

Speech/Data discrimination in Communication systems

Ashok Kumar Ginni¹, Dr. K. Padma Raju², Parthraj Tripathi³

¹(Electronics & communication Department, University College of Engineering, JNT University, Kakinada, India)

²(Pincipal, University College of Engineering, JNT University, Kakinada, India)

³(Parthraj Tripathi, Scientist -C, DLRL, DRDO, Ministry of Defense, Govt. of India, India)

ABSTRACT : This paper proposes a discrimination algorithm, which discriminates speech and data on a multiplexed input signal. Commercial communication networks may use single voice band channel for transmission of both speech and data. Also, for optimum utilization of channel, the pauses in voice signal are being utilized. At receiver side the speech and data should be separately extracted, in order to send information to the respective users. For above mentioned to happen with least error, sufficient measures are to be taken for identifying the type of the signal. The speech/data discriminator is the solution for above mentioned problem. This algorithm may also be useful in the analysis of intercepted signal, where speech/data discrimination may be performed to make sure that whether the communication channel carries data or voice. After discrimination, voice will be sent to voice codec and data to the data decoder for extraction of intelligence. In this paper we proposed a simple and low complexity algorithm for speech/data discrimination based on two parameters namely short time energy and zero crossing rates. This algorithm can serve the above two purposes and gives good results in discriminating between speech and data.

Keywords - Discrimination, intercepted signal, intelligence extraction.

I. INTRODUCTION

In order to the serve the growing demand for data transmission, need is to use the limited channel resources efficiently. For this purpose the communication network administrators uses techniques, like digital low rate encoding(LRE) and digital speech interpolation(DSI), these techniques are combined in digital circuit multiplication(DCM) and, due to these the communication system will get higher gain. In this process, the speech and data being sent through the same voice band channel in an interleaved fashion. The transmission of data signal happens during the pauses of voice signal in addition to the separate chunk of time for it. At receiver side the data and speech should be identified, for sending speech and data contents to the respective users.

In case of strategic intelligence gathering operations, the intercepted signal will be analyzed for information extraction. Here the intercepted signal goes through the several stages of signal processing like feature extraction, modulation identification, de-modulation, channel coding, de-multiplexing, speech/data discriminator and then to speech codec or data decoder, based on the content carried by the channel. This process is shown in Fig. 1 In this process the speech/data discrimination plays an important role, by identifying whether the content carried by the channel is speech or data. After this process speech is sent to voice codec and data to data decoder using classification analysis.

II. PROPOSED ALGORITHM FOR SPEECH/ DATA DISCRIMINATION

An accurate identification of voice-band data signals to discriminate them from speech signals is possible using techniques of statistical analysis. The incoming signal is analyzed within an observation window N samples wide, and typical parameters are extracted and then combined to provide a final decision on its nature (voice or data). These parameters are short-time energy and the zero crossing rates.

2.1 Short-time energy

Short-time energy reflects signal level during the observation window. The short-time energy at time n in an N sample window is given by (1)

$$(1) E(n, N) = \sum_{i=1}^N X^2(n - i)/N$$

X(n) being the incoming sample at time n The short-time energy for speech signals is on average lower than that for voice-band data signals because the average level of speech signals is lower than the average level of voice-band data signals (as specified in CCITT Recommendations). Additional information useful for Speech/Data discrimination can be given by the short-time energy calculated on a high-pass and low-pass

filtered signal. A speech signal has most of its energy concentrated at frequencies below 900 Hz, while voice-band data signals have a spectrum spread over 900 Hz. Let $F1(n)$ be the output sample of a low-pass filter and $F2(n)$ the output sample of a high-pass filter at time n , $X(n)$ being the input sample. The corresponding short-time energies at time n during an N sample observation window are $EF1(n, N)$ and $EF2(n, N)$.

