

國立交通大學

機械工程學系

碩士論文

以模型為基礎語言合成系統之研究及應用

A Study and Application of
a Model-based Speech Synthesis System

研究生：劉憲鎗

指導教授：陳大潘博士

中華民國九十四年六月

以模型為基礎語言合成系統之研究及應用

A Study and Application of
a Model-based Speech Synthesis System

研究生：劉憲鎗

Student：Sian-Chiang Liu

指導教授：陳大潘

Advisor：Da-Pan Chen

國立交通大學



A Thesis

Submitted to Department of Mechanical Engineering
College of Engineering

National Chiao Tung University

In Partial Fulfillment of the Requirements

For the Degree of

Master of Science

in

Mechanical Engineering

June 2005

Hsinchu, Taiwan, Republic of China

中華民國九十四年六月

以模型為基礎語言合成系統之研究及應用

研 究 生：劉憲鎗

指 導 教 授：陳大潘

國立交通大學機械工程學系（研究所）碩士班

摘 要

在本論文中，以一個可變形的彈性體--臉部模型為基礎，藉著有限單元法計算該模型承受施力所產生的變形，得到不同的唇型變化；並觀察語言中子音、母音所表現的不同特徵，做為該模型的變形目標，藉由最小應變能法，達到變形量控制，最後組合聲母與韻母，以動畫的方式表現說話時唇型的變化與快慢。希望藉由這樣的方式，在動畫製作的領域有不同的發展與應用。

A Study and Application of a Model-based Speech Synthesis System

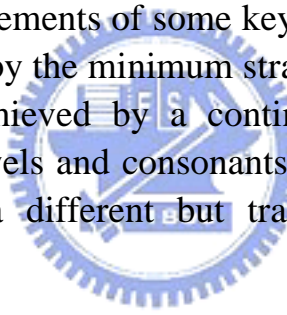
Student : Sian-Chiang Liu

Advisor : Da-Pan Chen

Department of Mechanical Engineering
National Chiao Tung University

Abstract

In this dissertation, we model a human face as an elastic thin shell structure of finite elements. We analyze the facial motion of speech by measuring the displacements of some key points around a lip and smooth it out to the entire face by the minimum strain energy principle. The animated human speech is achieved by a continuous interpolation between key frames of related vowels and consonants. This elastic-mechanical mode of human face offers a different but traceable means of human speech synthesis.



誌謝

感謝 神，引領不足的我完成這篇論文，透過指導教授陳大潘博士不斷地教導，使我瞭解研究中所要面對並解決的問題，在這段過程中的一直包容，也使我從原本的自我想法中脫離，定下心，認清自己的徬徨，拿著更肯定、更積極的心，更能主管自己的作為，走出自我設限，努力去挑戰、克服，才有如此成果，內心惟有不住地感謝。

感謝家人不斷地支持，使我能沒有擔憂地一路前進，感謝周遭所有朋友的鼓勵與扶持，加給我更多力量。爬過一山還有一山，在未來的人生中，自己還有許多要努力的，也有所當盡的責任，期許自己過著更踏實、更有價值的生活，使生命中的每一階段都有意義。



目 錄

| | |
|-------------------------|-----|
| 中文摘要 | i |
| 英文摘要 | ii |
| 誌謝 | iii |
| 目錄 | iv |
| 圖目錄 | vi |
| 一、緒論 | 1 |
| 二、文獻回顧 | 3 |
| 2.1 參數化的臉部模型 | 3 |
| 2.2 唇型的特徵 | 6 |
| 2.3 人類面部表情之有限元素分析 | 8 |
| 2.4 變形量控制 | 11 |
| 三、達成預定的變形 | 13 |
| 3.1 有限單元基礎 | 13 |
| 3.2 脸部模型的建立與相關設定 | 21 |
| 3.3 變形目標 | 29 |
| 3.4 模型的訓練 | 31 |
| 四、唇型的產生 | 34 |
| 4.1 外輪廓 | 34 |
| 4.2 內輪廓 | 37 |
| 4.3 下巴 | 39 |
| 4.4 臉形 | 40 |
| 4.5 其他母音、子音的結果 | 41 |
| 4.6 舌頭的搭配 | 43 |
| 五、動畫顯示 | 45 |

| | |
|---------------------|----|
| 5.1 動畫的產生 | 46 |
| 5.2 動畫播放的時間控制 | 47 |
| 六、結論 | 51 |
| 參考文獻 | 52 |



一、緒論

說話是由一連串複雜的動作所產生。從臉部肌肉的牽動、舌頭和牙齒的位置、聲帶的振動，發出不同的聲音。本論文的目標，即藉由一臉部模型（Facial Model）的視覺效果，表現出說話時嘴唇的不同樣式，如同人說話時所展現的特徵。唇語是一門由平時經驗的累積而用直覺判斷分析的學問，即使聽不到說話者的聲音，單單憑著觀察說話者嘴型的特徵變化，連貫前後的內容，大致上就可以瞭解說話的內容。因此，在人與人的談話中，除了聽覺之外，藉著觀察的結果，也可以是與人溝通上，很重要的一環。

基於觀察和分析的角度，在此，由一有限單元法（FINITE ELEMENT METHOD）建構的 Facial Model 上，表現出相同的變化模式。從力學的觀念--受力與變形量的關係，探討在嘴唇的不同部位，要施加多少力量，才能有預定、期盼的唇型變化。依照所觀察，量測出的變化量，或預定的變化量；計算出施加力量的大小和方向。所有的數據，是由不同視角所取得的實際變形投影量，作為此 3D 模型計算的依據。再根據這些量測得來或預定的數據，達到控制此 Facial Model 的目標。

不同於美工繪畫，每個美工都有自己風格的技巧和經驗，同樣的表情，從構圖到上色，成果不盡相同。在這裡，提供一個具有再現性

的方式，在操作上，透過部分關鍵點（key point）的變形量控制（著重在嘴唇的外輪廓上），即可獲得不同的表情。

接下來的章節，我們將做進一步的討論。

第二章--文獻回顧：由參數化模型的概念，不同參數設定代表不同的面貌，藉著比較真實說話狀態所表現的特徵(1)嘴巴的開合、(2)嘴巴的寬度及(3)嘴唇的突出，決定適當的參數值。另一方面，介紹如何運用有限單元法，操作臉部模型的變形。

第三章--預定目標的達成：操作有限單元臉譜模型，如何從部分的節點資料，獲得所要施加力量的大小與方向。藉著最小應變能量的方式，解決這部分的問題。

第四章--唇型的產生：觀察說話中聲母與韻母的基本特徵，產生語言合成所需的唇型。

第五章--動畫顯示：以動畫的方式，將一連串特定唇型的連續轉換過程，其中包含變化過程的產生和播放速度的控制。

第六章--結論

二、文獻回顧

2.1 參數化的臉部模型

在動畫製作中，電腦繪圖技術往往需要大量的美工技巧；在早期的 2-D 電腦動畫製作是以演算法產生影像；而在由 2-D 動畫往 3-D 動畫的發展上，key frame animation 面臨了資料量龐大的缺點。有別於 key frame animation 需要完整地訂定資料庫；參數化的臉部模型只需要改變一些參數設定，就可產生豐富變化的動畫。同樣地，一個完善的臉部模型，可以讓繪圖者只靠著適當的參數調配，給定適當的數值，就可以產生任何想要的臉部影像，而這些成果，只需要幾個參數的改變。



2.1.1 參數模型的基本概念

如果在一群組中的各個成員，都有著可辨識的不同，而這些差異可以找出一組相關聯的變數來表示，那麼，將變數轉換成可操控的參數，群組中的各成員，也都將有新的變化。一組完整的參數要使群組中的各個成員都有適切的參數值，這可能是離散的，也可能是有界的或是無界的連續數值。而最重要的，是要表現參數的唯一性，不同的參數組合，要排除有共同結果的可能性。

2.1.2 臉部參數

臉部模型的參數建立，也是依循相同的模式。最理想的情況，是找出一組完整的參數，可以產生不同的臉，而每一張臉孔都有著各式各樣的表情。但事實上卻不是如此盡善盡美，退而求其次，尋求品質上的適切及效率性即可。

參數的建立方式，最單純的方式是觀察臉部的表層特徵，進而將這些特徵數值化，做簡單的類比工作。其次，可以從臉部的解剖，由功能的觀點，瞭解每一臉部肌肉對表情變化的影響，建立表情變化的依據。混合以上兩種方式，結合外層和內部的結構，綜合出具有外觀及功能性的模型。如此，外在特徵的改變以及臉部的表情變化，均有相當的發揮空間。

以下，就參數的分類做介紹。

表情參數：最主要的部分在眼睛和嘴巴

- 眼睛部分—瞳孔擴張、眼瞼張開、眉毛擺動及視線。
- 嘴巴部分—嘴巴張開、嘴巴寬度、嘴角變化。

結構參數：

- 脖子的長度及形狀。
- 下巴、額頭、臉頰及顴骨的形狀。
- 眼瞼、眼球、虹膜及眼睛的位置。

- 鼻子長度、鼻梁寬度、鼻頭樣式。
- 各部位的比例關係。

藉由這個概念設計出的臉部模型，具有五類功能操作。

- (1) 眼球、瞳孔、虹膜的大小，虹膜的顏色，眼睛的位置等部位的參數操作，產生想要的眼睛。
- (2) 前額、顴骨、嘴巴及脖子等部位都有獨立的建構空間。
- (3) 下巴部位可以沿一軸線做轉動動作。
- (4) 臉部特徵的比例調控，下巴的突出，鼻子、嘴巴的大小等。
- (5) 嘴角、嘴唇的動作變化。

藉由以下幾張圖片，為這個模型做清楚的介紹[1]。



圖 2-1 眼瞼闔上的自然狀態



圖 2-2 眼瞼開張



圖 2-3 不滿的表情



圖 2-4 眼瞼張得更開、眉毛上揚、嘴寬縮窄、上唇抬起及瞳孔放大



圖 2-5 年輕化的臉



圖 2-6 較短的鼻子



圖 2-7 較短的下巴



圖 2-8 以改變後的模樣，重複圖 2-4 的操作

2.2 唇型的特徵

在大量的電腦動畫裡，說話是動畫人物除了肢體動作、表情變化外，經常表現的動作。比較嚴謹的，會考量真實與虛擬間的一致性；如何表現兩者的一致性，只有靠有系統地同化--收集大量說話時相關的表情變化，做為電腦動畫中，說話人物表現的依據。

在說話時嘴型的探討上，可以從較明顯的三個特性來訂定參數：(1) 嘴巴的開合，(2) 嘴唇的寬度及(3) 嘴唇的突出等。不同於表情可有較主觀的認定，考量個人說話的差異性，因此需有更客觀的比對；仔細量

測發音時的嘴型變化，得到各參數值的明確資料。接下來只需調整臉部模型的參數，就可以有同樣的動作表現。

接下來的幾張圖，可以看出不同發音的嘴型差異[2]。

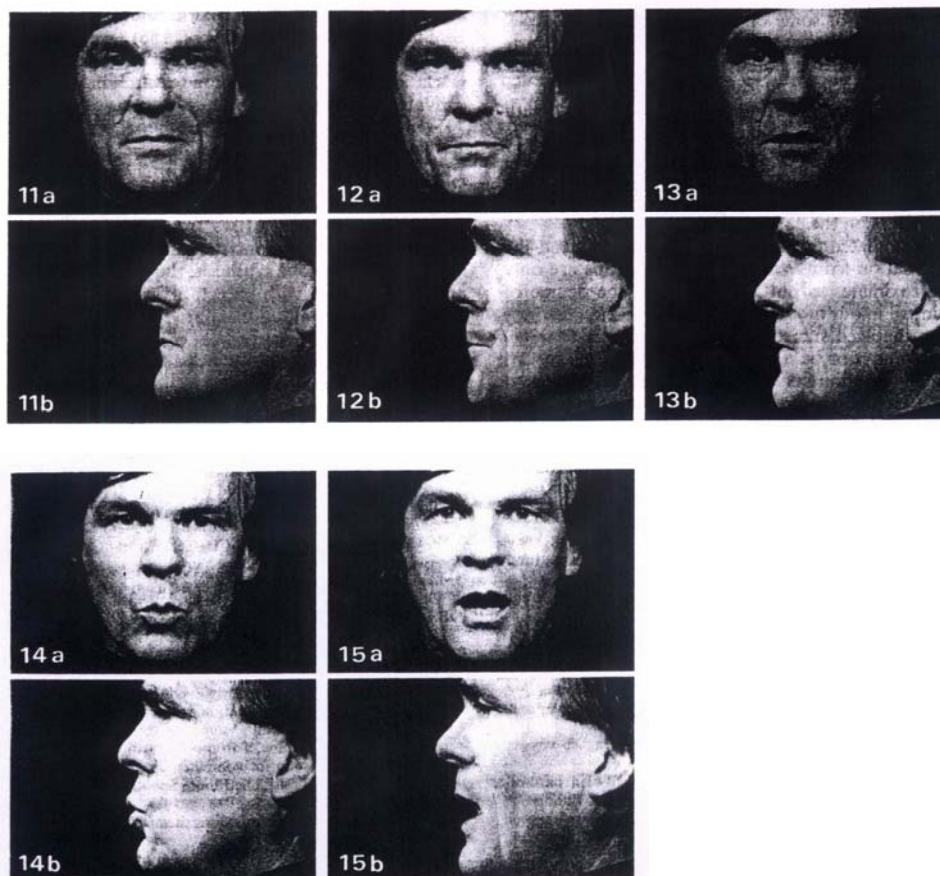


圖 2-9 不同發音的嘴型變化 (a 是前視圖，b 是側視圖)，其中--

11：無動作狀態。

12：/p/的發音。

13：/t/的發音，兩個子音只有側視圖中下唇的微小差異。

14：/u/的音，側視圖表現嘴唇的突出。

15：/a/的音，表現嘴唇的張大。

2.3 人類面部表情之有限元素分析

基於參數設定，改變參數值而產生的臉部模型，以有限單元法為基礎的臉部模型比較偏重在分析這方面。從解剖學的觀點，瞭解臉部肌肉的組成，相關的連結點、肌肉的運作模式；由力學的探討--受力與變形的關係，在肌肉與骨骼連結的地方，沿著肌肉紋理的方向，施加多大的力，會產生什麼樣的表現效果。

