行政院國家科學委員會專題研究計畫　成果報告

以磁振影像為基礎的立體舌頭圖譜之建構 II

計畫類別：個別型計畫
計畫編號：NSC92-2218-E-040-001-
執行期間：92 年 08 月 01 日至 93 年 07 月 31 日
執行單位：中山醫學大學語言治療與聽力學系

計畫主持人：吳炤民
共同主持人：王高倫

報告類型：精簡報告

處理方式：本計畫可公開查詢

中　華　民　國 93 年 11 月 2 日

行政院國家科學委員會補助專題研究計畫 ■ 成 果 報 告
□期中進度報告

以磁振影像為基礎的立體舌頭圖譜之建構 II

Construction of MRI-based 3D Atlas of the Human Tongue II

計畫類別：■ 個別型計畫　　□ 整合型計畫
計畫編號：NSC-92-2218-E-040-001
執行期間：九十二年八月一日至九十三年七月三十一日

計畫主持人：吳炤民
共同主持人：王高倫
計畫參與人員：

成果報告類型(依經費核定清單規定繳交)：■精簡報告　□完整報告

本成果報告包括以下應繳交之附件：
□赴國外出差或研習心得報告一份
□赴大陸地區出差或研習心得報告一份
□出席國際學術會議心得報告及發表之論文各一份
□國際合作研究計畫國外研究報告書一份

處理方式：除產學合作研究計畫、提升產業技術及人才培育研究計畫、列管計畫及下列情形者外，得立即公開查詢
□涉及專利或其他智慧財產權，□一年□二年後可公開查詢

執行單位：中山醫學大學 語言治療與聽力學系

中　華　民　國 九十三年十月二十七日

中文摘要

國內外有關構音及運動性言語障礙的研究大多集中在相關的聲學研究，然而隨著科技的日新月異，國外已漸將這方面的研究重心轉向和說話相關的生理系統來瞭解正常與說話障礙機轉。而在說話器官中以舌頭為最重要，舌頭在口腔內的運動可以發出不同的母音和子音；但是對舌頭的瞭解仍停留在基本的味覺及肌肉的解剖構造無法對說話時的舌頭變形提出有效的量化基礎。本計畫的主要目的在利用空間轉換技術來建立磁振影像為基礎的立體舌頭圖譜，做為舌頭在正常與說話障礙機轉的影響以及舌頭之形態分析研究的基礎。兩年的研究計畫中，第一年的工作內容以蒐集相關方法、文獻、以及建立初期的磁振影像資料庫（無構音障礙之十六名大專青年、男女各半），在研究發展上以確立磁振影像蒐集的步驟以及建立立體舌頭圖譜所需的影像登錄之空間轉換技術，並建立立體舌頭圖譜，其考量的重點是在運算的複雜性、準確度、及統計分析的可行性為主。第二年則擴大磁振影像資料庫的建立（無構音障礙四十名大專青年、男女各半），以建立立體舌頭的解剖界標點為主要工作，做為影像登錄及立體舌頭圖譜準確度評估的依據。本研究成果的立體舌頭的解剖界標點的選擇可以反應舌頭肌肉的收縮狀況為考量，可做為日後舌頭的生物力學模擬之應用。

關鍵詞：舌頭，空間轉換，磁振影像，說話，界標點

Abstract

Most of speech science research related to articulatory and motor speech disorders is based on the study of the recorded acoustic data to investigate the underlying mechanism of speech production. As modern technology advances, more researchers in western countries move their interests to physiological mechanism of speech production in normal and disordered speech. The tongue has been recognized to be the most important organ for the speech production process. Nevertheless, many reported studies of the tongue are either muscular anatomy oriented or related to the taste function. The main objective of this project is to build a MRI-based 3D tongue atlas with an established available spatial transformation technique. The 3D tongue atlas will be used as the basis for future research on tongue mophometrics and underlying physiological mechanism of normal and disordered speech production. The first year of the project will include literature search, MRI protocol and data acquisition -8 male and 8 female college students without speech disorders, and construction of the 3D tongue atlas with a selected suitable spatial transformation technique that is determined by its complexity, accuracy, and statistical capability. The second year of the project will include expansion of MRI data acquisition -20 male and 20 female college students without any speech disorders- based on the chosen MRI protocol and definition and selection of 3D tongue landmarks. The objectives of these landmarks are 1) to evaluate the mapping accuracy for the 3D tongue atlas; 2) to reflect tongue muscle contraction, and 3) to apply to the biomechanical modeling of the human tongue.

