Sound and Voice Verification and Identification A Brief Review of Töeplitz Approach

Khalid Saeed

Computer Engineering Department, Faculty of Computer Science,
Bialystok Technical University, Wiejska 45A, 15 351 Bialystok, Poland
ksaeed@wi.pb.edu.pl

Abstract. The main objective of the paper is to show how a new voice recognition approach is used for person recognition by identifying their voice. The tool used in this approach is a mathematical model based mainly on Töeplitz Forms and Burg's estimation models. After filtering, the recorded voice signal is processed to get its power spectrum estimation. The feature vector is derived from the power spectrum and its adjacent plots and Töeplitz matrices. This vector has proved to furnish a unique unrepeated print – each individual has his/her own voice print. In the work presentation, the author introduces some experiments performed in MATLAB to show how this voiceprint looks like and how it is verified for recognition. The basic idea is derived from applying Töeplitz matrix minimal eigenvalues algorithm to Burg's estimating model. This implies a graphical approach for feature extraction, selection and hence signal-image description confronting the conventional and traditional methods.

1 Introduction

The importance of sound verification has found and proved its practical application and significant role in many areas of our life and the nature around us. Apart from human voice verification for person identification, this research field has played an essential role in detecting many kinds of sounds and their sources. Example areas of some important key aspects are explosions, crashes, screams of people or animals, footsteps and many other sound-creating events and natural concepts. Identification of sound, and particularly human voice, has its noticeable effect when used as additional biometric data in a recognition system. Examples for such an application are the audio-video security approaches and their use in intelligent network surveillance systems. In their success rate, all systems of voice or sound verification and identification depend on how far the success of preprocessing stages is and if the stages preceding the classification are precise and accurate.

© Václav Snášel (Ed.): Znalosti 2008, pp. 22–27, ISBN 978-80-227-2827-0.
FIIT STU Bratislava, Ústav informatiky a softvérového inžinierstva, 2008.

Recently, the devices used in preprocessing are available (and cheap) and almost standard. However, still more precise and fast methods for signal feature extraction and signal description are desired. In this work, the author presents a new technique to cover both the preprocessing and data representation for voice classification and recognition with emphasis rather on the voice signal representation for the purpose of classification and recognition. A new Töeplitz Forms based algorithm has been developed to use the unique characteristics of the matrix minimal eigenvalues and their specific behavior that is suitable for image description, feature extraction and recognition [1]. It discusses the possibility of treating the speech signal graphically in order to extract the essential image features as a basic step for data mining applications in the biometric techniques. The experiments preliminarily performed according to the suggested algorithm are on the human-voice signal [2], [3], [4]. Figure 1 shows the position of Töeplitz-based signal descriptor in a standard voice signal recognition model from preprocessing to classification.

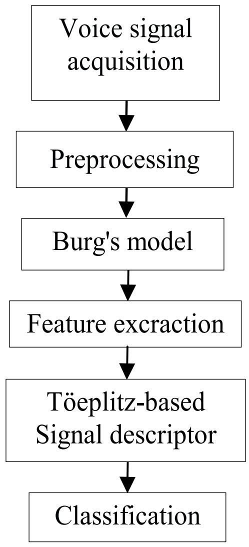

Fig. 1. Töeplitz-based signal descriptor in a voice signal recognition system.

2 Signal Estimation and Description

The suggested frequency spectral estimation (Burg's model [4], [5], [6]) is derived from *linear predictive coding principle*. The task of Burg's curves obtained when performing in MATLAB is to ensure accurate basic preprocessing for feature extraction. When performing together with Töeplitz matrix minimal eigenvalue-based approach, the model has proved the possibility of precise selection of signal features from the power spectral plot. This is achieved in a similar manner to the application of Töeplitz matrix minimal eigenvalues on images of written texts, signature, palm-print, face geometry or finger-prints, the topics that have shown a success rate of about 98% in general cases [7], [8], [9], [10], [11], [12]. The success rate was particularly high when the performing was in hybrid systems [13] by both Toeplitz matrices and Projection approach [14]. The extracted feature-carrying image comprises the elements of Töeplitz matrices to compute their minimal eigenvalues and present a set of feature vectors within a class of voices. Fig. 2 shows the plot of frequency spectral estimation as suggested by Burg when performing in MATLAB.

