

Comparação de modelos mecânicos a duas massas para produção da VOZ

Julien Mauprivez, Departamento de Engenharia Mecânica, PUC-Rio,
e-mail: mobmaup@gmail.com

Edson Cataldo, Departamento de Matemática Aplicada, programa de Mestrado em Engenharia de Telecomunicações, UFF, e-mail: ecataldo@im.uff.br, home page: <http://www.professores.uff.br/ecataldo/>

Rubens Sampaio, Departamento de Engenharia Mecânica, PUC-Rio,
e-mail: rsampaio@puc-rio.br

Introdução

A produção de vogais na fala tem como origem a indução da auto-oscilação das cordas vocais pelo fluxo de ar que provém dos pulmões (Figura 1). Os pulsos do fluxo glotal, assim gerados, vão se modificando ao passarem pelo trato vocal (região que vai desde a laringe até a boca e que atua como um filtro) gerando, finalmente, o som que ouvimos.



Figura 1: Ciclo glotal durante a produção de uma vogal, com o fluxo de ar direcionado de baixo para cima.

Existem vários modelos para simular esse comportamento particular das cordas vocais. Porém, os modelos de duas massas, embora sejam muito simplificados, permitem a reprodução de comportamentos típicos (Cataldo et al., 2006, 2008), com complexidade e custo computacional relativamente baixos.

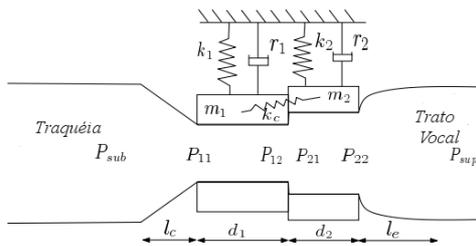


Figura 2: Modelo a duas massas da glote de Ishizaka e Flanagan (1972) com o fluxo circulando da esquerda para direita.

Nesse trabalho, compara-se o modelo de Ishizaka e Flanagan (1972), ilustrado na Figura 2, com o modelo de Lous (1998), ilustrado na Figura 3, que são os modelos de duas massas mais usados na literatura para síntese de voz e descrição de fenômenos relacionados à voz, a partir de modelos físicos.

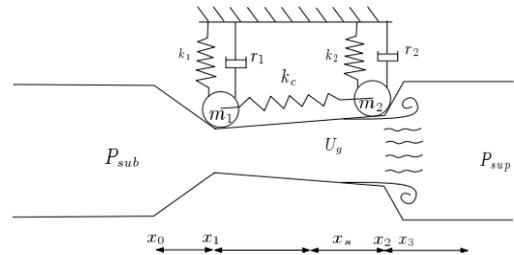


Figura 3: Modelo da glote de Lous et al. (1998)

Modelagem

O modelo de Ishizaka e Flanagan consiste em um modelo mecânico das cordas vocais, não-linear, de duas massas acoplado a um tubo de seção variável representando, por sua vez, o trato vocal. A partir da queda de pressão entre a traquéia e o trato vocal, assim como da geometria, a vazão na glote, suposta constante entre a entrada e a saída, é determinada usando a equação de Bernoulli, considerando viscosidade, inércia do fluido, e pressão acústica devida ao acoplamento com o trato vocal. A principal diferença entre os dois modelos está na geometria. As massas do modelo de Lous são vinculadas devido a três placas, para levar em conta a variação da posição do ponto de descolamento do fluxo em regime turbulento, para o fluxo de ar (Pelorson et al., 1994). As equações das dinâmicas das massas das cordas vocais, para os dois modelos, podem ser sintetizadas pela eq. (1):

$$m_i \frac{d^2 x_i}{dt^2} + r_i \frac{dx_i}{dt} + s(x_i) + k_c(x_i - x_j) = F_i \quad (1)$$

Para $i=1,2$, temos: m_i é a massa correspondente de cada corda vocal, r_i é o amortecimento, $s(x_i)$ é a rigidez, não-linear, x_i e x_j são os deslocamentos das massas ($j=1,2; i \neq j$) e F_i são as forças nas massas. No modelo de Ishizaka e Flanagan a rigidez é dada por:

$$s(x_i + x_{i0} > 0) = k_i(x_i + a_i x_{i0}^3) \quad (2)$$

$$s(x_i + x_{i0} < 0) = s(x_i + x_{i0} > 0) \quad (3)$$

$$+ K_i \left(\left(x_i + \frac{x_{0i}}{2} \right) + b_i \left(x_i + \frac{x_{0i}}{2} \right)^3 \right)$$

onde k_i, K_i, a_i e b_i são as constantes da rigidez e x_{i0} a abertura inicial. No modelo de Lous, as rigidez são dadas pelas constantes $s(x_i + x_{0i} > 0) = k_i$ e $s(x_i + x_{0i} < 0) = 4k_i$.

