# Fractal Dimensions of Voice Patterns and Voice Recognition

[1]Randy Salazar, [2]Rosario Cabillada, [3]Mark Borres and [4]Anthony Jagures


### *Abstract*

*The popularity and the convenience of using electronic communications have given rise to more transactions online. Despite the regular updates of safeguards, there are significant numbers of transactions that go awry. In the hotel business, forgeries and prank calls may be problematic, but there is nothing more distressing than to settle conflicts with guests. The lack of means to recognize, identify and verify callers exposes the transactions to pranks or to misunderstandings. In either case these frustrating transactions erode goodwill, which result in loss of future business. This study explores the use of fractal dimensions in characterizing the different facets of voice and speech dynamics. The different sinusoid samples intend to measure the physiological and the dynamics aspect of vocalization. Test results have shown that the differences of the group mean of the fractal dimensions of the voice wave patterns among the volunteers are significant. These also have shown the potential use of fractal dimensions in characterizing the voice patterns of different speakers and the eventual voice recognition or speaker identification.*


*Keywords: Voice recognition, voice patterns, fractals, fractal dimension, speaker identification*

**1.0 Introduction**

The advent of modern communication has made the lives of people in society more convenient. While in the past reservations are made personally on a face-to-face encounter today one reserves a booking by a telephone transaction. While these conveniences work well in an environment where forgeries and prank calls are non-existent the reality in business transactions proves otherwise. It is for this reason that authentication using voice recognition is an important field of investigation.

In the field of electronic communications functioning models of voice recognition and identification systems exist. Later studies propose the use of the Artificial Neural Network (ANN). ANN processes the speech information and speaker identification within the interconnected networks. The network interconnections increased the database and computing capacity, advance the decision making ability of the system (Hecht-Nielsen, 1989 )(Gershenson, 2003). Later applications of voice recognition have found their way in the mobile phone, call service centers and health care industries (Pogue, 2010, Knight, 2012, Chavez, 2013). However, the accuracy and the perspicacity of speech recognition system still something to be had. A widely used method to speech recognition is the Hidden Markov Model (HMM). This method measures the time variances in a spoken language and identifies a speaker by the statistical variance of the time dimension (Swee, 1998). The singular dimension of the characterization, that is time, limits the categorical variation of speech this will eventually create similarities and identification

[1,4]College of Engineering,
[2]College of Commerce,
[3]Center for Policy, Research and Development Studies

errors.

Fractal dimension measures how an object fills space, the approach can either be in two (2) or three (3) dimensions. Sinusoidal wave patterns are graphed in two or three dimensions. This offers more descriptive variations. This paper explores the use of fractal dimension to characterize the different voice wave patterns of a speaker for voice recognition.

## 2.0 Literature

The ability of humans to transfer concepts and ideas through spoken and written language is one of the greatest traits that separate humans from other animals (Perrachione, Del Tufo & Gabrieli, 2011).

### a. The Physics of sound

Sound is a disturbance of the air pressure that results from vibration. It is the center of speech communication. A sound wave is both the end product of the speech production mechanism and the primary source of raw material used by the listener to recover the speaker's message (Berg & Stork, 1982).

According to Furtună, (2008), sound travels through the environment as a longitudinal wave with a speed that depends on the environment density. The easiest way to represent a sound is a sinusoidal graphic. The graphic presents variation of air pressure depending on time. The shape of the sound wave depends on three factors: amplitude, frequency and phase.

The amplitude is the displacement of the sinusoidal graph above and below temporal axis (y = 0) and it corresponds to the energy the sound wave. Amplitude is measured in decibels (DB) and is the direct representation of how people hear sound volume. Frequency is the number of cycles the sinusoid makes every second and is measured in Hertz (Hz). The time needed for the sound wave to complete a cycle is a period. The phase measures the position of the start of a sinusoidal curve. Humans detect phase as a time delay between the two signals. This ability is how human sensorial system perceives a sound location, counting on different phases perceived by the ears (Furtună, 2008).

### b. Speaker Recognition

The ability to recognize individual conspecifics from their communicative vocalizations is an adaptive trait evinced widely among social and territorial animals, including humans. For humans, the ability to recognize one another by voice relies on the ability to compute the differences between the incidental phonetics of a particular vocalization and the abstract phonological representations of the words that vocalization contains (Perrachione, Del Tufo & Gabrieli, 2011).

According to Padmanabhan, (2012), speaker recognition is the process of identifying people from their voices. Further individuals will not sound alike because of physiological differences in their speech production mechanisms. People differ in their manner of speaking, which lead to differences in speaking rate, accent, intonation and others. Speaker recognition systems try to take advantage of these factors to discriminate between speakers.

According to Melim et al., (2006), speaker recognition methods can also be divided into text-dependent and text-independent methods. The former requires the speaker to say keywords or sentences that have the same text for both training and recognition trials, whereas the latter does not rely on a specific text being spoken.

Text-dependent methods recognize using a template matching techniques. In this method, the input utterance is represented by a sequence of spectral feature vectors. The time axes of the input utterance and each reference template or reference model of the registered speakers are aligned using a dynamic time warping (DTW) algorithm and the degree of similarity between them, accumulated from the start to the end of the utterance, is calculated (Melim et al., 2006).

### c. Approaches to speech recognition

According to Anusuya, & Katti, (2010), there are three approaches to speech recognition, a.) Acoustic Phonetic Approach; b.) Pattern Recognition Approach; and c.) Artificial Intelligence Approach.

*c.1 Acoustic phonetic approach*

Hemdal and Hughes (1967), postulate that there exist a finite, distinctive phonetic unit (phonemes) in spoken language. These units are broadly characterized by a set of acoustics properties that are manifested in the speech signal over time. This postulate is the basis of the acoustic phonetic approach. The earliest algorithms of speech recognition were based on segmenting and identifying speech sounds and providing appropriate labels to these sounds. The first step is the spectral analysis of the speech, combined with feature detection that converts spectral measurements to a set of characteristics that describe the broad acoustic properties of the different phonetic units. The next step is the segmentation and phonetic labeling of the stable acoustic region. The result is a phoneme lattice characterization of the speech (Anusuya, & Katti, 2010).

*c.2 Pattern Recognition approach*

The pattern-matching approach (Itakura 1975; Rabiner 1989; Rabiner and Juang 1993) involves two essential steps namely, pattern training and pattern comparison. This approach uses a well formulated mathematical framework and establishes consistent speech model representations. A speech model representation can be in the form of a speech template or a statistical model. A good example of the statistical model is the Hidden Markov Model (HMM) and can be applied to a sound (smaller than a word), a word, or a phrase. In the pattern-comparison, an assessment on the likeness and resemblance is made between an unidentified speech or speaker with each of the possible learned pattern in the training stage.

*c.3 Artificial Intelligence Approach (AI)*

The Artificial Intelligence approach is a hybrid of the acoustic phonetic approach and pattern recognition approach. Dijkstra, and de Smedt, (1996), define Artificial Intelligence (AI) as a branch of computer science in which develop methods and techniques that permit intelligent computer systems to be built. One field of study of AI is knowledge engineering design. Knowledge engineering involves the direct and explicit incorporation of the expert's speech knowledge into the recognition system. The limited success of pure knowledge engineering approach is mainly because of the difficulty of quantifying expert knowledge. Another difficult problem is the integration of many levels of human knowledge; phonetics, phonotactics, lexical access, syntax, semantics and pragmatics (Anusuya, & Katti, 2010).

**d. The Sound of Language**

In a lecture about approaches to speech recognition, the sounds of language are classified into what are called phonemes. A phoneme is a minimal unit of sound that has semantic content. e.g., the phoneme AE versus the phoneme EH captures the difference between the words "bat" and "bet." Not all acoustic changes change meaning. For instance, singing words at different notes does not change meaning in English. Changes in pitch do not lead to phonemic distinctions (Allen, 2003).

An important distinguishing feature of a phoneme is voicing. A voiced phoneme includes the sound from the vocal chords. A good example is the sound of a vowel. Vowel sounds stay steady over the time these are produced. The vowels IY (beat), IH (bit), EH (bat), and AE (bet) are vocalized by holding the tongue to the front and vary its height. To vocalize the vowels AA (Bob - Bahb), ER (Bird), AH (but), and AO (bought) hold the tongue in mid position. The tongue at the back position will resonate the sound of the vowels UW (boot), UH (book), and OW (boat).

Another class of vowels called the dipthongs, which change during their duration. Dipthongs starts with one vowel and ends with another. Some samples are AY (buy), AW (down, cf. AA W), EY (bait, cf. EH IY), and OY (boy, cf. AO IY).

Consonants fall into general classes, with many classes having voiced and unvoiced members. Table 1 shows the classes.

**e. Mapping a Continuous Speech**

A common approach to mapping a signal is by discrete events. These define a set of symbols that

Table 1. Consonant phonemes

| | Class | Voicing Description | consonants |
|---|---|---|---|
| a. | **Stops or Plosives** | These involve stopping the speech stream using the lips, tongue, and so on, and then rapidly releasing a stream of sound. These come in unvoiced, voiced pairs | P and B, T and D, and K and G. |
| b. | **Fricatives** | These involve "hissing" sounds generated by constraining the speech stream by the lips, teeth and so on. They also come in unvoiced, voiced pairs | F and V, TH and DH (e.g., thing versus that), S and Z, and finally SH and ZH (e.g., shut and azure) |
| c. | **Nasals** | These are voiced and involve moving air through the nasal cavities by blocking it with the lips, gums, and so on. | M, N, and NX (sing) |
| d. | **Affricatives** | These are like stops followed by a fricative | CH (church), JH (judge). |
| e. | **Semi Vowels** | These are consonants that have vowel-like characteristics | W, Y, L, and R. |
| f. | **Whisper** | Whisper | H |

correspond to useful acoustic features of the signal over a short time interval. A phoneme may look an ideal unit, but it is too complex and time is too long to be classified with simple signal processing techniques.

Discrete event mapping uses a much smaller interval of speech, typically 20 milliseconds. These intervals often overlap in time. Segments are then classified into a set of different types, each corresponding to a new symbol in a vocabulary called the codebook. In speech recognition, a word will contain many acoustic events. Mapping acoustic events to form words means the use of subword models. These models could be: a.) phoneme-based; b.) Syllable-based; c.) Demi-Syllable based, where each sound represents a consonant cluster preceding a vowel, or following a vowel; d.) Phonemes in context, or triphones: context dependent models of phonemes depend on what precedes and follows the phoneme (Allen, 2003).