有限單元法的計算，在處理彈性物體的變形上，有不錯的功效，不論它的形狀如何、有多複雜。由每一個 element 的區域勁度矩陣，組合成整個系統的彈性勁度矩陣；給予適當的邊界條件，整個物體的變形就有下列的關係：

$$[K]\{D\} = \{R\} \quad (1)$$

其中，K 是整個系統的彈性勁度矩陣；D 是變形量；R 是施力。

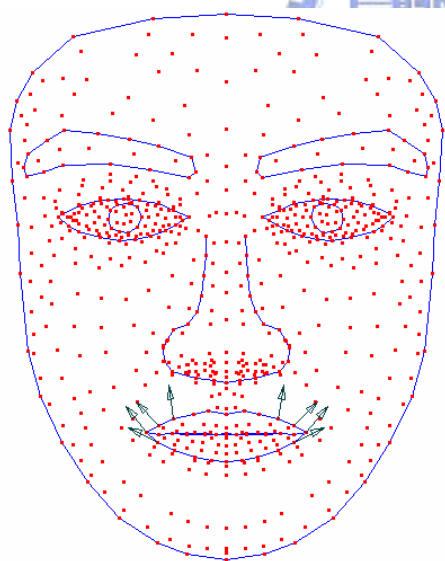
在這裡，整個系統選擇薄殼單元結構，它結合薄膜單元以及薄板單元的特性。其中，薄膜單元僅能承受沿著面方向的力，而薄板單元也僅能承受垂直於板面上的力和沿著面方向上的力矩。藉由兩者的結合，處理空間相關的問題。

另外，所必須提出的是，材料的所有特性，和真正的生理結構並不相同。皮膚、肌肉、其它的生理組織都不相干，只是做簡單的物理性的設定，如網格的剛度及 Poisson's ratio。

有別於電腦繪圖參數化臉部模型的參數設定，這些不同的臉部表情，從不同的節點施力，施力的方向、施力的大小，作用在有限單元模型上，做出各種不同程度的情緒表達；在經驗傳承上，可以比較有系統地延續下去。

觀察人類大部分的喜怒哀樂等情緒，大多都可以從臉部的表情變化看出端倪；例如，微笑的時候嘴角會上揚，盛怒的時候眉毛會上揚而眼睛會睜大，想哭的時候嘴角會下拉，大笑的時候嘴巴會張大而嘴角微微上揚，而眼睛容易眯成細細的下弦月，做鬼臉的時候表情更是千奇百怪。

接下來，以幾個操作的例子，表現臉部的表情變化，特別是在嘴巴動作的部分[3]。

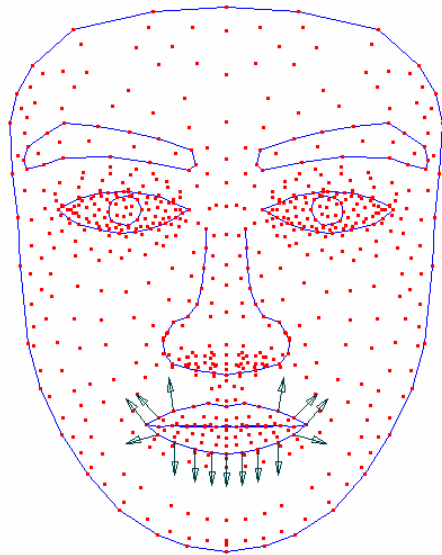


(a) 變形前



(b) 變形後

圖 2-10 微笑

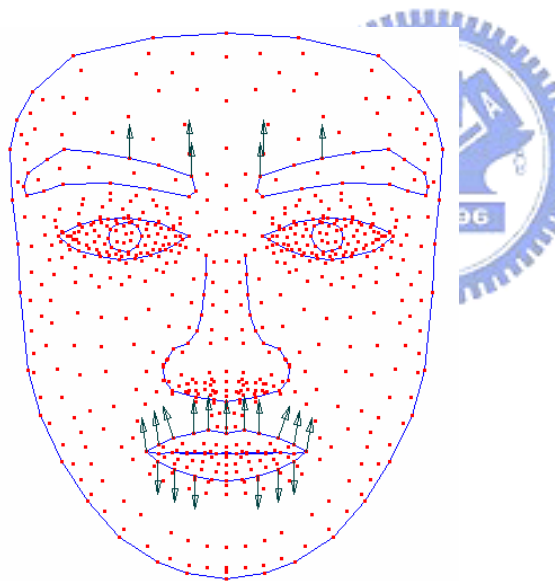


(a) 變形前



(b) 變形後

圖 2-11 大笑

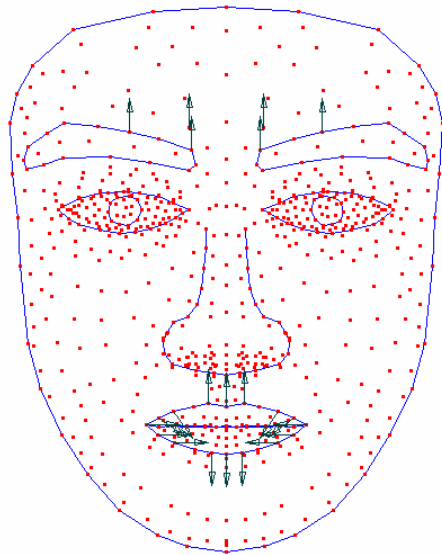


(a) 變形前



(b) 變形後

圖 2-12 吼叫



(a) 變形前



(b) 變形後

圖 2-13 驚訝

2.4 變形量的控制



一般而言，在有限單元法的求解上， K （系統的彈性勁度矩陣）依據邊界條件來決定， R （施力）是已知條件， D （變形量）是未知條件。只要給定 F 的大小和方向，即可求得 D 。如下：

$$\{D\} = [K]^{-1}\{R\} \quad (2)$$

若只是單純地想得到所有節點的變形量，上式(2)是非常便利的計算方式；不同的施力，將有不同的結果。相反地，只要有所有節點的變形量，當然也可以求得 R 。但是，當情況是只知道部分節點的變形量，亦即，上式(2)中，只有部分的變形量已知，又該如何求解？其關係式如下：

$$\{D_g\} = [P]\{R\} \quad (3)$$

其中， D_g 是已知變形量； P 是自 $[K]^{-1}$ 中，依據 D_g 在系統中的相對位置，整列取出。

在數學上，解方程式時，在已知條件少於未知數的數量時，會產生無限多組解的結果。而在這些結果之中，是否有合適的解，符合系統所需？文獻[4]當中，提出一個方法，以網格結構的最小應變能法（minimum strain energy solution）求得唯一的解，如下：

$$R' = P^T (PP^T)^{-1} D_g \quad (4)$$

雖然，以這個方式所得到的 R' ，施加於原系統所得到的變形量，在其他地方會有些許的誤差存在，可以保證的是，預定要達成的變形量是一致的。



三、達成預定的變形

經由前一個章節關於有限單元法的簡單介紹，在系統的節點上施力，可以做相關的變形量分析；如果，只是期望某些節點在變形之後，移動到特定的位置，我們可藉由嘗試的方式，經由輸入施力，計算輸出的結果，再判定是否是所期待的目標值。雖然，這個方式可以達到要求，但只能說是大略；首先，當目標不只一個，可能顧慮了這個，卻違背另一個，甚至更多，很難全面達到所有的要求。此外，當系統稍微複雜、稍大一些，可嘗試的方法也就更多，要歸納整理出一套可使人依循、重複操作的方式，就不是那麼容易了。基於上述的經驗，特別是難以找到一套有系統的操作方式的緣故，捨棄這樣的作法；尋求以控制變形量的方式，操作整個系統的變形。

在這一章節裡，介紹怎樣經由訓練的方式，得到我們所期望的模型變形；問題就在於如何更有系統、有規則性地得到一組施加於該系統的施力，使結果可以完全符合所要求的目標。相關的方法，在 2.4 變形量的控制中，已經大略地提出；運用最小應變能量的方法，藉著部分節點的變形量作為限制條件，計算出相對應的施力；不管限定條件的多少，其餘未經限定的節點變形量，也可一併得到。接下來，將一一說明整個流程。

3.1 有限單元基礎

有限單元法在使用上，運用的非常廣泛，資料非常充足且龐大。這裡只是提出一部份，針對我們要處理的力與變形的問題，做相關的介紹，如何在空間問題上，處理應力與應變之間的關係。

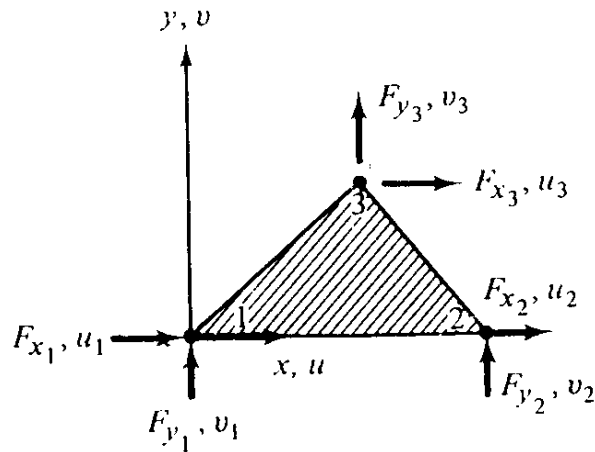


圖 3-1 CST 示意圖

3.1.2 薄板結構—平面彎矩問題

薄板結構僅能承受垂直於板面上的力及沿著面方向上的力矩，並因厚度遠小於長寬，所以不處理有關沿著面方向上的力；假設所處理的薄板面被設定為XY面，那受力的方向就僅能考慮Z方向上的施力，再考慮彎曲力矩對薄板的影響，則需加上對X軸及Y軸的彎曲力矩；這將把三度空間中每個點的六個自由度縮減成三個自由度 (w, θ_x, θ_y) ，因此可以知道變形將發生在Z方向上的位移與對X、Y軸的旋轉量。

在本論文中，這個部分使用DKT (Discrete Kirchhoff Triangle)，參考圖3-2，其負載與位移的關係可參考(5)式，其中：

$$\{d\}_i = [w_1 \ \theta_{x1} \ \theta_{y1} \ w_2 \ \theta_{x2} \ \theta_{y2} \ w_3 \ \theta_{x3} \ \theta_{y3}]^T \quad (9)$$

$$\{r\}_i = [fz_1 \ mx_1 \ my_1 \ fz_2 \ mx_2 \ my_2 \ fz_3 \ mx_3 \ my_3]^T \quad (10)$$

