

## 多言語対応・地域情報システムフレームワークの提案

陳 実<sup>†</sup> 清水 皓生<sup>‡</sup> 山田 敬三<sup>‡</sup> 佐々木 淳<sup>‡</sup>

岩手県立大学大学院 ソフトウェア情報学研究科<sup>†</sup> 岩手県立大学 ソフトウェア情報学部<sup>‡</sup>

### 1. はじめに

近年、日本の各地域への外国人観光客や居住者の急増に伴い、地域 Web サイトの多言語化と、そのサイトを低コストでリアルタイムに更新したいという要望が高まっている。

本研究では、地域におけるサイト運営者自身が操作することを前提とし、操作性に優れ翻訳精度の高い多言語対応・地域情報システムの構築を行うことをねらいとしている。

### 2. 現状の課題

通常、多言語対応の地域 Web サイトを構築する場合、地域における Web サイト運営者から外部サービス会社に発注することが多い。しかし、この方法によると Web サイトの構築にコストや時間がかかり、多言語情報更新が難しくなる。

一方、Web サイトを多言語対応にする場合、Google や Yahoo 等の既存翻訳サービスで、日本語 Web ページの URL を入力し、多言語に自動翻訳することが一般的である。しかし、地域 Web サイトにおいては固有用語が多いので、翻訳精度が低くなってしまふ。また、誤っている翻訳箇所を発見しても修正・編集が難しい。一方、Trados[1]等の翻訳ソフトを使用する場合、高コストであり、契約や更新が面倒等の問題がある。

そのため、本研究では、更新が頻繁な部分を、既存の API を活用するとともに、新たに構築した地域固有用語 DB と翻訳事例 DB と連携し、容易な作業で高い翻訳精度が得られるフレームワークを提案する。

### 3. 提案システム

本提案では、一般的なコンテンツ管理システム CMS (Contents Management System) のように、情報追加、編集、削除等のコンテンツ管理機能に加え、自動翻訳機能も有し、誤っている翻訳箇所を手動で修正・編集できるようにしたことが特長である。

#### 3.1. システム利用の流れ

本システムの構成と処理の流れを図 1 に示し、

A Proposal of Multi-language Regional Information System Framework

Shi CHEN<sup>†</sup>, Kousei SHIMIZU<sup>‡</sup>, Keizo YAMADA<sup>‡</sup>, Jun SASAKI<sup>‡</sup>,

<sup>†</sup>Graduate School of Software and Information Science, Iwate Prefectural University, <sup>‡</sup>Faculty of Software and Information Science, Iwate Prefectural University

図中(1)～(7)について説明する。

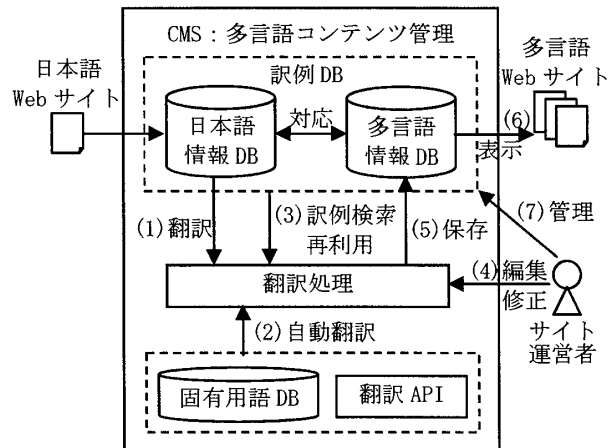


図 1 システム構成及び処理フロー

- (1) サイト運営者は日本語情報 DB から情報を選択し、翻訳を行う。
- (2) 予め構築しておいた特定地域やドメインに関する固有用語（地名やイベント名等）DB と既存翻訳 API を組合せ、自動翻訳を行う。
- (3) 訳例 DB を検索し、翻訳したい文章と一致または類似の訳文があれば、サイト運営者に提示する。訳例とは、過去の翻訳事例の原文と訳文をペアにすることである。日本語情報 DB と多言語情報 DB を組合せ、訳例 DB と共有する。
- (4) サイト運営者は自動翻訳あるいは訳例検索の結果を参照し、誤っている翻訳箇所を発見したら修正・編集して翻訳を行う。
- (5) 修正・編集した訳文を多言語情報 DB に保存する。
- (6) 多言語情報 DB から正しく翻訳された情報を用いて多言語 Web サイトとして表示し、動作を確認して更新が終了する。
- (7) サイト運営者は情報の編集や削除等の管理を行う。

#### 3.2. 主な機能

##### (1) コンテンツ管理機能

一般的な Web ページコンテンツ管理のように、サイト運営者自身によって情報の追加、編集、削除等の管理を行う。

##### (2) 日本語情報翻訳機能

翻訳したい日本語情報を、固有用語 DB と既存の翻訳 API を組合せて自動翻訳、あるいは過去の翻訳事例の訳例 DB から一致または類似度が高い訳例を検索し、翻訳を行う。誤っている翻訳箇所を発見すれば、人手で修正・編集して登録する。最後に、多言語 Web サイトで登録した情報を表示して結果を確認する。

