逆動力学的手法による人間調音運動の制御*

金 岡 秀^{*1}, 五味 裕 章^{*2}

Human Articulator Motor Control Based on Inverse Dynamics

Kangsoo KIM^{*3} and Hiroaki GOMI^{*3} NTT Communication Science Laboratories, Nippon Telegraph and Telephone Corporation,
3-1 Morinosato, Atsugi-shi, Kanagawa, 243-0198 Japan

The inverse dynamics based motor control strategy in human articulatory motion has been investigated. As the plant for this control application, dynamic model of human articulators consisting of lip soft tissue with related muscles and bones is introduced. The continuum of lip soft tissue is represented as a discrete model approximation composed of networked point masses interconnected via viscoelastic elements. Stiffness of every viscoelastic element is adjusted to ensure the compatibility in static deformation between discrete model and its continuum prototype. Simulation of forward dynamics by activating ingeniously selected muscle set induces a mimic of actual human speech articulation, called pseudo-speech motion. As a mathematical description of the human speech acquisition, articulatory control simulation based on iterative estimation of muscle motor command is presented. Inverse dynamics driving iterative estimation lets the articulator dynamic model put into practice an articulatory motion, induced by the estimated muscle motor command.

Key Words: Articulator, Biomechanics, Learning Control, Inverse Dynamics, Speech, Muscle and Skeleton

1. 結 言

人間の調音・発声は、上・下口唇、舌、顎や声道及び、これらに付属した筋や骨などで構成する調音器官の連成運動から発生する共鳴や、空気流れによる摩擦音などを使って生成される。人間の調音・発声を体系的に理解するためには、その構成要素の一つである調音運動を調べる必要があるが、その方法としては、MRI (Magnetic Resonance Imaging) や筋電信号、EMA (Electro-Magnetic Articulography) などの計測装置による直接計測と、動力学的拘束条件を満たすように作成した動力学系モデルによる、調音運動シミュレーションが考えられる。直接計測は求める情報がすぐ得られる長所があるが、調音運動は軟組織としての弾性体 (口唇、舌) が弾性変形しながら、下顎骨、舌骨などによる剛体運動と連動する極めて複雑なものであり、上記の計測装置はまだその詳細な観測に十分な分解能が提供できない。

い。また複雑に入り組んだ筋の活動は直接計測で観測することができない。本研究では調音動力学系モデルを構築し、直接計測では捉え難い調音運動メカニズムの解明を目指す。本研究で構築した動力学系モデルは上・下口唇、下顎骨、舌骨及び関連筋群をその構成要素とする。実際の人間調音運動では上・下口唇の他にも舌や軟口蓋などが協調して多様な発声音の生成に対応しているが、本研究では調音動力学系モデル開発の第一歩として口唇を中心とするモデルを開発し、口唇運動が支配的影響を持つと考えられる日本語母音の発声運動を、シミュレーションによって機械力学・制御の観点から考察した。

口唇軟組織は本来連続体であるが、本研究では Platt ら⁽¹⁾、や Dang ら⁽²⁾などが提案した離散系近似による連続体モデリングの手法を用いて、離散近似系としての軟組織モデルを導出した。離散近似系モデルにおいて連続体は複数の離散質点の分布として代表され、隣接した質点同士は離散の粘弾性要素を介してつながり合わせる。離散近似系モデルは例えば有限要素モデルなどと比べてモデルとしての精度は低いものの、計算負荷が小さく大変形運動の解析が容易であるなどの長所がある。

* 原稿受付 2006 年 10 月 10 日。

^{*1} 正員、日本電信電話(株)NTT コミュニケーション科学基礎研究所(〒243-0198 厚木市森の里宮宮 3-1)。^{*2} 日本電信電話(株)NTT コミュニケーション科学基礎研究所。

