A computational method for updating a probabilistic model of an uncertain parameter in a voice production model

E. Cataldo, Christian Soize, R. Sampaio

To cite this version:
E. Cataldo, Christian Soize, R. Sampaio. A computational method for updating a probabilistic model of an uncertain parameter in a voice production model. 1st International Symposium on Uncertainty Quantification and Stochastic Modeling (Uncertainties 2012), Feb 2012, Maresias, SP, Brazil. pp.1-8. hal-00698731

HAL Id: hal-00698731
https://hal-upec-upem.archives-ouvertes.fr/hal-00698731

Submitted on 17 May 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
A computational method for updating a probabilistic model of an uncertain parameter in a voice production model

Edson Cataldo¹, Christian Soize², Rubens Sampaio³

¹ Universidade Federal Fluminense, Applied Mathematics Department and Graduate program in Telecommunications Engineering, Rua Mário Santos Braga, S/N, Centro, Niterói, RJ, CEP: 24020-140, Brazil, ecataldo@im.uff.br
² Université Paris-Est, Laboratoire Modélisation et Simulation Multi Echelle, MSME UMR 8208 CNRS, 5, Bd Descartes, 77454, Marne-La-Vallée, France, christian.soize@univ-paris-est.fr
³ PUC-Rio, Mechanical Engineering Department, Rua Marques de São Vicente, 225, Gavea, Rio de Janeiro, RJ, CEP: 22453-900, Brazil, rsampaio@puc-rio.br

Abstract. The aim of this paper is to use Bayesian statistics to update a probability density function (p.d.f.) related to the tension parameter of the vocal folds, which is one of the main parameters responsible for the changing of the fundamental frequency of a voice signal, generated by a mechanical/mathematical model for producing voiced sounds. Three parameters are considered uncertain in the model used: the tension parameter, the neutral glottal area and the subglottal pressure. Random variables are associated to the uncertain parameters and their corresponding p.d.f.’s are constructed using the Maximum Entropy Principle. The Monte Carlo method is used to generate the voice signals, which are the outputs of the model. For each voice signal, the corresponding fundamental frequency is calculated and a p.d.f. for this random variable is constructed. Experimental values of the fundamental frequency are then used to update the p.d.f. of the fundamental frequency and, consequently, of the tension parameter, through the Bayes’ method.

Keywords. Uncertainties, Bayes, Probabilistic models, Voice production

1 INTRODUCTION

The production of voiced sounds (vowels are particular cases of voiced sounds) starts with the contraction-expansion of the lungs causing an airflow (due to the difference of pressure between the lungs and the mouth), which will induce the auto-oscillation of the vocal folds (located in the larynx). After passing through the glottis and due to the movement of the vocal folds, the airflow is transformed into pulses of air which are generated (quasi)-periodically. The pressure signal created is so called the glottal signal, which will further be filtered and amplified by the vocal tract to generate the sound we hear. The fundamental frequency of the voice signal, which is the frequency of the vocal folds oscillation, is the inverse of the period of the glottal signal. As the glottal signal is not exactly periodic, for each time interval corresponding to a complete cycle of the vocal folds, a different fundamental frequency is associated. So, the voice signals constitute a stochastic process and the fundamental frequency will be a random variable.

Some authors have modeled the vocal folds dynamics, mainly in a deterministic way (Koizumi et al., 1976; Lous et al., 1998; Zhang et al., 2005). One of these models is the well-known model proposed by Ishizaka and Flanagan (1972) and it will be used here because it has provided a simple and effective representation of the system for studying the underlying dynamics of voice production.

2 BRIEF DESCRIPTION OF THE ISHIZAKA AND FLANAGAN MODEL

A diagram of the model is shown in Fig 1.

