Voice pathology detection by means of acoustic analysis and nonlinear dynamics techniques

Juan Rafael Orozco Arroyave
Universidad de Antioquia – ARTICA research center
Medellín, Colombia

2’34’’ About Colombia!!!

http://www.youtube.com/watch?v=8kUU-DWOqmI

Voice pathology detection by means of acoustic analysis and nonlinear dynamics techniques

CONTENT

1. Introduction

- Most of the voice pathologies are suffered by people who continously work with voice.

  i.e. Professors, customer service, communications, etc.
- Voice pathologies due to neurodegenerative disorders like Parkinson disease are also common.
  - Some pathologies: Unilateral laryngeal paralysis, dysphonia, vocal fold paralysis, nodules, hyponasality, **hypernasality**, etc.
- Treatment costs for the health system are increasing.
  - Voice's treatment costs for professors in USA is about 2.5billion annually [1].

1. Introduction

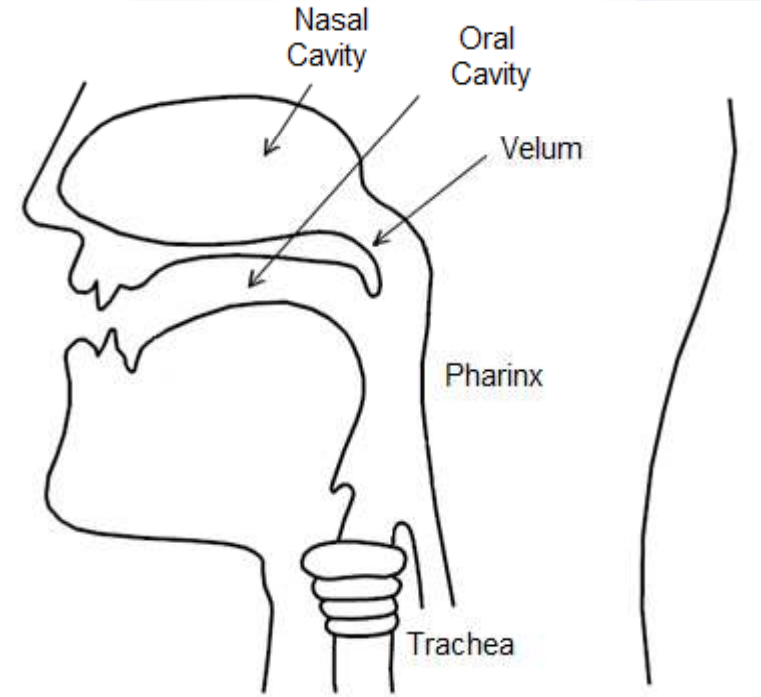
CLP: Cleft Lip and Palate

One of the most frequent congenital malformation around the world

HYPERNASALITY

About 1.5 of each 1.000 children (Europe)
About 1 of each 1.000 children (USA, Colombia)

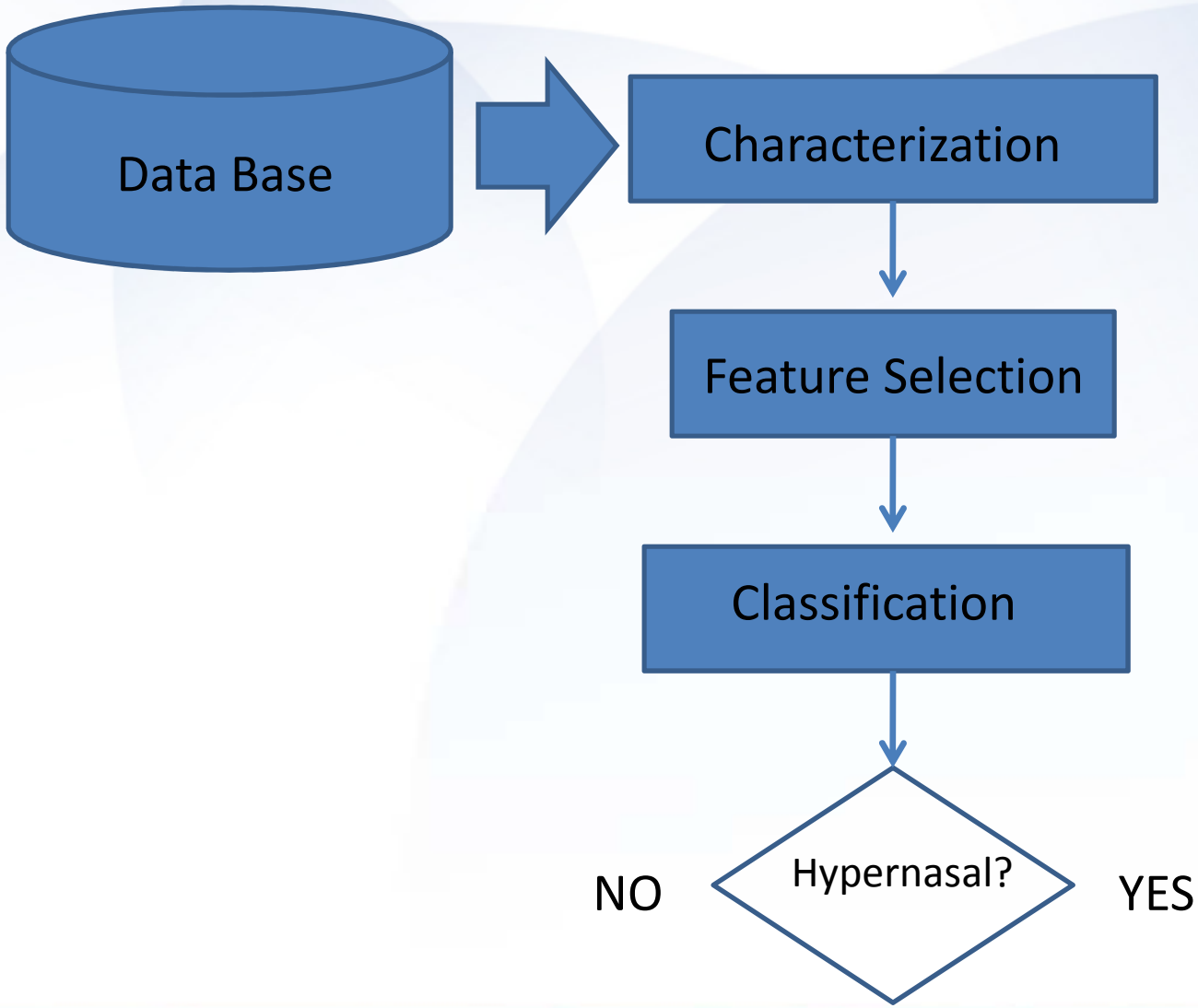
[2, 3]



SPEECH THERAPY

There isn't an objective measure to quantify the progress of the treatment.

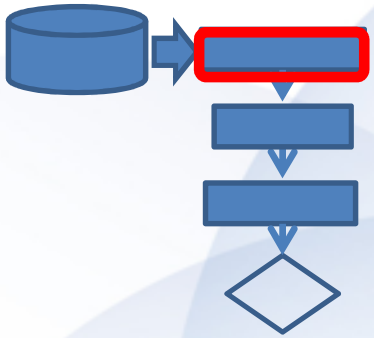
2. Methodology



2.1. Database

CG&PDS at Universidad Nacional de Colombia, branch Manizales.

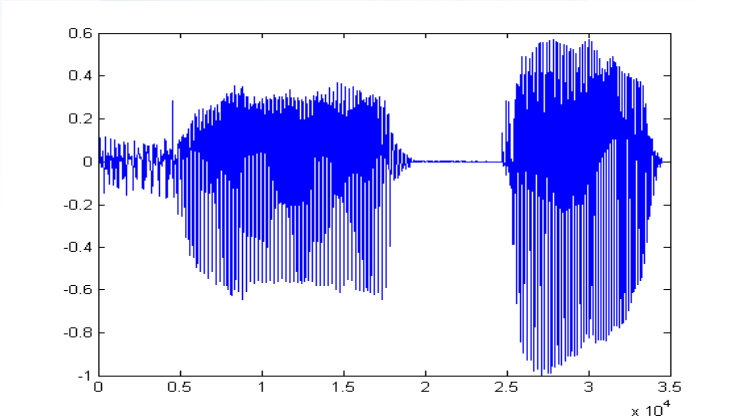
- Population: Children aged between 5 to 15
- Five Spanish vowels pronounced in sustained manner
- Frequency sample: 44.100Hz
- 156 hypernasal registers
- 110 healthy registers