E-mail: kskim@idea.brl.ntt.co.jp

離散系モデルが近似対象である連続体としての口唇軟組織の挙動を動力学的に正しく再現するためには、離散系と連続体系の間に、外力荷重による変形における整合性が取れていなければいけない。本研究では任意の外力荷重に対して、離散近似系が連続体系と等価の変形をもたらすよう、離散系モデルにおける粘弾性要素の弾性係数を最適調整する手法を開発・適用し、より高い精度で人間調音運動の模擬が可能な動力学系モデルを構築した。

調音動力学系モデルの構成要素の一つである筋群は、その収縮から発生する筋力でモデルを駆動するアクチュエータとしての役割を果たす。本研究では筋のモデルとして、筋群の配置及び軟組織や骨との接続関係が解剖学的根拠に基づいて構築された Nozoe ら⁽³⁾のモデルを用いる。

特定の発声音はそれに相応する調音運動によって生成される。目標とする調音運動に対して、それを実現する筋力の発生に必要な筋指令を決定し、発令する一連の過程を調音制御として定義する。調音制御において筋指令はコントローラとしての人間の脳が決定して発令するが、人間はある調音目標の達成に必要な筋指令を、学習過程を通じて獲得していくと考えられる。本研究では学習による人間の調音筋指令獲得過程を、逆動力学的手法による調音制御モデルとして表現し、人間の調音制御学習を模擬する数理モデルとして提案する。逆動力学的手法の適用で得られる調音筋指令は、それが誘発した調音運動による特定の発声音に対応する固有のものであり、人間の発話コミュニケーションにおいて音声とは別の形で表現される情報である。

2. 調音動力学系モデル

2・1 調音動力学系モデルの構成 本研究で用いている調音動力学系モデルの全体構成及び座標系を図1に示す。前述のように調音動力学系モデルは上・下口唇と関連筋群及び骨構造で構成されているが、図1で各線分要素はモデル化した筋や骨を表す。モデル化した筋の種類及びその配置は、Zemlin⁽⁴⁾による口唇周辺部の解剖図を参考にして決定した。

連続体である口唇軟組織はその幾何学的形状が良く表現できる位置に設けたノード上に分布する離散質点の集合として代表され、隣接した質点同士を離散の粘弾性要素でつなぎ合わせた離散近似系としてモデル化する。粘弾性要素は図2のようにばねとダッシュポットが並列で結合した機械力学系要素で、質点の運動に対して復元力及び減衰力を発生する。離散系モデル内の一つの質点は、一般に26個の粘弾性要素と接続を持

つが、表面など境界上に位置する質点はより少ない接続となる。

Fig. 1 Configuration and coordinate system of the human articulator dynamic model

Fig. 2 A viscoelastic element

2・2 連続体系適合離散近似系 離散系とそれが近似している連続体系が等価の外部荷重に対して同じ変形を引き起こす場合、その離散近似系を連続体系適合離散近似系として定義する。離散近似系の構築において、Lee ら⁽⁵⁾や Grzeszczuk ら⁽⁶⁾は離散近似系全体にわたって同一の弾性係数を持つ粘弾性要素を用いているが、このような構成は粘弾性要素の空間密度が一様な場合のみに連続体系適合性が成り立つ。一様な空間密度で粘弾性要素を配置するのは、直方体のような単純形状の対象物以外は極めて難しく、本研究では既に配置済みの粘弾性要素に対して、連続体系適合性をもたらすよう弾性係数の最適調整を行い、連続体系適合離散近似系を導出する手法を新たに開発して適用した。

離散の弾性要素（ばね）において断面積は定義されないが、本研究では離散要素の弾性係数を連続弾性体と同じく式(1)のように記述し、有効断面積 A_{eff} を定義する。

$$k = \frac{EA_{eff}}{l} \quad (1)$$

式 (1) でヤング率 E 及び離散要素長 l は既定であり、弾性係数の調整は有効断面積 A_{eff} をパラメータとして行われる。弾性係数の最適調整は図 3 のように隣接した 8 個のノードが構成する六面体の単位で行うが、本研究ではこの六面体を“単位キューブ”と呼び、単位キューブ内には総 28 個の粘弾性要素が存在する。