Keywords: tongue, spatial transformation, MRI, speech production, landmark

I Introduction

The use of magnetic resonance imaging (MRI) techniques has been popular for visualizing the soft-tissue of human body for diagnostic purposes because of its noninvasiveness. It has been applied to anatomical and pathological studies of the tongue and oropharynx [1][2]. The tongue has been recognized to be the most important organ for the speech production process. Nevertheless, many reported studies of the tongue are either muscular anatomy oriented or related to the taste function. Recently MRI related research has been moved from the study of the vocal tract configuration [3][4][5] to the internal musculature deformation of the human tongue [6][7][8]. Problems of these studies are 1) single subject study [6][7][8], 2) lack of statistical analysis capability [6][7], and 3) with limited tongue muscles being assessed [6][7]. Image registration (or spatial transformation) is often necessary when two images obtained by different modalities of medical imaging techniques (e.g. CT and MRI) are integrated to provide complementary information for clinical diagnosis. It requires a mapping function that transforms the coordinates of each point in one image into those of the corresponding point in the other. A mapping function in the context of this proposal is a spatial transformation that maps a set of points from a biological configuration into corresponding points in its warped representation. The mapping methods can be affine, perspective, and polynomial transformations. Since affine and perspective transformations can be expressed as first-order polynomials (see Wolberg, [9]), they are thus subsets of polynomial transformations. A more complete review of the mapping functions is available in Toga [10]. Toga's volume contains most of the updated warping approaches whose applications are mainly for the human brain atlas. The extreme complexity of the brain structure and function often needs cross-modality comparison and integration [11] to provide not only complimentary information but also the generation of atlases and maps [12][13][14]. Although these new brain imaging techniques allow better spatial resolution and interactive graphic tools for viewing and manipulating images, only few of them provide precision and reliability tests for their performance [15]. The main objective of this project is to build a MRI-based 3D tongue atlas with an established available spatial transformation technique. The selected suitable spatial transformation technique will be determined by its complexity, accuracy, and statistical capability. The second objective is to define and select landmarks for the 3D tongue. The purpose of these landmarks is 1) to evaluate the mapping accuracy for the 3D tongue atlas; 2) to reflect tongue muscle contraction, and 3) to apply to the biomechanical modeling of the human tongue. The 3D tongue atlas will then be used as the basis for future research on tongue mophometrics and underlying physiological mechanisms of normal and pathological speech production. The current proposal will provide a substantive methodological preparation for future research, in the form of computational tools and example data, to investigate the phonetic characteristics of various articulations. A typical application is comparison of different vowel gestures by the same tongue (such as the same speaker) and by different tongues (*i.e.* different speakers) in stationary (quasi-static) articulatory positions.

II Methods

A Subjects:

The MR images were acquired using a General Electrical SIGNA 1.5 Tesla scanner in the university hospital of Chung Shan Medical University. With previously described MRI acquisition protocol [16], we extend our MRI data to 40 subjects (20 female and male college students, 19-28 years old) who are native speakers of mandarin with Taiwanese accent and are without speech disorders were the subjects for the MRI data acquisition. The acquisition time for each subject was around 30 minutes. The subjects were given informed consents that were approved by Chung Shan Medical University Hospital Institutional Review Board.

B.MRI protocol and experiments:

The MR images were acquired using a General Electrical SIGNA 1.5 Tesla scanner in the university hospital of Chung Shan Medical University. In consideration of image resolution and acquisition time, each image of the Axial orientation was acquired with a TR of 400 ms and a TE of 10 ms, using an image matrix of 256 × 256 for 35 slices over a field of view of 24 cm (with a resolution of 0.94 mm). The slice thickness is 2 mm with zero inter-slice gaps. The scanning area covered the levels from the line that connects the ANS (anterior nasal spine) and dens of the Axis (the second cervical vertebrae) to the bottom of the tongue (see Fig. 1a). For the purpose of future evaluation of 3D reconstruction, images of the sagittal (TR, 416ms; TE, 10 ms, FOV 24 × 24; image matrix, 256 × 256 for 20 slices with 3 mm thickness, see Fig. 1b) and coronal (TR, 400ms; TE, 10 ms, FOV 24 × 24; image matrix, 256 × 256 for 14 slices with 5 mm thickness) orientations were also acquired for each subject, respectively. So far we have acquired MRI data for 16 subjects (8 female and 8 male college students).