在這部分，區域勁度矩陣的產生方法比較複雜，在文獻[3]的附錄一當中附上計算區域勁度矩陣的程式碼，從那裡可以了解整個矩陣產生的流程。而在上兩式中， w_n 是第i個單元中，各點在單元座標z向的位移量； fz_n 為第i個單元中，各點在單元座標z向上的負載； θ_{x_n} 、 θ_{y_n} 是第

i 個單元中，各點對單元座標 x 及 y 軸的旋轉量； m_{x_n} 、 m_{y_n} 為第 i 個單元中，各點對單元座標 x 及 y 軸的力矩。

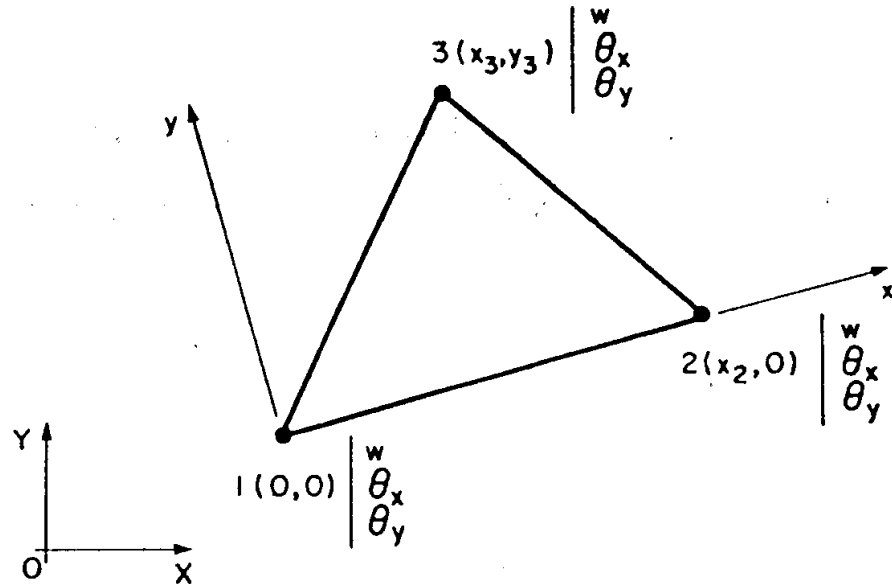


圖 3-2 DKT 示意圖

3.1.3 組合結構

當我們在處理空間中的一平板結構體時，它也許有著薄膜結構的特性，也有可能兼具薄板結構的特徵，所以為了處理這方面的問題，也就發展出一套結合薄膜及薄板特性的“平板結構”，又稱為“薄殼結構”。而在前面的討論中，我們可以發現 CST 的問題只考慮了 (u, v) 兩個自由度，DKT 的問題也只考慮了 (w, θ_x, θ_y) 三個自由度，但是在空間問題中，我們卻必須考慮六個自由度，因此我們必須加上一個式子來補足對所有自由度的考慮。(P.352, [5])

$$\begin{Bmatrix} m_{z_1} \\ m_{z_2} \\ m_{z_3} \end{Bmatrix} = EtA \mu \begin{bmatrix} 1 & -0.5 & -0.5 \\ -0.5 & 1 & -0.5 \\ -0.5 & -0.5 & 1 \end{bmatrix} \begin{Bmatrix} \theta_{z_1} \\ \theta_{z_2} \\ \theta_{z_3} \end{Bmatrix} \quad (11)$$

最後的三角單元中，其力與位移的關係如下：

$$\{d\}_i = [u_1 \quad v_1 \quad w_1 \quad \theta_{x_1} \quad \theta_{y_1} \quad \theta_{z_1} \quad u_2 \quad v_2 \quad w_2 \quad \theta_{x_2} \quad \theta_{y_2} \quad \theta_{z_2} \\ u_3 \quad v_3 \quad w_3 \quad \theta_{x_3} \quad \theta_{y_3} \quad \theta_{z_3}]^T \quad (12)$$

$$\{r\}_i = [fx_1 \quad fy_1 \quad fz_1 \quad mx_1 \quad my_1 \quad mz_1 \quad fx_2 \quad fy_2 \quad fz_2 \quad mx_2 \quad my_2 \quad mz_2 \\ fx_3 \quad fy_3 \quad fz_3 \quad mx_3 \quad my_3 \quad mz_3]^T \quad (13)$$

$[k]_i$ 擴充為 18x18 的對稱矩陣。

3.1.4 勁度矩陣的組合

在 3.1.1 到 3.1.3 節所提到的勁度矩陣都是屬於區域性的，其使用的點資料也都是在區域性座標系統的點資料，而真正在結構上的點資料，則必須經過一個旋轉矩陣[A]，其中 $\{d\}=[A]\{D\}$ ， $\{d\}$ 代表區域座標點資料， $\{D\}$ 代表系統實際的點資料。座標旋轉的目的在於配合區域性勁度矩陣的計算，我們必須先將空間中任意位置與任意方向的三角單元旋轉與平移，在計算勁度矩陣之前，必須讓三角單元三個頂點的座標依次為(0,0,0)、(a,0,0)、(b,c,0)，其中 $a>0, c>0$ 。之後才將旋轉後的點資料帶入來計算區域性勁度矩陣，而區域性勁度矩陣 $[k]_i$ 和全域勁度矩陣[K]有著以下的關係：

$$[K] = \sum_{i=1}^n [A]^T [k]_i [A] \quad (14)$$

而關於勁度矩陣的組合，我們可以先參考圖 3-3，圖中顯示每個單元(Element)的頂點都有各自的編號 i,j,k，其相當於前幾章所提到的三角單元的頂點 1,2,3，現在我們用這個簡單的例子來說明勁度矩陣的組合。

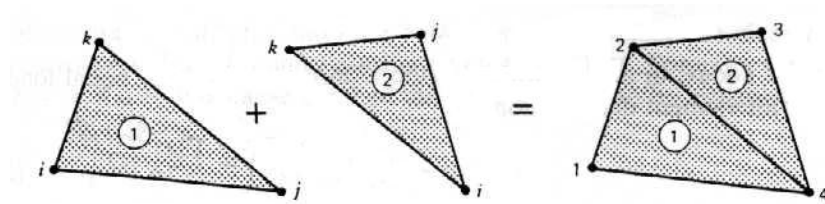


圖 3-3 三角單元組合示意圖

假設圖 3-3 裡的三角單元每個點都只有一個自由度，則每個勁度矩陣大小就是 3×3 ，則我們可以得到這樣的關係：

$$[k]_1 \{d\}_1 = \begin{bmatrix} a_1 & a_2 & a_3 \\ a_4 & a_5 & a_6 \\ a_7 & a_8 & a_9 \end{bmatrix} \begin{Bmatrix} d_i \\ d_j \\ d_k \end{Bmatrix}, \quad [k]_2 \{d\}_2 = \begin{bmatrix} b_1 & b_2 & b_3 \\ b_4 & b_5 & b_6 \\ b_7 & b_8 & b_9 \end{bmatrix} \begin{Bmatrix} d_i \\ d_j \\ d_k \end{Bmatrix} \quad (15)$$

現在我們考慮編號 1 的三角單元，並分別用區域和全域的編號來代入 $\{r\} = [k]\{d\}$ 這個公式，我們可以得到

$$\begin{aligned} r_i &= a_1 d_i + a_2 d_j + a_3 d_k & r_1 &= a_1 D_1 + a_2 D_4 + a_3 D_2 \\ r_j &= a_4 d_i + a_5 d_j + a_6 d_k & r_4 &= a_4 D_1 + a_5 D_4 + a_6 D_2 \\ r_k &= a_7 d_i + a_8 d_j + a_9 d_k & r_2 &= a_7 D_1 + a_8 D_4 + a_9 D_2 \end{aligned} \quad (16)$$

由於節點 3 與編號 1 的三角單元沒有接觸，再令 $r_3 = 0$ ，增加了這項之後，(16)式可以改寫成：

$$\begin{Bmatrix} r_1 \\ r_2 \\ r_3 \\ r_4 \end{Bmatrix} = \begin{bmatrix} a_1 & a_3 & 0 & a_2 \\ a_7 & a_9 & 0 & a_8 \\ 0 & 0 & 0 & 0 \\ a_4 & a_6 & 0 & a_5 \end{bmatrix} \begin{Bmatrix} D_1 \\ D_2 \\ D_3 \\ D_4 \end{Bmatrix} \quad (17)$$

(17)式中的方陣代表的就是 $[k]_1$ ，同樣的編號 2 的單元也可以做相同的運算，然後我們就可以得到兩的大小相同且對同樣的向量 $\{D\}$ 做運算的兩個矩陣 $[k]_1$ 和 $[k]_2$ ，並且可以推演出 $[K]\{D\} = (\sum [k])\{D\}$ ，其中：

$$[K] = [k]_1 + [k]_2 = \begin{bmatrix} a_1 & a_3 & 0 & a_2 \\ a_7 & a_9 & 0 & a_8 \\ 0 & 0 & 0 & 0 \\ a_4 & a_6 & 0 & a_5 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & b_9 & b_8 & b_7 \\ 0 & b_6 & b_5 & b_4 \\ 0 & b_3 & b_2 & b_1 \end{bmatrix} \quad (18)$$

從(18)式我們可以觀察出一般三角單元的勁度矩陣組合方式。同樣

的，即使組合的並非純粹的三角單元，只要找出對應的節點，就可以做類似的運算，就像圖 3-4 所顯示的一樣。

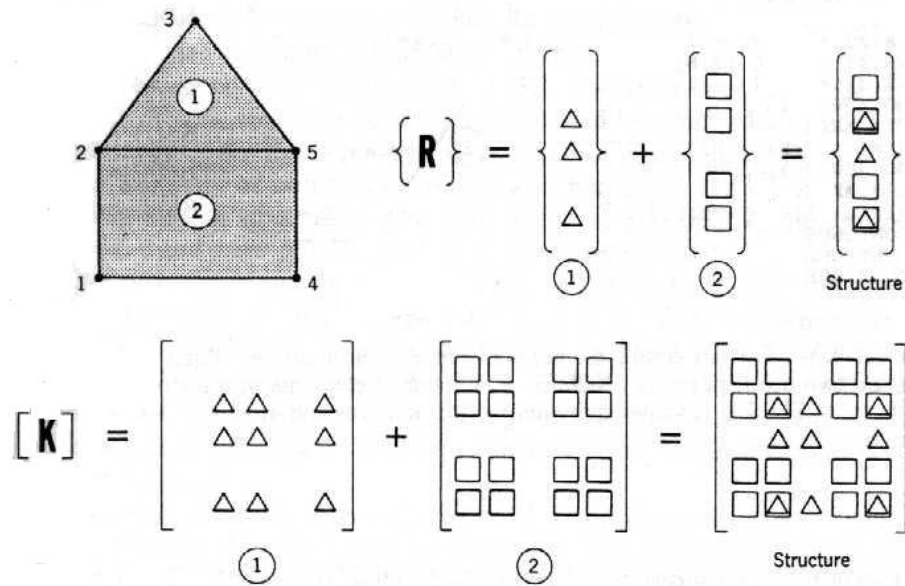


圖 3-4 不同單元組合示意圖

3.1.5 ID Array



在介紹完勁度矩陣之後，我們要開始對計算方法加以討論。

首先，在處理 $\{R\} = [K]\{D\}$ 這個問題中，我們必須知道，如果我們不給予結構體邊界條件的限制，則當我們施予外力的時候，結構體將會產生剛體運動而不是變形，而在本篇論文中，我們是考慮施予外力來產生所期盼的變形，因此我們必須限制結構體若干的自由度，而限制的方法通常就是指定某些節點的變形量（通常是將變形量指定為零）。

當我們利用 $\{D\} = [K]^{-1}\{R\}$ 來求得所期望的變形量時，由於 $\{D\}$ 的某些成員是已知的，所以我們不應該將這些已知的成員放入 $\{D\} = [K]^{-1}\{R\}$ 這個計算式中，假設原來有 n 個自由度，則 $\{D\}$ 與 $\{R\}$ 的大小是 $n \times 1$ ， $[K]$ 的大小則是 $n \times n$ ，如果我們限制了 m 個自由度，則新的 $\{D\}$ 與 $\{R\}$ 的大

小將會是 $(n-m) \times 1$ ，[K]的大小則會是 $(n-m) \times (n-m)$ ，也就是我們必須將陣列中所有被限制的成員移除。方法是直接宣告新的陣列，而新陣列的大小與成員的對應位置就由 ID Array 來決定。

假設我們要處理一個具有八個點的結構，而每個點有六個自由度，則我們應該宣告一個 6×8 的 ID Array，根據邊界條件，如果自由度被限制則在 ID Array 的對應位置填入 0，如果沒有被限制則填入 1，如(19)式所示。在這個陣列中，有 1 的部分是我們需要計算的，為了知道新陣列的成員該如何組成，我們必須給予編號，也就是我們要將(19)式中的 1 照順序給予編號，如(20)式所示。在這樣的編號之後，我們就可以知道{D}與{R}將會是 7×1 的陣列，而[K]則會是 7×7 的陣列。

