Studies in Engineering and Exact Sciences
ISSN: 2764-0981

Verification and identification of the speech signal using algebraic analysis and DWT

Verificação e identificação do sinal de fala usando análise algébrica e DWT

DOI: 10.54021/seesv5n1-017

Recebimento dos originais: 01/04/2024
Aceitação para publicação: 01/26/2024

Amraoui Abdelkader
PhD in Electronic Engineering
Institution: Ziane Achour University of Djelfa
Address: City Daia 33/74, Djelfa, Algéria
E-mail: kader2717@yahoo.fr

Saadi Slami
PhD in Electronic Engineering
Institution: Ziane Achour University of Djelfa
Address: City 5 Juillet, Djelfa, Algéria
E-mail: saadisdz@gmail.com

ABSTRACT
We presented a flowchart, of a new approach that relies on the mathematical operations of matrices and to verify and identifies the speaker, this is a new method difficult to detect by attackers, and this is due to the difficulty to determine its base matrix used in the speaker verification and identification process, Although there are many artificial intelligence programs that can imitate voices and generating false conversations, which poses security and hacking problems.

Keywords: verification, identification, LPC, matrix, DWT, speech signal.

RESUMO
Apresentamos um fluxograma de uma nova abordagem que se baseia nas operações matemáticas de matrizes para verificar e identificar o locutor. Esse é um novo método difícil de ser detectado por invasores, e isso se deve à dificuldade de determinar a matriz de base usada no processo de verificação e identificação do locutor, embora existam muitos programas de inteligência artificial que podem imitar vozes e gerar conversas falsas, o que apresenta problemas de segurança e de hacking.

Palavras-chave: verificação, identificação, LPC, matriz, DWT, sinal de fala.

1 INTRODUCTION

Since speech is a highly personal and sensitive form of communication, it is essential to prevent unauthorized access to the vocal signal or its content. Many different systems have been developed in terms of methods used to extract features and recognize speech, and it is divided into two parts: speech recognition and speaker verification.

The objective of this research is to find new ways and methods to identify and verify the speaker, and also to diversify the protection of the speech signal against different signal attacks, using matrices as a means of protection and identification, with new calculations and measurement of the signal. In the 1980s, many companies developed voice recognition systems. And in the 1990s, companies that worked on voice, such as IBM, developed a system for the commercial market. According to the American company – International Biométrique Groupes (IBG), this biometric technology held 4.3% of the market in 2000 [1]. The EURASIP magazine, According to their primary objectives, covered a subject on information security in 2006, which highlighted the dynamism of the field.

1.1 PREVIOUS RESEACH AND STUDIES

Numerous algorithms have been developed for extracting acoustic features from speech signals, which constitute a key element in detection and speaker identification. The fundamental frequency, f0, is one of the most critical of these characteristics. Our new approach, which we presented in our " A Novel Approach for Speaker Gender Identification and Verification using DWT First Level Energy and Zero Crossing [2].

We can claim that our research is both easier and more secure by a higher percentage than 93.75% when it comes to comparing and identifying the speaker signal, compared to all previous research. Our research is founded on the algebraic analysis of a fundamental algorithm for estimating speech signal frequencies.

2 STUDY AND DISCUSSION

2.1 PRIMARY SIGNAL PROCESSING

The highest frequency of the speech signal in humans is 5 khz, since the process of converting the signal requires twice its frequency, therefore the speech signal requires frequency sampling at least twice that, this is a say 11025hz, which is one of the values known in the processes applied to the speech signal either for the recording of audio signals or other applications on the speech signal.

2.2 FILTER THE SIGNAL

In this experiment, we apply a Butterworth filter of order n=8 to the speech signal with a cutoff frequency of fc=600 hz to remove high frequencies associated with noise.

