DEIM Forum 2017 F8-3

英日言語横断検索におけるクエリ拡張結果の詳細分析

玉置　賢太†　　酒井　哲也†

† 早稲田大学基幹理工研究科情報理工・情報通信専攻酒井研究室　〒 169–0072 東京都新大久保 3–4–1
E-mail: †madao@akane.waseda.jp, tetsuyasakai@acm.org

あらまし　言語横断情報検索とは，クエリとは異なる言語の文書を検索する技術のことである．本研究では，クエリ翻訳による言語横断情報検索において，異なる言語間の分散表現の類似性を利用したクエリ拡張を行い，その有効性について論ずる．また，比較対象として，クエリ翻訳の前後の擬似適合性フィードバックによるクエリ拡張を用いる．今回の実験では分散表現を用いたクエリ拡張の優位性を示すことはできなかったが，本論文では両クエリ拡張方式の詳細な分析を行い，将来への課題を明確にする．

キーワード　自然言語処理，言語横断情報検索，word2vec，擬似適合性フィードバック

1. 導　　入

言語横断情報検索とは，ユーザの入力したクエリとは異なる言語の文書を検索する技術のことである [10]．その主要なアプローチは，クエリのほうを文書の言語に翻訳するものである．このアプローチでは，クエリ翻訳の精度がシステム全体の検索有効性を大きく左右することとなる．そこで，クエリ翻訳を補強する手法として，クエリ拡張を組み合わせる手法が研究され，この組み合わせにより高い検索有効性が実現できることが知られてきた [7] [19]．クエリ拡張を適用するタイミングには 2 つある．[23]．1 つは，クエリを翻訳する前に拡張する pre-translation expansion で，もう 1 つは翻訳後に拡張をする post-translation expansion である．

筆者らの研究では [21]，クエリ翻訳に対し辞書ベースなアプローチとして，Mikolov らが提案した手法 [12] を用いて，言語横断情報検索の有効性向上を図っている．Mikolov らは，word2vec [2] により得られる分散表現に，異言語にわたり線形変換によって関連付けられた類似性があることに着目した．ここでの word2vec とは，skip-gram による分散表現手法 [13] を実装したツールである．Mikolov らはこの類似性を利用するため，一方の言語の単語ベクトルを，他方の単語ベクトルに変換するための行列，翻訳行列を求めた．この翻訳行列によって翻訳を実現するのが，Mikolov らが提案した手法である．結果として，翻訳行列をクエリ翻訳に用いる手法では，ベースラインである Google 翻訳 [5] と比べて，言語横断情報検索の有効性を向上することができなかったと述べられている．

筆者らの従来研究では分散表現の類似性をクエリ翻訳に用いていたが [21]，本研究ではクエリ拡張に適用する．図 1 は従来手法と提案手法の違いを説明した図である．この図 1 のように，クエリ翻訳は Google 翻訳によって行い，翻訳行列によって得られる単語を拡張クエリタームとして用いるのが提案手法である．このクエリ拡張手法が有効であった場合，翻訳前後のコーパスを用いたクエリ拡張ではない，新しいアプローチとして利用することができる．これにより，今までは頭打ちだったシステムの性能向上に繋がる可能性がある．

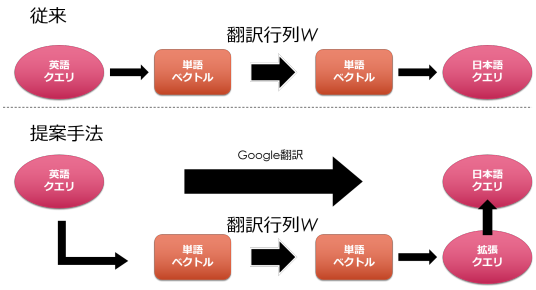

図 1　従来手法と提案手法の比較

2 章では関連研究について述べ，3 章では提案手法について述べる．また，4 章では実験方法について述べる．5 章では実験結果について述べ，6 章では結論・展望を述べる．

2. 関 連 研 究

2.1　言語横断情報検索

本研究では，言語横断情報検索の一つのアプローチとして，クエリ翻訳による情報検索を用いている．クエリ翻訳による情報検索では，翻訳に伴う語義曖昧性により，単言語検索よりも性能が劣ってしまう場合が多い [7]．Ballesteros らの研究では [7]，翻訳の前後にそれぞれの言語においてクエリ拡張を行うことで，言語横断情報検索の性能が向上するかどうか検証されている．結果として，翻訳前後にクエリ拡張を行うことで，翻訳に伴う語義曖昧性を減らし，性能が向上したと述べられている．なお，Ballesteros らは [7]，local feedback [6] と local context analysis [20] の 2 つのクエリ拡張手法による実験を行っている．