2.2. Characterization

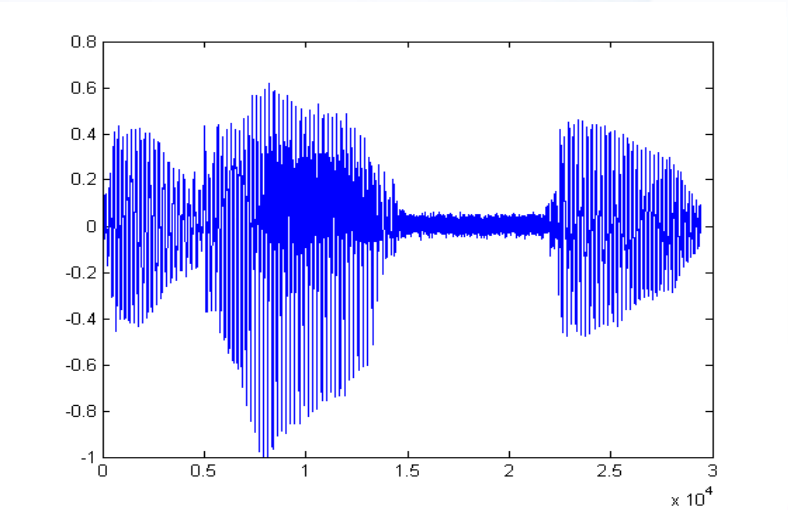
Acoustic, Cesptral and Noise Features

Voice signal in time domain: Jitter and Shimmer

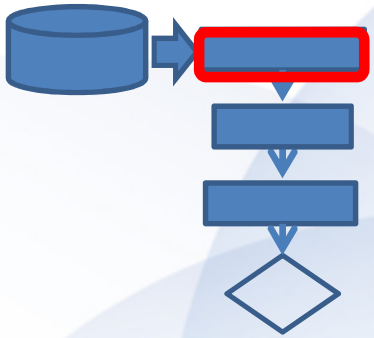


Healthy Voice /gato/

Hypernasal Voice /gato/

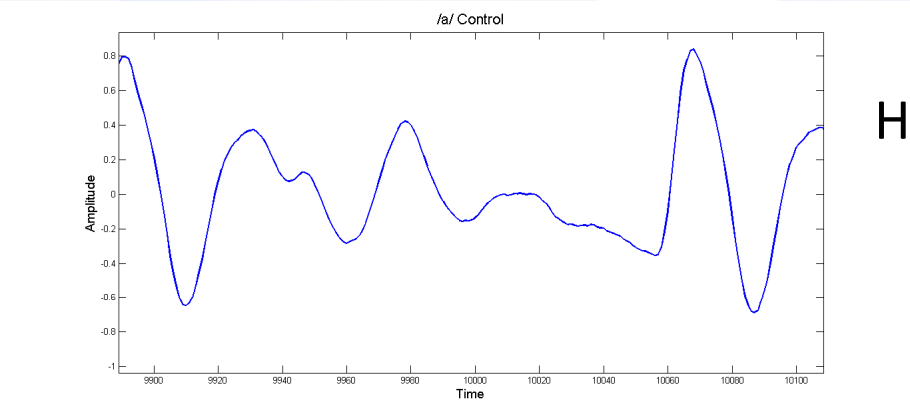


Features used in [4 - 5] to evaluate hypernasality in vowels.



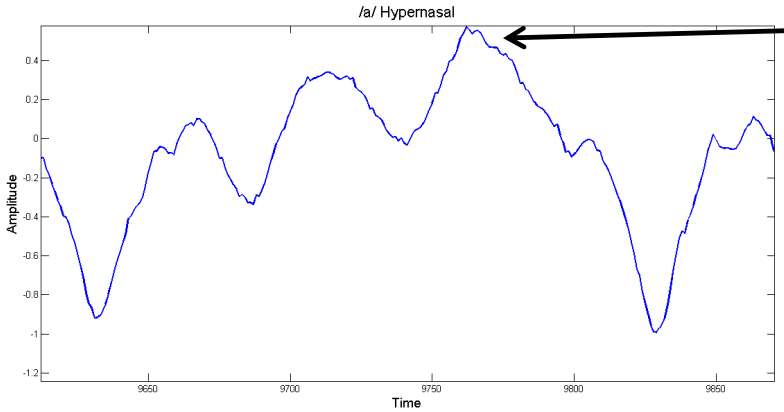
2.2. Characterization

Acoustic, Cesptral and Noise Features



Healthy Voice
/a/

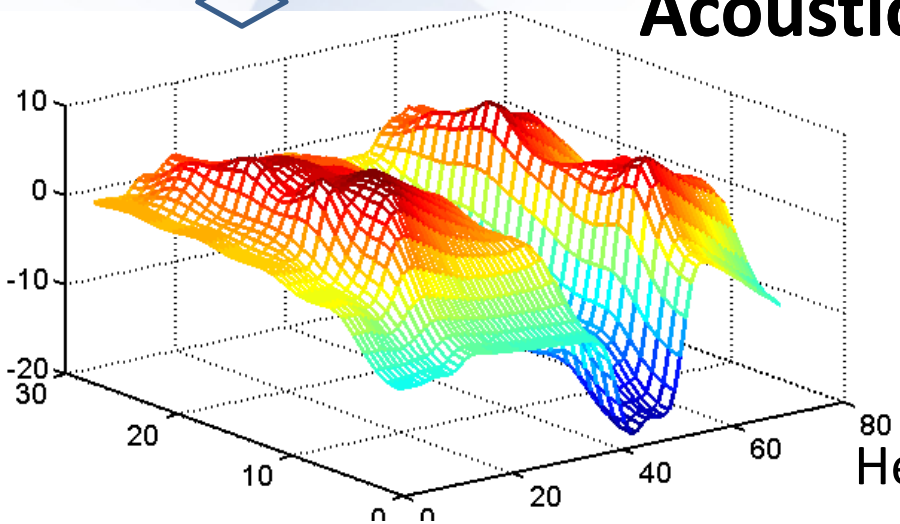
Measure of
Instability in Voice



Hypernasal Voice
/a/

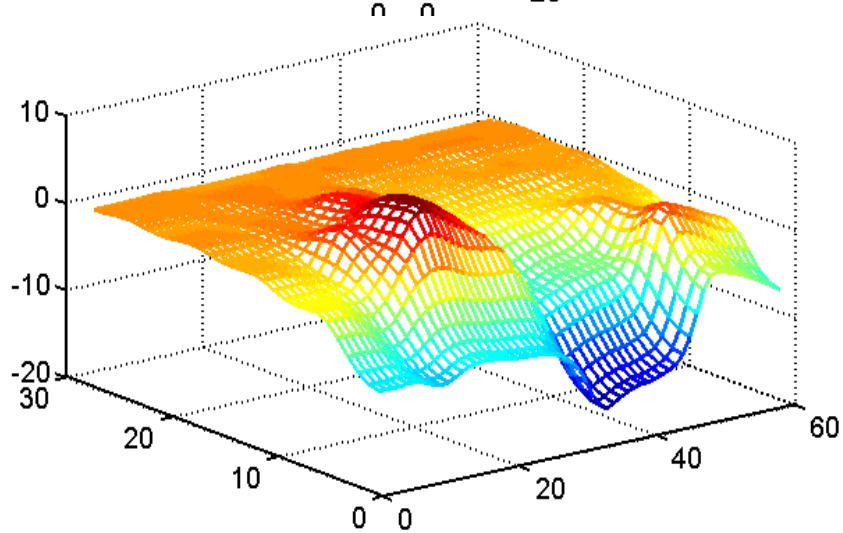
2.2. Characterization

Acoustic, Cesptral and Noise Features



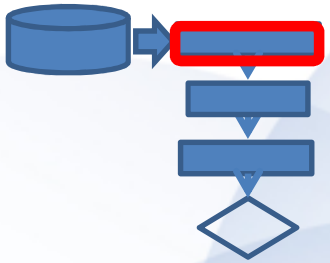
Healthy Voice
/gato/

Energy in more frequency bands of the voice signal



Hypernasal Voice
/gato/

Features used in [6] to evaluate hypernasality in vowels and words.

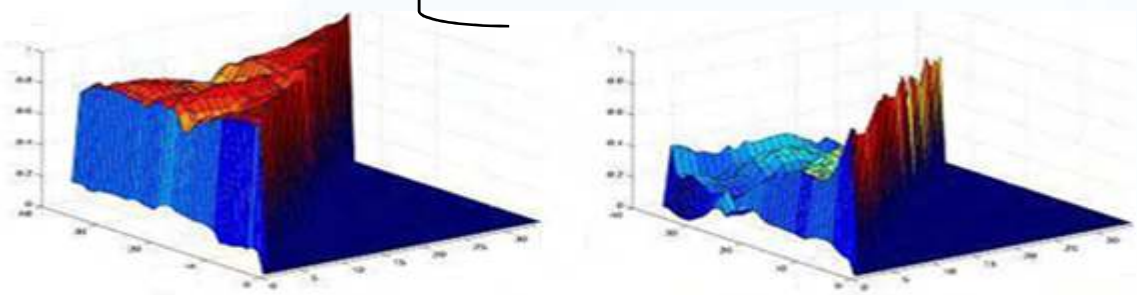


2.2. Characterization

Acoustic, Cesptral and Noise Features

Noise Measures

- HNR – Harmonics to Noise Ratio [7]
- CHNR – Cepstral HNR [8]
- NNE – Normalized Noise Energy [9]
- GNE – Glottal to Noise Excitation Ratio [10]



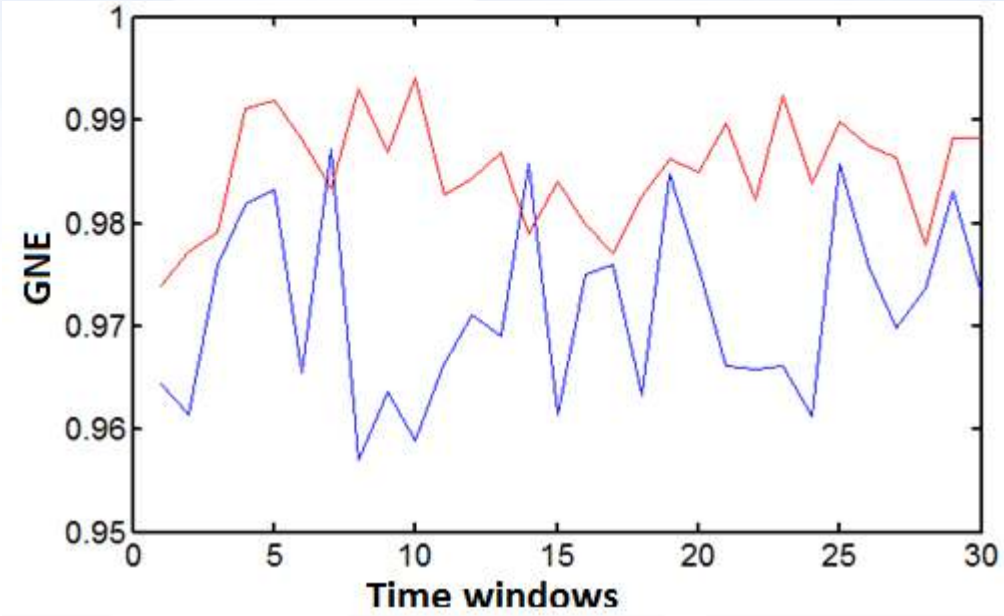
Healthy Voice Pathologic Voice

In [11], the importance of using noise measures to evaluate hypernasality is highlighted.