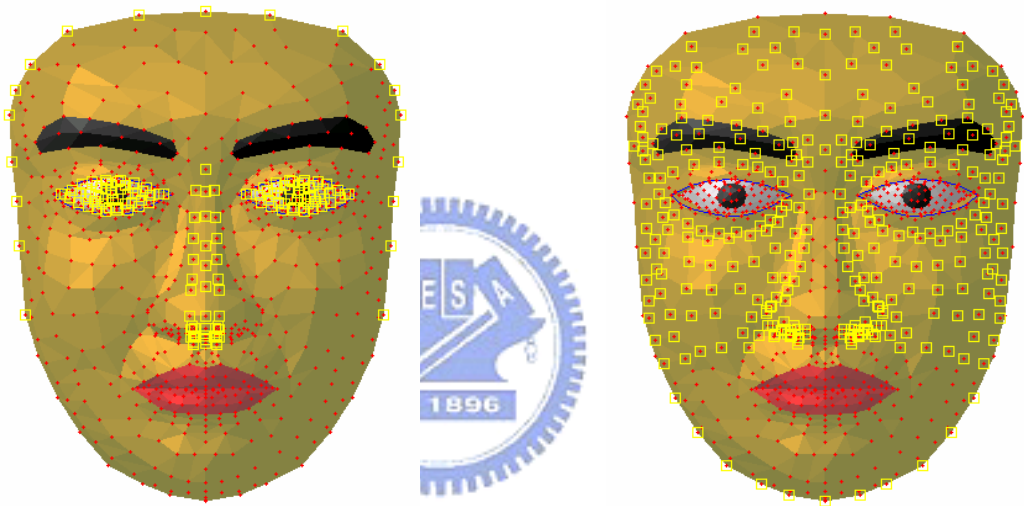
$$\begin{aligned}
 [ID] &= \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{matrix} u \\ v \\ w \\ \theta_x \\ \theta_y \\ \theta_z \end{matrix} \quad (19)
 \end{aligned}$$

$$\begin{aligned}
 [ID] &= \begin{bmatrix} 1 & 2 & 0 & 0 & 0 & 5 & 0 & 7 \\ 0 & 3 & 4 & 0 & 0 & 0 & 6 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{matrix} u \\ v \\ w \\ \theta_x \\ \theta_y \\ \theta_z \end{matrix} \quad (20)
 \end{aligned}$$

以這個例子來說，如果我們以 $R_{(i,j)}$ 、 $D_{(i,j)}$ 、 $K_{(i,j),(i,j)}$ 來代表原來陣列的成員，其中 i 代表 Node 的編號(1~8)，j 代表自由度的編號(1~6)，則新陣列應該如(21)式所示，以這樣的規則就可以處理所有問題了。

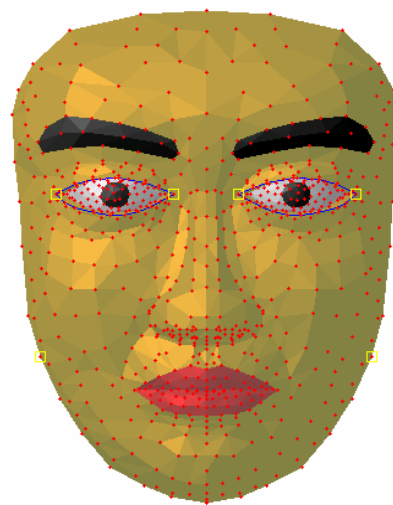
3.2.1 邊界條件

在圖 3-5 中，(a)所選取的點代表鼻樑、眼球及下頷骨以上的臉皮底部，是設定為完全不動的部分，固定所有的自由度；(b)所選取的點代表下頷骨以上，眼皮和(a)所選取的點以外的部分，以及下頷骨以下的臉皮底部，設定為不做深度方面的變形，也就是固定 Z 向的自由度；(c)在眼角部位固定 X、Y、Z 向的自由度。(本論文中以右手邊為 X 軸正向，上面為 Y 軸正向，出紙面為 Z 軸正向，採右手座標系)



(a)限制所有自由度的點

(b)限制 Z 向自由度的點



(c)限制 X、Y、Z 向自由度的點

圖 3-5 邊界限制

在邊界條件的上，大多是按照原先的設定，只有少部分變動。要特別說明的是，整個有限單元模型被分為上半部與下半部。在上半部，本應有頭骨做為支撐，限制部分方向的運動；雖然，在實際上，臉的上半部皮膚也是可以有向外隆起的能力，但為了維持臉部的外觀、輪廓，選擇限制上半部的Z向自由度。至於下半部，因有下頷骨的運動，並不作特別的限制。

3.2.2 網格特性

在網格特性上，大多仍是依循原先的設定，不去變動；有部分區域為配合使用上的需要，則做了一些變更。原先的設定中，蒲松比定為 0.3，這方面不做變動；關於楊氏係數，大致上分為眼睛、嘴巴及其他部位等三部分，絕大部分的區域都定為 100 psi。

而在使用的過程中，考量到嘴唇開合的問題，臉部上，皮膚能移動的範圍各有不同，遂分成三部分--臉頰、腮及下巴。臉頰和下巴部位因皮膚附著於頭骨之上，和骨頭間比較少相對運動，但特別的是，下巴部位的下頷骨會繞著頭骨轉動；而處於臉頰和下巴之間的腮，則隨著下頷骨的開合而伸縮。因應三者之間不同的變形趨勢，變更彼此之間相對的楊氏係數，作為分別。經由三者之間的關係性，將腮和下巴部位的楊氏係數作調整，臉頰和其他部位則維持不變，當成是比較的標準。

如同之前提到的，本系統所使用的有限單元模型，缺少頭骨結構和可轉動的下頷骨。雖然，無法使系統達到轉動的要求，但可以換一個方式，使下巴部位的楊氏係數變大，剛性比其餘部位更強，變形時下巴部位節點間彼此的差異性縮小，更加一致，圖 3-6(a)所顯示，線段所圍的區域，楊氏係數定為 1000。相對之下，腮的部位，是屬於較柔

軟的部位，受上下兩端力的拉扯，有較明顯的變化產生，在圖 3-6(b)所顯示，線段所圍的區域，楊氏係數定為 30。

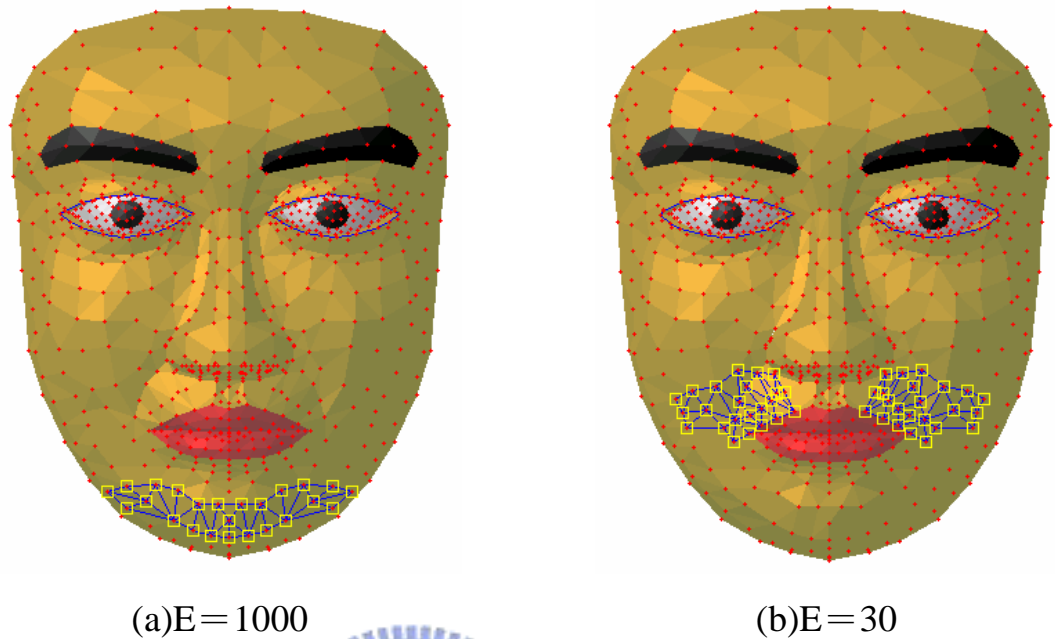


圖 3-6 楊氏係數的設定

從上面的圖可看出，更改網格楊氏係數的位置，並沒有完整包含整個所提到的區域。最主要的原因在於，本模型所涵蓋的範圍是『臉』的部分，並不是整個『頭部』，所以在靠近邊界的地方，保留部分空間，增加操控上的彈性。更加看重的目標是，為了使此模型的變形，在表現上有更好的效果，與真實的說話動作更相符合而所做的調整。這也是因應模型操作上所面臨的問題，相關的調整，可以使嘴唇的外輪廓在變形之後，看起來與真實更相似。

藉由接下來的操作，來顯示調整前後的差異。由原先的設定分別與圖 3-6 調整楊氏係數後的兩個模型(a)和(b)做比較。首先，在下唇緣的節點上，如圖 3-7，都給定相同的施力 $(0,-10,-10)$ 。

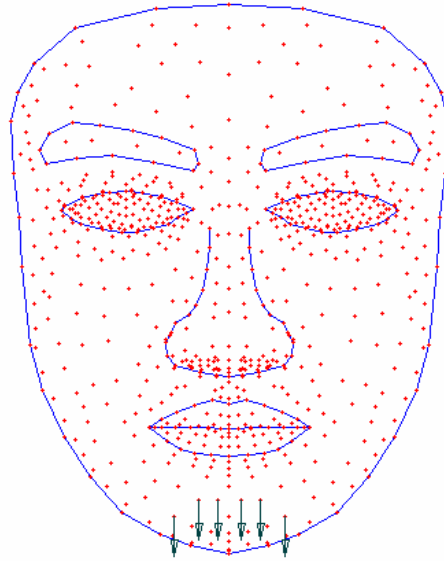


圖 3-7 施力節點的位置

在操作原本的模型時，首先發現，因上唇與下唇之間嘴裂的存在，使得嘴裂以下的區域，不像其他區域的網格，有相連接的網格；在缺少足夠支持的情況下，很容易因負 Y 軸方向的力，產生明顯的變形。情況如同在兩端點被固定住的一根樑，因橫向的力所產生的變形；同樣大小的力，在愈接近樑的中點位置施力，變形量也就愈大。而所得到的變形，特別是下唇緣的輪廓，和所期待的目標有很大的誤差。因為先要模擬的是因下巴向下，嘴唇張開的狀況，所以將力施加在下巴位置的幾個節點上，也因為下巴是一體的狀態，將施力假設為均佈力。在下巴部位的網格楊氏係數增大之後，可以整體帶動下巴，如此所得到的結果--下唇緣的輪廓，在嘴唇張大的同時，不會有太大的曲率變化。

另外，在比較上唇和下唇高度上的比例（Y 軸方向的長度）也發現，通常的狀況應當是兩者的比例值很接近 1；卻因上唇所能張開的高度不足，且為要修正這樣的結果，需要在局部給更大的力來調整。為要使上唇能有更大的高度變化，將上唇上方區域的網格調整為較小的楊氏係數。如此，在相同的施力下，上唇能有更大的高度變化，而對

