

令和 5 年 5 月 29 日現在

機関番号：33910

研究種目：基盤研究(B) (一般)

研究期間：2018～2021

課題番号：18H01062

研究課題名(和文) 英語講義音声の音声翻訳と音声要約に基づく日本語字幕付き教材の自動生成に関する研究

研究課題名(英文) Automatic generation of lecture's materials with Japanese caption based on English lecture's speech translation and speech summarization

研究代表者

中川 聖一 (Nakagawa, Seiichi)

中部大学・工学部・客員教授

研究者番号：20115893

交付決定額(研究期間全体)：(直接経費) 13,400,000円

研究成果の概要(和文)：本研究では、英語講義・講演音声の音声認識、英日音声翻訳、音声要約の要素技術を開発し、それらを統合した日本人の学習用の字幕表示システムを開発した。本研究の対象として、主にTED Talksとした。

音声認識に関しては、TED英語講演に対して、単語音声認識精度約88%が得られた。音声翻訳に関しては、テキスト入力に対して約15のBLEU値、音声入力に対して約14のBLEU値が得られた。人手による評価では、「まずまず内容が理解でき、意図が伝わる。一部に誤訳がある」という結果が得られた。音声要約に関しては、重要文抽出要約に対して、音声入力からの要約は、テキスト入力からの要約と遜色ない結果が得られた。

研究成果の学術的意義や社会的意義

大学講義のオープンコースウェア等により、手軽に有用な講義・講演音声学習に利用できるようになった。しかし、英語音声のコンテンツを日本人学生が理解するのは困難である。例えば、TOEIC700点程度の学生でも英語講義の正しい聞き取り率は、単語換算で50%程度である。

本研究は英語の講義・講演音声から重要文を抽出し、英語音声と同期して日本語で字幕として表示するシステムを開発した。テキスト入力による重要文抽出や日本語への翻訳精度と比べて、音声入力に対して重要文抽出の精度を維持したまま、翻訳精度の低下は10%程度に抑えることができた。英語音声コンテンツを学習に利用できることを示した社会的意義は大きい。

研究成果の概要(英文)：In this study, we developed fundamental technologies for English speech recognition, English-to-Japanese speech translation, and speech summarization of English lecture audio, and integrated them into a subtitling system for Japanese learners. The main target of this study for various lectures was TED Talks.

For speech recognition, we obtained a word recognition accuracy of approximately 88% for TED English talks. For speech translation, about 15 BLEU values were obtained for text input and about 14 BLEU values for speech input. The human evaluation showed that "the content is understandable at first, and the intent is conveyed. However, some mistranslations were found". As for audio summarization, the results showed that the summary from the audio input was not inferior to the summary from the text input in terms of the summary based on the important sentence selection.

研究分野：知能情報学

キーワード：英日音声翻訳 英日機械翻訳 音声要約 テキスト要約 英語音声の認識 英語講義 英語講演 字幕

1. 研究開始当初の背景

我国の文部科学省のスーパーグローバル事業等により、英語による授業やバイリンガル授業が注目され、実施している大学が増えている。英語がネイティブである教師、もしくは、発音がネイティブに近い日本人教師が英語で講義を行うのが望ましいが、人材確保が難しく、また日本人学生にとっては、日本人の英語よりもネイティブの英語による講義の方が理解が難しくなる。

一方、最近の e-learning の普及とともに、インターネット上に大学の講義映像を公開する動きが活発である。このように、ネット上で良い英語コンテンツの講義映像が公開されている。これらを授業の一部に取り入れるのは、グローバル人材を育成するには有用であると考えられる。

