

Proceeding Series of the Brazilian Society of Computational and Applied Mathematics

Técnica de extração da seção de *Poincaré* para análise de sinais de voz

Fernando Araujo de Andrade Sobrinho¹

Universidade de São Paulo, Escola de Engenharia de São Carlos, Departamento de Engenharia Elétrica. Instituto Federal de Educação, Ciência e Tecnologia do Sul de Minas Gerais.

Maria Eugênia Dajer²

Universidade Tecnológica Federal do Paraná. Campus Cornélio Procópio. Departamento de Engenharia Elétrica.

Luís Fernando Costa Alberto³

Universidade de São Paulo, Escola de Engenharia de São Carlos, Departamento de Engenharia Elétrica.

Resumo Um sinal de voz pode ser representado no espaço de fase tridimensional tomando amostras atrasadas desse sinal. A seção de *Poincaré* permite estudar, dentre outras características, a dispersão dos pontos nessa seção que podem auxiliar no diagnóstico de patologias. Esse artigo propõe uma técnica para extração da seção de *Poincaré* de um sinal de voz.

Palavras-chave. Seção de Poincaré, análise acústica de voz, reconstrução do espaço de fase, análise não linear, dinâmica não linear.

1 Introdução

Diversos estudos foram realizados para detecção de patologias na laringe. Esses estudos usualmente exploram as alterações de amplitude e período característicos de sinais de vozes patológicas através de técnicas como análise acústica, padrões visuais e análise não linear. Essas técnicas são não invasivas e auxiliam no diagnóstico de patologias realizados por fonoaudiólogos e otorrinolaringologistas, além de quantificar a evolução de terapias da voz. Apesar de promissoras, não substituem as técnicas tradicionais de análise de voz. Mas a importância de se utilizar técnicas não lineares para análise de voz é extrair novas informações que não podem ser obtidas com as técnicas convencionais.

Na análise não linear, vários parâmetros podem ser extraídos para o estudo de sinais de vozes como dimensão de correlação [1], expoente de *Lyapunov* [8] [10], entropia [2],

¹fsobrinho@usp.br; fernando_sobrinho@yahoo.com

²medajer@utfpr.edu.br

³lfcaberto@usp.br

padrões visuais da dinâmica vocal [3] e seção de Poincaré [12] [7] [5]. Na técnica proposta em [7] para extração da seção de *Poincaré*, não há informação de medidas quantitativas sobre a dispersão dos pontos na seção, enquanto que em [5], limita-se a extrair a seção nos pontos de máximo e mínimo do sinal temporal. Em [12] são extraídas apenas 10 seções de *Poincaré* do espaço de fase.

Esse trabalho tem como propósito explorar as características de não linearidade do sinal de voz apresentando um algoritmo para extrair a seção de *Poincaré* de seu espaço de fase de dimensão 3. A contribuição dessa técnica é que o algoritmo proposto armazena num vetor o ponto correspondente na série temporal ao ponto em que a trajetória corta o hiperplano de *Poincaré*, e com essa informação pode-se determinar a frequência fundamental e o *jitter* do sinal de voz. Vale ressaltar que o *jitter* e a frequência fundamental são importantes parâmetros para análise acústica de um sinal de voz. A técnica proposta também permite extrair a seção de *Poincaré* em todo o espaço de fase e calcular a média e o desvio padrão da dispersão dos pontos no plano, além de se obter o espaço de fase médio reconstruído.

2 Seção de *Poincaré*

Uma técnica clássica para análise de sistemas dinâmicos é devido a Henry Poincaré. Essa técnica transforma um sistema dinâmico contínuo no tempo em um sistema discreto utilizando o mapa de *Poincaré* ou seção de *Poincaré*.

Dado um sistema dinâmico autônomo de dimensão n apresentando órbitas periódicas em seu espaço de fase, a seção de *Poincaré* reduz a dimensão desse espaço de fase para um sistema de dimensão $(n - 1)$. Ela é determinada realizando-se um corte perpendicular às trajetórias do espaço de fase e tomando somente os pontos que interceptam essa seção.

A Figura 1 ilustra a seção de Poincaré realizada numa trajetória \mathbf{x} . Nessa Figura, x_i é um ponto da trajetória e Σ é o plano de dimensão $(n - 1)$ transversal à trajetória e que contém x_i . A trajetória iniciando em x_i cortará o plano Σ em T segundos no ponto x_{i+1} .