2.3 APPLY THE DISCRETE WAVELET TRANSFORM (DWT)

After completing the preprocessing and filtering of the speech signal, we apply the discrete wavelet transform (DWT) at three levels to the resulting signal using equations (1) and (2) to determine both the approximate coefficients and the detailed coefficients.which we presented in our as documented [3]

$$a_1 = \sum_{k=-\infty}^{\infty} X(k) g[2n-k] \quad (1)$$

$$d_1 = \sum_{k=-\infty}^{\infty} X(k) h[2n-k] \quad (2)$$

a_1 : approximate third level coefficients of DWT.
d_1 : detailed third level coefficients of DWT.
g_n: low pass filter
h_n : high pass filter

2.4 CUT THE SIGNAL WITH FRAMES, AND WINDOWING

Instead of taking the entire signal, we divide the speech signal (the approximate third-level coefficients) into a set of frames. Each frame is composed of a group of samples that express the characteristics of that frame, and each frame has a duration of 30 milliseconds and a length of 330 samples.

Nbrs of samples = 11025 hz * 30 ms =11025hz * 0.030 s = 330,75 of samples == 330 of samples (3)

2.5 AUTOCORRELATION

The idea of using the autocorrelation function is to determine how similar two successive samples of a signal are.The autocorrelation function of an ergodic and stationary signal on a frame is the case from a window of our speech signal.The result of the auto-correlation is a symmetric vector of length 2N with a maximum in the middle. Then we limit the calculation of the maximum in a single part of the autocorrelation signal, either right or left part.

2.6 ANALYZES THE LPC

The objective of LPC is to provide a limited number of linear prediction coefficients and prediction error energy to directly represent the speech signal approaching the speech system.

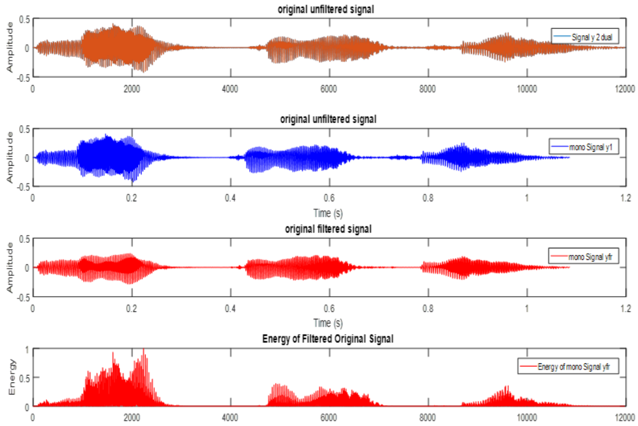
2.7 CHARACTERISTIC EXTRACTION BY LPCC ALGORITHMS

LPCC (Linear Predictive Cepstral Coefficients) ,They are generally used as features for speech recognition and speaker identification systems, that is to say that LPCC is obtained by the derivation from the LPC coefficients [4].

2.8 COMPARISON OF DATA

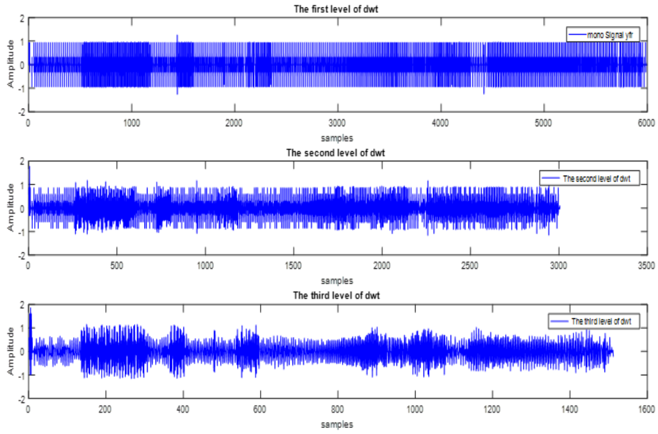
To determine the speaker recognition rate, we calculate the two metrics, which correspond to FAR (False Acceptance Rate) and FRR (False Rejection Rate). These metrics involve elements from two square matrices: the database matrix and the speaker's speech signal matrix, which need to be compared.

Figure 1 – The signal before and after filtering with its energy



Source: Authors.

