INPE – National Institute for Space Research
São José dos Campos – SP – Brazil – May 16-20, 2016

ESTIMATION OF PITCH PERIOD IN VOICE SIGNALS USING POINCARE SECTION

Fernando Araujo de Andrade Sobrinho[1], *Maria Eugênia Dajer*[2], *Luis Fernando Costa Alberto*[3]

[1]Instituto Federal de Educação, Ciência e Tecnologia do Sul de Minas Gerais, Poços de Caldas, Brazil, fernando.sobrinho@ifsuldeminas.edu.br
[2]Universidade Tecnológica Federal do Paraná , Cornélio Procópio, Brazil, medajer@utfpr.edu.br
[3]Universidade de São Paulo, São Carlos, Brazil, lfcalberto@usp.br

Abstract: Researchers reported instability in children voices due to physiological changes in laryngeal structures. The aim of this paper is to detect the Fundamental frequency F_0 of healthy children voices using Poincaré Section technique (PST). The statistic T-test indicates that measures of F_0 using PST are compatible with those obtained with Fourier transform. Results indicate that PST is a promising tool for F_0 analysis with the advantage of determining a fundamental frequency even in high nonlinear and non-periodic signals.

keywords: Time Series Analysis, Analysis and Control of Nonlinear Dynamical Systems with Practical Applications, Synchronization in Nonlinear Systems, Poincare Section, Nonlinear analysis.

1. INTRODUTION

One of the most important parameters to pathologic voice analysis is referred to fundamental frequency (F_0). A classical tool do determine F_0 of a voice signal uses Fourier Transform [5] where voice signal is represented in its frequency spectrum. The fundamental frequency is given by the first peak of the spectrum.

In general, Fourier transform is applied to the analysis of periodic signals with several harmonics components. In the last decades, several researches discovered that voice signals present non-linearity [2–4] and the application of Fourier transform may not extract the exact fundamental frequency.

In case of strong non-linearity, frequency spectrum may present two close peaks, as indicate in Figure 1, and the question is "what's the correct fundamental frequency?".

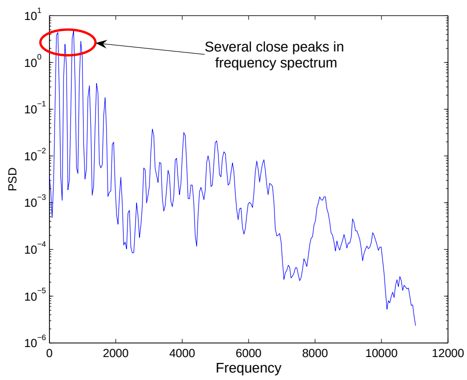

Figure 1 – Frequency spectrum with two close peaks

Using concepts of nonlinear analysis, a voice signal can be represented in a multi-dimensional phase space taking several delay samples of this signal. According to [1], the flow of trajectory in phase space repeats itself after one pitch period, T_0. The pitch period may be marked by means of a cross section through the phase space or *Poincaré* section. The purpose of this paper is presenting a tool to extract fundamental frequency of a voice signal.

The method used to data analysis was the statistical T-test at the default of 5% significance level. T-test is used to verify if 2 groups belong to the same sample. This means that if *p-value* is less than 5%, the samples belong to the different samples and are statistically different. Otherwise, the samples are statistically equal.

2. EXTRACTING POINCARÉ SECTION

In this section, a basic *Poincaré* section algorithm to extract F_0 is described.