**f. Signal Conditioning**

Signal conditioning, in electronic communications includes functions such as signal amplification, filtering, electrical isolation, and multiplexing. In addition, it also may require excitation currents or voltages, bridge completion, linearization, or high amplification for proper and accurate operation or input for the next process (Soloman, 2009).

Signal amplification performs two important functions: increase the resolution of the input signal, and increase its signal-to-noise ratio. Filtering is the removal of the unwanted noise or frequency. Noise, in communication systems, is an error or undesired random disturbance of a useful information signal in a communication channel (Scherz, 2006). Signal isolation is often used to isolate possible sources of signal interference (Austerlitz, 2002).

**g. Fractals**

Fractal is a general term used to describe both the geometry and the processes which exhibit self-similarity, scale invariance, and fractional dimension (Mandelbrot, 1983). Mandelbrot shows that a classical geometry deals with shapes or objects described in integral dimension. A point is 0-dimensional, a line having 1-dimension, a plane figure with 2-dimensions and the 3-dimensional solid. He further showed that there are many phenomena that are appropriately described in terms of dimensions that are between two dimensions. A straight line has dimension d = 1 and a zigzag of this will have a dimension between one and two for the curve is more than a line but less than a 2-dimensional figure. Here, the dimension manifested is referred to as a fractional dimension - a dimension whose value lies between integral values.

Fractal image analysis according to Zmeskal et al. (2013) means the determination of fractal dimension and fractal measure of the image. Fractal dimensions and fractal measures are obtained by using the Box Counting Method. Traditionally, box counting method works by laying meshes of different sizes r, and counting the number of boxes N needed to cover tested object completely. Using these values of r and N, the regressed slope D of the linear portion of the function log N(r) = D (log (1/r)) + log k is the box fractal dimension of the tested image, and the regressed intercept k is the fractal measure.

The use of fractals in the study of voice recognition is relatively few. Fractal analysis numerically characterizes images and patterns. The method is aptly suitable for the different sinusoidal patterns of sound waves.

**3.0 Methodology**

This paper look into the fractal dimensions of the sinusoidal patterns of the sound wave of the phoneme "AH" and the syllable "PASS" in the sample word "PASSWORD." The word "PASSWORD" includes several phonemes "P", "AH", "SS", "W", "AO", "R" and "D". Figure 1 shows the segmentation of "PASSWORD" to its basic phonemes. A sound editor generates the sinusoidal patterns of the voice samples. The wave patterns are identified and extracted for the phoneme "AH" and "PASS" segments. The sample segment sizes of "AH" are in a 5-wave cycle and a 30-millisecond segment and "PASS" in a syllable segment. Fractal image analysis software measures the fractal dimension for each of the extracted segment of sinusoid image. The phoneme "AH" 5-cycle segment and "PASS" segment are sampled based on wave pattern features. The "AH" 30-millisecond segment is sampled on a fixed time and is characterized by the frequency of voices spoken.

Ten voice samples from ten individual volunteers were tested. The samples consist of ten vocalizations of the sample word. The voices were communicated and recorded through two types of media one direct and the other via the intercom, with five samples from each medium see figure 2. The following were the parameters of the data gathering; a.) Voices were recorded in normal vocalization, b.) Voice segment samples were extracted from a common phoneme, (Allen, 2003) except for "PASS" c.) Amplitude modification was proportionately applied to the segment samples (Scherz, 2006, Soloman, 2009). Noise is kept minimal while recording to avoid filtering. The filtering procedure requires more studies and may distort the actual voices of the volunteers in the process.

The paper uses Free Audio Editor (Version 2014.8.6.1) for the sinusoid pattern generation and sound signal conditioning. Figures 3, 4 and 5, are the sample images of the wave patterns. Fract3 software is also used for the fractal analysis and characterization of the wave patterns. Figures 3A, 4A and 5A are the fractal analysis sample results of the sinusoid images. The fractal dimensions of the voice segments are then subjected to an ANOVA or variance analysis. This procedure intends to determine if the difference of the characterization is significant. Fract3 and Free Audio Editor are freeware online, can be freely downloaded with necessary downloading permissions.

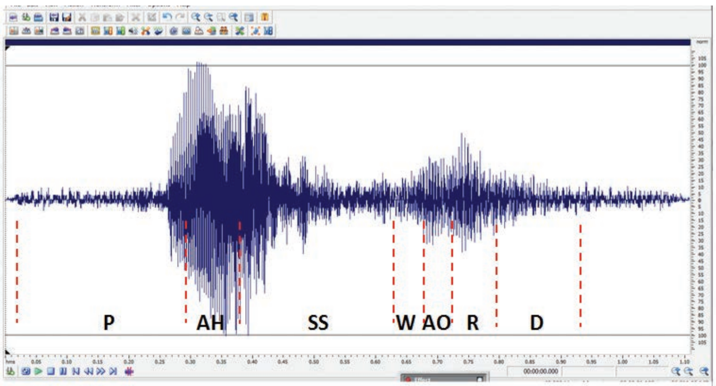
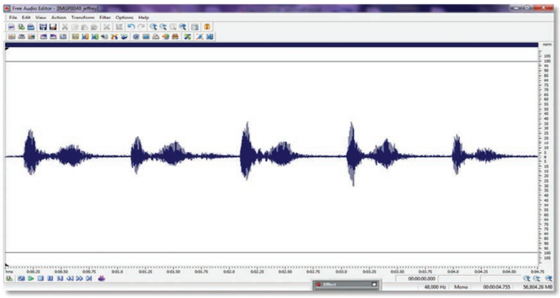
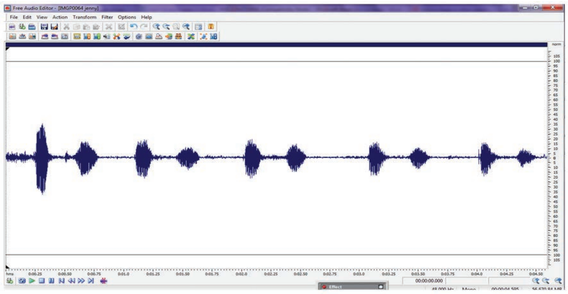



**Figure 1. Phoneme Map of "PASSWORD"**



a.) Voice Data, five samples direct recordings



b.) Voice Data, five samples through intercom

**Figure 2. Voice Data, ten samples of "PASSWORD" records**

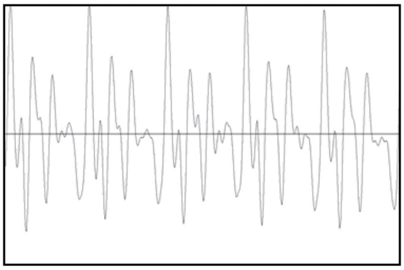


**Figure 3. "AH" 5-cycle segment wave pattern**

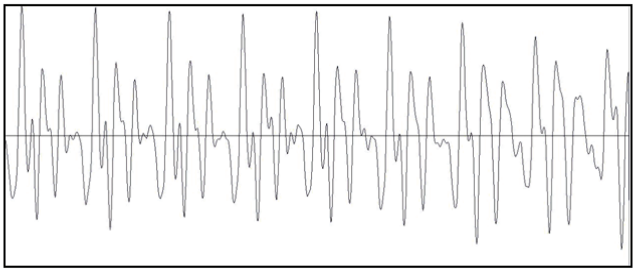


**Figure 4. "AH" 30 millisecond segment wave pattern**

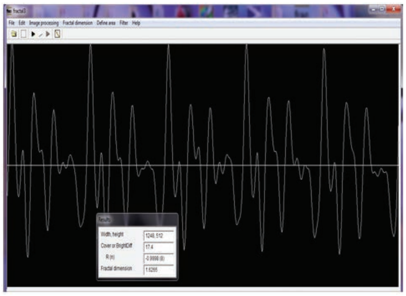


**Figure 3A. "AH" 5-cycle segment FRACTAL ANALYSIS (FD = 1.6265)**

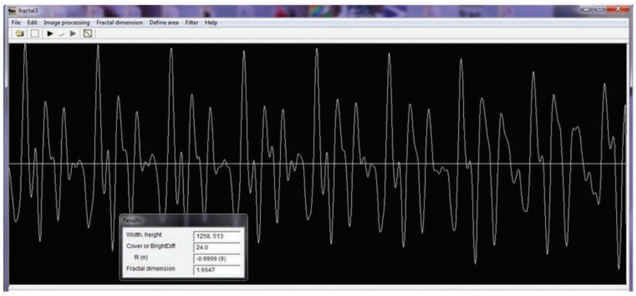


**Figure 4A. "AH" 30 millisecond segment wave pattern FRACTAL ANALYSIS (FD = 1.6647)**

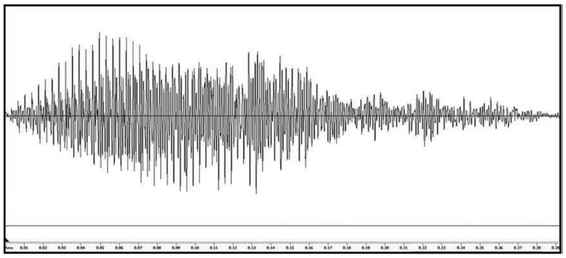


**Figure 5. "PASS" Typical Wave Pattern sample**

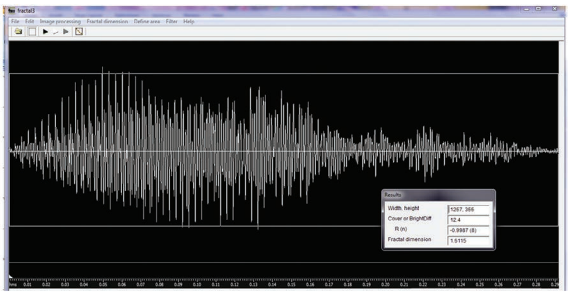


**Figure 5A. "PASS" Typical Wave Pattern FRACTAL ANALYSIS (FD = 1.6115)**

**4.0 Results and Discussion**

To develop a speaker recognition machine according to Reynolds, (1995), one must first ask, "How humans recognize one another by voice alone?" Traditional methods of speaker identification correlate the speaker identity with physiological and behavioral characteristics of the speaker. These features exist in both the vocal source and vocal tract characteristics of the person and the dynamics of the speech (Melim et al., 2006). These features will eventually be depicted in the sound waves vocalized during a speech.