2.2　異言語にわたる分散表現の類似性

ここでの分散表現とは，word2vec によって得られたベクトル表現をいう．このベクトル空間は，コーパス中の単語の異なり数 T 次元のビットベクトルを，より低い次元で表現したベクトルの集まりとなっている [13]．また，word2vec により得られたベクトル空間は，コーパス内の単語の意味的な遠近関係を表現しているとされている．

近年，Mikolov らの研究 [12] によって，このベクトル空間に，異言語にわたる類似性が見られることがわかった．この類似

性というのは，異なる言語のコーパスによりそれぞれ学習を行ったとき，生成されるベクトル空間内の，同じ意味を持つ単語間の位置関係に類似が見られるというものである．図 2 に Mikolov ら [12] の論文から転載した類似性の例を示す．

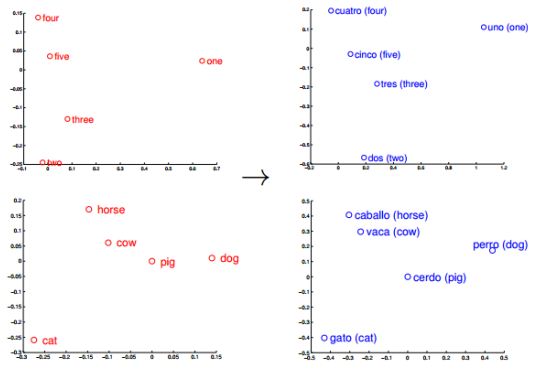

図 2　英語とスペイン語それぞれのコーパスによる分散表現の類似性 (左が英語の分散表現，右がスペイン語の分散表現．Mikolov ら [12] より転載．)

Mikolov ら [12] の論文では，この類似性を利用した機械翻訳の手法が示されている．異なる言語間の同じ意味を持つ単語同士をマッピングすることで，翻訳を実現するというものである．このため Mikolov らは，翻訳行列と呼ばれる行列を求めている．翻訳行列は，一方の言語の単語ベクトルを，同じ意味をもつ他方の単語ベクトルに変換するための行列で，以下の確率的勾配降下法 [9] の問題を解くことで求めている．

$$\min_{W} \sum_{i=1}^{n} \| Wx_i - z_i \|^2 \qquad (1)$$

なお，W は求めている翻訳行列であり，x_i は一方の言語に属する単語 i のベクトルであり，z_i は他方の言語に属する単語 i のベクトルである．x_i と z_i は同じ意味を持つ単語であるとし，一方の単語ベクトルに翻訳行列 W を掛けた時に，他方の単語ベクトルとの差が最小になるように学習を行っている．

この研究では，語族的に近い英語・スペイン語の組み合わせのほかに，英語・ベトナム語の組み合わせ等でも実験を行い，一定の効果があることが示されている．

2.3　言語の分散表現を用いた言語横断情報検索

筆者ら [21] は，上記の翻訳行列による機械翻訳を，クエリ翻訳に利用した言語横断情報検索の検証を行った．この研究では，クエリ翻訳の前後に擬似適合性フィードバックによるクエリ拡張を行った場合の実験も行われている．また，翻訳行列による翻訳に対する比較対象として，Google 翻訳 API [5] をクエリ翻訳に用いた場合についても実験を行っている．図 3 に筆者ら [21] の論文より転載した実験の概要を示す．

また，筆者らの実験の結果を表 1 に示す．筆者らの実験では，言語横断情報検索の実験・適合性判定用のデータとして，NTCIR-5，6 CLIR タスク [11] のデータセットを用いている．また，評価には NTCIR のいくつかのタスクで用いられてきた NTCIREVAL [17] を利用している．これは，様々な評価指標に

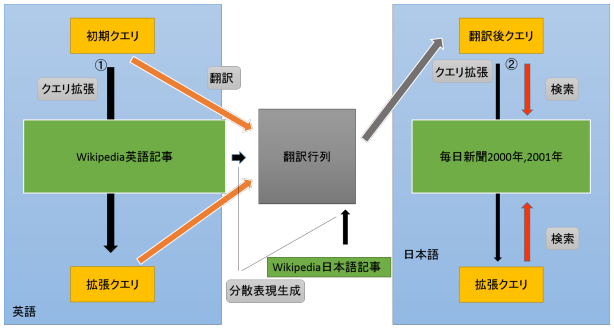

図 3　実験の概要 (筆者ら [21] より転載)

表 1　各手法における平均 MSnDCG 値の比較 (筆者ら [21] より引用)

	拡張無し	翻訳前	翻訳後	翻訳前&翻訳後
翻訳行列	0.1577	0.2240	0.1627	0.1926
Google 翻訳	0.3363	0.2053	0.3269	0.3000

より検索エンジンを評価するツールである．表 1 の MSnDCG とは，NTCIREVAL に実装されている，Burges ら [8] によって再定義されたマイクロソフト版の nDCG である [22]．

結論として，翻訳行列をクエリ翻訳に用いた言語横断情報検索では，Google 翻訳を用いた場合より性能が上がることがなかったと述べられている．原因としては翻訳精度の差が挙げられ，その差を改善する手段として次元数等のパラメータの調整が挙げられている．

