INTERSPEECH 2020
October 25–29, 2020, Shanghai, China

Improved Model for Vocal Folds with a Polyp with Potential Application

Jônatas Santos[1], *Jugurta Montalvão*[1], *Israel Santos*[1]

[1]Electrical Engineering Department, Federal University of Sergipe, São Cristóvão, Brazil

jonatascrs@academico.ufs.br, jmontalvao@ufs.com, israelfilhocpd@gmail.com

Abstract

A new model for vocal folds with a polyp is proposed, based on a mass-spring-damper system and body-cover structure. The model was used to synthesize a wide variety of sustained vowels samples, with and without vocal polyps. Analytical conjectures regarding the effect of a polyp on synthesized voice signals corresponding to sustained vowels were performed. These conjectures are then used to estimate intrinsic dimension and differential entropy. These parameters were used to implement a naive classifier with the samples of the public *Saarbruecken* Voice Database, as a proof of concept. The results obtained suggests that the model presented in this paper might be a useful tool for tuning actual polyp detectors.

Index Terms: models of speech production, vocal fold polyp, intrinsic dimensional analysis

1. Introduction

Polyps are one of the most common lesions on vocal folds, being unilateral in most cases [1]. Its origin is related to trauma, caused by high mechanical stress related to abuse, misuse, and overuse of voice [2] or by voice-unrelated violent physical activities with intense respiratory behavior [3]. The presence of a polyp affects folds vibration, potentially causing incomplete closure during phonation and leading to irregular oscillation patterns on both affected and unaffected folds [4]. The vocal polyp voice might be perceived as rough and, in some cases, breathy [3], due to motion aperiodicity on the glottis and occasional air leakage, respectively [1].

Mathematical models for the vocal folds vibration, usually based on mass-spring-damper systems and fluid mechanics, are powerful tools to analyze specific behaviors of vocal folds vibration in different conditions. Relying on the myoelastic-aerodynamic theory of phonation [5], several mathematical models were proposed to represent glottal vibration, applied to normal and pathological phonation [6]. Since the diagnosis of vocal polyps requires auditory experience and invasive procedures to the patient [2], the application of mathematical models might yield useful complementary tools for a specialist to analyze and identify characteristics of vocal polyp on voice [7, 8]. A low-dimensional non-linear model of vocal folds with a unilateral polyp was proposed in [9], adding an extra mass to a previous simplified two-mass model [10]. It is noteworthy that more refined finite element models were proposed in [11] and [12], and that previous modal analysis in a finite element model suggests that low-dimensional models may represent satisfactorily the oscillation patterns of vocal folds [13].

The intrinsic dimension (ID) analysis of data is a growing field in pattern recognition and has become increasingly useful due to the dimensionality reduction of raw and processed data. For instance, the ID is related to the minimum number of latent variables (free parameters). For instance, the method presented in this work uses the broadest definition that the ID in question exists only locally, and, in our analysis [14], the intrinsic dimension and differential entropy are jointly estimated, under the assumption that, projecting signals in low-dimensional and high entropy spaces, tasks like classification, regression, clustering, and data visualization are better performed [15, 16, 17, 18, 19]. This paper presents a modification in a model of vocal folds with a polyp, based on the body-cover structure [20], and makes use of intrinsic dimension and entropy analysis to exemplify some alternatives in automatic detection of vocal polyp.

2. Mathematical body-cover model of vocal folds with a polyp

The proposed model adapts the structure presented in [9] into the three-mass body-cover model of vocal folds described in [21]. For each vocal fold, the model uses one mass for the body, two mass for the cover, coupled by a linear spring, and, on the fold affected by the lesion (the right fold was arbitrarily chosen), one mass for the polyp, as presented in Figure 1.