1. Construct a 2-dimensional trajectory $X(t)$ from the scalar speech waveform x(t) over a stationary time window using eq. 1.

$$X = (x(t_k), x(t_k - \tau)) \tag{1}$$

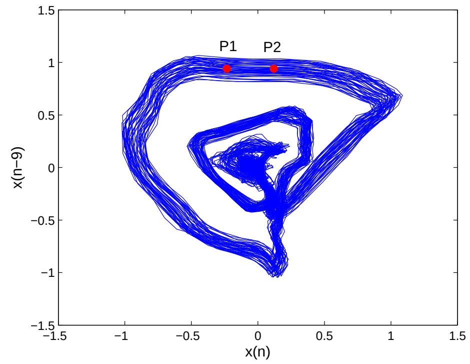

Figure 2 – Phase space of a voice signal

3. Select five points indicated by $P_3(x_3; y_3)$, $P_4(x_4; y_4)$, $P_5(x_5; y_5)$, $P_6(x_6; y_6)$ and $P_7(x_7; y_7)$ in the neighborhood of P_1 and P_2, as indicated in Figure 3.

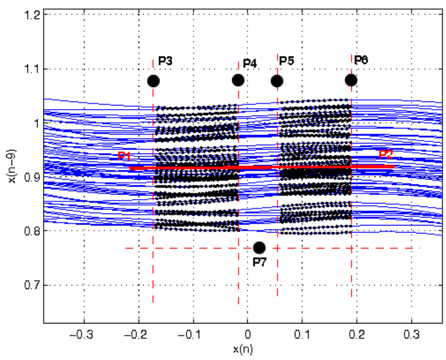

Figure 3 – An amplified phase space

4. Define a vector $\overrightarrow{OA}$ wich is tangent to the average trajectory. O is the medium point limited by P_3,P_4 and P_7 and A is the medium point limited by P_5,P_6 and P_7, the dark areas of Figure 3.

5. Define B_{ixj} as a matrix that contains the coordinates of the points in the phase space delimited by P_3, P_4 and P_7. Here, i is the position of the points and j is the segment of trajectory, as shown in Figure 4.

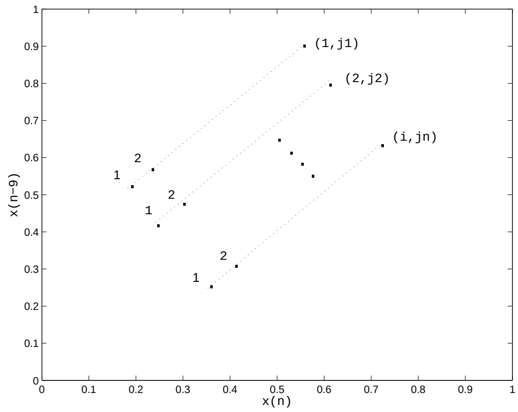

Figure 4 – Segment of trajectory

6. Make an inner product among vectors $\overrightarrow{OA}$ and $\overrightarrow{OB(i,j)}$. The point with the least inner product for each segment of trajectory, in module, will be the one that is closest to the *Poincaré* section.

7. Define the third coordinate $x(t_k - 2\tau)$. The third coordinate is necessary to represent the *Poincaré* section in 3 dimensions, as showed in Figure 5.

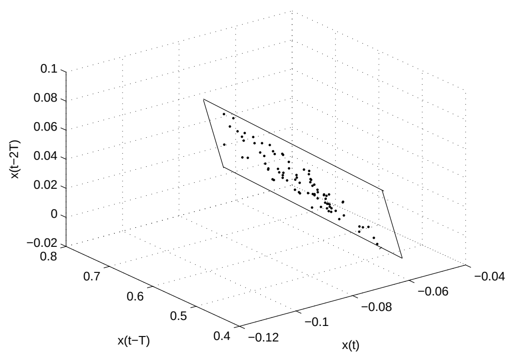

Figure 5 – Points of Poincaré section in 3-dimensional phase space

Once we have several cycles in phase space, we can determine the T_0 using eq. 2.

$$T_0 = \frac{\sum_{i=1}^{n} T_i}{n} \tag{2}$$

where n is the number of cycles.

Fundamental frequency will be $F_0 = T_0^{-1}$.

Graphic of Figure 6 show 10 cycles of trajectory. In this case the fundamental frequency is about 120 Hz.

