

ハイブリッド機械翻訳技術による 日中英音声翻訳システム

Japanese/ Chinese/ English Hybrid Speech Translation System

知野 哲朗 釜谷 聡史

■ CHINO Tetsuro

■ KAMATANI Satoshi

音声入力での自動通訳は大きく拡大する音声翻訳の新市場であるが、まだユーザーの満足できる性能が実現されていない。

東芝は、“言語の壁”を解消して、自由な異言語コミュニケーションを可能にするために、任意の文を正確かつ自然に翻訳できるハイブリッド機械翻訳技術を開発した。日本語、中国語、及び英語の間の相互翻訳が可能で、日本、中国、及びオーストラリアの3か国で実施したフィールドテストによって、異言語話者とのコミュニケーションで、旅行における典型的な対話の目的の約70%を、2分程度の時間で達成できることを確認した。

Toshiba has proposed a new hybrid machine translation (MT) method to overcome the language barrier in cross-linguistic communication. The proposed method utilizes both of two complementary methods of MT; namely, the example-based MT (EBMT) method that can produce natural translations within restricted domains, and the rule-based MT (RBMT) method that produces relatively halting translation with wide coverage.

We have developed an experimental hybrid speech translation system for Japanese, Chinese, and English, and confirmed a task achievement rate of about 70% within about two minutes in typical tasks in travel situations through field tests conducted in Japan, China, and Australia.

1 まえがき

音声入力による自動通訳は、人類にとって長年の夢であった。近年の技術進歩によって、それははるか遠い夢ではなくなり、既に携帯電話などを利用した商用サービスが開始されているが、現実はまだユーザーを満足させる性能は実現できていない。将来大きく開花する可能性を持つ音声翻訳の新市場を現実のものとするためには、ユーザーの真のニーズを満たす“実用的な音声翻訳”の実現が必要である。

東芝は、任意の文を正確かつ自然に翻訳することが可能なハイブリッド機械翻訳 (HybridMT) 技術を開発した。ここでは、この技術の概要と特長について述べる。

2 日中英ハイブリッド音声翻訳システムの概要

実用的な音声翻訳を実現するためには、“任意の表現を翻訳”と“自然な表現で翻訳”の二つの要求を同時に満たす必要がある。まず前者に関し、もし入力できる表現に制約が課せられると、ことばの最大の利点である“無限の拡張性”を生かすことができなくなり、自然言語を利用する必然性が失われてしまう。そのため、自然言語を受け付けるアプリケーションでは、任意の文を扱えることが非常に重要である。一方後者は、細かなニュアンスを正確に伝えることも、コミュニケーションの円滑さをもたらすと考えられるため、同様に重要である。

現在利用可能な翻訳技術は、次の二つに大別することができる。

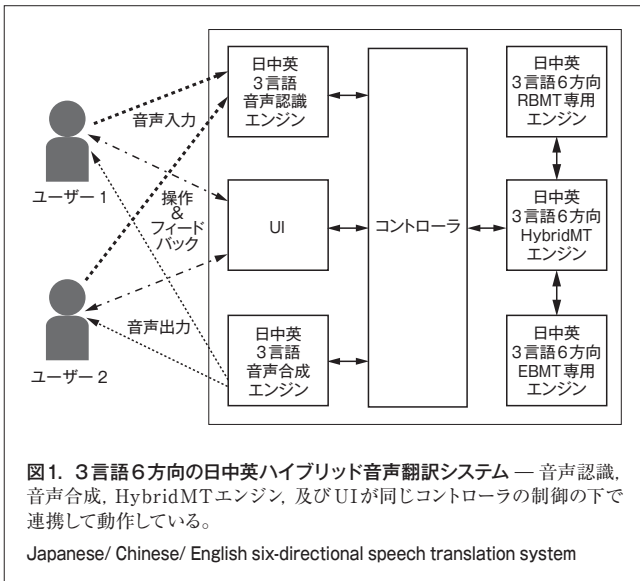
- (1) 規則に基づく翻訳技術
RBMT (Rule-based Machine Translation)
- (2) 用例に基づく翻訳技術
EBMT (Example-based Machine Translation)
SMT (Statistical Machine Translation)
TM (Translation Memory)