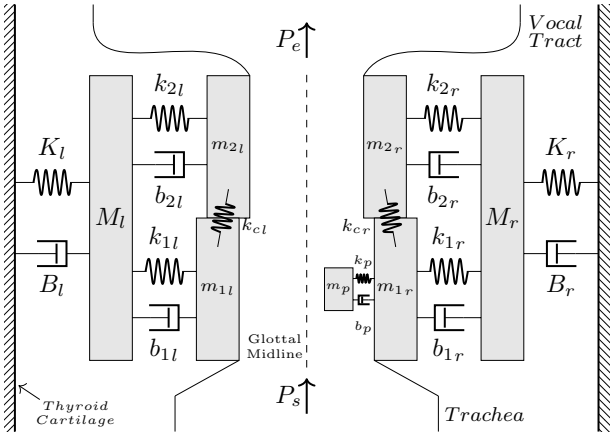

Figure 1: *Body-Cover Model of Vocal Folds with a Polyp.*

2.1. Motion equations

The equations of motion for the three-masses of each vocal fold α ($\alpha = r,\ l$) and the polyp mass are written as:

$$m_{2\alpha}\ddot{x}_{2\alpha} = F_{k_{2\alpha}} + F_{b_{2\alpha}} - F_{k_{c\alpha}} + F_2 + F_{Col_{2\alpha}} \quad (1)$$

$$m_{1r}\ddot{x}_{1r} = F_{k_{1r}} + F_{b_{1r}} + F_{k_{cr}} + F_1 + F_{Col_{1r}} - F_{k_p} - F_{b_p} \quad (2)$$

$$m_{1l}\ddot{x}_{1l} = F_{k_{1l}} + F_{b_{1l}} + F_{k_{cl}} + F_1 + F_{Col_{1l}} \quad (3)$$

$$m_p\ddot{x}_p = F_{k_p} + F_{b_p} + F_{Col_p} \quad (4)$$

$$M_\alpha\ddot{X}_\alpha = F_{K_\alpha} + F_{B_\alpha} - F_{k_{1\alpha}} + F_{d_{1\alpha}} + F_{k_{2\alpha}} + F_{d_{2\alpha}} \quad (5)$$

where:

http://dx.doi.org/10.21437/Interspeech.2020-3049

- $x_{i\alpha}$ ($i = 1,\ 2$) are the displacements of the cover masses, X_α is the displacement of each body mass and x_p is the displacement of the polyp mass;
- $F_{k_{i\alpha}}$, F_{k_p} and F_{K_α} are the spring forces acting on the masses and $F_{k_{c\alpha}}$ is the coupling force between cover masses, given by:

$$\begin{aligned} F_{k_{i\alpha}} = -k_{i\alpha}\big(&((x_{i\alpha} - x_{i\alpha_0}) - (X_\alpha - X_{\alpha_0})) \\ &+ \eta_{i\alpha}\left((x_{i\alpha} - x_{i\alpha_0}) - (X_\alpha - X_{\alpha_0})\right)^3\big) \end{aligned} \quad (6)$$

$$\begin{aligned} F_{k_p} = -k_p\big(&((x_p - x_{p_0}) - (x_{1r} - x_{1r_0})) \\ &+ \eta_p\left((x_p - x_{p_0}) - (x_{1r} - x_{1r_0})\right)^3\big) \end{aligned} \quad (7)$$

$$F_{K_\alpha} = -K_\alpha\left((X_\alpha - X_{\alpha_0}) + H\left(X_\alpha - X_{\alpha_0}\right)^3\right) \quad (8)$$

$$F_{k_{c\alpha}} = -k_p\left((x_{1\alpha} - x_{1\alpha_0}) - (x_{2\alpha} - x_{2\alpha_0})\right) \quad (9)$$

where $x_{i\alpha_0}$, x_{p_0} e X_{α_0} are the initial position of the masses, $k_{i\alpha}$, k_p, K_α and $k_{c\alpha}$ are the spring constants and $\eta_{i\alpha}$, η_p and H are the nonlinear coefficients of the springs;

