

CONTROL STRATEGY OF
PHYSIOLOGICAL ARTICULATORY MODEL FOR
SPEECH PRODUCTION

Xiyu Wu*

Jianwu Dang†

Peking University

Tianjin University

Japan Advanced Institute of Science and Technology

ABSTRACT

In speech production the articulatory apparatus includes the organs that execute efferent motor commands from the central nervous system. In order to simulate the behavior of the articulatory apparatus, computational modeling is a commonly-used method. We have constructed a full 3D physiological articulatory model that includes the tongue, jaw, hyoid bone and vocal-tract wall based on the continuum finite element method. This model comprises articulatory muscles with realistic properties and geometrical arrangements. The muscle activation patterns are used to control the movements of the model. In order to use the model to investigate the speech motor control mechanism, generate speech sounds, predict the effect of surgical operation, etc., we have to realize an automatic control strategy. The task of the control strategy is given a particular target, how to generate muscle activation patterns that can control the model to achieve the target. There are two main control strategies for the physiological articulatory model: feedforward control and feedback control. Feedforward control is a kind of mapping used to directly find muscle activation patterns according to the desired target, and feedback control is used to adjust muscle activation patterns to reduce the distance between the

* **Xiyu Wu** 吴西愉 teaches at Peking University and conducted part of this study at Japan Advanced Institute of Science and Technology (JAIST).

† **Jianwu Dang** 党建武 is a professor of Tianjin University, and Japan Advanced Institute of Science and Technology (JAIST). He is the corresponding author of this article: 1-1, Asahidai, Nomi City, Ishikawa, Japan, 923-1292; [jdang@jaist.ac.jp]

desired target and the realized position. In speech production, feedforward mapping is used to rapidly generate muscle activation patterns to control the articulators to produce fluent speech. Feedback control plays the role of learning and maintaining the feedforward mapping. When the degree of accuracy using feedforward mapping cannot satisfy the requirement, feedback control can be used to realize fine motor control. In this paper, we describe how to use feedback control as a learning loop to construct feedforward mapping. The constructed feedforward mapping was assessed through an open-set test, and reasonable articulatory positions were obtained by comparison with the desired targets. Furthermore, the ability of feedback control to improve control accuracy was proved by a large quantity of simulations.

SUBJECT KEYWORDS

Speech production Physiological Articulatory Model Muscle activation pattern Speech motor control

面向言語產生的發音生理模型控制方法

吳西愉*

北京大學

党建武†

天津大學

日本北陸先端科學技術大學院大學

提要

在言語產生的過程中，發音器官是執行來自中樞神經系統的運動指令的終端組織。電腦建模的方法常常被應用於模擬發音器官的行為。為此，我們使用連續體有限元的方法建立了一個三維的發音器官的生理模型，該模型包含舌、下頷、舌骨以及聲道壁等發音器官。該模型還包括了根據其生理解剖屬性建立的用於控制發音器官運動的肌肉模型。為了將該生理模型應用於探索發音器官的運動控制機制，產生自然流暢的語音，對發音器官的手術後功能進行預測等，我們需要對模型建立一個自動的控制機制。控制機制的任務在於給定一個發音目標，如何自動的產生肌肉激活模式去控制模型到達目標。對於發音器官的生理模型而言，有兩種主要的控制方式：前饋控制和回饋控制。前饋控制是一種從發音控制目標到肌肉激活模式的映射，用於根據發音目標產生肌肉激活模式；而回饋控制主要用於調整肌肉激活模式來減少模型實現的位置到目標之間的距離，最終控制模型到達目標。在言語產生的過程中，前饋控制用於快速的產生肌肉激活模式控制發音器官產生流暢的語音。回饋控制主要用於學習前饋控制的映射，並且維持前饋控制可行性。當前饋控制的精度不能滿足需求時，回饋控制可以用於提高控制精度以實現精確控制。在本文中，我們將重點介紹如何使用回饋控制作為學習回路，建立前饋控制的映射。對已建立的前饋映射的開集測試表明該映射可以用於控制模型在誤差允許範圍內達到目標。而且，大量的模型模擬表明回饋控制可用於改善前饋控制的控制精度。

主題詞

言語產生 發音生理模型 肌肉激活模式 發音器官運動控制

* 吳西愉從事教研工作於北京大學，並求學於日本北陸先端科學技術大學院大學。

† 党建武為天津大學教授和日本北陸先端科學技術大學院大學教授，本文通訊作者：日本石川縣能美市旭台 1-1，郵編 923-1292，電郵[jdang@jaist.ac.jp]