

## Chapter 28

# Vocal Acoustic Analysis: ANN Versos SVM in Classification of Dysphonic Voices and Vocal Cords Paralysis

**João Paulo Teixeira**


 <https://orcid.org/0000-0002-6679-5702>

*Research Centre in Digitalization and Intelligent Robotics (CEDRI) and Applied Management  
Research Unit (UNIAG), Instituto Politécnico de Bragança, Bragança, Portugal*

**Nuno Alves**

*Instituto Politécnico de Bragança, Bragança, Portugal*

**Paula Odete Fernandes**

 <https://orcid.org/0000-0001-8714-4901>

*Applied Management Research Unit (UNIAG), Instituto Politécnico de Bragança, Bragança, Portugal*

### ABSTRACT

*Vocal acoustic analysis is becoming a useful tool for the classification and recognition of laryngological pathologies. This technique enables a non-invasive and low-cost assessment of voice disorders, allowing a more efficient, fast, and objective diagnosis. In this work, ANN and SVM were experimented on to classify between dysphonic/control and vocal cord paralysis/control. A vector was made up of 4 jitter parameters, 4 shimmer parameters, and a harmonic to noise ratio (HNR), determined from 3 different vowels at 3 different tones, with a total of 81 features. Variable selection and dimension reduction techniques such as hierarchical clustering, multilinear regression analysis and principal component analysis (PCA) was applied. The classification between dysphonic and control was made with an accuracy of 100% for female and male groups with ANN and SVM. For the classification between vocal cords paralysis and control an accuracy of 78,9% was achieved for female group with SVM, and 81,8% for the male group with ANN.*

DOI: 10.4018/978-1-6684-2408-7.ch028

## **1. INTRODUCTION**

Vocal Acoustic Analysis is often used for voice disorders assessment and diagnose (Bielamowicz et al., 1996; Brockmann-Bauser, 2011; Pylypowich, & Duff, 2016; Salhi, Mourad, & Cherif, 2010; Teixeira & Fernandes, 2015). The advantage of such techniques relies on the non-invasive character of the exam when compared with current practice in medicine, for example, laryngoscopy or stroboscopic exams (Brockmann-Bauser, 2011).

Both laryngoscopy and stroboscopic exam consist in inserting a thin tube into the throat or into the nostrils. Stroboscopy is painless, an office-based procedure done with topical anaesthesia. It is a special method used to visualize vocal fold vibration (Hirano, 1974). It uses a synchronized, flashing light passed through a flexible or rigid telescope. The flashes of light from the stroboscope are synchronised to the vocal fold vibration at a slightly slower speed, allowing the examiner to observe vocal fold vibration during sound production in what appears to be slow motion. The resulting video depicts video-stroboscopic examination of the vocal folds.

This incision technique will always be necessary to confirm or even support chirurgical operations on the vocal folds or in the larynx/pharynx.

Although voice disorders may be diagnosed by an auditory perceptual analysis made by the otolaryngologist, this may lead to different results depending on the practitioner experience (Teixeira & Fernandes, 2014).

It is common in daily life of primary care facilities the people complain about hoarseness in their voices. The dysphonia affects 30% of adults and 50% of older adults. This disease modifies voice quality and has a significant impact on life quality. This also represents a significant economic burden. In patients with a progressive pathology, it is important to do a diagnosis as fast as possible for the sake of having access to better treatment and prognosis (Pylypowich & Duff, 2016).

There are several acoustic parameters extracted from speech signal processing useful to identify the vocal pathology, yet no parameter alone is able to classify between healthy or pathologic voice.

Teixeira and Fernandes (2015) analysed the statistical significance of Jitter, Shimmer and HNR parameters for dysphonia detection. A statistical analysis was performed over the three parameters for the vowels /a/, /i/ and /u/ at three different tones, high, low and normal. In this work, Jitter and Shimmer are suggested as good parameters to be used in an intelligent diagnosis system of dysphonia pathologies.

To test this analysis, it is necessary to apply an intelligent tool and some reduction dimension and feature selection techniques. Feature selection is intended to select the best subset of predictors. The feature selection problem arises from large datasets who may contain redundant information and variables that have little or no predictive power (May, Dandy, & Maier, 2011). The correct choice of input features leads to a small subset that may boost/improve the performance when intelligent tools are used.

Henríquez et al. (2009) studied the usefulness of six nonlinear chaotic measures based on nonlinear dynamics theory in the discrimination between two levels of voice quality: healthy and pathological. The studied measures are first and second order Rényi entropies, the correlation entropy and the correlation dimension. The values of the first minimum of mutual information function and Shannon entropy were also studied. Two databases were used to assess the usefulness of the measures: a multi-quality and a commercial database (MEEI Voice Disorders). A classifier based on standard neural networks was implemented in order to evaluate the measures proposed. Global success rates of 82.5% (multi-quality database) and 99.7% (commercial database) were obtained. This difference in performance highlights the importance of having a controlled speech acquisition process.

15 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

[www.igi-global.com/chapter/vocal-acoustic-analysis/288977](http://www.igi-global.com/chapter/vocal-acoustic-analysis/288977)

## Related Content

---

### Adaptive Hybrid Higher Order Neural Networks for Prediction of Stock Market Behavior

Sarat Chandra Nayak, Bijan Bihari Misra and Himansu Sekhar Behera (2016). *Applied Artificial Higher Order Neural Networks for Control and Recognition* (pp. 174-191).

[www.irma-international.org/chapter/adaptive-hybrid-higher-order-neural-networks-for-prediction-of-stock-market-behavior/152103](http://www.irma-international.org/chapter/adaptive-hybrid-higher-order-neural-networks-for-prediction-of-stock-market-behavior/152103)

### Hierarchical Neural Networks for Modelling Adaptive Financial Systems

Masoud Mohammadian and Mark Kingham (2006). *Artificial Neural Networks in Finance and Manufacturing* (pp. 109-123).

[www.irma-international.org/chapter/hierarchical-neural-networks-modelling-adaptive/5351](http://www.irma-international.org/chapter/hierarchical-neural-networks-modelling-adaptive/5351)

### Transfer Learning in 2.5D Face Image for Occlusion Presence and Gender Classification

Sahil Sharma and Vijay Kumar (2019). *Handbook of Research on Deep Learning Innovations and Trends* (pp. 97-113).

[www.irma-international.org/chapter/transfer-learning-in-25d-face-image-for-occlusion-presence-and-gender-classification/227846](http://www.irma-international.org/chapter/transfer-learning-in-25d-face-image-for-occlusion-presence-and-gender-classification/227846)

### Neuroglial Behaviour in Computer Science

Ana B. Porto and Alejandro Pazos (2006). *Artificial Neural Networks in Real-Life Applications* (pp. 1-21).

[www.irma-international.org/chapter/neuroglial-behaviour-computer-science/5361](http://www.irma-international.org/chapter/neuroglial-behaviour-computer-science/5361)

### Neuromorphic Adiabatic Quantum Computation

Shigeo Sato and Mitsunaga Kinjo (2009). *Complex-Valued Neural Networks: Utilizing High-Dimensional Parameters* (pp. 352-375).

[www.irma-international.org/chapter/neuromorphic-adiabatic-quantum-computation/6775](http://www.irma-international.org/chapter/neuromorphic-adiabatic-quantum-computation/6775)