The sample "AH" 5-cycle by observation shows the natural voice patterns and features of the vocal chords and tract of the person for "AH" phoneme. These sinusoidal patterns differ from one person to the other, figure 6. The sample "AH" 30-millisecond segment, exhibit the frequency characteristics of the voice sample. Depending on the pitch of a person's normal voice, the image shows how the sinusoid fills the 30-milliseconds time space. The higher the pitch the more waves fill the 30-millisecond space, figure 7. The sample "PASS" presents the speech dynamics pattern of a person. The emphases and stresses in the vocalization of the inclusive phonemes are depicted in the amplitudes ratios and the length of time of vocalization, Figure 8.

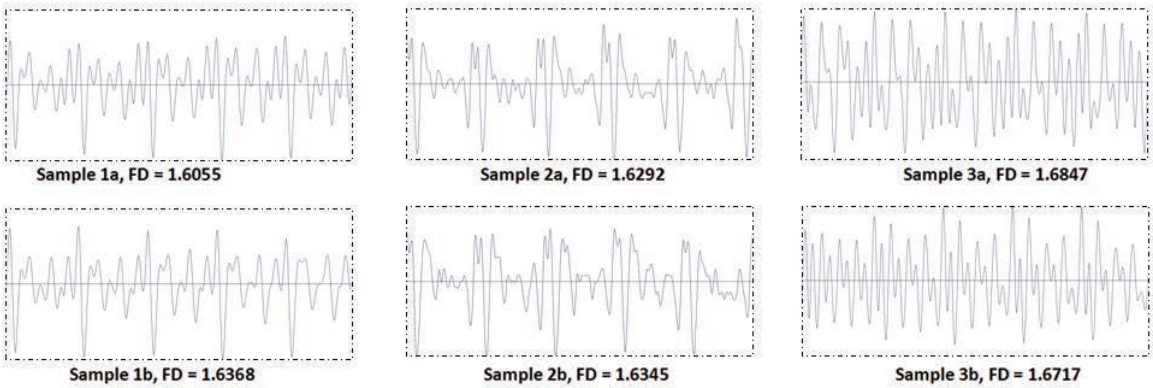



**Figure 6. "AH" 5- cycle Samples**

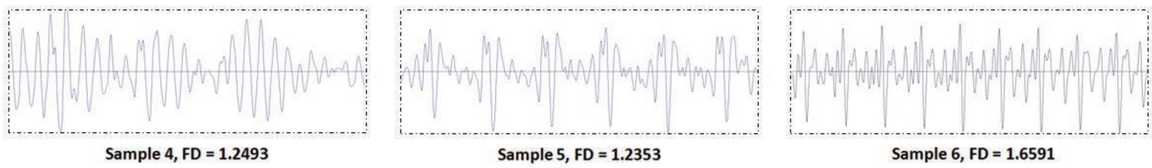



**Figure 7. "AH" 30 millisecond Samples**

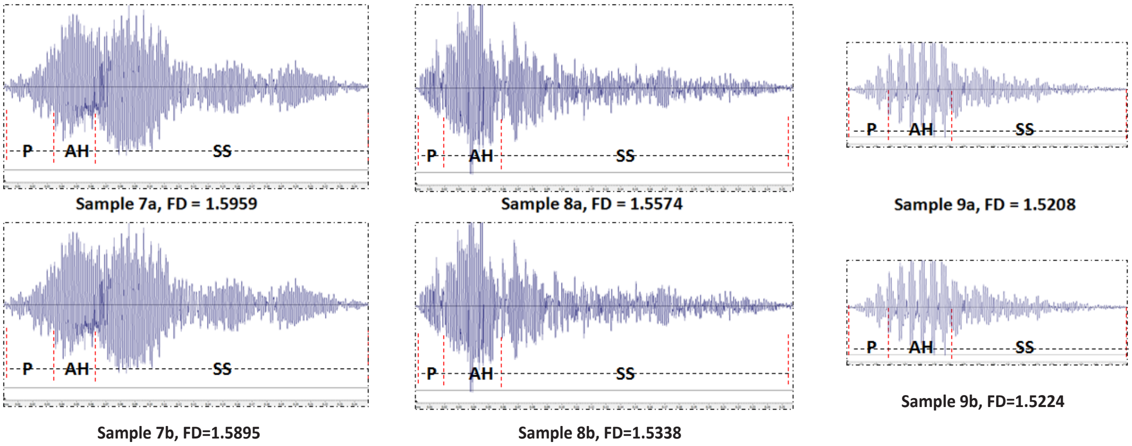



**Figure 8. "PASS" Samples**

The study explores the use of fractals in characterizing the voice patterns of the different test volunteers. Based on the concepts presented the following are the assumptions:

A 1: Every person has a different physiological speech production mechanism. (Melim, et al., 2006), (Padmanabhan, 2012). This difference creates a unique voice signature wave pattern. The fractal dimension of the "AH" 5-wave cycle sample is unique to a person and different from each other.

A 2: Voices transmitted via the electronic media are a recreation of the original voice and are altered by the electronic media system (Scherz, 2006). These will have different wave pattern as with the voice directly communicated. The fractal dimensions of these wave patterns will differ from the other.

A 3: Under normal communications, each person has a distinctive voice frequency. This distinction can be observed on a fixed time segment. The fractal dimension of 30 millisecond segment will differ between each others.

A 4: Under normal communications, the dynamics of speech or vocalization is unique to the person and will differ with the others. The fractal dimensions of the sinusoidal pattern of vocalization of the syllable "PASS" will be different from one person to the others.

From these assumptions the following null hypotheses are tested:

*Ho 1. "There is no significant difference on the fractal dimensions among the test volunteer's "AH" 5-cycle voice signature wave patterns."*

*Ho 2. "There is no significant difference on the fractal dimensions between the test volunteer's voice wave patterns that are communicated directly and those communicated via the intercom."*

*Ho 3. "There is no significant difference on the fractal dimensions of the sinusoidal patterns sound of the "AH" 30-millisecond sample among the test volunteers."*

*Ho 4. "There is no significant difference on the fractal dimensions of the sinusoidal patterns of the sound created on the syllable "PASS" among test volunteers."*

On the null hypotheses:

*Ho1: "There is no significant difference on the fractal dimensions among the test volunteer's "AH," 5-cycle voice signature wave patterns."*

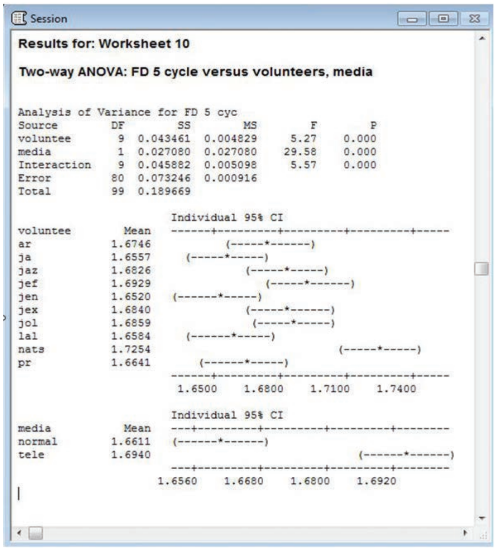
Figure 9 is the two-way ANOVA test of the fractal dimensions of the "AH" 5-cycle samples. One hundred voice samples were tested. There are ten individual volunteers each with ten voice samples taken. Five voice samples were communicated directly and recorded and five voice samples communicated and recorded via the intercom. The test shows the effect of the two factors “volunteers” and “media” to voice identification. The factor “volunteer” refers to person factor and “media” refers to medium factor, that is direct or via intercom.

For each volunteer, ten voice samples were grouped. Each volunteer with a group mean of the voice pattern fractal dimensions. At 95% confidence index, the null hypothesis has a 0.000 probability of likelihood to be true. The null hypothesis *Ho1* is rejected. Hence the group mean of the fractal dimensions of the “AH” wave pattern from a volunteer is significantly different from the other wave patterns of the other volunteers. Therefore the fractal dimension of the “AH” wave pattern is uniquely attributable to one person or volunteer.

*Ho2: “There is no significant difference on the fractal dimensions between the test volunteer’s voice wave patterns that are communicated directly and those communicated via the intercom.”*

Figure 9, also revealed the effect of the second factor “media”. There are two types of media tested. “Normal” refers to the directly communicated voices and “tele” refers to voices communicated through the intercom. The critical value F(1, 80) is 3.9604. Results shows an F(1, 80) = 29.58, hence, the difference of the group mean of the fractal dimensions of the voices communicate directly and through the intercom is significant. The same results can also be found in the tests conducted for the 30-millisecond and “PASS” samples, figures 10 and 11. This means that the fractal dimensions of the voice communicated via intercom is different from the fractal dimensions of the voice communicated direct. That voices via intercom may not be identifiable if the database used is that of the voice directly communicated. Voice recognition will need sets of database one for each media. The alternative is a filter system that can isolate the voice and remove the effect of the media system.

The media effect is also shown in the two-way ANOVA result on interaction among the factors “volunteers” and “media”, figures 9, 10 and 11. There is a significant interaction between these two factors. The volunteers submitted two sets of samples, direct recording and via intercom. The interference of the electronic media in the recreation of the sound has a significant effect on the wave patterns and its fractal


Session

Results for: Worksheet 10

Two-way ANOVA: FD 5 cycle versus volunteers, media

Analysis of Variance for FD 5 cyc

| Source | DF | SS | MS | F | P |
|---|---|---|---|---|---|
| voluntee | 9 | 0.043461 | 0.004829 | 5.27 | 0.000 |
| media | 1 | 0.027080 | 0.027080 | 29.58 | 0.000 |
| Interaction | 9 | 0.045882 | 0.005098 | 5.57 | 0.000 |
| Error | 80 | 0.073246 | 0.000916 | | |
| Total | 99 | 0.189669 | | | |

Individual 95% CI

| voluntee | Mean |
|---|---|
| ar | 1.6746 |
| ja | 1.6557 |
| jaz | 1.6826 |
| jef | 1.6929 |
| jen | 1.6520 |
| jex | 1.6840 |
| jol | 1.6859 |
| lal | 1.6584 |
| nats | 1.7254 |
| pr | 1.6641 |