Fig. 3 An "unit cube" in discrete lip soft tissue model

単位キューブが近似している連続体において、各単位キューブ内に定義される局所座標系 $O'-x'y'z'$ (図 3) におけるある i 軸断面に、 i 軸方向の外力荷重 f_{Ci} が加わった場合、 i 軸断面の変位は次式で求められる。

$$\delta_{Ci} = \frac{f_{Ci} l_{Ci}}{EA_{Ci}} \quad (2)$$

$i=1, 2, 3$ であり、それぞれ局所座標系 $O'-x'y'z'$ における x', y', z 軸方向を表す。 l_{Ci}, A_{Ci} は連続体の i 軸方向長及び断面積を表す。連続体と等価の荷重条件下の単位キューブにおいて、 i 軸断面の変位は次式で求められる⁷⁾。

$$\delta_i = K_{cb}^{-1} f_i \quad (3)$$

K_{cb} は単位キューブの剛性マトリクスで、キューブを構成している各粘弾性要素の弾性係数及びその方向性によって作られる。 f_i は i 軸断面に作用する外力荷重ベクトルであり、連続体に作用する f_{Ci} と等価のものになるように、 i 軸断面上 4 つのノードに割り当てられる。 δ_i は i 軸断面上 4 つのノードの変位を成分とするベクトルであり、局所座標系 $O'-x'y'z'$ で記述される。 δ_i において i 軸方向成分を 4 つのノードで平均した値を $\bar{\delta}_i$ とすれば、 $\bar{\delta}_i$ は単位キューブにおいて、連続体における荷重条件 f_{Ci} と等価の荷重条件による i 軸断面変位の代表値となる。 $\bar{\delta}_i$ の定義から式(4)の P が最小になるように最適調整された弾性係数すなわち有効断面積を持つ粘弾性要素で構成された単位キューブは、等価の外力荷重に対して最小誤差で連続体変形を近似する変形をもたらすものであり、連続体系適合離散近似系としての条件を満たす。

$$P = \sum_{i=1}^3 |\delta_{Ci} - \bar{\delta}_i| \quad (4)$$

本研究では上記の手法で口唇軟組織の離散近似系を構成する全ての単位キューブに対して、粘弾性要素の弾性係数の最適調整を行い、連続体系適合離散近似系を導出した。複数の単位キューブに共有される要素に対しては、各キューブから導出した連続体系適合弾性係数の平均値を適用する。

図 4 に六面体の連続体系及びそれを全て異なった寸法の 4 個の単位キューブで近似した離散近似系モデルを示す。

Fig. 4 (a) A hexahedral continuum (b) Its discrete model approximation composed of 4 unit cubes

図 4b の離散近似系に、 x 軸断面上一様応力分布荷重を加えた場合の変形を 図 5 に示す。図 5a は全て一定の弾性係数で構成した離散近似系における変形である。この場合、各ノードの変位が全て異なっており、一様断面変形を引き起こす連続体系の挙動を再現していない。図 5b は上記の手法による弾性係数の最適調整を通じて得た離散近似系における変形で、極めて小さい誤差で連続体系変形を近似しており、開発した手法が連続体系適合性を実現していることが分かる。

Fig. 5 Forced deformations of two discrete models : (a) Model of constant element stiffness (b) Model of continuum-compatible element stiffness

2・3 筋動力学系モデル 筋は調音動力学系において力を発生し、調音運動を引き起こすアクチュエータとしての役割を果たす。本研究で用いている筋の動力学系モデルは、Hill⁸⁾が提案したものを基本としており、図 6 に示す三つの個別ユニットが並列に結合した形となっている。