![Figure 1: Two-mass model of the vocal folds.](image)

The dynamics of the system is given by Eqs. (1) and (2) (Cataldo et al., 2008, 2009):
\[ \psi_1(w)\dot{u}_g + \psi_2(w)|u_g|\dot{u}_g + \psi_3(w)u_g + \frac{1}{c_1} \int_0^\tau (u_g(\tau) - u_1(\tau))d\tau - y = 0 \] (1)

\[ [M]\ddot{w} + [C]w + [K]w + h(w, w, u_g, \dot{u}_g) = 0 \] (2)

where \( w(t) = (x_1(t), x_2(t), u_1(t), u_2(t), u_3(t))^T \), the functions \( x_1 \) and \( x_2 \) are the displacements of the masses, \( u_1 \) and \( u_2 \) describe the air volume flow through the (two) tubes that model the vocal tract and \( u_3 \) is the air volume flow through the mouth. The subglottal pressure is denoted by \( y \) and \( \dot{u}_g \) is the function that represent the glottal pulses signal. The function output radiated pressure \( p_r \) is given by \( p_r(t) = u_r(t)r_r \), in which \( r_r = \frac{128\rho_v}{9\pi^2} \), \( \rho \) is the air density, \( v_r \) is the sound velocity, and \( y_2 \) is the radius of the second tube. The description of the other quantities that appear in the equation and a detailed discussion of the model, including its implementation, can be found in (Cataldo et al., 2009).

The process to generate a voiced sound is complex and its modeling involves a lot of quantities which should be controlled. The interest here is in the changing of the fundamental frequency. The three main parameters responsible for these changings, as discussed in (Cataldo et al., 2008, 2009; Ishizaka and Flanagan, 1972) are described in the following:

- \( a_{g0} \): the area at rest between the vocal folds, called the neutral glottal area.
- \( y \): the subglottal pressure.
- \( q \): the tension parameter which controls the fundamental frequency of the vocal-fold vibrations because vocal fold abduction and tension are the main factors used by a speaker to control phonation. In order to control the fundamental frequency of the vocal folds, parameters \( m_1, k_1, m_2, k_2, k_c \) are written as \( m_1 = m_1/q, k_1 = qk_1, m_2 = m_2/q, k_2 = qk_2, k_c = qk_c, \) in which \( m_1, k_1, m_2, k_2, k_c \) are fixed values.

These three parameters will be considered as uncertain and random variables will be associated to them. It means that for each realization of the three random variables a different voice signal is produced, characterizing that the voice production process generates a stochastic process.

The probability density functions associated to the random variables corresponding to the chosen uncertain parameters will be constructed by using the Maximum Entropy Principle (or better, the Jaynes’s Maximum Entropy Principle) (Jaynes, 1957a, 1957b).

The measure of uncertainty (entropy) used here was proposed by (Shannon, 1948) and it is given by Eq. (3):

\[ S(p_X) = -\int_{-\infty}^{+\infty} p_X(x)\ln(p_X(x))\, dx. \] (3)

in which \( p_X \) is the p.d.f. of the random variable \( X \).

The goal is to maximize the measure \( S(p_X) \), subject to the constraints given by Eq. (4):

\[ \int_{-\infty}^{+\infty} p_X(x)\, dx = 1 \quad \text{and} \quad \int_{-\infty}^{+\infty} p_X(x)g_i(x)\, dx = a_i, \quad i = 1, \ldots, m \] (4)

in which \( a_i \) are usable information related to the functions \( g_i \).

According to the first part of the principle, only probability distributions consistent with the constraints given should be used. However, an infinity of probability distributions compatible with the constraints may exist. The second part of the principle states the way to choose one among the many p.d.f.’s that satisfies the constraints, the (unique) probability distribution that maximizes the entropy.

3 PRIOR PROBABILISTIC MODEL OF THE UNCERTAIN PARAMETERS

The three parameters \( a_{g0}, y, \) and \( q \) are modeled by random variables \( A_{g0}, Y, \) and \( Q, \) respectively. Consequently, parameters \( m_1, k_1, m_2, k_2, \) and \( k_c \) become random variables denoted by \( M_1, K_1, M_2, K_2, \) and \( K_c \) given by \( M_1 = m_1/Q, K_1 = Qk_1, M_2 = m_2/Q, K_2 = Qk_2, \) and \( K_c = Qk_c. \) The probability models derived here are particular cases of those described in (Soize, 2000, 2001). Since no information is available concerning cross statistical moments between random variables \( A_{g0}, Y, Q, \) they will be considered independent. The details about the construction of the p.d.f.’s related to these three random variables can be found in (Cataldo et al., 2009). The expressions of the p.d.f.’s will be described in the following.

