# 研究計画提案書（どちらかを○で囲む：修士論文／ 課題研究 ）（修士） 平成 27 年 10 月 28

日

Research Proposal for Master’s Thesis / Research Project (Circle one)

|  |  |  |  |
| --- | --- | --- | --- |
| 氏名 Name | DO KHAC PHONG | 学生番号  Student Number | 1410220 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 主指導教員  Supervisor | Jianwu Dang | 印  Seal | 副テーマ指導教員  Advisor for Minor Research | Ryuhei Uehara | 印  Seal |
| 副指導教員  Second Supervisor | Hirokazu Tanaka | 印  Seal |  | | |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| ＜現在の単位修得状況＞ IS courses I have obtained credits | | | | | | |
|  | 導入講義科目  Introductory courses | 基幹講義科目  Basic courses | 専門講義科目  Technical courses | コース専門講義科目  Specialized Technical courses | 先端講義科目  Advanced courses | その他  Other courses |
| 科目数  Number of courses |  | 1 | 2 |  |  |  |
| 単位数  Number of credits |  | 2 | 4 |  |  |  |

＜研究テーマ＞ Research Title

A study of phonemic restoration using deep neural network

＜研究の目的＞ Research Aim

目的を書く

State your aim.

In daily life, sometimes our speech is more likely to be distorted by extraneous transient sounds such as cough, traffic noise, and so on. Surprisingly, in many cases, although speech sound is missing partially by the interference, we can understand speech’s content without any problem. This phenomenon is known as phonemic restoration - an outstanding and sophisticated capacity of our brain which can create speech signals to compensate for the missing part. To fully the mechanism involved in the phonemic restoration, we evaluate several hypotheses on the phonemic restoration use acoustic and articulatory data of speech utterances based on deep neural network.

＜研究の背景・特色・重要性＞ Research Background, Originality, Significance

（修士論文）研究の背景・特色を書く

(Master’s Thesis) State your research background and its originality.

In daily communication, in case spoken words are masked by background noise, we can still hear these words and comprehend the content of the speech as if they had not been masked. The auditory system has the capacity to restore the disrupted portions of the spoken words, making speech more meaningful. This mechanism uses a great deal of intrinsic and extrinsic elements such as expectations, linguistic knowledge, syntactic, semantic, and lexical constrains and context in noisy environments. This phenomenon is known as phonemic restoration. In order to simulate this sophisticated capacity of brain, many preceding research were conducted in this field. Some of them added noises with suitable magnitude in missing words, others filled in missing words by spectrum in frequency domain. However, results of these research is not enough objective and use a small amount of data.

At the moment, Deep Neural Network (DNN) is perceived as a potential approach to solve speech and speaker recognition issues. Therefore, I make a decision to use DNN to restore missing words of the speech. In addition, I also integrate two various types of data, namely acoustic data and articulatory data. According to some papers, using articulatory data increases the intelligibility and continuity of speech rather than utilizing acoustic data solely.

＜研究計画・方法＞Research Plan, Method

（修士論文）本研究では、以下の研究結果を論文にまとめる。

(Master’s Thesis) State your research plan and method. For example using this format:

This research comprises

1. System survey (1 November – 31 December)

In this period of time, I have an intention to gather general information in terms of signal and speech processing. First of all, myriads of fundamental aspects of speech (frequency, frame rate, spectrum…) and some basic transform functions of signal. Subsequently, I am going to conduct a research some approaches with respect to phonemic restoration.

At the same time, I enroll on a subject titled “Analysis of Information Science” where I can derive essential background of signal transformation and description.

1. Data analysis (1 January – 31 January)

There are two different types of data used in this research, namely acoustic data and articulatory data. Both of them are collected by using Electromagnetic Midsagittal Articulographic (EMA) system. Articulatory data consist of three dimensional data related to the position of the upper lip, the lower lip, the jaw, the tongue bottom, the tongue tip and tongue dorsum. According to some research, articulatory data contribute greatly to the intelligibility of the speech with missing sounds.