下唇的高度變化沒有什麼影響。

藉著接下來的圖，一一解說所得到的改變。

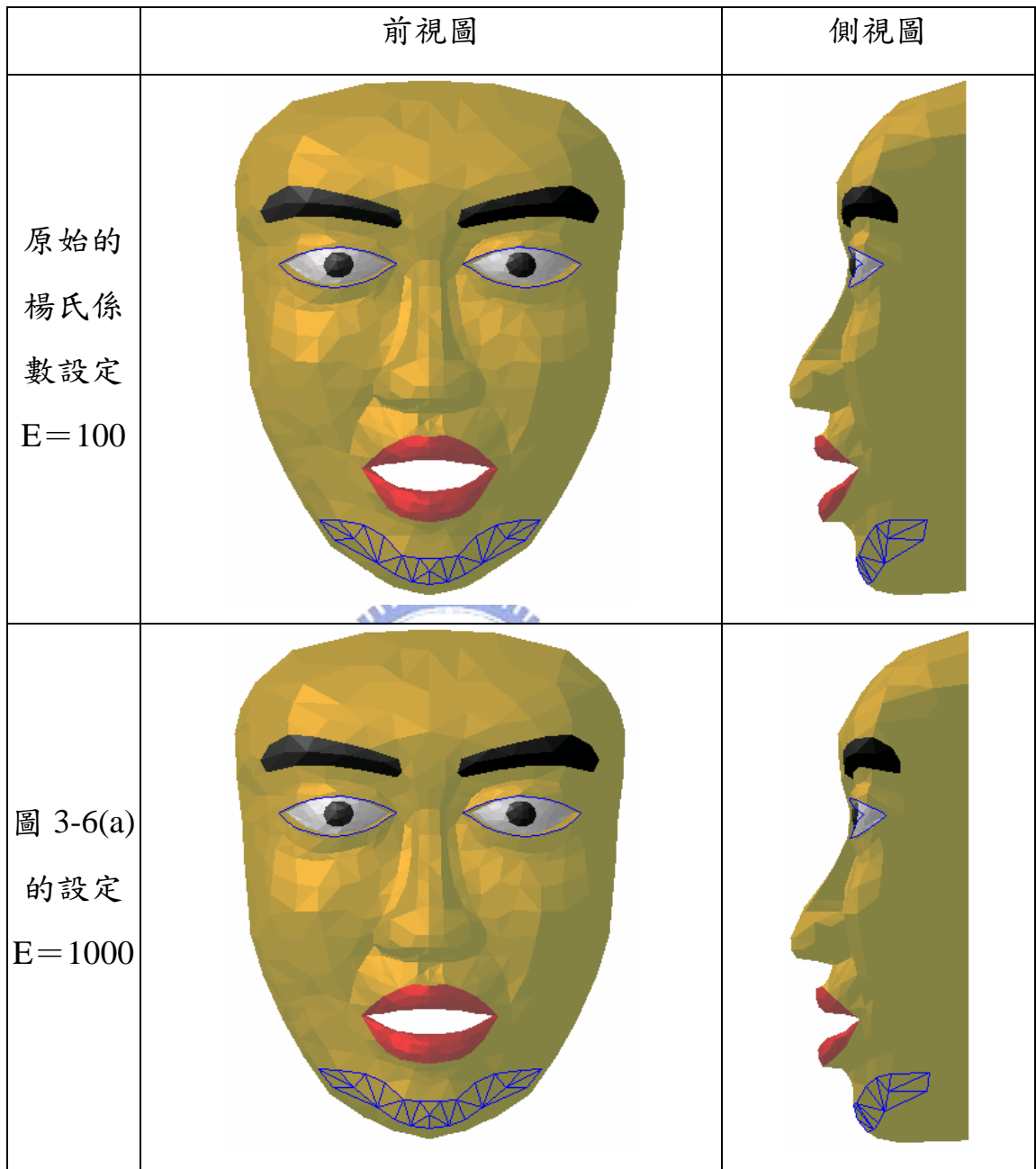


圖 3-8 原始設定與圖 3-6(a)模型的比較

原先的設定中，隨著嘴唇的愈張愈大，下唇所佔的比例會愈來愈大，下唇緣的輪廓則顯得過於突出。由圖 3-8 上下兩組圖片可比較出，下巴部位的楊氏係數調整，可改善這個問題，不會使變形量過於集中在下唇的中央。

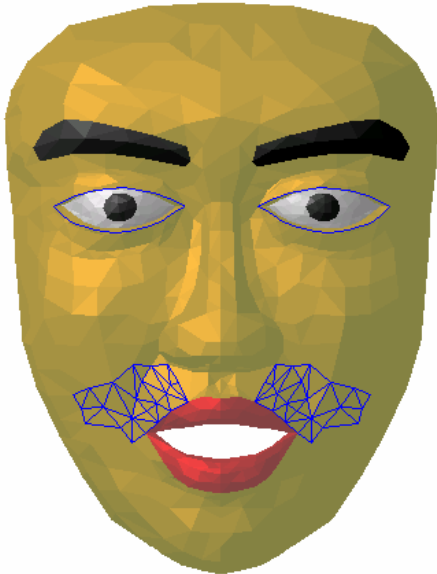

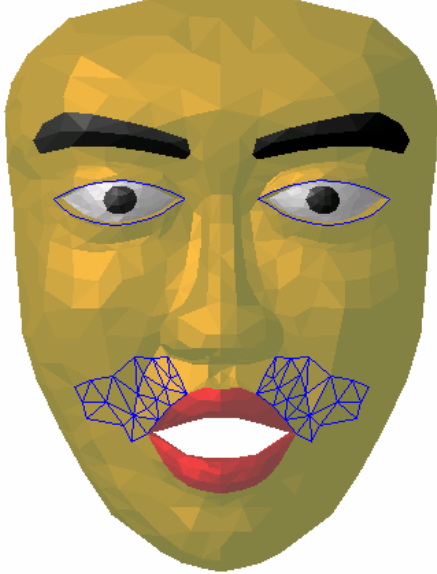

| | 前視圖 | 側視圖 |
|------------------------------|--|--|
| 原始的 楊氏係 數設定 $E=100$ |  |  |
| 圖 3-6(b) 的設定 $E=30$ |  |  |

圖 3-9 原始設定與圖 3-6(b)模型的比較

在圖 3-9 這一部份，在改變楊氏係數之後，使得上唇能張得更大，也較容易調整嘴角所張開的角度，比原先顯得更柔軟。

綜合圖 3-6(a)、(b)兩個模型的設定，得到新的模型，結合了為上、下兩唇所做的改變，如下圖 3-10 所做的比較：





| | 前視圖 | 側視圖 |
|----------|--|--|
| 原始 設定 |  |  |
| 綜合 模型 |  |  |

圖 3-10 原始設定與綜合模型的比較

當然，整個系統更深入去探討，要更改的還有，不只是以上所提到的這些。實際上，人的臉部還要更複雜，肌肉動作的細膩度很難去模擬，尤其是嘴唇的部分，分為內、外兩層肌肉，靈活度更高。只提出這部分的主要原因是，因下頷骨轉動所帶動的嘴唇張開，是所有說話動作的根本；先從這裡改善，簡化系統在操作上的要求，至於嘴唇形狀的控制，則配合嘴唇周邊的節點施力而調整。

3.3 變形目標

經由前面章節的準備，得到整個系統的勁度矩陣後，開始訓練此臉部模型的變形。同時，另一個步驟是必須設定變形目標，這部分完全決定在操作者的手上；可以是一個想要去模擬的對象，也可以取決於自己的設計。重點是使模型和對象體之間要有關連性，從結構上來看，在對象體和模型相對應的節點位置上，必須有一組點資料可當參考；依據這組點資料的位置變化，成為模型的變形目標。想要模擬得更完整，所要取的點資料也就愈多；理論上點資料的設定可以和模型的節點一樣多，可以達到完全近似的效果，但如此做很花費功夫。通常，需要進行模擬的模型所包含的節點不在少數，要一一設定這些相對應的點資料，反而失去模擬的意義。

本論文的主題是語言合成系統之研究，變形目標的優先選擇是實際上的說話動作。在這部分，採用數位相機的影像輸出，擷取由靜止狀態到動作完成狀態的過程，再從中比較點資料的位移量。

3.3.1 點資料

因著嘴唇動作是我們研究的重點，所選取的點資料自然而然以嘴唇及周邊區域做為對象；為了方便觀察，嘴唇與臉部交界的上、下唇緣是我們最好的觀察對象，有著視覺上的便利性，藉著兩者間顏色的不同，產生非常明顯的輪廓。而另一方面，為了之後量測位移量的工作，必須要有一個基準點；固定不動的條件對臉部而言，是很難達到的要求，退而求其次，選擇鼻頭這個點，比較少有動作變化，位置也較顯著。

只是這樣，資料仍然不足，除了嘴唇的形狀之外，下巴張開是主

要的动作，它的位置也要列入观察范围。不同於之前所提到的點，這部分面臨不易觀察的困難，比較可行的方式是從側邊看，有一個向前的突出，取當中的最高點。

綜合前面所述，基準點 1 個，上、下唇緣共 20 點，及下巴的參考點 1 個，共有 22 個點。他們的位置如下圖 3-11 所示。

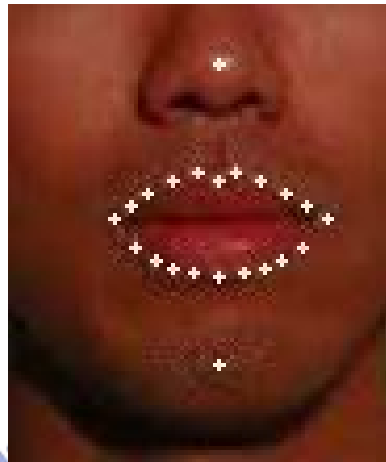


圖 3-11 點資料的選取

3.3.2 位移量

關於有限單元模型的討論中，每一個節點都是 6 個自由度，其中 3 個沿著座標軸方向的位移，其餘 3 個繞著座標軸的轉動。沿著座標軸方向的位移，可從照片中去比較、測量，但轉動的部分，並不是能力所及，放棄這方面的測量；為訓練模型所給的資料，將只限定在 (u, v, w) 。此外，一個方向所拍攝的照片，從平面 2D 的資料中，只能獲得其中 2 項，必須再從不同的方向獲得更多資料；由正視圖的資料，可以得到 (u, v) ，由側視圖的資料，可以得到 (v, w) ，從兩個方向去整合，完整得到所需的資訊，如下圖 3-12 所示。

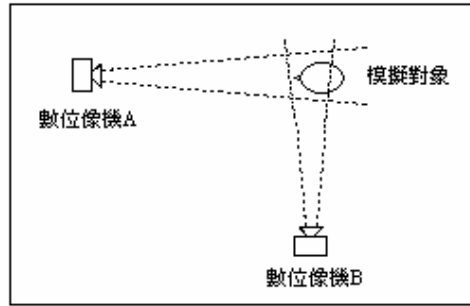


圖 3-12 取樣示意圖

在拍攝照片的過程中，若只是某一個狀態，只要對象體保持靜止不動，拍攝時間的先後不會有影響。但為了得到整個動作的變化過程時，如何達到同步化，還有待克服；雖然，以相同型號的數位相機，設定上可以一致，但操作上的時間差控制，仍略嫌粗糙。除了同步化的問題，對象體的姿勢，特別是頭的動作，難免有一些不自覺的擺動，或者前後，或者左右，甚至轉動，對位移量的測量結果，存在程度不一的誤差，必須去注意。

3.4 模型的訓練

經由 3.1 到 3.3 節的內容，完成所有的準備工作，接著進入資料整合的程序。在這裡，藉由圖 3-13 工作流程圖的表達，更清楚瞭解接下來的工作。

首先是在計算整個系統的勁度矩陣之後，經由高斯消去法得到它的反矩陣；這部分的運算雖然簡單，但因著矩陣龐大，花費最多時間。接著回顧 2.4 節變形量的控制中方程式(2)和(3)，在反矩陣中，藉由變形目標的條件限制，建立矩陣 P。建立的方式，要以 3.1.5 節 ID Array 中的(20)式所做的順序編號，找到對應的關係，假設經重新編排後(2)式如下：

$$\begin{Bmatrix} d_1 \\ d_2 \\ d_3 \\ d_4 \\ d_5 \\ d_6 \\ d_7 \end{Bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} & a_{17} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} & a_{26} & a_{27} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} & a_{36} & a_{37} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} & a_{46} & a_{47} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} & a_{56} & a_{57} \\ a_{61} & a_{62} & a_{63} & a_{64} & a_{65} & a_{66} & a_{67} \\ a_{71} & a_{72} & a_{73} & a_{74} & a_{75} & a_{76} & a_{77} \end{bmatrix} \begin{Bmatrix} r_1 \\ r_2 \\ r_3 \\ r_4 \\ r_5 \\ r_6 \\ r_7 \end{Bmatrix} \quad (22)$$

這時若在第 2、5、6 號位置給定變形目標，則(3)式的關係如下：

$$\{D_g\} = \begin{Bmatrix} d_2 \\ d_5 \\ d_6 \end{Bmatrix} = \begin{bmatrix} a_{21} & a_{22} & a_{23} & a_{24} & a_{25} & a_{26} & a_{27} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} & a_{56} & a_{57} \\ a_{61} & a_{62} & a_{63} & a_{64} & a_{65} & a_{66} & a_{67} \end{bmatrix} \begin{Bmatrix} r_1 \\ r_2 \\ r_3 \\ r_4 \\ r_5 \\ r_6 \\ r_7 \end{Bmatrix} \quad (23)$$

矩陣 P 依變形目標在系統中的位置，整列取出反矩陣中相對應位置的成員而建立。

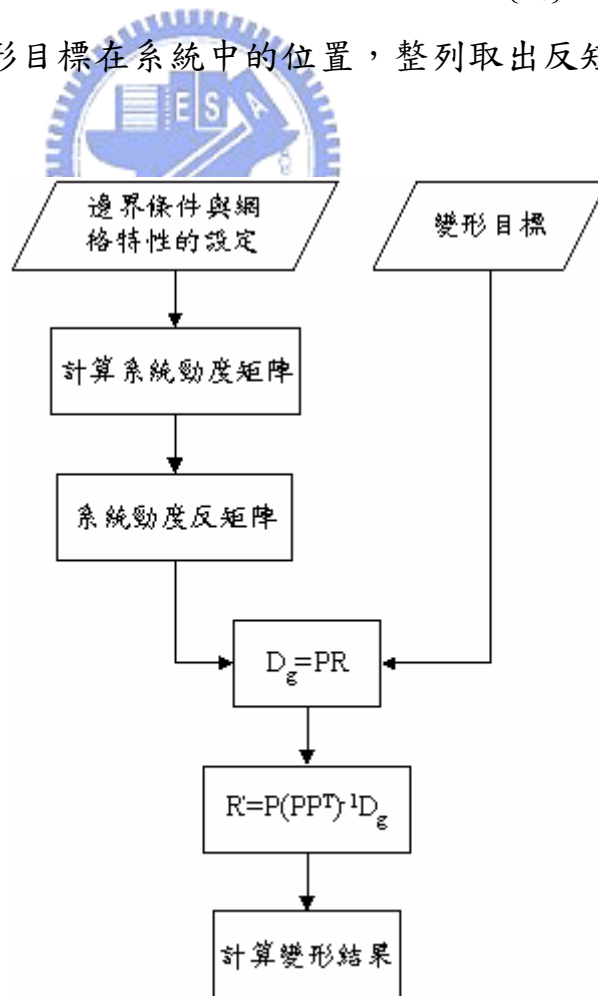


圖 3-13 工作流程圖

接下來純粹是計算的步驟，經由(4)式的矩陣運算，得到所需要的施力，最後以這組力施加於模型上，獲得變形結果。重複上述的程序，給定不同的變形目標，進行訓練模型的工作，產生各個變形圖。