- $F_{b_{i\alpha}}$, F_{b_p} e F_{B_α} are the damping forces acting on each mass, given by:

$$F_{b_{i\alpha}} = -b_{i\alpha}\left(\dot{x}_{i\alpha} - \dot{X}_\alpha\right) \quad (10)$$

$$F_{b_p} = -b_p\left(\dot{x}_p - \dot{x}_1\right) \quad (11)$$

$$F_{B_\alpha} = -B_\alpha\dot{X}_\alpha \quad (12)$$

- $F_{Col_{i\alpha}}$ and F_{Col_p} are the forces on the masses due to collision with other masses, given by:

$$F_{Col_{i\alpha}} = -\Theta\left(-a_i\right)k_{Col_{i\alpha}}\left(x_{i\alpha} + \eta_{Col_{i\alpha}}x_{i\alpha}^3\right) \quad (13)$$

$$F_{Col_p} = -\Theta\left(-a_p\right)k_{Col_p}\left(x_p + \eta_{Col_p}x_p^3\right) \quad (14)$$

where Θ is the collision function, proposed in [10], a_i and a_p are the areas between masses of each fold and between the polyp and the lower mass of left fold, respectively, a_{i_0} and a_{p_0} are the initial areas, $k_{Col_{i\alpha}}$ and k_{Col_p} are the collision stiffness constants and $\eta_{Col_{i\alpha}}$ and η_{Col_p} are collision nonlinear coefficients;

- F_i are the aerodynamic forces acting on cover masses, due to air pressure at the glottis.

The areas between masses of both sides of vocal folds are defined as follows:

$$a_i = l\left(x_{ir} + x_{il}\right) \quad (15)$$

$$a_p = l_p\left(x_p + x_{1l}\right) \quad (16)$$

where l is the length of the glottis and l_p is the length of the polyp.

2.2. Glottal aerodynamics

The external forces acting on vocal fold, due to air pressure, are obtained by adapting the equations presented in [22] for the three-masses model. For an open glottis condition, i.e., $a_1 > 0$ and $a_2 > 0$, the pressures within the glottis are computed as follows:

$$P_1 = \begin{cases} P_s - \frac{a_d}{a_1}P_{kd}, & a_2 > a_1 \\ P_s - \left(\frac{a_2}{a_1}\right)^2 P_{kd}, & a_2 \leq a_1 \end{cases} \quad (17)$$

$$P_2 = \begin{cases} P_s - P_{kd}, & a_2 > a_1 \\ P_s - \frac{a_2}{a_1}P_{kd}, & a_2 \leq a_1 \end{cases} \quad (18)$$

where P_s is the subglottal pressure, a_d and P_{kd} are, respectively, the area and the kinetic pressure at flow detachment point, both defined as proposed in [22].

For the glottal closure, pressures acting on the vocal folds are obtained as follows:

$$P_1 = \begin{cases} P_s, & a_2 \leq 0 \text{ and } a_1 > 0 \\ P_h, & a_1 \leq 0 \end{cases} \quad (19)$$

$$P_2 = \begin{cases} P_e, & a_1 \leq 0 \text{ and } a_2 > 0 \\ P_h, & a_2 \leq 0 \end{cases} \quad (20)$$

where P_h is the hydrostatic pressure, defined as the mean between P_s and P_e, the pressure at epilarynx tube, immediately above the glottis, dependent of vocal tract coupling.

The aerodynamic forces that act on the folds are given by:

$$F_i = lT_iP_i \quad (21)$$

where T_i are the thickness of each cover mass.

2.3. Glottal flow and vocal tract coupling

The glottal flow is computed according to the Titze's equation [23]. Turbulent flow on the glottis are modelled using the Reynolds number as proposed in [24].

The vocal tract and trachea are coupled to the vocal folds using a wave-reflection analog model for the acoustics [25]. The model is implemented as proposed in [26], using the formulation for a lossless tube and applying an attenuation factor to represent losses. The areas of trachea and vocal tract were obtained and discretized as in [27] and are available with the open source software *LeTalker 1.22* [28]. For this study, only vocal tract shapes for vowels /a/, /i/ and /u/ were used.