(3) 固有用語登録機能

サイト運営者のみならず、一般ユーザーでも特定地域やドメインに関する語彙を登録して管理できるようにする。

4. プロトタイプシステムの開発

上記の機能を有するプロトタイプシステムを開発し、著者の所属する研究室で構築した地域食材・料理提供システムをケーススタディとして、動作の検証を行った。管理者側の多言語情報登録と一般利用者側多言語 Web ページ更新の画面例を図 2, 図 3 に示す。

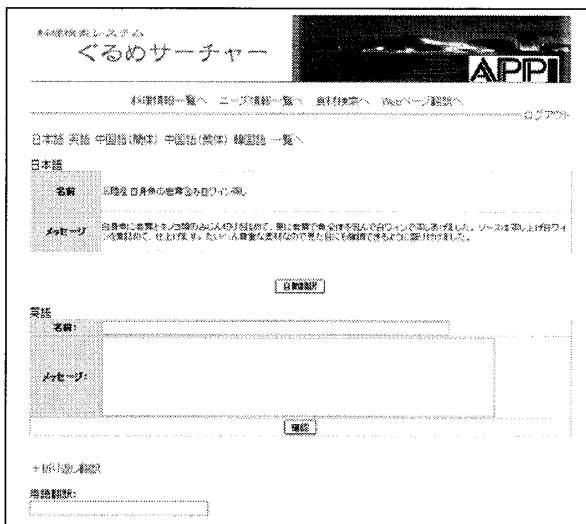


図 2 管理者向け多言語情報登録画面

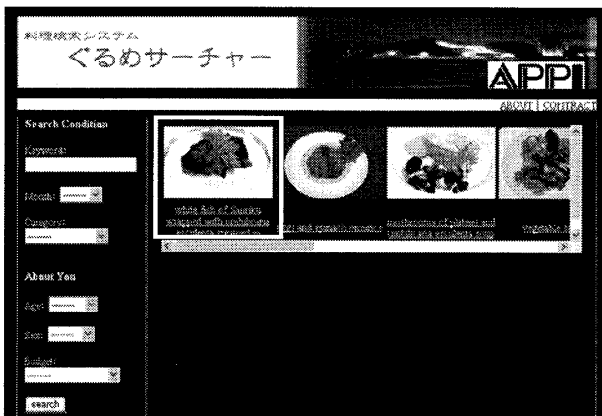


図 3 一般利用者向け多言語 Web ページ

5. 実験評価

プロトタイプシステムを用いて、(1)翻訳精度と(2)システムの操作性・満足度という2つ方面で評価を行った。

(1) 翻訳精度の評価

本研究では、翻訳自動評価尺度 BLEU[2] (範囲: 0~1) によって提案システムの翻訳精度と既存の Google 翻訳精度との比較を行った。その結果を表 1 に示す。

表 1 翻訳精度 BLEU 評価

	英語	中国語 (簡体)	中国語 (繁体)
Google 翻訳	0.3832	0.3329	0.3290
システム翻訳	0.5255	0.4782	0.4950

この実験の結果、提案手法で翻訳した結果は、既存翻訳サイトの代表的な1つである Google 翻訳よりも、精度が高くなることを確認した。

(2) システム操作性・満足度の評価

本システムはホテル安比グランドの職員の方に使っていただき、システム操作性・満足度についてアンケート調査を行った。現在、対象者2名であるが、システムの有効性を確認した。また、「4ヶ国語全部翻訳するようによければ良い」等のシステムに対しての意見が得られた。今後はシステムを改良し、より多くの方に評価していただく予定である。

6. おわりに

本研究では、地域 Web サイト運営者自身可以利用できる操作性および翻訳精度に優れた多言語 Web サイト更新用のフレームワークを提案した。また、プロトタイプシステムを構築し、翻訳精度等についての検証を行った。今後は、広く利用実験を行い、評価及び改良を行っていく予定である。

本研究は観光や教育などのみならず、外国人が必要とする生活情報や商品情報のサイトにも幅広く応用できると考えられる。

終わりに、本研究について貴重な助言をいただきました(株)岩手ホテルアンドリゾートの皆様へ感謝申し上げます。

参考文献

[1] TRADOS, <http://www.trados.com/>.  
 [2] NIST Open Machine Translation, <http://www.nist.gov/speech/tests/mt>.  
 [3] 馬場真知子, 福田豊, “外国人支援から見た地方自治体の Web サイト—多文化共生と ICT,” 日本社会情報学会学会誌, vol.21, no.1, pp.5-17, Sept.2009.