Using Wave Equation to Extract Digital Signal Features

Ayman Al-Rawashdeh
Department of Electric Power Engineering
Faculty of Engineering Technology
Al-Balqa Applied University
Amman, Jordan
dr.ayman.rawashdeh@bau.edu.jo

Ziad Al-Qadi
Department of Computer Engineering
Faculty of Engineering Technology
Al-Balqa Applied University
Amman, Jordan
dr.ziad.alqadi@bau.edu.jo

***Abstract*—Voice signals are one of the most popular data types. They are used in various applications like security systems. In the current study a method based on wave equation was proposed, implemented and tested. This method was used for correct feature array generation. The feature array can be used as a key to identify the voice signal without any dependence on the voice signal type or size. Results indicated that the proposed method can produce a unique feature array for each voice signal. They also showed that the proposed method can be faster than other feature extraction methods.**

Keywords-wave signal; wave equation; feature array; voice parameters

I. INTRODUCTION

Digital signals (DS) are one of the most important types of data currently used in different engineering applications. Examples of these applications include audio and speech processing and compact disc (CD) generation, target detection, position and velocity determination, radar and sonar signal processing, spectral estimation, image processing in addition to biomedical signal processing. To obtain these digital signals, special transducers are employed. The output of the transducers is subjected to sampling and quantization in order to obtain the required DS. Figure 1 illustrates the process of capturing the analogue voice signal which is then converted to a digital signal using a digital convertor connected to the output of the loudspeaker.

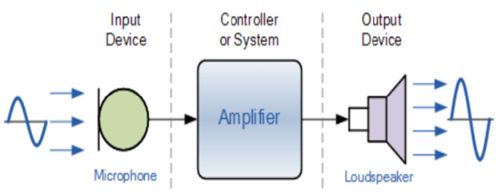


Fig. 1. Obtaining analogue version of the DS

A. Voice Signal

Audio digital sound signal (DSS) refers to a digital representation of the audio waveform for digital signal processing, storing and transmitting. Digital sound can be obtained from analogue audio signal by applying sampling and quantization using the audio file as an input and the digital audio file as an output (Figure 2). When analogue sound waves are stored in a digital form, each digital audio file can be decomposed into a series of samples [1-3].

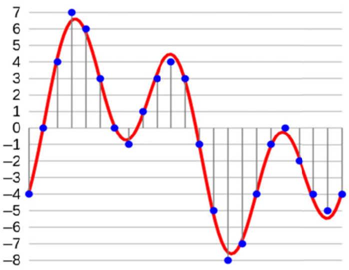
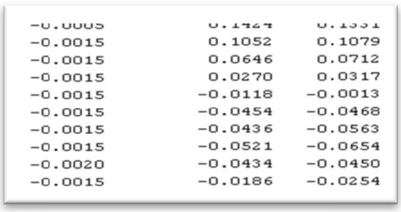


Fig. 2. Sampling and quantization to produce digital signal

There are two types of DSS, mono and stereo signals. Mono signal is a one-channel signal and its amplitude values are arranged in one column array. Stereo file is a two-channel signal and its values are arranged in a two column array as shown in Figure 3.



-0.0005	0.1424	0.1331
-0.0015	0.1052	0.1079
-0.0015	0.0646	0.0712
-0.0015	0.0270	0.0317
-0.0015	-0.0118	-0.0013
-0.0015	-0.0454	-0.0468
-0.0015	-0.0436	-0.0563
-0.0015	-0.0521	-0.0654
-0.0020	-0.0434	-0.0450
-0.0015	-0.0186	-0.0254

(a) Mono signals (b) Stereo signals

Fig. 3. First 10 samples of (a) S8 and (b) cow wave files

B. Voice Signal Analysis and Voice Signal Parameters

The voice parameters mentioned in [4], can be easily calculated to form a voice key or signature. The obtained voice key or signature is then used to identify or retrieve the voice signal. The set of parameters includes:

- Estimated Mu of the population [4].