2.2. Characterization

Acoustic, Cesptral and Noise Features

GNE – Healthy (red) vs Hipernasal (blue)



To measure noise due to velopharyngeal incompetence and compensatory movements on vocal tract (pharynx and vocal tract)

2.2. Characterization

Non Linear Dynamics Features

Human voice production system can be described nonlinearly, so it is a Nonlinear Dynamic System [12 - 14].

A way of analyzing nonlinear dynamic systems are Complexity Measures (Invariants).

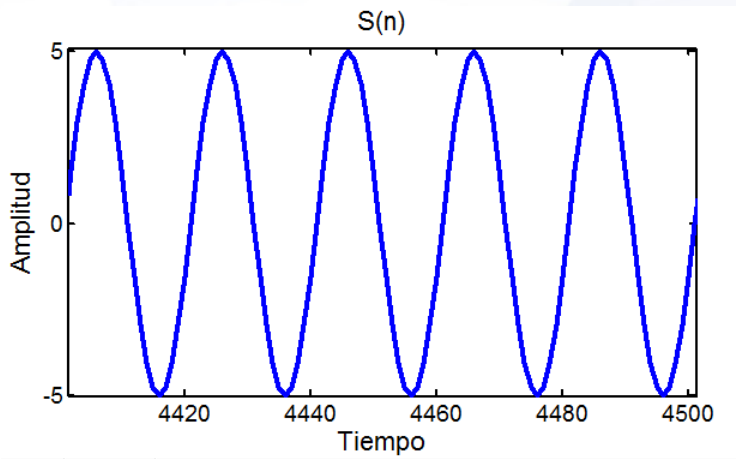
- ❑ Correlation Dimension (CD)
- ❑ Largest Lyapunov Exponent (LLE)
- ❑ Hurst Exponent (H)
- ❑ Lempel-Ziv Complexity (LZC)

2.2. Characterization

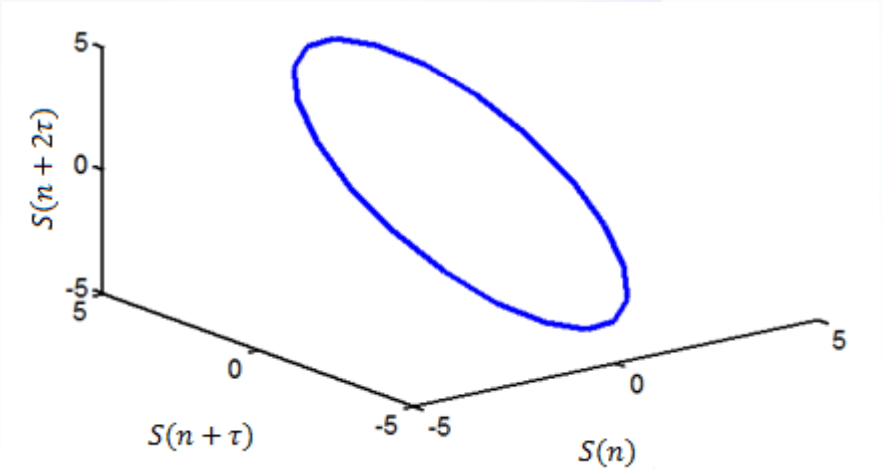
Non Linear Dynamics Features

Embedding space reconstruction – Diffeomorphic and Strange Attractors

Periodic Signal



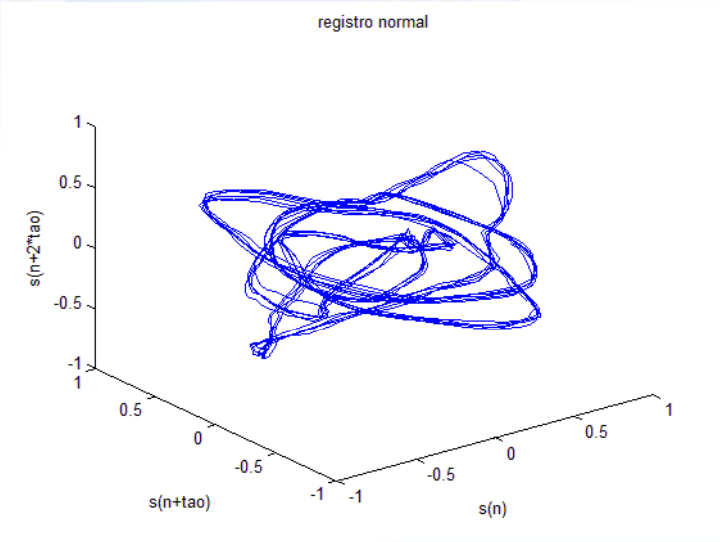
Periodic signal's attractor



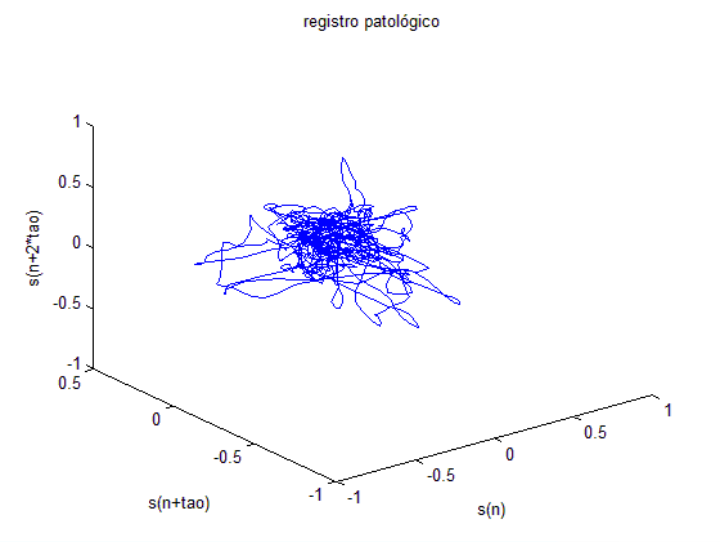
2.2. Characterization

Non Linear Dynamics Features

Taken's Theorem [15]



Healthy Voice



Pathologic Voice

Complexity measures can quantify the grade of pathology in voice

2.2. Characterization

Non Linear Dynamics Features

Correlation Dimension (D_c)

Estimated by Grassberger and Procaccia Method [16]

D_c Gives a measure of how complex is the attractor of a signal.

According to embedding theory, more complex signals will generate irregular attractors, so counting points on consecutive “spheres” is possible to know how irregular and complex is the time series (signal).

Other important information in Correlation Dimension is the dependency between points in the same sphere.

2.2. Characterization
Non Linear Dynamic Features

Largest Lyapunov Exponent (LLE)

Estimated by Rosenstein's Method [17]

Based on Oseledec's Theorem [18]:
Separation rate between points in a phase space trajectory is given by:

$$d(t) = Ce^{\lambda_1 t}$$

λ_1 is LLE , $d(t)$ is the mean divergence in time t and C is a constant used for normalization proposes.

2.2. Characterization
Non Linear Dynamic Features

Hurst Exponent (H)

Generalization of the description of Brownian Movement, based on range scaling method proposed by Hurst [19].

Hurst Exponent allows to estimate long term dependences between points of the signal.

Due to its wide applicability in time series forecasting and complexity measurement, this feature is considered appropriate to identify normal and pathological voices.

2.2. Characterization
Non Linear Dynamic Features

Lempel-Ziv Complexity (LZC)

This feature is wide used to estimate the complexity of a binary series.
Its computation allows to know the number of patterns needed to represent a given sequence [20].

For practical purposes on signal processing, is necessary to assign 0 to when the difference between two consecutive samples is negative and 1 when is positive or null.

2.3. Feature Selection

Sorting features according to its discriminant capacity is necessary to get stable and consistent results, which is reflected in the overall performance of the system.

Two methods were tested for Features Selection.