3.　提 案 手 法

本研究では，2 章で述べた異言語にわたる分散表現の類似性を，クエリ翻訳としてではなく，クエリ拡張の手法として適用する．また，翻訳前に擬似適合性フィードバックを行ったうえで，翻訳行列によるクエリ拡張を行う．これらの手法により英日言語横断情報検索を行い，その効果を検証する．

3.1　クエリ翻訳

クエリ翻訳の手法としては，全文翻訳等に用いられてきた機械翻訳に加え，辞書ベース翻訳やパラレルコーパスを用いた翻訳などが考えられる [15]．また，機械翻訳の中にはルールベースな翻訳と統計的な翻訳が存在し，時としてパラレルコーパスを利用しているものもある．

本研究では，クエリ翻訳には焦点を当てていないため，既存の機械翻訳機として Google 翻訳 API [5] をクエリ翻訳に用いる．本研究中のいずれの実験でも，クエリ翻訳部分は Google 翻訳によるものとする．

3.2　翻訳行列を用いたクエリ拡張

Mikolov ら [12] は，クエリ言語の単語ベクトルに，翻訳行列 W を掛け合わせて得られたベクトルから，最も距離の近いベクトルを持つ単語を翻訳単語としていた．本研究では，図 4 のように，翻訳単語として扱っていた単語を，拡張クエリタームとしてターゲット言語の文書を検索するために用いる．また，最も距離の近い 1 語のみではなく，図 4 の提案手法のように，付近の単語 n 個を拡張クエリタームとした場合についても実験を行い，その効果を検証する．

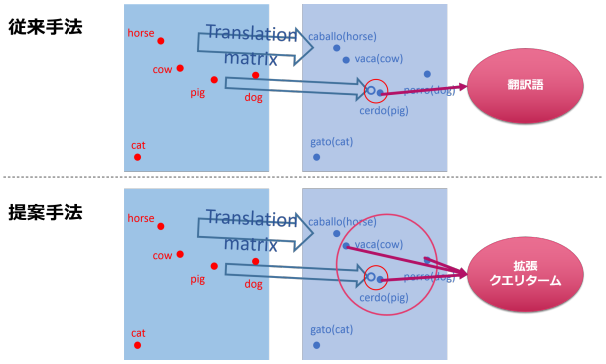

図 4　翻訳行列を用いたクエリ拡張

翻訳行列の生成には，Wikipedia の英日それぞれの記事 [4] をコーパスとして用いる．教師データの生成については林ら [23] に従い，コーパス内から 5000 語を取り出し，Google 翻訳によって同じ意味を持つ単語対を生成している．

3.3　擬似適合性フィードバック

本研究では，翻訳行列によるクエリ拡張に加えて，翻訳前に擬似適合性フィードバックによるクエリ拡張を行った場合についても実験を行う．また，ベースラインとして翻訳前のみにクエリ拡張を行う手法と，翻訳後のみにクエリ拡張を行う手法についても実験を行い，その性能の違いについて分析を行う．擬似適合文書を検索するためのコーパスには，英日それぞれの言語のコーパスを用いる．

本研究では，擬似適合性フィードバックの手法として，Robertson らによる Offer Weight [16] により拡張クエリタームを選出する．以下に Offer Weight の計算式を示す．

$$OW = r * \log\left(\frac{(r+0.5)(N-n-R+r+0.5)}{(n-r+0.5)(R-r+0.5)}\right) \qquad (2)$$

ここで，N は検索対象コーパスに含まれる全文書数であり，n はその中で着目した単語が生起する文書数である．また，R は適合文書として扱う文書の数であり，r はその中で着目した単語が生起する文書の数である．この Offer Weight を，擬似適合文書中全ての単語について算出し，降順で上位の単語を拡張クエリタームとする．

実験では，擬似適合文書数 R を 10 とし，拡張クエリターム数を Offer Weigh の上位 10 単語とした．

4.　実 験 方 法

本研究では，5 通りの手法について実験を行い，その結果を分析する．これらは，大きくベースラインと提案手法の 2 つに分けられる．図 5 に概観を示す．

ベースライン手法としては，3 通りの手法を実験する．図 5 における上 3 本の矢印が示している．1 つ目はクエリ拡張を一切行わない，クエリ翻訳のみによる言語横断情報検索である．2 つ目は翻訳前のみにクエリ拡張を行う手法で，3 つ目は翻訳後のみにクエリ拡張を行う手法である．

提案手法としては 2 通りの手法について実験を行う．図 5 の下 2 本の矢印がこの手法を示している．1 つ目は，翻訳行列によるクエリ拡張のみを行う手法である．2 つ目は，翻訳前に擬似適合性フィードバックによるクエリ拡張を行い，その拡張クエリタームを含めて翻訳行列によるクエリ拡張を行う手法である．翻訳行列によるクエリ拡張に関しては，付近の単語ベクトル n 個を拡張クエリタームとするが，$n = 1, 10, 20$ について実験を行う．