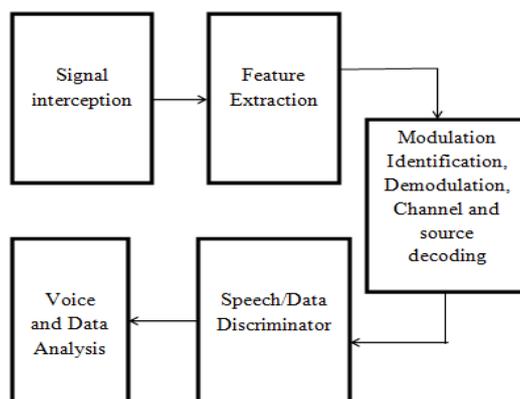


Fig. 1. Signal Interception analysis System Model

Setting the filter cut-off frequencies close to 900Hz, i.e. the low pass filter pass band is 0 to 900 Hz and high pass filter pass band is 900 to 4000 Hz with ripple pass band is 1db and the minimal stop band attenuation is 20 db. Now we can tell simply that weather the input observation window is speech or data based on the below conditions.

If $EF2(n, N) \geq EF1(n, N)$

$X(n), X(n-1) \dots X(n-N+1)$ are speech samples

If $EF1(n, N) \geq EF2(n, N)$

$X(n), X(n-1) \dots X(n-N+1)$ are data samples

The short time energy of voice-band data signals is roughly constant being formed by sinusoids of the same amplitude.

2.2 Zero crossing rates

Zero crossing rate measurement allows us to obtain spectral information on a signal. The zero crossing count at time n is the number of sign changes in an N Sample observation window. This count can be defined as:

$$(3) \text{ZOX}(n, N) = \sum_{i=1}^N | \text{sign}[X(n - N + i)] - (\text{sign}(X(n - N + i - 1))) |$$

Where the sign function is defined as:

$$(4) \text{Sign}(a) = 0 \text{ if } a < 0, 1 \text{ otherwise.}$$

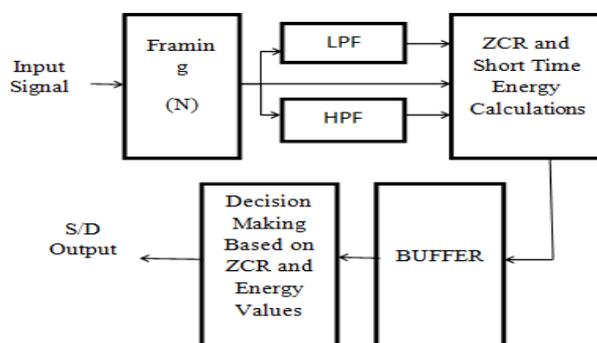


Fig. 2.S/D Discriminator Block Diagram

$\text{ZOX}(n, N)$ ranges from 0 to N and roughly reflects the dominant frequency in the signal. The number of extrema of the signal at time n in the same window $\text{ZIX}(n, N)$, can be represented by the zero crossing count of the difference signal $Y(n) = X(n) - X(n-1)$. $\text{ZIX}(n, N)$ roughly reflects the high frequency component

of the signal, being the difference signal equivalent to the output of a high-pass filter which amplifies the high frequency components of the signal. Both parameters $ZOX(n, N)$ and $ZIX(n, N)$ are insensitive to the signal amplitude and as a consequence the results obtained from their analysis remain valid for the whole amplitude range the plane (ZOX, ZIX) , N being fixed, the region of the points of voice-band data signals is essentially separated from the region of speech signal points, because the spectral characteristics of voice and data signals differ from each other. An example of this is shown in Fig. 3, which shows a scatter diagram of ZOX and ZIX for speech and a 8000 bit/s modem signal, with $N=320$. Speech occupies a crescent shape around the modem region. This is due to the fact that this type of modem uses the full voice bandwidth.