四、唇型的產生

接下來，我們進入最後的準備的工作。藉由唇語(Lip-reading)的技巧，不只是可用來分辨對方所說的話；從另一面來說，因著使用者的需要，它本身也是一種表達的方式。在語言的學習過程中，『看』也是很重要的學習步驟，必須試著從外觀上去模仿。本論文的目標，即從學習語言的相關經驗，試著以臉部模型，站在『被看』的角色，作為一個表達語言的工具。

這裡的工作是從基本的唇型開始準備。首先，大抵語言的發音，都有類似『聲母』與『韻母』的分別，本論文選擇英語為範例，它的發音基礎是母音，如/a/、/e/、/i/、/o/、/u/，以及子音，如/b/、/d/、/f/、/g/、/h/...等。掌握了各個子音及母音的發音，其次，即可以母音和子音組成每一個音節，進而串接各音節，完成一個字，更進一步，一整句話，甚至一段話的唇型，均可完整表達。

4.1 外輪廓

在這方面，藉由先前所準備的變形目標，每一個『音』的特徵憑著所拍攝的照片顯現，取得不同發音的位移量，控制模型變化，分別獲得初步的變形圖--這部分只是決定最終變形圖的開端。接下來，將以圖 4-1/ o/ 這個音為例子，一一介紹如何取得最終的變形圖。



圖 4-1 /o/的發音圖

在取得發音圖中的點資料之後，經由和起始狀態(rest state)的比較，如下圖 4-2，測量位移量，作為控制模型的輸入值。

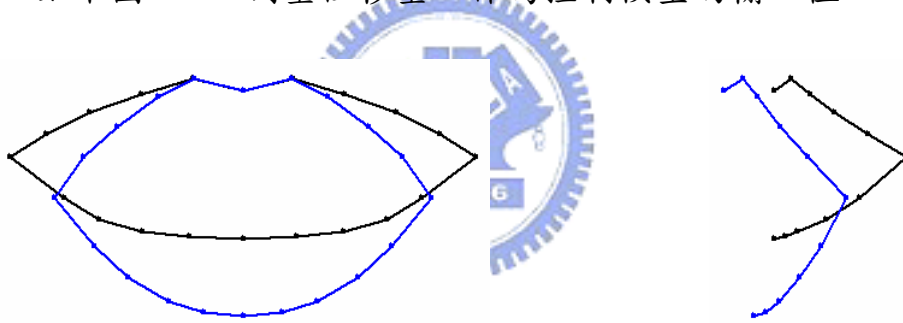


圖 4-2 發音前後比對圖

在如下所示的圖 4-3 中，經變形量限制的節點，除方向箭頭外，另外以較大的點作為標記，以避免和其它的節點產生混淆。此外，在介紹節點的變形量限制時，為簡化說明的步驟，將只介紹右半臉，左半臉取對稱值；也就是如果右半臉某節點的位移量為(a, b, c)，則其對稱於左半臉之節點其位移量應該為(-a, b, c)。

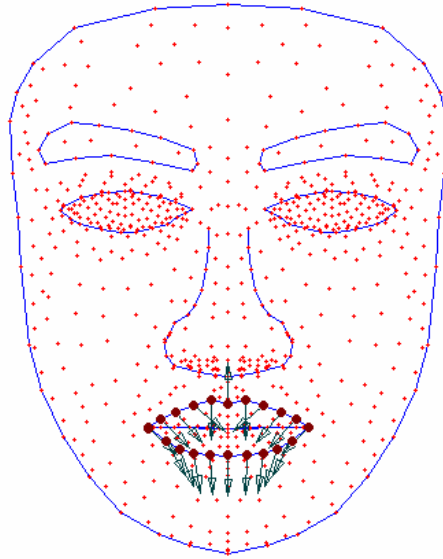


圖 4-3 外輪廓的變形量限制

在外輪廓上，依順時針方向，右半側節點的變形量限制分別是：

(0,0.02,0.23)、(-0.0001,-0.007,0.23)、(-0.08,-0.07,0.23)、(-0.17,-0.2,0.2)、
 (-0.32,-0.33,0.2)、(-0.48,-0.48,0.2)、(-0.33,-0.61,0.04)、(-0.25,-0.74,-0.06)、
 (-0.14,-0.86,-0.13)、(-0.08,-0.97,-0.08)、(0,-0.98,-0.08)



圖 4-4 外輪廓結果

由上圖 4-4 的結果，和圖 4-1 做比較，先不論之前章節中所提到的下巴位置及形狀，在嘴巴的內輪廓上，存在明顯的問題；分別由正視圖和側視圖看，未達成所期望的目標。

4.2 內輪廓

原本最初的構想是，以外輪廓的控制，希望能直接得到不錯的成果。但大多時候，單單靠著外輪廓的控制，仍不足以表現整個發音的特徵，整體上還有很大的改進空間，必須做進一步的調整。在這裡面臨的問題，原本希望藉由更改網格特性，減少操作上的手續，但仍找不到很好的設定來改善；只有繼續增加變形量限制。

在微調的過程中，這部分的工作並不像前一節有『點資料』可以參考，『點資料』也不容易取得；微調的方法不是很固定：希望哪裡的形狀要加強，就在附近的點增加額外的變形量限制。因為缺少可以在事先準備的工作，只能比對照片，靠著一步一步的調整來修正。

因微調所增加的變形量限制，可以有比較彈性的選擇，不必 3 個方向都給變形量；基本上，調整正視圖的形狀，通常只需正向的位移量(u, v)，而調整側視圖時，只需給側向的位移量(w)。之所以減少變形量的自由度限制，避免因不瞭解而產生彼此間互相干涉的情形。

內輪廓的調整是微調程序的第一個步驟，針對嘴型形狀的不理想，所進行的修正程序。

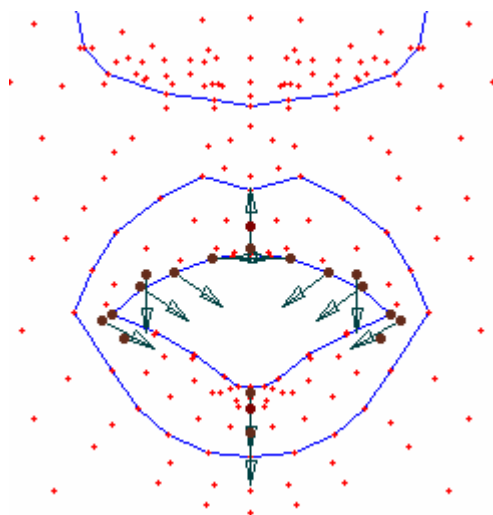


圖 4-5 內輪廓的變形量限制

由上而下，右半側節點的變形量限制分別是：

$(0,0,0.27)$ 、 $(0,0.01,0.24)$ 、 $(-0.02,0,0)$ 、 $(-0.17,-0.18,0)$ 、 $(0,-0.25,0.35)$ 、 $(-0.36,-0.31,0)$ 、
 $(0,0,0.19)$ 、 $(-0.82,-0.48,0.62)$ 、 $(0,0,0.11)$ 、 $(0,-0.96,0.023)$ 、 $(0,-1,0.028)$ 、 $(0,-0.98,-0.02)$

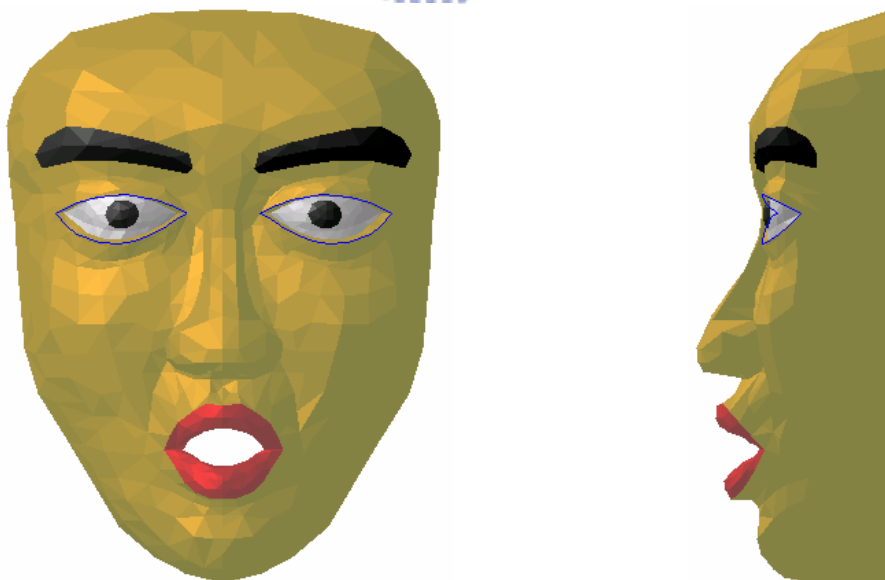


圖 4-6 內輪廓修正結果

4.3 下巴

在修正內輪廓之後，針對上圖 4-6 的結果，輸入尚未運用的下巴點資料，進行下巴部位的調整。在下巴部位，當初只取 1 個參考點，結果不夠理想，引用調整內輪廓的微調方式，加入其他節點的變形量限制來加強效果。

此一階段的工作，主要藉由模型側視圖的顯示，可以清楚表示下巴的位置及形狀，照著參考目標的樣式，觀看其中的差別，比對要修正的部位。在模型的對稱軸(Y 軸)上，即沿著側視圖前緣的節點，如下圖 4-7 所示，作為選擇的對象，調整下巴部位的變形。修正的結果如圖 4-8。

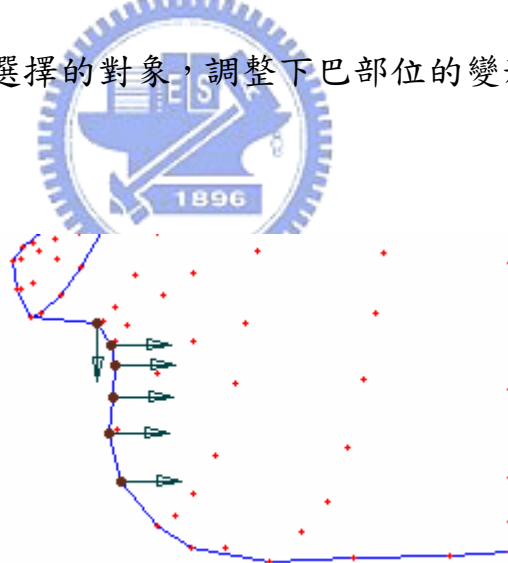


圖 4-7 下巴的變形量限制

由上而下，節點的變形量限制分別是：

$(0, -0.89, 0)$ 、 $(0, 0, -0.08)$ 、 $(0, 0, -0.1)$ 、 $(0, 0, -0.14)$ 、 $(0, 0, -0.27)$ 、 $(0, 0, -0.41)$