一方、一般の日本人学生が英語の講義を理解するのは困難であることを忘れてはならない。我々の以前の評価実験では、TOEIC 400 点から 750 点の 15 人の学生を被験者として MIT のプログラミング科目の講義映像を視聴してもらい、2 回の視聴の繰り返しと、任意の時点での映像停止・再生を認めた条件での英語の正しい聞き取り率(書き起こし率)は、単語換算で 40% ~ 65% 程度であった(TOEIC の平均 582 点、単語聞き取り率の平均 55%; 全国の日本人学生の TOEIC の平均は 500 点程度)。この書き起こし文を 1 文当たり 2 分 30 秒以内の制限で翻訳してもらったところ、翻訳の尺度である BLEU 値(正解翻訳文の単語の並び方と翻訳結果の単語の並び方の類似性を求める尺度)で 3~8 程度の惨憺たる結果であった(平均 4.7)。実際の講義では、オンライン・リアルタイムで理解して行く必要があり、もっと厳しい条件である。ちなみに、英語が完全に正しく聞き取れたとした場合(書き起こし正解率が 100% の場合)は、学生の翻訳結果の BLEU 値は約 10~16 程度であった(平均 12.9)。これに対して、我々の申請時の音声認識システムでは、単語正解率は約 80~85% 程度(Google の音声認識精度とほぼ同等)、この音声認識結果文に対する翻訳の BLEU 値は約 9 であった。これらの結果は TOEIC 750 点の学生の聞き取り率 65%、その翻訳結果の BLEU 値 8 を上回っている。しかし、残念ながら、我々の申請時のシステムでは、正しく英語講義を理解するための字幕精度としてはまだ不十分であった。

2. 研究の目的

上述したように、我々の申請時のシステムは、TOEIC 800 点相当の学生以上の英語音声認識と TOEIC 600 点相当の英語 - 日本語テキスト翻訳を実現できているが、音声翻訳の精度は BLEU 値で 9、テキスト翻訳で約 12 程度である。これは、ルールベースの翻訳ツールである EXCITE 翻訳と同程度、大量コーパスで学習している Google 翻訳(BLEU 値 約 15)と比べれば性能はまだ低い。我々の経験では、英語講義をほぼ正しく理解できるためには BLEU 値で 20 以上(TOEIC 800 点以上の学生のテキスト翻訳に相当)、理解支援でも 15 以上(TOEIC 700 点以上相当)が必要と考えている。この値を達成するための英語音声認識システム(英語講義中の単語聞き取り率・認識率 90% 程度)とその日本語への音声翻訳システムを開発することが第一の研究目的である。

英語の講義に対して、すべて正しく書き起こしてきたとしても、その読解に注力し過ぎて、講師のスライドを指すジェスチャーや表情をみる時間的余裕がなくなり、全体の理解の妨げになると考えられる。我々の以前の研究では、全文字幕と重要文だけの字幕、及び重要句だけの字幕を英語表示と日本語表示で比較し、ともに講義の理解にこの 3 者の差は、ほとんどないことを明らかにしている。もちろん、好まれる字幕表示は、英語能力とも関係している。音声認識結果の整形機能も取り入れ、読みやすく理解しやすい字幕表示方法、重要文や重要箇所の抽出にもとづく素早く内容が理解できる字幕方法など、学生の英語能力に応じた字幕の表示方法を開発する

ことが第二の研究目的である。

3. 研究の方法

英語講義音声日本人学生が理解できるため、あるいは理解支援するためには、英語書き起こし正解率 85%~90%程度、日本語への翻訳精度の BLEU 値 20 以上が望まれる。これを達成するための英語音声認識システム(英語講義中の単語認識率 90%程度)とその日本語への音声翻訳システム(BLEU 値 20 以上)を開発するのが目的である。

音声認識に関しては、我々が開発を進めているディープニューラルネットワークに基づく音声認識システムの特徴パラメータに人間の抹消系聴覚に基づいたフィルタバンク特性を自動学習する手法を発展させる。

音声翻訳に関しては、ニューラル翻訳では、{大量の原文の英語 その日本語訳の文}のペア集合(パラレルコーパス)が必要であるが、これを少量コーパスからの適応技術で対処する。話し言葉である英語講義音声に近い発話スタイル・ドメインで利用可能な代表的なパラレルコーパスは TED Talks であるが、現在のところ利用できる TED Talks の英語 日本語パラレルコーパスは 20 万文ペア程度である。少ないパラレルコーパスで翻訳精度をあげるために、新聞などの他の大量のパラレルコーパスから TED Talks のような話し言葉への発話スタイル適応技術(transfer learning)を開発する。また、話し言葉の単言語コーパスを翻訳し、疑似的なパラレルコーパスを利用する方法を開発する。