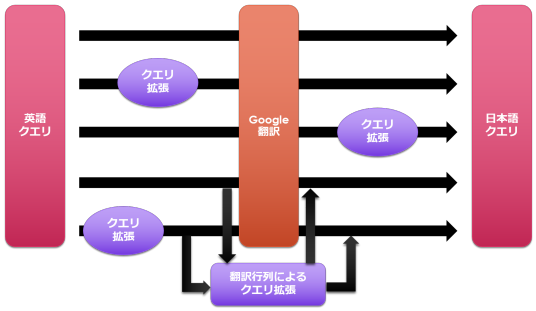

図 5　実 験 方 法

以下に実験で用いるデータ及びツールについて示す．

4.1　データセット

本研究では，検索対象記事及び適合性判定データとして NTCIR-5，6 CLIR タスク [11]・NTCIR-7 ACLIA タスク [14] のデータセットを用いている．検索対象は，毎日新聞の 2000 年及び 2001 年の新聞記事であり，約 20 万記事が含まれる．林ら [23] に従い，C0 タグで表される索引記事番号を ID とし，T2 タグのみを本文として扱う．また，トピックはデータセットに含まれる合計 97 トピックとし，クエリとして扱うのは各トピックのタイトルフィールドとする．

評価データとして用いたのが NTCIR-5，6 CLIR タスクであり，NTCIR-7 ACLIA タスクのデータセットは訓練用データとして用いた．なお ACLIA タスク用のデータセットは，検索対象コーパスが毎日新聞 1998 年から 2001 年の新聞記事となっている．このため本研究では，適合性判定データから 1998 年 1999 年の記事 ID を取り除いて利用した．

4.2　ベクトル空間の生成

翻訳行列の生成に用い，またクエリ拡張のためにも用いる英日それぞれの言語のベクトル空間生成には，Wikipedia の英日記事 [4] を用いる．Wikipedia 英語記事は，筆者取得時で約 460 万記事が含まれ，日本語記事はこちらも筆者取得時で約 90 万記事が含まれる．また，word2vec で学習を行う際の次元数のパラメータは，英語側で 800 次元，日本語側で 200 次元としている．この値は Mikolov ら [12] が，ソース言語である英語を 800 次元，ターゲット言語のスペイン語を 200 次元としているのに従っている．

4.3　擬似適合性フィードバック

擬似適合性フィードバックによるクエリ拡張では，翻訳前後それぞれの言語のコーパスから擬似適合文書を検索する．本研究では，翻訳前の言語である英語のコーパスとして Wikipedia 英語記事を用いる．これは，4 章 2 節で用いるものと同様に約 460 万記事を含んでいる．翻訳語の日本語のコーパスとしては，検索対象コーパスである毎日新聞 2000 年及び 2001 年の記事を

用いている．

4.4 検索エンジン

本研究では，擬似適合性フィードバックとターゲットコーパスの検索に，検索エンジン Indri [3] を用いている．Indri の機能により，indexing 及び検索を行っている．また，Indri の検索機能により，初期クエリの重みを 1.0 とし，拡張クエリタームの重みは 0.2 として実験を行った．

4.5 評価指標

本研究では，実験結果の評価のために，NTCIREVAL [17] を利用した．特に本研究では，検索上位 1000 件までの文書で評価された MSnDCG [8] の値を用いる．なお，nDCG は正規化された値である．このため，最大値は 1 となる．

5. 実験結果

5.1 手法ごとの平均値

以下の表 2 に，各手法ごとの平均 MSnDCG を示す．表 2 より，平均値のみを見ると提案手法の中では，n=1 とした場合の翻訳行列拡張が最も高い値を示した．しかし，ベースラインとなる擬似適合性フィードバックによるクエリ拡張には大きく差をつけられ，拡張を行わない場合と比べても僅かに下がるという結果であった．このことから，翻訳行列を用いたクエリ拡張では，言語横断検索の性能を向上させることができなかったと考えられる．

表 2　各手法における平均 MSnDCG

	手法	MSnDCG
1	クエリ拡張無し	0.4400
2	翻訳前拡張	0.5031
3	翻訳後拡張	0.5082
4	翻訳行列拡張 (n=1)	0.4347
5	翻訳行列拡張 (n=10)	0.4076
6	翻訳行列拡張 (n=20)	0.3251
7	2 + 翻訳行列拡張	0.1247