Fig. 6 Modified Hill's model for muscle structure

図6のユニット1は定数弾性係数のばねで、初期の釣り合い状態よりその長さが増加した場合のみに力を発生する非線形性を持つ。ユニット2は粘性力及び弾性力を発生する要素で、ユニット3は受信した筋指令に基づいて筋力を発生する能動要素である。本研究ではユニット3の発生筋力をHill⁽⁸⁾が本来提案したモデルの代わりに、Itoら⁽⁹⁾が提案した次式の近似モデルを用いて計算する。

$$f_{m3} = \mu \alpha A_m [\exp(-\varepsilon^2) - \exp(-\varepsilon_{\max}^2)] \quad (5)$$

式(5)で f_{m3} は能動筋要素による発生筋力、 μ はゲインで、 A_m は当該筋要素の断面積を意味する。また α は個別の筋に伝わる筋指令である。 ε は筋要素のひずみで、その最大値 $\varepsilon_{\max}$ は0.6として設定されており、 f_{m3} は $-0.6 \leq \varepsilon \leq 0.6$ の範囲内のみで有効な値を持ち、その範囲外では0となる。上記の筋モデルによる発生筋力は、筋モデルと接続を持っている口唇軟組織や骨モデルのノードに作用する。

2・4 運動方程式 調音動力学系モデルの運動方程式は各質点における線形運動量保存則から導出される。個々の質点は粘弾性要素によって隣接の質点と接続しているため、モデルの動的挙動を表す運動方程式は次のような連立方程式の形となる。

$$m\ddot{R} = -K(R - R_0) - C\dot{R} + F_a \quad (6)$$

m は対角要素として各ノードの質量を持つマトリクスで、 R 及び R_0 は各ノードの現在及び初期位置を表すベクトルである。 K 、 C は剛性及び粘性マトリクスで、モデルの静的復元及び減衰の特性を表現する。 F_a は作用外力ベクトルで、モデルに外力として加えられる力はここに現れる。

位置ベクトルの時間微分が速度ベクトルであることから、運動方程式は式(7)のように表せる。本研究では速度ベクトルを状態ベクトルとする式(7)を動力学系モデルとして用い、時間領域において数値的にその解を求める。

$$\dot{V} = -m^{-1} [K(R - R_0) + CV - F_a] \quad (7a)$$

$$V = \dot{R} : \text{Velocity Vector} \quad (7b)$$

2・5 モデル剛性及び粘性

式(6)の運動方程式に表されている剛性マトリクスは、トラス構造などの解析によく使われるマトリクス変位法⁽⁷⁾で定義されるもので、その作成法を簡略に説明する。要素番号 ne の粘弾性要素がノード番号 $nn0$ 及び $nn1$ のノードを両端としてまたがっており、その弾性係数が $k^{(ne)}$ である場合、全系モデルの弾性における要素 ne の寄与分は、式(8b)のサブマトリクス $k_0^{(ne)}$ を介して、剛性マトリクス K の中に式(8a)のように展開される。

$$K(k^{(ne)}) = \begin{bmatrix} \cdots & nn0 & \cdots & nn1 & \cdots \\ \vdots & 0 & \cdots & 0 & \cdots & 0 \\ nn0 & \vdots & k_0^{(ne)} & \vdots & -k_0^{(ne)} & \vdots \\ \vdots & 0 & \cdots & 0 & \cdots & 0 \\ nn1 & \vdots & -k_0^{(ne)} & \vdots & k_0^{(ne)} & \vdots \\ \vdots & 0 & \cdots & 0 & \cdots & 0 \end{bmatrix} \quad (8a)$$

$$k_0^{(ne)} = k^{(ne)} \begin{bmatrix} e^2 & ef & eg \\ ef & f^2 & fg \\ eg & fg & g^2 \end{bmatrix} \quad (8b)$$