- Estimated sigma [5].
- Peak factor (crest factor (CF)) [5].
- Dynamic range [6-8].
- Power spectral density (PSD) [4, 9].

A Matlab code was written to calculate the above mentioned parameters [4] for different voice signals and the implementation results are shown in Table I. The code used was:

```
%computing sigma, mu, peak fact, dynamic range
 % get a section of the sound file
%[x,fs] = waveread('ziadyes.wav', [20000 65001]);
[x,fs] = audioread('s8.wav');
xxx=x;
% Sigma and Mu estimation
[u s] = normfit(x);
sigma = num2str(s);
zs1=sigma
mu = num2str(u);
zs2=mu
% computing of the peak (crest) factor
rms = sqrt(mean(x.^2));
x = abs(x);
peak = max(x);
Q = 20*log10(peak/rms);
Qstr = num2str(Q);
zs3=Qstr
% computing of the dynamic range
maxval = peak;
x(x==0)=[];
minval = min(x);
D = 20*log10(maxval/minval);
Dstr = num2str(D);
zs4=Dstr
%calculating power spectral density
Fs = 32e3;
t = 0:1/Fs:2.96;
x = audioread('dog.wav');
Pxx = periodogram(x);
Hpsd = dspdata.psd(Pxx,'Fs',Fs); % Create a PSD data object.
psd1=mean(Hpsd.data);zs5=psd1
```

Table I, shows the implementation results.

II. METHODOLOGY

A. Wave Equation

The wave equation [10, 11] takes the form shown in (1):

$$\frac{\partial^2 u(x,t)}{dt^2} = c^2 \frac{\partial^2 u(x,t)}{dx^2} \qquad (1)$$

where $c = \sqrt{E/\rho}$ is the velocity. Usually a discrete equation is considered in the form of the equation with finite difference of second order as shown in (2):

$$\frac{\partial^2 u_n(t)}{dt^2} = \frac{c^2}{h^2}\left(u_{n+1}(t) - 2u_n(t) + u_{n-1}(t)\right) \qquad (2)$$

If c=1 and h=1, then (3) can be applied:

$$\frac{\partial^2 u_n(t)}{dt^2} = \left(u_{n+1}(t) - 2u_n(t) + u_{n-1}(t)\right) \qquad (3)$$

Equation (3) can be solved by applying convolution between the array x=[1 -2 1] and the voice signal (Laplace operator).

B. Method

The proposed method is based on the wave equation and the following algorithm shows the sequence of operations to be implemented to generate voice features.

1. Get the original voice signal data matrix (one column matrix for mono voice signal, or two column matrix for stereo voice signal).
2. Reshape the voice signal data to one row array.
3. Apply the convolution between Laplace operator and the row array.
4. Check each value in the convolution results.
5. If the value is greater than zero, add 1 to a local minimums count.
6. If the value is equal to zero, add 1 to a stable count.
7. If the value is less than zero, add 1 to a local maximums count.
8. Save the 3 counts as a features array for the voice signal.

III. IMPLEMENTATION AND EXPERIMENTAL RESULTS

A Matlab code was written to implement the proposed method using several experiments.

```
clc
close all
clear all
[x,fs] = audioread('cow.wav');
[s1 s2]=size(x)
tic
zz=reshape(x,s1*s2,1);
dd=[1 -2 1];
zz1=filter2(dd,zz);
ff1=sum([zz1<0]);
ff0=sum([zz1==0]);
ff2=sum([zz1>0]);
toc
```

In experiment 1, the word "zero" spoken by a person was taken and its wave file was then recorded using the same sampling frequency (12500), but with a different number of samples. Table II, shows the implementation results of experiment 1. Figure 4 shows (a) one of the voice signals wave and (b) the obtained second derivative of this signal. In the experiment 2, different animal voices with different sizes and different sampling frequencies were taken and manipulated. Results of this implementation are shown in Table III. From the results shown in Tables I-III, the following are noted:

- Each voice file was obtained from a unique feature array, and thus, such a unique array can be easily used to identify the voice signal. Each obtained feature for each voice file is unique, thus each feature array can be used as a key to identify the voice signal.