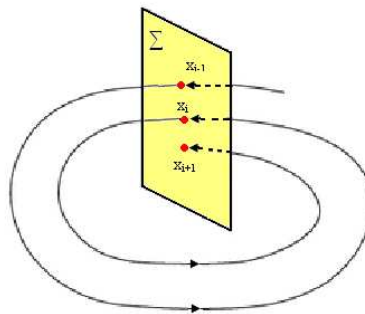


Figura 1: Plano Σ contendo os pontos pertencentes à seção de *Poincaré*. Adaptado de [11].

Quando houver uma órbita periódica γ , uma trajetória iniciando no plano $\Sigma \cap \gamma = p$ irá retornar ao plano Σ no ponto p após decorrido um tempo igual ao período da órbita. Órbitas iniciadas numa vizinhança suficientemente pequena nas vizinhanças de p irão interceptar o plano Σ nas vizinhanças de p . Daí pode-se dizer que Σ define um

mapeamento P de alguma vizinhança V contida em Σ de p , dentro de outra vizinhança V contida em Σ de p . P é chamada de mapa de *Poincaré* ou primeiro mapa de retorno [11].

Pela seção de *Poincaré* é possível verificar que as trajetórias de alguns sistemas dinâmicos no espaço de fase nunca repetem o mesmo percurso, mas estão confinadas dentro de um feixe bem limitado [7]. Em sistemas caóticos, a seção de *Poincaré* apresenta dimensão fractal e autosimilaridade. A presença de estruturas regulares na seção de *Poincaré* de sinais de vozes indica a presença de uma dinâmica de baixa dimensão [6]. Em sinais de vozes fortemente soprosas, por exemplo, a seção de *Poincaré* apresentou um aspecto nebuloso, indicando um comportamento caótico. Também é possível diferenciar grupos de vozes normais de patológicas através da dispersão dos pontos na seção de *Poincaré* [12]. Um dificuldade encontrada para extração da seção de *Poincaré* é a determinação de um plano transversal ao espaço de fase em 3 dimensões.

3 Método

Nesta seção, o método proposto para extração da seção de *Poincaré* do espaço de fase tridimensional é apresentado. Para se extrair a seção de *Poincaré* a partir de um sinal de voz amostrado $x(i)$, normaliza-se este sinal em relação ao seu valor máximo. Em seguida, seleciona-se um trecho do sinal no tempo e a partir desse, constrói-se um vetor de dimensão 3 contendo amostras atrasadas do sinal. O tempo de atraso é determinado pelo método da informação mútua [4]. A partir desse vetor gera-se o espaço de fase reconstruído, indicado na Figura 2a.

Na Figura 2a, selecionam-se 2 pontos $P_1(x_1, y_1)$ e $P_2(x_2, y_2)$ com o *mouse* de modo a formar uma reta tangente ao fluxo. A partir desses 2 pontos rotaciona-se todo o espaço de fase de um ângulo θ , de acordo com a expressão (1). Essa rotação é feita para se trabalhar com o eixo x como referência.

$$\theta = \arctan\left(\frac{y_1 - y_2}{x_1 - x_2}\right) \quad (1)$$

Após a rotação, é feita uma ampliação do trecho limitado pelos pontos P_1 e P_2 , resultando na Figura 2b. Nessa Figura, utilizando o *mouse*, selecionam-se 3 pontos indicados por $P_3(x_3; y_3)$, $P_4(x_4; y_4)$ e $P_7(x_7; y_7)$ entre os pontos P_1 e P_2 para delimitar a região do corte de onde se vai extrair a seção de *Poincaré*; e mais 2 pontos, indicados por $P_5(x_5; y_5)$ e $P_6(x_6; y_6)$, para definir a região de orientação da trajetória.

O Algoritmo seleciona a região do corte de acordo com a expressão (2) e (3).

Na direção do eixo x:

$$x_3 \leq x(i) \leq x_4 \quad (2)$$

Na direção eixo y:

$$y_7 \leq y(i) \leq \text{Min}(y_3; y_4) \quad (3)$$

E a região de orientação da trajetória de acordo com a expressão (4) e (5).

