Universidade de Brasília
Instituto de Ciências Exatas
Departamento de Ciência da Computação

Sintese de Voz

Leandro Ramalho Motta Ferreira

Monografia apresentada como requisito parcial
para conclusão do Curso de Computação — Licenciatura

Orientador
Prof. Dr. Jorge Carlos Lucero

Brasília
2016

Universidade de Brasília — UnB
Instituto de Ciências Exatas
Departamento de Ciência da Computação
Curso de Computação — Licenciatura

Coordenador: Prof. Dr. Coordenador

Banca examinadora composta por:

Prof. Dr. Jorge Carlos Lucero (Orientador) — CIC/UnB
Prof. Dr. Professor I — CIC/UnB
Prof. Dr. Professor II — CIC/UnB

CIP — Catalogação Internacional na Publicação

Ferreira, Leandro Ramalho Motta.

Sintese de Voz / Leandro Ramalho Motta Ferreira. Brasília : UnB, 2016.
73 p. : il. ; 29,5 cm.

Monografia (Graduação) — Universidade de Brasília, Brasília, 2016.

1. Sintese, 2. Voz, 3. Saúde

CDU 004.4

Endereço: Universidade de Brasília
Campus Universitário Darcy Ribeiro — Asa Norte
CEP 70910-900
Brasília–DF — Brasil

Universidade de Brasília
Instituto de Ciências Exatas
Departamento de Ciência da Computação

Sintese de Voz

Leandro Ramalho Motta Ferreira

Monografia apresentada como requisito parcial
para conclusão do Curso de Computação — Licenciatura

Prof. Dr. Jorge Carlos Lucero (Orientador)
CIC/UnB

Prof. Dr. Professor I
CIC/UnB

Prof. Dr. Professor II
CIC/UnB

Prof. Dr. Coordenador
Coordenador do Curso de Computação — Licenciatura

Brasília, 10 de maio de 2016

Dedicatória

Dedico a....

Agradecimentos

Agradeço a....

Resumo

AINDA Não tem

Palavras-chave: Sintese, Voz, Saúde

Abstract

Still there isn't.

Keywords: Synthesis, Voice, Health

Sumário

Lista de Figuras

Lista de Tabelas

Capítulo 1

Introdução

1.1 Motivação

A UnB comporta pessoas de qualidade altíssima que focam seu trabalho na pesquisa ná area da simulação de voz, portanto tenho a oportunidade de desenvolver pesquisa com maior apoio possível o que é indispensável. Apesar dos esforços e trabalho serem na simulação de voz no contexto na área da saúde, simulação de voz com patologias, exista oportunidade para a simulação de voz no entretenimento musical que pode ser explorada. Coloco também como motivação o sonho pessoal de simular a voz que foi despertada ao longo do curso.

Capítulo 2

Conceitos Básicos de Sintese de Voz

2.1 Anatomia da Voz

Para estudar a produção e a síntese da voz, é necessário ter um conhecimento acerca da anatomia e do funcionamento físico da voz (10). Sendo assim, as subseções seguintes descreverão brevemente detalhes da anatomia do sistema fonador humano e como o som é produzido, moldado e inuenciado por este sistema.

2.1.1 A Anatomia

A Figura 2.2 (10), mostra os órgãos associados com a produção da voz. Dentro das condições normais, a voz é produzida quando um uxo de ar vindo dos pulmões é convertido em energia acústica através da vibração das pregas vocais, localizadas na laringe. Os padrões de vibrações resultantes são moldados acusticamente quando o som passa pelo trato vocal acima da laringe. O sistema respiratório serve como uma

fonte de potência para a produção do som, sendo responsável por movimentar o ar através do trato vocal. A laringe atua como um oscilador convertendo a potência aerodinâmica produzida em energia sonora, sendo frequentemente retratada como a fonte da voz. No entanto, a mais importante função da laringe não é a produção de som, e sim, vedar as vias aéreas aos pulmões completamente, protegendo-as de objetos estranhos ou líquidos, principalmente durante a deglutinação. De maneira análoga, a laringe serve como uma válvula de acesso às vias respiratórias e por essa característica, atua também no controle do uxo de ar que por elas passam. Sendo assim,é fácil notar que há uma necessidade de mobilidade para toda estruturada laringe, logo é de se esperar que sua estrutura seja formada em sua maioria por cartilagens. De fato o é, com exceção de um osso chamado de Hioide, a laringe é basicamente formada por cartilagens e músculos. A seguir, analisaremos brevemente a dinâmica dos músculos e cartilagens da laringe.

2.1.2 Músculos e Cartilagens

Os músculos e cartilagens atuam diretamente no processo de abdução e adução das pregas vocais. Estas estão localizadas dentro da laringe e devido à dinâmica das cartilagens e dos músculos, podem executar os movimentos citados de forma a produzir som.

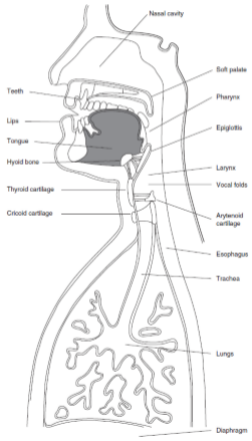


Figura 2.1: Aparelho Fonador

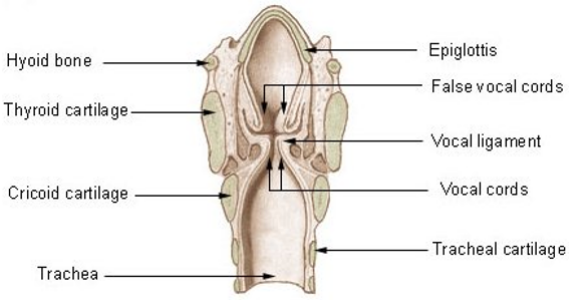


Figura 2.2: : Secção coronal da laringe e parte superior da traquéia. (10)

Cartilagens da Laringe

De maneira sucinta, estas cartilagens servem como base de interconexão para os músculos intrínsecos ao redor da laringe. Dentre as cartilagens acima, a epiglote é responsável por vedar as vias respiratórias movimentando-se sobre a entrada das mesmas. O resto das cartilagens garantem a mobilidade da laringe em conjunto com outras estruturas como por exemplo o sternum.

Os músculos na laringe podem ser divididos em dois grupos, os intrínsecos e os extrínsecos [1]. Os músculos intrínsecos interconectam as cartilagens da laringe, ao passo que, os extrínsecos conectam a laringe à outras estruturas externas, como o osso hióide. A Figura 2.3 detalha alguns dos músculos intrínsecos da laringe. Alguns desses músculos têm inuência direta em algumas características da voz. Por exemplo, o músculo cricotiroideo é o músculo primário utilizado no controle do tom da voz. Por sua vez, o músculo cricoaritenoideo posterior atua na abdução das pregas vocais, ao passo que o músculo interaritenoideo atua como adutor das pregas vocais.

Os músculos extrínsecos, Figura 2.3, atuam basicamente no movimento da laringe, agindo como depressor e elevador da estrutura laríngea. Além disso também conectam estruturas do trato vocal à estrutura laríngea, como por exemplo a língua ao osso hioide.

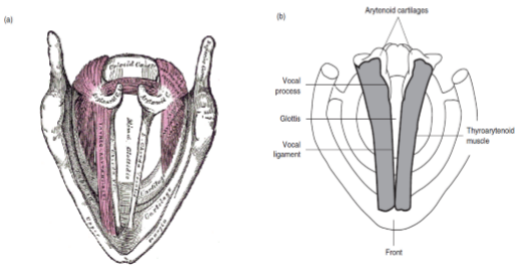


Figura 2.3: : Músculos Intrínsecos da Laringe. (1)

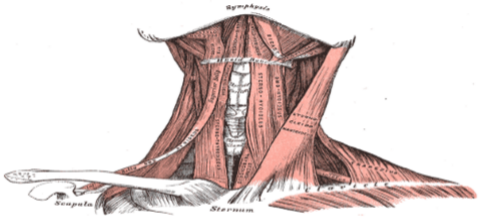


Figura 2.4: :Músculos Extrínsecos da Laringe. (1)

2.1.3 Pregas Vocais

As pregas vocais, como dito anteriormente, estão localizadas dentro da laringe, mais especicamente na parte superior da traqueia. Elas estão posteriormente ligadas às cartilagens aritenoides, e anteriormente ligadas à cartilagem tireoide. As suas bordas exteriores estão ligadas a músculos na laringe, enquanto as suas bordas interiores são livres.