これら二つの技術は相補的な性質を持っている。すなわち、RBMTは、文法規則や辞書情報に基づいて任意の文を翻訳できるため、前述の“任意の表現を翻訳”の解決に有効であるが、一般化あるいは抽象化された規則に基づいて原文の解析と訳文の生成を行うために、その翻訳結果は流暢(りゅうちょう)な訳文にならず、“自然な表現で翻訳”の解決には向かない。一方、EBMTでは、各言語間での意味的に等価な実際の表現例の対から成る用例に基づいて、訳文が生成される。そのため、正確で自然な訳文を得やすく、“自然な表現で翻訳”の解決を得意とするが、あらかじめ用意されている対訳と大きく異なる入力文を正しく翻訳できない。そのために、適用可能な範囲が限定されるという欠点があり、“任意の表現を翻訳”の解決には向かない。

今回の開発では、前述の二つの要求を同時に満たすために、RBMTとEBMTの技術を同時に利用して、両者の長所を引き出すHybridMT技術を採用ことにした。

試作した3言語6方向の日中英ハイブリッド音声翻訳システムの構成を図1に示す。

このシステムでは、3言語の音声認識を行うエンジン、同じく3言語の音声合成を行うエンジン、3言語6方向のHybridMT



を行うエンジン、更にUI (User Interface) が、コントローラの制御の下で連携して動作している。ここで、音声認識エンジンは、HMM (Hidden Markov Model) 方式^(注1)による多言語・大語彙 (ごい)・連続音声認識エンジンであり、一方、音声合成エンジンは、複数素片選択融合方式⁽⁴⁾による多言語 Text-To-Speech エンジンで実現されている。

3 HybridMTエンジン

3.1 RBMT専用エンジン

音声翻訳のための規則に基づく翻訳では、話しことばを適切に解析し、翻訳、生成する必要がある。しかし、話しことばは、従来の翻訳技術の対象である書きことばと比較して、より“断片的”であり、“非文法性”が高く、“文脈並びに知識依存性”が高いといった特徴がある。更に、音声翻訳への応用では、“音声認識誤り”の発生を避けることが困難で、翻訳処理の対象となる入力から最初から誤りを含んでいる可能性が高い。これら四つはすべて、翻訳処理をより困難にする要因であり、その結果は翻訳処理の各フェーズにおけるあいまい性の増大を招き、従来の技術では現実的な時間内に適切な翻訳結果を得ることが難しかった。

このあいまい性の増大という問題を解決するために、複数の統語構造候補を一つの構造で表現できる圧縮共有統語森構造を利用した、統語森駆動・規則ベース翻訳技術を開発し、これによってRBMT専用エンジンを実現した。

3.2 EBMT専用エンジン

このシステムのEBMT専用エンジンは、TM技術と、文単

(注1) 不確定な時系列データをモデル化するための統計的手法。
(注2) 文の中の意味的なまとまりを意味し、日本語では“文節”などがチャンクに相当する。

位及び部分文単位の照合処理に加えて、穴あき定型文によって翻訳を行うパターンベース機械翻訳技術を組み合わせ、用例ベース翻訳技術を用いて実現されている。ここで行われる用例ベース翻訳処理では、統計ベースの文分割処理が用いられており、EBMT結果候補には“確信度スコア”が付与される。確信度スコアは、“翻訳対象となる原言語文と用例対の原言語文の類似度”及び、“翻訳結果である目的言語における訳文の単語の並びの統計的なもっもらしさ”の重み付けの和として算出される。

更に、言語ごとに個別に実施されるチャンキング処理^(注2)の結果を基本単位として、この用例ベース翻訳エンジンでの入力と用例の間の照合処理や、訳文生成における置換処理を動作させることによって、高精度化を図っている。

3.3 統合型HybridMTエンジン

HybridMTエンジンは、前述のRBMT専用エンジンとEBMT専用エンジンを併用して、話しことばの翻訳を行う。処理フローを図2に示す。ここでは、例を用いて説明する。