2.4. Muscular rules to set model parameters

The rules proposed in [29] to relate muscular activities and model parameters were used. Stiffness, mass, damping, thickness and adduction ($x_{i\alpha_0}$) of each mass and the length of the folds were defined by normalized activation levels (ranging from 0.0 to 1.0) of cricothyroid (a_{ct}), thyroarytenoid (a_{ta}) and lateral cricoarytenoid (a_{lc}) muscles.

For this study, samples were generated using a_{ct} and a_{ta} varying from 0.05 to 0.85 with steps of 0.1 and a_{lc} varying from 0.425 to 0.575 with steps of 0.25. The nonlinear coefficients assume the values proposed in [21, 30], k_p is fixed at $140\ kg/s^2$ and other polyp parameters were defined as proposed in [9]. The resulting model was used to generate normal (removing the polyp mass) and pathological samples. The samples from simulations with parameters that didn't produced sustained oscillation were discarded.

3. Vocal polyp detection from local intrinsic dimension and differential entropy

As an illustration of a practical utility of the model developed in Section 2, we implemented the model in a computer program, and we simulated more than 800 male and female speech signals[1], with and without polyp, in a wide range of fundamental frequency, according to the muscular rules [29].

Then we proceeded a careful study of these synthetic signals, from many different temporal and spectral points of view.

[1]These samples are available at biochaves.com/en/polyp-samples

Eventually, we noticed that a polyp induces irregular (random) amplitude fluctuations through time, as illustrated in Fig. 2, which prompted us to work out a polyp detection approach based on signal entropy and local intrinsic dimension, as explained in this Section.

Let $s(n)$, $n = 0, 1, 2, \ldots, N$, with $s \in \mathbb{R}, N \in \mathbb{N}$ represent N sound signal samples acquired at a constant f_s sampling rate of samples per second. For a proper approach of speech signal modelling, we assume that $\{s(0), s(1), \ldots, s(N)\}$ is an instance of a stochastic process, where the presence of a polyp is expected to disturb the stochastic model. Indeed, as illustrated in Fig. 2, it is clear that after the speech onset, a random amplitude fluctuation is caused by a simulated polyp in the vocal fold.

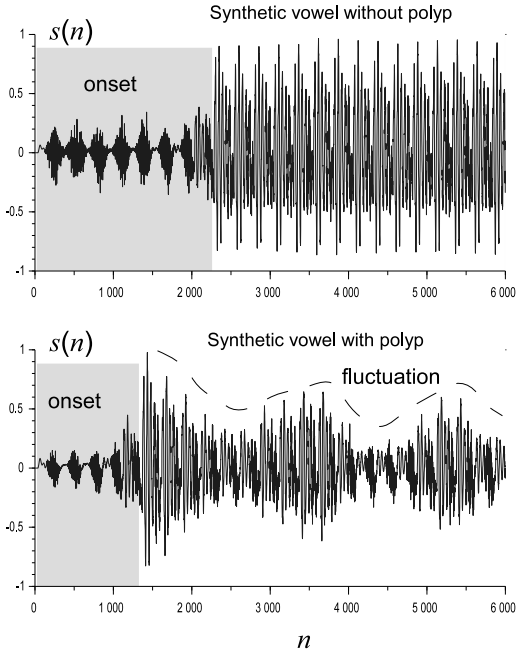

Figure 2: *Signals representing instances of $s(n)$ from synthetic vowels without (up) and with (bottom) polyp. Speech onset interval are indicated, as well as the random amplitude fluctuation caused by a simulated polyp in the vocal fold.*

For the purpose of this work, we assume that the model perturbation imposed by a polyp is enough to impress a detectable change in the produced speech, even when the polyp is just a tiny perturbation in the vocal fold structure. This assumption is inspired by the skill of some trained speech therapists that are able to detect abnormalities, as the ones induced by polyps, in patients through the listening of their voices.