(a)

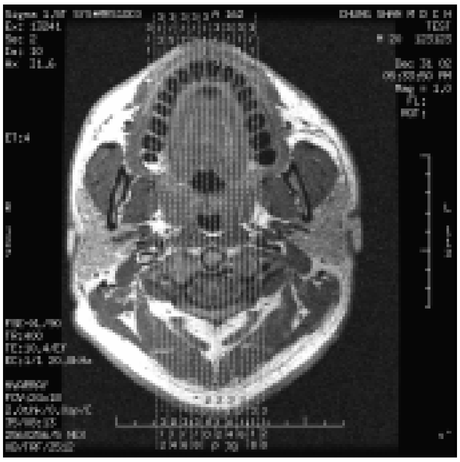
(b)

Fig. 1 (a) the Axial slices localizer; (b) the sagittal slices localizer.

III Results and discussion

Landmarks are the homologous points that define locations having anatomical significance and identifiable geometric coordinates. Finding the homologous points in the paired data sets for tongue images is a difficult task, but it plays a crucial role in the mapping process. How good the mapping depends on how good the landmark points are. It is desirable that all landmarks are robust and easily recognizable with respect to anatomical structures or their readily computable and physically meaningful representative points (such as the centers of gravity).

The first step to select landmarks from the paired data sets is to find common features that can be identified in both data sets. Typical features used for the selection of landmarks are extrema on a contour, intersection among different anatomical structures, maxima of curvatures, and centroids of closed-boundary regions, etc. These features were further classified by Bookstein [17] into three principal types of landmarks that are frequently used in morphometrics of biological organs. Typical features used for landmark selection are intersection among different anatomical structures (Type 1), maxima of curvatures, and centroids of closed-boundary regions (Type 2), and extrema on a contour (Type 3), etc.

In principle, the selection of landmark should spread over an entire target image or object and be from anatomic structure with classified features (Types 1, 2, and 3 landmarks). In addition, lingual arteries, nerves, and phonetic definition such as the tongue blade (the part closest to the upper alveolar ridge) can be candidates of landmarks even though they are used in the current study. The selected landmarks should be able to reflect individual tongue muscle contraction. The intended landmarks are listed in Table 1.

Insufficiency in anatomical details (e.g. boundaries between muscles) of the MRI data within the tongue pose difficulties in identifying landmarks. This difficulty is primarily caused by the course slice sampling in the vertical position in the MRI data. Therefore slice thickness of 2 mm were used during the period of data acquisition. In addition, a complete image set of all orientations (i.e. the coronal, the transverse, and the sagittal orientation), as a good source of information is needed for landmark identification and 3-d reconstruction. However, this step is not completed yet at this moment.

IV Conclusions

The intended work to be finished at the end of the project are: 1) a MRI database of 40 college students specialized for speech organs; 2) a final 3D tongue atlas; 3) definition and selection of 3D tongue landmarks; and 4) methods to evaluate the selected landmarks. Due to circumstances, we have acquired a MRI database of 40 college students specialized for speech organs. However, a final 3D tongue atlas construction is yet to be accomplished. Nevertheless, the current proposal will provide a substantive methodological preparation for future research, in the form of computational tools and example data, to investigate the phonetic characteristics of various articulations.

V References

[1] Lufkin, R.B.et al. (1986) *Radiology*, 161: 69-75.

[2] Mckenna, K.M., B.A, Jabour, R.B. Lufkin, and W.N. Hanafee (1990) *Topics of Magnetic Resonance Imaging*, 2: 49-59.

[3] Narayanan, S. S., A.A. Alwan, and K. Haker (1995) *J. Acoust. Soc. Am.*, 98: 1325-1347.

[4] Story, B., I. Titze, and E. A. Hoffman (1996) *J. Acoust. Soc. Am.*, 100: 537-554.

[5] Narayanan, S., D. Byrd, and A. Kaun (1999) *J. Acoust. Soc. Am.*, 106: 1993-2007.

[6] Stone, M. et al. (2001), *J. Acoust. Soc. Am.*, 109: 2974-2982.

[7] Takano, S. and K. Honda (2001) *Spring meeting, J. Acoust. Soc. Japan.*

[8] Wu, Chao-Min (1996) *Ph.D. thesis*, The Ohio State University, Columbus, OH.