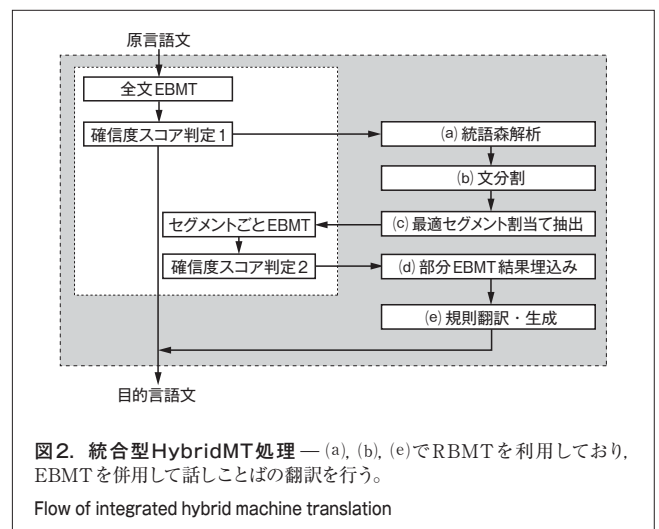
まず、下記の日本語入力が正しく認識され、統合型Hybrid-MTエンジンにもたらされたとする。

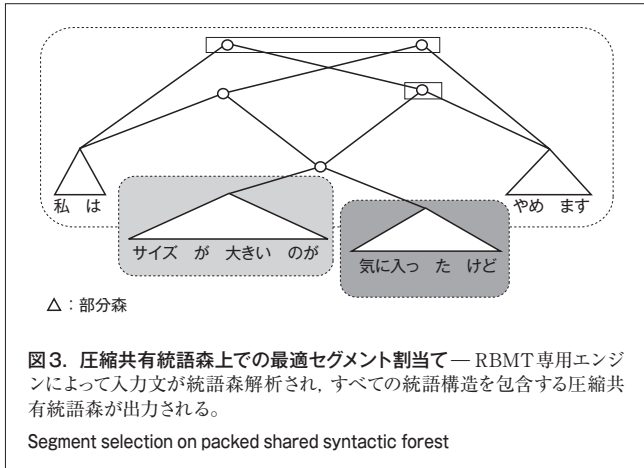
「私はサイズが大きいのが気に入ったけどやめます」

この文は、最初EBMT専用エンジンによって全文に対する用例ベース翻訳が試みられ、結果の確信度スコアに応じて結果の採否が判定される。ここでは、十分な確信度スコアを持つ全文EBMT結果が得られなかったことにする。その結果、図2(a)でRBMT専用エンジンによって入力文が統語森解析され、すべての統語構造を包含する圧縮共有統語森が出力される(図3)。

更に、図2(b)で翻訳の単位となる“セグメント”候補が抽出され、入力文は、以下の三つのセグメントに分解される。

「サイズが大きいのが / 気に入ったけど / 私はやめます」





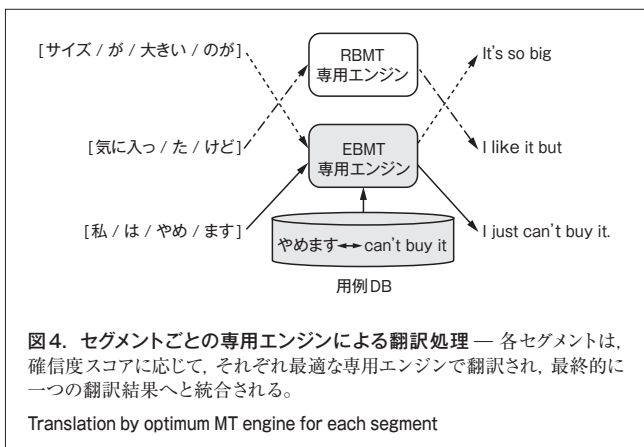
最後のセグメントで、元の文で遠く離れた位置にある“私は”と“やめます”とが、正しく一つのセグメントへとまとまっているのは、統語森構造に基づくこの技術の効果である。

この三つのセグメントは、以降の処理によって、それぞれRBMT専用エンジン又はEBMT専用エンジンで翻訳される。各々のセグメントは、確信度スコアに応じて、それぞれ最適な専用エンジンで翻訳され、最終的に一つの翻訳結果へと統合される(図4)。

以上の処理によって、RBMTとEBMTの2種類の専用エンジンを併用した以下の翻訳結果が得られる。

「It's so big. I like it, but I just can't buy it.」

この例では、2番目のセグメントだけがRBMT専用エンジンで翻訳され、最初と最後のセグメントはEBMT専用エンジンによって翻訳がなされた。特に最後のセグメントに関しては、“やめる”という日本語が、旅行会話に特化した用例データベース(DB)の働きによって、“買うことをやめる”の意味に意識されている。この効果は、用例DBを利用環境に合わせて適宜切り替えることによって、より自然で適切な翻訳の実現に活用できる。