文の要約に関しては、英語講義・講演音声と時間的に同期した字幕を表示することが望ましいため、重要文抽出に基づく要約法を開発する。字幕の読解速度を考慮して要約率(重要文抽出率)は 40~50%程度とする。ニューラルネットワークによる sequence to sequence (end to decode)写像に基づく文単位の embedding による意味表現を利用する。文の意味表現ベクトル間の類似度により冗長な文の除去など、従来の重要文要約法には対処出来ない補完機能が実現でき、要約性能の向上が得られものと考えられる。

英日翻訳と要約の研究は、最初はテキスト入力(音声認識率 100%に相当)で行い、最終年度で、音声入力に対する音声翻訳と音声要約に適用し、システムの評価を行う。

研究代表者と研究分担者の役割分担は次の通りである。研究代表者の中川は、システム全体の計画・指導と下記の分担者の研究成果を取り入れて、音声入力に対する音声翻訳・音声要約システムを開発する。研究分担者の秋葉は、統計的機械翻訳や音声ドキュメント検索に関する研究を長く携わり、最近ではニューラル機械翻訳に力を入れ、日英機械翻訳システムの改善を図る。もう一人の分担者である山本は、実環境下での音声認識の研究に長く携わってきており、英語音声の認識システムの開発を行う。

4. 研究成果

(1) 音声認識

最近のニューラルネットワークに基づく音声認識技術の進展は目覚ましく、様々なツールが公開されている。本研究では、TED 講演音声をテキストに書き起こすための音声認識ツールキットとして Kaldi を用いた。Kaldi では認識システムとして DNN-HMM を用いており、これは隠れマルコフモデル(Hidden Markov Model)から出力される音素の状態確率を DNN(Deep Neural Network)を用いて算出する手法である。

本研究では DNN の隠れ層の層数は 13 層とし、隠れ層の各ユニット数は 1024 とした。出力層のユニット数は共有トライフォン数に対応している。ネットワークへの入力特徴量は MFCC(メル周波数ケプストラム係数)40 次元を使用した。特徴量の前処理として fMLLR、LDA、

SAT の特徴量変換を行い、不特定話者用の特徴抽出を行った。

使用した音声データは約 450 時間の TED 講演音声コーパスの TEDLiumv3 である。学習データ、検証データ、評価データでセットが分かれており、それぞれ重複する講演は含まれていない。また、大量の読み上げ音声データである Librispeech の約 1000 時間を使用して初期モデルを学習した。音声認識の評価には、IWSLT-2016 の 10 講演と TEDLiumv3 のテストセット 16 講演を使用した。

以下(a), (b),(c)の三通りの手法でモデルを作成し、音声認識率を比較した。

(a) TED 講演音声による単独モデル

TED 講演 450 時間の学習データで DNN-HMM を学習した。言語モデルは TED 講演 450 時間のテキストから作成し語彙サイズは 15 万語程である。

(b) Librispeech による単独モデル

Librispeech 960 時間の学習データで DNN-HMM を学習した。このモデルのみ、DNN の規模を、層数 17 層、隠れ層のユニット数は 1536 とした。言語モデルは Librispeech 960 時間のテキストから作成し、語彙サイズは 20 万語程である。

(c) TED 講演音声と Librispeech の混合学習によるモデル

Librispeech 960 時間の学習データと TED 講演 450 時間の計 1410 時間を用いて DNN-HMM を学習した。言語モデルは TED 講演 450 時間と Librispeech 960 時間を合わせたテキストから作成し、語彙サイズは 28 万語程である。

TED+Librispeech 混合モデルでの認識率は TED 単独モデルよりやや良く、音声認識率は IWSLT-2016 で約 88%、TEDLiumv3 で約 89%であった。Librispeech 単独モデルから TED データによる適応を行うとさらに良くなると思われる。なお、(b)のモデルを用いた Librispeech テストデータでの認識実験では単語正解精度は 94.8%で、最先端の性能レベルであった。

以降、IWSLT-2016 に対して一番認識率が高かった TED 単独モデルを用いて音声翻訳と音声要約を行った。このために、音声認識結果の数値の読み表現を数値表現に後処理した(例: thirteen 13)。なお、認識結果には句読点はないので、音声翻訳・音声要約の評価時の参照テキストも句読点を除去した。