Using the simulator described in Section 2, through experiments with signals such as those shown in Fig. 2, we were able to pinpoint a statistical difference in synthetic voices with and without polyps, which is a consequence of the random amplitude fluctuation shown in that figure. More specifically, it is known that a perfectly periodic signal produces a phase-space trajectory corresponding one-dimensional closed loop, whereas random noises of fluctuations are expected to deform this trajectory. Figure 3 illustrates two 2D projections of such trajectories for two synthetic sustained vowels. For a proper illustration, we set a time lag of $\tau = 0.001$ s for both signals, therefore points correspond to coordinate pairs $(s(n), s(n + \tau f_s))$, in a signal segment of 60 ms (i.e. a signal chunk of 3000 samples, at $f_s = 44.1KHz$).

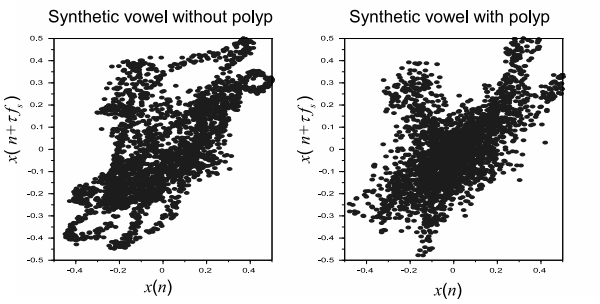

Figure 3: *Dots represent coordinate pairs $(s(n), s(n + \tau f_s))$ from synthetic vowels without (left) and with (right) polyp.*

For a better comparison, both sets of points were enclosed in a square of unit edge. It is noteworthy that it corresponds to a signal normalization where all sample values fall inside the interval [-0.5, 0.5]. This normalization is also applied to all signals considered in this work.

Even from the restrained perspective of those 2D projections, as in Figure 3, where signal trajectories are too convoluted to a proper planar visualization, it is clear that trajectories associated to vocal folds with polyps are less detached from each other (subfigure on the right) than trajectories from speech without polyps (subfigure on the left). Therefore, two assumptions form the basis of the method proposed in this work to polyp detection from speech signals, namely:

(a) After onset, sustained clean vowels are almost periodic, therefore they form trajectories in the phase space that are locally 1D, in other words, they produce (convoluted) funicular loops.

(b) By contrast, sustained vowels where vocal fold dynamic is disturbed by the presence of a polyp produce trajectories less well defined in the phase space, possibly collapsing on each other.

A relevant consequence of (b) is that one should no longer expect locally 1D trajectories, for collapsing trajectories are likely to locally occupy more than a single dimension of the space. To take advantage of it, we adapted the method proposed in [14], which is able to jointly estimate the local dimension and effective volume (through the differential entropy) of a given instance of a stochastic signal, such as a speech segment. In this adaptation, 6D patterns (vectors) are obtained from consecutive signal samples as:

$$\mathbf{p}(n) = [s(n), s(n+1), \ldots, s(n+5)],\ n = 0, 1, 2, \ldots, N-5,$$

but only patterns with signal samples after onset are retained for further analyses.

For a set of P patterns, where P corresponds to about 50 ms of signal (e.g. $P = 2500$ patterns for $f_s = 44100$ samples per second), from a given speech sample, we proceed the joint analysis as explained in [14], where two statistical measurements in logarithmic scale are obtained through straightforward coincidence counting. The two measurements in logarithmic scale are $log_2(r)$ and $log_2(C(r))$, where r corresponds to the edge length of a cube (or a hypercube) inside which two patterns are said coincident, and $C(r)$ stands for the coincidence rate for this definition of coincidence.

For the purpose of this manuscript, let $u(r) = log_2(r)$ and $v(r) = log_2(C(r))$ be just two functions of the edge length r, therefore local intrinsic dimension can be estimated as $\hat{d}(r) = \frac{dv(r)}{du(r)}$. Consequently, the differential entropy, for a given analysis scale r, can be estimated as $\hat{h}(r) = \hat{d}(r) - v(r)$. In Figure 4, both estimates are illustrated for two synthetic signals.

