

# Processamento digital de sinais no diagnóstico diferencial de doenças laríngeas benignas

## *Digital signal processing in the differential diagnosis of benign larynx diseases*

---

IUBERI CARSON ZWETSCH<sup>1</sup>  
RUBEM DUTRA RIBEIRO FAGUNDES<sup>2</sup>  
THAIS RUSSOMANO<sup>3</sup>  
DIOGO SCOLARI<sup>4</sup>

---

### RESUMO

**Objetivos:** Apresentar um novo procedimento em Processamento Digital de Sinais para avaliar distúrbios da voz, especificamente os relacionados às doenças benignas da laringe.

**Métodos:** Foram utilizados sinais acústicos gravados de vozes normais e com alterações, de pacientes atendidos em consultório, que realizaram videolaringoscopia. O sinal acústico foi digitalizado para o estudo das alterações acústicas vocais. Sobre estes sinais foi aplicada uma janela e análise cepstral.

**Resultados:** Observando-se as diferentes alterações das vozes em estudo, foi possível desenvolver um conjunto de parâmetros de análise, que permitiu um diagnóstico coerente das alterações laríngeas, baseado nos achados cepstrais. Casos escolhidos aleatoriamente, e que não participaram do estudo, foram identificados somente com as informações do conjunto de parâmetros com 80% de acerto, validando significativamente o método.

**Conclusões:** Este método computacional é muito eficiente e não invasivo, identificando com sucesso alterações vocais pela análise cepstral.

**DESCRIPTORIOS:** DOENÇAS LARÍNGEAS; CEPSTRUM; ANÁLISE CEPSTRAL.

### ABSTRACT

**Aims:** To present a new Digital Signal Processing procedure to evaluate voice disorders, specifically those related to the benign larynx diseases.

**Methods:** Recorded acoustic signals of normal and abnormal voices of outpatients subjected to videolaryngoscopy had been digitalized for the study of the alterations in vocal acoustics. Window and cepstral analysis were applied on these signals.

**Results:** Through the observation of the different alterations in the voices, it was possible to develop a set of parameters of analysis that allowed a coherent diagnosis of larynx alterations, based in the cepstral findings. Cases randomly selected, who had not participated of the study, had been identified only through the information of the set of parameters, with 80% of rightness, significantly validating the method.

**Conclusions:** This computing method is very efficient and non-invasive, successfully identifying vocal alterations by cepstral analysis.

**KEY WORDS:** LARYNX DISEASES; CEPSTRUM; CEPSTRAL ANALYSIS.

---

<sup>1</sup> Médico Otorrinolaringologista. Autor responsável.

<sup>2</sup> Professor PhD da Faculdade de Engenharia Elétrica e do Mestrado em Engenharia Elétrica PUCRS. Coordenador do Laboratório de Sistemas, Sinais e Computação da Faculdade de Engenharia - SiSC/FENG-PUCRS.

<sup>3</sup> Professora PhD das Faculdades de Medicina e Ciências Aeronáutica e do Mestrado em Engenharia Elétrica/Biomédica da PUCRS. Coordenadora do Laboratório de Microgravidade/IPCT-PUCRS.

<sup>4</sup> Engenheiro Eletricista, Ênfase em Telecomunicações.

## INTRODUÇÃO

Atualmente o diagnóstico das alterações laríngeas benignas é realizado pelo exame de videolaringoscopia, que é um exame da laringe, onde se utiliza um sistema de óptica endoscópica para a visualização das pregas vocais. Esta óptica é conectada a uma microcâmara e o exame é gravado com o sinal de voz através de um microfone. Certas alterações laríngeas, mesmo para médicos especialistas experientes, podem trazer dificuldade diagnóstica, pois às vezes são muito parecidas no aspecto, apesar de apresentarem origens e alterações fisiopatológicas diferentes. Essas dificuldades também são encontradas na realização de técnicas computacionais de processamento de sinais que, em determinados casos, não são eficientes ou suficiente para a diferenciação das alterações laríngeas.<sup>1</sup>

As seguintes alterações laríngeas: nódulo vocal, cisto vocal, pólipos vocais, edema de Reinke e sulco vocal, representam a grande maioria dos atendimentos de pacientes com alteração da voz, que não sejam alterações transitórias devido a infecções das vias aéreas superiores, onde temos uma inflamação das pregas vocais que geralmente cedem em alguns dias, com a melhora do quadro geral.<sup>2-5</sup> Na maioria dos casos, as alterações laríngeas citadas produzem uma rouquidão com características típicas de cada uma, principalmente quando analisadas por ouvintes mais experientes, tais como médicos otorrinolaringologistas ou profissionais da área da fonoaudiologia.<sup>3, 4</sup> Várias técnicas de análise de sinais da voz são estudadas para a identificação de alterações laríngeas.<sup>1, 6-14</sup>