(2) Transformer 翻訳モデルによる音声翻訳

翻訳モデルは Transformer モデルを使用した。これはエンコーダとデコーダからなり、再帰型ニューラルネットワークの様に時系列データを用いて学習を行う。出力を求める際は自己注意機構を用いる。エンコーダは同じ構成のエンコーダの積み重ねによって構成されており、それらのエンコーダは自己注意とフィードフォワードネットワーク(FFNN)により構成される。デコーダも同じ構成のデコーダが積み重ねられてできている。一つのデコーダは自己注意と FFNN に加えてその間に注意機構が入っている。標準モデルはエンコーダ 6 層、デコーダ 6 層であるが、最適な数は学習データ量に依存する。22 万文の英日パラレルコーパスで学習した場合は、3 層 - 3 層が一番良かったので、本実験ではエンコーダとデコーダは共に 3 層とした

本実験で使用する TED 講演の IWSLT のコーパスには英語と日本語の対訳コーパスが 22 万文と少ないため、英語または日本語の単言語コーパスをベースモデルで翻訳/逆翻訳することで英語と日本語の疑似対訳コーパスを作成し、翻訳モデル学習の追加の学習データとすることでデータ拡張を行った。データ拡張には IWSLT2018 英語-スペイン語ペアの英語側コーパス(約 8 万文)と CSJ 日本語コーパスに含まれている模擬講演(約 22 万文)を用いた。また、論文抄録 100 万文ペアからなる書き言葉 ASPEC コーパスにより英日単/双方向翻訳モデルを学習し、そのパラメータを初期値として IWSLT+データ拡張したデータセットで学習した。なお、ASPEC

コーパスによる翻訳モデル学習の際には、TED 語彙に合わせて学習を行った。双方向モデルとは、英日と日英の翻訳モデルが同一のことである。

音声翻訳結果の BLEU 値を表 1 に示す。音声翻訳はテキスト翻訳と比べて、約 14% BLEU 値が低下した。10%程度の低下に留まるためには、音声認識精度が 90%以上必要と思われる。翻訳性能はテキスト翻訳および音声翻訳共に初期の高い目標には届かなかったが、人手による評価では、テキスト翻訳に対して、一部誤りがあるが、ほぼ内容が理解できるという結果が得られ、授業支援としての目標は達成できた(5点満点で 2.9、人手の翻訳結果に対しては 4.1 の評価)。

表 2 音声要約結果 (Rouge3) : IWSLT-2016

表 1 音声翻訳結果 (BLEU)

テストデータ : IWSLT-2016

モデル 入力	Transformer 単方向モデル	Transformer 双方向モデル
テキスト	15.14	15.52
音声	12.44	13.35

手法	テキスト	音声入力	
	原文	認識結果	原文に変換
MMR	40.4		
MMR+BERT	54.8		
BertSum-CNN	53.8	42.9	55.1
BertSum-TED	47.8	35.3	47.5
BertSum-CNN/TED	62.1	48.2	62.1

(3) 音声要約

従来の代表的な重要文抽出要約法である MMR は、ドキュメント全体との関連度と、情報の新規性に基づいて抽出する文を順に決定していくことで、全体としてドキュメントとの関連が高くかつ冗長性の低い文集合を抽出することを目指す手法である。本研究では、文間の類似度に BERT による分散表現のベクトル間距離を使用した。さらに、BERT による文ベクトルで重要文と非重要文を識別するネットワークを学習し、この出力値(重要文の事後確率)も付加して重要文を決定した(MMR+BERT法)。

また、最近のニューラルネットワークをベースにした代表的な重要文抽出法である BertSumExt も使用した。この手法は、複数文の文境界記号付き入力単語列(512 トークン以下の制限あり)に対して BERT によって文ベクトル列を得て、これを Transformer に入力して、重要文/非重要文のラベル列を抽出する手法である。本手法を TED 講演に利用するために、講演を 512 トークン以下に分割し、それぞれで重要文抽出を行い、最後に講演全体で設定された重要文数になるようにスコアで選定した。また、オリジナルな BertSumExt は、CNN/Daily news 約 30 万記事の 1000 万文から、各記事から 3 文の重要文を抽出するように学習されており、このモデル(CNN)と、MMR+BERT と同じく TED の 51 講演の約 4500 文から学習したモデル(TED)、CNN のモデルから TED データで適応したモデル(CNN/TED)を使用した。