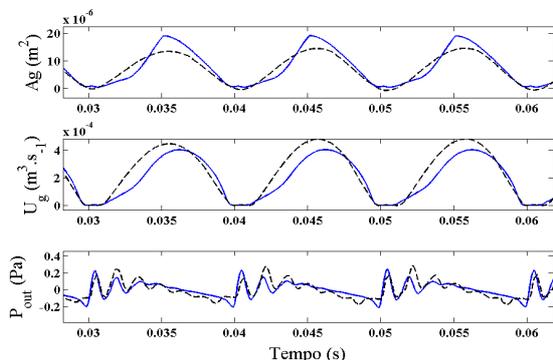


Figura 4: Área glotal (Ag), Vazão da glote (Ug) e pressão acústica (Pout). Linha cheia: modelo de Ishizaka para $x_{01}=x_{02}=0,5$ mm e Linha tracejada: modelo de Lous para $x_{01}=x_{02}=0,2$ mm.

No modelo de Ishizaka e Flanagan, as duas massas têm valores diferentes e a assimetria de cada pulso glotal ocorre devido à diferença entre os valores das massas. No Modelo de Lous, as massas têm o mesmo valor e a assimetria do pulso glotal ocorre devido ao ponto de descolamento que, principalmente durante o fechamento, modifica a distribuição de pressão. A simetria mecânica do modelo de Lous é uma vantagem quando se quer modelar uma prótese de glote (Lous,1998) ou uma maquete de glote *in-vitro* (Ruty, 2007). Porém, seus parâmetros adicionais dificultam seu ajuste, sempre necessário com modelos simplificados. Por exemplo, o fenômeno de histerese, em relação ao valor de pressão subglotal necessária para a fonação, estudado em Lucero(1999) usando um modelo de Ishizaka, não foi reproduzido em Rutu(2007) usando um modelo de Lous.

Os dois modelos foram simulados numericamente, através do método das diferenças finitas e um exemplo dos gráficos obtidos é mostrado na Figura 4. Todos parâmetros usados para simulação, da ordem dos parâmetros reais, são adaptados dos modelos de Lous e Ishizaka. Ajustemos a frequência dos modelos através da razão k_1/m_1 para cada modelo. Para obter resultados similares, um ajuste na área neutra dos modelos foi necessária. Consideramos o mesmo trato vocal para os dois casos, com oito tubos acústicos, modelando a produção da vogal /a/. Os sons obtidos podem ser acessados na página:

www.professores.uff.br/ecataldo/ENEBI2009.htm.

Apesar de os sinais apresentados, no tempo, serem diferentes (Figura 4), o conteúdo freqüencial apresenta formantes da mesma ordem, significando que, em ambos os casos, estamos gerando uma vogal /a/ (Fant, 1960), embora usando modelos diferentes.

Conclusões

Dois modelos de duas massas para a produção da voz são comparados: o modelo de Ishizaka e Flanagan e o modelo de Lous. O modelo de Lous apresenta novidades quanto à geometria e, também, quanto ao modelo para o fluxo de ar. Porém, com qualquer um dos dois modelos apresentados pode-se gerar a mesma vogal, embora algumas diferenças apareçam. Há uma preocupação também com a equivalência dos parâmetros de um e de outro modelo, pois os seus valores não são exatamente iguais. Embora o modelo de Lous apresente-se como mais moderno, há alguns fenômenos que ele não consegue simular, como o da histerese, caso em que o modelo de Ishizaka e Flanagan representa de forma satisfatória e coerente com dados experimentais. Por outro lado, o modelo de Lous apresenta-se mais coerente em relação ao fluxo de ar em regime turbulento, já que discute um ponto de descolamento, o que não é previsto pelo modelo de Ishizaka e Flanagan. Mostramos, então, que os modelos não são substitutos, mas complementares e, dependendo da aplicação, devemos usar um ou outro modelo.

Agradecimentos

O autor agradece o suporte do CNPq e da FAPERJ no desenvolvimento desta pesquisa.

Referências bibliográficas

- Cataldo E., Leta F.R., Lucero J.C., Nicolato L. "Synthesis of voiced sounds using low dimensional models of the vocal cords and time-varying subglottal pressure", *Mechanics Research Communications* 33, 250-260, 2006.
- Cataldo E., Soize C., Sampaio R., Desceliers C., "Probabilistic modeling of a nonlinear dynamical system used for producing voice", *Computational Mechanics* 43, 265-275, 2008.
- Fant G., "Acoustic theory of speech production", Mouton, The Hague, 1960.
- Ishizaka K., Flanagan, J.L., "Synthesis of voiced sounds from a two-mass model of the vocal cords", *Bell System Technical Journal* 51, 1233-1267, 1972.
- Lous N.J.C., Hofmans G.C.J, Veldhuis N.J., Hirschberg A., "A symmetrical two-mass model coupled to a vocal tract and trachea, with application to prothesis design", *Acta Acustica* 84, 1135-1150, 1998.
- Lucero J.C., "A theoretical study of the hysteresis phenomenon at vocal fold oscillation onset-offset", *Journal of the Acoustical Society of America* 105, 423-431, 1999.
- Pelerson X.; Hirschberg A.; Van Hassel R.R.; Wijnands A.P.J.; Auregan Y., "Theoretical and experimental study of quasisteady flow separation within the glottis during phonation", *Journal of the Acoustical Society of America* 96, 3416-3431, 1994.
- Ruty N., Pelerson X., Van Hirtum A., Lopes Arteaga I., Hirschberg A., "An in-vitro setup to test the relevance and the accuracy of low-order vocal folds models", *Journal of the Acoustical Society of America* 121, No. 1, 2007.