

Title	音声信号と調音状態の一对多の関係の分析及びその応用に関する検討
Author(s)	錦戸, 信和
Citation	
Issue Date	2011-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/9605
Rights	
Description	Supervisor: 党建武, 情報科学研究科, 博士

A Model-based Investigation and Application of One-to-Many Relationship between Speech Sound and Articulation

Akikazu Nishikido
School of Information Science,
Japan Advanced Institute of Science and Technology

March, 2011

Abstract

In process of speech production, a speech sound within the same category can be produced by various articulations with different positions or configurations of speech organs. Such a relationship between speech sound and articulation brings an ill-posed problem, namely the one-to-many problem, in inverse estimation of articulation from speech sound. In other words, there may be countless solutions from an inverse estimation for a given speech sound. In previous studies, the one-to-many relationship has been reduced to some extent by employing morphological, dynamic and physiological constraints. Nevertheless, the one-to-many problem is far from being solved. Here, observations from ventriloquism or articulatory compensation indicate that there exist two types of basic articulations that can produce speech sounds with the same category. Both of them can be physiologically realized by humans. But, one kind of articulation appears in natural speech, while the other does not appear in natural speech. The former is referred to as “natural articulation”, and the latter is “unnatural articulation”. The unnatural articulations cause the one-to-many problem in the inverse estimation since some unnatural articulations satisfy the above constraints. Therefore, it is necessary to exclude unnatural articulations from candidates of the inverse estimation for solving the one-to-many problem.

The purpose of this study is to exclude unnatural articulations from candidates of the inverse estimation of articulation. For this purpose, we generated a great variety of possible articulations using a physiological articulatory model, and visualized the articulatory structure based on analyzing the articulations generated. Moreover, we proposed a method for discriminating between natural and unnatural articulations, and excluded the unnatural ones from candidates in the inverse estimation by applying the proposed method to the inverse estimation system.

The following outcomes are indicated from this study.

1. the structure of articulations in the one-to-many relationship was clarified by clusters in non-linear space.
2. a method for discriminating between natural and unnatural articulations was proposed, and natural articulations could be discriminated with accuracy more than 97%, and unnatural articulations could be discriminated with accuracy more than 99%.
3. the inverse estimation system for articulations was constructed with applying the proposed method for the inverse estimation.
4. 90% of unnatural articulations from inverse estimation was excluded.

The above outcomes are not only benefit for the inverse estimation of articulation from speech sound, but also help for people with hearing difficulties, or for acquiring a second language.

Key Words: one-to-many problem, inverse estimation, unnatural articulation, articulatory model, vowel production