- ❑Principal Component Analysis (PCA). [21]

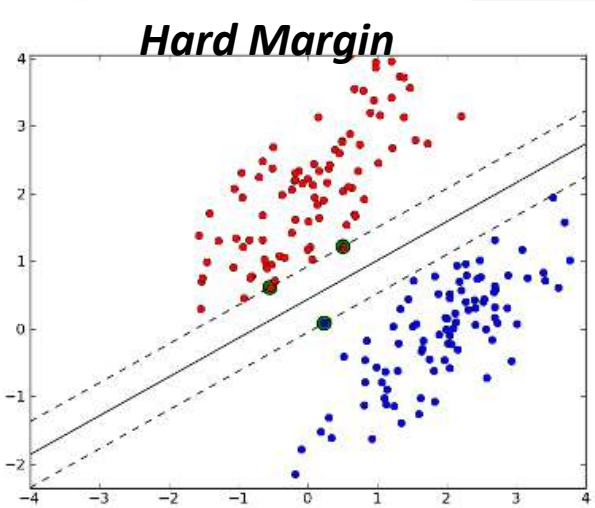
- ❑Sequential Floating Feature Selection (SFFS). [22]

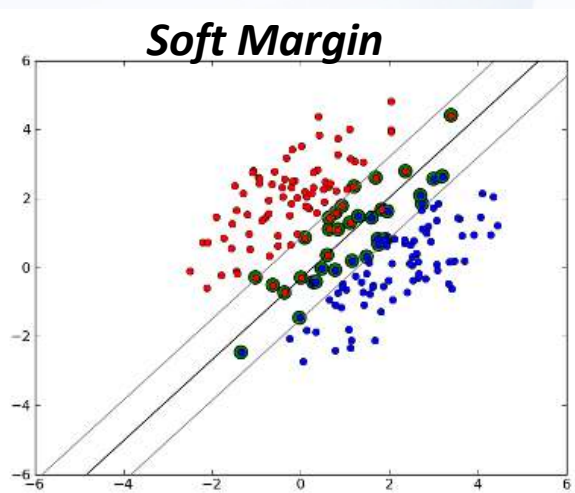
2.3. Classification

- Different Classifiers have been tested:
  - Linear Bayesian
  - Quadratic Bayesian
  - K – Nearest Neigbors
  - Soft Margin - Support Vector Machine (SM-SVM)

2.3. Classification

Soft Margin Support Vector Machine (SM-SVM) is used for deciding whether a register is pathologic or healthy. [23]





3. Experimental Results

EXPERIMENT Nº1: Automatic Detection of Hypernasality in Children by means of AAV

Database: 110 healthy and 156 hypernasal registers. Recorded in low noise conditions and all sampled at 44.100Hz.
Five Spanish Vowels.

Characterization

Feature	Jitter	Shimmer	HNR	CHNR	NNE	GNE	11MFCC
Mean Index	1	2	3	4	5	6	7 to 17
Std. Dev. Index	18	19	20	21	22	23	24 to 34
Variance Index	35	36	37	38	39	40	41 to 51

3. Experimental Results

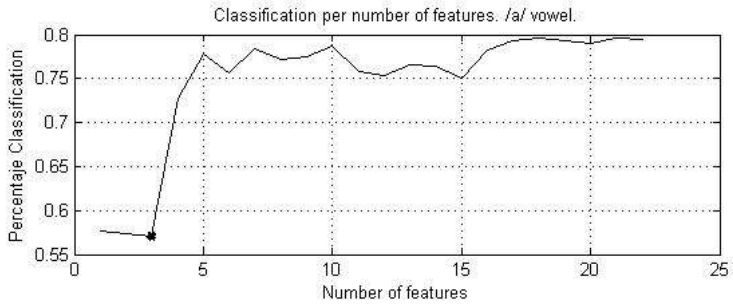
EXPERIMENT Nº1: Automatic Detection of Hypernasality in Children by means of AAV

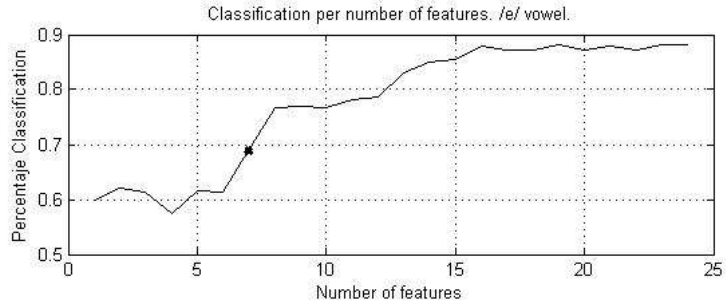
Automatic Feature Selection: Using PCA transformation, features are selected and sorted according to their relevance weight, given by the proper value associated to each proper vector in the representation space.

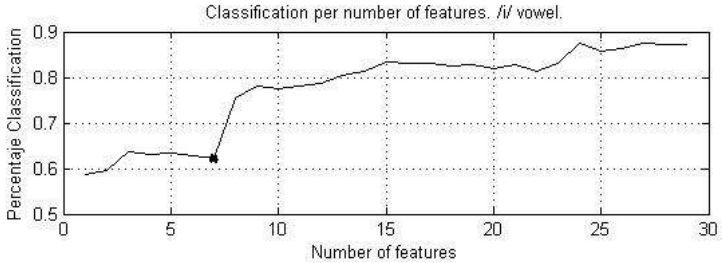
Classification: A Bayesian Classifier is used.
- Linear Bayesian

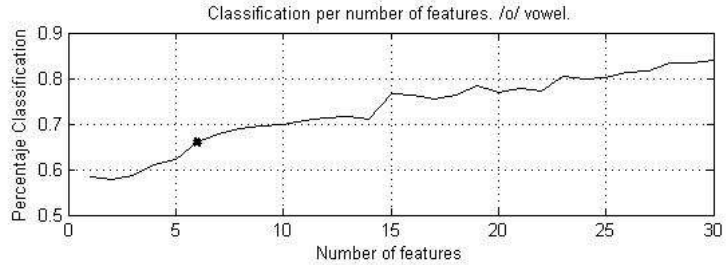
3. Experimental Results

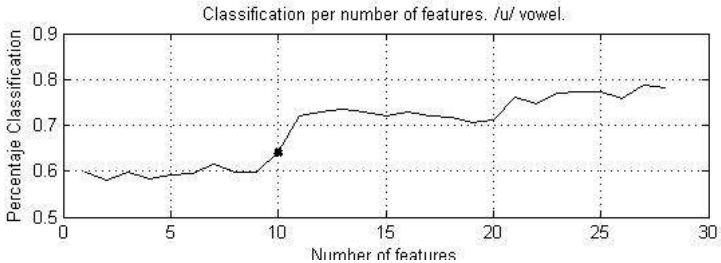
EXPERIMENT Nº1: Automatic Detection of Hypernasality in Children by means of AAV











3. Experimental Results

EXPERIMENT Nº1: Automatic Detection of Hypernasality in Children by means of AAV

Success rates increase up to 20% when acoustic and noise features are considered (NNE, GNE, HNR and CHNR).

VOWEL	Success Rates Before Acoustic and Noise	Success Rates After Acoustic and Noise
/a/	57,06%	79,57%
/e/	68,89%	88,82%
/i/	62,08%	87,49%
/o/	65,95%	84,10%
/u/	64,13%	78,86%

Initial space dimensionality is reduced in 45% after PCA transformation.

3. Experimental Results

EXPERIMENT Nº2: Automatic Selection of Acoustic and Non-linear Dynamics Features in Voice Signals for Hypernasality Detection

Database: 110 healthy and 156 hypernasal registers. Recorded in low noise conditions and all sampled at 44.100Hz.
Five Spanish Vowels.

Characterization – Acoustic Features

Feature	Jitter	Shimmer	HNR	CHNR	NNE	GNE	11MFCC
Mean Index	1	2	3	4	5	6	7 to 17
Std. Dev. Index	18	19	20	21	22	23	24 to 34
Variance Index	35	36	37	38	39	40	41 to 51

3. Experimental Results

EXPERIMENT Nº2: Automatic Selection of Acoustic and Non-linear Dynamics Features in Voice Signals for Hypernasality Detection

Characterization – Nonlinear dynamic Features

Feature	CD	LLE	H	LZC
Mean Index	1	2	3	4
Std. Dev. Index	5	6	7	8

Feature Selection Results: Acoustic feature spaces

Vowel	Initial Dimension	Reduced - PCA	Reduced - SFFS
/a/	51	21	23
/e/	51	23	27
/i/	51	32	18
/o/	51	26	20
/u/	51	28	20
UNION	147	97	99

3. Experimental Results

EXPERIMENT Nº2: Automatic Selection of Acoustic and Non-linear Dynamics Features in Voice Signals for Hypernasality Detection

Nonlinear Dynamic Features Spaces

Vowel	Initial Dimension	Reduced - PCA	Reduced - SFFS
/a/	8	8	3
/e/	8	6	6
/i/	8	2	5
/o/	8	8	5
/u/	8	4	7
UNION	36	23	16

3. Experimental Results

EXPERIMENT Nº2: Automatic Selection of Acoustic and Non-linear Dynamics Features in Voice Signals for Hypernasality Detection

Best AccuracyResults

With Acoustic Features

Vowel	Accuracy – Without Selection (%)	Accuracy – Reduced with PCA (%)	Accuray – Reduced with SFFS (%)
/a/	86.11 ± 8.88	88.28 ±8.38	86.03 ± 6.27
/e/	89.87 ± 10.06	92.45 ± 4.06	93.258 ± 4.54
/i/	87.58 ± 6.27	91.28 ± 5.77	92.45 ± 6.83
/o/	89.87 ± 3.15	89.13 ± 4.86	89.49 ± 9.62
/u/	86.44 ± 6.77	87.22 ± 6.70	89.83 ± 4.74
UNION	93.28 ± 4.12	93.73 ± 5.28	92.86 ± 5.63

3. Experimental Results

EXPERIMENT Nº2: Automatic Selection of Acoustic and Non-linear Dynamics Features in Voice Signals for Hypernasality Detection

