LIKENESS FUNCTIONS OF THE ACOUSTIC PATTERNS AS AN INDEX FOR OBJECTIVE ESTIMATION OF SPEECH TRANSMISSION QUALITY

Czesław Bąsztura

Institute of Telecommunication and Acoustics Technical University of Wrocław ul.B.Prusa 53,55 50-317 Wrocław POLAND

ABSTRACT

The purpose of the paper is a presentation of a new objective measure for estimation of speech transmission quality and to perform a preliminary evaluation of conformity between the results obtained by means of the proposed method and subjectively measured speech intelligibility. The new method uses likeness functions of the acoustic patterns as an index for the evaluation of speech transmission quality. Eight likeness functions as distance or proximity meanures, i.e. Hamming, Euclidean, Minkowski, Chebyshev, Camber, Chi-square, Tanimoto and derectional cos were investigated. As the test signal three key phrases of natural speech were used. The preliminary results indicate the possibility of good estimation of speech transmission quality by measuring and counting the likeness functions, especially by means of Hamming, Euclidean, Minkovski and Chebyshev distance measures.

INTRODUCTION

Speech intelligibility as a measure of speech transmission quality may be classified as either subjective or objective. The subjective measurement is a procedure for determining the communication channels intelligibility using a predetermined vocabulary and selected speakers and listeners panel.
Subjective measurement techniques generaly attempt to determine intelligibility for an information presented in one of the following three forms:
a) nonsense syllables (logatoms) list
b) limited list of words,
c) list of sentences.
In subjective measurement methods the inteligibility is determined by the ability of the listeners to identify spoken (or recorded) syllables, logatoms, words or sentences. A number of subjective methods have been devised with the desirable results. However, the reguirements for listeners panels greatly restricts the utility of these methods, and a long-sought goal is to replace the subjective scoring with objective measurements. An objective measure for the fidelity of a speech communication system is a measure that is computed from data which contain no human subjective response.

There is a hypothesis that it is possible to design a relatively compact objective measures which are in a good correlation with subjective results over a subset of distortions and disturbances introduced by speech transmission channels [1,2,5,6]. Over the years some number of papers contained informations about evaluation of speech transmission systems by objective measures [1,2,7,8].

These measures include signal to noise ratios, arithmetic and geometric spectral distance measures (Viswanathan et al. [7], cepstral distance measures (Barnwell et al. [1,2]), various parametric distance measures such as pseudo-area functions and log are a fuctions from LPC analysis (Gray and Markel [5]), MTF (Steeneken and Houtgast [6]), and many more [7,8].

The task of comparing and contrasting the validity of such measures is immense. To check the validity of a particular objective measure over a given class of distortions and disturbances, a researcher must create a data base of distorted speech and corresponding data base of subjective results.

The essential features of computation of a relationship between subjective and objective measures are illustrated in Fig. 1

METHOD

In further analysis a following basic definitions and presumptions were mode:

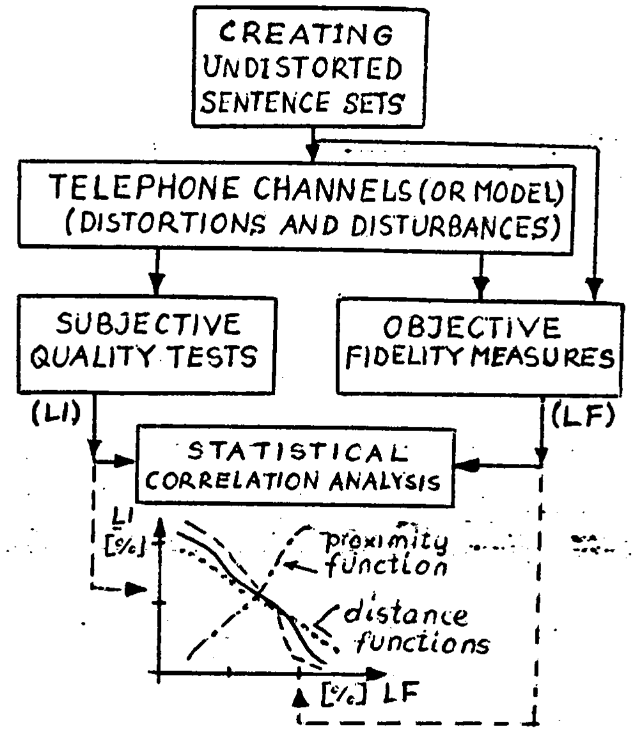

Fig.1 The computation of a relationship between objective and subjective quality measure.