As bordas das pregas vocais são construídas de epitélio, sendo compostas também de algumas bras musculares. As pregas vocais são bandas triangulares planas de cor branca e acima de ambos os lados destas, se encontram as pregas vestibulares ou falsas pregas vocais. O espaço entre as pregas vocais é chamado de glote, sendo que o que está acima da glote é denominado supraglotal e o que está abaixo é denominado subglotal. A Figura 2.5 mostra em mais detalhes a anatomia das pregas vocais, os componentes musculares e as cartilagens atuantes.

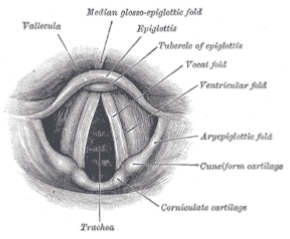


Figura 2.5: :Cordas Vocais e Componente (1)

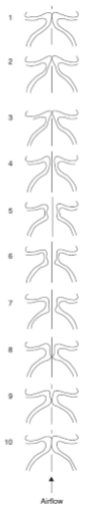


Figura 2.6: :Movimento das Cordas Vocais (10)

2.1.4 Overview: Geração de Som e Ressonadores

A produção do som da voz é composta de duas etapas importantes que ocorrem na região subglotal/glotal e supraglotal. A primeira é a transformação da energia aerodinâmica em energia sonora, pelo movimento e vibração das pregas vocais.

O segundo é a transformação do som primitivo gerado em voz, através da atuação dos ressonadores e formantes na região supraglotal.

A vibração das pregas vocais é extremamente complexa e, diversos músculos em união com a pressão exercida pelo ar atuam para tornar esse movimento possível. De maneira sintetizada, as pregas vocais vibram do topo ao fundo de maneira que não vibram como se fosse um bloco, mas sim de forma ondulatória, conforme mostrada na Figura 2.6. Essa vibração é responsável por mudanças depressão entre a região subglotal e a região glotal, causando uma diferença de pressão(condensação e rarefação) nessa área durante o movimento das pregas vocais, o que ocasiona a geração de som.

Entretanto, o som produzido é um som primitivo, conforme dito anteriormente, e para se transformar na voz característica humana, ele deve ser ltrado e moldado pelos ressonadores(formantes) no trato vocal(Figura 2.7). Todas as cavidades mostradas no trato vocal atuam como ressonadores para a onda sonora produzida pelas pregas vocais. Um ressonador pode entrar em estado de vibração através de uma força aplicada ao mesmo em inércia ou por interação com algo que já esteja em estado de vibração.

Neste segundo caso, as vibrações produzidas pelo ressonador serão amplicações ou atenuações dependendo de quão próximas ou distantes, em termos de frequência, são as vibrações da onda sonora em contato com o ressonador. Caso a onda sonora possua vibrações cujas frequências se assemelhem às frequências do ressonador,estas então serão amplicadas pelo ressonador.

Entretanto,caso as frequências do som gerado vibrem em uma frequência distante da frequência natural do ressonador, então estas serão abafadas. A voz passa por esse processo ao ser formada. Um som primitivo advindo das pregas vocais entra em contato com os ressonadores no trato vocal, estes por sua vez em conjunto

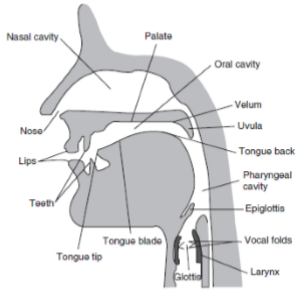


Figura 2.7: :Trato Vocal (10)

Os conceitos e propriedades descritos acima são extremamente importantes para se entender a manutenção da produção do som. As pregas vocais são músculos e músculos são compostos por bras, logo, as pregas vocais consistem de uma grande concentração de bras. Além disso, entre as bras que compõe as pregas vocais existem também uidos atuantes, o que caracteriza as pregas vocais como um material viscoelástico.

Para se entender a capacidade de absorção e regeneração das pregas vocais, em detrimento das vibrações de alta frequência e as pressões do ar, deve-se primeiramente estudar as propriedades absorcivas do material que as compõe. Ou seja, em outras palavras, deve-se estudar as propriedades mecânicas do tecido viscoelástico, e uma ferramenta que facilita o entendimento é o estudo da curva força-alongamento de um material. Entretanto, construir uma curva de força-alongamento depende essencialmente da geometria da amostra do material (16) e, por se tratar de uma material biológico, é difícil obter uma geometria precisa pois as bras estão constantemente se reorientando em detrimento de lesões e cortes brosos. Para viabilizar este estudo, Titze (16) sugere normalizar as forças atuantes e as deformações resultantes para que não haja a dependência direta da geometria. Essa normalização se dá através da substituição da curva força-alongamento por um curva tensão-deformação. A Figura 3.1 retirada do estudo feito por Titze (16) demonstra uma curva hipotética de tensão-deformação para os tecidos que compõe as pregas vocais humanas.

Esta Figura ilustra o comportamento das bras das pregas vocais através da relação entre uma força atuante e a deformação gerada por esta. A importância desta analises e deve ao fato de que é possível estabelecer uma relação direta entre nódulos vocais e uma fonação prolongada, alta(em termos de frequência) e intensa. A partir da análise desta curva é então possível estabelecer um precedente para a formação de nódulos vocais: a frequência e amplitude da vibração estão diretamente ligadas ao surgimento de um nódulo vocal e consequentemente o de uma fenda pois a força de impacto entre as pregas vocais é proporcional à altura tonal quando acima do tom natural e à intensidade durante a fonação.

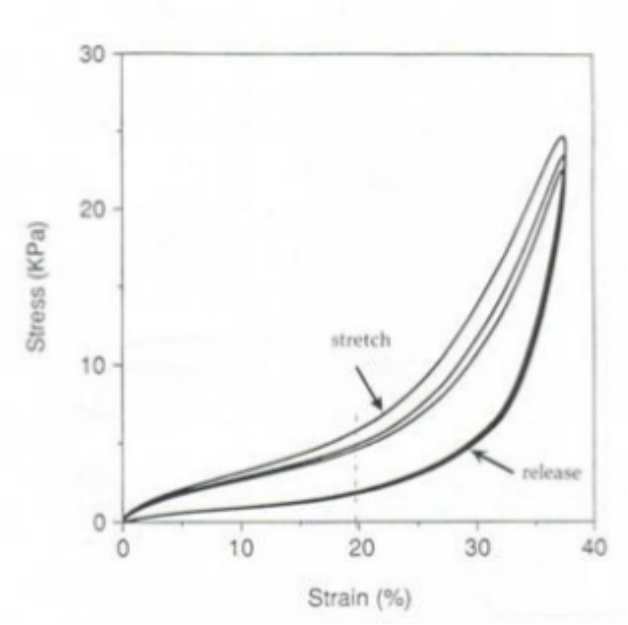


Figura 2.8: :Curva Hipotética Tensão-Deformação das Cordas vocais Humanas (16)

Capítulo 3

Fundamentos Básicos para a Produção da Voz

Neste capitulo serão introduzidos princípios físicos e biológicos da produção da voz. Além dos conceitos introduzidos aqui, serão analisados os processos mecânicos da geração da voz por parte da laringe em conjunto com o trato vocal pois são de extrema importância para a concepção dos modelos computacionais e matemáticos utilizados para se desenvolver o sintetizador.

