

## 論文審査の結果の要旨

学位申請者

OHNMAR HTUN

本論文は「Development of Cross-Lingual Phonetic Similarity Metrics (言語横断型音声類似度指標の開発)」と題し、7章より構成されている。第1章「Introduction」では、機械翻訳や他言語検索などの分野において何故音声類似度指標の開発が求められているのかなど、本研究の背景、目的について述べ、関連先行研究のレビューを行っている。

第2章「Definition of Basic Terms」では、音声類似度、訳語形成方法の様々な類型、転写規則、SOUNDEX法、編集距離など、本研究で使用する幾つかの基本的な用語や概念についての定義や概念の説明を行っている。

第3章「Methodology」では、音声類似度計算の基礎となる言語間の音声対応表 Cross-Language Sound Grouping (CLSG) 及びこれを用いた音声類似度指標の計算方法について述べる。本研究ではアジアの主要8言語（中国語、韓国語、ミャンマー語、マレー語、タイ語、ベトナム語、インドネシア語）について特に子音文字を中心に音声対応表を作成した。また、本研究で対象とするのは当該言語の文字で書かれた単語であるので、非ラテン文字使用言語の場合にはラテン文字への転写、類似子音のグループ化、国際音標記号による表現、編集距離の計算法、単語の長さを考慮した正規化などの詳細について述べる。

第4章「Experiment of Classical Edit Distance Measures」では、8言語についての92種類の元素名のデータを用いた実験結果について述べ、第3章で提案した音声類似度指標が音訳語の検出をどの程度正確に行えたか評価を行っている。

第5章「Experiment of Stochastic Edit Distance Measures」では、編集距離による判定の過程に確率的なモデルを用いた類似度指標の学習実験結果について述べ、第4章で行った実験との比較を通じた評価結果について述べる。

第6章「Discussion」では、音声対応表 CLSG の有効性に関する評価、学習モデルの有効性に関する評価、第4章、第5章の実験結果についての考察を行っている。

第7章「Conclusion and Future Research」では、以上の結果をまとめて、本研究の達成した成果について述べる。

本研究は、広範囲（語族のベースでは5種類をカバー）にわたる8種類の言語を対象として、音訳語を一定の正確性を持って抽出することのできる言語間音声類似度指標が開発したものであり、その成果は、音訳語の一致などから機械翻訳における対応テキストを抽出する技術 (transliteration mining) 、多言語情報検索などに応用されることが期待される。よって、本論文は工学上及び工業上貢献するところが大きく、博士（工学）の学位論文として十分な価値を有するものと認める。

審査委員主査 三 上 喜 貴