Na direção do eixo x:

$$x_5 \leq x(i) \leq x_6 \quad (4)$$

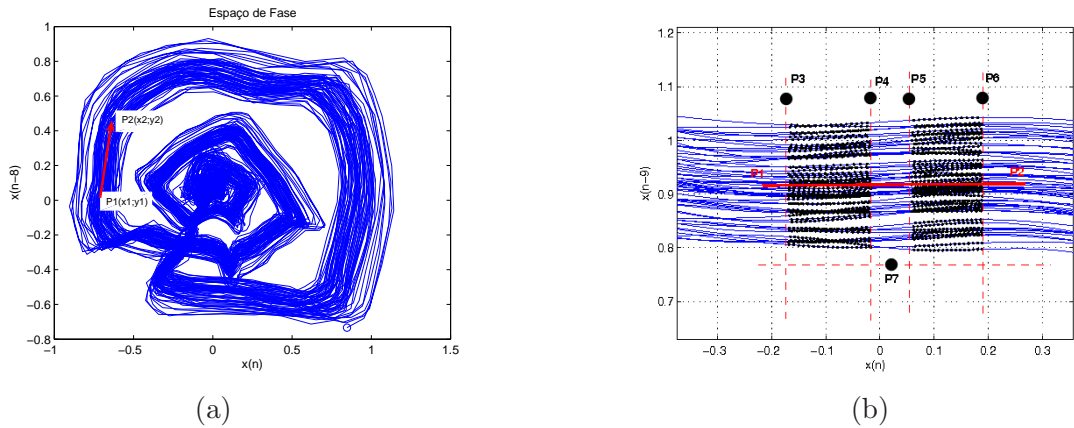


Figura 2: (a) Espaço de fase em 2 dimensões de uma voz saudável. (b) Espaço de fase ampliado. As barras tracejadas limitam a região do corte e o trecho de orientação da reta tangente ao fluxo das trajetórias.

Na direção eixo y :

$$y_7 \leq y(i) \leq \text{Min}(y_5; y_6) \tag{5}$$

Os pontos médios da região do corte entre os pontos P_3 e P_4 e da região de orientação da trajetória entre os pontos P_5 e P_6 formarão a reta normal ao plano de *Poincaré*, indicada pelo vetor \vec{OA} , ilustrado na Figura 3a. Em seguida, determinam-se os pontos do eixo z ($x(n - 18)$) e calcula-se o produto escalar entre o vetor \vec{OA} e o vetor $\vec{OB}(i)$, sendo $B(i)$ o i -ésimo ponto pertencente à região do corte, de acordo com a expressão (6).

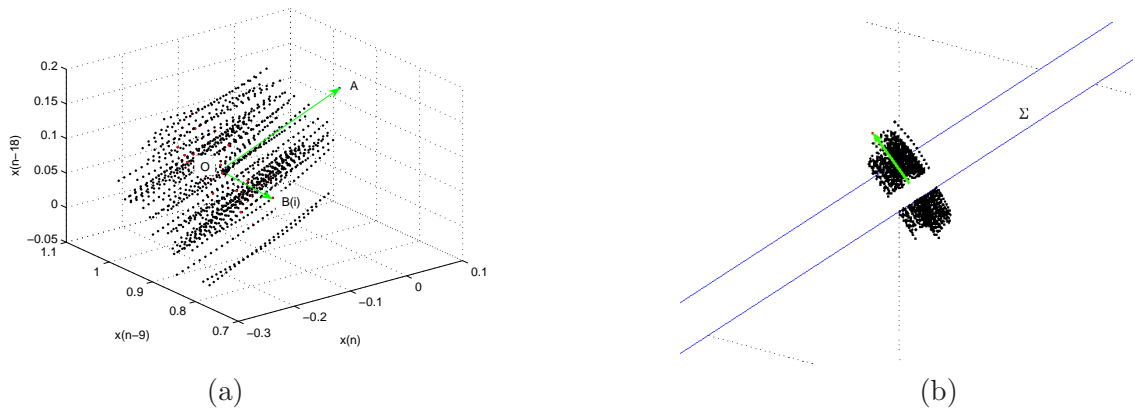


Figura 3: (a) Determinação dos pontos pertencentes ao plano de Poincaré através do produto escalar. Os pontos $B(i)$ pertencentes à seção são aqueles cujo resultado do produto escalar é igual a zero. (b) Visualização do plano Σ cortando o espaço de fase.

$$\vec{OA} \cdot \vec{OB}(i) \tag{6}$$

A condição para $B(i)$ pertencer ao plano de *Poincaré* (Figura 3b) é $\vec{OA} \cdot \vec{OB}(i) = 0$.