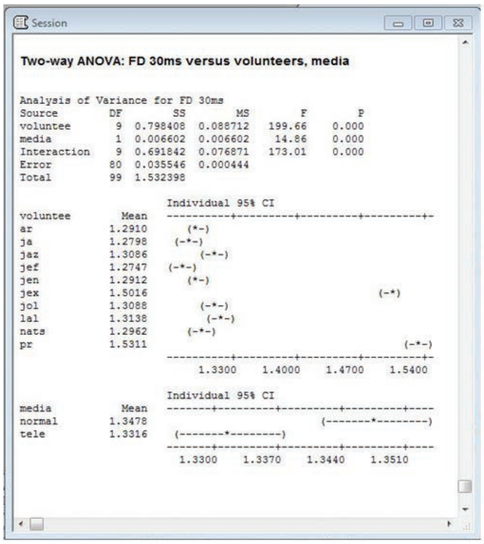
1.6500 1.6800 1.7100 1.7400

Individual 95% CI

| media | Mean |
|---|---|
| normal | 1.6611 |
| tele | 1.6940 |

1.6560 1.6680 1.6800 1.6920

**Figure 9. Two-Way ANOVA of “AH” 5-wave cycle samples Fractal Dimension**


Session

Two-way ANOVA: FD 30ms versus volunteers, media

Analysis of Variance for FD 30ms

| Source | DF | SS | MS | F | P |
|---|---|---|---|---|---|
| voluntee | 9 | 0.798408 | 0.088712 | 199.66 | 0.000 |
| media | 1 | 0.006602 | 0.006602 | 14.86 | 0.000 |
| Interaction | 9 | 0.691842 | 0.076871 | 173.01 | 0.000 |
| Error | 80 | 0.035546 | 0.000444 | | |
| Total | 99 | 1.532398 | | | |

Individual 95% CI

| voluntee | Mean |
|---|---|
| ar | 1.2910 |
| ja | 1.2798 |
| jaz | 1.3086 |
| jef | 1.2747 |
| jen | 1.2912 |
| jex | 1.5016 |
| jol | 1.3088 |
| lal | 1.3138 |
| nats | 1.2962 |
| pr | 1.5311 |

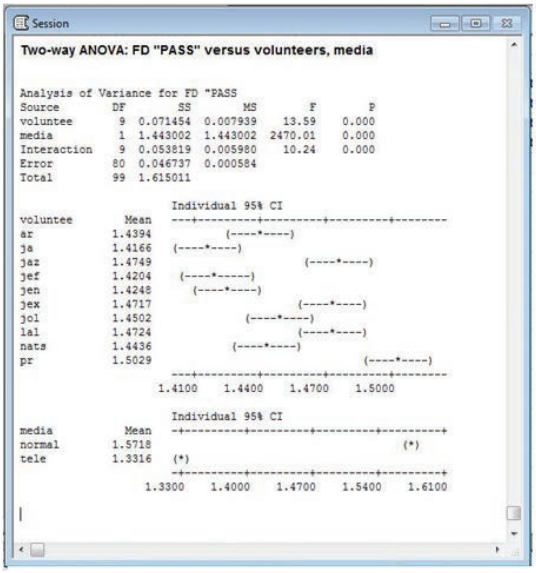
1.3300 1.4000 1.4700 1.5400

Individual 95% CI

| media | Mean |
|---|---|
| normal | 1.3478 |
| tele | 1.3316 |

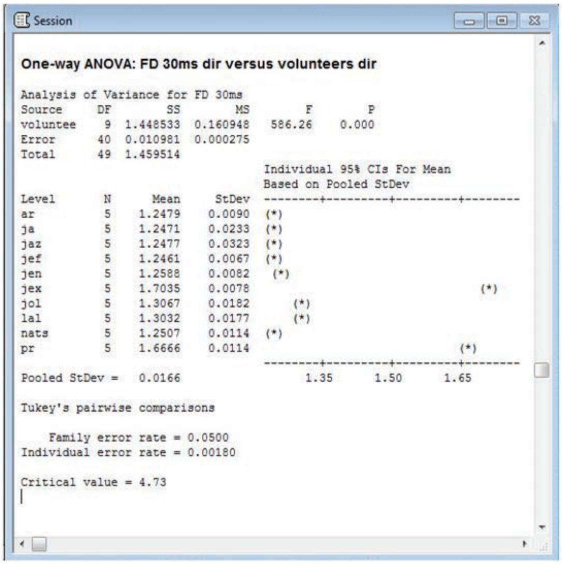
1.3300 1.3370 1.3440 1.3510

**Figure 10. Two-Way ANOVA of “AH” 30-millisecond samples Fractal Dimension**



```
Session
Two-way ANOVA: FD "PASS" versus volunteers, media

Analysis of Variance for FD "PASS
Source        DF        SS        MS        F        P
voluntee       9  0.071454  0.007939    13.59    0.000
media          1  1.443002  1.443002  2470.01    0.000
Interaction    9  0.053819  0.005980    10.24    0.000
Error         80  0.046737  0.000584
Total         99  1.615011

                       Individual 95% CI
voluntee      Mean   ---+---------+---------+---------+--------
ar          1.4394           (----*----)
ja          1.4166   (----*----)
jaz         1.4749                        (----*----)
jef         1.4204    (----*-----)
jen         1.4248      (----*----)
jex         1.4717                       (----*----)
jol         1.4502               (----*----)
lal         1.4724                       (----*----)
nats        1.4436             (----*----)
pr          1.5029                                 (----*----)
                     ---+---------+---------+---------+--------
                      1.4100    1.4400    1.4700    1.5000

                       Individual 95% CI
media         Mean   -+---------+---------+---------+---------+
normal      1.5718                                       (*)
tele        1.3316   (*)
                     -+---------+---------+---------+---------+
                   1.3300    1.4000    1.4700    1.5400    1.6100
```

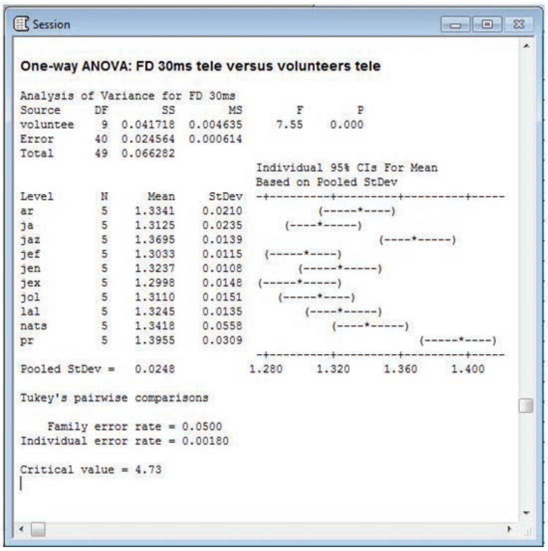
**Figure 11. Two-Way ANOVA of "PASS" samples Fractal Dimension**



```
Session
One-way ANOVA: FD 30ms dir versus volunteers dir

Analysis of Variance for FD 30ms
Source     DF        SS        MS        F        P
voluntee    9  1.448533  0.160948   586.26    0.000
Error      40  0.010981  0.000275
Total      49  1.459514
                                   Individual 95% CIs For Mean
                                   Based on Pooled StDev
Level       N      Mean     StDev  --------+---------+---------+--------
ar          5    1.2479    0.0090  (*)
ja          5    1.2471    0.0233  (*)
jaz         5    1.2477    0.0323  (*)
jef         5    1.2461    0.0067  (*)
jen         5    1.2588    0.0082   (*)
jex         5    1.7035    0.0078                                (*)
jol         5    1.3067    0.0182      (*)
lal         5    1.3032    0.0177      (*)
nats        5    1.2507    0.0114  (*)
pr          5    1.6666    0.0114                             (*)
                                   --------+---------+---------+--------
Pooled StDev =   0.0166                   1.35      1.50      1.65

Tukey's pairwise comparisons

    Family error rate = 0.0500
Individual error rate = 0.00180

Critical value = 4.73
```

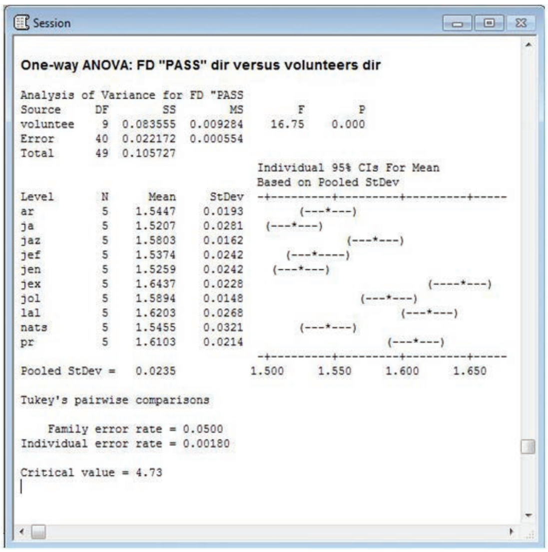
**Figure 12. One-Way ANOVA of Fractal Dimension of "AH" 30-ms direct recorded samples**

dimensions. The volunteer's voice patterns will be affected by the medium of communication, in the way the intercom system influence the sinusoidal patterns of the voice as it transmits and recreates the voices of the volunteers. Hence the fractal dimensions of these voices will be affected. This is an important consideration to include a noise filter system in voice recognition. This is also the reason why a filter system is a standard feature on existing voice recognition systems.



```
Session
One-way ANOVA: FD 30ms tele versus volunteers tele

Analysis of Variance for FD 30ms
Source     DF        SS        MS        F        P
voluntee    9  0.041718  0.004635     7.55    0.000
Error      40  0.024564  0.000614
Total      49  0.066282
                                   Individual 95% CIs For Mean
                                   Based on Pooled StDev
Level       N      Mean     StDev  -+---------+---------+---------+-----
ar          5    1.3341    0.0210           (-----*----)
ja          5    1.3125    0.0235       (----*-----)
jaz         5    1.3695    0.0139                     (----*-----)
jef         5    1.3033    0.0115    (-----*----)
jen         5    1.3237    0.0108         (-----*-----)
jex         5    1.2998    0.0148   (-----*-----)
jol         5    1.3110    0.0151      (-----*----)
lal         5    1.3245    0.0135          (----*-----)
nats        5    1.3418    0.0558              (----*-----)
pr          5    1.3955    0.0309                          (-----*----)
                                   -+---------+---------+---------+-----
Pooled StDev =   0.0248           1.280     1.320     1.360     1.400

Tukey's pairwise comparisons

    Family error rate = 0.0500
Individual error rate = 0.00180

Critical value = 4.73
```

