ISSN: 2413-6999
Journal of Information, Communication, and Intelligence Systems (JICIS)
Volume 2, Issue 1, February 2016

Implementation of Log Likelihood (Itakura) Method for Distortion Measuring between Human Sound Vectors

Oday K. Hamid

***Abstract*— Measuring distance is used widely in speaker recognition or speech recognition system because both systems need to find the distance between the testing vectors and training vectors (how much they are closer).This method are useful for solving a variety of signal processing problems such as detection , segmentation , classification , recognition , or coding. The advantage of distance measure is to compare the shapes of two spectral vectors we need to inter them into distance measure program in order to indicate the similarity or difference between them. The distance between two feature sound vectors in LPC spectral shape (linear predictive coding) with order P=16 were measured by using Log-Likelihood distance measure method and an accuracy result was obtained.**

Index Terms— Itkura, Log Liklihood, Sound vectors.

I. Introduction

Distance measures between statistical models or between a model and observations are widely used concepts in signal processing (and in automatic control) for solving various problems such as detection, automatic segmentation, classification, pattern recognition, coding (model validation, choice of optimal input signals for system identification).

The studies concerning distance measures are basically of two types, apart from those of probabilities and statistication, on one hand there are general studies for the computing of error probabilities in classification problems (of any objects characterized by any measurements) . without taking into account neither the nature of the parameters which characterized the probability laws nor the way by which they have been estimated. On the other hand there are a lot of specific studies in the speech processing domain (coding, recognition) where refinements of ITAKURA or spectral distance measures still emerge now [1].

II. Theory

The methods used in classification could be categorized as geometric . topology and probabilistic: the three methods are best illustrated when the test and reference patterns are viewed as point in multi-dimensional space. The methods are explained with an example in a 2-dimensional space as in figure (1)

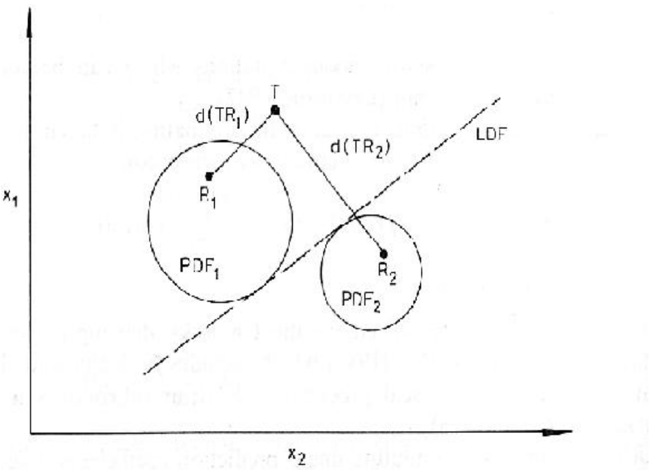

Fig.1. 2-dimentional space of training vectors

Geometric method: divides space into regions (with each class in one region) with bounder. These boundaries are defined by linear discriminate functions because it line on the same side of the linear discriminate function (LDF) as R1.
In topological method: each class is represented by one or more points in the space. The distance between the test vector point and each class is determined and the test vector is assigned to the class with the shortest distance T is classified as R1 because the distance from T to R1 is less than distance to R2.

In probabilistic method: a probability density function: is defined for each point in the space. The test pattern is assigned to the class which has the greatest PDF at that point is classified as R1 because the probability density function PDF1 at T is greater than PDF2. Maximum likelihood estimation endeavors to find the most “likely" values of distribution parameters for a set of data by maximizing the value of what is called the "likelihood function”. This likelihood function is largely based on the probability density function (pdf) for given distribution. As an example, consider a generic [2].

Oday K .Hamid, Dept. of Computer Techniques Engineering, Dijlah University College, (e-mail: qwera9222@gmail.com). Baghdad, Iraq.

III. DISTORTION MEASURE

A key component of most pattern-comparison algorithm is a prescribed measurement of dissimilarity between two feature vectors. This measurement of dissimilarity can be handled with mathematical rigor if the patterns are visualized in a vector space.

Assume we have two feature vectors x,y and the distance between them is d(x,y) ,so the following properties should be satisfied [3]&[4]:

A.symmetry d(x,y)=d(y,x)

B.positive definiteness d(x,y)>0 for x≠y

d(x,y)=0 for x=y

C. the distance measure d(x,y) should have physically meaningful interpretation in the frequency domain and should be analytically tractable.