- Any change (even very small) will be reflected on the corresponding feature array, and consequently will produce another new feature array.
- Figure 4 clearly indicates that the proposed method can be used for voice signal smoothing.
- Finally, the proposed method can be used to generate a key for any type of voice signal (mono or stereo) with any size and any sampling frequency.

The third experiment was carried out in order to compare the efficiency between the proposed method and the voice signal parameters. Different wave files were taken and processed by both methods. Feature array generation time was calculated using CPU processing time. Table IV, shows the implementation results of experiment 3. The results strongly reveal that the proposed method can reduce the feature extraction processing at least 28 times and accordingly can be considered to be more efficient than the voice signal parameters methods described in the literature [4-5].

TABLE I. CALCULATED VOICE SIGNAL PARAMETERS

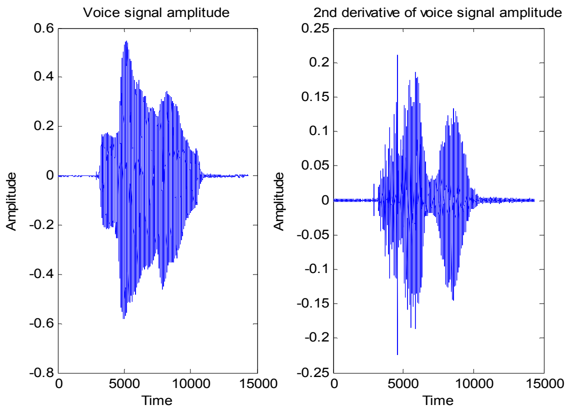
Voice	Sigma	Mu	Peak (Crest Factor, CF)	Dynamic range	Power Spectral Density (PSD)	Calculation time (s)
S1	0.090572	0.001204	15.4313	60.7982	0.0026	0.2366
S2	0.17885	-0.000242	14.0642	65.3405	0.0102	0.2376
S3	0.14246	-0.000229	14.4525	63.7525	0.0065	0.2391
S4	0.10539	0.0012252	13.5841	60.2668	0.0035	0.2346
S5	0.12479	0.0012236	13.2751	61.4252	0.0050	0.2342
S6	0.088082	-0.000240	16.2567	61.3809	0.0025	0.2359
S7	0.11217	0.0012188	13.9156	61.14	0.0040	0.2356
S8	0.12832	-0.000714	13.1212	61.5132	0.0052	0.2329

TABLE II. RESULTS OF EXPERIMENT 1

Voice signal	Number of samples	Features		
		Number of local max	Stable	Number of local min
S1	12544	5185	2423	4936
S2	14080	5743	2450	5887
S3	18944	8091	2675	8178
S4	15872	7087	1795	6990
S5	16384	6915	2503	6966
S6	16896	6788	3074	7034
S7	14592	6142	2477	5973
S8	14336	6193	1615	6528

TABLE III. RESULTS OF EXPERIMENT 2

Voice signal	Size number of samples	Frequency Sampling (FS)	Features		
			Number of local max	Stable	Number of local min
Cow	30964*2	16000	30689	5	31234
Bird	78588*2	44100	63759	21867	71550
Dog	46069*2	44100	41300	7590	43248
Dolphin	39294*2	44100	39972	3968	34648
Donkey	118994*2	44100	112979	10988	114021
Duck	205536*2	44100	192594	27020	191458
Elephant	22517*2	44100	21957	1769	21308



(a) (b)

Fig. 4. Voice signal and 2nd derivative

TABLE IV. RESULTS OF EXPERIMENT 3

Voice	Calculation time (s) Parameters (1)	Calculation time (s) 2nd derivative (2)	Speedup=(1)/(2)
S1	0.2366	0.008166	28.9738
S2	0.2376	0.007583	31.3332
S3	0.2391	0.007616	31.3944
S4	0.2346	0.007489	31.3259
S5	0.2342	0.007429	31.5251
S6	0.2359	0.007413	31.8225
S7	0.2356	0.007441	31.6624
S8	0.2329	0.007394	31.4985
Cow	0.7429	0.008937	83.1263
Dog	0.7674	0.009850	77.9086