Figure 2 – The three signal DWT levels



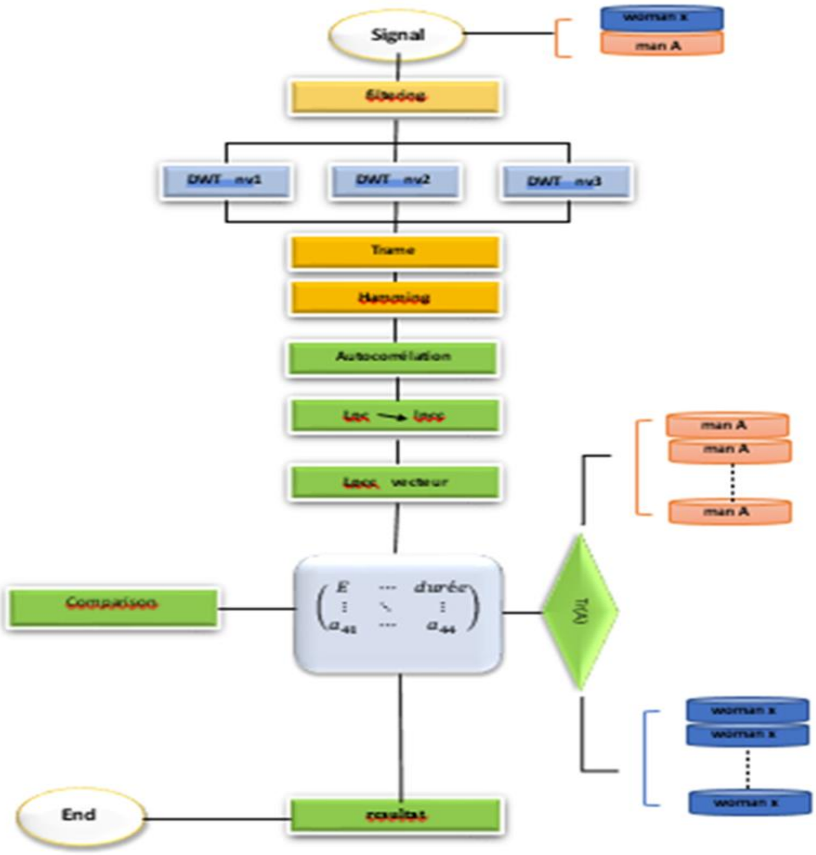
Source: Authors.

2.8.1 Figure: Organization Chart

Below, we propose a scheme for speech signal processing and speaker recognition and verification using a new approach based on the algebraic analysis and matrix calculations and matrix traceability.

STUDIES PUBLICAÇÕES
Studies in Engineering and Exact Sciences
ISSN: 2764-0981

Figure 1 – The proposed scheme for processing, identifying, and verifying a speaker's speech signal



Source: Authors.

Table 1 – Table comparing the original signal with the duplicate signal and the imitated signal

	Women B701	Women B702	Women B703	Women D001	Women B302 Imitated voice	Level of DWT
Nbrs of Sample	11978	11971	11976	11973	11975	Signal original
Energy	530	530	530	300	160	//
Zero Crossing Numbre	867	867	867	735	752	//
Zero Crossing Rate	0.07	0.07	0.07	0.06	0.06	//
Coefficient 2 of LPC	1	1	1	0.4	0.5	Level 1 (DWT)
Coefficient 3 of LPC	0.3	0.3	0.3	0	0	//
Coefficient 4 of LPC	0	0	0	0	0	//
Coefficient 5 of LPC	0	0	0	0	0	//
Coefficient 2 of LPC	1.5	1.4	1.4	1.8	1.7	Level 2 (DWT)
Coefficient 3 of LPC	1	1	1	2	2	//
Coefficient 4 of LPC	1	1	1	2	2	//
Coefficient 5 of LPC	1	1	1	2	2	//
Coefficient 2 of LPC	4	3	4	3	4	Level 3 (DWT)
Coefficient 3 of LPC	9	5	9	6	9	//
Coefficient 4 of LPC	14	6	14	9	14	//