Depois de selecionar todos os pontos da seção, duas rotações são feitas em torno do eixo y e z , respectivamente. A rotação é feita utilizando as expressões (7), tomando-se a reta \vec{OA} como referência.

$$R_y(\theta) = \begin{bmatrix} \cos \theta & 0 & -\sin \theta \\ 0 & 1 & 0 \\ \sin(\theta) & 0 & \cos \theta \end{bmatrix}; R_z(\theta) = \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (7)$$

Em seguida, projetam-se os pontos no plano $x = 0$, resultando na seção de *Poincaré* (Figura 4a), e os respectivos pontos da seção no sinal temporal (Figura 4b). A importância de se obter um plano perpendicular à trajetórias é que, ao rotacionar os pontos através das equações 7, os pontos estarão num plano paralelo ao plano $x = 0$; por isso é possível projetá-los sem que haja distorção das medidas.

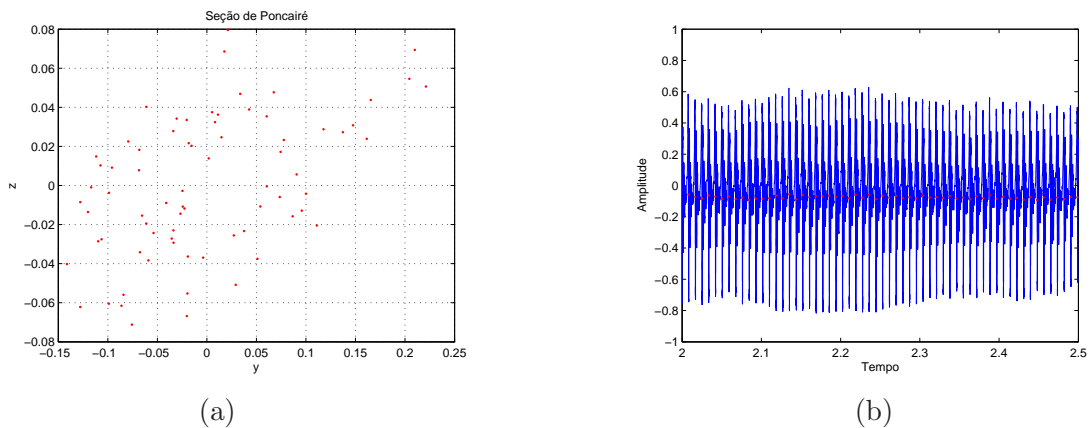


Figura 4: (a) Seção de *Poincaré* extraída da região do corte representado na Figura 3b. (b) Os pontos destacados na série temporal do sinal de voz representam os pontos em que a trajetória cruza o plano de *Poincaré*.

Resultados e Discussões

O resultado obtido com a técnica proposta mostra que é possível extrair a seção de Poincaré de todo o espaço de fase e ainda obter o conhecimento da posição do ponto da seção no sinal temporal. Com base nessa informação é possível extrair parâmetros do sinal de voz como média e desvio padrão da dispersão dos pontos na seção, *jitter*, mapeamento desses pontos além de se medir a frequência fundamental, importante parâmetro para análise de voz. Nesse artigo, o algoritmo foi aplicado num trecho de 500ms de um sinal de voz saudável da vogal sustentada 'a'. A frequência fundamental medida foi de 118

Hz, a média e o desvio padrão da dispersão foram 0.0571 e 0.0109, respectivamente. A projeção dos pontos no plano é útil para verificar que os ciclos das trajetórias do espaço de fase não se repetem devido ao *jitter* e *shimmer*, apesar de estarem confinadas dentro de um feixe limitado. Os resultados que podem ser obtidos com a técnica proposta podem contribuir com outras medidas como a dimensão de correlação, expoente de *Lyapunov*, entropia, etc. Essas medidas podem também alimentar uma rede neural para classificação de vozes saudáveis e patológicas, e diferenciação de patologias.