4 評価実験

日本語、中国語、及び英語の3言語間の相互音声翻訳処理に関して、EBMT専用エンジン、RBMT専用エンジン、及びHybridMTエンジンを評価対象翻訳システムとして評価実験を実施した。

4.1 評価方法

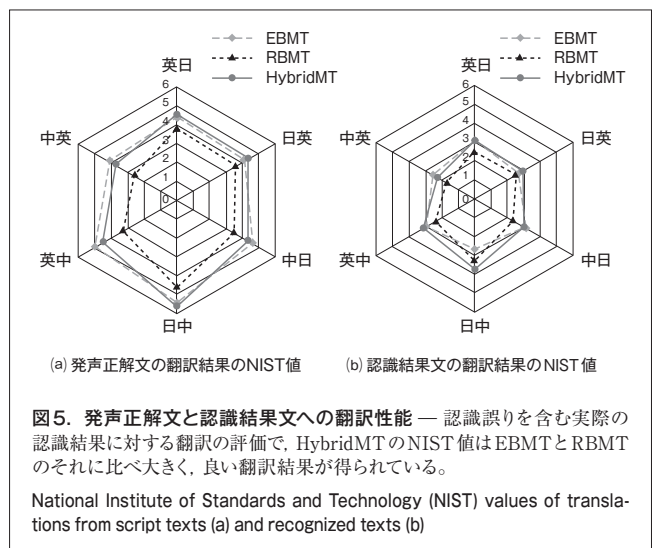
この実験では、この領域で広く使われている自動評価手法であるNIST法^(注3)を用いた。この方法では、大量の入力文に対し、それぞれの模範訳をあらかじめ用意し、各翻訳システムからの翻訳結果と、模範訳の表層上の統計的な類似性を算出することで、各翻訳結果のよしあしが数値で算出されるようになっており、NIST値が大きいほど良い評価であることを表す。

4.2 評価データ

この実験では、音声翻訳システムが実際に利用されることを想定し、ホテル、レストラン、駅という音響環境が互いに異なる三つの典型的な実環境において、日中英の三つの言語ごとに複数のネイティブ話者である被験者を用いてデータを収集した。

4.3 発声正解文への翻訳性能

全6翻訳方向の、発声正解文セットの翻訳結果に対する各翻訳エンジンのNIST値を図5(a)に示す。ここでは、英日、日英、日中の三つの翻訳方向で、HybridMTがもっとも高い評価値を得ており、提案方式の有効性が確認された。一方、全6方向平均では、HybridMTエンジンは、EBMT専用エンジンにわずかに及ばなかったもののRBMT専用エンジンは大きく引き離しており、このハイブリッド方式によって、この実験で優勢であったEBMT専用エンジンの能力の大部分を引き出せていたことが確認できた。



(注3) National Institute of Standards and Technology (米国標準技術局)が提案した機械翻訳の自動評価法。

4.4 認識結果文への翻訳性能

図5(b)は、全6翻訳方向の、音声認識結果文セットの翻訳結果に対する各翻訳エンジンのNIST値を示している。ここでも、日英、日中、英中の三つの翻訳方向でハイブリッド機械翻訳がもっとも高い評価値を得た。更に、特筆すべきこととして、六つの全翻訳方向の平均において提案方式がもっとも高い評価値を得ている点が挙げられる。図5(a)は、音声認識誤りがまったく生じない条件での結果であるが、図5(b)は、認識誤りを含む実際の認識結果に対する翻訳の評価であり、実際の利用条件に即した評価である。つまり、全6方向の平均で、HybridMTがEBMT専用エンジン及びRBMT専用エンジンを大きく上回る評価値を得たことは、開発した技術の優位性と実効性を示すものであると言える。

5 国内外でのフィールドテスト

音声翻訳システムは音声認識技術を用いているため、使用される場所の環境雑音や利用者の声の調子などによって、その性能が大きく左右される。このため、実用化にあたっては、使用が想定される場所や環境下での評価実験や改良を繰り返し実施し、性能改善を行っていくことが必須となる。