a) The typical telephone channels, represented by an adjustable model of the the telephone channel as an investigated object, are used.

b) The criterion reference for objective measure is subjectively measured logatoms speech intelligibility.

c) The measurement conditions for both subjective and objective procedures are the same.

d) As a test signal for objective measurements three following Polish key phrases were used:

1. ALO (part of word "Hello"),
2. JUTRO BĘDZIE ŁADNY DZIEN ("Tomorrow will be a fine day")
3. SPRAWDZENIE PRZYDATNOSCI FUNKCJI PODOBIENSTWA DO OCENY JAKOSCI TRANSMISJI SYGNAŁU MOWY. ("Verification of the likeness functions usefulness to evaluation of speech signal transmission quality".)

e) As an objective measure, i.e. distance and proximity measures, eight likeness functions: Hamming, Euclidean, Minkovski, Chebyshev, Camber, Chi-square, Tanimoto and directional cos were examined.

The likeness function have a form :

$$d^{MIN}(X,Y)=\left[\sum_{p=1}^{P}|X_p-Y_p|^r\right]^{\frac{1}{r}} \quad r\geq 1 \qquad 1/$$

were: p=1,2,.. P p - dimensionality of vector parameters from speech signal.

Xp- p- th element of reference vector (from the undistorted speech signal)

Yp- p- th element of vector (from a distorded speech signal)

For r=1 d^{MIN} - Hamming distance
r=2 d^{MIN} - Euclidean distance.

Chebyshev distance:

$$d^{CZE}(X,Y) = \max_{p} (X_p-Y_p) \qquad 2/$$

Camber distance:

$$d^{CAM}(X,Y) = \sum_{p=1}^{P}\frac{|X_p-Y_p|}{|X_p+Y_p|} \qquad 3/$$

Chi-square distance:

$$d^{CHI}(X,Y)=\sum_{p=1}^{P}\frac{1}{X_p+Y_p}\left[\frac{X_p}{\sum_p X_p}-\frac{Y_p}{\sum_p Y_p}\right] \qquad 4/$$

Directional cos proximity:

$$b^{COS}(X,Y)= \frac{XY^{Tr}}{|X||Y|} \qquad 5/$$

Tanimoto proximity:

$$b^{TAN}(X,Y)=\frac{XY^{Tr}}{XX^{Tr}+YY^{Tr}-XY^{Tr}} \qquad 6/$$

EXPERIMENT AND RESULTS

First step in the experiment is a choise of an adequate test signals. Second problem relies on finding an effective set of parameters representing the test signal and presence of distortions and disturbances in these signals.

Next problems depend on the assumed method (Compare "METHOD").

The analysis of the previous investigations [3, 4] shows a fairly large effectiveness of representing the voice and speech features by parameter set with being the distribution of the time intervals between the zero - crossing of a speech signal [Fig.2]

$$Y \stackrel{col}{=} \{Y_1, Y_2, \dots Y_p \dots Y_P\} \qquad 7/$$

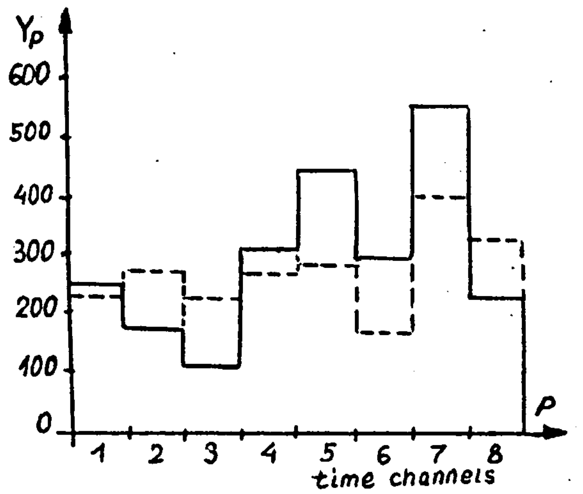

Fig.2 An example of the time intervals distribution between the zero-crossing of a speech signal (for two voices, the same text), P=8).