D. it should be possible to efficiently evaluate d(x,y).

E. if Z is another arbitrary data vector, then it should satisfy the triangular inequality:

d(x,y)≤ d(x,z)+d(y,z)

That is it should be metric. If a measure of a difference (d) satisfies only the positive definiteness property we customarily call it a distortion measure particularly when the vectors are representations of signal spectra [3] as shown in figure (2).

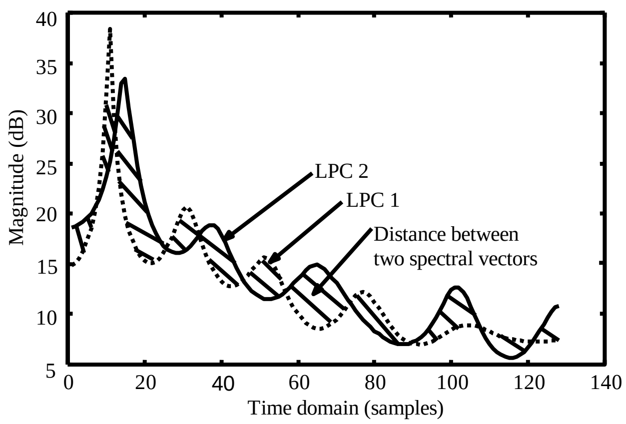

Fig.2. Frequency domain difference between two spectral vectors bandwidths

In many researches that need a large number of comparisons between vectors so we must choose a good distortion measure drive the processes to the goal [5]. To be useful, a distortion measure must process to a certain degree the following properties [6].

1. It must be subjectively meaningful, in the sense that small and large distortion correspond to good and bad subjective quality, respectively.

2. It must be tractable, in the sense that it is amenable to mathematical analysis and leads to practical design techniques.

3. It must be computable, in the sense that the actual distortions resulting in a real system can be efficiently computed.

IV. DISTORTION MEASURES TYPES:

There are several types of spectral distortion measures that are designed to measure the dissimilarity or distance between two spectra of speech and in this research some types are introduced:

1. Ecludien Distance Measure

This distance measure can only be applied if all variables are continuous. The Ecludien distance between two point is clearly defined .the distance between two vectors is here defined by the Ecludien distance between two vectors centers .A vector center is defined as the vector of cluster means of each variable .

This type widely used in distortion measure were distortion between training vector (Ar) and test vector (At) can be represented as [4]:

$$DS_{EC}(At,Ar) = [\sum_{i=1}^{N} |At_i - Ar_i|^q]^{1/q} \quad (1)$$

In most application that use this type of distortion measure makes the value of (q) equal (2) then the equation become:

$$DS_{EC}(At,Ar) = \left[\sum_{i=1}^{N} |At_i - Ar_i|^2\right]^{1/2} \quad (2)$$

But if q=1 then the distance , called , city block and equation become:

$$DS_{EC}(At,Ar) = \sum_{i=1}^{N} |At_i - Ar_i| \quad (3)$$

Since the square root of number is proportional with this number , then eq.can be simplified to:

$$DS_{EC}(At,Ar) = (At-Ar)^T (At-Ar) = \sum_{i=1}^{N} |At_i - Ar_i|^2 \quad (4)$$

The last equation reduces the time that needs in computation and it is called mean square distance between two shapes.

In recognition test ,it has been shown that poor results were obtained using Ecludien distance measure directly on linear predictive coding (LPC) coefficients or on the reflection coefficients[7] .

2. Maximum Likelihood Distance Measure

Its also defined ITAKURA –Satio distance measure method.

It was represented in this form [7]:

$$DS_{IS}(At,Ar) = \left[\frac{a_r' v_t a_r}{a_t' v_t a_t} - 1\right] + \ln\left[\frac{G_t^2}{G_r^2}\right] \quad (5)$$

Where:

Ar,At=Training and testing vectors respectively.

ISSN: 2413-6999
Journal of Information, Communication, and Intelligence Systems (JICIS)
Volume 2, Issue 1, February 2016

a_r, a_t=Coefficients of linear predictive coding of training and testing vectors.

v_t=Autocorrelation matrix

$$\left[v(i)=\frac{1}{N}\sum_{n=1}^{N-i} x(n)x(n+i) \qquad 0\le i\le p\right] \qquad (6)$$

G_r, G_t=Gain of training and testing vectors.

The MLE method has many large sample properties that make it attractive for use. It is asymptotically consistent, which means that as the sample size gets larger, the estimates converge to the true values. It is asymptotically efficient, which means that for large samples, it produces the most precise estimates. It is also asymptotically unbiased, which means that for large samples one expects to get the true value on average. The estimates themselves are normally distributed, if the sample is large enough, These are all excellent large sample properties [7] & [8].