[9] Wolberg, G. (1990) *Digital Image Warping.* IEEE Computer Society Press, Los Altos, CA.

[10] Toga, A. W. (1999) *Brain Warping*. Academic Press, San Diego, CA.

[11] Christensen, G.E, S.C. Joshi, and M.I. Miller (1997) *IEEE Trans. Med.Imaging*, 16:864-877.

[12] Thompson, M., R.P. Woods, M.S. Mega, and A.W. Toga (2000) *Human Brain Mapping*, 9:81-92.

[13] Wood, R.P., M.Dapretto, N.L. Sicotte, A.W. Toga, and J.C. Mazziotta (1999) *Human Brain Mapping*, 8:73-79.

[14] Missimer, J,U. Knorr, R.P. Maguire, H. Herzog, R.J. Seitz, L. Tellman, and K.L.Leenders (1999) *Human Brain Mapping*, 8:245-258.

[15] Fiez, J.A., H. Damasio, and T.J. Grabowski (2000) *Human Brain Mapping*, 9:192-211.

[16] 吳炤民(2003). *以磁振影像為基礎的立體舌頭圖譜之建構*。行政院國家科學委員會專題研究計畫成果報告。

[17] Bookstein, F.L. (1991) *Morphometric tools for landmark data*.Cambridge: Cambridge University Press, pp. 55–87.

Table 1 Landmarks for 3D tongues

Number	**Name**	**Type**	**Description**
1	Ant. Tip Pt.	3	Ant. most pt. on the tongue
2	SG-HG Pt.	1	Intersecting pt. of HG and SG
3	Ant. Apex of LS	1	Location where IL, HG, and SG unite
4	Pos. Apex of PS	1	Location where IL, GG, V & T unite
5	Lat. SL-VT Pt.	1 & 3	Location where SL, V &T meet laterally
6	Ant. SL-VT Pt.	1	Location where SL, V & T meet anteriorly
7	Pos. SL-VT Pt.	1	Location where SL, V & T meet posteriorly

8	Top End SL Pt.	2	Location of SL's most superoposterior pt.

Ant.: Anterior; Pos.: Posterior; Lat.: Lateral; Pt.: Point; LS: Lateral Septum; PS: Paramedian septum; V: Verticalis; T: Transversus; SL: Superior Longitudinalis; IL: Inferior Longitudinalis; GG: Genioglossus; HG: Hyoglossus; SG: Styloglossus muscle.

V 計畫成果自評：

本計畫之研究成果依計畫已蒐集四十位大專生（男女各半）的口腔 MRI 資料，可是我們尚未運用我們所蒐集到的 MRI 資料去初步建立一個 3D 的舌頭。針對前述之差距，我們有以下之說明：

1. 前述的差距因計畫主持人之前服務的學系中山醫大聽語系並無研究生（研究所），而且大學部學生（聽語系）並無具備適合執行此研究計畫的工程背景及軟體書寫的能力，因此無法由本系的學生來兼任，雖然在第二年聘用校內相關之醫學影像系之大學部學生來擔任學生助理。以為從此可以趕上進度，不過限於大學部學生能力與時間的限制，再加上計畫主持人自己在系上的教學負擔很重（4.5 門課/每學期），也因為如此在建立一個 3D 舌頭的研究進度上並無太大的突破。不過我們是如期蒐集四十位大專生（男女各半）的口腔 MRI 資料。正因為沒有適當的研究助理的協助，計畫主持人並沒有（也不敢）在 93 年度的提出任何國科會研究計畫。
2. 在我們原先的計畫中，我們並無意要發展新的 3D 重建的方法，我們只想借用適合我們 MRI 資料既有的方法，並運用到我們的資料。即使如此，我們仍需要具備工程背景及軟體書寫的能力的助理來協助。相信以計畫主持人目前所任教的國立中央大學電機系在教學負擔較輕且有五位研究生（四位碩士生及一位博士生）的環境下，會利用所蒐集到 MRI 資料將差距的部分逐漸完成。以磁振影像為基礎的立體舌頭圖譜之建構與相關說話機制之研究一直是計畫主持人努力的方向，希望不會因之前的進度落後而影響到後續研究的可能性，相信在目前具有更多與更好的研究環境下，定能在此研究上有進一步的突破，造福聽語障礙的族群。