式(8b)で e, f, g は要素 ne の方向余弦を表す。動力学系モデル全体の総要素数を Ne とすれば、剛性マトリクス K は全 Ne 個の粘弾性要素に対して式(8)の展開を行い、個々の展開結果を重ね合わせて完成する。

$$K = \sum_{ne=1}^{Ne} K(k^{(ne)}) \quad (9)$$

粘性マトリクス C は、式(8)の展開において要素弾性係数 $k^{(ne)}$ の代わりに要素粘性係数 $c^{(ne)}$ を入れ替えることで導出される。

3. 調音運動シミュレーション

3・1 模擬発声運動のシミュレーション 構築した調音動力学系モデルを用いて、筋収縮による動力学系モデルの運動応答を時間領域で計算し、模擬発声運動のシミュレーションを行った。本シミュレーションでは発声音でなく、口唇動力学系モデルの時間領域における運動応答からシミュレーションの結果を評価するため、シミュレーションから得た口唇運動を模擬発声運動として定義する。このシミュレーションは、まだモデル化していない舌や声道運動からの影響が少なく、口唇運動のみでも運動特性がよく表れる日本語母音の発声動作を対象として行った。

日本語母音/a/の発声を模擬した口唇運動のシミュレーションの結果を図7に示す。/a/の模擬発声動作で上・下口唇はシミュレーション開始後0.34sに最大変

位で開口しており、そのあと筋や軟組織の弾性要素による復元力によって元の閉口位置に戻っていく。

Fig. 7 Simulated lip motion of speaking /a/ (t=0.00, 0.16, 0.34 and 0.40s)

母音/a/の発声は顎の運動と深い係わりがあるが、顎運動は下顎骨の回転・並進と下唇周囲軟組織の変形が複合した運動特性を持つ。下顎骨は顎上突起を介して固定軌道に拘束されており、軌道上を滑りながら顎上突起を中心とした剛体回転を行う。本研究では下顎骨運動の記述において、Laboissiere^ら(10)やDang^ら(2)の顎運動モデルで無視されている顎上突起に作用する垂直反力としての求心力を取り入れ、より精巧なモデリングによるリアルなシミュレーションの実現を図る。

次に/a/の発声動作とは異なった運動特性を持つ発声動作に対する本動力学系モデルの有効性の検証のため、母音/u/ (図8) 及び母音/o/ (図9) の模擬発声運動をシミュレーションした。

Fig. 8 Simulated lip motion of speaking /u/ (t=0.00, 0.16, 0.26 and 0.40s)

Fig. 9 Simulated lip motion of speaking /o/ (t=0.00, 0.16, 0.26 and 0.40s)

母音/u/の場合、/a/の発声運動とは異なって下顎骨の動きは少なく、口唇の突き出しや狭めが重要な構成要素である。母音/o/の場合は、口唇の突き出しや狭めと共に、下顎骨運動も係わってくるため、より巧妙に組み合わせさせた筋指令の活性化が必要となる。

模擬発声運動のシミュレーションにおいて、本研究で開発した動力学系モデルは人間と同様に筋活動によって多様な口唇表情を生成しており、人間調音運動の力学的解明のための道具としてその有効性を確認した。

4. 調音運動の逆動力学的制御

4・1 逆力学問題の定義 逆力学問題とはある動力学系において与えられた出力から、それを引き起こした原因としての入力を推定するものである。このような定義から調音運動における逆力学の問題は目標とする発声音に対して、その生成に必要な筋指令を推定する問題となる。人間は発声器官による発声音の生成はもちろん聴覚器官による音声受信も行い、音声情報を送受信の両方向においてコミュニケーションツールとして用いている。幼年期の発話獲得の過程では、聴覚器官から受信した自分の発声音を目標としている音声に参照させ、両者の差分をなくしていく仕組みで発話学習を行うと考えられる。図10はこのような概念をフィードバック制御の観点に基づいて表したブロック図である。図10から分かるように筋指令は、人間の発話獲得を一種の制御過程として表現した場合、制御入力として働いているものであり、学習を通じて人間が覚えていく発話技能の実体もこの筋指令である⁽⁴⁾。従って逆動力学的手法による調音筋指令の推定プロセスを、人間の発話学習による調音制御獲得の数理モデルの一つとして考えることができる。