IWSLT-2016 の 10 講演のテキストデータで行った要約実験の結果を表 2 に示す。重要文の割合は約 46%である。TED データだけの学習では MMR+BERT が最も優れているが、CNN のモデルから TED データで適応した BertSumExt の CNN/TED は、これを上回った。次に、BertSumExt のみで音声入力の要約実験を行った。要約結果を表 2 に示す。音声要約では抽出された文の認識結果に対応する原テキストデータを用いての評価も行い ROUGE 値欄の一番右欄に示す。これは重要文箇所を音声で再生した場合は音声認識の誤りに関係なくテキスト内容が正しく聴講できるための尺度である。テキスト入力と比べて音声入力では音声認識誤りの影響により Rouge-3 で 15~20%の性能低下がみられた。しかし、重要文箇所の抽出の尺度(音声認識結果を元の原文に変換後の ROUGE)では、音声認識誤りの影響はなく性能の低下はみられなかった。

5. 主な発表論文等

〔雑誌論文〕 計3件（うち査読付論文 3件/うち国際共著 0件/うちオープンアクセス 2件）

1. 著者名 森田智熙、秋葉友良、塚田元	4. 巻 106-D, No.4
2. 論文標題 ニューラル機械翻訳の反復的逆翻訳に基づくデータ拡張のための混成サンプリング法	5. 発行年 2023年
3. 雑誌名 電子情報通信学会論文誌	6. 最初と最後の頁 298, 306
掲載論文のDOI（デジタルオブジェクト識別子） 10.14923/transinfj.2022PDP0005	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Hiroshi Seki, Kazumasa Yamamoto, Tomoyosi Akiba, Seiichi Nakagawa	4. 巻 E102-D, No.2
2. 論文標題 Discriminative learning of filterbank layer within deep neural network based speech recognition for speaker adaptation	5. 発行年 2019年
3. 雑誌名 IEICE Transaction, INF.&SYST.	6. 最初と最後の頁 364, 374
掲載論文のDOI（デジタルオブジェクト識別子） 10.1587/transinf.2018EDP7252	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

1. 著者名 中川聖一	4. 巻 54
2. 論文標題 最近の音声言語処理研究の動向 - 筆者の音声認識、音声翻訳、話者認識の研究を中心として -	5. 発行年 2019年
3. 雑誌名 中部大学工学部紀要	6. 最初と最後の頁 7, 20
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

〔学会発表〕 計36件（うち招待講演 0件/うち国際学会 5件）

1. 発表者名 K. Masuda, Y. Hayakawa, K. Yamamoto, S. Nakagawa
2. 発表標題 Summarization of spoken lectures based on MMR method and important/unimportant sentence using BERT
3. 学会等名 Proc. Global Conference on Consumer Electronics（国際学会）
4. 発表年 2022年

1. 発表者名 K. Takagi, T. Akiba, H. Tsukada
2. 発表標題 Semi-supervised ASR based on iterative joint training with discrete speech synthesis
3. 学会等名 Proc. APSIPA (国際学会)
4. 発表年 2022年

1. 発表者名 足立十一郎、山本一公、中川聖一
2. 発表標題 TED講演の英日翻訳と日英翻訳の検討
3. 学会等名 言語処理学会、第29回年次大会
4. 発表年 2023年

1. 発表者名 坂野晴彦、桜井陽生、足立十一郎、山本一公、中川聖一
2. 発表標題 TED英語講演の音声認識・音声翻訳・音声要約の検討
3. 学会等名 言語処理学会、第29回年次大会
4. 発表年 2023年

1. 発表者名 山本優紀、秋葉友良、塚田元
2. 発表標題 ニューラル機械翻訳におけるIterative back-translationを利用したコンパラブルコーパスの活用
3. 学会等名 言語処理学会、第29回年次大会
4. 発表年 2023年

1. 発表者名 谷川琢磨、秋葉友良、塚田元
2. 発表標題 双方向翻訳モデルの相互学習による対訳語彙の教師なし獲得過程の調査
3. 学会等名 言語処理学会、第29回年次大会
4. 発表年 2023年

1. 発表者名 B. T. Thanh、秋葉友良、塚田元
2. 発表標題 中間言語を介した2つの対訳コーパスを用いた対訳文のない言語対のNMTの検討
3. 学会等名 言語処理学会、第29回年次大会
4. 発表年 2023年