The p.d.f. for \( A_{g0} \) is given by Eq. (5):

\[ p_{A_{g0}}(a_{g0}) = I_{[0,+\infty)}e^{-\lambda_0 - \lambda_1 a_{g0} - \lambda_2 (a_{g0})^2}, \] (5)

where \( \lambda_0, \lambda_1 \) and \( \lambda_2 \) are the solution of the three equations defined by Eqs. (6), (7) and (8):
\[
\int_{-\infty}^{+\infty} p_{A_0}(a_{0}) \, da_{0} = 1 ,
\]
\[
\int_{-\infty}^{+\infty} a_{0} \, p_{A_0}(a_{0}) \, da_{0} = \Delta_{A_0} ,
\]
\[
\int_{-\infty}^{+\infty} a_{0}^2 \, p_{A_0}(a_{0}) \, da_{0} = c ,
\]

Since the constant \( c \) is unknown, a new parametrization expressing it as a function of the coefficient of variation \( \delta_{A_0} \) of the random variable \( A_0 \) is given by \( c = \Delta_{A_0}^2 \left( 1 + \delta_{A_0}^2 \right) \).

The p.d.f. for \( Y \) is given by Eq. (9):
\[
p_Y(y) = 1_{|0,+\infty|}(y) \frac{1}{\delta_Y} \left( \frac{1}{\delta_Y} \right)^{\frac{1}{\delta_Y}} \frac{1}{\Gamma\left(\frac{1}{\delta_Y}\right)} \left( \frac{y}{\delta_Y} \right)^{\frac{1}{\delta_Y} - 1} \exp\left( -\frac{y}{\delta_Y} \right),
\]
in which \( \delta_Y = \sigma_Y / Y \) is the coefficient of variation of the random variable \( Y \) such that \( 0 \leq \delta_Y < 1/\sqrt{2} \) and \( \sigma_Y \) is the standard deviation of \( Y \). In this equation \( \alpha \mapsto \Gamma(\alpha) \) is the Gamma function defined by \( \Gamma(\alpha) = \int_{0}^{\infty} t^{\alpha-1} e^{-t} \, dt \).

The p.d.f. for \( Q \) is given by Eq. (10):
\[
p_Q(q) = 1_{|0,+\infty|}(q) \frac{1}{\delta_Q} \left( \frac{1}{\delta_Q} \right)^{\frac{1}{\delta_Q}} \frac{1}{\Gamma\left(\frac{1}{\delta_Q}\right)} \left( \frac{q}{\delta_Q} \right)^{\frac{1}{\delta_Q} - 1} \exp\left( -\frac{q}{\delta_Q^2} \right),
\]
in which the positive parameter \( \delta_Q = \sigma_Q / Q \) is the coefficient of variation of the random variable \( Q \) such that \( \delta_Q < 1/\sqrt{2} \) and \( \sigma_Q \) is the standard deviation of \( Q \).

### 4 STOCHASTIC SYSTEM WITH THE PRIOR PROBABILISTIC MODEL

As explained above, the stochastic system is deduced from the deterministic one substituting \( a_{0} \), \( y \), \( q \) by the random variables \( A_0 \), \( Y \), \( Q \). Consequently, the random variable associated to the fundamental frequency \( F_0 \) is given by \( F_0 = \mathcal{M}(A_0, Y, Q) \). However, the nonlinear mapping \( \mathcal{M} \) is not explicitly known and it is implicitly defined by Eqs. (1) and (2) substituting \( a_{0}, y, q \) by random variables \( A_0, Y, Q \). The fundamental frequency associated to each realization of the voice signal is calculated through the glottal signal, calculating the inverse of its period. In order to validate the development presented here, voice signals produced by one person have been analyzed and their statistics have been compared with simulations. A voice signal corresponding to a sustained vowel /a/ has been recorded from one person and 1,800 frames were obtained from this signal, each one with 0.01s of length. For each frame, the corresponding fundamental frequency was calculated. So, a corresponding p.d.f., the so-called experimental, can be constructed. Figure 2 shows the p.d.f. of the fundamental frequency constructed from the experimental data.