3. Likelihood Distance Measure

This method derived from previous method where it resulted from equaling G_r and G_t and it is also called gain-normalized method , and the distance calculated from this equation [7]&[8]:-

$$DS_L(At,Ar)=\frac{a_r' V_T a_r}{a_t' V_T a_t}-1 \qquad (7)$$

4. LOG-LIKELIHOOD DISTANCE MEASURE

This method used in this paper, it is also named as ITAKURA distance measure where distance can be expressed as [9] :-

$$DS_{LL}(At,Ar)=C_r+\log_2[(b_r.r_t)/(a_t.r_t)] \qquad (8)$$

$$C_r=\log_2(a_r.a_r) \qquad (9)$$

$$b_r=\left[2\sum_{j=0}^{p-i} a_r(j)a_r(i+j)\right]/(a_r.a_r) \qquad (10)$$

$$r_t=\frac{V_t(i)}{V_t(0)} \qquad 0\le i\le p \qquad (11)$$

The log-likelihood distance measure can handle both continuous and categorical variables. It is a probability based distance. The distance between two clusters is related to the decrease in log-likelihood as they are combined into one cluster. In calculating log-likelihood, normal distributions for continuous variables and multinomial distributions for categorical variables are assumed. It is also assumed that the variables are independent of each other, and so are the cases [6] & [7].

5. The Modified Log-Likelihood Distance Measure

Where distance (DS) can be found by simple modification on ITAKURA method to reduce computation processing and space region with respect to a small energy (E_p) in LPC process and the modification can be as [8]:

*modified autocorrelation coefficients by dividing it on (E_p):

$$r_e(i)=\frac{r(i)}{E_p} \qquad 0\le i\le p \qquad (12)$$

*Squared gain term =GNT= $2\sum_{i=1}^{p} a(i,t).r(i,t)$ (13)

$$DS_{MD}(t,r)=cc_r+\log_2\left|\frac{\sum_{i=0}^{p} b(i,r).r(i,t)}{GNT}\right| \qquad (14)$$

* If GNT =E_p

$$DS_{MD}(t,r)=cc(r)+\log_2\left|\frac{\sum_{i=0}^{p} b(i,r).r(i,t)}{E_p}\right| \qquad (15)$$

And if we substitute the value of E_p from equation (12) then:

$$DS_{MD}(t,r)=cc(r)+\log_2\left|\sum_{i=0}^{p} b(i,r).r_e(i,t)\right| \qquad (16)$$

V. PRACTICAL WORK

In recognition tests the important step is the choice of a pattern similarity measure (or distortion measure) which quantitatively tells how close a reference temple is to unknown speaker.

For the LPC coefficients, the modified log-likelihood distance measure is a very powerful distortion measure based on assumed of the LPC parameter set and this method used in this paper.

VI. DATA BASE

There are two feature sound vectors in LPC spectral shape (linear predictive coding) with order P=16 which used widely in extracting or representation speaker's sound features and, each vector have size (1*16) points or numbers which represent the order of LPC. The goal is to find the distortion between them by using log likelihood distance measure method. The sound features represented in LPC vectors are extracted by special method the researcher did not explain it because it is not the goal of this paper so these vectors are interred to the distance measure program.

ISSN: 2413-6999
Journal of Information, Communication, and Intelligence Systems (JICIS)
Volume 2, Issue 1, February 2016

The proposed work examined through theoretical analysis and computer simulation using MATLAB version 6 programming language under MICROSOFT WINDOWS XP operating system.

VII. Results

In the beginning the distance between any vector and same vector was calculated to see the accuracy of this method, When each vector with order of LPC=16.

A tiny distance almost equal to zero between the vector and itself was found that mean this method gives us correct results as shown in figure (3) below:

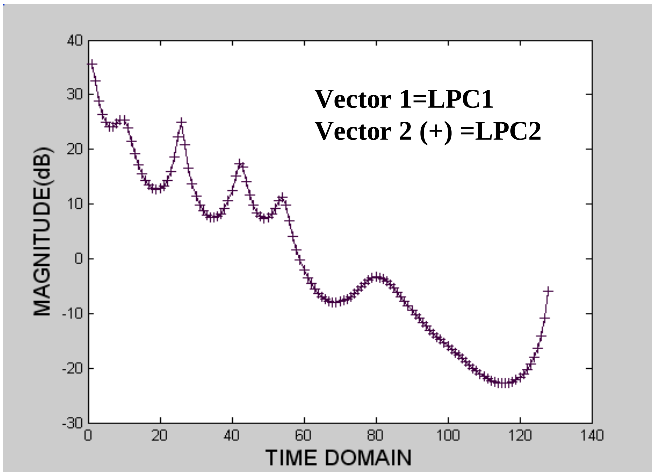

Fig.3. Frequency domain difference between 2 spectral vectors

The distortion between the vector and itself =2.4567e-012 **(near zero)**

And now the distortion between another two vectors was calculated as shown in figures (4) and (5).