A análise cepstral do sinal de voz para o estudo das alterações laríngeas é muito útil, permitindo trabalhar com o sinal da glote (excitação) separadamente das repercussões ressonantes do trato vocal, facilitando o entendimento das modificações que ocorrem nas pregas vocais. A aplicação desta técnica no estudo do sinal acústico de vozes alteradas poderia detectar modificações nas ondas que se relacionem com as alterações laríngeas e, conseqüentemente, identificar modelos para uma classificação, permitindo a obtenção de uma ferramenta de diagnóstico não-invasiva.<sup>8-12</sup>

A qualidade da voz depende do modo de fechamento e abertura da glote e da vibração das pregas vocais. Certas alterações laríngeas impedem que pregas vocais tenham uma vibração

glotal harmônica. Os principais fatores que determinam a vibração vocal são:

1. posição da prega vocal, ou a extensão em que as pregas vocais são aduzidas ou abduzidas;
2. mioelasticidade, ou o grau de elasticidade das pregas vocais (determinado pela posição e grau de tensão decorrente da contração do músculo vocal);
3. nível de pressão do ar através das pregas vocais.<sup>15</sup>

As alterações das pregas vocais também podem determinar que estas não vibrem em concordância, resultando em uma área vocal onde o trato vocal é excitado em duas frequências fundamentais diferentes.<sup>2,15</sup>

A análise cepstral do sinal de voz permite trabalhar com o sinal da glote (excitação) e do trato vocal (ressonância) separadamente, pelas suas propriedades homomórficas, separando das características do filtro do trato vocal da seqüência de excitação, o que facilita o estudo das alterações das pregas vocais. Dentro das propriedades matemáticas envolvidas no processo, salienta-se principalmente as transformadas de Fourier e funções logarítmicas que resultará em uma função chamada cepstral ou cepstro, responsável pela dissociação do sinal de voz.

O processo resumido de estimação do cepstro pode ser visto no diagrama de blocos apresentado na Figura 1.

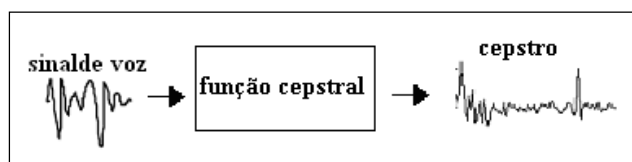


Figura 1 - Diagrama de blocos da estimação do cepstro.

A concentração das componentes aparece no cepstro como picos. O eixo horizontal da função cepstral tem dimensões temporais e o nome de quefrências. Com isto, na voz se obtém uma clara distinção entre a componente de excitação e a contribuição do trato vocal, que aparece como um aglomerado de componentes aos baixos valores de quefrência afastado da componente do período fundamental que aparece com valores mais altos de quefrência.

Na Figura 2, vê-se o cepstro de um segmento de voz onde o pico correspondente ao período

fundamental (excitação) está próximo da que-freência de 10 ms, separado das componentes do trato vocal às de baixas quefrências (trato vo-cal).<sup>1,6-8,16,17</sup>

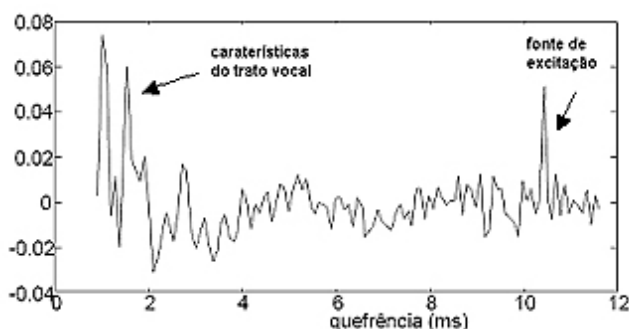


Figura 2 - Cepstro de um segmento de fala.

A função de transferência do trato vocal e a função de excitação da voz aparecem em partes separadas da escala de quefrências, pelo que podem ser separadas as duas funções, ou removida uma delas por um processo de “liffteragem” (*lifftering*), facilitando o estudo individualizado das modificações ocorridas na excitação e da parte ressonantal. Assim, o estudo do sinal acústico de vozes alteradas, através da aplicação da técnica de análise cepstral na detecção de modificações no sinal relacionadas às alterações laríngeas e na identificação de modelos para a classificação destes distúrbios em estudo, permitirá obter uma ferramenta não-invasiva de diagnóstico.