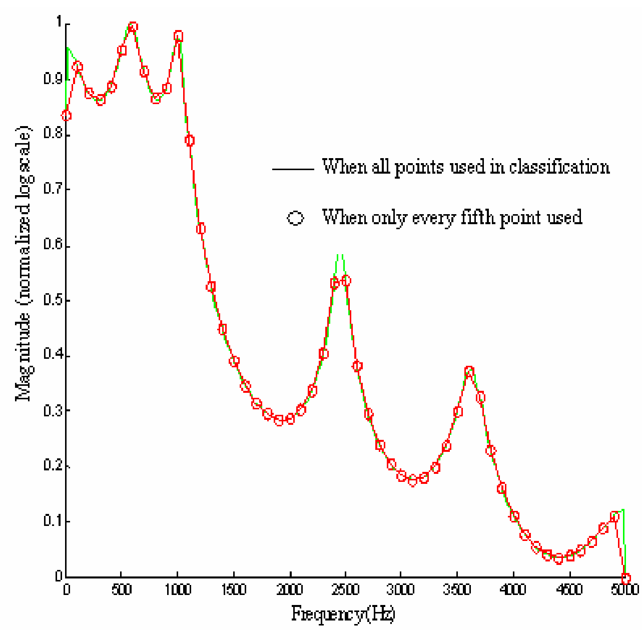

Fig. 2. Frequency spectral estimation – Burg's model. Computations are in MATLAB.

Equation (1) shows the definition of Töeplitz matrix elements C's in their modified form [12].

$$C=\begin{bmatrix} c_0 & c_1 & c_2 & \cdots & c_{n-1} \\ c_1 & c_0 & c_1 & \ddots & \vdots \\ c_2 & c_1 & c_0 & \ddots & c_2 \\ \vdots & \ddots & \ddots & \ddots & c_1 \\ c_{n-1} & \cdots & c_2 & c_1 & c_0 \end{bmatrix} \qquad (1)$$

The C's are the coefficients of a series developed from the relations between the characteristic points in Burg's model of Fig. 2. The series is of Taylor's type [1], [12], [15]. The details of Töeplitz forms and their minimal eigenvalues behavior are beyond the structure of this summarized work, but are given in other related works, see for example [1] and [15].

Figure 3, however, introduces the behavior of the minimal eigenvalues λ of the submatrices in (1) for three different voices. It can be seen how the three plots (the series of λ) in the figure differ in their behavior from one voice-descriptor to another.

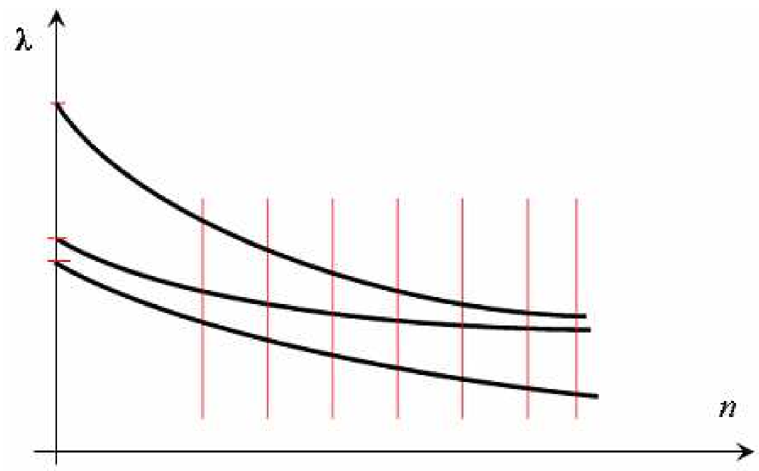

Fig. 4. Töeplitz minimal eigenvalues series as descriptors for three different voices. Notice how the three series of λ differ in their behavior from one voice-descriptor to another – the starting points, their successive values in given intervals and their series limit for higher values of n.

The required computations for both Burg's model and Töeplitz approach were performed in MATLAB proving voice-signal image recognition in an easy-to-use for classification way. This comes from the fact that the presented problem solution and its Matlab implementation do not require to implement any special hardware and can be used in tandem with other biometric technologies, like face recognition, in hybrid systems for multi-factor verification.

3 Conclusions

Voice identification in terms of both speech and speaker authentication has its unique significance and role in almost all known biometric techniques. The reason is simply that people seek for an easy-to-use, reliable and safe system to show and prove their authenticity whenever needed. Each biometric type of technology demands a user to play their active part in the process of being identified and authenticated. "*As known,* *it is obvious that almost all biometric systems involve user's writing on paper, stamping their fingerprints, showing their open eye to a camera, pressing their hand to show its geometry, and the most common way of identifying the user by their recalling a code or a password to enter the identifying machine with*." Voice biometric methods of human identification, however, need nothing but the human utterance to obtain their *voiceprint*, the term used in most voice biometric solutions as a template of human unique voice features displayed when entering the identifying system. The brief review given in this paper manifests a new approach to the voiceprint as a means for simple data description. The current and future applications of this approach will be on general sound print verification for people and animals or any other important sound causing beings or events in our life and environment.

Acknowledgements

I would like to express my deep gratitude to Professor Vaclav Snasel, VSB-Technical University of Ostrava, Czech Republic, for his kind invitation to introduce the details of this abstract-paper as a keynote at the International Conference Znalosti 2008.