Fig. 10 Schematic of human articulatory control

4・2 調音筋指令の逆推定 本研究では発声音の代わりに調音運動の運動学的情報に基づいて調音筋指令の逆推定を行う。運動学的情報としては、少ない数で有効な調音運動観測ができるように、動力学系モデル内に設けた調音運動観測点の位置ベクトルを用いる。Gomi ら⁽¹¹⁾によると逆動力学の問題では観測点の数が多ければ、収束が早く推定解の精度も高くなる。しかし実問題を考えると、多数の観測点に対する実時間追跡計測は計測装置の容量などに制約があり、更にモデルの表面でなく内部に設けられた観測点に対する時間連続位置計測は、実質上極めて難しいものである。本研究では開口部口唇形状の拘束を考慮し、図 11 に示す開口部境界上の 4 点を口唇運動観測点として選定し、推定動作の観測点における位置誤差を評価する。

Fig. 11 Lip motion observation points

逆動力学問題の解は、式(10)のように定義した評価関数を最小にする筋指令の時系列として求められる。

$$J = \frac{1}{2} (\mathbf{x}_{ref} - \mathbf{x})(\mathbf{x}_{ref} - \mathbf{x})^T \quad (10)$$

式(10)で $\mathbf{x}_{ref}$ 及び $\mathbf{x}$ は口唇運動観測点の位置ベクトルで、それぞれ目標値及び推定値を表している。従って評価関数 J は目標動作に対する推定動作の位置誤差の二乗を意味しており、 $\mathbf{x}_{ref}$ 及び $\mathbf{x}$ が毎時間ステップごとに更新されるため、 J は各時間ステップにおける位置誤差の時系列として得られる。

逆動力学問題の解は Gradient Descent Search (GDS) のアルゴリズムによって逐次的に求めていく。Miyamoto ら⁽¹²⁾はけんだまロボットの動作学習に GDS を用いて

おり、本研究では彼らの手法を口唇動力学系モデルに適合するよう再構成し、逆動力学問題の解を求める。

GDS による調音運動筋指令の逆推定過程を以下に説明する。ある時刻 t_i において j 番目の繰り返しステップにおける筋指令は、 $j-1$ 番目の繰り返しステップの同じ時刻における筋指令にその修正値を足して、式(11)のように求めていく。

$$\alpha_i^j = \alpha_i^{j-1} + \Delta\alpha \quad (11)$$

上式で上添字 j は繰り返しステップ、下添字 i は時間ステップを表すインデックスであり、 α は筋指令ベクトル、 $\Delta\alpha$ はその修正値である。式(11)から求めた筋指令 α_i^j を式(5)に適用すれば発生筋力が計算される。発生筋力を入力として式(6)の運動方程式を数値的に解けば、GDS の j 回繰り返し適用による時刻 $t_i \sim t_{i-1}$ の区間における推定運動応答が得られる。

GDS の適用において繰り返しステップを一つの独立過程としてみなして変数 τ で表せば、評価関数 J の τ に対する変化率は次式のように展開される。

$$\begin{aligned} \frac{dJ}{d\tau} &= \frac{dJ}{dx} \frac{dx}{dV} \frac{dV}{dF_a} \frac{\partial F_a}{\partial \alpha} \frac{d\alpha}{d\tau} \\ &= -(\mathbf{x}_{ref} - \mathbf{x}) \mathbf{G} \frac{d\alpha}{d\tau} \end{aligned} \quad (12a)$$

$$\mathbf{G} = \frac{dx}{dV} \frac{\partial V}{\partial F_a} \frac{\partial F_a}{\partial \alpha} \quad (12b)$$