翻訳前と翻訳後に擬似適合性フィードバックによるクエリ拡張を行った場合では，ともに拡張無しの場合に比べて言語横断検索の性能が向上していることが分かる．

5.2 翻訳行列を用いたクエリ拡張

翻訳行列を用いた中で最も平均値が高かった n=1 の場合について詳しく見ていく．以下の図 6 と図 7 に，拡張無しの場合の結果と，翻訳行列拡張 (n=1) の結果を比較したグラフを示す．この 2 つのグラフは，2 つの手法の MSnDCG の値の差 (拡張有り - 拡張無し) をトピック毎に算出した結果である．図 6 が NTCIR5 に含まれるトピックの算出結果であり，図 7 が NTCIR6 に含まれるトピックの算出結果である．図 6 及び図 7 を見ると，全体として差が小さいことが分かる．NTCIR5 で 1 つのトピックが 0.7 以上下がったことを除くと，その他のトピックでは差の絶対値が 0.3 未満となっている．また，全体を通して性能が上昇しているわけではないため，表 2 のように平均値として差が出なかったと考えられる．なお，翻訳行列拡張 (n=1) を行ったときに，MSnDCG の値が拡張無しと比べて下がったトピックは，全 92 トピック中 52 トピックであった．

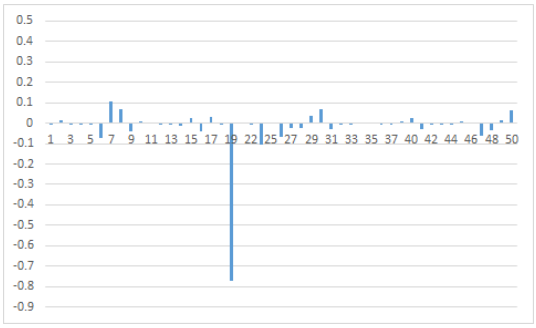

図 6　拡張無しと翻訳行列拡張 (n=1) の比較 (NTCIR5)

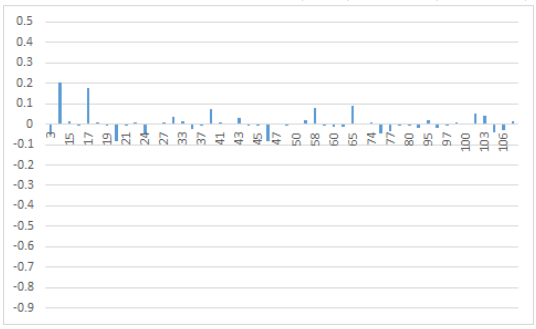

図 7　拡張無しと翻訳行列拡張 (n=1) の比較 (NTCIR6)

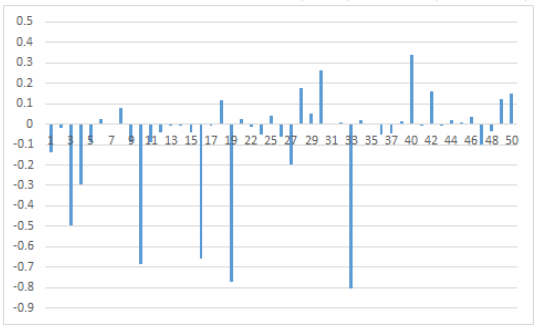

図 8　拡張無しと翻訳行列拡張 (n=10) の比較 (NTCIR5)

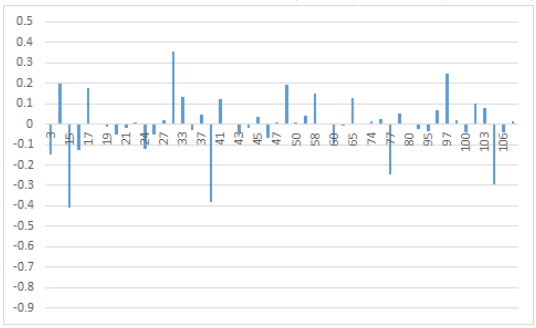

図 9　拡張無しと翻訳行列拡張 (n=10) の比較 (NTCIR6)

次に，n=10 の場合の結果について示す．図 6 と同様に，拡張無しの結果と n=10 の場合の結果をトピック毎に比較したグラフを図 8 と図 9 に示す．図 8 及び図 9 より，拡張クエリタームとして扱う語を 10 単語まで増やした場合，拡張無しと比べたときの変化の幅が大きくなっていることがわかる．

5.3 擬似適合性フィードバックを用いたクエリ拡張

まず，翻訳前拡張の結果を詳しく示す．以下の図 10 と図 11 に，拡張無しの場合と翻訳前拡張を行った場合の比較を，グラフにして示す．この 2 つの図より，翻訳前の拡張では多くのトピッ

クで MSnDCG の値が上昇していることが分かる．MSnDCG の値が上昇したトピックの数は 97 トピック中 75 トピックとなり，実際に約 77%のトピックで値が伸びていた．また今回用いたトピックでは，特定のトピックにおける MSnDCG の上昇具合が特に高いことが見られた．