Best AccuracyResults

With Nonlinear Dynamic Features

Vowel	Accuracy – Without Selection (%)	Accuracy – Reduced with PCA (%)	Accuray – Reduced with SFFS (%)
/a/	86.78 ± 5.10	86.01 ± 5.73	87.16 ± 4.37
/e/	87.19 ± 6.05	87.96 ± 7.21	87.57 ± 6.56
/i/	87.98 ± 3.40	86.42 ± 8.33	86.86 ± 5.6
/o/	86.48 ± 8.65	86.14 ± 4.31	85.73 ± 5.96
/u/	86.11 ± 6.42	86.15 ± 8.23	86.85 ± 8.20
UNION	91.16 ± 7.24	92.08 ± 8.21	92.05 ± 5.71

3. Experimental Results

EXPERIMENT Nº2: Automatic Selection of Acoustic and Non-linear Dynamics Features in Voice Signals for Hypernasality Detection

Acoustic analysis

Specificity:

$$\frac{TN}{TN+FP}$$

Sensitivity:

$$\frac{TP}{TP+FN}$$

Vowel	SM	NF	Accuracy	Sensitivity	Specificity
/a/	WS	51	86.11±8.88	81.90±8.30	92.58±12.71
	PCA	21	88.28±8.38	90.77±12.06	83.31±10.90
	SFFS	23	86.03±6.27	92.93±6.26	76.48±9.31
/e/	WS	51	89.87±10.06	89.54±12.30	90.46±8.86
	PCA	23	92.45±4.06	95.51±4.46	89.00±8.94
	SFFS	27	93.28±4.54	97.76±5.62	87.17±10.33
/i/	WS	51	87.58±6.27	86.66±6.11	88.85±8.90
	PCA	32	91.28±5.77	99.33±2.11	80.69±9.45
	SFFS	18	92.45±6.83	97.45±4.55	86.35±11.95
/o/	WS	51	89.87±3.15	88.30±4.39	91.35±8.48
	PCA	26	89.13±4.86	88.41±7.98	90.35±10.27
	SFFS	20	89.49±9.62	87.11±14.78	93.82±9.34
/u/	WS	51	86.44±6.77	84.05±7.53	88.66±9.26
	PCA	28	87.22±6.70	95.45±4.64	76.01±12.76
	SFFS	20	89.83±4.74	92.01±7.17	83.25±19.20
Union	WS	147	93.28±4.12	92.99±6.41	93.30±7.72
	PCA	97	93.73±5.28	92.86±9.19	93.71±11.69
	SFFS	99	92.86±5.63	92.28±9.96	94.11±9.96

SM: Selection Method, WS: Without Selection, NF: Number of Features

3. Experimental Results

EXPERIMENT Nº2: Automatic Selection of Acoustic and Non-linear Dynamics Features in Voice Signals for Hypernasality Detection

Non-linear dynamics analysis

Specificity:

$$\frac{TN}{TN+FP}$$

Sensitivity:

$$\frac{TP}{TP+FN}$$

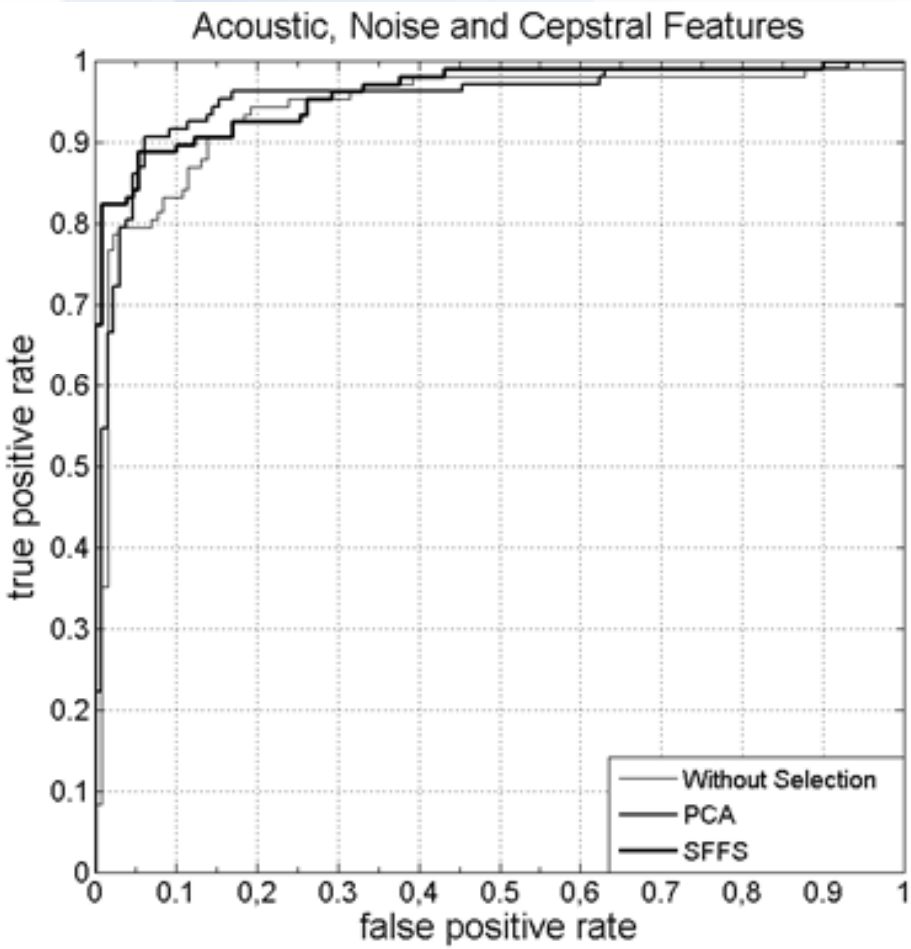
Vowel	SM	NF	Accuracy	Sensitivity	Specificity
/a/	WS	8	86.78±5.10	86.77±7.44	87.90±9.32
	PCA	8	86.01±5.73	83.43±9.06	89.73±8.57
	SFFS	3	87.16±4.37	87.01±6.70	88.57±8.92
/e/	WS	8	87.19±6.05	85.62±9.06	89.86±5.24
	PCA	6	87.96±7.21	86.40±9.13	90.67±6.69
	SFFS	6	87.57±6.56	85.60±9.06	90.86±6.14
/i/	WS	8	87.98±3.40	87.54±8.92	88.98±9.94
	PCA	2	86.42±8.33	84.64±12.32	87.69±7.02
	SFFS	5	86.86±5.62	82.99±8.45	93.09±5.18
/o/	WS	8	86.48±8.65	84.85±12.49	88.71±10.02
	PCA	8	86.14±4.31	86.85±5.27	87.07±11.23
	SFFS	5	85.73±5.96	83.49±8.90	88.34±8.96
/u/	WS	8	86.11±6.42	87.84±6.90	84.42±13.45
	PCA	4	86.15±8.23	87.26±8.40	83.95±14.21
	SFFS	7	86.85±8.20	85.83±9.61	87.84±10.60
Union	WS	36	91.16±7.24	90.84±11.00	91.28±10.52
	PCA	23	92.08±8.21	95.49±7.21	88.05±12.73
	SFFS	16	92.05±5.71	93.58±7.39	90.06±9.78

SM: Selection Method, WS: Without Selection, NF: Number of Features

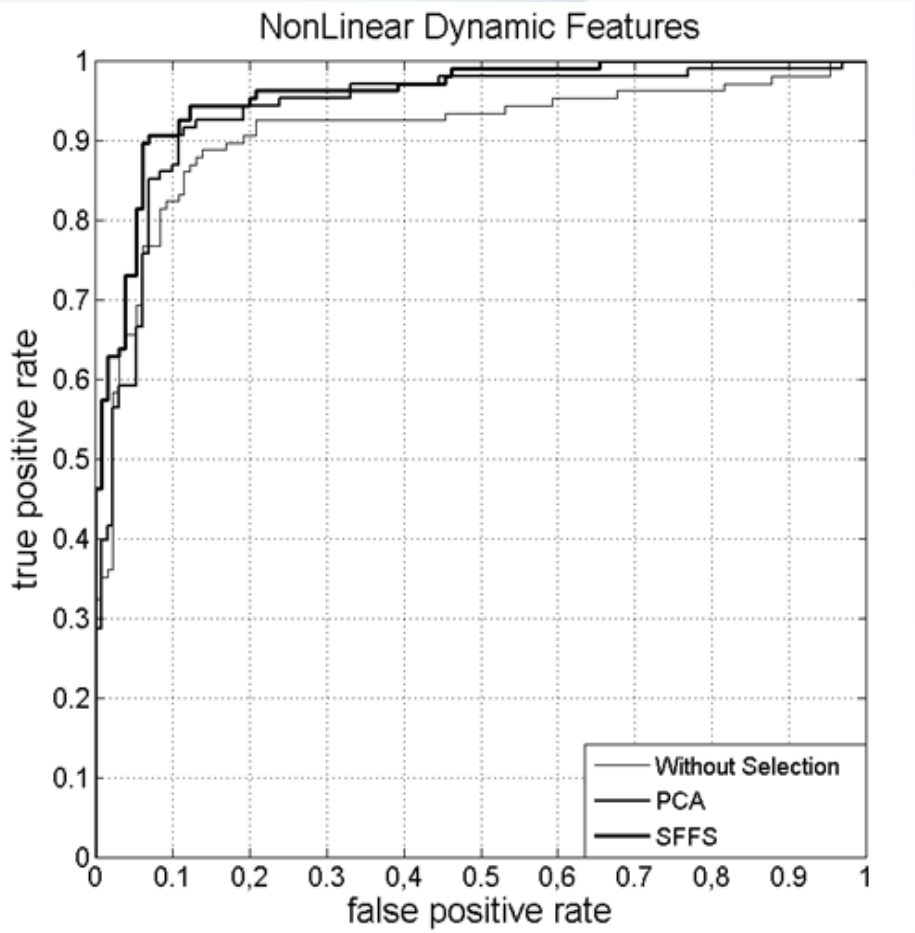
3. Experimental Results

EXPERIMENT Nº2: Automatic Selection of Acoustic and Non-linear Dynamics Features in Voice Signals for Hypernasality Detection

AUC = 0.9616



AUC = 0.9578



3. Experimental Results

EXPERIMENT Nº3: Automatic Selection of Non-linear Dynamics Features for Voice Pathology Detection in Running Speech

Database: A subset of the Kay Elemetrics database.
36 healthy and 36 pathologic registers of people reading "the rainbow passage".