**Figure 13. One-Way ANOVA of Fractal Dimension of "AH" 30-ms via intercom samples**



```
Session
One-way ANOVA: FD "PASS" dir versus volunteers dir

Analysis of Variance for FD "PASS
Source     DF        SS        MS        F        P
voluntee    9  0.083555  0.009284    16.75    0.000
Error      40  0.022172  0.000554
Total      49  0.105727
                                   Individual 95% CIs For Mean
                                   Based on Pooled StDev
Level       N      Mean     StDev  -+---------+---------+---------+-----
ar          5    1.5447    0.0193         (---*---)
ja          5    1.5207    0.0281    (---*---)
jaz         5    1.5803    0.0162                (---*---)
jef         5    1.5374    0.0242       (---*----)
jen         5    1.5259    0.0242     (---*---)
jex         5    1.6437    0.0228                           (----*---)
jol         5    1.5894    0.0148                  (---*---)
lal         5    1.6203    0.0268                        (---*---)
nats        5    1.5455    0.0321         (---*---)
pr          5    1.6103    0.0214                      (---*---)
                                   -+---------+---------+---------+-----
Pooled StDev =   0.0235           1.500     1.550     1.600     1.650

Tukey's pairwise comparisons

    Family error rate = 0.0500
Individual error rate = 0.00180

Critical value = 4.73
```

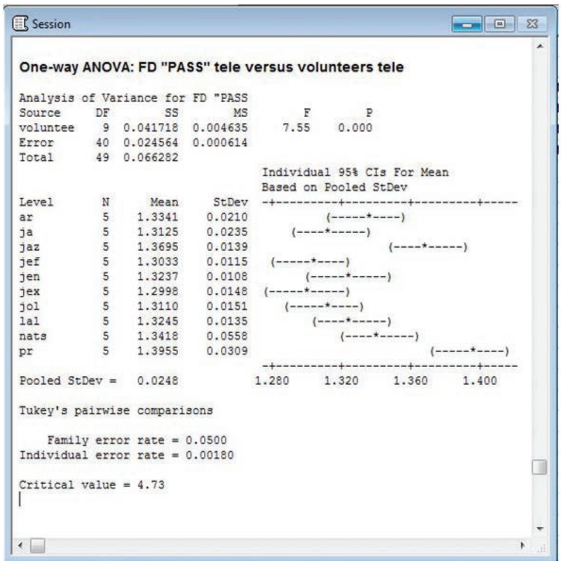
**Figure 14. One-Way ANOVA of Fractal Dimension of "PASS" direct recorded samples**

*Ho3: "There is no significant difference on the fractal dimensions of the sinusoidal patterns of the sound of "AH" 30 milliseconds sample among the test volunteers."*

Figure 12 and 13 show the one-way ANOVA result of the fractal dimensions of 30-millisecond samples of direct recording and via intercom respectively. The one way variance analyses shows F(9, 40) = 586.26. p = 0.00 and F(9, 40) = 7.55, p = 0.00 respectively. On both occasions, the group mean of the fractal dimensions of the "AH" 30-millisecond samples from an individual is significantly different from the others. The null hypothesis *Ho3* is rejected. On normal communications, the voice frequency wave pattern of the volunteer is significantly different from the other voice frequency patterns. Hence, the fractal dimension of person's voice frequency wave pattern can be associated to one person. There may be some caution however, the tone of a person's voice is always affected by the person's emotional disposition or it can also be intentionally altered.

*Ho4: "There is no significant difference on the fractal dimensions of the sinusoidal patterns of the sound created on the syllable "PASS" among test volunteers."*

Figures 14 and 15 are the one-way ANOVA tests of the fractal dimensions of the "PASS" wave pattern samples, both direct recorded and via intercom, respectively. The probability for the null hypothesis to be likely true is 0.000. The null hypothesis *Ho4* is



```
Session

One-way ANOVA: FD "PASS" tele versus volunteers tele

Analysis of Variance for FD "PASS
Source     DF        SS        MS        F        P
voluntee    9  0.041718  0.004635     7.55    0.000
Error      40  0.024564  0.000614
Total      49  0.066282
                                   Individual 95% CIs For Mean
                                   Based on Pooled StDev
Level       N      Mean     StDev  -+---------+---------+---------+-----
ar          5    1.3341    0.0210           (-----*----)
ja          5    1.3125    0.0235       (----*-----)
jaz         5    1.3695    0.0139                    (----*-----)
jef         5    1.3033    0.0115   (-----*----)
jen         5    1.3237    0.0108        (-----*-----)
jex         5    1.2998    0.0148  (-----*-----)
jol         5    1.3110    0.0151    (-----*----)
lal         5    1.3245    0.0135         (----*-----)
nats        5    1.3418    0.0558             (----*-----)
pr          5    1.3955    0.0309                         (-----*----)
                                   -+---------+---------+---------+-----
Pooled StDev =   0.0248           1.280     1.320     1.360     1.400

Tukey's pairwise comparisons

    Family error rate = 0.0500
Individual error rate = 0.00180

Critical value = 4.73
```

**Figure 15. One-Way ANOVA of Fractal Dimension of "PASS" via intercom recorded samples**

rejected. Under normal communications, the group mean of the fractal dimensions of the "PASS" sample wave patterns from a volunteer is significantly different from the other volunteers' "PASS" wave pattern. Hence, the speech dynamic wave pattern is different and the fractal dimension of this wave pattern can be associated to a single person.

## 5.0 Findings

Table 2 is the summary of the hypotheses test results.

**Table 2. Summary of ANOVA results**

| null hypotheses | indicators / tests | P -value | DF | F | Fcritical | Decision | interpretation |
|---|---|---|---|---|---|---|---|
| *Ho1: there is no significant difference in the Fractal Dimensions among the test's volunteer's "AH" 5-cycle voice signature wave patterns* | uFD51 = uFD52 = uFD53 = uFD54= uFD55= uFD56= uFD57= uFD58= uFD59= uFD510 | 0.000 | 9 | 5.27 | 1.9991 | **F > Fcritical** & the probability of the Ho1 to be likely true is 0.000; **Ho1 is rejected;** | the group mean of the fractal dimensions of the "AH" wave patterns is significantly different for each person, the fractal dimension of this signature wave pattern can be attributed to the person. |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| *variance analysis of the 2nd factor "media" Ho2: "There is no significant difference on the fractal dimensions between the test volunteer's voice wave patterns that are communicated directly and those communicated via the intercom."* | uFD5dir = uFD5tele | 0.000 | 1 | 29.58 | 3.9604 | in 3 tests: **F > Fcritical**; 0.000 probability that the hypothesis will likely to be true; the hypothesis is **rejected** | voices transmitted through an electronic media like the telephone or the intercom have different wave patterns than those directly received and the fractal dimensions of the wave patterns are different |
| | uFD30msdir = uFD30mstele | 0.000 | 1 | 14.86 | 3.9604 | | |
| | uFdpassdir = uFDpasstele | 0.000 | 1 | 2470.01 | 3.9604 | | |
| *Ho: there are no significant interactions among the factors volunteers and media* | C = (ud1 ut1-ud1 ut2)-(ud2 ut1-ud2 ut2)...: CFD5=0 | 0.000 | 9 | 5.57 | 1.9991 | in the 3 tests : **F > Fcritical** and the P-value is 0.000 the null hypothesis is also rejected | The interactions of the factors "volunteers" and "media" are significant. The fractal dimension of the wave patterns are affected by the type of media during communications. This is a major factor that must be considered in the implementation of noise filter in voice recognition. |
| | C = (ud1 ut1-ud1 ut2)-(ud2 ut1-ud2 ut2)...: CFD30=0 | 0.000 | 9 | 173.01 | 1.9991 | | |
| | C = (ud1 ut1-ud1 ut2)-(ud2 ut1-ud2 ut2)...: CFDpass=0 | 0.000 | 9 | 10.24 | 1.9991 | | |
| *Ho3: There is no significant difference on the fractal dimensions of the sinusoidal wave patterns of the "AH" 30 milliseconds sample among the test volunteers.* | for 30 ms direct; uFD1 = uFD2 = uFD3= uFD4= uFD5= uFD6= uFD7= uFD8= uFD9= uFD10 | 0.000 | 9 | 586.26 | 4.0847 | in both testings made: **F > Fcritical** and there is a 0.000 probability that Ho3 will likely to be true; **Ho3 is rejected** | under normal communications, the group mean of the fractal dimensions of the voice frequency patterns is significantly different among volunteers. The fractal dimension of this voice frequency pattern can be distinctly correlated to the person. |
| | for 30 ms tele; uFD1 = uFD2 = uFD3= uFD4= uFD5= uFD6= uFD7= uFD8= uFD9= uFD10 | 0.000 | 9 | 7.55 | 4.0847 | | |
| *Ho4: There is no significant difference on the fractal dimensions of the sinusoidal wave pattern of the sound generated on the syllable "PASS" among test volunteers.* | for "PASS" dir; uFD1 = uFD2 = uFD3= uFD4= uFD5= uFD6= uFD7= uFD8= uFD9= uFD10 | 0.000 | 9 | 16.75 | 4.0847 | in both testings made: **F > Fcritical** and there is a 0.000 probability that Ho4 will likely to be true; **Ho4 is rejected** | Under normal communications, the group mean of the fractal dimensions of the "PASS" wave pattern samples are significantly different among volunteers. The fractal dimension of the "PASS" sample can be distincly attributed to the person. |
| | for "PASS" tele; uFD1 = uFD2 = uFD3= uFD4= uFD5= uFD6= uFD7= uFD8= uFD9= uFD10 | 0.000 | 9 | 7.55 | 4.0847 | | |

## 6.0 Conclusions

This paper explores the use of fractals dimension in characterizing the different voice wave patterns of a speaker for voice recognition. The study tested hypotheses base on the phonemic samples "AH" 5-cycle, "AH" 30-millisecond and "PASS" wave sample. The phonemic samples in this study intend to define the different wave patterns of vocalization. The fractal dimensions intend to measure and characterize the wave patterns for voice recognition. The "AH" 5–cycle sample intend to illustrate the sinusoid pattern created by the vocal source and tract of a person as it vocalized the phoneme "AH". The "AH" 30-millisecond sample intends to exhibit the voice frequency pattern during normal communications of the phoneme "AH". The "PASS" wave pattern samples intend to show the speech dynamics as it is vocalized by the person. In the respective tests, the following conclusions are forwarded:

1. *The fractal dimensions of "AH" 5-cycle wave pattern as result of the first hypothesis (Ho1) tests:*

The group mean of the fractal dimensions of the "AH" 5-cycle samples of a volunteer is significantly different from the others. **Hence the fractal dimension measures and distinctly differentiates the "AH" 5-cycle wave pattern of the individual from other wave patterns of different source. The fractal dimension of "AH" 5-cycle wave pattern signifies to a vocal source and tract of a person.**

To apply this to real practice and for real time voice recognition, a voice data bank consisting of fractal dimensions of the voice wave patterns from different individuals is necessary for verification and identification reference (Rabiner, 1989). At any specific time any word will be spoken and phonemes may be vocalized. Different phonemes have different wave patterns. The vowel phonemes are an ideal identifying element. These are always vocalized in any speech. A phoneme lattice characterization of speech (Anusuya, M. A., & Katti, S. K., 2010) using the fractal dimensions, will be highly recommended as an identifying signature of the speaker.