次に，図 10・図 11 と同様に，翻訳後拡張と拡張無しの場合の比較を図 12 と図 13 に示す．図 12 及び図 13 より，翻訳後拡張においても多くのトピックにおいて MSnDCG の値が上昇していることが分かる．実際に値が上昇したトピックは，97 トピック中 82 トピックであり，約 85%のトピックで値が上昇していた．また，上昇した 82 トピックのうち 28 トピックが 0.1 以上上昇していることから，翻訳前拡張よりも安定して MSnDCG の値が上昇していることがわかる．

5.4 ランダム化 Tukey HSD 検定

本論文では，実験結果の統計的有意性を確かめるために，ランダム化 Tukey HSD 検定 [18] によって検定を行った．全システム中任意の 2 システムの p 値を以下の表 3 に示す．なお，翻訳前+翻訳行列拡張は，他のどのシステムとの対でも p 値が 0 であったため割愛する．

表 3　任意の 2 システムの p 値

	拡張無し	翻訳前	翻訳後	n=1	n=10	n=20
拡張無し	-	0.2518	0.1706	1	0.9028	0.0004
翻訳前	-	-	1	0.1668	0.0104	0
翻訳後	-	-	-	0.1084	0.0058	0
n=1	-	-	-	-	0.9584	0.0014
n=10	-	-	-	-	-	0.0448
n=20	-	-	-	-	-	-

表 2 より，提案手法の n=1，10 の場合では拡張無しと比べて大きな差が無かったが，これらの p 値を確認してみると有意水準 5%において有意な差が得られていないことが分かる．このことから，この 2 つの手法では拡張無しと比べて有意な変化を得られないと考えられる．また，n=20 の場合では，ベースラインと提案手法を含めてどの手法との組み合わせでも有意差が確認できる．ただし，MSnDCG の平均値ではどの手法も下回っているため，この手法では有意な性能低下を招いてしまっていると考えられる．

5.5 トピックごとの分析

前節まででは，全トピックの平均値の観点から結果を示したが，本節では各手法の中で特徴的な結果を残したトピックに関して結果を示す．

NTCIR-5 : 019 ‘supersonic airliner, Concord, airplane crash’

このトピックでは，翻訳行列拡張 (n=1, 10) の両手法で MSnDCG 値が拡張無しを，他のトピックと比べても大きく下回った．元クエリの翻訳結果は「超音速 旅客機 コンコルド 飛行機 クラッシュ」となっているが，翻訳行列拡張の結果は「モーターグライダー 暴走事故 飛行速度 旅客機 マイル (21km)」となった．n=10 の場合においても以上の単語に加えて「ジェット機 ダグラス DC-3」等の関連する単語が拡張されてはいるが，

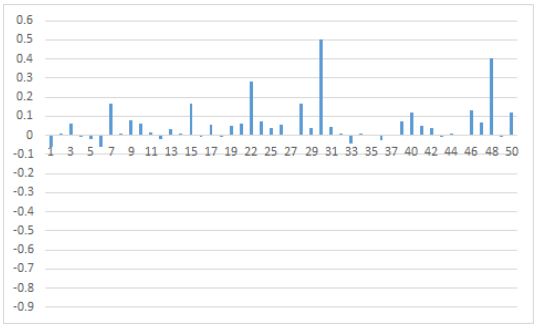

図 10　拡張無しと翻訳前拡張の比較 (NTCIR5)

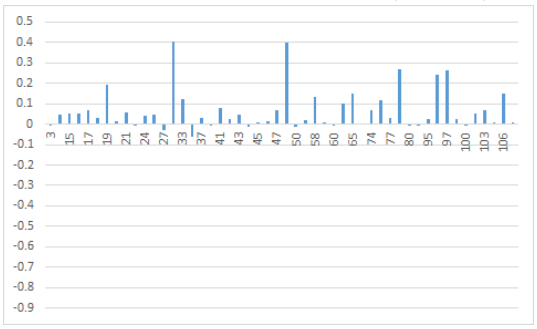

図 11　拡張無しと翻訳前拡張の比較 (NTCIR6)

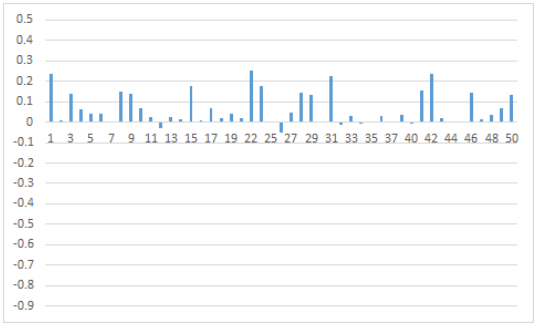

図 12　拡張無しと翻訳後拡張の比較 (NTCIR5)