Figure 2: Probability density function corresponding to experimental fundamental frequency

The problem to be solved to update the p.d.f. of \( Q \), using Bayesian statistics, will be divided into two parts: at first, an inverse problem will be solved in order to obtain, from simulations, a p.d.f. of the fundamental frequency near to the experimental one. Then, at the second part, the p.d.f. obtained in the first part will be updated, using experimental data and the Bayesian method. Consequently, the updated p.d.f. of \( Q \) will be obtained.
4.1 Description of the first part

The idea is to identify the mean values $A_{g0}$, $\gamma$, $Q$, and also the coefficients of dispersion $\delta_{A_{g0}}$, $\delta_\gamma$, $\delta_Q$ such that the experimental mean value of the fundamental frequency $m_{F0} = 117.0580$ Hz and also the experimental coefficient of dispersion of the fundamental frequency $\delta_{F0} = \frac{\sigma_{F0}}{m_{F0}} = 0.0084$ can be achieved.

**Step 1**: Values of $a_{g0}$, $\gamma$, and $q$ are chosen, in the corresponding deterministic model, such that an output radiated pressure signal with fundamental frequency $f_0 = 117.0580$ Hz is obtained.

**Step 2**: The values of $a_{g0}$, $\gamma$, and $q$ found in Step 1 are used as the mean values $A_{g0}$, $\gamma$, and $Q$ in the corresponding stochastic problem.

**Step 3**: With the mean values described in Step 2, values of $\delta_{A_{g0}}$, $\delta_\gamma$, and $\delta_Q$ are chosen such that the value of $\delta_{F0} = \frac{\sigma_{F0}}{m_{F0}} = 0.0084$.

Clearly, in order to identify the parameters as described, many tests were performed. If the number of cases is large, a strategy to solve this inverse problem can be to create an adequate cost function. The values obtained in each step were:

**Step 1**: $a_{g0} = 5 \times 10^{-2}$ m$^2$, $\gamma = 750$ Pa, and $q = 0.63$.

**Step 2**: $A_{g0} = 5 \times 10^{-2}$ m$^2$, $\gamma = 750$ Pa, and $Q = 0.63$ which will be used in the corresponding stochastic problem.

**Step 3**: With the mean values described in Step 2, the mean value of the fundamental frequency obtained, considering 700 realizations and using the Monte Carlo method, was $m_{F0} = 117.1603$ Hz. With the values of the coefficients of dispersion $\delta_{A_{g0}} = 0.03$, $\delta_\gamma = 0.03$, $\delta_Q = 0.006$, the value obtained for the coefficient of dispersion of the fundamental frequency was $\delta_{F0} = 0.0077$.

Figure 3 shows the probability density function constructed from experimental data and the probability density function constructed from simulations. The function *ksdensity* from MATLAB was used.

4.2 Description of the second part

Let $f_0^{\exp} (\theta_1), \ldots, f_0^{\exp} (\theta_{\nu^{\exp}})$ be the $\nu^{\exp}$ realizations of the random variable $F_0^{\exp}$, which correspond to values of the fundamental frequency obtained experimentally (here, $\nu^{\exp} = 1,800$).