Characterization – Nonlinear dynamic Features

Feature	CD	LLE	H	LZC
Mean Index	1	2	3	4
Std. Dev. Index	5	6	7	8
Skewness	9	10	11	12
Curtosis	13	14	15	16

3. Experimental Results

EXPERIMENT Nº3: Automatic Selection of Non-linear Dynamics Features for Voice Pathology Detection in Running Speech

Segmentation regardless intonation content.

Feature Selection: Using Sequential Floating Feature Selection (SFFS) .

Classification:

-Soft Margin Support Vector Machine (SM-SVM)

- Neural Net (NN)

- K-Nearest Neigbors (K-NN)

With only 6 of those NLD features is possible to detect pathologic voices

Accuracy with SM-SVM	Accuracy with NN	Accuracy with K-NN
95.0% ± 3.54%	87.9.% ± 7.74%	89.15% ± 6.50%

4. Questions

References

[1] K. Verdolini, and L.O. Ramig. “Review: occupational risks for voice problems”. L*ogopedics, Phoniatrics, Vocology*, 26 (1): 37-46.
[2] “Congenital malformations worldwide”, *International Clearinghouse for Birth Defects Monitoring Systems*, Amsterdam, Holland, Tech. Rep. 1991.
[3] A. M. Duque, B. A. Estupiñán and P. E. Puertas. “Labio y paladar fisurados en niños menores de 14 años”. *Colombia Médica*, 33: 108-112, 2002.
[4] Lee, G.S., Wang C.P. and Fu S., ”Evaluation of hypernasality in vowels using voice low tone to high tone ratio'‘, *The Cleft Palate Journal*, 46(1):47--52, 2009.
[5] Wertzner H.F., Schreiber S., and Amaro L., “Analysis of fundamental frequency, jitter, shimmer and vocal intensity in children with phonological disorders”. *Rev. Bras. Otorrinolaringol*, 71(5):582--588, 2005.
[6] Maier, A., Hönig, F., Hacker C., Shuster M. and Nöth E., “Automatic Evaluation of Characteristic Speech Disorders in Children with Cleft Lip and Palate'',
Proc. Of 11th INTERSPEECH, Brisbane, Australia, pp. 1757--1760, 2008.
[7] E. Yumoto, W. J. Gould and T. Baer. “Harmonics to Noise Ratio as hoarseness index of degree of hoarseness”. *Journal of the Acoustical Society of America*. (71):6. 1982
[8] P. Murphy and O. Akande. “Cepstrum-based Harmonics to Noise Ratio Measurement in voiced speech”. *Lecture Notes in Computer Science*. Springer-Verlag, Berlin, pp. 199-218, 2005.

References

[9] H. Kasuya, S. Ogawa and Y. Kikuchi. “An adaptive comb filtering method as applied to acoustic analysis of pathological voice”. *Proc. of the ICASSP*, 1986.
[10] D. Michaelis, T. Gramss and H.W. Strube. “Glottal to Noise Excitation Ratio – a new measure for describing pathological voices”. *Acustica/Acta*. Vol. 83, pp. 700-706. 1997.
[11] Setsuko, I., “Effects of Breathy Voice Source on Ratings of Hypernasality”, *The Cleft Palate Journal*, 42(6):641--648, 2005.
[12] Giovanni, A., Ouaknine, M., Guelfucci, R., Yu, T., Zanaret, M. and Triglia, J.M., “Nonlinear behavior of vocal fold vibration: the role of coupling between the vocal folds”, *Journal of Voice*, 13(4):456--476, 1999.
[13] Little, M.A., McSharry, P.E., Roberts, S.J., Costello, D.E. and Moroz, I.M., “Exploiting nonlinear recurrence and fractal scaling properties for voice disorder detection”, *Biomedical Engineering Online*, 6(23), 2007.
[14] Ghazaleh Vaziri, Farshad Almasganj, Roozbeh Behroozmand “Pathological assessment of patients’ speech signals using nonlinear dynamical analysis”. Journal of Computers in biology and medicine, (4):1, pp.54-63, 2010.
[15] F. Takens, "Detecting strange attractors in turbulence," *Dynamical Systems and Turbulence*, Lecture Notes in Mathematics, Springer, 1981, pp. 366-381.

References

[16] P. Grassberger and I. Procaccia, “Measuring the strangeness of strange attractors”, *Physica D*, Vol. 9, pp. 189-208.
[17] M. Rosenstein, J. Collins and C. De Luca "A practical method for calculating largest Lyapunov exponents from small data sets“ *Physica*, vol. 65, pp. 117-134. 1993.
[18] V. A. Oseledec, “A multiplicative ergodic theorem. Lyapunov characteristic numbers for dynamical systems” *Transactions of Moscow Mathematic Society*.
Vol. 19, pp. 197–231, 1968.
[19] H. E. Hurst, R. P. Black and Y. M. Simaika, “Long-term storage: an experimental study” London, 1965.
[20] F. Kaspar and H. G. Shuster, “Easily calculable measure for complexity of spatiotemporal patterns” *Physical Review A,* (36):2, pp. 842-848, 1987.
[21] I. T. Jolliffe, “Principal Component Analysis”, 2nd ed., Springer.
Series in Statistics. New York, NY, USA, 2002.
[22] P. Pudil, J. Novovicova and J. Kittler,”Floating Search Methods in Feature Selection” *Pattern Recognition Letters*, 15(11):1119--1125, 1994.
[23] Scholk¨opf, B. and Smola, A.J., *Learning with Kernels, The MIT* Press, 2002.

ANEX I – CD estimation

Correlation Dimension (CD)

Estimated by Grassberger and Procaccia Method [15]

$$D_c = \lim_{\varepsilon \to 0} \left(\frac{\log(C_m(\varepsilon))}{\log(\varepsilon)} \right)$$

Where: $C_m(\varepsilon) = \lim_{N \to \infty} \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} \Theta\left(\varepsilon - \left\| x_i - x_j \right\|\right)$

And: $\Theta(z) = \begin{cases} 0 & \text{for} \quad z \leq 0 \\ 1 & \text{for} \quad z > 0 \end{cases}$ is the Heaviside function.

N Is the number of points in the signal

$\left\| x_i - x_j \right\|$ Is the distance between points i and j.

ANEX I – CD estimation

D_c Gives an idea of how complex is the signal.

According to embedding theory, more complex signals will generate irregular attractors, so counting points on consecutive “spheres” is possible to know how irregular and complex is the time series (signal).

Other important information in Correlation Dimension is the dependency between points in the same sphere.

ANEX II – LLE estimation

Largest Lyapunov Exponent (LLE)

Estimated by Rosenstein's Method [16]

Based on Oseledec's Theorem [17]:
Separation rate between points in a phase space trajectory is given by:

$$d(t) = Ce^{\lambda_1 t}$$

λ_1 is LLE , $d(t)$ is the mean divergence in time t and C is a constant used for normalization proposes.

ANEX II – LLE estimation

LLE estimates the divergence rate of points in the states space (attractor).

This property is important considering that periodic signals generate closed attractors while non-periodic generate irregulars.

As vocal fold movement in healthy humans is quasi-periodic, and hypernasal people have problems with their vocal phonation due to its velo-pharingeal incompetence, their voice signals are non-periodic, thus LLE can be considered as a good estimator for automatic detection of hypernasality in voice.

ANEX III – H estimation

Hurst Exponent (H)

Generalization of the description of Brownian Movement, based on range scaling method proposed by Hurst [18].

Einstein's proof shows that distance traveled by a particle is proportional to the square root of the time:

$$R = cT^{0.5}$$

Where R is the range for particle movement, c is a constant and T is the time.

ANEX III – H estimation

For the generalization proposed by Hurst, signal must be transformed considering its accumulated deviations respect to the mean:

$$X(i) = \sum_{n \in M_i} \left(x(n) - \bar{x}(n)_{M_i} \right)$$

Where M_i is the $i-th$ segment of the signal $x(n)$, which contains M points, and $\bar{x}(n)$ is the mean of the set of points in this segment.

Hurst's proof generalizes the expression found by Einstein:

$$\frac{R}{S} = cT^H$$

Where R is the variation range of the signal evaluated for each segment M and is expressed as:

$$R(M) = \max_i \{X(i)\} - \min_i \{X(i)\}$$

ANEX III – H estimation

$S = \sigma$ is the standard deviation of the signal and c is a scaling constant.

Thus H is calculated as the slope of the straight line formed in the curve $\frac{R(M)}{\sigma} vs M$, when is plotted in logarithmic scale, i.e.

$$(cM)^H = \frac{R(M)}{\sigma}$$

Then,

$$H = \frac{\log\left(\frac{R(M)}{\sigma}\right)}{\log(cM)}$$

ANEX III – H estimation

Hurst Exponent allows to estimate long term dependences between points of the signal.

Due to its wide applicability in time series forecasting and complexity measurement, this feature is considered appropriate to identify normal and pathological voices.

ANEX IV – LZC estimation

Lempel-Ziv Complexity (LZC)

This feature is wide used to estimate the complexity of a binary series**.**
Its computation allows to know the number of patterns needed to represent a given sequence [19].