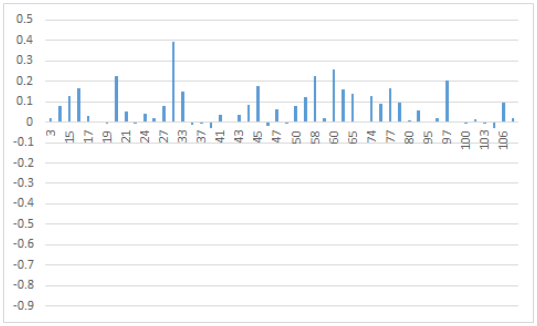

図 13　拡張無しと翻訳後拡張の比較 (NTCIR6)

翻訳結果を向上させる結果には至らなかった．これは，元クエリの意味よりもより一般的な「飛行機」の意味に引っ張られてしまい，結果として元クエリの意味が薄まってしまったことが原因であると考えられる．

NTCIR-5 : 030 ‘space tourism, international space station, Dennis Tito’

このトピックでは，翻訳前拡張により MSnDCG 値が高い値を示し，拡張無しと比べると 3 倍近い差が付いた．翻訳前拡張の結果を翻訳すると「ソユーズ 宇宙飛行 商業の 軌道 宇宙船 ISS 航空宇宙 ミッション 米航空宇宙局 (NASA)」となった．

翻訳前拡張により，国際宇宙ステーションの略称や，デニス・チトー氏が搭乗した「ソユーズ」ロケット [1] 等の単語により，MSnDCG 値が大きく向上したと考えられる．

NTCIR-6：014 ‘Environmental Hormone’

このトピックでは，翻訳行列拡張 (n=1，10) の両手法でも MSnDCG 値が拡張無しの値を上回った．元クエリの翻訳結果は「環境の ホルモン」となるが，n=1 の場合の拡張クエリタームは「胸腺 野生動物保護」となる．この場合，元クエリの意味が抽象的であるため,「野生動物保護」のようなクエリタームによって検索結果が向上したと考えられる．ただし，n=10 の場合では「腫瘍化 胎児期」といったより具体的なクエリタームも得られるが「地球温暖化問題 環境保護」等の抽象的なクエリタームも増えていた．このため，n=1 と n=10 の場合の差はほとんど無いという結果になった．

NTCIR-6：030 ‘Animal Cloning Technique’

このトピックでは，翻訳行列拡張 (n=1, 10) 及び，翻訳後拡張で高い MSnDCG 値が確認できた．Google 翻訳による翻訳結果では「動物 クローニング 技術」となる．これによる翻訳後拡張では「クローン ヒト 応用 卵子 細胞 移植 受精卵 実験 核 妊娠」という拡張クエリタームが得られ，クローン技術と関わる単語が複数得られているため，検索結果の向上に繋がったと考えられる．

翻訳行列拡張では，n=1 の場合「動物 フィールド単位 遺伝子操作」というクエリタームが得られ，拡張無しの場合と比べて MSnDCG が 0.0374 のみ上昇している．これらの単語に加えて，n=10 の場合では「細胞融合 ヒトゲノム計画」等のクエリタームが得られたため，MSnDCG 値が 0.3569 上昇している．

6. 結　　論

6.1 翻訳行列をクエリ拡張に用いることの有効性

本節では，翻訳行列によるクエリ拡張について論じる．表 2，表 3 より，翻訳行列を用いたクエリ拡張では，拡張無しと比べて有意に性能を向上させることができなかった．既存のクエリ拡張手法である擬似適合性フィードバックを用いたクエリ拡張と比べても，MSnDCG の平均値で 10%以上の差が付いた．

図 6 と図 7 及び図 8 と 9 を見てみると，翻訳行列拡張でクエリタームとして扱う単語を増やすと，拡張無しと比べたときの振れ幅が広がることが分かる．ただし，全体として拡張無しよりも結果が悪くなるトピックが多いため，単語を増やすごとに平均値が下がってまっている．よって，翻訳行列を有効にクエリ拡張に利用するためには，翻訳行列が有効である場面を判断して利用する必要があると考えられる．翻訳行列による拡張が成功している場面としては，元クエリが具体的ではない場合やクエリターム数が少ない場合があげられる．ただし，今回の実験の結果では成功事例自体が少ないため，より多くのデータを用いた評価を行う必要がある．

6.2 擬似適合性フィードバックによるクエリ拡張

本節では，擬似適合性フィードバックによるクエリ拡張について論じる．表 2 より，擬似適合性フィードバックによって，翻訳前後双方の MSnDCG の平均値が 10%以上向上していることが分かる．この結果は，擬似適合性フィードバックによるクエリ拡張の有効性を示しているが，表 3 のようにどちらの手法でも有意水準 5%において有意差は無かった．

6.3 翻訳行列を用いた言語横断検索

筆者らの従来の研究では [21]，2 つの言語ベクトル空間にまたがる翻訳行列をクエリ翻訳に用いていたが，既存の機械翻訳ツールである Google 翻訳に比べて高い性能を出すことができなかった．そこで本研究では，異なるアプローチとして，翻訳行列によるクエリ拡張で，既存の手法を用いたシステムの性能を向上させることを提案した．しかし，表 2 より，元のクエリと翻訳前拡張クエリタームを同時に用いて翻訳行列拡張を行ったシステムでは，大きく性能が低下してしまうことが分かった．

言語の分散表現の類似性を用いた言語横断検索の今後は，翻訳行列による言語間のマッチングの精度自体を向上させるアプローチと，既存の翻訳手法やクエリ拡張手法との組み合わせ方でシステム全体の性能を向上させるアプローチが考えられる．

文　　献

[1] BBC News — SCI/TECH — Profile: Tito the spaceman. http://news.bbc.co.uk/2/hi/science/nature/1297924.stm.