Coefficient 5 of LPC	20.36	7.06	20.56	12.01	19.5	//
Speech Matching rate	100%	87.5%	93.7%	12.5%	43.75%	

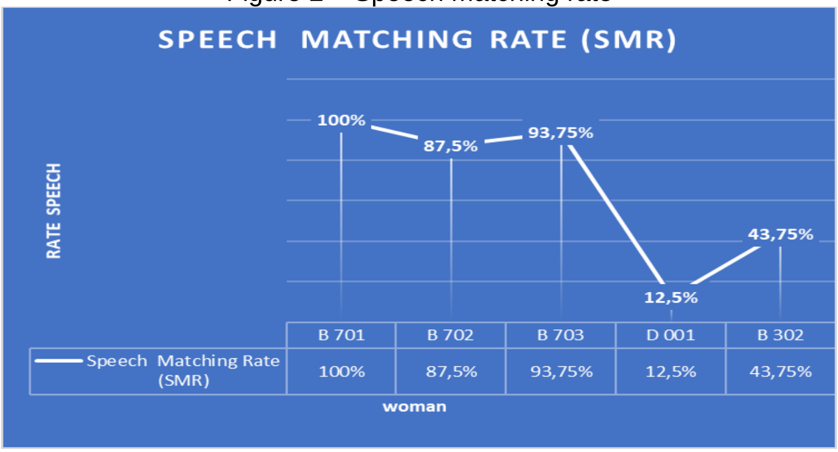
Source: Authors.

The comparison table (1) presents the results obtained after applying these metrics to the same sentence, " NICE TO MEET YOU " which has been used consistently in all stages of the analysis. These results are for speakers of both genders across all three levels of discrete wavelet transform (DWT).

4 DISCUSSION AND CONCLUSION

To verify the speaker's identity, we refer to our stored database by comparing the matrix generated in previous steps ,This matrix is composed of characteristics extracted from the audio signal using the LPCC standard and the approximate coefficients of the discontinuous wavelet transform (DWT) for the three levels. Subsequently, we compare this matrix with the list of matrices in our stored database. This is to authenticate, confirm, and verify the speaker, as shown in the flow chart in Figure (2).

Figure 2 – Speech matching rate



Source: Authors.

In the initial phase, we carefully selected the mother wavelet and identified the most suitable branch for preserving the signal characteristics at the third level of the discrete wavelet transform (DWT), Specifically, we opted for branch (bior3.9), belonging to the orthogonal mother wavelet. This decision was informed by from a prior study (4) that highlighted this particular branch as the most effective in retaining the characteristics of the third level in the discrete wavelet transform ,Then we created a matrix with dimension (4*4) whose elements in the first row consist of the characteristics of the original signal, which are the energy, the number and zero-crossing rate, and the number of samples of this signal, respectively, As for the elements of lines (2) to (4), they are, respectively, the first four coefficients of the spectral coefficients for the linear prediction of the signal segment after applying the Hemming windows and the discrete wavelet transform (DWT) for the three levels of this signal.

REFERENCES

CHASSÉ, M. La biométrie au Québec: Les enjeux. document d'analyse. **Commission d'Accès à l' Information du Québec**, v. 89, Jul 2002. Retrieved from https://www.ethique.gouv.qc.ca/media/xsybl3kk/biometrie-reflexion-fr.pdf

ABDELKADER, A.; SALAMI, S. A Novel Approach for Speaker Gender Identification and Verification using DWT First Level Energy and Zero Crossing. **Engineering, Technology & Applied Science Research**, v. 12, n. 06, p. 9570-9578, 2022. doi: https://doi.org/10.48084/etasr.5269

ABDELKADER, A.; SAADI, S. Conseil Supérieur de la Langue Arabe. **Revue Science et Technologie**, n. 06, p. 11-23, 2023. Available at: https://hcla.dz/media/magazines/pdf/6ST.pdf

ALMAADEED, N.; AGGOUN, A.; & ABBES, A. (2016). **Text-Independent Speaker Identification Using Vowel Formants**. p. 345–356. doi: 10.1007/s11265-015-1005-5