For practical purposes on signal processing, is necessary to assign 0 to when the difference between two consecutive samples is negative and 1 when is positive or null.

ANEX IV – LZC estimation

LZ estimation is based on the reconstruction of a sequence by X copying and insertion of symbols inside a new series.

Consider a sequence $X = x_1 x_2 \ldots x_n$ which shall be analyzed from left to right, take the first bin in the binary chain and insert it by default as starting point.

Define S as a variable that holds bits inserted until the moment, so at the beginning S only has x_1.

Define Q as a variable that accumulates every bit that is analyzed from left to right inside binary train.

ANEX IV – LZC estimation

On each iteration S and Q are joined to form SQ.

$SQ\pi$ Is the resulting sequence after remove last digit in SQ

When $Q \notin v(SQ\pi)$ the process of bits insertion is done.

Complexity c will be the number of sub sets in which original sequence is divided.

ANEX IV – LZC estimation (example)

Set $X = 0001001$

1. 0 is the first bit inserted, thus is represented as 0*. Where * indicates that there finish a block of bits and a new one must be start.

2. $S = 0$, and $Q = 0$, is the second bit in the sequence X. So, $SQ = 00$ and $SQ\pi = 0$
   Note that $Q \in v(SQ\pi)$ where $v(SQ\pi)$ denotes the vocabulary of the set $SQ\pi$ Thus, the second block is not finished yet.

3. $S = 0$ because the bits insertion process have not finished. $Q = 00$ is the sequence found continuing with the analysis of the bits train.
   $SQ\pi = 00$, and $Q \in v(SQ\pi)$ so the block of bits that is in process will be 0*00.

4. $S = 0$, $Q = 001$, $SQ = 0001$ and $SQ\pi = 000$. Note that $Q \notin v(SQ\pi)$, so here the second block is finished, ie: 0*001*

ANEX IV – LZC estimation (example)

LEMPEL ZIV COMPLEXITY. (An ilustrative example).

5. $S = 0001, Q = 0$, $SQ = 00010$ and $SQ\pi = 0001$. But $Q \in v(SQ\pi)$, so 0*001*0

6. $S = 0001$, $Q = 00$, $SQ = 000100$ and $SQ\pi = 00010$, $Q \in v(SQ\pi)$, thus 0*001*00

7. $S = 0001$, $Q = 001$, $SQ = 0001001$ and $SQ\pi = 000100$. Since $Q \in v(SQ\pi)$, then here the division of the sequence X is not finished.
0*001*001…

As three blocks (patterns) were necessary to represent the sequence, the complexity is $c = 3.$

ANEX V – ROC curves

ROC Curves Construction

TP: # of patterns of class 0 correctly classified as class 0.
FN: # of patterns of class 0 missclassified as class 1.
FP: # of patterns of class 1 missclassified as class 0.
TN: # of patterns of class 1 correctly classified as class 1.

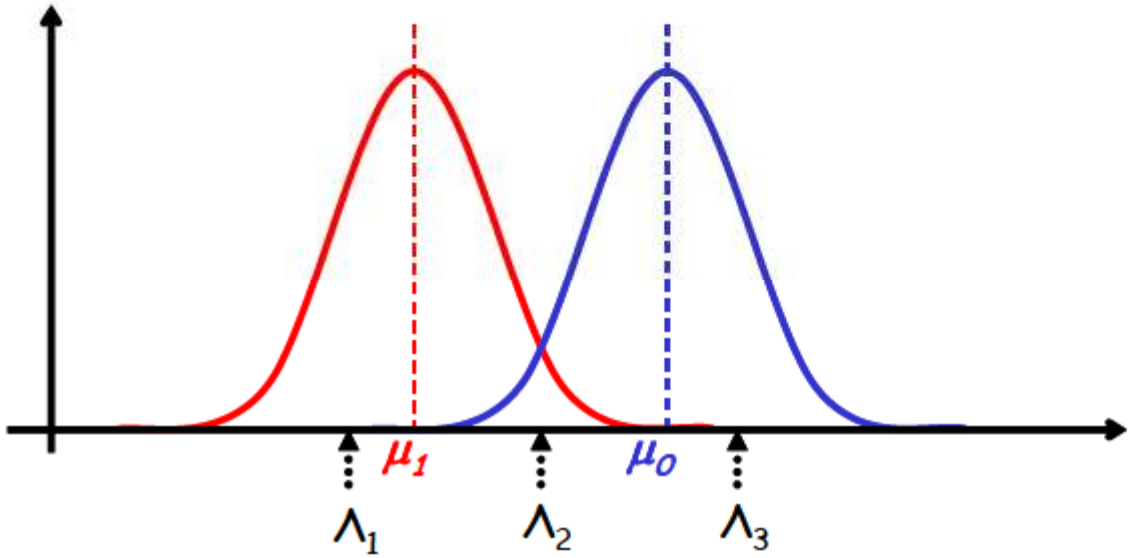
Confusion Matrix

		True Class	
		Class 0	**Class 1**
Estimated Class	**Class 0**	TP	FP
	Class 1	FN	TN

Specificity: $\frac{TN}{TN+FP}$

Sensitivity: $\frac{TP}{TP+FN}$

ANEX V – ROC curves



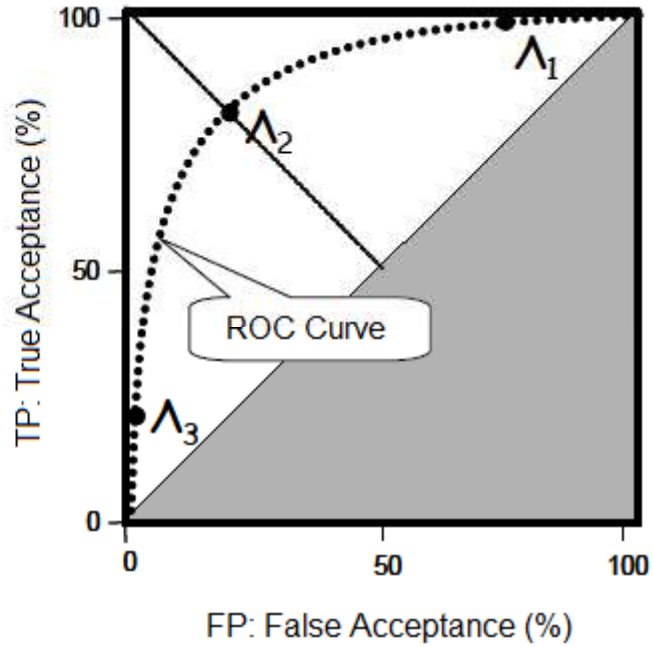


Figure taken from: “Contribuciones Metodológicas para la evaluación objetiva de patologías laríngeas” Cap. 5.

$$x(n) \rightarrow [x_1, x_2, \ldots, x_L] \rightarrow \mathbf{P} = [p_1, p_2, \ldots, p_L]$$

$$\frac{|p_{i+1} - p_i|}{\overline{\mathbf{P}}} \times 100\% = j_i$$

$$i = 1, 2, \ldots, L-1$$

$$\mathbf{Jitt} = [j_1, j_2, \ldots, j_{L-1}] \rightarrow \left\{ \begin{array}{l} \overline{\mathbf{Jitt}} \\ STD(\mathbf{Jitt}) \\ VAR(\mathbf{Jitt}) \end{array} \right.$$

$$x(n) \rightarrow [x_1, x_2, \ldots, x_L] \rightarrow \mathbf{A} = [a_1, a_2, \ldots, a_L]$$

$$\frac{|a_{i+1} - a_i|}{\overline{\mathbf{A}}} \times 100\% = s_i$$

$$i = 1, 2, \ldots, L-1$$

$$\mathbf{Shim} = [s_1, s_2, \ldots, s_{L-1}]$$

$$\left.\begin{array}{l} \overline{\mathbf{Shim}} \\ STD(\mathbf{Shim}) \\ VAR(\mathbf{Shim}) \end{array}\right\}$$