Uma outra vantagem dessa técnica é que, além do baixo tempo de processamento e armazenamento de informações, pode ser aplicada por qualquer profissional da área de voz após treinamento. Uma limitação é que, dependendo do tamanho do sinal selecionado no tempo, pode não ser possível selecionar os pontos da região do corte e os pontos da orientação da trajetória manualmente, pois os *loops* do espaço de fase se entrelaçam. Nesse caso, há necessidade de se estabelecer um tamanho máximo do sinal temporal antes de aplicar o algoritmo, ou refinar o algoritmo para sanar esse problema.

Determinado o plano de *Poincaré*, uma subrotina do algoritmo faz esse plano percorrer todo o espaço de fase gerando novas seções. De cada seção extrai-se se ponto médio e, dessa forma, determina-se o espaço de fase médio reconstruído ilustrado na Figura 5.

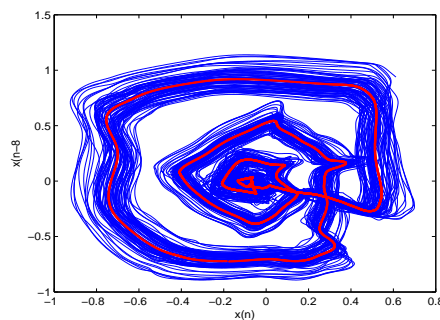


Figura 5: Espaço de fase médio obtido através do ponto medio de cada seção de *Poincaré* ao longo de todo espaço de fase.

4 Conclusão

A técnica proposta se mostrou eficaz para extração da seção de *Poincaré* de sinais de vozes saudáveis e patológicas de todo o espaço de fase. A principal aplicação dessa técnica é a extração de medidas quantitativas para classificação dessas vozes ou para acompanhamento de terapias de voz.

Agradecimentos

Agradeço ao departamento de engenharia elétrica da Escola de Engenharia de São Carlos (EESC/USP). E ao grupo CNPq de Pesquisa em Engenharia Médica por disponibilizar

o banco de dados de vozes saudáveis e patológicas para análises.

Referências

- [1] J. B. Alonso, F. D. Maria and C.M. Travieso, Using nonlinear features for voice disorder detection, In: ISCA Tutorial and Research Workshop (ITRW) on Non-Linear Speech Processing, (2005).
- [2] S. C. Costa, S. Correa, H. Falcão, N. Almeida and F. Assis, Uso da Entropia na Discriminação de Vozes Patológicas, In: II Congresso de Inovação da Rede Norte Nordeste de Educação Tecnológica, João Pessoa, Paraíba, (2007).
- [3] M. E. Dajer, Vocal Dynamic Visual Pattern for voice characterization, In: Journal of Physics: Conference Series, IOP Publishing, p. 012026, (2011).
- [4] A. Fraser and H. L. Swinney, Independent coordinates for strange attractors from mutual information, Physical review A, v. 33, n. 2, p. 1134, (1986).
- [5] R. Hegger, H. Kantz and T. Schreiber, Practical implementation of nonlinear time series methods: The TISEAN package, Chaos: An Interdisciplinary Journal of Non-linear Science, v. 9, n. 2, p. 413-435, (1999).
- [6] H. Herzel and J. Wendler, Evidence of chaos in phonatory samples, In: Second European Conference on Speech Communication and Technology, (1991).
- [7] G. Kubin, Poincaré section techniques for speech. In: Speech Coding For Telecommunications Proceeding, IEEE Workshop on, IEEE, p. 7-8, (1997).
- [8] A. Kumar and S. K. Mullick, Nonlinear dynamical analysis of speech, The Journal of the Acoustical Society of America, v. 100, n. 1, p. 615-629, (1996).
- [9] I. Mann and S. McLaughlin, Poincare maps and pitch detection in speech, In: Signals Systems and Chaos (Ref. No. 1997/393), IEE Colloquium on, IET, p. 5/1-5/5, (1997).
- [10] S. Narayanan and A. Alwan, A nonlinear dynamical systems analysis of fricative consonants, The Journal of the Acoustical Society of America, v. 97, n. 4, p. 2511-2524, (1995).
- [11] T. S. Parker and L. O. Chua, Chaos: A tutorial for engineers, Proceedings of the IEEE, v. 75, n. 8, p. 982-1008, (1987).
- [12] F. A. A. Sobrinho, Medida da dispersão da periodicidade de um sinal de voz normal e voz patológica através da seção de Poincaré, Dissertação de mestrado, Universidade de São Paulo - São Carlos, (2011).