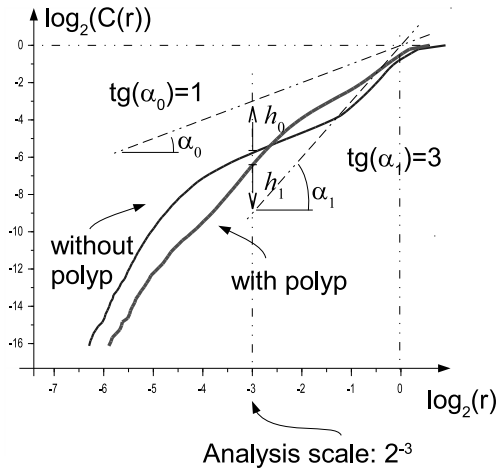

Figure 4: *Two curves from which joint estimates of intrinsic dimensions (curves slopes) and differential entropies (distances from corresponding constant slope lines) can be obtained for a range of scale analyses. For instance, at a scale analysis of 2^{-3}, intrinsic dimensions and differential entropies for signals without and with vocal fold polyp are estimated as tangents of angles α_0 and α_1, and distances h_0 and h_1, respectively.*

Thanks to a laborious study of curves similar to those presented in Figure 4, for all synthetic sustained vowel, we observed that at a scale corresponding to $\log_2(r) \approx -3$, therefore for $r \approx 2^{-3}$, a difference in terms of intrinsic dimension and differential entropy consistently appears between signals from vocal folds with and without polyps. As explained before, this difference was analytically expected, and we conjecture that it can be attributed to the random amplitude fluctuation, which induces an increase of local intrinsic dimension (at small scale of analysis r), and forces a decrease of differential entropy (effective volume, as explained in [14]) for that higher dimension. Both effects are illustrated in Figure 4.

To briefly test this new approach (although it plays more the role of a concept proof, in this work), we used the *Saarbruecken* Voice Database [31]. Accordingly, in Fig. 5 points correspond to real speech samples from vocal folds with (dots) and without (crosses) polyp. Each point coordinate is given by the estimated differential entropy and intrinsic dimension for an analysis scale of $r = 2^{-3}$. Only sustained vowel /a/ in *neutral* tone were used in this experiment. Moreover, only samples from $s(3000)$ on were analysed, in agreement with the observation (with synthetic signal) that signal segments just after transitory emission interval are carriers of more discriminative features.

Finally, a very simple classifier was implemented according to the following rules:

- If the differential entropy of a given signal is greater than T_h and its intrinsic dimension is lower than T_d, then the speech segment is classified as normal;

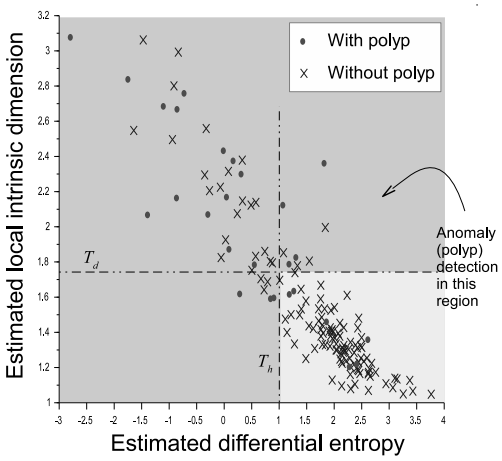

Figure 5: *Illustration of points corresponding to real speech samples from vocal folds with (dots) and without (crosses) polyp. Each point coordinate is given by the estimated differential entropy and intrinsic dimension for an analysis scale of 2^{-3}. Two thresholds for entropy and dimension, $T_h = 1$ and $T_d = 1.75$, respectively are indicated, yielding a simple anomaly (polyp) classification rule.*

- otherwise, it is classified as anomalous (with polyp, in this case).

For illustration purposes, as shown in Figure 5, for $T_h = 1$ and $T_d = 1.75$ about 22 % of speech with polyps were missed (classified as normal), whereas about 23 % of normal signals were wrongly classified as anomalous.

4. Discussion and Conclusions

In this paper a mathematical body-cover model of vocal folds with a polyp was developed, and a computational implementation of this model was obtained.