## MÉTODOS

Foram utilizados sinais acústicos gravados de pacientes com alterações vocais e de normais, atendidos em consultório, que realizaram videolarinoscopia. A distribuição das alterações foi a seguinte: 17 cistos, 5 edemas de Reinke, 9 nódulos, 9 sulcos e 8 pólipos. Também foram obtidos 13 sinais de vozes sem lesões nas pregas vocais e sem alteração acústica perceptível auditivamente, que compõem o grupo normal.

O sinal acústico de foi digitalizado em modo mono, com frequência de amostragem de 22 kHz e 16 bits de resolução de vocalizações específicas que permitissem obter-se informações sobre as modificações acústicas que as alterações laríngeas causam no sinal em estudo.<sup>1,7,16-18</sup>

Os dados obtidos foram submetidos à análise de Processamento Digital de Sinais em pro-

grama rotina de análise cepstral no Matlab<sup>19</sup> que consta dos seguintes passos (Figura 3):

1. obtenção de fragmento da vocalização e enjanelamento;
2. análise cepstral.

A pré-ênfase que permite a filtragem de sons labiais não foi usada, porque altera também o sinal de excitação (da glote), fenômeno descrito na literatura e constatado também neste trabalho. O enjanelamento foi ajustado para a análise do frame da vocalização em estudo e com o uso da análise cepstral.<sup>16,17</sup>

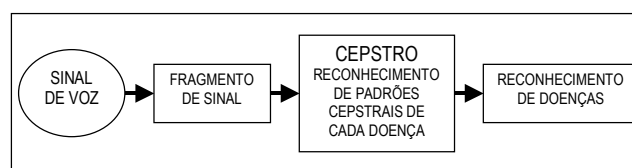


Figura 3 - Esquema em blocos da análise do sinal acústico.

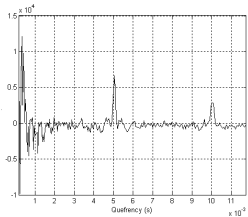
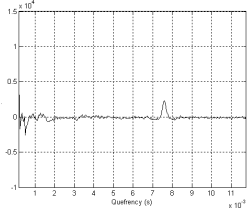
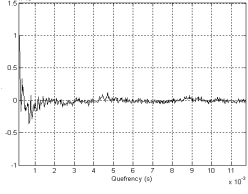
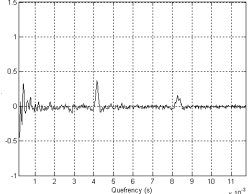
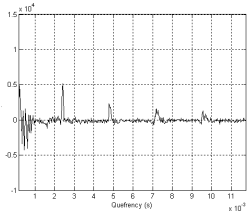
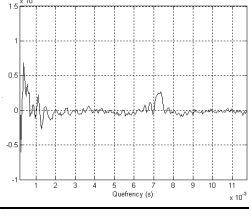
## RESULTADOS

Os achados cepstrais das vocalizações foram analisados. Com a ajuda das características fisiopatológicas já conhecidas de cada alteração laríngea foram descritas as modificações percebidas, as quais podem ser particularmente atribuídas às diferentes alterações laríngeas.

Um aspecto relevante percebido é quando se analisa os achados produzidos por três vocalizações separadamente, para a mesma alteração laríngea (e mesmo paciente) e para um caso normal. Nos achados do caso normal, as modificações das três vocalizações são muito discretas, mas nos casos das alterações laríngeas temos modificações diferentes em cada vocalização testada, provavelmente porque os distúrbios vibratórios provocadas pela alterações laríngeas se manifestam acusticamente de forma diferente, de acordo com o tipo de emissão vocal e de acordo com a alteração.

Na Tabela 1, apresenta-se, de maneira condensada, os dados das alterações da laringe e os achados do método proposto, comparando as cinco alterações e um caso normal com determinados fragmentos da vocalização. É possível observar-se que existem diferenças das alterações laríngeas no perfil cepstrográfico, quando comparadas com o normal e entre si. Estas diferenças se repetem na maioria dos casos analisados preliminarmente neste estudo.

TABELA 1 – Doenças das pregas vocais, suas características principais e aspectos diagnósticos pelos métodos convencional e o proposto na pesquisa.