式(12)で V は速度ベクトル、 F_a は作用外力ベクトルを表す。次に式(12a)に表れている筋指令の繰り返しステップに対する変化率 $d\alpha/d\tau$ を式(13)のように設定しておく。

$$\frac{d\alpha}{d\tau} = [(\mathbf{x}_{ref} - \mathbf{x}) \mathbf{G}]^T \quad (13)$$

式(12a)に式(13)を代入することで、 $dJ/d\tau$ は次式のように表すことができる。

$$\frac{dJ}{d\tau} = -(\mathbf{x}_{ref} - \mathbf{x}) \mathbf{Q} (\mathbf{x}_{ref} - \mathbf{x})^T \quad (14a)$$

$$\mathbf{Q} = \mathbf{G} \mathbf{G}^T \quad (14b)$$

式(14)で $dJ/d\tau$ は正定値行列 $\mathbf{Q}$ を挟んだ位置誤差の二次形式となっており、常に負の値を持つ。従って式(13)の $d\alpha/d\tau$ を用いて繰り返しステップの適用による筋指令 α の修正値 $\Delta\alpha$ を式(15)のように決定すれば、式(11)による筋指令の更新は観測点における位置誤差

をしだいに減らしていき、推定動作を目標動作に収束させる。

$$\Delta\alpha = \eta \frac{d\alpha}{d\tau} = \eta [(x_{ref} - x)G]^T \quad (15)$$

式(15)で η はゲインで、収束性の確保や収束速度の観点から最適の値を選択する必要がある。本研究では数回の試行実行によって η を決定し、適用している。

4・3 逆推定筋指令による調音運動制御 調音筋指令の逆推定による調音動力学系モデルの制御結果を以下に示す。調音制御によって達成する目標動作としては、日本人男性被験者による実際の発声運動の計測から得た時系列の口唇運動情報を用いる。初期条件としては動きのない停止状態、また全ての筋における0の初期指令を設定する。

目標動作の運動学的情報は光学式3次元運動計測装置による測定から取り込む。時系列の口唇運動情報を求める観測点には計測装置による認識のための半球状のマーカーを貼り付ける。本計測実験では図9に示した4つの口唇運動観測点及び、3つの補助観測点にマーカーを貼り付けておく(図12)。

Fig.12 Markers attached on 4 observation points (OP) and 3 points for auxiliary measurement

図13に逆推定筋指令によって目標動作を追従する口唇運動制御の結果を示す。目標動作としては被験者による日本語母音/a/の発声動作が用いられている。時刻 $t = 0.00, 0.08, 0.12, 0.34, 0.48$ 及び $0.60s$ における目標動作(左列)が推定動作(右列)と比較されているが、8回の繰り返しを経て求めた推定動作が全時間領域にわたって目標動作へ良く収束している。

$t=0.00s$

Fig.13 Comparison of lip motion sequences at $t = 0.00, 0.08, 0.12, 0.34, 0.48$ and $0.60s$:

Reference (left column) vs. Estimated (right column)

図14に逆推定から求めた口唇運動制御における入力としての筋指令を示す。GDSによる逆推定は筋Jaw Openerへの指令を有意義なレベル以上に活性化し、発生した筋力による収縮運動が図13の推定動作を引き起こしている。

Fig. 14 Transition of estimated muscle motor command issued to jaw opener

図 15 に繰り返し推定の進行に伴う評価関数の推移を示す。8 回の繰り返し推定から得た筋指令が誘発した推定動作が、十分小さい誤差で目標動作へ収束していることが分かる。

Fig. 15 Transition of performance index

5. 考察とまとめ

本研究では人間の発話学習過程において働いていると考えられる調音制御について、動力学系モデルを用いた逆動力学的手法で考察した。調音動力学系モデルは、人間の調音運動を計算力学的手法で解明するために開発されたものであり、口唇軟組織と関連筋群及び骨構造を要素として構成されている。離散近似系として表現した口唇軟組織モデルが、静的変形において連続体系と適合性を持つよう、離散近似系を構成する粘弾性要素の剛性最適調整法を開発し、適用した。