The work was supported by the Rector of Bialystok Technical University (grant number W/WI/10/07) and was partially presented *but not printed* as a keynote at ICEEDT'07 - International Conference on Electrical Engineering Design and Technologies ICEEDT Conference, Hammamet, Tunisia, 2007.

References (*Further Readings*)

1. Saeed K., Nammous M. K., "A Speech-and-Speaker Identification System: Feature Extraction, Description, and Classification of Speech-Signal Image," IEEE Trans. on Industrial Electronics, vol. 54, no. 2 (2007) 887-897
2. Saeed K.: A Note on Biometrics and Voice Print: Voice-Signal Feature Selection and Extraction – A Burg-Töeplitz Approach. Proc. 10th IEEE Workshop on Signal Processing-SP'2006, Poznan, Poland (2006) 7-12
3. Nammous M., Saeed K.: Voice-Print and Text-Independent Speaker Identification. International Conference on Electrical Engineering Design and Technologies - ICEEDT'08, Tunisia (2007). *Proceedings available on CD.*
4. Saeed K. and Kozłowski M.. An Image-Based System for Spoken-Letter Recognition. 10th Int. Conference CAIP'03 on Computer Analysis of Images and Patterns, August 2003, Groningen. Proceedings published in: Lecture Notes in Computer Science, Petkov and Westenberg (Eds.), LNCS 2756, Springer-Verlag Heidelberg: Berlin (2003) 494-502
   Available: http://aragorn.pb.bialystok.pl/~zspinfo/arts/KSaeedCAIP'03.pdf
5. Saeed K., Nammous M. K.. Heuristic Method of Arabic Speech Recognition. 7th International Conf. on Digital Signal Processing and its Applications – IEEE-DSPA'05, Tom II, Moscow (2005) 528-530
   Available: http://aragorn.pb.bialystok.pl/~zspinfo/arts/KSaeedDSP'05.pdf

6. Saeed K., Nammous M. K., "A New Step in Arabic Speech Identification: Spoken Digit Recognition," Information Processing and Security Systems, K. Saeed and J. Pejas (Eds), Springer Science+Business Media, New York (2005) 55-66
Available: http://aragorn.pb.bialystok.pl/~zspinfo/arts/KSaeedSpringer.pdf
7. Saeed K., Tabędzki M., "Intelligent Feature Extract System for Cursive-Script Recognition," 4th IEEE International Workshop on Soft Computing as Transdisciplinary Science and Technology (WSTST'05), Springer Heidelberg, Muroran, Japan (2005) 192-201
Available: http://aragorn.pb.bialystok.pl/~zspinfo/arts/2005%20WSTST.pdf
8. Adamski M., Saeed K.: Signature image Recognition by Shape context Image matching. Journal of Medical Informatics and Technologies, vol. 11 (2007) 89-95
9. Saeed K., Werdoni M.: A New Approach for Hand-Palm Recognition. In: Pejaś J. and Piegat A. (Eds), Enhancement Methods in Computer Security – Biometric and Artificial Intelligence Systems, Springer Science + Business Media, New York (2004) 185-194
10. Saeed K.: Minimal-Eigenvalue-Based Face Feature Descriptor. In: Dramiński M., Grzegorzewski P., Trojanowski K., Zadrożny S. (Eds), Issues in Intelligent Systems Models and Techniques. Institute of System Research, Polish Academy of Sciences, Akademicka Oficyna Wydawnicza EXIT, Warsaw, Poland (2005) 185-196
11. Saeed K., Leoniuk K.: A Method for Fingerprint FeatureExtraction. Proc. 13th Intern. MultiConf. on Advanced Computer Systems – ACS-AIBITS/CISIM'06, Miedzyzdroje, Poland (2006) 31-42
12. Saeed K., "Image Analysis for Object Recognition," Publications of Bialystok Technical University, Poland (2004)
13. Saeed K., Tabędzki M.: Hybrid System for Letter Recognition. Intern. 3rd Conf. on Neural Networks and Artificial Intelligence - ICNNAI'03, Minsk, Belarus (2003) 44-48
14. Saeed K.: A Projection Approach for Arabic Handwritten Characters Recognition. Proc. of ISCI - International Symposium on Computational Intelligence, Aug. 31 – Sep. 1, New Trends and App. in Comp. Intelligence, Springer-Verlag, Kosice, Slovakia (2000) 106-111
15. Saeed K.: Computer Graphics and Analysis: A Method for Arbitrary Image Shape Description. Intern. Journal of Machine Graphics and Vision - MGV, vol. 10, no. 2, Institute of Computer Science, Polish Academy of Sciences, Warsaw, Poland (2001) 185-194