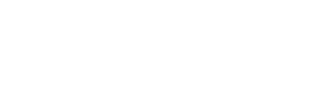
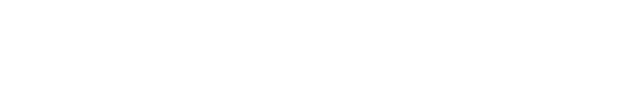
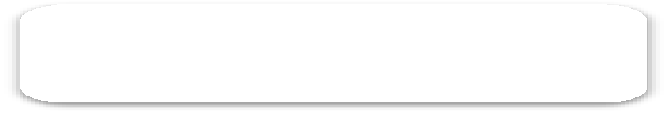
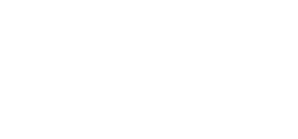
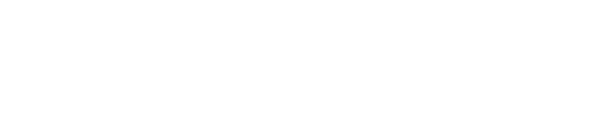
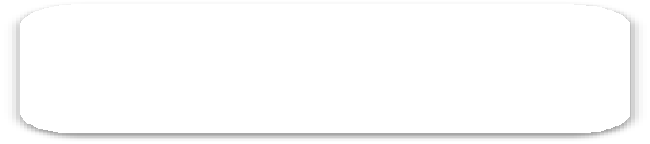
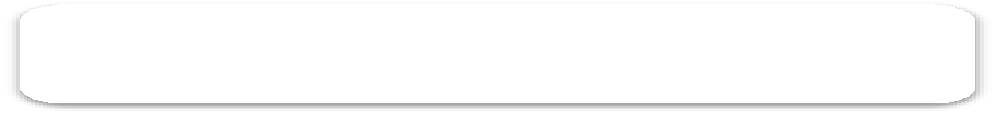
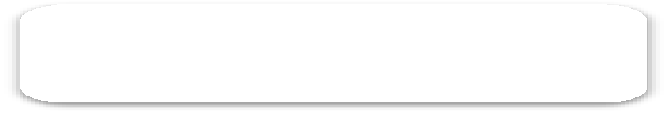
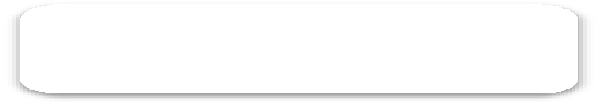
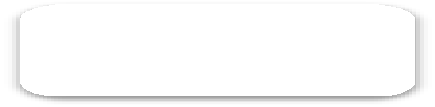
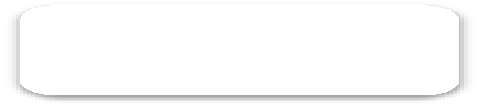
I intend to apply Mel Frequency Cepstral Coefficients (MFCCs) which are a feature widely used in automatic speech and speaker recognition.

Deep Neural Network is also ultilised in this research because of its outperformance in recognition.

1. System design (1 February – 15 March)

- Training data using Deep Neural Network

At the first stage of training data, MFCC of disrupting acoustic signal will be extracted, including delta and delta-delta. Meanwhile, the articulatory data provide position of lips, jaw and tongue in three dimensional space. These data will be integrated into a feature vector to train a DNN. The output will be a MFCC describing the original acoustic signal.



Disrupting

Acoustic data

Articulatory

data

MFCC (36 features)

Extract position, velocity, acceleration

+ Vector Quantization (6 \* 3 \* 36)

Integrate features into

a vector

Deep Neural Network

MFCC of original

acoustic data

4. Experiment and evaluation (16 March– 15 May)

Signal with missing words will be created to test my model. Particularly, each missing signal and its articulatory data will test DNN. The stage of feature extractions is the same as training data. After that, this feature vector will be tested by DNN and release a MFCC feature as an original feature of signal. After that, we will reconstruct a signal from that MFCC feature and the evaluation will be based on its

continuity and intelligibility.

5. Writing the thesis (16 May – 1 July)

Eventually, it takes me one and a half month to synthesize my experiment result and write Master’s thesis.