To illustrate the potential usefulness of such model, more than 800 sustained vowels from females and males, with and without polyps were simulated, and these samples were studied to produce a preliminary draft of a polyp detector from speech signal only (whose practical interest would be to avoid some invasive clinical exams [2]). In this brief illustration, a classifier was adjusted from synthetic data, yielding a classification result which is far from chance. Nevertheless, this classifier is presented here just as a proof of concept, and further development is necessary for it to be taken as a properly validated polyp detector.

On the other hand, this illustration suggests that the model presented in this paper can be used for, for instance, data augmentation for the purpose of fine tuning of actual polyp detectors from speech in future works. It is particularly useful if one considers that real databases, such as the one used in this work, are relatively rare, and typically hard to obtain (which possibly justifies why they are relatively small too).

5. Acknowledgements

This work has been supported by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) to J.S, with a graduate scholarship, and by The Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) to J.M., grant 308319/2018-4.

6. References

[1] D. de Vasconcelos, A. d. O. C. Gomes, and C. M. T. de Araújo, "Vocal fold polyps: literature review," *International archives of otorhinolaryngology*, vol. 23, no. 01, pp. 116–124, 2019.

[2] M. M. Johns, "Update on the etiology, diagnosis, and treatment of vocal fold nodules, polyps, and cysts," *Current opinion in otolaryngology & head and neck surgery*, vol. 11, no. 6, pp. 456–461, 2003.

[3] M. Behlau, G. Madazio, and P. Pontes, "Disfonias organofuncionais," in *VOZ - O Livro do Especialista*, M. Behlau, Ed. Livraria e Editora Revinter Ltda., 2001, vol. I, ch. 3, pp. 295–341.

[4] M. Hirano, W. J. Gould, A. Lambiase, and Y. Kakita, "Vibratory behavior of the vocal folds in a case with a unilateral polyp," *Folia Phoniatrica et Logopaedica*, vol. 33, no. 5, pp. 275–284, 1981.

[5] J. Van den Berg, "Myoelastic-aerodynamic theory of voice production," *Journal of speech and hearing research*, vol. 1, no. 3, pp. 227–244, 1958.

[6] B. D. Erath, M. Zañartu, K. C. Stewart, M. W. Plesniak, D. E. Sommer, and S. D. Peterson, "A review of lumped-element models of voiced speech," *Speech Communication*, vol. 55, no. 5, pp. 667–690, 2013.

[7] T. Koizumi and S. Taniguchi, "A novel model of pathological vocal cords and its application to the diagnosis of vocal cord polyp," in *First International Conference on Spoken Language Processing*, 1990.

[8] T. Koizumi, S. Taniguchi, and F. Itakura, "An analysis-by-synthesis approach to the estimation of vocal cord polyp features," *The Laryngoscope*, vol. 103, no. 9, pp. 1035–1042, 1993.

[9] Y. Zhang and J. J. Jiang, "Chaotic vibrations of a vocal fold model with a unilateral polyp," *The Journal of the Acoustical Society of America*, vol. 115, no. 3, pp. 1266–1269, 2004.

[10] I. Steinecke and H. Herzel, "Bifurcations in an asymmetric vocal-fold model," *The Journal of the Acoustical Society of America*, vol. 97, no. 3, pp. 1874–1884, 1995.

[11] J. J. Jiang, C. E. Diaz, and D. G. Hanson, "Finite element modeling of vocal fold vibration in normal phonation and hyperfunctional dysphonia: implications for the pathogenesis of vocal nodules," *Annals of Otology, Rhinology & Laryngology*, vol. 107, no. 7, pp. 603–610, 1998.

[12] R. Greiss, J. Rocha, and E. Matida, "Validation of a finite element code for a continuum model of vocal fold vibration under the influence of a sessile polyp," *Canadian Acoustics*, vol. 43, no. 1, pp. 13–23, 2015.