1. 発表者名 河合輝也、秋葉友良
2. 発表標題 答弁の種類に着目した抽象型要約に基づく議会議録質問応答
3. 学会等名 言語処理学会、第29回年次大会
4. 発表年 2023年

1. 発表者名 K. Yamamoto, A. Ishiki, S. Nakagawa
2. 発表標題 Improvement of elderly speech recognition using Gammatone filterbank adaptation
3. 学会等名 Global Conference on Consumer Electronics (国際学会)
4. 発表年 2021年

1. 発表者名 T. Kawai, T. Akiba, S. Masuyama
2. 発表標題 Development of political QA systems targeted as assembly minutes based abstractive summarization
3. 学会等名 Internatinal Conference on Advanced Informatics: Concepts, Theory and Applications (国際学会)
4. 発表年 2021年

1. 発表者名 B. T. Thanh, 秋葉友良、塚田元
2. 発表標題 双方向翻訳モデルと反復的逆翻訳を用いた低資源言語に対するニューラル機械翻訳の性能向上
3. 学会等名 言語処理学会、第28回年次大会
4. 発表年 2022年

1. 発表者名 山岸勇輝、秋葉友良、塚田元
2. 発表標題 入力側単言語資源と転移学習の利用による講演字幕を対象とした英日ニューラル機械翻訳の改善
3. 学会等名 言語処理学会、第28回年次大会
4. 発表年 2022年

1. 発表者名 森田知てる、秋葉友良、塚田元
2. 発表標題 翻訳の品質評価に基づく動的な混成サンプリングによるNMTの双方向反復逆翻訳手法の改善
3. 学会等名 言語処理学会、第28回年次大会
4. 発表年 2022年

1. 発表者名 西田悠斗、秋葉友良、塚田元
2. 発表標題 ニューラル機械翻訳のための日本語膠着語的性質を考慮したマルチタスク学習
3. 学会等名 言語処理学会、第28回年次大会
4. 発表年 2022年

1. 発表者名 紺谷優志、秋葉友良、塚田元
2. 発表標題 ラウンドトリップ翻訳を用いたニューラル機械翻訳のデータ拡張
3. 学会等名 言語処理学会、第28回年次大会
4. 発表年 2022年

1. 発表者名 谷川琢磨、秋葉友良、塚田元
2. 発表標題 Iterative Back Translationは対訳語彙を獲得できるか？
3. 学会等名 言語処理学会、第28回年次大会
4. 発表年 2022年

1. 発表者名 高木景矢、秋葉友良、塚田元
2. 発表標題 Iterative Back Translationと離散音声表現を用いた音声認識のためのデータ拡張
3. 学会等名 日本音響学会春季研究発表会
4. 発表年 2022年

1. 発表者名 民谷慎一郎、秋葉友良、塚田元
2. 発表標題 複数の音声認識結果を用いた系列変換モデルによる音声翻訳システムの検討
3. 学会等名 日本音響学会、秋季研究発表会
4. 発表年 2019年

1. 発表者名 高木景矢、秋葉友良、塚田元
2. 発表標題 ボトルネック特徴量の合成に基づく音声認識のためのデータ拡張の検討
3. 学会等名 日本音響学会、春季研究発表会
4. 発表年 2020年

1. 発表者名 民谷慎一郎、秋葉友良、塚田元
2. 発表標題 End-to-Endとカスケード方式のアンサンブルによる音声翻訳の検討
3. 学会等名 日本音響学会、春季研究発表会
4. 発表年 2020年

1. 発表者名 早川由倭、山本一公、中川聖一
2. 発表標題 文の分散表現を利用したMMR法に基づく講義・講演ドキュメントの要約
3. 学会等名 言語処理学会、第26回年次大会
4. 発表年 2020年

1. 発表者名 高橋竜、秋葉友良、塚田元
2. 発表標題 汎用分散表現BERTを用いたニューラル機械翻訳の検討
3. 学会等名 言語処理学会、第26回年次大会
4. 発表年 2020年

1. 発表者名 藤澤謙太、秋葉友良、塚田元
2. 発表標題 ニューラル機械翻訳における双方向反復的教師なし適応の改善
3. 学会等名 言語処理学会、第26回年次大会
4. 発表年 2020年

1. 発表者名 山岸勇輝、秋葉友良、塚田元
2. 発表標題 逆翻訳を用いたデータ拡張と転移学習を利用した英日講演字幕翻訳の改善
3. 学会等名 言語処理学会、第26回年次大会
4. 発表年 2020年