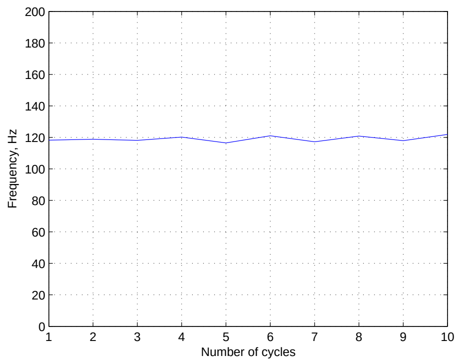

Figure 6 – Fundamental frequency versus number of cycles

3. RESULTS

We analyze 13 voice signals children obtained from the Medical Engineer Research Group [6]. Subjects were asked to sustain vowel sound $\backslash a \backslash$, in a comfortable conversation level, from brazilian portuguese, during 5 seconds. Voices were recorded at a sampling rate of 22,050 samples per second. We extracted segments of 200ms from this signal wich presented stationarity in amplitude. Fundamental frequency F_0 was extracted using Fourier Transform and *Poincaré* section. Results are shown in Table 1.

Table 1 – Fundamental frequency extract using Fourier transform and Poincaré section

	Fundamental frequency, F_0	
Children	Fourier transform [Hz]	Poincaré section [Hz]
1	172.3	191.49
2	258.4	260.14
3	301.5	324.93
4	258.4	256.4
5	215.3	201.75
6	301.5	297.00
7	258.4	244.73
8	215.3	202.65
9	215.3	198.33
10	258.4	236.59
11	215.3	222.49
12	215.3	228.02
13	258.4	263.48

Medium and standard deviation of both frequencies are shown in Table 2.

Table 2 – Medium and standard deviation of F_0

	Medium [Hz]	Standard deviation [Hz]
μ	241.83	240.62
σ	37.47	39.91

3.1 Statistical Test

We applied the statistical T-test and the *p-valeu* obtained was 0.94%. This means that the frequencies extracted from Fourier Transform and *Poincaré* section are statistically equal.

4. DISCUSSION

Although the Fourier transform is commonly applied to extract the frequency spectrum of voice signals, that is considered almost periodic, problems may occurs in very high nonlinear signals. *Poincaré* section deals with nonlinearity and eliminates the problems of computing F_0 using Fourier transform. We suggest that extracting F_0 using *Poincaré* section improves and complements the Fourier transform technique.

5. CONCLUSION

This paper presents a tool based in nonlinear dynamic to extract fundamental frequency of a voice signal. The technique is based on the extraction of a *Poincaré* section of the voice signal and has shown to be suitable to extract the F_0 of a voice signal.

6. ACKNOWLEDGMENTS

Thanks to Medical Engineer Research Group and CNPq Foundation for making data available for our study.

References

[1] KUBIN, Gernot. Poincaré section techniques for speech. In: Speech Coding For Telecommunications Proceeding, 1997, 1997 IEEE Workshop on. IEEE, 1997. p. 7-8.

[2] ALONSO, Jesús B. et al. Using nonlinear features for voice disorder detection. In: ISCA Tutorial and Research Workshop (ITRW) on Non-Linear Speech Processing. 2005.

[3] KUMAR, Arun; MULLICK, S. K. Nonlinear dynamical analysis of speech. The Journal of the Acoustical Society of America, v. 100, n. 1, p. 615-629, 1996.

[4] NARAYANAN, Shrikanth S.; ALWAN, Abeer A. A nonlinear dynamical systems analysis of fricative consonants. The Journal of the Acoustical Society of America, v. 97, n. 4, p. 2511-2524, 1995.

[5] RABINER, Lawrence R.; SCHAFER, Ronald W. Digital processing of speech signals. Prentice Hall, 1978.

[6] http://dgp.cnpq.br/dgp/espelhogrupo/6246151242882953