The posterior probability density function $p_Q^{\text{post}}$, related to the random variable $Q$, can be calculated using Bayesian Statistics by Eq. 11:

$$p_Q^{\text{post}}(q) = L^{\text{bayes}}(q) \cdot p_Q^{\text{prior}}(q)$$  \hspace{1cm} (11)

in which $p_Q^{\text{prior}}$ is the prior p.d.f., given by Eq. (10), and $L^{\text{bayes}}$ is the likelihood function given by Eq. (12):

$$L^{\text{bayes}}(q) = \frac{\prod_{\ell=1}^{\nu^{\exp}} p_{F_0|Q}(f^{\exp}_{\ell}|q)}{E_Q \left\{ \prod_{\ell=1}^{\nu^{\exp}} p_{F_0|Q}(f^{\exp}_{\ell}|Q^{\text{prior}}) \right\}}.$$  \hspace{1cm} (12)
The posterior p.d.f. \( p_{F_0}^{\text{post}} \), related to the fundamental frequency, is given by Eq. (13):

\[
p_{F_0}^{\text{post}}(f_0) = \int_{\mathbb{R}} p_{F_0|Q}(f_0|q) p_{Q}^{\text{post}}(q) dq.
\]  

(13)

in which \( p_{F_0|Q} \) is the conditional p.d.f. of \( F_0 \), given \( Q = q \).

Using Eq. (11), Eq. (13) can be rewritten as Eq. (14) (Soize, 2010b):

\[
p_{F_0}^{\text{post}}(f_0) = E_{\text{bayes}}\{ L_{\text{bayes}}(Q_{\text{prior}}) p_{F_0|Q}(f_0|Q_{\text{prior}}) \}.
\]  

(14)

4.3 Computational aspects

Considering a number (\( \nu \)) sufficiently large of the realizations of the random variable \( Q_{\text{prior}} \), for each frequency \( f_0 \), the value or \( p_{F_0}^{\text{post}}(f_0) \) can be estimated by Eq. (15) (Soize, 2010b):

\[
p_{F_0}^{\text{post}}(f_0) \approx \frac{1}{\nu} \sum_{\ell=1}^{\nu} L_{\text{bayes}}(Q_{\text{prior}}(\theta_\ell)) p_{F_0|Q}(f_0|Q_{\text{prior}}(\theta_\ell)).
\]  

(15)

For each realization \( Q_{\text{prior}}(\theta_\ell) \), the conditional probability density function \( p_{F_0|Q} \) is constructed and the corresponding value \( p_{F_0|Q}(f_0|Q_{\text{prior}}(\theta_\ell)) \) is then calculated.

4.3.1 Construction of \( p_{F_0|Q} \)

To construct the function \( p_{F_0|Q} \), 100 deterministic values of \( q \) (from 0.6153 up to 0.6442) were considered and, for each value of \( q \), a p.d.f. of the fundamental frequency was obtained by simulation. The corresponding conditional p.d.f.'s are shown in fig. 4.

4.3.2 Evaluation of \( L_{\text{bayes}} \)

For a specific value of \( q \), \( L_{\text{bayes}}(q) \) is calculated by Eq. 12, with the estimation given by Eq. 16:

\[
E_Q \left\{ \prod_{\ell=1}^{\nu_{\text{exp}}} p_{F_0|Q}(f_0^{\text{exp}},\ell|Q_{\text{prior}}(\theta_\ell)) \right\} = \frac{1}{\nu} \sum_{n=1}^{\nu} \prod_{\ell=1}^{\nu_{\text{exp}}} p_{F_0|Q}(f_0^{\text{exp}},\ell|Q_{\text{prior}}(\theta_n)).
\]  

(16)

Figure 4: Probability density functions \( p_{F_0|Q} \) (100 plots) and the probability density function of the fundamental frequency obtained from experimental data (thick line).

So, corresponding values of \( p_{F_0|Q}(f_0|q) \) can now be calculated, for giving values of \( f_0 \) and \( q \). Here, values of \( f_0 \) were considered from 109 Hz up to 125 Hz, with 0.1 Hz of spacing.
5 RESULTS

Let \( F_0^{\text{exp}} \) be the random variable associated to the fundamental frequencies obtained experimentally and \( F_0^{\text{sim}} \) the random variable associated to the simulated fundamental frequencies. The aim is to update the probability density function of \( F_0^{\text{sim}} \), using experimental values of the fundamental frequency, applying the Eq. 15. Fig. 5 shows the updated p.d.f. \( (p_{F_0}^{\text{upd}}) \) of the fundamental frequency considering \( \nu_{\text{exp}} = 1, 10, 100, 800 \) and 1,800.