[13] D. A. Berry, H. Herzel, I. R. Titze, and K. Krischer, "Interpretation of biomechanical simulations of normal and chaotic vocal fold oscillations with empirical eigenfunctions," *The Journal of the Acoustical Society of America*, vol. 95, no. 6, pp. 3595–3604, 1994.

[14] J. Montalvão, "Noise variance estimation through joint analysis of intrinsic dimension and differential entropy," *IEEE Signal Processing Letters*, vol. 26, no. 9, pp. 1330–1333, 2019.

[15] S. T. Roweis, "Nonlinear dimensionality reduction by locally linear embedding," *Science*, vol. 290, no. 5500, pp. 2323–2326, Dec. 2000. [Online]. Available: https://doi.org/10.1126/science.290.5500.2323

[16] J. B. Tenenbaum, "A global geometric framework for nonlinear dimensionality reduction," *Science*, vol. 290, no. 5500, pp. 2319–2323, Dec. 2000. [Online]. Available: https://doi.org/10.1126/science.290.5500.2319

[17] L. v. d. Maaten and G. Hinton, "Visualizing data using t-sne," *Journal of machine learning research*, vol. 9, no. Nov, pp. 2579–2605, 2008.

[18] P. Grassberger and I. Procaccia, "Characterization of strange attractors," *Physical Review Letters*, vol. 50, no. 5, pp. 346–349, Jan. 1983. [Online]. Available: https://doi.org/10.1103/physrevlett.50.346

[19] J. A. Costa and A. O. Hero, "Determining intrinsic dimension and entropy of high-dimensional shape spaces," in *Statistics and Analysis of Shapes*. Birkhäuser Boston, 2006, pp. 231–252.

[20] M. Hirano, "Morphological structure of the vocal cord as a vibrator and its variations," *Folia Phoniatrica et Logopaedica*, vol. 26, no. 2, pp. 89–94, 1974.

[21] B. H. Story and I. R. Titze, "Voice simulation with a body-cover model of the vocal folds," *The Journal of the Acoustical Society of America*, vol. 97, no. 2, pp. 1249–1260, 1995.

[22] I. R. Titze, "Regulating glottal airflow in phonation: Application of the maximum power transfer theorem to a low dimensional phonation model," *The Journal of the Acoustical Society of America*, vol. 111, no. 1, pp. 367–376, 2002.

[23] ——, "Parameterization of the glottal area, glottal flow, and vocal fold contact area," *The Journal of the Acoustical Society of America*, vol. 75, no. 2, pp. 570–580, 1984.

[24] R. A. Samlan and B. H. Story, "Relation of structural and vibratory kinematics of the vocal folds to two acoustic measures of breathy voice based on computational modeling," *Journal of Speech, Language, and Hearing Research*, 2011.

[25] J. Kelly and C. Lochbaum, "Speech synthesis," in *Proceedings of the Speech Communication Seminar*, 1962, pp. 583–596.

[26] B. H. Story, "Physiologically-based speech simulation using an enhanced wave-reflection model of the vocal tract." Ph.D. dissertation, University of Iowa, 1995.

[27] B. H. Story, I. R. Titze, and E. A. Hoffman, "Vocal tract area functions from magnetic resonance imaging," *The Journal of the Acoustical Society of America*, vol. 100, no. 1, pp. 537–554, 1996.

[28] B. Story, "Letalker 1.22," 2016, disponível em: http://sal.arizona.edu/node/26 (último acesso em 7 de março de 2020).

[29] I. R. Titze and B. H. Story, "Rules for controlling low-dimensional vocal fold models with muscle activation," *The Journal of the Acoustical Society of America*, vol. 112, no. 3, pp. 1064–1076, 2002.

[30] K. Ishizaka and J. L. Flanagan, "Synthesis of voiced sounds from a two-mass model of the vocal cords," *Bell system technical journal*, vol. 51, no. 6, pp. 1233–1268, 1972.

[31] M. Pützer and W. J. Barry, "Saarbruecken voice database," available at: http://stimmdb.coli.uni-saarland.de/, 2007, Institut für Phonetik - Universität des Saarlandes.