

携帯電話試作機上で動作する旅行会話向け 自動通訳システムの開発

長田 誠也 花沢 健 磯谷 亮輔 山端 潔 奥村 明俊

NEC

1. はじめに

我々は、常に携帯していつでもどこでも手軽に使える自動通訳システムの実現を目指してきた[1][2]。自動通訳システムとは、例えばユーザが日本語を発声するとそれを英語に翻訳するシステムを指し、今回は市販されている携帯電話と同等のスペックを持つ試作機上で動作する日本語から英語と英語から日本語への旅行会話向けの自動通訳システムを開発したのでこれを報告する。

2. システムの構成

今回開発した自動通訳システムは、図 1 に示すように、日本語・英語の大語彙連続音声認識エンジン、日英・英日機械翻訳エンジンを有し、これらを通訳統合部が連携・制御している。

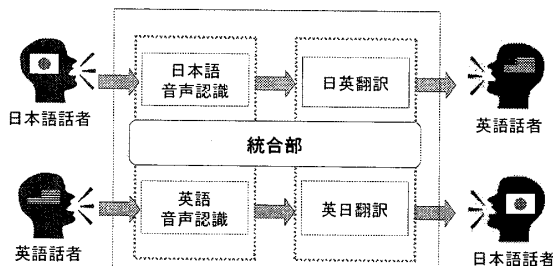


図 1：自動通訳システムの構成

英語話者が英日自動通訳システムに向けて発声すると、英語音声認識エンジンがこの発声内容を音声認識し、英日機械翻訳エンジンが音声認識結果を受け取り日本語へ翻訳する。この音声認識結果と翻訳結果を図 2 のように画面に表示する。日本語から英語方向も同様である。

このシステムをコンパクトで高性能なものにすることで、サーバと通信することなく、市販されている携帯電話と同等のスペックを持つ試

作機単体で動作するものになっている。

以下で、携帯電話試作機上で動作する音声認識エンジン、翻訳エンジンについて説明する。

2.1 音声認識エンジン

音声認識エンジンは、コンパクトでスケラブルな音声認識エンジン[3]を使用しており、これにより日本語の音声、英語の音声を認識している。本エンジンは統計的言語モデルを用いる大語彙連続音声認識エンジンとなっており、エンジン自体は言語非依存で、モデルや辞書などのリソースを切り替えることで言語の切り替えが可能になっている。

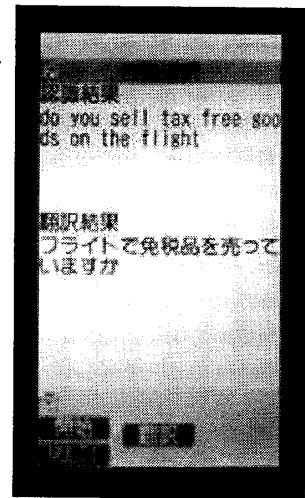


図 2：自動通訳画面例

携帯電話試作機に搭載するにあたって、各計算処理の固定小数点化、コンパクトな環境での精度劣化低減、言語モデルなどの大きなリソースのプログラムへの埋め込みなどを行なった[4]。

シミュレーションでの音声認識精度は、単語正解精度で日本語が 94%、英語が 92%となり、従来のシステム[1][2]と比較して同等の結果となった。また使用メモリ量は日本語認識、英語認識ともにワークメモリ込みで 10MB 程度となった[4]。

2.2 機械翻訳エンジン

機械翻訳エンジンは、語彙規則型のコンパクト機械翻訳エンジン[5]を使用することで、旅行会話の日本語から英語、英語から日本語の機械翻訳を行なっている。本エンジンは、文法知識が単語辞書中に局所化されているため、翻訳対象文に必要な文法のみをメモリ上に展開するだけでよく、すべての文法をメモリ上に持つ必要がないためコンパクト化における利点がある。また、エンジン自体は言語非依存で、辞書とその語彙が持つ文法、リソースを切り替えること

