほか、さらに画像における物体認識技術(柳井, 2007)の利用を検討する。

#### 参考文献

- Alm, C. O., et al. 2006. Challenges for Annotating Images for Sense Disambiguation. *Proc. of Workshop on Frontiers in Linguistically Annotated Corpora*, pp.1-4.
- Barnard, K., and Johnson, M. 2005. Word Sense Disambiguation with Pictures. *Artificial Intelligence*, Vol.167, pp.13-30.
- Cruse, A. 2004. *Meaning in Language (second edition)*. Oxford University Press.
- Frank, S. 2000. Cross-Language Multimedia Information Retrieval. *Proc. of the Sixth Conference on Applied Natural Language Processing*, pp.13-20.
- Oard, D. W. 2001. Interactive Cross-Language Information Retrieval. *SIGIR Forum* Vol.35, No.1, pp. 1-3.
- Sanderson, M., et al. 2004. Measuring a Cross Language Image Retrieval System. *Proc. of ECIR 2004*, pp. 353-363.
- 藤井 敦, 石川徹也. 2005. テキスト処理による画像の多義性解消と事典検索サイトへの応用. 第11回言語処理学会年次大会, B5-1.
- 柳井啓司. 2007. 一般物体認識の現状と今後. 情報処理学会論文誌: コンピュータビジョンとイメージメディア, Vol.48, No. SIG16 (SVIM 19), pp.1-24.