An especial computer program counted likeness functions (as objective measures) and made statistical correlation analysis of figure of merit LI=function of (LF).
Some of the results are shown in Fig. 3.

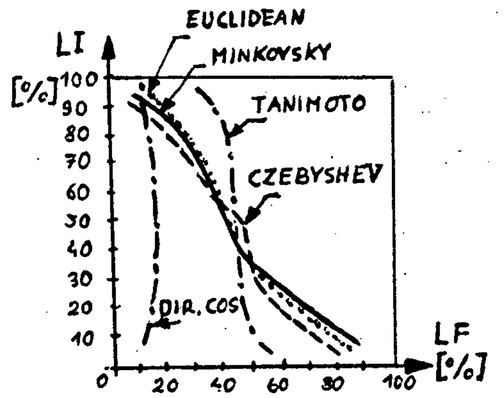

Fig.3 Examples of the statistical dependences of LI and LF.

Comparison of the results of the experiments on the 40 different telephone channels implemented by a physical model of telephone channel permits the following observations and conclusions to be given:

a) Hamming distance gives a good correlation LI with LF for all of the test signals (especially for 3 degrees of polynomial regression)

For example:

$$LI = 92{,}91 - 0.0121\, d^{HA} + 0{,}151 \cdot 10^{-5} (d^{HA})^2 - 0.631 \cdot 10^{-10} (d^{HA})^3 \qquad 8/$$

b) Euclidean and Minkovski (r=3) distance (likewise the Hamming distance) give a good correlation for third key phrase. For example:

$$LI = 90{,}2 - 0{,}0184\, d^{EU} + 0{,}364 \cdot 10^{-5} (d^{EU})^2 - 0{,}24 \cdot 10^{-9} (d^{EU})^3 \qquad 9/$$

$$LI = 89{,}6 - 0{,}0203\, d^{MIN} + 0{,}448 \cdot 10^{-5} (d^{MIN})^2 - 0{,}33 \cdot 10^{-9} (d^{MIN})^3 \qquad 10/$$

c) Chebyshev distance gives a good correlation (especjally for third key phrase and for higer degree of polynomial regression).

d) Camber distance and Tanimoto and directional cos proximity did not give satisfactory results.

The future investigations will concentrate mainly on the selection of new test signals and other parameters of the speech signal.

LITERATURE

1. Barnwell T.P." Objective measures for speech quality testing", J.A.S.A, 66 (6), Dec. 1979,pp.1658-1663.

2. Barnvell T.P. " Correlation analysis of subjective and objective measure for speech quality". Proc. ICASSP 80,New york, 1980.

3. Basztura Cz., Majewski W. " The application of long term analysis of the zero-crossing of a speech signal in automatic speaker identification " Archives of Acoustic, 3.1,1978,pp.3-15.

4. Basztura Cz., Jurkiewicz J., " The zero - crossing analysis of a speech signal in the short term method of automatic speaker identification " Archives of Acoustic, 3, 3, 1978, pp. 185-195.

5. Gray A.H., Markel J.D., "Distance measures for speech processing" IEEE Trans. on Acoustic, speech and Signal Process, ASSP - 24, No.5, 1976,pp.380-391.

6. Steeneken H.J.M., Houtgast T.," A physical method for measuring speech transmission quality ", JASA, 67 (1). Jan. 1980,pp. 318-326.

7. Viswanathan R., " Objective speech quality evaluation of narrowband LPC vocoders" Proc. IEEE, No.6, 1978, pp. 591-594.

8. Voiers W.D., " Diagnostic acceptability measure for speech communication systems". Proc. IEEE ICAPP. 1977,pp. 204-207.

9. Second Draft Proposal ISO/DP 4870 for Acoustics - Recommended method for the construction and calibration of speech intelligibility tests -October. 1976.