口唇運動情報を出力とする逆力学問題の解として筋指令を求め、人間調音制御のシミュレーションを行った。求められた筋指令は良い精度で推定動作を目標動作に収束させており、逆動力学的手法が人間の調音制御学習過程を表現する数理モデルの一つとして有効であることを確認した。しかし本逆力学問題にはその解析にあたって不良設定や局所解の問題が潜在して

おり、その解決への有力な手段としては、口唇運動において物理的妥当性を持つ拘束条件の発掘及びその適用が考えられる。

現段階の調音動力学系モデルには口唇のみがモデル化されているが、今後はその他の重要な調音器官である舌及び声道の動力学系モデルも構築し、口唇モデルと結合することで自己完結型の調音動力学系モデルの完成を目指している。また現在は発声音の代わりに口唇運動の運動学的情報を発声運動の評価指標や調音制御の出力として用いているが、真の人間調音・発声を模擬するためには、調音運動情報から発声音を生成する音声生成モデルを構成要素として含む調音動力学系モデルの開発が必要と考えられる。

文 献

- (1) Platt, S.M. and Badler, N.I., Animating Facial Expressions, *Proc. SIGGRAPH81, ACM SIGGRAPH*, (1981), pp.245-252.
- (2) Dang, J. and Honda, M., A Physiological Model of a Dynamic Vocal Tract for Speech Production, *Acoust. Sci. and Tech.*, Vol. 22, No.6 (2001), pp. 415-425.
- (3) Nozoe, J., Gomi, H., Dang, J. and Honda, K., A Physiological Model of Perioral Dynamics for Realistic Deformation during Speech, *The IEICE Transactions on Information and System*, Vol. J88-D-II, No.9 (2005), pp.1944-1953 (in Japanese).
- (4) Zemlin, W.R., Speech and Hearing Science, (1998), Allyn and Bacon.
- (5) Lee, Y., Terzopoulos, D. and Waters, K., Realistic Modeling for Facial Animation, *Proc. SIGGRAPH95, ACM SIGGRAPH*, (1995), pp. 55-62.
- (6) Grzeszczuk, R. and Terzopoulos, D., Automated Learning of Muscle-Actuated Locomotion through Control Abstraction, *Proc. SIGGRAPH95, ACM SIGGRAPH*, (1995), pp. 63-70.
- (7) Dawe, L.J., Matrix and Finite Element Displacement Analysis of Structures, (1984), Oxford University Press.
- (8) Hill, A.V., The Heat of shortening and the Dynamic Constraints of Muscle, *Proc. Roy. Soc. London*, B 126 (1938), pp. 136-195.
- (9) Ito, T., Murano, E.Z. and Gomi, H., Fast Force-Generation Dynamics of Human Articulatory Muscles, *J. Appl. Physiol.*, Vol.96, No.6 (2004), pp. 2318-2324.
- (10) Laboisiere, R., Ostry, D.J., and Feldman, A.G., The Control of Multi-Muscle Systems: Human Jaw and Hyoid Movements, *Biol. Cybern.*, Vol. 74, No. 4 (1996), pp. 373-384.
- (11) Gomi, H., Piao, X., Nozoe, J., Dang, J. and Honda, M., Movement Control of Articulatory Dynamics Model by Imitation Learning, *Technical Report of IEICE*, NC2003-124 (2004), pp.31-36 (in Japanese).
- (12) Miyamoto, H., Schaal, S., Gandolfo, F., Gomi, H., Koike, Y., Osu, R., Nakano, E., Wada, Y. and Kawato, M., A Kendama Learning Robot Based on Bi-Directional Theory, *Neural Networks*, Vol. 9 (1996), pp. 1281-1302.