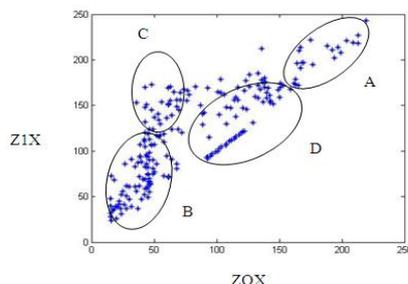


Fig. 3. Scatter diagram of ZOX (zcr of input signal), ZIX (zcr of difference signal) for speech fax signals with 320 sample observation windows. Region A with high values of ZOX is characteristic of unvoiced sounds. Region C with low ZOX and high ZIX is characteristic of closed vowels such as "I" with a considerable gap between first and second format frequencies. Region B is characteristic of open vowels. The modem signal occupies the region D.

2.3 discriminator performance

If in the N sample observation window only one type of signal (speech or voice band data) is present, a greater value for N reduces the probability of false detection of data as speech or speech as data. In fact, with a wider window, the short time energy reflects the average signal level much better and the region of voice band-data signals can be better distinguished from the regions of speech signals in the scatter diagrams. On the other hand, during a transition from speech to data or vice versa, both signals are present in the observation window for $125*(N-1)$ is (125 ps is the sample rate). During this time the S/D discriminator output is unpredictable. Thus the greater N is the longer the maximum transition time is. Other factors that limit the value of N are relevant to the processor used to implement the S/D discriminator algorithm in terms of memory size and processor speed.

2.4 Voice-band data transmission in DCM equipment

The presence of a certain number of voice-band data signals on incoming channels affects DSI gain because continuous data signals cannot be interpolated (voice-band data signals have a 100% activity). In the presence of data signals causes an obvious reduction of the achievable DSI gain for speech channels. A control of voice band channels is then necessary to avoid the reduction of the DSI gain for interpolated speech channels. The data channels should be detected and routed through digital non-interpolated (DNI) channels within the DSI or through alternative routes. The reserved data channels can be pre-allocated or dynamically assigned in function of speech traffic. Speech and voice-band data signals have quite different characteristics (in terms of level, frequency, time correlation), thus a low rate speech encoder (i.e. an ADPCM) customized for speech signal is not suitable for voice-band data signals. In fact any low rate encoding technique makes maximum use of the correlation properties of speech signals to reduce the required bitrates, while data signals are not so correlated.

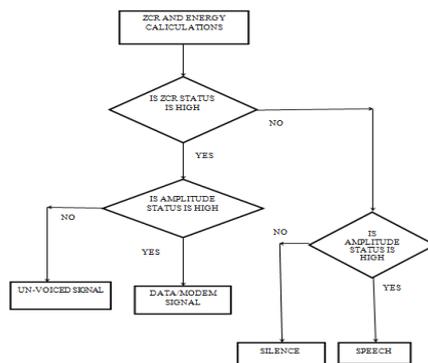


Fig. 4. Flow chart of discrimination algorithm

A good analysis of the characteristics of voice-band data signals (speech and data) as well as the effects of the encoding of voice band signals can be obtained. The 32 Kbit/s ADPCM algorithm specified in the CCITT Rec. G.721 is a compromise to meet the requirements for speech and voice-band data signals, with the following exceptions: inability of transmitting high-speed voice-band data (24800 bit/s) in satisfactory performances of CCITT V.23 1200 bit/s FSK modem data transmission in a full-duplex mode overflow oscillations in the decoder when particular code words are continuously received (i.e. all "zero's"). The problem of voice-band data transmission through DCM equipment has several solutions C61. Transparent transmission through a clear channel is a possible solution. The voice and data channel should be detected by an S/D discriminator and routed through a non-interpolated channel as per the above protocol.

Our Mat lab simulation results shown in Fig. 5 and Fig. 6. In Fig. 5 we have taken a cascaded data and speech signal our input signal contains Data, speech, un-voice and silence our proposed algorithm is succeeded in discriminating among data ,speech unvoice and silence. And in Fig. 6 we taken an interpolated fashioned speech and data signal here we interpolated speech and data only. So the is no presence of unvoice or silence. Here the decision making procedure in discriminating between speech and data is well explained in above flow chart shown in Fig. 4.