Definição <sup>2,3,5</sup>	Características Acústicas <sup>2,3,5</sup>	Sistema cepstral de análise vocal (proposto)
<p><b>Normal</b> Pregas vocais sem lesões</p>	Som perceptivelmente adequado	
<p><b>Cisto</b> Lesão: cística, com fluido em geral unilateral</p>	Efeito: aumento da massa vibratória, em geral mais rigidez que nódulos/pólipos, assimetria vibratória Sensação acústica: pitch agravado, diplofonia, aspereza e rouquidão moderada	
<p><b>Edema</b> Lesão: edema da mucosa das pregas</p>	Efeito: aumento da massa e profundidade vibratória da onda, aperiódica, rigidez diminuída Sensação acústica: pitch muito agravado e rouquidão intensa	
<p><b>Sulco</b> Lesão: falha em forma de sulco, uni ou bilateral</p>	Efeito: diminuição da massa vibratória, onda “trava” no sulco e a rigidez é intensa Sensação acústica: pitch agudizado, soprosidade, pode apresentar bitonalidade e rouquidão	
<p><b>Nódulo</b> Lesão: quase sempre bilateral e simétrica</p>	Efeito: interferem na vibração dependendo de sua firmeza, aumento de massa Sensação acústica: rouquidão e pitch variáveis, soprosidade, aspereza, pode ter diplofonia	
<p><b>Pólipo</b> Lesão: em geral unilateral</p>	Efeito: aumento da massa vibratória, aumento da rigidez em geral, vibração assimétrica e aperiódica Sensação acústica: pitch agravado e rouquidão dependendo do tamanho	

Observando-se as diferentes alterações, foi possível desenvolver um conjunto de parâmetros de análise, agrupados em uma outra tabela, com os quais pôde-se diagnosticar de forma coerente as alterações, tomando como base os achados cepstrais de diferentes vocalizações. Para testar e validar este procedimento de lingüística, foram estudados casos escolhidos aleatoriamente e que não participaram do estudo, os quais foram

diagnosticados somente com as informações da tabela. O resultado obtido, quando os resultados dos gráficos foram submetidos aos valores da tabela, foi de 80% de acerto, validando significativamente o método proposto e contribuindo de forma efetiva para a elaboração de um protocolo de diagnóstico não-invasivo das alterações laringeas, baseado na técnica de análise cepstral.

## DISCUSSÃO

Observando as diferenças entre os resultados obtidos, resumidos na Tabela 1, percebe-se que existem diferenças entre as alterações laríngeas analisadas com o cepstro das diferentes vocalizações coletadas, sendo que algumas alterações laríngeas geram modificações significativas.

Por exemplo, os achados do edema de Reinke são os que mais se diferenciam de todas outras alterações laríngeas, pois esta alteração deixa o pitch muito agravado. O sulco vocal é uma alteração laríngea que deixa a prega vocal rígida e este é um dos fatores que produzem modificações significativas na voz. Todas as alterações laríngeas tendem se diferenciar significativamente dos normais, colaborando com a idéia de identificar as alterações pelas modificações por elas produzidas no sinal acústico.

Também se destaca que algumas alterações apresentam características e constantes que servem como método diagnóstico destas. Outras alterações nem sempre são constantes ou, às vezes, estão presentes em mais de uma alteração laríngea, e isto deve ser mais aprofundado em trabalhos futuros. Acredita-se que este método inédito, por ser não invasivo, de custo mais baixo e fácil execução e com resultados promissores, como demonstrado nestes dados prévios, será uma ferramenta de diagnóstico de triagem muito útil.

A Tabela 2 demonstra um agrupamento resumido das características principais obtidas no estudo.

TABELA 2 – Resumo das principais características de cada método.

	Diagnóstico com videolaringoscopia (convencional)	Método cepstral de análise da vocalização (sistema proposto)
Invasivo	Sim (pouco invasivo)	Não
Execução	Somente por profissional médico	Por qualquer pessoa com treinamento adequado
Triagem	Pouco prático para triagem	Muito adequado para triagem
Custos	Iguais aos de um procedimento em consultório	Mais baixos que o convencional

Esta técnica de análise gráfica cepstral fornece informações sobre as alterações laríngeas que não são demonstradas com outros métodos, principalmente no que se refere à visualização de regiões alteradas específicas e à possibilidade de

relacioná-las com as características anatomo-fisiológicas já conhecidas para as alterações laríngeas, sendo isto uma inovação nesta área.

Com os resultados da proposta de análise cepstral do sinal da voz nas alterações laríngeas iniciados neste trabalho, permite-se prever que este método será uma ferramenta diagnóstica muito útil e promissora para este propósito, pois é um método não-invasivo, de custo mais baixo e fácil execução. Para uma maior qualificação deste método, deve-se realizar trabalhos futuros com a inclusão de mais amostras, de outras alterações laríngeas mais raras e da utilização de um delineamento adequado para a validação, como um estudo prospectivo duplo cego.