3.0.1 A Biomecanica da Laringe

Primeiramente, devemos ter em mente que o principal papel da laringe não é a produção de voz e sim a proteção das vias respiratórias. Dito isto, podemos fazer uma simples analise, visto que apesar de seu papel principal, a laringe também atua como um instrumento da fala humana. Se analisarmos os instrumentos criados pelo homem podemos notar que estes dependem basicamente de sua geometria, do material que o compõe e da interação de suas partes acústicas. Do mesmo modo, a laringe possui uma determinada geometria, é composta por tecido humano e em conjunto como trato vocal compõe a parte acústica do nosso corpo. Entretanto, nada é tão simples, a sua geometria e as propriedades do material humano envolvidos na produção do som são bastante irregulares (16). Outra analogia interessante é sobre o instrumento e quem o utiliza. Um bom pianista por exemplo, sua musica é boa porque ele é habilidoso como instrumento? Ou sua musica é boa por que o instrumento é bem feito e o som gerado por este é agradável? Ou os dois? Essas perguntas também podem ser feitas com respeito a voz. Para entendermos o que inuencia na qualidade da síntese da voz é necessário analisar a biomecânica da voz, que nada mais é analisar o movimento do material vivo e as forças atuantes sobre ele (16).

3.0.2 Fatores Biológicos que Afetam a Produção de Som na Laringe

A parte da (bio)mecânica que se relaciona diretamente com a atuação da laringe na produção do som é a mecânica dos meios contínuos, que é a parte da mecânica que lida com a matéria distribuída sobre uma determinada região no espaço, e consequentemente, se contrapõe à mecânica de partículas.

Dentro da mecânica de meios contínuos, mais especicamente, a parte que irá nos auxiliar no estudo do comportamento da laringe se chama mecânica de sólidos e uidos. Dito isto, analisaremos a seguir alguns conceitos físicos que tem forte ligação com os processos que ocorrem na região da laringe durante a produção do som:

Tensão e Deformação

São características de forças atuantes em superfícies, como por exemplo a resistência do ar.

Tensão

é quantidade de força por unidade de área (16), podemos escrever na forma da equação 3.0.2 :

$$\sigma = \frac{f}{A}$$

(3.1)Onde é a tensão, f é a força aplicada e A a área de aplicação desta força.

Deformação

é a medida de deformação de um meio após a aplicação de uma tensão (16) e pode ser escrito na forma da equação :

$$\epsilon = L - \frac{L_0}{L_0}$$

(3.2)

Onde é a medida de deformação, L é o comprimento após a tensão e L0 é o comprimento antes da tensão.

Normalmente uma dada deformação em uma dimensão resulta em uma deformação oposta em outra dimensão em um dado meio. Se uma deformação é uniforme por todo o corpo de um objeto, então chamamos de compressão, se o volume diminui por conta desta deformação, e expansão, se o volume aumenta.

Viscosidade

É a velocidade de deformação(consequentemente, de restauração) de um determinado uido quando atuam forças de tensão no mesmo. Matematicamente pode ser expresso conforme a equação 3.0.2:

$$\epsilon = \eta \frac{d_e}{d_t}$$

(3.3)

Para viscosidade e t tempo. Quanto maior a viscosidade, mais devagar será a deformação de um meio.

3.0.3 Modelo Massa-Mola Auto Sustentável

: A fechadura e abertura da glote num sistema massa mola de apenas um lado

$$P = (1 - \frac{a2}{a1}) * (Ps - Pi) + Pi$$

3.0.4 Elasticidade

É uma propriedade do meio que determina quão completa será a restauração do meio após uma dada deformação.

Os conceitos e propriedades descritos acima são extremamente importantes para se entender a manutenção da produção do som. As pregas vocais são músculos e músculos são compostos por bras, logo, as pregas vocais consistem de uma grande concentração de bras. Além disso, entre as bras que compõe as pregas vocais existem também uidos atuantes, o que caracteriza as pregas vocais como um material viscoelástico. Para se entender a capacidade de absorção e regeneração das pregas vocais, em detrimento das vibrações de alta frequência e as pressões do ar, deve-se primeiramente estudar as propriedades absorcivas do material que as compõe.

Ou seja, em outras palavras, deve-se estudar as propriedades mecânicas do tecido viscoelástico, e uma ferramenta que facilita o entendimento é o estudo da curva força-alongamento de um material. Entretanto, construir uma curva de força-alongamento depende essencialmente da geometria da amostra do material (16) e, por se tratar de uma material biológico, é difícil obter uma geometria precisa pois as bras estão constantemente se reorientando em detrimento de lesões e cortes brosos. Para viabilizar este estudo, Titze (16) sugere normalizar as forças atuantes e as deformações resultantes para que não haja a dependência direta da geometria. Essa normalização se dá através da substituição da curva força-alongamento por um curva tensão-deformação. A Figura 3.1 retirada do estudo feito por Titze (16) demonstra uma curva hipotética de tensão-deformação para os tecidos que compõe as pregas vocais humanas. Esta Figura ilustra o comportamento das bras das pregas vocais através da relação entre uma força atuante e a deformação gerada por esta. A importância desta analise se deve ao fato de que é possível estabelecer uma relação direta entre nódulos vocais e uma fonação prolongada, alta(em termos de frequência) e intensa. A partir da análise desta curva é então possível estabelecer um precedente para a formação de nódulos vocais: a frequência e amplitude da vibração estão diretamente ligadas ao surgimento de um nódulo vocal e consequentemente o de uma fenda pois a força de impacto entre as pregas vocais é proporcional à altura tonal quando acima do tom natural e à intensidade durante a fonação.

3.0.5 Fluxo de Ar nas Vias Respiratórias

Em seções posteriores serão apresentados modelos para representar a vibração das pregas vocais, entretanto para entender por completo esses modelos, devemos antes entender o papel do uxo de ar nesse processo e para isso devemos analisar primeiro como a pressão de ar é gerada e exercida.

À primeira vista, imagine-se que o uxo de ar é gerado somente pelos pulmões, entretanto, se formos analisar como inspiramos e expiramos percebemos que existem outros fatores contribuintes para a criação e manutenção de um uxo de ar nas vias respiratórias.

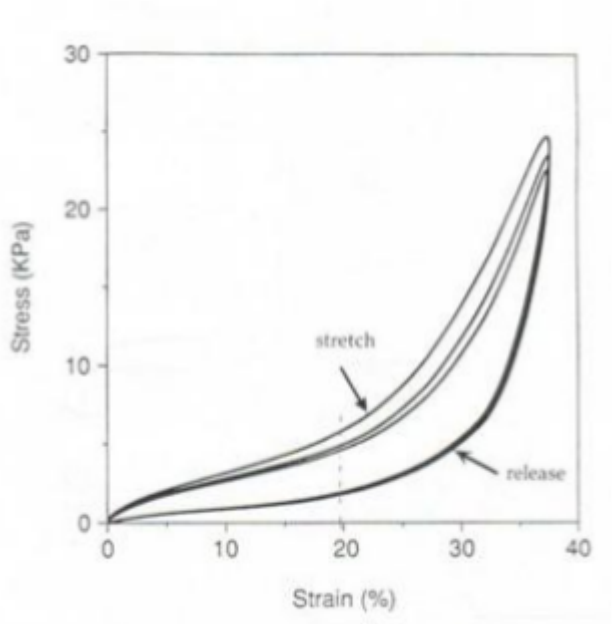


Figura 3.1: Curva Hipotética Tensão-Deformação das Cordas vocais Humanas (16)

O pulmão, como é de se imaginar, é responsável pelo armazenamento do ar e pela troca de oxigênio entre os tecidos que o compõe. Aqui, porém, somente nos interessa esse papel de armazenamento que o pulmão exerce. Ao inspirar o ar, o pulmão aumenta a pressão interna gerando uma tensão aplicada ao tecido elástico que o compõe, a expansão desse tecido gera uma mudança de pressão devido ao aumento do volume pulmonar e à deformação do tecido pulmonar. Em conjunto com essa mudança de pressão, o diafragma, músculo localizado abaixo dos pulmões, é comprimido levando a uma outra mudança de pressão. O diafragma, por sua vez, é um músculo que pode ser controlado, e em virtude disso, é possível controlar a pressão exercida nos pulmões, algo que cantores utilizam para sustentar a técnica vocal. Assim como o diafragma, a caixa torácica também sofre mudanças, expandindo devido ao aumento de volume do pulmão o que gera uma pressão torácica sobre os pulmões. Devido ao fato de ser uma estrutura óssea e portanto ter uma maior rigidez e menor controle se comparado à parede abdominal, a pressão exercida nos pulmões pela caixa torácica é signicantemente menor do que a pressão exercida pela aplicação de uma força à parede abdominal.