2. *The result of the two-way ANOVA of the second factor media (Ho2), and the result of the interaction of the factors from same tests,*

The group mean of the fractal dimensions of the voices communicated directly and the voices communicated through the intercom is significantly different and accordingly the result of the interactions of the two factors is also significant. **Hence the fractal dimensions of voice patterns communicated via the intercom may not identify the speaker if the database used is solely that of the fractal dimensions of the voice patterns directly communicated.**

To apply this real practice, additional measures must be taken for voice recognition to work. One option is for voice recognition to have a set of database one for each communication media like mobile phones, landline or radio, and others. The alternative is a filter system that can isolate the voice and remove the effect of the communication media. The effect of the communication media in the fractal dimension of the voice patterns makes it necessary to filter the communication system noise. This is also the reason why a filter system is a standard feature on existing voice recognition systems. The filter process will allow the verification of fractal dimensions of voice patterns to the fractal dimensions of voice patterns directly recorded database.

3. *The fractal dimensions of "AH" 30-millisecond wave pattern, as result of the hypothesis (Ho3) tests:*

The group mean of the fractal dimensions of the "AH" 30-millisecond samples from an individual is significantly different from the others. **Hence under normal conditions of communications, the fractal dimension of voice frequency wave pattern can be associated to one person.** These can be used to identify speaker.

4. *The fractal dimensions of "PASS" wave pattern samples, as result of the hypothesis (Ho4) tests:*

The group mean of the fractal dimensions of the "PASS" sample wave patterns from an individual is significantly different from the other "PASS" wave patterns of the other volunteers. **Hence, under normal conditions of communications, the "PASS" wave pattern fractal dimension is distinctive and can be associated to a person. The speech dynamic pattern is unique to the individual.** However, the large diversity and variations of speech dynamics will require a considerable number of categorizations and a large database for recognition.

This paper has shown the effective use of fractal dimensions in measuring and categorizing wave patterns. To develop a voice recognition system will require an algorithm for voice characterization and identification. This would also require a database system that can handle the different variations of wave patterns due to the effect of the communication system or otherwise a filter system that can remove its effects. An added challenge to real time voice recognition is the provision of real-time fractal analysis of speech.

**References:**


Allen, J. F. (2003).CSC248, Lec 12: Approaches to Speech Recognition. (Retrieved date April 20, 2014, from: Hajim School of Engineering and Applied Science Department of Computer Science: http://www.cs.rochester.edu/u/james/CSC248/lec12.pdf)

Anusuya, M. A., & Katti, S. K. (2010). Speech recognition by machine, a review. arXiv preprint arXiv:1001.2267.

Austerlitz, H. (2002). Data acquisition techniques using PCs. Academic press.

Berg, R. E., & Stork, D. G. (1982). The physics of sound. Pearson Education India.

Charlton, G. "Problems Cancelling a Hotel Booking." http://www.telegraph.co.uk/travel/travel-advice/9847658/Problems-cancelling-a-hotel-booking.html. Travel. February 4, 2013. Web. September 20, 2013

Charlton, G. "Resolving Hotel Booking Issues With Expedia Call-Centre Staff." http://www.telegraph.co.uk/travel/columnists/gillcharlton/9333560/Resolving-hotel-booking-issues-with-Expedia-call-centre-staff.html. Travel. June 15, 2012. Web. September 20, 2013

Chavez, S. (2013, April ,). Speech Recognition: A Work in Progress. For the Record, 25(Special Showcase Edition). Spring City, California, USA: Great Valley Publishing Co., Inc. Retrieved june 28, 2014, from http://www.fortherecordmag.com/archives/0413bonusp10.shtml

De Smedt, K. (1996). Computional models of incremental grammatical encoding. In A.

Dijkstra & K. de Smedt (eds.) (1996). Computational psycholinguistics: AI and connectionist models of human language processing (pp. 24-48). London: Taylor & Francis, 1996.

Fractal Analysis System ver 3.4.7 (Fractal3E) downloaded from http://cse.naro.affrc.go.jp/ sasaki /index-e.html

Free Audio Editor version: 2014 8.6.1 downloaded from http://www.free-audio-editor.com/

Furtună, T. F. (2008). Dynamic programming algorithms in speech recognition. Revista Informatica Economică nr, 2(46), 94.

Gershenson, C. (2003). Artificial neural networks for beginners. arXiv preprint cs/0308031.

Hemdal, J. F., & Hughes, G. W. (1967). A feature based computer recognition program for the modeling of vowel perception. Models for the Perception of Speech and Visual Form, Wathen-Dunn, W. Ed. MIT Press, Cambridge, MA.

Hecht-Nielsen, R. (1989, June). Theory of the backpropagation neural network. In Neural Networks, 1989. IJCNN., International Joint Conference on (pp. 593-605). IEEE.

Itakura, F. (1975). Minimum prediction residual principle applied to speech recognition. Acoustics, Speech and Signal Processing, IEEE Transactions on, 23(1), 67-72.

Knight, W. (2012, May 29). Business Report :Where Speech Recognition is Going. Retrieved june 28, 2014, from MIT Technology Review: http://www.technologyreview.com

Mandelbrot, B. B. (1983). The fractal geometry of nature. Macmillan.

Melim, P., Urias, J., Solano, D., Soto, M., Lopez, M., & Castillo, O. ( 2006). Voice Recognition with Neural Networks, Type-2 Fuzzy Logic and Genetic Algorithms. Engineering Letters, 13:2.

Melim, P., & Castillo, O. (2005). Voice recognition with neural networks, fuzzy logic and genetic algorithms. In Hybrid Intelligent Systems for Pattern Recognition Using Soft Computing (pp. 223-240). Springer Berlin Heidelberg.

Moore, R. K. (1994, September). Twenty things we still don't know about speech. In by H. Niemann, R. De Mori, and G. Hanrieder (infix, St. Augustin) (Vol. 9, p. 17).

Padmanabhan, R. (2012). Studies on voice activity detection and feature diversity for speaker recognition (Doctoral dissertation, INDIAN INSTITUTE OF TECHNOLOGY, MADRAS).

Perrachione, T. K., Del Tufo, S. N., & Gabrieli, J. D. (2011). Human voice recognition depends on language ability. Science, 333(6042), 595-595.

Pogue, D. (2010, November 17). Talk to the Macnihes:Progress in Speech Recognition Software. Retrieved June 28, 2014, from Scientific American: http://www.scientificamerican.com/article/talk-to-the-machine/

Rabiner, L. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. Proceedings of the IEEE, 77(2), 257-286.

Rabiner, L. R., Juang, B. H., & Lee, C. H. (1996). An overview of automatic speech recognition. In Automatic Speech and Speaker Recognition (pp. 1-30). Springer US.

Reynolds, D. A. (1995). Automatic speaker recognition using Gaussian mixture speaker models. In The Lincoln Laboratory Journal.

Scherz, Paul. (2006, Nov 14) Practical Electronics for Inventors. ed. McGraw-Hill.

Soloman, S. (2009). Sensors handbook. McGraw-Hill, Inc..

Swee, L. H. (1998). Implementing Speech-Recognition Algorithms on the. USA: Texas Instrument.

Zmeskal, O. Bzatec, T., Nezadal, M. & Buchnicek, M. (2013), Image Fundamentals: HarFA Basics, (Retrieved November 11 2013 from: http://www.fch.vutbr.cz/lectures/imagesci.)

## APPENDIX

**Table 3a. Fractal analysis and wave properties of the "AH" 5-cycle samples direct**

| | wave properties | | | | FD 5 cycle direct | |
|---|---|---|---|---|---|---|
| sample | period ms | mean f \| std dev | frequency | mean f \| std dev | fd / 5 cyle | mean \| std dev |
| ar1 | 5.4700 | 5.542000 | 182.815356 | 180.537937 | 1.638300 | 1.655200 |
| ar2 | 5.7600 | 0.144983 | 173.611111 | 4.667523 | 1.622200 | 0.024060 |
| ar3 | 5.6000 | | 178.571429 | | 1.678100 | |
| ar4 | 5.3800 | | 185.873606 | | 1.673800 | |
| ar5 | 5.5000 | | 181.818182 | | 1.663600 | |
| ja1 | 4.4600 | 4.754000 | 224.215247 | 210.704534 | 1.575900 | 1.611080 |
| ja2 | 4.5900 | 0.216056 | 217.864924 | 9.775834 | 1.634600 | 0.029835 |
| ja3 | 4.8600 | | 205.761317 | | 1.629200 | |
| ja4 | 4.9400 | | 202.429150 | | 1.634500 | |
| ja5 | 4.9200 | | 203.252033 | | 1.581200 | |
| jef1 | 7.1000 | 7.210000 | 140.845070 | 138.743022 | 1.676500 | 1.679780 |
| jef2 | 7.0500 | 0.148535 | 141.843972 | 2.837982 | 1.643400 | 0.021927 |
| jef3 | 7.2000 | | 138.888889 | | 1.698300 | |
| jef4 | 7.4250 | | 134.680135 | | 1.693700 | |
| jef5 | 7.2750 | | 137.457045 | | 1.687000 | |
| jen1 | 4.1200 | 4.532000 | 242.718447 | 221.466320 | 1.601900 | 1.623660 |
| jen2 | 4.3000 | 0.302192 | 232.558140 | 15.250193 | 1.628500 | 0.020579 |
| jen3 | 4.7000 | | 212.765957 | | 1.602200 | |
| jen4 | 4.7600 | | 210.084034 | | 1.642800 | |
| jen5 | 4.7800 | | 209.205021 | | 1.642900 | |
| jex1 | 4.2600 | 5.108000 | 234.741784 | 198.447491 | 1.684700 | 1.695780 |
| jex2 | 5.0000 | 0.661604 | 200.000000 | 25.968743 | 1.671700 | 0.020168 |
| jex3 | 4.7800 | | 209.205021 | | 1.718200 | |
| jex4 | 5.5400 | | 180.505415 | | 1.715200 | |
| jex5 | 5.9600 | | 167.785235 | | 1.689100 | |
| jol1 | 4.1800 | 4.238000 | 239.234450 | 236.019766 | 1.659000 | 1.666620 |
| jol2 | 4.3100 | 0.074967 | 232.018561 | 4.198155 | 1.625700 | 0.033742 |
| jol3 | 4.2600 | | 234.741784 | | 1.678600 | |
| jol4 | 4.1400 | | 241.545894 | | 1.716600 | |
| jol5 | 4.3000 | | 232.558140 | | 1.653200 | |
| lal1 | 4.2000 | 4.754000 | 238.095238 | 211.582282 | 1.661500 | 1.681720 |
| lal2 | 4.4900 | 0.398472 | 222.717149 | 18.412333 | 1.633500 | 0.035282 |
| lal3 | 5.1400 | | 194.552529 | | 1.688000 | |
| lal4 | 5.0600 | | 197.628458 | | 1.724500 | |
| lal5 | 4.8800 | | 204.918033 | | 1.701100 | |
| p1 | 3.5200 | 3.437600 | 284.090909 | 293.866041 | 1.626500 | 1.638900 |
| p2 | 4.0800 | 0.400348 | 245.098039 | 31.973710 | 1.671200 | 0.025300 |
| p3 | 3.0200 | | 331.125828 | | 1.654500 | |
| p4 | 3.2900 | | 303.951368 | | 1.605500 | |
| p5 | 3.2780 | | 305.064063 | | 1.636800 | |
| jaz1 | 3.6000 | 4.188000 | 277.777778 | 240.268028 | 1.644900 | 1.635020 |
| jaz2 | 4.1200 | 0.353086 | 242.718447 | 22.126491 | 1.617800 | 0.010597 |
| jaz3 | 4.3500 | | 229.885057 | | 1.639700 | |
| jaz4 | 4.4400 | | 225.225225 | | 1.632500 | |
| jaz5 | 4.4300 | | 225.733634 | | 1.640200 | |
| nats1 | 7.5250 | 8.179000 | 132.890365 | 123.798155 | 1.709100 | 1.723240 |
| nats2 | 7.3100 | 1.057222 | 136.798906 | 14.864054 | 1.733900 | 0.021626 |
| nats3 | 9.7750 | | 102.301790 | | 1.694300 | |
| nats4 | 8.7500 | | 114.285714 | | 1.749300 | |
| nats5 | 7.5350 | | 132.714001 | | 1.729600 | |