Também é necessário desenvolver modelos de alterações laríngeas com o diagnóstico mais preciso possível para o treinamento do sistema, através de diagnóstico de certeza das alterações. O diagnóstico de certeza é obtido somente com uma biópsia da doença, mas pode-se obter um diagnóstico confiável com o acompanhamento da mesma e sua evolução com o tratamento, obtendo o diagnóstico evolutivo, sem ser invasivo. Outro aspecto relevante é determinar as possibilidades de que o sexo da pessoa testada possa influenciar o resultado, já que as alterações laríngeas tendem a alterar o pitch da voz, como observado na pesquisa.

E, ao ser obtido uma modelagem ideal das alterações com este método inovador, será possível usá-lo para treinamento mais preciso de um sistema de reconhecimento automático como, por exemplo, as redes neurais artificiais ou modelos de Markov,<sup>18</sup> tornando-se um importante equipamento de diagnóstico.

## REFERÊNCIAS

1. Rosa MO, Pereira JC, Greller M, Carvalho A. Signal processing and statistical procedures to identify laryngeal pathologies. *IEEE Transactions on Biomedical Engineering IEEE/EESC-USP*. 1999;423-6.
2. Dedivitis RAB. Métodos de avaliação e diagnóstico de laringe e voz. 2ª ed. São Paulo: Lovise; 2002.
3. Minoru H, Diane MB. Exame videostroboscópico da laringe. Porto Alegre: Artes Médicas; 1997.
4. Erich CM, Lupercio LB, Osiris CB, Behlau M, Melo DM. Incidência de lesões laríngeas não neoplásicas em pacientes com queixas vocais. *Revista Brasileira de Otorrinolaringologia*. 2001;67:788-94.
5. Khul I. Manual prático de laringologia. Porto Alegre: Editora da Universidade; 1982.
6. Martinez CE, Rufiner HL. Acoustic analysis of speech for detection of laryngeal pathologies. *International Conference, EMBS: Proceedings of the 22nd Annual EMBS, Chicago, IL, July 2000*. p. 23-8.

7. Rabiner RL, Schafer R. Digital processing of speech signals. New Jersey: Prentice-Hall; 1978.
8. Rabiner RL, Gold B. Theory and application of digital processing. New Jersey: Prentice Hall; 1975.
9. Gomez P, Godino JL, Rodriguez F. Evidence of vocal cord pathology from the mucosal wave Cepstral contents. Madrid: IEEE/Universidade Politécnica de Madrid; 2004. p. 45-52.
10. Wilpon JG, Rabiner LR, Lee CH. Automatic recognition of keywords in unconstrained speech using Markov Models. IEEE Transactions on Acoustic, Speech and Signal Processing. 1990;11:1870-8.
11. Hadjittodorov S, Mitev P. A computer system for acoustic of pathological voices and laryngeal diseases screening, Technical note. Medical Engineering & Physics. Bulgaria. 2002;24:419-29.
12. Manfredi C, D'aniello M, Brusaglioni P, Ismaelli A. A comparative analysis of fundamental frequency estimation methods with application to pathological voices. Medical Engineering & Physics. 2000;22:135-47.
13. Wszolek W, Tadeusiewicz R, Izvorski A, Wszolek T. Automated understanding of selected voice tract pathologies based on the speech signal analysis. International Conference, EMBS: Proceedings of the 23<sup>rd</sup> Annual EMBS, Istanbul, October 2001. p. 25-8.
14. Mitev P, Hadjttodorov S. Fundamental frequency estimation of voice of patients with laryngeal disorder. Information Sciences, Bulgaria. 2003;156:3-19.
15. Hansen JH, Gavdida-Ceballos L, Kaiser RJF. A non-linear operator-based speech feature analysis method with application to vocal fold pathology assessment. IEEE Transactions on Biomedical Engineering March. 1998;45:937-40.
16. Furui S. Digital speech processing, synthesis and recognition. New York: Marcel Dekker; 2001.
17. O'Shaughnessy D. Speech Communications: Human and Machine. New York: Institute of Electrical and Electronics Engineers; 2000.
18. Fagundes RDR. Reconhecimento de voz, linguagem contínua, usando modelos de Markov. São Paulo, 1993. [Dissertação de Mestrado].
19. Ingle VK, Proakis J. Digital signal processing using MATLAB. Canadá: Brooks/Cole; 2000.

**Endereço para correspondência:**  
IUBERI CARSON ZWETSCH  
Av. Ipiranga, 6681, Prédio 30 - PUCRS  
90619-900, Porto Alegre, RS, Brasil  
Fone: 9328-3996 - Fax: 3026-4652  
E-mail: iuberi@uol.com.br