A Figura 3.2 ilustra essa transmissão de pressão por todo o torso citada acima:

O Sistema Pulmonar

O sistema pulmonar consiste dos pulmões e vias respiratórias, compostas pela traqueia, glótis e trato vocal. Conforme descrito anteriormente, o tórax e o abdômen em conjunto com o pulmão atuam na geração do uxo de ar. A glótis, por sua vez, atua como reguladora do uxo de ar através de variações em seu fechamento que permitem mudanças na pressão envolvida com o movimento do uxo do ar pelas vias respiratórias. Isso signica que é possível dar constância ao uxo de ar(Figura).

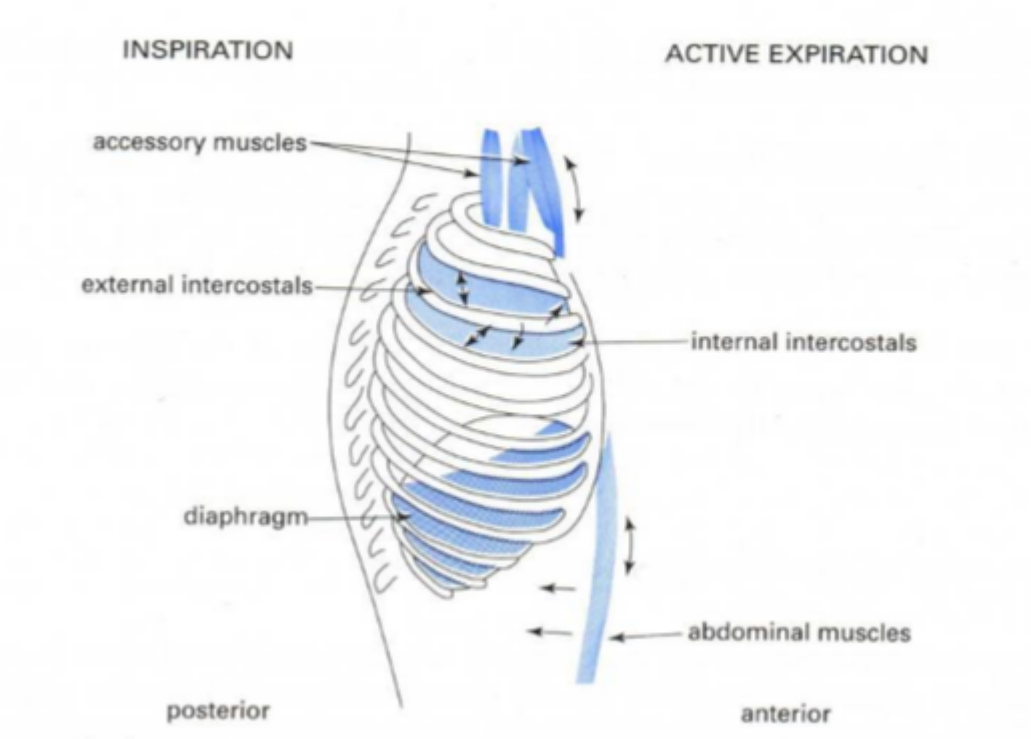


Figura 3.2: Pressões Atuantes no Torso (16)

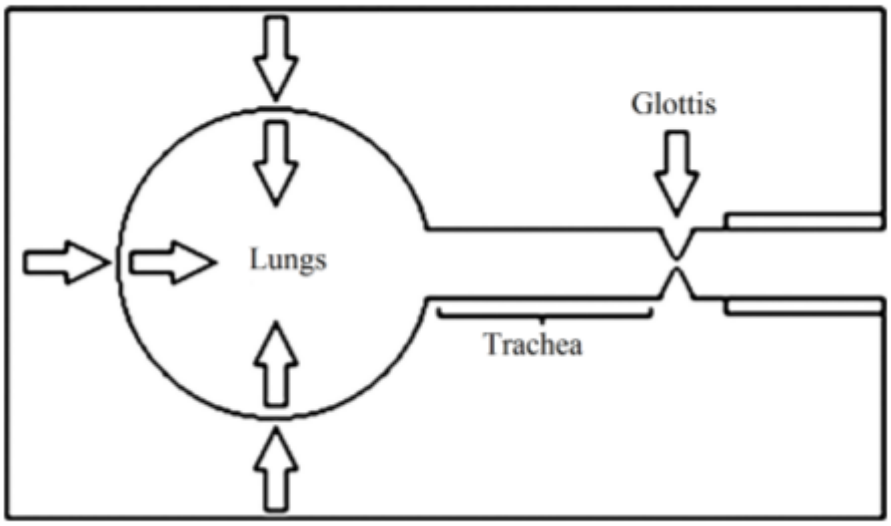


Figura 3.3: Sistema Pulmonar

O Proceso Físico da Respiração

Uma importante lei que nos auxilia no entendimento da relação entre o volume pulmonar e a pressão pulmonar é a Lei de Boyle. A lei nos diz que em um ambiente cujas paredes não são rígidas e em uma temperatura constante, pressão e volume são inversamente proporcionais.

Se aplicarmos a lógica dessa lei à respiração humana, ca fácil visualizar o processo físico que envolve a respiração. Quando inspiramos, o diafragma é contraído o que aumenta o volume do pulmão e diminui a pressão interna. Consequentemente, o pulmão se enche de ar. Relacionada à pressão atmosférica, a pressão pulmonar se torna menor, o que gera uma busca pelo equilíbrio das pressões no meio. Esse equilíbrio é alcançado ao se expelir o ar, diminuindo o volume pulmonar e então aumentando a pressão interna novamente. Na verdade é um pouco mais complexo do que isso, entretanto, esse conceito de equilíbrio de pressões é fundamental para o entendimento da abdução e adução das pregas vocais, o que permite a transformação da energia aerodinâmica em energia acústica.

3.0.6 Leis de Conservação para Fluxos em Dutos

Porém, antes de abordar a transformação de energia aerodinâmica em acústica, é importante se familiarizar com alguns conceitos que regem o comportamento de uidos em dutos. Isso porque trataremos o uxo de ar na síntese de voz como um uído e a modelagem da glote e do trato vocal como a concatenação de pequenos dutos.

Lei da Continuidade para um Fluxo Incompressível

Quando dizemos que um uido é incompressível signica dizer que sua densidade não se altera quando este é forçado a passar por uma constrição. Imagine agora um uido connado a um duto que possui uma mudança, ao longo de seu comprimento, em sua área transversal formando uma constrição(Figura 2.4). Se não for permitido ao uido vazar pelas paredes do duto, então todas as partículas do uido devem ser mantidas mesmo durante a mudança de área transversal ao longo do duto. Para que isso seja possível, as partículas devem acelerar durante a constrição, mantendo-se constante o número de partículas em movimento por unidade de área. Essa relação pode ser expressa pela equação 3.0.6.

$$v_1 * A_1 = v_2 * A_2 = constante = U \tag{3.4}$$

Onde v e A são respectivamente a velocidade da partícula e a área em que se encontra. U é o uxo. Matematicamente então, a Lei da Continuidade expressa que um uxo incompressível em um duto é constante, independente do que acontece com a área transversal ao longo do mesmo (16).