**Table 3b. Fractal analysis and wave properties of the "AH" 5-cycle samples, intercom**

| | wave properties | | | | FD 5 cycle tele | |
|---|---|---|---|---|---|---|
| sample | period ms | mean f \| std dev | frequency | mean f \| std dev | fd / 5 cyle | mean \| std dev |
| ar1 | 5.4200 | 5.252000 | 184.501845 | 190.790224 | 1.681100 | 1.693920 |
| ar2 | 5.3900 | 0.256066 | 185.528757 | 9.912230 | 1.756700 | 0.057869 |
| ar3 | 5.3300 | | 187.617261 | | 1.725000 | |
| ar4 | 5.3200 | | 187.969925 | | 1.603200 | |
| ar5 | 4.8000 | | 208.333333 | | 1.703600 | |
| ja1 | 4.5200 | 4.718000 | 221.238938 | 212.058521 | 1.693200 | 1.700380 |
| ja2 | 4.7200 | 0.115412 | 211.864407 | 5.330346 | 1.667600 | 0.024389 |
| ja3 | 4.8100 | | 207.900208 | | 1.735600 | |
| ja4 | 4.7800 | | 209.205021 | | 1.703600 | |
| ja5 | 4.7600 | | 210.084034 | | 1.701900 | |
| jef1 | 7.3500 | 7.886000 | 136.054422 | 127.000837 | 1.685100 | 1.706080 |
| jef2 | 7.8000 | 0.339363 | 128.205128 | 5.633873 | 1.652200 | 0.036191 |
| jef3 | 8.0250 | | 124.610592 | | 1.735300 | |
| jef4 | 8.2500 | | 121.212121 | | 1.728400 | |
| jef5 | 8.0050 | | 124.921924 | | 1.729400 | |
| jen1 | 4.1200 | 4.576000 | 242.718447 | 219.255202 | 1.702000 | 1.680340 |
| jen2 | 4.5100 | 0.287106 | 221.729490 | 14.435636 | 1.704800 | 0.048971 |
| jen3 | 4.6300 | | 215.982721 | | 1.643000 | |
| jen4 | 4.7600 | | 210.084034 | | 1.616400 | |
| jen5 | 4.8600 | | 205.761317 | | 1.735500 | |
| jex1 | 5.1800 | 5.584000 | 193.050193 | 179.702259 | 1.650900 | 1.672180 |
| jex2 | 5.3700 | 0.368822 | 186.219739 | 11.723741 | 1.712100 | 0.040189 |
| jex3 | 5.4200 | | 184.501845 | | 1.704700 | |
| jex4 | 5.9400 | | 168.350168 | | 1.678600 | |
| jex5 | 6.0100 | | 166.389351 | | 1.614600 | |
| jol1 | 4.4600 | 4.762000 | 224.215247 | 210.348029 | 1.686200 | 1.705160 |
| jol2 | 4.6900 | 0.216956 | 213.219616 | 9.671437 | 1.714200 | 0.012109 |
| jol3 | 4.7300 | | 211.416490 | | 1.708600 | |
| jol4 | 4.9000 | | 204.081633 | | 1.715900 | |
| jol5 | 5.0300 | | 198.807157 | | 1.700900 | |
| lal1 | 5.0800 | 5.052000 | 196.850394 | 199.189374 | 1.622500 | 1.635160 |
| lal2 | 5.1700 | 0.435511 | 193.423598 | 18.159088 | 1.618800 | 0.013912 |
| lal3 | 5.0600 | | 197.628458 | | 1.640900 | |
| lal4 | 5.5800 | | 179.211470 | | 1.642200 | |
| lal5 | 4.3700 | | 228.832952 | | 1.651400 | |
| p1 | 3.8900 | 3.716000 | 257.069409 | 270.080923 | 1.692700 | 1.689240 |
| p2 | 3.9500 | 0.246840 | 253.164557 | 18.339418 | 1.711800 | 0.035826 |
| p3 | 3.4200 | | 292.397661 | | 1.712500 | |
| p4 | 3.4800 | | 287.356322 | | 1.626800 | |
| p5 | 3.8400 | | 260.416667 | | 1.702400 | |
| jaz1 | 3.7400 | 4.490000 | 267.379679 | 224.696760 | 1.711900 | 1.730120 |
| jaz2 | 4.5400 | 0.447996 | 220.264317 | 24.869563 | 1.730100 | 0.013132 |
| jaz3 | 4.5700 | | 218.818381 | | 1.747000 | |
| jaz4 | 4.9400 | | 202.429150 | | 1.736800 | |
| jaz5 | 4.6600 | | 214.592275 | | 1.724800 | |
| nats1 | 4.5000 | 4.456000 | 222.222222 | 225.003442 | 1.679400 | 1.727540 |
| nats2 | 4.3500 | 0.250460 | 229.885057 | 13.063395 | 1.753800 | 0.029414 |
| nats3 | 4.0800 | | 245.098039 | | 1.726200 | |
| nats4 | 4.6900 | | 213.219616 | | 1.748700 | |
| nats5 | 4.6600 | | 214.592275 | | 1.729600 | |