そこで、小型パソコンで動作するシステム(図6)を用いて、2008年4月～9月の間に、国内、中国、及びオーストラリアの3か国4都市で、合計6回のフィールドテストを実施した。

ここでは、計130余人の日本語、中国語、英語の各ネイティブ話者である被験者を用い、室内(会議室、ホテル、レストラン、店舗、地下広場など)と屋外(公園、路上、駅など)の計15種類の環境の下で、4種類の実験を行い、合計1万発話以上のデータを収集した。このデータの分析結果から、現在のシステムを用いた異言語話者とのコミュニケーションで、旅行における典型的な対話の目的の約70%を、2分程度の時間で達成

できることを確認した。

6 あとがき

ここでは、当社の研究開発センター、中国研究センター、及び欧州研究所が共同開発した日中英3言語6方向のハイブリッド音声翻訳システムの概要と特長、及び評価実験結果について述べた。更に、日本、中国、オーストラリア3か国で実施したフィールドテストの概要と得られた知見も合わせて述べた。

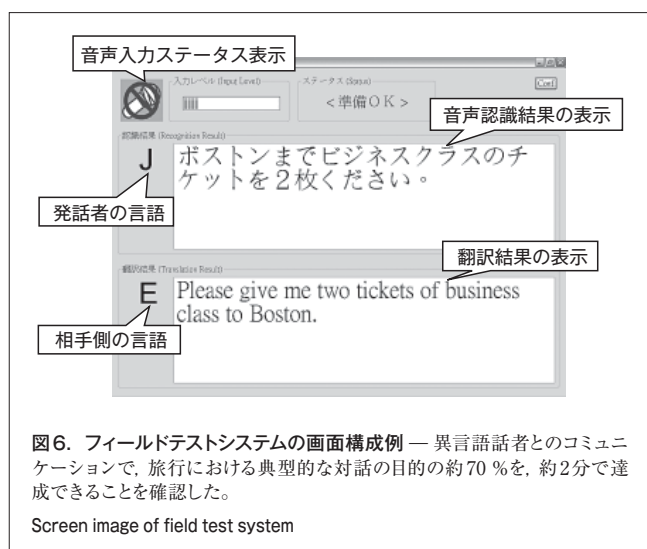
今後、フィールドテストで得られた知見に基づいて改良を施すことで、実際の利用場面で本当に有益な音声翻訳システムの実現を目指す。

謝辞

中国でのフィールドテスト実施に際し、多大な協力を頂いた北京言語文化大学(BLCU)の関係各位に感謝の意を表します。

文献

- (1) Chino, T., et al. Partial forest transfer for spoken language translation. Proc. of RANLP (Recent Advances in Natural Language Processing) '05. Borovets, Bulgaria, 2005-09, p.157 - 161.
- (2) Wu, H., et al. Improving translation memory with word alignment information. Proc. of MT Summit X. Phuket, Thailand, 2005-09, p.364 - 371.
- (3) Kamatani, S., et al. Forest driven dependency analysis enhanced by Japanese clause structure estimation. Proc. of The 20th Pacific Asia Conference on Language, Information and Computation, Wuhan, China, 2006-11, p.265 - 273.
- (4) Mizutani, T., et al. Concatenative speech synthesis based on the plural unit selection and fusion method. IEICE Trans. E88-D, 11, 2005, p.2565 - 2572.
- (5) 知野哲朗, ほか. 日中英3言語6方向音声翻訳システム. 情報処理学会第185回自然言語処理研究会+第71回音声言語処理研究会. NL-185+SLP-71, 横浜, 2008, p.15 - 22.
- (6) 釜谷聡史, ほか. 統語森係り受け解析による最尤構文木の推定. 情報処理学会第163回自然言語処理研究会. NL-163, 東京, 2004, p.99 - 106.



知野 哲朗 CHINO Tetsuro

研究開発センター 知識メディアラボラトリー主任研究員。
自然言語処理、ヒューマンインタフェースの研究・開発に従事。情報処理学会会員。
Knowledge Media Lab.



釜谷 聡史 KAMATANI Satoshi

研究開発センター 知識メディアラボラトリー。
自然言語処理、機械翻訳の研究・開発に従事。情報処理学会、言語処理学会会員。
Knowledge Media Lab.