1. 発表者名 Yuuki Yamagishi, Tomoyosi Akiba, Hajime Tsukada
2. 発表標題 English-Japanese Machine Translation for Lecture Subtitles Using Back-Translation and Transfer Learning
3. 学会等名 Proc. IEEE 9-th Global Conf. on Consumer Electronics
4. 発表年 2020年

1. 発表者名 森田知、秋葉友良、塚田元
2. 発表標題 Fine-Tuningと混成的な逆翻訳サンプリングに基づくNMTの双方向反復的教師なし適応の改善
3. 学会等名 言語処理学会、第27回年次大会
4. 発表年 2021年

1. 発表者名 藤澤謙太、秋葉友良、塚田元
2. 発表標題 NMTの双方向反復的教師なし適応手法における初期対訳コーパスサイズの影響と翻訳モデル獲得に関する調査
3. 学会等名 言語処理学会、第27回年次大会
4. 発表年 2021年

1. 発表者名 高木景矢、秋葉友良、塚田元
2. 発表標題 End-to-End音声翻訳のためのデータ拡張の検討
3. 学会等名 言語処理学会、第27回年次大会
4. 発表年 2021年

1. 発表者名 Hiroshi Seki, Kazumasa Yamamoto, Tomoyosi Akiba, Seiichi Nakagawa
2. 発表標題 Rapid speaker adaptation of neural network based filterbank layer for automatic speech recognition
3. 学会等名 IEEE on Spoken Language Technology Workshop (国際学会)
4. 発表年 2018年

1. 発表者名 佐橋広也、西村友樹、秋葉友良、中川聖一
2. 発表標題 統計的機械翻訳とニューラル翻訳による翻訳候補の文の分散表現に基づくリスコアリングの検討
3. 学会等名 情報処理学会、音声言語情報処理研究会
4. 発表年 2018年

1. 発表者名 森田知熙、秋葉友良、塚田元
2. 発表標題 双方向の逆翻訳を利用したニューラル機械翻訳の教師なし適応の検討
3. 学会等名 情報処理学会、第5回自然言語処理シンポジウム
4. 発表年 2018年

1. 発表者名 渡邊拓斗、高田凌平、佐橋広也、山本一公、秋葉友良、中川聖一
2. 発表標題 局所的トピック情報を利用した論文抄録(ASPEC)の英日機械翻訳
3. 学会等名 言語処理学会、第25回年次大会
4. 発表年 2019年

1. 発表者名 佐橋広也、秋葉友良、中川聖一
2. 発表標題 科学技術論文抄録と講義音声の英日機械翻訳のリスコアリングの検討
3. 学会等名 言語処理学会、第25回年次大会
4. 発表年 2019年

1. 発表者名 高田凌平、秋葉友良、塚田元
2. 発表標題 ニューラル機械翻訳におけるトピック情報の利用
3. 学会等名 言語処理学会、第25回年次大会
4. 発表年 2019年

1. 発表者名 中島貴太、関博史、山本一公、中川聖一
2. 発表標題 フィルタバンクと活性化関数の出力値の話者適応に基づくDNN-HMMによる音声認識
3. 学会等名 電子情報通信学会、総合大会
4. 発表年 2019年

1. 発表者名 関博史、山本一公、秋葉友良、中川聖一
2. 発表標題 Encoder-decoderネットワークの枠組みにおけるフィルタバンク層の雑音適応の検討
3. 学会等名 日本音響学会、春季研究発表会
4. 発表年 2019年

〔図書〕 計1件

1. 著者名 中川聖一（編著）、小林聡、峯松信明、宇津呂武仁、秋葉友良、北岡教英、山本幹雄、甲斐充彦、山本一公、土屋雅稔（共著）	4. 発行年 2018年
2. 出版社 コロナ社	5. 総ページ数 288
3. 書名 音声言語処理と自然言語処理（増補）	

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究分担者	秋葉 友良 (AKIBA TOMOYOSI) (00356346)	豊橋技術科学大学・工学(系)研究科(研究院)・准教授 (13904)	
研究分担者	山本 一公 (YAMAMOTO KAZUMASA) (40324230)	中部大学・工学部・教授 (33910)	

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関