**Table 4. Fractal analysis and wave properties of the "AH" 30-millisecond samples**

| | wave properties | | FD 30 milliseconds direct | | wave properties | | FD 30 milliseconds tele | |
|---|---|---|---|---|---|---|---|---|
| sample | cycles | mean \| std dev | FD | mean \| std dev | cyles | mean \| std dev | FD | mean \| std dev |
| ar1 | 5.484460 | 5.416138 | 1.244600 | 1.247880 | 5.53505 | 5.723706 | 1.348400 | 1.334140 |
| ar2 | 5.208330 | 0.140028 | 1.233600 | 0.009005 | 5.56586 | 0.297368 | 1.343000 | 0.020954 |
| ar3 | 5.357140 | | 1.253200 | | 5.62852 | | 1.331000 | |
| ar4 | 5.576210 | | 1.255800 | | 5.63910 | | 1.349300 | 0.049078 |
| ar5 | 5.454550 | | 1.252200 | | 6.25000 | | 1.299000 | |
| ja1 | 6.726460 | 6.321136 | 1.274400 | 1.247120 | 6.63717 | 6.361756 | 1.313800 | 1.312520 |
| ja2 | 6.535950 | 0.293277 | 1.269300 | 0.023313 | 6.35593 | 0.159911 | 1.278900 | 0.023550 |
| ja3 | 6.172840 | | 1.221600 | | 6.23701 | | 1.343300 | |
| ja4 | 6.072870 | | 1.235000 | | 6.27615 | | 1.305000 | |
| ja5 | 6.097560 | | 1.235300 | | 6.30252 | | 1.321600 | |
| jef1 | 4.225350 | 4.162290 | 1.243800 | 1.246100 | 4.08163 | 3.810024 | 1.287100 | 1.303320 |
| jef2 | 4.255320 | 0.085141 | 1.250200 | 0.006700 | 3.84615 | 0.169015 | 1.309400 | 0.011536 |
| jef3 | 4.166670 | | 1.254800 | | 3.73832 | | 1.306800 | |
| jef4 | 4.040400 | | 1.237200 | | 3.63636 | | 1.296700 | |
| jef5 | 4.123710 | | 1.244500 | | 3.74766 | | 1.316600 | |
| jen1 | 7.281550 | 6.643988 | 1.267600 | 1.258800 | 7.28155 | 6.577656 | 1.305700 | 1.323680 |
| jen2 | 6.976740 | 0.457504 | 1.267700 | 0.008184 | 6.65189 | 0.433068 | 1.333700 | 0.010790 |
| jen3 | 6.382980 | | 1.254300 | | 6.47948 | | 1.328200 | |
| jen4 | 6.302520 | | 1.253600 | | 6.30252 | | 1.322600 | |
| jen5 | 6.276150 | | 1.250800 | | 6.17284 | | 1.328200 | |
| jex1 | 7.042250 | 5.953424 | 1.715200 | 1.703500 | 5.79151 | 5.391068 | 1.310900 | 1.299760 |
| jex2 | 6.000000 | 0.779061 | 1.703800 | 0.007819 | 5.58659 | 0.351712 | 1.319300 | 0.014822 |
| jex3 | 6.276150 | | 1.696400 | | 5.53505 | | 1.295300 | |
| jex4 | 5.415160 | | 1.705800 | | 5.05051 | | 1.288800 | |
| jex5 | 5.033560 | | 1.696300 | | 4.99168 | | 1.284500 | |
| jol1 | 7.177030 | 7.080592 | 1.311200 | 1.306680 | 6.72646 | 6.310444 | 1.335900 | 1.311000 |
| jol2 | 6.960560 | 0.125945 | 1.295800 | 0.018162 | 6.39659 | 0.290143 | 1.297600 | 0.015114 |
| jol3 | 7.042250 | | 1.303400 | | 6.34250 | | 1.312500 | |
| jol4 | 7.246380 | | 1.335200 | | 6.12245 | | 1.308200 | |
| jol5 | 6.976740 | | 1.287800 | | 5.96422 | | 1.300800 | |
| lal1 | 7.142860 | 6.347468 | 1.333200 | 1.303180 | 5.90551 | 5.975680 | 1.332300 | 1.324500 |
| lal2 | 6.681510 | 0.552370 | 1.294200 | 0.017676 | 5.80271 | 0.544774 | 1.318600 | 0.013525 |
| lal3 | 5.836580 | | 1.291100 | | 5.92885 | | 1.314700 | |
| lal4 | 5.928850 | | 1.292300 | | 5.37634 | | 1.312500 | |
| lal5 | 6.147540 | | 1.305100 | | 6.86499 | | 1.344400 | |
| p1 | 8.522730 | 8.815982 | 1.664700 | 1.666640 | 7.71208 | 8.102428 | 1.396100 | 1.395460 |
| p2 | 7.352940 | 0.959213 | 1.666100 | 0.011371 | 7.59494 | 0.550182 | 1.413800 | 0.030940 |
| p3 | 9.933780 | | 1.657400 | | 8.77193 | | 1.415200 | |
| p4 | 9.118540 | | 1.685900 | | 8.62069 | | 1.341800 | |
| p5 | 9.151920 | | 1.659100 | | 7.81250 | | 1.410400 | |
| jaz1 | 8.333330 | 7.208040 | 1.298330 | 1.247686 | 8.02139 | 6.740902 | 1.378100 | 1.369540 |
| jaz2 | 7.281550 | 0.663793 | 1.255700 | 0.032328 | 6.60793 | 0.746088 | 1.388500 | 0.013884 |
| jaz3 | 6.896550 | | 1.212500 | | 6.56455 | | 1.353800 | |
| jaz4 | 6.756760 | | 1.231600 | | 6.07287 | | 1.360400 | |
| jaz5 | 6.772010 | | 1.240300 | | 6.43777 | | 1.366900 | |
| nats1 | 3.986710 | 3.713944 | 1.249300 | 1.250660 | 6.66667 | 6.750104 | 1.259200 | 1.341820 |
| nats2 | 4.103970 | 0.445924 | 1.270000 | 0.011446 | 6.89655 | 0.391900 | 1.352000 | 0.055829 |
| nats3 | 3.069050 | | 1.249600 | | 7.35294 | | 1.398800 | |
| nats4 | 3.428570 | | 1.240800 | | 6.39659 | | 1.382500 | |
| nats5 | 3.981420 | | 1.243600 | | 6.43777 | | 1.316600 | |

**Table 5. Fractal analysis and wave properties of the "PASS" samples**

| | | | "PASS" | |
|---|---|---|---|---|
| sample | millisec | mean\|std dev | FD "pass" | mean\|std dev |
| ar1 | 253.50000 | 0.266700 | 1.557700 | 1.544740 |
| ar2 | 277.00000 | 0.009162 | 1.523600 | 0.019265 |
| ar3 | 272.00000 | | 1.523700 | |
| ar4 | 269.00000 | | 1.559500 | |
| ar5 | 262.00000 | | 1.559200 | |
| ja1 | 252.00000 | 0.271200 | 1.544000 | 1.520720 |
| ja2 | 267.00000 | 0.018213 | 1.534100 | 0.028097 |
| ja3 | 264.00000 | | 1.472300 | |
| ja4 | 301.00000 | | 1.530000 | |
| ja5 | 272.00000 | | 1.523200 | |
| jef1 | 243.00000 | 0.205600 | 1.538500 | 1.537400 |
| jef2 | 182.00000 | 0.023215 | 1.526400 | 0.024213 |
| jef3 | 206.50000 | | 1.578900 | |
| jef4 | 191.50000 | | 1.520800 | |
| jef5 | 205.00000 | | 1.522400 | |
| jen1 | 254.00000 | 0.284750 | 1.557100 | 1.525900 |
| jen2 | 308.00000 | 0.045305 | 1.543900 | 0.024160 |
| jen3 | 248.00000 | | 1.507200 | |
| jen4 | 260.00000 | | 1.500100 | |
| jen5 | 353.75000 | | 1.521200 | |
| jex1 | 326.00000 | 0.334700 | 1.671100 | 1.643680 |
| jex2 | 362.50000 | 0.022742 | 1.622500 | 0.022817 |
| jex3 | 355.00000 | | 1.640000 | |
| jex4 | 311.00000 | | 1.663200 | |
| jex5 | 319.00000 | | 1.621600 | |
| jol1 | 225.50000 | 0.233200 | 1.579800 | 1.589360 |
| jol2 | 233.00000 | 0.018016 | 1.570200 | 0.014831 |
| jol3 | 218.00000 | | 1.599300 | |
| jol4 | 264.00000 | | 1.607300 | |
| jol5 | 225.50000 | | 1.590200 | |
| lal1 | 289.00000 | 0.271400 | 1.649400 | 1.620300 |
| lal2 | 298.00000 | 0.029228 | 1.625000 | 0.026790 |
| lal3 | 289.00000 | | 1.641700 | |
| lal4 | 250.00000 | | 1.595900 | |
| lal5 | 231.00000 | | 1.589500 | |
| p1 | 292.00000 | 0.323650 | 1.611500 | 1.610340 |
| p2 | 353.75000 | 0.027043 | 1.631300 | 0.021438 |
| p3 | 347.50000 | | 1.610600 | |
| p4 | 322.50000 | | 1.623100 | |
| p5 | 302.50000 | | 1.575200 | |
| jaz1 | 246.50000 | 0.287700 | 1.590400 | 1.580340 |
| jaz2 | 292.00000 | 0.025631 | 1.597700 | 0.016180 |
| jaz3 | 317.00000 | | 1.575200 | |
| jaz4 | 288.00000 | | 1.582800 | |
| jaz5 | 295.00000 | | 1.555600 | |
| nats1 | 258.00000 | 0.235000 | 1.533800 | 1.545460 |
| nats2 | 258.00000 | 0.026514 | 1.557400 | 0.032138 |
| nats3 | 227.00000 | | 1.580600 | |
| nats4 | 238.00000 | | 1.559300 | |
| nats5 | 194.00000 | | 1.496200 | |

| | | | "PASS" telephone | |
|---|---|---|---|---|
| sample | ms | mean\|std dev | FD "pass" | mean\|std dev |
| ar1 | 247.000000 | 227.600000 | 1.480300 | 1.524740 |
| ar2 | 226.000000 | 12.239281 | 1.508800 | 0.037130 |
| ar3 | 227.000000 | | 1.531900 | |
| ar4 | 225.000000 | | 1.581500 | |
| ar5 | 213.000000 | | 1.521200 | |
| ja1 | 300.000000 | 288.200000 | 1.490600 | 1.509940 |
| ja2 | 277.000000 | 17.253985 | 1.468300 | 0.032271 |
| ja3 | 267.000000 | | 1.514600 | |
| ja4 | 287.000000 | | 1.553100 | |
| ja5 | 310.000000 | | 1.523100 | |
| jef1 | 316.000000 | 299.600000 | 1.456800 | 1.442220 |
| jef2 | 335.000000 | 81.174696 | 1.442600 | 0.019405 |
| jef3 | 358.750000 | | 1.416400 | |
| jef4 | 331.250000 | | 1.464500 | |
| jef5 | 157.000000 | | 1.430800 | |
| jen1 | 256.000000 | 322.550000 | 1.492900 | 1.504960 |
| jen2 | 350.000000 | 40.470823 | 1.516500 | 0.016955 |
| jen3 | 353.750000 | | 1.482600 | |
| jen4 | 340.000000 | | 1.523800 | |
| jen5 | 313.000000 | | 1.509000 | |
| jex1 | 390.000000 | 376.500000 | 1.742000 | 1.729760 |
| jex2 | 392.500000 | 20.811655 | 1.722000 | 0.007409 |
| jex3 | 355.000000 | | 1.728800 | |
| jex4 | 392.500000 | | 1.729100 | |
| jex5 | 352.500000 | | 1.726900 | |
| jol1 | 303.500000 | 326.200000 | 1.510700 | 1.453920 |
| jol2 | 352.500000 | 27.664508 | 1.460400 | 0.034297 |
| jol3 | 340.000000 | | 1.425600 | |
| jol4 | 345.000000 | | 1.432800 | |
| jol5 | 290.000000 | | 1.440100 | |
| lal1 | 178.000000 | 293.600000 | 1.731300 | 1.734240 |
| lal2 | 352.500000 | 83.840175 | 1.748400 | 0.018679 |
| lal3 | 355.000000 | | 1.742000 | |
| lal4 | 352.500000 | | 1.703000 | |
| lal5 | 230.000000 | | 1.746500 | |
| p1 | 287.000000 | 334.450000 | 1.779900 | 1.787280 |
| p2 | 269.000000 | 53.598741 | 1.775600 | 0.014513 |
| p3 | 352.500000 | | 1.778700 | |
| p4 | 373.750000 | | 1.791200 | |
| p5 | 390.000000 | | 1.811000 | |
| jaz1 | 327.500000 | 328.500000 | 1.542600 | 1.526480 |
| jaz2 | 342.500000 | 25.285866 | 1.527600 | 0.017128 |
| jaz3 | 345.000000 | | 1.522600 | |
| jaz4 | 342.500000 | | 1.499700 | |
| jaz5 | 285.000000 | | 1.539900 | |
| nats1 | 232.000000 | 292.900000 | 1.684800 | 1.732260 |
| nats2 | 306.000000 | 45.571921 | 1.793100 | 0.043552 |
| nats3 | 357.500000 | | 1.746000 | |
| nats4 | 292.000000 | | 1.741900 | |
| nats5 | 277.000000 | | 1.695500 | |