![Figure 5: Updated p.d.f.'s of the fundamental frequency for different values of \( \nu_{\text{exp}} \).](image)

Starting from \( \nu_{\text{exp}} = 1000 \), the p.d.f. does not change anymore. It means that the same p.d.f. is obtained considering \( \nu = 1,000 \) or more. It should be observed that the p.d.f. obtained for \( \nu_{\text{exp}} = 1,800 \) is almost the same of the p.d.f. constructed with experimental values.

Let \( p_{F_0}^{\text{exp}}, p_{F_0}^{\text{sim}} \) and \( p_{F_0}^{\text{upd}} \) be the p.d.f.'s related to the fundamental frequencies obtained experimentally, related to the fundamental frequencies simulated and updated, respectively. In order to compare the three p.d.f.'s, Fig. 6 shows the plots of the functions \( |p_{F_0}^{\text{exp}} - p_{F_0}^{\text{sim}}| \) and \( |p_{F_0}^{\text{exp}} - p_{F_0}^{\text{upd}}| \).

![Figure 6: Functions \( |p_{F_0}^{\text{exp}} - p_{F_0}^{\text{sim}}| \) (dashed line) and \( |p_{F_0}^{\text{exp}} - p_{F_0}^{\text{upd}}| \) (continuous line).](image)

Calculating the area under the plots of the Fig. 6, the values found were:

\[
\int_{-\infty}^{+\infty} |p_{F_0}^{\text{exp}}(f_0) - p_{F_0}^{\text{sim}}(f_0)| \, df_0 = 0.0258 \quad \text{and} \quad \int_{-\infty}^{+\infty} |p_{F_0}^{\text{exp}}(f_0) - p_{F_0}^{\text{upd}}(f_0)| \, df_0 = 0.0108.
\]

Although the difference between the values is not so big, it can be noted that the \( p_{F_0}^{\text{upd}} \) is nearer to the \( p_{F_0}^{\text{exp}} \) than \( p_{F_0}^{\text{sim}} \).
Now, it is possible to obtain the p.d.f. of $Q$, using Eq. (11). Figure 7 shows the plots of the prior probability density function of $Q$ and the posterior probability density function of $Q$, obtained with $v_{\text{exp}} = 1, 100, \text{and} 1,800$.

![Figure 7: Prior p.d.f. of $Q$ and the corresponding updated p.d.f.'s.](image)

The p.d.f. of $Q$ is near a delta function located in $Q = 0.633$. With this value of the parameter $q$, and considering the random variables $A_g$ and $Y$, it is possible, using the model, to obtain the p.d.f. of the fundamental frequency constructed with experimental values.

6 CONCLUSIONS

Using Bayesian statistics, the p.d.f. of the random variable $Q$ related to an important parameter which takes part in a mathematical model for producing voice, was updated after obtaining new experimental data. It should be observed that the first prior p.d.f. for $Q$ was obtained using the Maximum Entropy Principle, and there is difficulty to obtain real values for this parameter, because it is related to a biological quantity. Using Bayes Theorem, the p.d.f. of the tension parameter was updated, without getting values directly for this parameter, but from other observable quantity (the fundamental frequency). From the posterior p.d.f. obtained for $Q$, it was possible to simulate voice signals and to construct a p.d.f. for the fundamental frequency which is also the same of the p.d.f. constructed with experimental values.

7 ACKNOWLEDGEMENTS

This work was supported by Fundação de Amparo à Pesquisa no Rio de Janeiro (FAPERJ - programa Jovens Cientista do Nosso Estado), by CAPES (CAPES/COFECUB project N. 672/10) and by the Brazilian Agency Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

8 REFERENCES


RESPONSIBILITY NOTICE

The author(s) is (are) the only responsible for the printed material included in this paper.