Lei de Bernoulli para Conservação de Energia

Outra importante lei de aplicação para o entendimento da síntese de voz é a Lei de Bernoulli para Conservação de Energia. Essa lei foi desenvolvida a partir do reconhecimento de que a energia total atuante no uido em qualquer ponto durante o trajeto no duto tem duas componentes, a saber, uma energia potencial e uma energia cinética. A

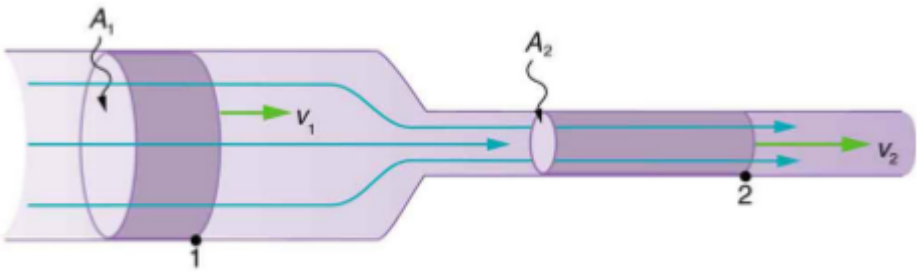


Figura 3.4: Dinâmica do Fluido em um Duto com Mudança de Área Transversal

energia potencial está relacionada diretamente com a pressão no duto e a energia cinética é proporcional ao quadrado da velocidade da partícula. Portanto, podemos expressá-la conforme a equação

$$P + \frac{1}{2}\rho * v^2 = constante$$

(3.5)

Onde ρ é a densidade do uido, v a velocidade da partícula e P a pressão no duto. Segue disso o Princípio de Bernoulli: Se a energia em uma corrente de uido é constante, um aumento na velocidade da partícula deve ser acompanhado de uma queda na pressão (16).

Analisando a Figura **??**, a pressão na área de constrição deve ser menor que a pressão na área maior do duto, considerando que não hajam perdas de energia no processo. Esse conceito é importante para se compreender o caráter oscilatório das pregas vocais.

Resistência Glotal

A resistência a um uxo é uma característica do sistema de transporte, ou seja, do meio, e pode ser descrita como a razão entre pressão e uxo(equação 3.0.6)

$$R = \frac{P}{U}2$$

(3.6)

Em sistemas de transporte, as constrições atuam como ponto de resistência a um uxo. No caso das vias respiratórias humanas, isso normalmente ocorre na glote ou em alguma parte do trato vocal com pouco espaço. Podemos então denir que a resistência glotal será a pressão na glote dividida pelo uxo que a atravessa. A resistência glotal está diretamente ligada a qualidade vocal pois é responsável por auxiliar no controle do uxo de ar.

3.0.7 Oscilação das Pregas Vocais

Nesta seção analisaremos os processos e princípios físicos que ocorrem na laringe para que as pregas vocais oscilem, tendo como objetivo destrinchar tais processos a m de entendermos a reprodução dos mesmos por modelos computacionais utilizados para simulação da voz humana.

Os primeiros estudos sobre vibração das pregas vocais especulavam que as pregas vocais se ajuntavam por um efeito de pressão negativa(Princípio de Bernoulli) na glote. A esta descrição de vibração nas pregas vocais se deu o nome de Teoria Mioelástica Aerodinâmica da Vibração das Pregas Vocais, conhecida também pelo nome de seu criador como Teoria de van der Berg (18). A teoria de van der Berg serviu como marco para o desenvolvimento teórico do campo de estudos da voz e consequentemente o aparecimento dos primeiros modelos matemáticos de representação do funcionamento das pregas vocais (8).

Entretanto, apesar disso, sua teoria é inadequada para explicar a vibração auto sustentável das pregas vocais. Isso porque as forças de Bernoulli por si só não são capazes de distinguir entre os movimentos interiores e posteriores das pregas vocais. Assim sendo, são necessários mecanismos para prover um aumento ou decréscimo às forças de Bernoulli durante a abertura e o fechamento das pregas vocais respectivamente, caso contrário, as oscilações serão amortecidas.

3.0.8 Critérios para Oscilação

3.0.9 Critérios para Oscilação

Alguns critérios devem ser atendidos para que um determinador padrão de movimento seja considerado como uma oscilação mecânica, a saber:

No sistema onde ocorre o movimento deve haver uma posição de equilíbrio estável, que é caracterizada por uma força restaurativa que sempre acelera o corpo em movimento de volta para a sua posição de repouso. Deve haver inércia(no caso do sistema mecânico, a massa atua como propriedade de inércia) no sistema para superar esta posição de equilíbrio. A perda, em excesso, de energia por ciclo de oscilação deve ser zero.. . .

3.0.10 Tipos de Oscilação

De acordo com Titze (16), os tipos de oscilação são:

- Oscilação Natural: Quando um sistema que se encaixa nos critérios anteriores se move sem interferência após um distúrbio inicial.
- Oscilação Natural: Quando um sistema que se encaixa nos critérios anteriores se move sem interferência após um distúrbio inicial.
- Oscilação Forçada: Requer uma fonte externa de condução que por si só é um oscilador. Dita grande parte do padrão de vibração do sistema.
- Oscilação Auto-Sustentável: Requer uma fonte de energia estável e uma interação não-linear entre os componentes internos ao sistema. As perdas de energia são compensadas, mantendo o padrão oscilatório.

3.0.11 Tensao

Conceito de Tensão For por unidade de ar

$$\sigma = \frac{F}{A}$$

Sendo: f : força aplicada.
A: área de aplicação desta força.

3.0.12 Curva Força e Alongamento

Utiliza-se para não ser dependente da geometria do material. Utilizamos nas cordas vocais(?) por serem materias biológicos. Cria uma figura ilustrando comportamento da deformação das pregas vocais

3.0.13 Viscosidade

É a velocidade de deformação(consequentemente, de restauração) de um determinado fluido quando atuam forças de tensão no mesmo. Matematicamente pode ser expresso conforme a equação seguinte: **?**

$\sigma = \eta * \frac{d*\epsilon}{dt}$

3.0.14 Reflexao de Som

Um fenômeno ligado a rigidez e amortecimento entre um meio e outro.(9) Ondas quando tentam penetrar em um segundo meio, sendo o segundo meio rígido, as partículas primeiro meio se aglomeram tentando passar porém falham, seu acúmulo de particulas geram pressão que acabam criando uma outra onda no primeiro meio decorrente da primeira onda.(6)

O mesmo ocorre com o meio 2 sendo totalmente não rígido e o primeiro meio sendo bem rígido, Exaurindo excesso de particulas do meio 1 no meio 2 criando rarefação no meio 1, o que cria uma outra onda de pressão negativa (2). A propagação é sempre em direção oposta à fonte, no caso é na direção contrária à coluna de ar(meio 1).

3.0.15 Geração e Propagação do Som

Fisicamente, o som é gerado quando há uma perturbação do equilíbrio da densidade(ou pressão) de um gás, liquido ou sólido (16). Se a perturbação na pressão é positiva, então haverá uma condensação no meio e portanto um aumento na densidade do ar. Se a perturbação for negativa, haverá uma rarefação no meio o que ocasionará um decréscimo na densidade do ar. Considere um pistão em um tubo cilíndrico preenchido de ar. Pelo movimento oscilatório do pistão, o ar tende a condensar ou rarefazer. Sendo assim, as partículas em contato com o pistão tendem a seguir o movimento oscilatório do mesmo, porém, por não estarem rigidamente conectadas nem à superfície do pistão nem entre si mesmas, elas estabelecem um padrão próprio de movimento em detrimento das colisões. Dessa forma, quantidade de movimento é transmitida às partículas adjacentes, criando uma nova pertubação em um ponto mais longe ao pistão. A este fenômeno se dá o nome de propagação. Da mesma maneira, acontece com a propagação do som na glote.

Podemos pensar a respeito da geração do som na glote com a seguinte analogia: o movimento das pregas vocais pode ser comparado ao movimento de bater palmas, a cada choque das mãos, uma perturbação é gerada e consequentemente, som é gerado. Se pensarmos no papel da glote na geração do som, podemos pensar que a glote é como se

fosse uma sirene, o uxo de ar que passa por ela é espremido no canal o que leva a uma perturbação e consequentemente a geração do som.

A partir do conceito de propagação de ondas sonoras é possível deduzir o conceito de reexão sonora, que está ligado à propagação da onda em um meio com propriedades de impedância acústica. Devido a estas propriedades de impedância, a propagação da onda ao entrar em contato com a superfície do meio, é alterada.

Capítulo 4

Computação e Síntese de Voz

4.0.1 Modelo Mono-Massa

Os modelos de uma massa são os mais simples para se representar as pregas vocais,com essa única massa representando toda a estrutura das pregas vocais. Como é representado por uma massa somente, o grau de liberdade neste sistema é 1, ou seja, o deslocamento é dado somente em uma direção (4). Conforme foi citado no Capítulo 2, para um modelo viscoelástico das pregas vocais apresentar auto-oscilação, energia deve ser periodicamente adicionada ao sistema de modo que supere as fricções internas do tecido laríngeo. Tal energia é acrescida ao sistema por uma carga externa quando a mesma está em fase com a velocidade de vibração das pregas vocais.