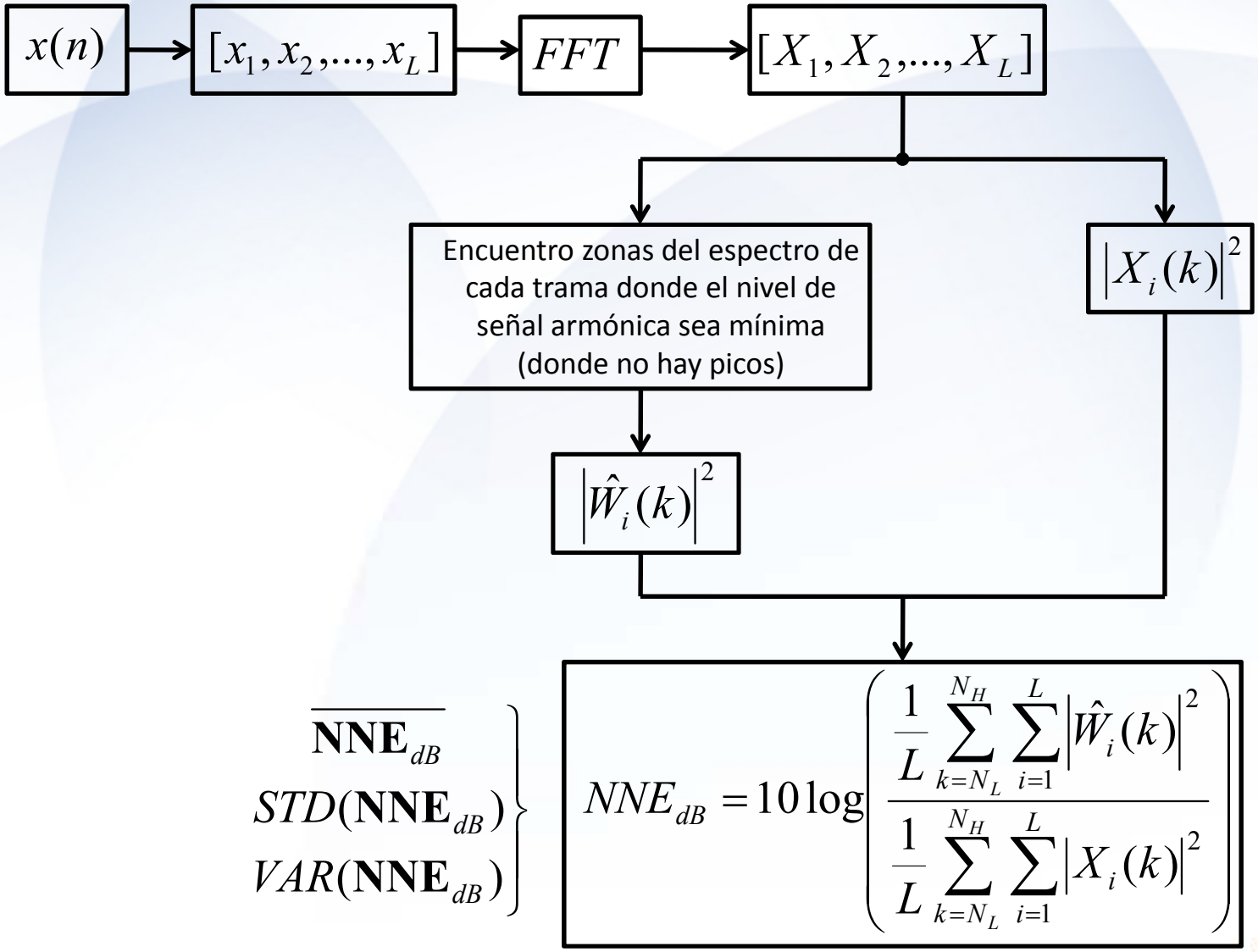
$$x(n) \rightarrow [x_1, x_2, \ldots, x_N] \rightarrow x_A = \sum_{i=1}^{N} \frac{x_i}{N} \rightarrow \mathbf{H} = N \sum_{n=0}^{T} x_A^2(n)$$

$$\downarrow$$

$$\mathbf{N} = \sum_{i=1}^{N} \sum_{n=0}^{T} \{x_i(n) - x_A(n)\}^2$$

$$\downarrow$$

$$HNR_{dB} = 10 \log\left(\frac{\mathbf{H}}{\mathbf{N}}\right) \rightarrow \left\{ \begin{array}{l} \overline{\mathbf{HNR}_{dB}} \\ STD(\mathbf{HNR}_{dB}) \\ VAR(\mathbf{HNR}_{dB}) \end{array} \right.$$



$$\left.\begin{array}{c}\overline{\mathbf{NNE}_{dB}}\\ STD(\mathbf{NNE}_{dB})\\ VAR(\mathbf{NNE}_{dB})\end{array}\right\}$$

$$NNE_{dB} = 10\log\left(\frac{\frac{1}{L}\sum_{k=N_L}^{N_H}\sum_{i=1}^{L}\left|\hat{W}_i(k)\right|^2}{\frac{1}{L}\sum_{k=N_L}^{N_H}\sum_{i=1}^{L}\left|X_i(k)\right|^2}\right)$$

$$x(n) \rightarrow [x_1, x_2, \ldots, x_N] \rightarrow w_i \times x_i = \hat{x}_i \rightarrow C_{\hat{x}_i}(n) = IFFT\left[\log\left|FFT(\hat{x}_i)\right|\right]$$

Ubicación y filtrado de rahmónicos

FFT

N_i

$$H_i = \log\left|FFT(\hat{x}_i)\right| - N_i$$

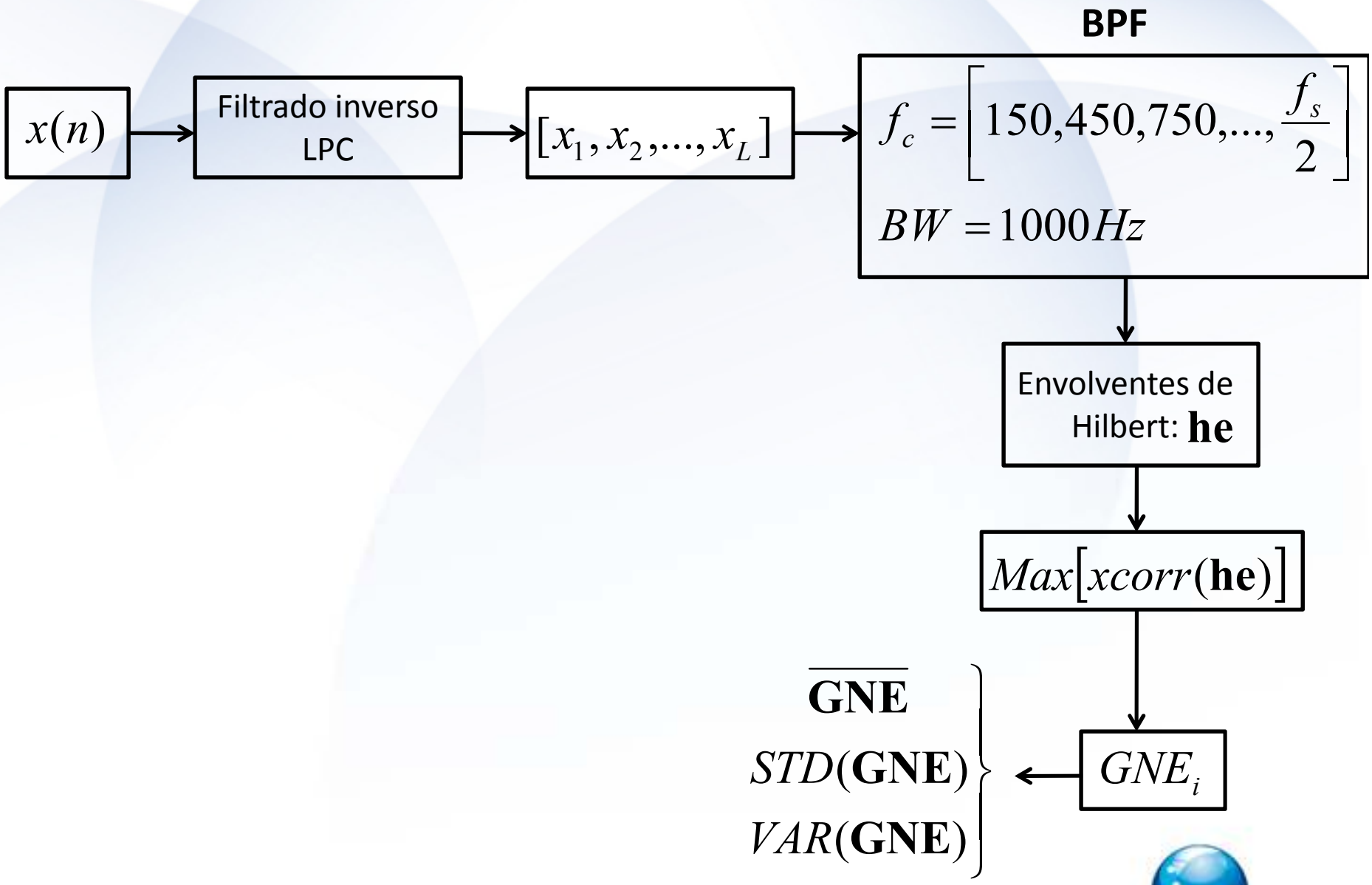
Corrección del nivel de ruido

Se ubican los mínimos entre armónicos suscesivos de H_i para formar $\mathbf{m}_{\mathbf{H_i}}$

$$Ruido_i = N_i - \mathbf{m}_{\mathbf{H_i}}$$

$$Señal_i = \left|FFT(\hat{x}_i)\right|$$

$$CHNR_{dB} = 10\log\left(\frac{Ruido_i}{Señal_i}\right) \begin{cases} \overline{\mathbf{CHNR}_{dB}} \\ STD(\mathbf{CHNR}_{dB}) \\ VAR(\mathbf{CHNR}_{dB}) \end{cases}$$


BPF
$x(n)$
Filtrado inverso LPC
$[x_1, x_2, \ldots, x_L]$
$f_c = \left[150, 450, 750, \ldots, \frac{f_s}{2}\right]$
$BW = 1000Hz$
Envolventes de Hilbert: **he**
$Max[xcorr(\mathbf{he})]$
GNE_i
$\overline{\mathbf{GNE}}$
$STD(\mathbf{GNE})$
$VAR(\mathbf{GNE})$

$$x(n) \rightarrow [x_1, x_2, \ldots, x_L] \rightarrow \hat{x}_i(n) = x_i(n) w_i(n) \rightarrow FFT$$

$w_i(n)$: Ventana

$$F(mel) = 2595 \log\left(\frac{f}{700} + 1\right)$$

$$\hat{X}_i(k) H(k) \rightarrow DCT \rightarrow \mathbf{MFCC}_i$$

$$\left.\begin{array}{c} \overline{\mathbf{MFCC}} \\ STD(\mathbf{MFCC}) \\ VAR(\mathbf{GNE}) \end{array}\right\} \mathbf{MFCC}_i$$