[2] Google Code Archive - Long-term storage for Google Code Project Hosting. http://code.google.com/p/word2vec.

[3] Indri. http://sourceforge.net/p/lemur/wiki/Home/.

[4] Wikipedia 記事コーパス. http://dumps.wikimedia.org/.

[5] ビジネス向けツール Google 翻訳. https://translate.google.co.jp/about/intl/ja_ALL/forbusiness.html.

[6] Rony Attar and Aviezri S Fraenkel. Local feedback in full-text retrieval systems. *Journal of the ACM (JACM)*, Vol. 24, No. 3, pp. 397–417, 1977.

[7] Lisa Ballesteros and W Bruce Croft. Phrasal translation and query expansion techniques for cross-language information retrieval. In *ACM SIGIR Forum*, Vol. 31, pp. 84–91. ACM, 1997.

[8] Chris Burges, Tal Shaked, Erin Renshaw, Ari Lazier, Matt Deeds, Nicole Hamilton, and Greg Hullender. Learning to rank using gradient descent. In *Proceedings of the 22nd international conference on Machine learning*, pp. 89–96. ACM, 2005.

[9] William A Gardner. Learning characteristics of stochastic-gradient-descent algorithms: A general study, analysis, and critique. *Signal Processing*, Vol. 6, No. 2, pp. 113–133, 1984.

[10] David A Hull and Gregory Grefenstette. Querying across languages: a dictionary-based approach to multilingual information retrieval. In *Proceedings of the 19th annual international ACM SIGIR conference on Research and development in information retrieval*, pp. 49–57. ACM, 1996.

[11] Kazuaki Kishida, Kuang-hua Chen, Sukhoon Lee, Kazuko Kuriyama, Noriko Kando, and Myaeng Sung Hyon Chen, Hsin-Hsi. Overview of CLIR task at the fifth NTCIR workshop. 2005.

[12] Tomas Mikolov, Quoc V Le, and Ilya Sutskever. Exploiting similarities among languages for machine translation. *arXiv preprint arXiv:1309.4168*, 2013.

[13] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S Corrado, and Jeff Dean. Distributed representations of words and phrases and their compositionality. In *Advances in neural information processing systems*, pp. 3111–3119, 2013.

[14] Teruko Mitamura, Eric Nyberg, Hideki Shima, Tsuneaki Kato, Tatsunori Mori, Chin-Yew Lin, Ruihua Song, Chuan-Jie Lin, Tetsuya Sakai, Donghong Ji, et al. Overview of

the NTCIR-7 ACLIA tasks: Advanced cross-lingual information access. In *Proceedings of NTCIR-7*, 2008.

[15] Jian-Yun Nie. *Cross-Language Information Retrieval*. Morgan Claypool, 2010.

[16] Stephen E Robertson and Karen Sparck Jones. Simple, proven approaches to text retrieval. Technical Report UCAM-CL-TR-356, University of Cambridge, 1994.

[17] Tetsuya Sakai. NTCIREVAL: A generic toolkit for information access evaluation. In *Proceedings of the forum on information technology*, Vol. 2, pp. 23–30. Citeseer, 2011.

[18] Tetsuya Sakai. Metrics, statistics, tests. In *Bridging Between Information Retrieval and Databases*, pp. 116–163. Springer, 2014.

[19] Tetsuya Sakai, Makoto Koyama, Tatsuya Izuha, Akira Kumano, Toshihiko Manabe, and Tomoharu Kokubu. Toshiba bridje at NTCIR-6 CLIR: The head/lead method and graded relevance feedback. In *Proceedings of NTCIR-6*, pp. 36–43, 2007.

[20] Jinxi Xu and W Bruce Croft. Query expansion using local and global document analysis. In *Proceedings of the 19th annual international ACM SIGIR conference on Research and development in information retrieval*, pp. 4–11. ACM, 1996.

[21] 玉置賢太, 林佑明, 酒井哲也. 言語の分散表現と擬似適合性フィードバックを用いた英日言語横断検索. *DEIM Forum 2016*, 2016.

[22] 酒井哲也. 情報アクセス評価方法論 検索エンジンの進歩のために. コロナ社, 2015.

[23] 林佑明, 酒井哲也. 言語の分散表現による文脈情報を利用した言語横断情報検索. *DEIM Forum 2015*, 2015.