A modicação na pressão aerodinâmica devido às mudanças de orientação convergente divergente da glote é o mecanismo primário pelo qual o supracitado ocorre siologicamente. Por denição, os sistemas mono-massa não podem desenvolver por si só a variação(no tempo) glotal necessária para que as pressões supraglotal e glotal sejam carregadas assimetricamente, permitindo que o efeito de Bernoulli aja em reconhecimento à abertura e fechamento da glote. Para que isso ocorra neste sistema, deve haver uma interação da glote com um tubo acústico supraglotal ou subglotal 16.

Considere agora o sistema representando cada prega vocal, cujos componentes são uma massa m, uma constante de rigidez k e uma constante de amortecimento b. A constante de amortecimento representa a viscosidade do tecido, ou seja, atua como um absorvedor de energia. A constante k representa as propriedades elásticas do tecido ao passo que m é a massa do tecido em movimento, Figura 4.1.

A pressão P na glote atua perpendicular à superfície do tecido. Se esta pressão se alterar de acordo com a direção da velocidade, então energia será transmitida ao tecido pelo uido, e esta pode superar a energia perdida com a viscosidade do tecido. Entretanto, como essa pressão pode ser diferente ou as simétrica no caminho de volta do que foi ao sair? Ou seja, como a pressão pode ser alterada e consequentemente acompanhar e inuenciar a abertura e o fechamento da glote?

Titze derivou uma expressão para a pressão intraglotal na superfície das pregas vocais[26] e forneceu uma versão simplicada dessa expressão para o modelo em questão (16), conforme mostra a equação 4.0.1

$$P = (1 - \frac{a_2}{a_1}) * (P_s - P_i) + P_i$$

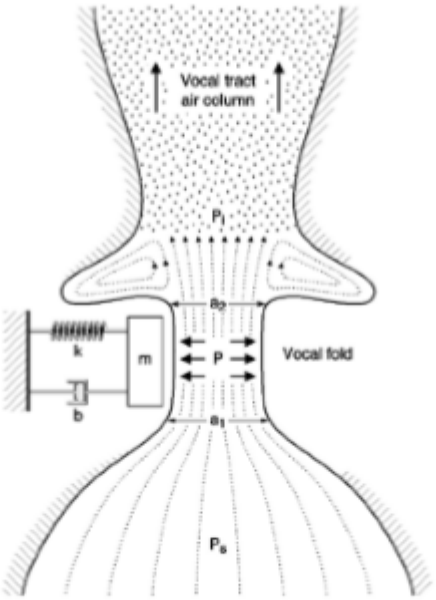


Figura 4.1: Modelo Mono-Massa (16)

(4.1)

Onde a_1 e a_2 são as áreas de entrada e saída da glote respectivamente, Ps é a pressão subglótica e Pi é a pressão exercida sobre o trato vocal, chamada de pressão de input. O primeiro termo da equação é um fator geométrico que descreve o formato da glote. O segundo termo, a diferença de pressões, descreve a pressão transglotal, ou seja, aquela que percorre a glote. Para o modelo mono-massa, temos que a1 = a2 o que faz com que a pressão intraglotal seja simplesmente equivalente à pressão supraglotal.

A relação resultante de que a pressão condutora P é igual à pressão supraglotal sugere que algo deve acontecer acima da glote para que essa pressão seja alterada durante o ciclo glotal. O elemento chave aqui é a inércia do ar no trato vocal. A lentidão em resposta da coluna de ar acima das pregas vocais causa uma outra condição de superação que auxilia na vibração, isto porque quando a glote está abrindo e o uxo de ar aumentando, a coluna de ar está sendo acelerada pelo uxo glotal. Isto cria uma pressão positiva Pi na entrada do trato vocal, fazendo com que as pregas vocais se separem. Essa pressão positiva gera também um aumento no ímpeto(ou na quantidade de movimento) da coluna de ar. Quando a glote fecha, o ímpeto da coluna de ar continua e o uxo na glote não se sustenta com o uxo da coluna de ar, gerando uma pressão negativa(sucção) acima das pregas vocais auxiliando no fechamento das pregas vocais. Portanto, Pi conduz as pregas vocais em sincronia com seu movimento natural.

Essa explicação é muito similar ao efeito de Bernoulli em ação, entretanto, um termo adicional muito importante está envolvido: o atraso na resposta à pressão da coluna de ar no trato vocal.

Dessa maneira, ca claro que em decorrência do grau de liberdade do modelo mono massa, não há como este modelo reproduzir uma movimentação ondular das pregas vocais, o que auxiliaria na criação e manutenção da auto-oscilação. Em decorrência disso, o trato

vocal atua suprindo esta falta de modo que o mesmo possibilita a alteração nas pressões durante as fases de movimento das pregas vocais o que induz uma auto-oscilação no sistema mono-massa, conforme explicado.

4.0.2 Modelo Computacional para Representação do Trato Vocal

Para a representação do trato vocal, Fraj (5), Titze (16) descrevem um modelo físico, onde o trato vocal é representado como uma concatenação de pequenos tubos cilíndricos de tamanhos variados e consequentemente de áreas diferentes. As condições de continuidade da pressão acústica e da velocidade do uido em cada junção de tubos permite o mapeamento das ondas acústicas depressão e de velocidade de entrada com as de saída. Aplicando isso interativamente, obtêm-se a simulação da propagação da onda no trato vocal. Levando em consideração que o tempo de propagação em um duto equivale à uma amostra no sistema, é possível aproximar o tamanho necessário dos tubos cilíndricos de acordo com a frequência da amostra e a velocidade do som.

Existem dois modelos para a representação do trato vocal como um concatenação de tubos cilíndricos. O primeiro, foco deste trabalho, é chamado de modelo de reexão de onda (15) (5). O segundo, caso haja interesse em pesquisa, é chamado de modelo linha de transmissão e pode ser também encontrado no livro de Titze (16).

4.0.3 Modelo de Reflexão

Este modelo tem seus fundamentos na reexão das ondas sonoras no trato vocal, conforme citado no Capítulo 2. Enquanto a glote produz uma onda acústica com muitas frequências, o trato vocal amplica um subconjunto dessas frequências para radiar pela boca. O ltro desses subconjuntos está diretamente ligado às características de ressonância e reexão do trato vocal, e é isto que este modelo busca implementar.

Impedância Acústica em um Tubo

Se a velocidade média de uma partícula for multiplicada pela área transversal do tubo, um uxo médio pode ser obtido. Esse uxo médio é conservado quando um tubo se expande ou contrai. Dessa maneira, podemos denir a impedância acústica como a razão entre a pressão acústica p e o uxo acústico u. Logo temos a eq 4.0.3:

$$Z = \frac{p}{u}$$

(4.2)

Se não há reexões no tubo, a razão entre pressão e uxo adota a seguinte forma:

$$Z = \frac{\rho * c}{u}$$

(4.3)

Onde ρ é a densidade do ar, c é a velocidade do som e A é a área transversal do tubo.Seção de reexão do som, foi denido uma fórmula para a impedância acústica livre como c. A partir disso, é fácil notar que o coeciente de impedância acústica no tubo nada

mais é do que o coeciente de impedância acústica livre dividido pela área transversal do tubo.

