

Speech-Based Clinical Diagnostic Systems

Jesús Bernardino Alonso Hernández

University of Las Palmas de Gran Canaria, Spain

Patricia Henríquez Rodríguez

University of Las Palmas de Gran Canaria, Spain

INTRODUCTION

It is possible to implement help systems for diagnosis oriented to the evaluation of the **fonator system** using speech signal, by means of techniques based on expert systems. The application of these techniques allows the early detection of alterations in the **fonator system** or the temporary evaluation of patients with certain treatment, to mention some examples. The procedure of measuring the **voice quality** of a speaker from a digital recording consists of quantifying different acoustic characteristics of speech, which makes it possible to compare it with certain reference patterns, identified previously by a “**clinical expert**”.

A speech acoustic quality measurement based on an auditory assessment is very hard to assess as a comparative reference amongst different voices and different human experts carrying out the assessment or evaluation.

In the current bibliography, some attempts have been made to obtain objective measures of speech quality by means of multidimensional clinical measurements based on auditory methods. Well-known examples are: GRBAS scale from Japon (Hirano, M., 1981) and its extension developed and applied in Europe (Dejonckere, P. H. Remacle, M. Fresnel-Elbaz, E. Woisard, V. Crevier-Buchman, L. Millet, B., 1996), a set of perceptual and acoustic characteristics in Sweden (Hammarberg, B. & Gauffin, J., 1995), a set of phonetics characteristics with added information about the excitement of the vocal tract. The aim of these (quality speech measurements) procedures is to obtain an objective measurement from a subjective evaluation.

There exist different works in which objective measurements of speech quality obtained from a recording are proposed (Alonso J. B., 2006), (Boyanov, B & Hadjitodorov, S., 1997), (Hansen, J. H. L., Gavidia-Ceballos, L. & Kaiser, J. F., 1998), (Stefan Hadjitodorov & Petar Mitev, 2002), (Michaelis D.; Frohlich M. & Strube H.

W., 1998), (Boyanov B., Doskov D., Mitev P., Hadjitodorov S. & Teston B., 2000), (Godino-Llorente, J. I.; Aguilera-Navarro, S. & Gomez-Vilda, P., 2000).

In these works a voiced **sustained sound** (usually a vowel) is recorded and then used to compute speech quality measurements. The utilization of a voiced **sustained sound** is due to the fact that during the production of this kind of sound, the speech system uses almost all its mechanisms (glottal flow of constant air, vocal folds vibration in a continuous way, ...), enabling us to detect any anomaly in these mechanisms. In these works different sets of measurements are suggested in order to quantify speech quality objectively. In all these works one important fact is revealed; it is necessary to obtain different measurements of the speech signal in order to compile the different aspects of acoustic characteristics of the speech signal.

BACKGROUND

A **speech recording** gives different characteristics of the speech quality of a speaker. The recorded speech signal can be represented in different domains. Each domain shows some of the speech characteristics in a preferential way. The main domains studied in speech processing are:

- *Time Domain*
- *Spectral Domain*
- *Cepstral Domain*
- *Inverse Model Domain*

Most works in digital speech signal processing are based on these domains. However, other works use new domains derived from the former ones.

In the following section the most important features of each domain are described.

Time Domain

A high quality speech signal possesses a more regular envelope than a low quality speech signal. This fact is more evident in short time intervals. The main phenomena that enable us to distinguish between high quality speech and low quality speech are:

The energy of the speech signal in a short time interval changes considerably between two consecutive intervals in low quality speech whereas in high quality speech there is a less change in energy.

In low quality speech unperiodicity (without periodicity) intervals during voiced sustained speech appear.

Spectral Domain

A low quality speech (a voiced **sustained sound**) has the following characteristics:

- Less regularity of the spectral envelope, mainly in low frequencies.
- More percentage of energy in low frequencies with regard to the total energy.
- Energy blocks in high frequencies. These blocks are caused by glottal noise.
- A great change of the power spectrum from a frame with regard to contiguous frames.

Figure 1. Speech signal in time domain: the five Spanish and sustained vowels are illustrated. The upper figure is a speaker with high quality speech. The lower figure is a speaker with low quality speech.

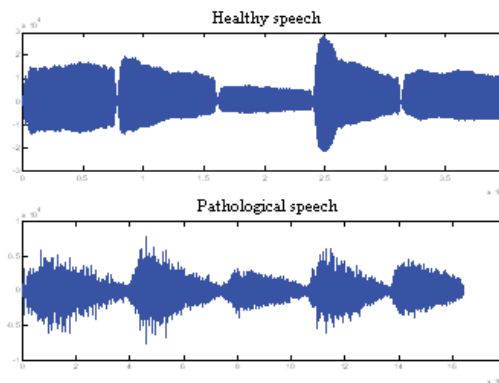
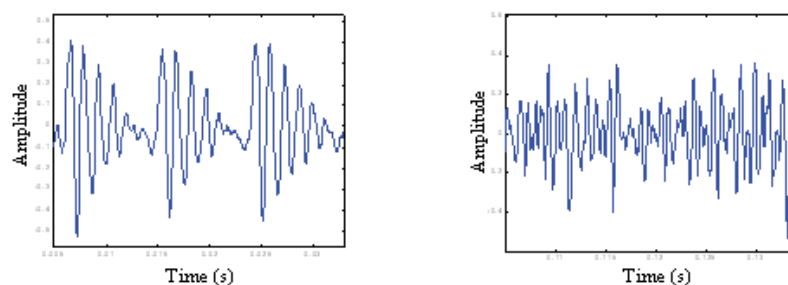


Figure 2. Sustained voiced sound during a short time interval from a high quality speech (left) and from a low quality speech (right).



6 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/speech-based-clinical-diagnostic-systems/10428

Related Content

Analysis on Detecting Cyber Security Attacks Using Deep Ensemble Learning on Smart Grids

K. Vanitha, M. Mohamed Musthafa, A. M. J. Md Zubair Rahman, K. Anitha, T. R. Maheshand V. Vinoth Kumar (2023). *Handbook of Research on Advancements in AI and IoT Convergence Technologies* (pp. 229-246).

www.irma-international.org/chapter/analysis-on-detecting-cyber-security-attacks-using-deep-ensemble-learning-on-smart-grids/330068

Optimization of Queuing Theory Based on Vague Environment

Verónica Andrea González-López, Ramin Gholizadehand Aliakbar M. Shirazi (2016). *International Journal of Fuzzy System Applications* (pp. 1-26).

www.irma-international.org/article/optimization-of-queuing-theory-based-on-vague-environment/144201

Generative AI Soaring in the Skies: New Horizons in Aviation English

Nazmi Dinçerand Gökhan Demirdöken (2024). *Transforming Education With Generative AI: Prompt Engineering and Synthetic Content Creation* (pp. 388-410).

www.irma-international.org/chapter/generative-ai-soaring-in-the-skies/338547

Ambient Interface Design (AID) for the Ergonomically Challenged

Rosaleen Hegarty, Tom Lunney, Kevin Curranand Maurice Mulvenna (2010). *International Journal of Ambient Computing and Intelligence* (pp. 57-64).

www.irma-international.org/article/ambient-interface-design-aid-ergonomically/43863

Machine Learning Approaches for Sentiment Analysis

Basant Agarwaland Namita Mittal (2017). *Artificial Intelligence: Concepts, Methodologies, Tools, and Applications* (pp. 1740-1756).

www.irma-international.org/chapter/machine-learning-approaches-for-sentiment-analysis/173400