IV. CONCLUSION

In the current research, a method based on the wave equation designed to extract voice signal features was proposed. The experimental results showed that the extracted feature array for each voice signal is unique and can be used as a key to identify the voice signal. The proposed method suits any type of voice signal with any size and any sampling frequency. Moreover, it was shown that the feature array is very sensitive to any changes in the voice signal. The comparison between the proposed method's experimental results and other methods, like the voice signal parameters method described in [4-5], proved that our new proposed method is faster.

REFERENCES

[1] R. Szabo, A. Gontean, I. Lie, "Sound Based Coin Recognition and Clapper", 16th International Conference on Soft Computing (MENDEL), pp. 509–516, June 23-25, 2010.

[2] A. K. Paul, D. Das, M. M. Kamal, "Bangla Speech Recognition System Using LPC and ANN", Seventh International Conference on Advances in Pattern Recognition (ICAPR), Kolkata, India, pp. 171–174, February 4-9, 2009

[3] R. Szabo, A. Gontean, "Human Voice Signal Synthesis and Coding", IFAC Proceedings Volumes, Vol. 46, No. 28, pp. 336-341, 2013

[4] K. M. Matrouk, A. Alhasanat, H. Alasha'ary, Z. Alqadi, H. M. Al-Shalabi, "Speech Fingerprint to Identify Isolated Word-Person", World Applied Sciences Journal, Vol. 31, No. 10, pp. 1767-1771, 2014

[5] R. M. Stern, F. H. Liu, Y. Ahshima, T. M. Sullivan, A. Acero, "Multiple Approaches to Robust Speech Recognition", DARPA Speech V Natural Language Workshop, pp. 274–279, Harriman, 1992

[6] P. Chatzimisios, C. Verikoukis, I. Santamaría, M. Laddomada, O. Hoffmann, Mobile Lightweight Wireless Systems, Second International ICST Conference, Mobilight, Barcelona, Spain, 2010, Springer, 2011

[7] B. Wu, J. Zhu, F. N. Najm, "Dynamic Range Estimation", IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, Vol. 25, No. 9, pp. 1618-1636, 2006

[8] B. Wu, J. Zhu, F. N. Najm, "An analytical approach for dynamic range estimation", 41st Design Automation Conference, San Diego, USA, July 7-11, 2005

[9] S. Miller, D. Childers, Probability and random processes, Academic Press, pp: 370-375, 2012

[10] F. Gouyon, F. Pachet, O. Delerue, "Classifying percussive sounds: a matter of zero-crossing rate", COST G-6 Conference on Digital Audio Effects (DAFX-00), Verona, Italy, December 7-9, 2000

[11] A. Fichtner, S. M. Hanasoge, "Discrete wave equation upscaling", Geophysical Journal International, Vol. 209, No. 1, pp. 353–357, 2017

[12] V. E. Tarasov, "Discrete Wave Equation with Infinite Differences", Applied Mathematics & Information Sciences Letters, Vol. 5, No. 2, pp. 41-44, 2017

AUTHORS PROFILES

Dr. Ayman Y. Al-Rawashdeh was born on 01 January 1970 in Jordan. He received his diploma degree in 1995 and PhD degree in 2008 in the field of Mechatronics Engineering. Currently he works as an Assistant Professor at the Electric Power Engineering Department, Faculty of Engineering Technology, Al-Balqa Applied University, Jordan. His main interests include renewable energy, drive system analysis, simulations and signal processing.

Ziad AL-Qadi was born on 09 March 1955 in Jordan. He received his diploma degree in 1980 and PhD degree in 1986 in the field of Computer Engineering. Currently he is a Professor at the Computer Engineering Department, Faculty of Engineering Technology, Al-Balqa Applied University, Jordan. His main interests include signal processing, pattern recognition, algorithms, modelling and simulations.