Capítulo 5

Synpath

5.0.1 Synpath

O SynPath é um sintetizador computacional, desenvolvido em linguagem Python [13, 20, 28], criado por Lucero (?). Este software é uma extensão do sintetizador concebido por Fraj[19], incorporando um modelo de vibração para as pregas vocais. O seu propósito é aumentar a delidade siológica do sintetizador de Fraj e permitir o controle direto dos sons sintetizados em termos de parametrização da laringe. Para se obter um simples controle sobre o sintetizador e facilitar o seu uso para aplicações práticas, é necessário que a representação das pregas vocais seja simples. Além disso, o modelo das pregas vocais deve garantir variações suaves no uxo gerado na glote. A falta de suavidade gera timbres não naturais e consequentemente a perda da delidade siológica buscada. Sendo assim, o modelo multi-massa para representação das pregas vocais não pode ser utilizado visto que produz variações não suaves e é um modelo matematicamente muito complexo levando a instabilidades numéricas, afetando o uso para aplicações práticas. O SynPath tomou como base para a representação das pregas vocais o modelo de onda mucosa desenvolvido por Titze (16). Basicamente, o modelo é um oscilador mono-massa, conforme descrito no Capitulo 3, incorporando a transferência de energia do uxo de ar para as pregas vocais. Entretanto, o modelo de Titze possui duas restrições, uma que foi solucionada e aplicada no desenvolvimento do SynPath e a outra que ainda não foi solucionada e portanto é também uma restrição do modelo computacional do SynPath. A primeira restrição é que o modelo de Titze foi concebido para o estudo em pequenas oscilações, sendo que para oscilações de maior amplitude, este não é apropriado. Entretanto, esse modelo foi extendido por Lucero (?) para abranger maiores oscilações utilizando um mecanismo limitador de amplitude durante as oscilações (?). Mesmo com essa extensão para maiores amplitudes, o modelo ainda apresentava uma outra restrição, um atraso pequeno para o deslocamento da onda no canal glotal. A consequência desse atraso é que a pressão limite para que ocorra a vibração se torna independente da frequência de vibração (?). Porém, sabe-se que um esforço maior é necessário para que tons mais agudos sejam vocalizados, ou seja, a pressão para que se ocorra vibração em frequências maiores(tons maiores) é maior. Essa restrição ainda não foi solucionada tendo em vista a diculdade de se realizar o supracitado computacionalmente. Mesmo com essa restrição, o modelo de Lucero (?) é o modelo utilizado para a representação do caráter oscilatório das pregas vocais no software SynPath, sendo que esta restrição não solucionada não tem forte inuência no produto nal.

5.0.2 Concepção Matemática e Modelagem do Sistema

O trato vocal no SynPath é representado como uma sequencia de tubos cilíndricos concatenados pelos quais se propaga uma onda acústica planar com componentes anteriores e posteriores (**?**). Tanto o trato vocal subglótico quanto o trato vocal supraglótico estão incluídos na representação, sendo que no caso do trato supraglótico, ainda estão inclusos as cavidades nasais e os sínus paranasais. Neste modelo, as perdas por vibração, viscosidade do ar e condução térmica são consideradas.

A propagação da onda acústica foi resolvida numericamente para um valor de amostra de 88.2 kHz, o que xa o comprimento dos tubos cilíndricos em 0,4 cm. Sendo assim, a traqueia foi modelada com 36 tubos cilíndricos cujas áreas transversais foram xadas em 2,5 cm2.

5.0.3 Pregas Vocais

O modelo de Titze (16) para representação das cordas vocais pode ser visto na Figura 4.1. Este modelo assume completa simetria entre as pregas vocais e somente permite o movimento horizontal das mesmas. As propriedades biomecânicas do tecido são aglomeradas no ponto médio da glote, levando à equação 4.1, de movimento das pregas vocais:

Onde x é o deslocamento do tecido, M,B e K são a massa, o amortecimento e a rigidez respectivamente por unidade de área da superfície das pregas vocais, é o coeciente de amortecimento não linear e Pg é a pressão intraglotal. Essa equação representa o movimento de cada prega vocal dentro do modelo estabelecido.

Requisitos Funcionais

O Synpath é consistido também dos seguintes requisitos funcionais, os requistos funcionais são as funcionalidades que o sistema executará(**?**)

- 1 - Validação do Parâmetros passados pelo Usuário, se condizem com restrições do programa.
- 2 - Plotar um gráco inicial do trato vocal de acordo com os parâmetros do usuário.
- 3 - Plotar três grácos referentes as propriedades da voz simuladas com os parâmetros fornecidos pelo usuário. – O primeiro gráco refere-se às posições adotadas pelas cordas vocais, a área da glótis, ao uxo de ar nessa área e às características desse uxo. – O segundo gráco refere-se às características do som gerado pela simulação física do aparato fonador pelo programa. – O terceiro gráco refere-se ao espectro de frequência do som gerado e do uxo da glótis.
- 4 - Gerar um arquivo de texo com as características de voz gerada, frequencia,amplitude, ruído da voz, entre outros
- 5 Gerar um arquivo de som de voz simulada.

Requisito Não-Funcionais

O Synpath consiste também dos seguintes requisitos não funcionais, requisitos não funcionais são requisitos são parametros de qualidade, requisitos que limitam as funcionalidades do sistema(**?**).

- 1 - O Sistema deve produzir os gráficos que os requisitos funcionais delimitaram em um intervalo de 1(um) minuto.
- 2 - Após gerar os gráficos e os exibi-los o arquivo texto e o arquivo de audio deverão ser exibidos
- 3 - Para que o sistema esteja funcional é necessário ter instalado os pacotes: MatPlotLib e NumPY
- 4 - O sistema deve ser executado em plataformas de um sistema operacionais como Windows, Linux ou MacOs, Versões recentes de acordo com a data desse documento.

5.0.4 HMMs

Minera-se de várias partituras musicais para treino. Os dados minerados dessas músicas são fonemas, altura, intensidade e os intervalos, isto é relação com outras notas.Esses dados são convertidos e mapeados em "labels"dependentes de contexto (3).Após isso as HMM's são treinados através dos dados de treinamento usando o algoritmo EM.(7). Após isso ocorre a fase de sintesis, usa-se outra partitura para ser convertida em "labels"dependentes de contexto e estima-se quais "labels"preprocessadas são correspondentes.(12)

5.0.5 MHRSMM

Uma variação HSMM. Modelo de multipla regressão HSMM. Parametros importantes são i e mi dos outputs pdfs

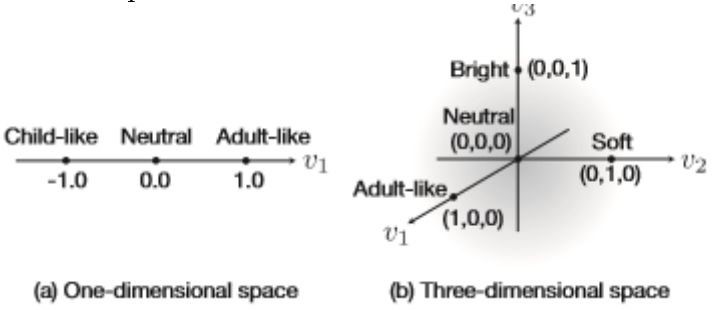
$$\mu_i = H_b i\xi$$

$$m_i = H_p i\xi$$

$$\xi = [1, v_1, v_2, ..., v_L]^T$$

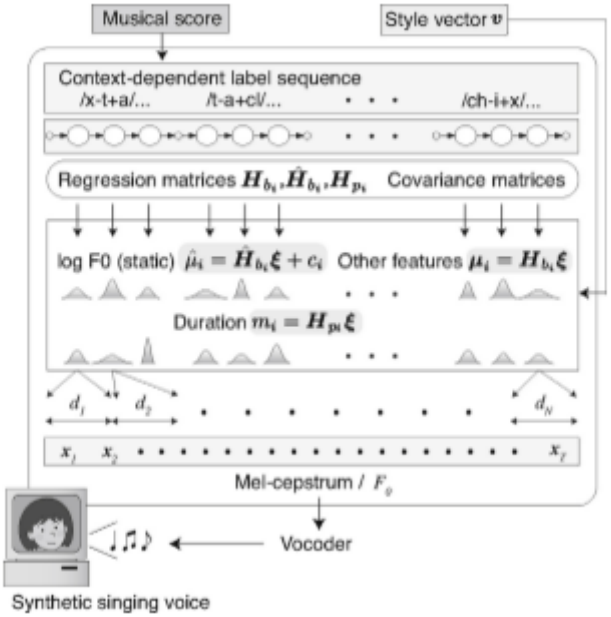
$$\xi = [1, v^T]^T$$

Onde L é a dimensão do vetor de estilo e vi é a intensidade do enésimo estilo de canto. Um exemplo de um vetor de estilos de canto de tamanho L = 2 e L = 3.