で言語の切り替えが可能になっている。

携帯電話試作機に搭載するにあたって、音声認識エンジンで使用するモデルを学習するためのコーパスを使って、翻訳用の辞書やリソースを音声認識結果として出現するものにあわせることで、辞書サイズやリソースのサイズを削減した。同時に、これによりエンジンの起動速度の高速化も行うことができた。

翻訳精度は上記コーパスからランダムに 500 文を抽出して表 1 の a~d の 4 段階で主観評価を行なったところ、日英翻訳で b 評価以上 73%、c 評価以上 91%、英日翻訳で b 評価以上 77%、c 評価以上 91%となり、従来のシステム[1][2]と比較して同等の結果となった。また使用メモリ量はワークメモリ込みで日英翻訳、英日翻訳ともに 10MB 程度となった。

表 1：翻訳精度の評価値と評価基準

評価値	評価基準
a	原文と照らし合わせて自然で適切
b	言い回しなどに不自然さはあるが統語的に正しく、原文の意味も正しく理解できる
c	文法的に誤りがあるなど b 評価にはならないが、原文の意味は伝わる
d	原文の意味が理解できない/誤解を生じる

3. オンライン評価

今回開発した自動通訳システムが実際に動作している携帯試作機（動作周波数 500MHz）で評価した。評価方法は、日本語の旅行会話文 20 文を 2 回ずつ計 40 文発声したときの音声認識文正解率と、音声認識結果が正しかったときの翻訳正解率で、日本人 8 人で評価を行なった。

表 2：オンライン評価結果（文正解率）

	計算方法 1	計算方法 2
音声認識率	73.4%	79.4%
翻訳正解率	80.0%	100.0%

音声認識率は、計算方法 1 では発声した文とまったく同じ文が出力されたときの割合、計算方法 2 では助詞などは異なるが意味は変わらない文が出力されたときの割合の 2 種類の結果を表 2 に示している。

また翻訳正解率は、音声認識結果が正しかったときの翻訳結果に対して、表 1 の a~d の 4 段階で主観評価したときに、a 評価と b 評価の文の合計の割合を表 2 の計算方法 1 に示し、a 評価と

b 評価と c 評価の文の合計の割合を表 2 の計算方法 2 に示している。

起動時間は音声認識エンジンと機械翻訳エンジンともにそれぞれ 2 秒程度となった。

発声後から音声認識結果が出力されるまでと、音声認識結果出力後から翻訳結果が出力されるまでの時間はそれぞれ 1 秒以下となっており、体感で待たされている感じはなく結果が出力されることを確認した。

4. おわりに

携帯電話試作機単体で動作する日英方向・英日方向の双方向の旅行会話向け自動通訳システムを開発した。

本システムは、今まで開発を続けてきた日英方向・英日方向の自動通訳機能を、コンパクトかつ高速に動作させることで、携帯電話試作機上に搭載したものである。これにより、今までに開発してきた自動通訳システムと同等の精度を持ち、十分にレスポンスが速いシステムとなっていることが確認できた。また、ミドルウェアとして実装しているため、携帯電話機の各種アプリケーションから呼び出せるようになっており、メーラやブラウザなど他のアプリケーションからも使用することが可能になっている。

今後、実用可能性について更に検討を続けていきたい。

参考文献

- [1] 山端他, “低消費電力マルチコアプロセッサで動作する日英自動通訳システム”, 情処全国大会, 4B-2, 2006.
- [2] 山端他, “PDA で動作する旅行会話向け日英双方向音声翻訳システム”, 情処研報, 2002-NL-150-9, 2002 年 7 月
- [3] 磯谷他, “話し言葉認識に向けた基本技術と応用”, 情処研報, 2005-NL-169, 2005.
- [4] 花沢他, “携帯電話試作機上で動作する旅行会話向け音声認識システム”, 情処全国大会, 2D-3, 2009.
- [5] 山端他, “語彙化されたツリーオートマトンに基づく会話文翻訳システム”, 言語処理学会第 6 回年次大会講演論文集, pp.264-267, 2000 年 3 月.