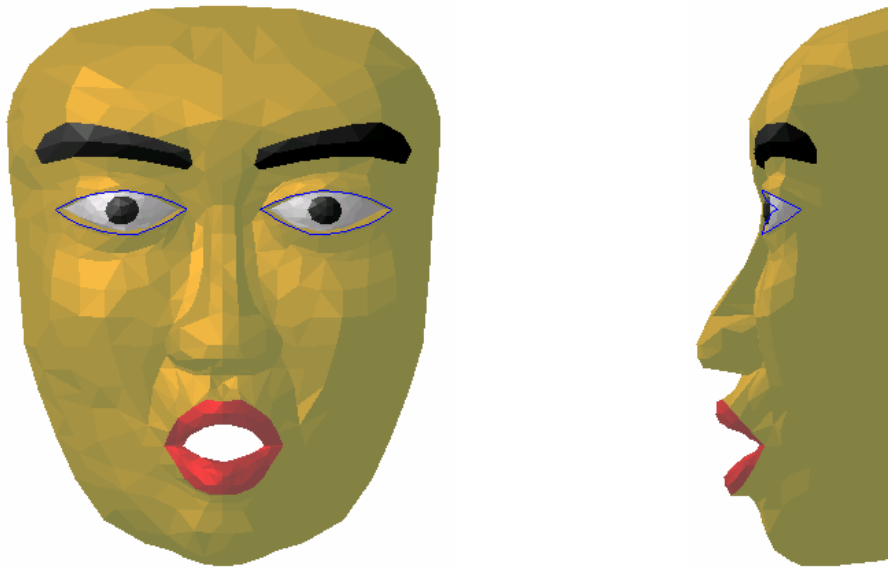


圖 4-8 下巴修正結果

4.4 臉形

在前面的工作結束之後，臉形因著遷就於整個嘴型的變化，特別是在第三個步驟之後--下巴形狀的調整，形狀往往顯得不自然。

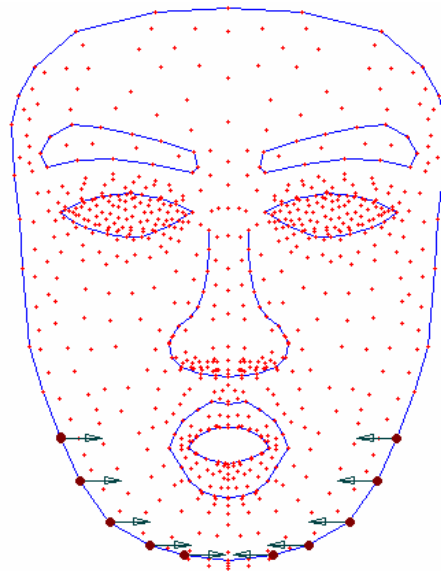


圖 4-9 臉形的變形量限制

由上而下，右半側節點的變形量限制均是(-0.05,0,0)

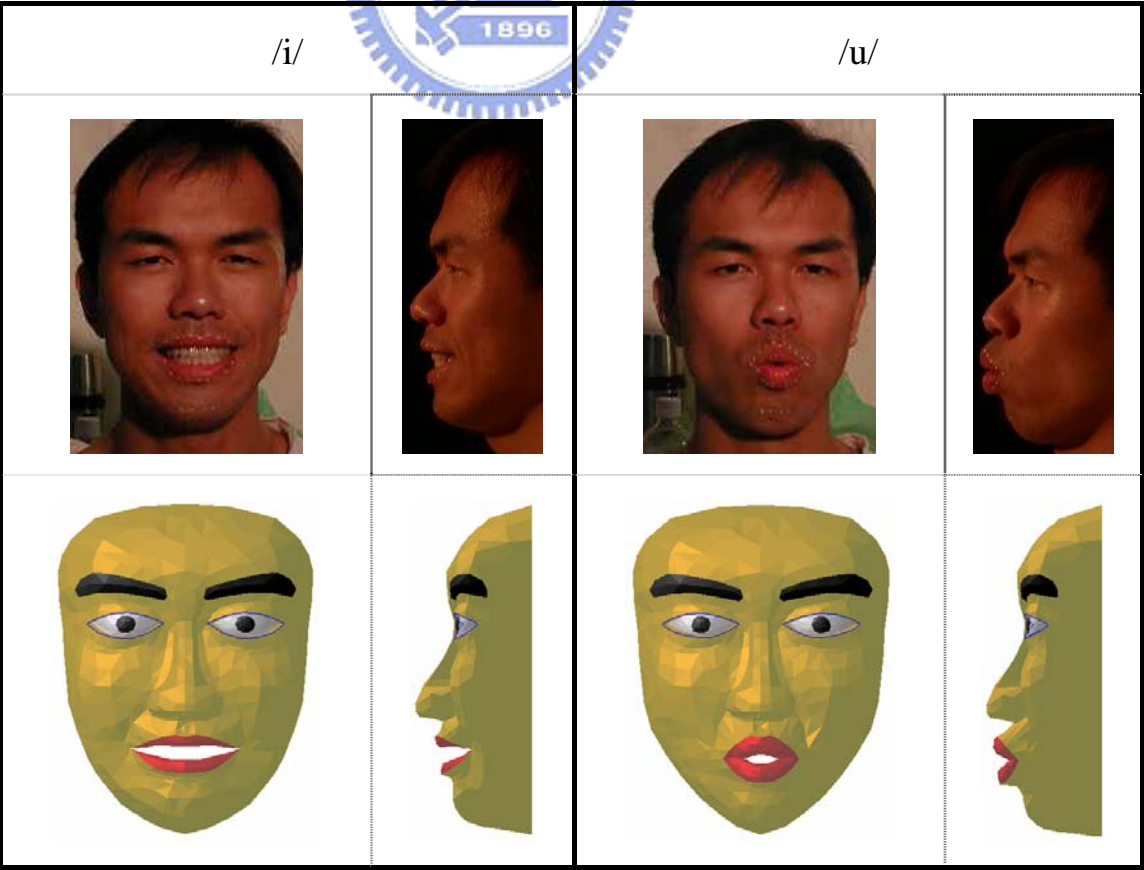
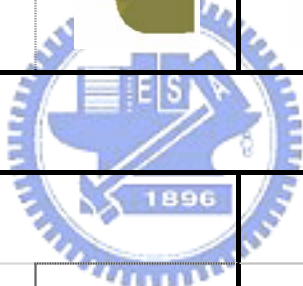
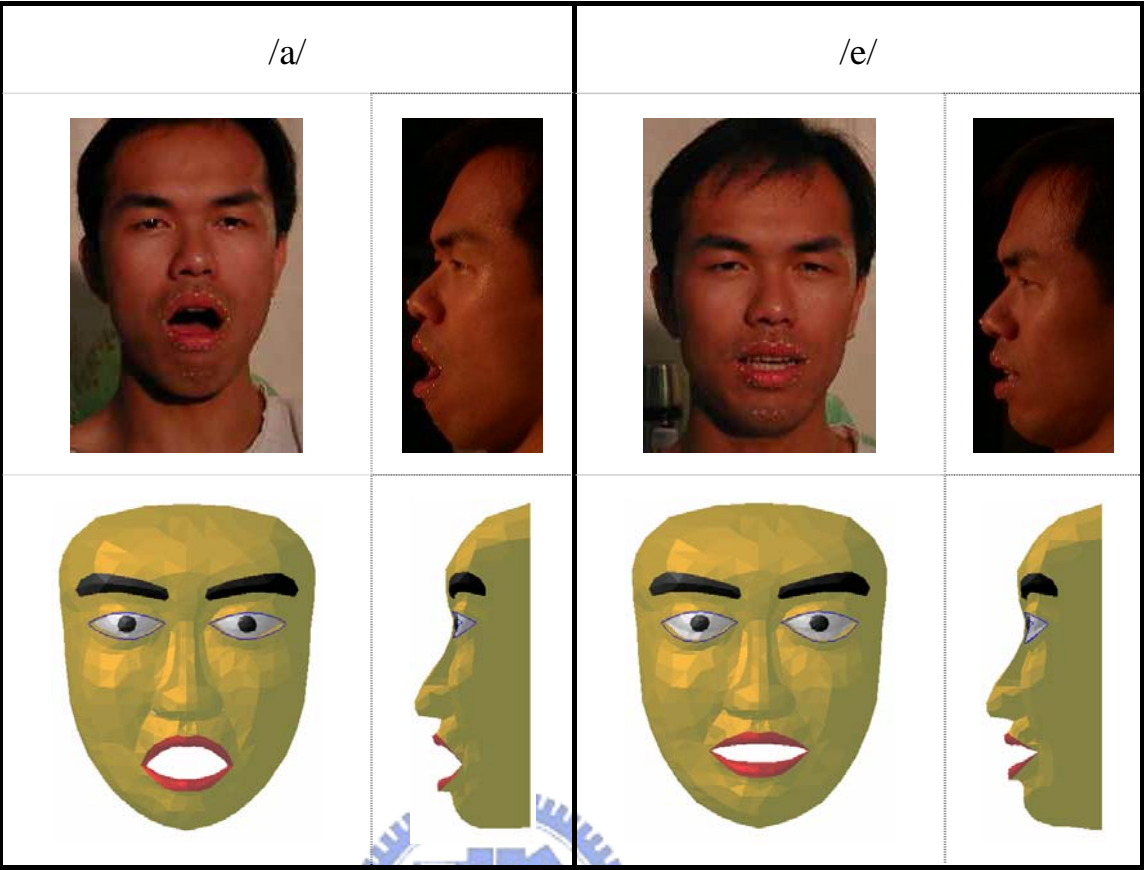


圖 4-10 臉形修正結果

經由上圖 4-9 的變形量限制，在臉部模型的下半部，選擇外緣上的節點，修正結果如圖 4-10。不同的發音，面臨的情況並非都相同，有的如 /o/ 這個例子，臉會張得太大，有的則縮得過小；實際上的臉型並不會有這麼大的變化，為解決這種情況，必須再增加限制，限制條件只侷限在 X 方向給微小的變形量限制，這裡所做的只是略加修飾的工作。到這裡，一個『發音』完成。

4.5 其他母音、子音的結果

透過相同的工作流程，可以獲得其他母音或子音的唇型，這裡主要是提供特徵上的不同，事實上，結果會隨著不同的人、不同語氣而有變化。



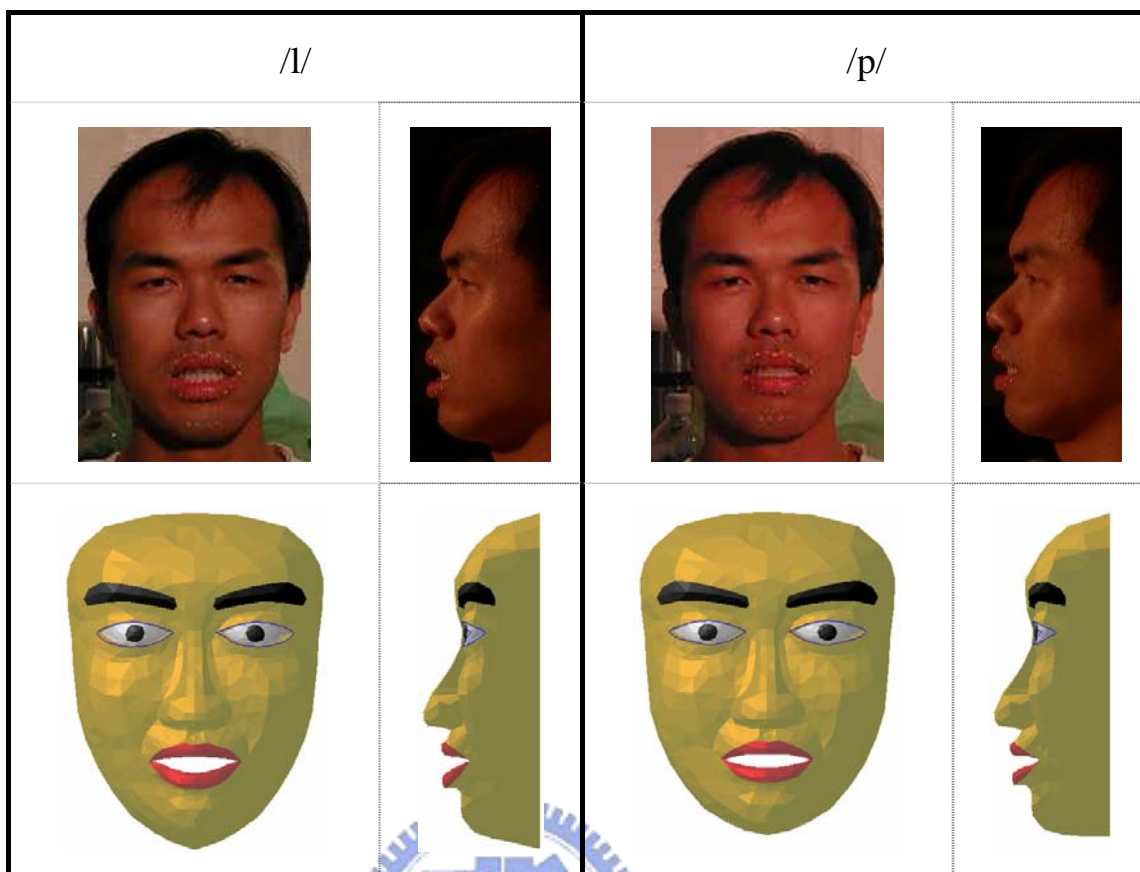


圖 4-11 其他母音、子音的結果

4.6 舌頭的搭配

考慮到除了唇型的不同，實際上舌頭在說話中也佔了一個重要的地位，舌頭的靈活，大大影響一個人說話的清晰與否。在這模型之中加入舌頭，藉由舌頭的動作，與唇型互相搭配，更清楚地產生不同『音』之間的分別。

對於舌頭的設定，在人的口腔構造中，舌頭的根部會隨著下巴的開闔而跟著上下移動，但是這樣的要求--舌頭與臉部相連結，與模型的球殼結構無法結合。退而求其次，額外加入的舌頭，和原本的顏面部