(a) One-dimensional space (b) Three-dimensional space

Controle do Sintentisador de voz cantada baseado em MRHSMM



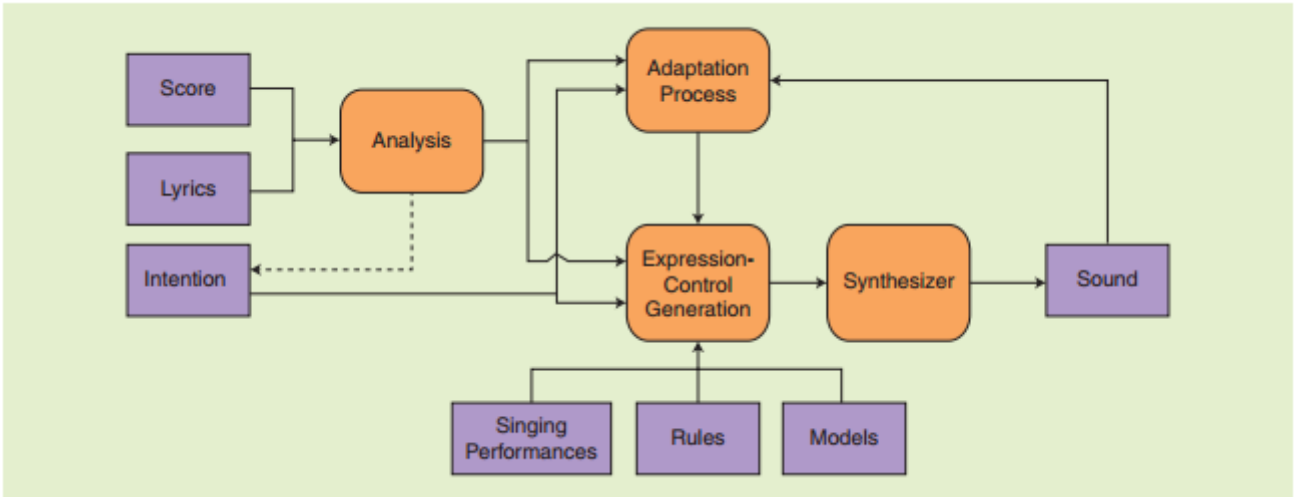
Durante a fase de síntese o usuário do programa adiciona vetores de estilos de acorodo com a intenção e a expressividade pretendida. Parametros de output como duração são gerados pelos vetores de estilos dados e matrizes de regressões treinadas usando MRHSMMs

Resultado de todo esse processo é um sequência HSMM usando parametros de geração de fala

MRHSMM possui uma dificuldade de gerar contorno F0 que acompanhe o contexto de mudança de altura das notas o author TAKASHI NOSE, propõe um treinamento de HSMM e HMM nos parametros

5.0.6 FrameWorks Sintetisador de Voz

Frame Work de um Sistema Sintetisador de Voz:



Input

Consiste da partitura, letra e emoção. o input é analisado e derivado em uma transcrição fonética, alinhamento com a performance alvo ou dados contextuais.(17)

Expressão

Expressão músical é um conceito intuitivo porém dificil de se definir. A expressão é chave na percepção da qualidade e naturalidade musical. No caso da voz cantada implica-se usar vários outros parametro além de frequencia e amplitude. Psicológicamente contorno do timbre, vibrato, tremolo, timing fonético.(17)

5.0.7 Envoltoria F0

Envoltórias F0 são usadas para expressar informação linguistica, para-linguística e não-linguistica.(13)

As Envoltória F0 apresentam três (3) características importantes que fazem diferenciar uma voz falada a uma voz cantada.(14)

- 1 - O alcance dinâmico de uma envoltória F0 é mais largo que o de uma voz falada
- 2 - A envoltória F0 corresponde e tende a se manter estável em uma nota. A mudança de nota de uma envoltória F0 corresponde a melodia da música
- 3 - Existem muita flutuações f0 que são apenas observadas em apenas vozes cantadas

5.0.8 Sintese de Voz em Mandarím

Utiliza-se a técnica HNM para a sintese da voz cantada em mandarím. HNM significa , "harmonic plus noise model". O modelo HNM divide o espectro de um sinal em dois(2) com larguras não iguais para modelagem melhor do espectro.(11)

Referências

[1] John & Charles Bell. The anatomy and physiology of the human body, 1826. [Down, pre-B, ED]. vii, 4

[2] Ferdinand de Saussure. *Self-oscillating source for vocal-tract synthesis.* IEEE Tran. Audio Eletroacoust, Audio Eletroacoust., 1968. 16

[3] Gilles Degottex, Pierre Lanchantin, Axel Roebel, and Xavier Rodet. Mixed source model and its adapted vocal tract filter estimate for voice transformation and synthesis. *Speech Communication*, 55(2):278 – 294, 2013. 24

[4] Byron D. Erath, Matías Zañartu, Kelley C. Stewart, Michael W. Plesniak, David E. Sommer, and Sean D. Peterson. A review of lumped-element models of voiced speech. *Speech Communication*, 55(5):667 – 690, 2013.

[5] Jean Schoentgen Samia Fraj and Francis Grenez. *Development and perceptual assessment of a synthesizer of disordered voices.* Journal of Acoustical Society of America, 2012.

[6] Henry Gray. *Anatomy of the Human Body.* IEEE Tran. Audio Eletroacoust. 16

[7] Jon Gudnason, Mark R.P. Thomas, Daniel P.W. Ellis, and Patrick A. Naylor. Data-driven voice source waveform analysis and synthesis. *Speech Communication*, 54(2):199 – 211, 2012. 24

[8] James L. Flanagan Jont B. Allen Mark A. Hasegawa-Johnson. *Speech Analysis Synthesis and Perception.* 2008.

[9] Matias Zanartu Kelley C. Stewart Michael W. Plesniak David E. Sommer Sean D. Peterson Byron D. Erath. *A review of lumped-element models of voiced speech.* Speech Communication. 16

[10] Jody Kreiman and Diana Van Lancker Sidtis. Foundations of voice studies, 2011. vii, 2, 3, 5, 6

[11] Chyi-Yeu Lin, Li-Chieh Cheng, Chang-Kuo Tseng, Hung-Yan Gu, Kuo-Liang Chung, Chin-Shyurng Fahn, Kai-Jay Lu, and Chih-Cheng Chang. A face robot for autonomous simplified musical notation reading and singing. *Robotics and Autonomous Systems*, 59(11):943 – 953, 2011. 26

[12] Takashi Nose, Misa Kanemoto, Tomoki Koriyama, and Takao Kobayashi. Hmm-based expressive singing voice synthesis with singing style control and robust pitch modeling. *Computer Speech Language*, 34(1):308 – 322, 2015. 24

[13] Takeshi Saitou, Masashi Unoki, and Masato Akagi. Development of an {F0} control model based on {F0} dynamic characteristics for singing-voice synthesis. *Speech Communication*, 46(3–4):405 – 417, 2005. Quantitative Prosody Modelling for Natural Speech Description and GenerationInternational Conference on Speech Prosody. 26

[14] Takeshi Saitou, Masashi Unoki, and Masato Akagi. Development of an {F0} control model based on {F0} dynamic characteristics for singing-voice synthesis. *Speech Communication*, 46(3–4):405 – 417, 2005. Quantitative Prosody Modelling for Natural Speech Description and GenerationInternational Conference on Speech Prosody. 26

[15] BradH Story. *Tubetalker*. Dept. of Speech, Language, and Hearing, Tucson, AZ, 2010. 20

[16] Ingo Titze. *Principles of Voice Production*. Prentice Hall, New Jersey 07632, 1994. vii, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 18, 19, 20, 22, 23

[17] M. Umbert, J. Bonada, M. Goto, T. Nakano, and J. Sundberg. Expression control in singing voice synthesis: Features, approaches, evaluation, and challenges. *IEEE Signal Processing Magazine*, 32(6):55–73, Nov 2015. 26

[18] Janwillem van den Berg. Myoelastic-aerodynamic theory of voice production. *Journal of Speech, Language, and Hearing Research*, 1(3):227–244, 1958.