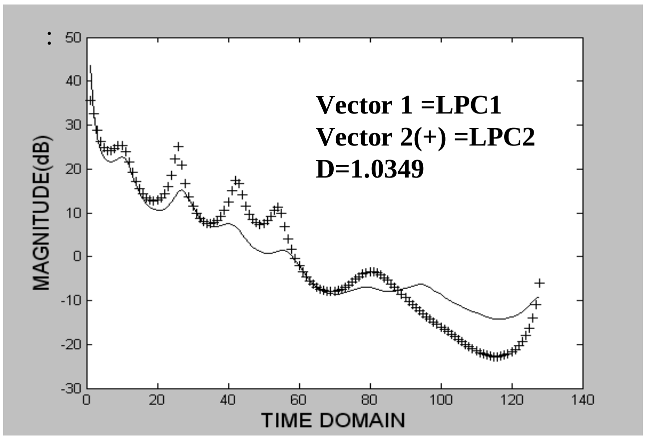

Fig.4. Frequency domain difference between 2 spectral vectors

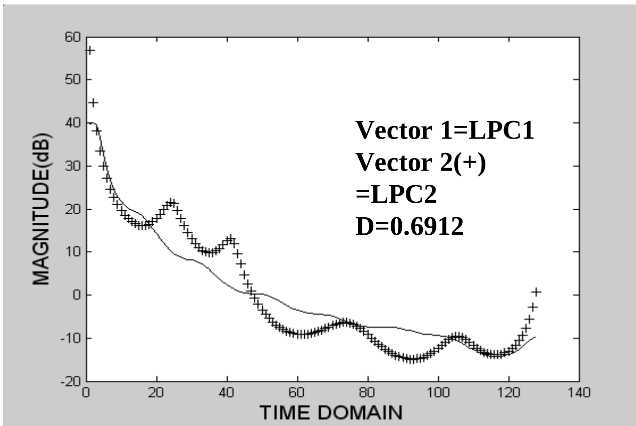

Fig.5. Frequency domain difference between 2 spectral vectors

VIII. CONCLUSION

The most important conclusions that have been reached are:

- The proposed method of measuring distortion give us a high accuracy in distortion measure between spectral shapes as shown in the results of previous section.
- The employment of the programming language (Matlab) becomes wealthy in the shorthand on the program volume and the ease of its correction.

IX. SUGGESTIONS FOR FUTURE WORK

Evaluate the distance between vectors using another distance measure method like such as Ecludien distance measure.

REFERENCES

[1] M.Bassaville ,"***Distance Measure for signal processing and pattern recognition***".

[2] A. Buzo, A. H. Gray, R. M. Gray and J. D. Markel, ***"Speech Coding Upon Vector Quantization"***, IEEE, Trans. Acoust. , Speech and Signal Process. , Vol. ASSP-28, October 1980.

[3] L. R. Rabiner and B.H. Juang "Fundamentals of speech recognition", M. N. AL-Trfi ***"Speaker recognition based"*** , prentice-Hell, New Jersey 1993.

[4] A. H. Al-Nakkash, ***"A Novel Approach For Speakers Recognition Using Vector Quantization Technique",*** M.Sc. Thesis, University of Technology, Baghdad, 2001.

[5] A. Buzo, A. H. Gray, R. M. Gray and J. D. Markel, ***"Speech Coding Upon Vector Quantization"***, IEEE, Trans. Acoust. , Speech and Signal Process. , Vol. ASSP-28, October 1980, pp. 562-574.

[6] M. J. Haider, ***"Speech Compression and Recognition Using Wavelet Transform"***, M.Sc. Thesis, College of Engineering, University of Baghdad, Baghdad, 1999.

[7] L. R. Rabinar, M. M. Sondhi and S. E. Levinson, ***"A Vector***

Quantization Combining Energy and LPC Parameters and Its Application To Isolated Word Recognition", AT&T Bel Laboratories Technical Journal, Vol. 63, No. 5, June 1984.

[8] W. A. Mahmoud, ***"Quantization Techniques For The Classification and Recognition of Speech Signals"***, Ph.D. Thesis, University of Swansea, England, 1986.

[9] F. Itakura, ***"Minimum Prediction Residual Principle Applied To SpeechRecognition"***, IEEE Trans. Acoust. , Speech and Signal Process. , Vol. ASSP-23, Feb. 1975, pp. 67-72.