位，必須可以形成分離的兩組，對於彼此的動作不會互相干涉，考量這兩組的獨立性，新增加的舌頭，獨自擁有新的邊界條件。而獨立的邊界條件，同時也限制了要使整個舌頭都能移動的可能性，在有限單元法中，分析的對象體必須擁有完全限制自由度的節點。在條件的限制下，給予以下的設定：

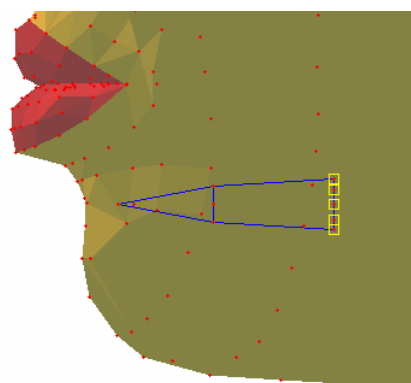


圖 4-12 舌頭的位置

在圖 4-12 中，直線所圍的部分即舌頭的位置，使其位於此處的主要考量是強調舌尖的動作，並且不希望舌頭根部影響到唇型的顯示。小正方形所框的點即限制位移量的節點，其中最上面和最下面的節點限制 X、Y、Z 向的自由度，其餘的只限制 Y 向的自由度。並簡單地分成前後兩段，達到能伸展與彎曲的功能，如下圖 4-13 變形前後的結果。



圖 4-13 舌頭的功能

五、動畫顯示

在上一章裏得到的唇型，只是得到母音和子音的 end-state，對於表現一個『音』的分別上，還不是很足夠。在第四章所做的五個母音的結果/a/、/e/、/i/、/o/、/u/，各有獨特的特色供人辨識；但對於子音，唇型和唇型之間的分際，並不是那麼突出，試著去分析子音的分別，主要的不同，是因著不同的發聲方式而有不同的聲音，這方面要藉著加強特別的發音動作來區別。大多數的子音在發聲時，因著肌肉的施力，唇型會有些微變化，但幅度並不大，可以歸類於靜態的部分，如/d/、/f/、/k/、/H/、/J/等；另外像是/b/、/p/靠著嘴唇的動作發音，唇型有很明顯的變化，在發音上必須考量該子音的位置做區分。

以第四章最後的/l/、/p/這兩個音做介紹，像是/l/這個音是在唇型上屬於比較靜態的，在嘴唇微開的狀態，由舌頭輕輕頂住上顎和喉部發聲來配合；/p/這個音是從抵住上下嘴唇開始，嘴唇放鬆後將氣向外噴出。雖然這兩個音在最後的 end-state 沒有很大的差別，卻可以從動作的變化上做區分。再以/l/、/n/做比較，兩者在外觀上幾乎相同，但/n/的發音方式是舌頭平放，以鼻腔發音，只能靠著舌頭的位置來分別是發什麼音。

另外，以/p/、/m/這兩個音做比較，如果只是單獨一個音，在字的最後音節出現，/m/是一直維持在抵嘴的狀態，用鼻腔發音，和/p/的嘴

唇動作，放鬆後被空氣衝開，可以有所區隔；但是，當/p/、/m/各和母音組成一個音節時，都是從原先抵嘴的狀態，搭配各個母音的唇型發音；子音的動作，只佔整個發音動作一開始的一小段時間，而且幾乎都相同，所能指出的差別，是/p/這個音有將嘴唇稍稍向前、向上下分開的動作，但其實是很細微的一個差別，在配合母音發音的同時，很難察覺出來。

其他比較特別的例子，像是/f/和/v/藉由上牙齒輕咬下唇、/s/和/z/必須將牙齒緊閉、/h/和/b/將舌頭置於上下牙齒中間來發音，但這些關於牙齒和舌頭相配合的動作，還無法在這個模型中實現，少了牙齒和舌頭的輔助顯示，也少了很多顯示上的分別。



5.1 動畫的產生

本論文的主要目的，即是藉由動畫播放的方式，將所要表示的字句，盡可能地清楚表現出來。按著所需要的各個子音和母音，排定動作順序，一一置入。和普遍的動畫製作所不同的是，為節省製作上所必須耗費的準備工作，這裡將所產生的唇型設定為 key frame，整個動作被分為若干階段，每個階段各有起始狀態和結束狀態，在不同的狀態間，利用線性內插的方式，按著相對應的節點位置及網格，計算出變化過程的節點位置，得到各個階段的連續變化圖。

以數學模型做說明：

$$P = (1-t)P_a + tP_b; 0 \leq t \leq 1 \quad (24)$$

在每一階段中， P_a 代表起始狀態的形狀參數， P_b 代表結束狀態的形狀參數， t 為時間控制。理所當然地，使用這方法有一個很大的限制，所使用的 key frame，必須是相同的模型所得到的，具有一致的網格與顏色。全部的動畫，從第一個階段的開始，轉換到結束狀態；在新的階段中，原先的結束狀態成為新的開始，繼續新的變化過程，直至終了。

在每一個階段的安排上，有時不能直接置入唇型的 end-state 做為 key frame，必須考量動作上的分別，選擇合適的 state。又因著變化過程的簡化，同時也失去了部分的真實性，有時唇型的變化可以用如此的方法近似，但並不適用於全體，為了修正這方面的不足，必須為部分的變化過程額外安排過渡性質的 key frame，使變化過程可以有更多一點的轉折、變化性。

5.2 動畫播放的時間控制

在前一節中提到，每一個階段都是以線性內插的方式來控制變化的過程，其中，要完成每一個動作的時間又不盡相同，為了有更真實的呈現，進一步在時間控制上加強處理，使每一階段可以表現不同的

時程控制，必須為(24)式的時間控制有更多說明。動畫是以短時間內快速切換具有微小變形量的連續畫面而產生，要決定的項目有兩項：

1. frame 與 frame 之間所插入的圖片張數
2. 圖片的暫停時間

只要調整其中一項，均可使動畫播放可影響時程的控制；在第一項中，調整每一階段所插入的圖片張數，轉換速度將隨著所插入圖片張數的增加而變慢；而在第二項中，調整每一階段中圖片的暫停時間，轉換速度將隨著圖片暫停時間的增加而變慢。也可以兩者合併使用，但為避免操作上的繁雜，而且這部分也受到應用程式的限制，在播放動畫的過程中，要中途變更暫停時間的設定不是容易；所以在這裡採用第一種方式，先整理過程中每一階段所需要的時間，再決定圖片的暫停時間，以此換算需要插入的圖片張數。

藉由圖 5-1 解說，若整個過程共有 4 張 frame，在第一階段 frame1 到 frame2 的過程需 1.0 秒，第二階段 frame2 到 frame3 的過程需 0.3 秒，第三階段 frame3 到 frame4 的過程需 0.5 秒；所以，在設定暫停時間時，必須以這三者的公因數為單位，且換算每一階段所需要插入的圖片張數時，一般而言，為使整體的動畫播放有平順的效果，每秒鐘的播放速度太多或太少都不是好的選擇，透過現有的繪圖函式，得到很大的便利，不必將變化過程切分的太細就有很好的效果，所以在這例子中，

以 1.0、0.3、0.5 的最大公因數 0.1(秒)為圖片的暫停時間，事實上以更小的單位也很難分辨差異性，所以這三個階段各別換算得到跨度 10、3、5，跨度代表每一階段從起始狀態到結束狀態所經過的間距，扣除結束狀態的 frame，即三個階段中插入圖片張數各為 9、2、4。以上只是舉例，圖片的暫停時間決定整體的播放速度，若要再變更某一階段的轉變時程，可依變動的時間來決定跨度的增減。

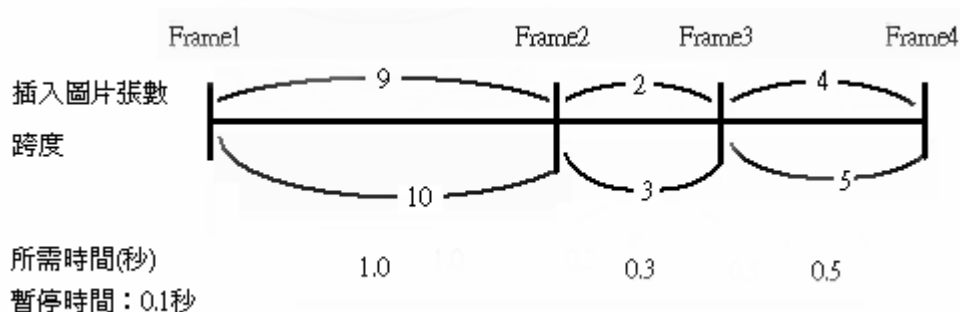


圖 5-1 速度控制

語言合成就是將一連串特定唇型的連續轉換過程，隨著內容，按著字的發音，依序置入所需要、合適的 frame，並搭配時間的控制，決定圖片的暫停時間與插入圖片的張數，以動畫播放的方式來呈現說話時嘴唇動作、特徵的表現及變化的快慢。接下來的部分，以 lip、loop、lull 這三個字介紹成果，受限於文稿的方式，只顯示主要的 frame，並以文字略做說明，三個例子都是將暫停時間設定為 0.1 秒，各階段的跨度可按所需時間設定，需特別注意的是，像是強調 lip 與 loop 最後音節『p』這類音的時候，給較短的跨度比較能看出效果。



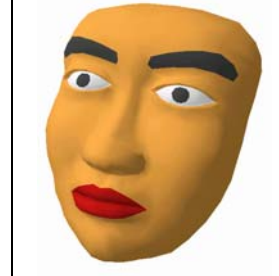
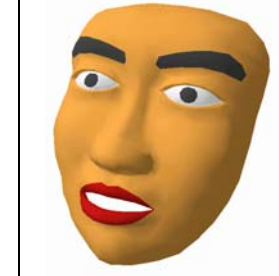
| | | | |
|---|---|--|---|
|  |  |  |  |
| Frame1 | Frame2 | Frame3 | Frame4 |
| l | i | p-起始狀態 | p-結束狀態 |

圖 5-2 『lip』



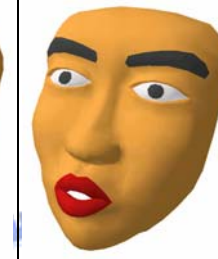
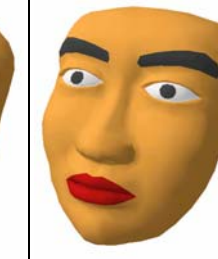
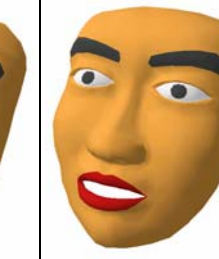
| | | | | |
|--|--|--|---|--|
|  |  |  |  |  |
| Frame1 | Frame2 | Frame3 | Frame4 | Frame5 |
| l | u-過渡 | u | p-起始狀態 | p-結束狀態 |

圖 5-3 『loop』




| | | |
|---|--|---|
|  |  |  |
| Frame1 | Frame2 | Frame3 |
| l | a | l |

圖 5-4 『lull』

六、結論

以一個可承受施力而產生變形的彈性體所建立的臉部模型，在美工之外，發展出不同的方法，經由變形量的控制，產生所要表現的效果；經由觀察並分析說話時的唇型特徵，決定語言合成中所必須的動作，並藉由動畫的動態模擬，完成此一研究。

本系統中，雖然在建立與修改模型上有很大的自由度，可以隨使用者的需要做不同的變化，但整個過程是非常繁複的工作，使用者必須耐心、仔細地逐一完成，若要呈現不同的臉部特徵、輪廓，仍有許多的不便；2.1 節所提到的參數化模型，和本系統結合，可以是一個不錯的研究，將大大增加使用上的便利性。

在研究中也明瞭，雖然模型已經可以達到一些基本要求，但要完美達到如實際上臉部、嘴唇的動作，還是有很大的距離，模型的相關設定，如何能更真實還需要再討論；說話的動作也不只是嘴巴的開合，配合人的情緒表現、表情的變化，其中所具備的細膩性，還有很大的可發揮空間。

參考文獻

1. Frederic I. Parke, "Parameterized Models for Facial Animation", IEEE, pp.61-66, November 1982.
2. David R. Hill, Andrew Pearce, and Brian Wyvill, "Animating Speech: an automated approach using speech synthesised by rules", The Visual Computer, 3, pp.277-289, 1988.
3. 詹凱富, "人類面部表情控制之有限元素分析", 國立交通大學, 碩士論文, 民國八十九年。
4. Sumit Basu and Alex Pentland, "A Three-Dimensional Model of Human Lip Motions Trained from Video", IEEE, 16, pp.46-53, June 1997.
5. Robert D. Cook, David S. Malkus, Michael E. Plesha, Concepts and Applications of Finite Element Analysis, Wiley, New York, 1989.