III. RESULTS

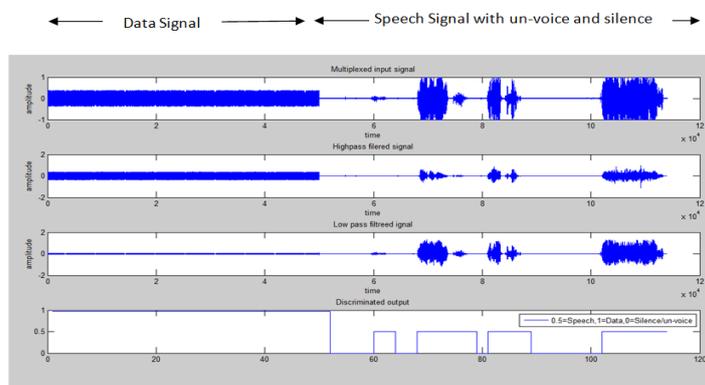


Fig. 5. Final decision on a typical composite signal. (a) Original multiplexed signal (speech and modem). (b) High pass filtered multiplexed signal (c) Low pass filtered multiplexed signal (d) discrimination output, here output level 1 represent data, 0.5 represents speech and 0 is un-voice or silence.

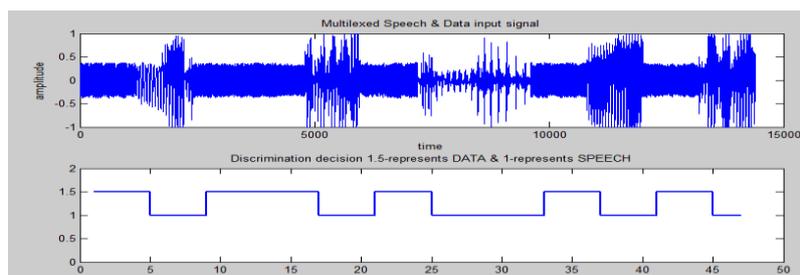


Fig. 6. Discrimination decision between data and speech. (a) Original multiplexed signal (speech and data). (b) Output of discrimination here 1.5-level represents data and 0.5-level represents speech.

IV. CONCLUSIONS

The request for data circuits on networks is continuously growing, and the need for intelligence gathering also became a necessity in military applications. So at the receiver for the above two applications, the speech/data discrimination is necessary. Our proposed method offers a reliable, simple and low complexity algorithm for speech/ voice-band data discrimination. Due to the virtue of low complexity it's the hardware implementation also become easy. Its performance has been evaluated on data and speech (fax for data, recorded tape for speech), and this algorithm give good results in discriminating between speech and data.

REFERENCES

- [1] S.Casale, C. Giarrizzo, A. LA Corte. "A DSP implemented speech/voice band data discriminator" IEEE Trans. Commun. , Apr. 1986.
- [2] C.Roberge and J.P. Adoul, "Fast on-line speech/voiceband data discriminator for statistical multiplexing of data with telephone channels," IEEE Trans. Common., vol. COM-34, pp. 744-751, Aug. 1986.
- [3] A pattern recognition approach to voiced-unvoiced-silence classification with applications to speech recognition IEEE transactions on accostics, speech, and signal processing, vol. assp-24, no. 3, June 1976.
- [4] Digital signal processing using matlab-Third Edition-vinay k.Ingle & John g.Proakis.
- [5] Signal Processing Toolbox For Use with MATLAB- Mathworks User guide 4.2.
- [6] N. Benvenuto and T. W. Goeddel, "Classification of voiceband datasignals using the constellation magnitude," IEEE Trans. Commun., vol.43, pp. 2759-2770, Nov. 1995.
- [7] J.P. Adoul and F. Daaboul, "Parametric segmentation of speech into voiced, unvoiced and silence intervals," in Proc. IEEE, Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP '77.