International Journal of Scientific Research & Engineering Trends
Volume 8, Issue 3, May-Jun-2022, ISSN (Online): 2395-566X

A Review Article Novel Approach Ofdm 5g Communication and Papr Minimization Using Modified Clipping Method

Preeti Tembhurne, Swati Nigam
Dept. Of ECE
Sanghvi institute of Management & science Indore,
SIMS pigdambar, Rau behind IIM Indore pin code 453332

Abstract – To improvement of voice signal quality and reducing its noise level to implementation of various type method like filter and Compression etc. Hence observation of various article obtained maximum methods is not compatible of real time speech signal. So our proposed implementation gives a topological method to formulation of noise problem and real time speech signal processing. Its numerical technique are called of thresholding based DWT (Discrete wavelet Transform). This approach available of multiple decomposition and Composition to reduce of noise level in terms of db.

Keywords- MATLAB, DWT, 1 Directional DWT, Signal processing, noise, SNR.

I. INTRODUCTION

Digital filter-banks are an integral part of many speech and audio processing algorithms used in today's communication systems. They are commonly employed for adaptive subband filtering, for example, to perform acoustic echo cancellation in hands-free communication devices or multichannel dynamicrange compression in digital hearing aids, e.g., [1]. Another frequent task is speech enhancement by noise reduction, e.g., [2].

This eases the communication in adverse environments where acoustic background noise impairs the intelligibility and fidelity of the transmitted speech signal. A noise reduction system is also beneficial to improve the performance of speech coding and speech recognition systems, e.g., [3]. The choice of the filter-bank has a significant influence on the performance of such systems in terms of signal quality, computational complexity, and signal delay. Such distortions can occur when the hearing aid user is talking. In this case, the processed speech can interfere with the original speech signal, which reaches the cochlea with minimal delay via bone conduction or through the hearing aid vent.

To prevent this, the algorithmic signal delay of the filter-bank used for the signal enhancement must be considerably lower than the tolerable processing delay, i.e., the latency between the analog input and output signal of the system. In addition, a filter-bank with non-uniform time-frequency resolution, which is similar to that of the human auditory system, is desirable to perform multi-channel dynamic-range compression and noise reduction with a small number of frequency bands [4].

II.SPEECH PROCESSING

Speech processing has many applications like wireless communication, teleconferencing system, long distance communication and radar communication [1]. In all these application a noise free speech is highly desired. So to have a noise free speech we need to perform certain task to remove the noise before processing it. Basically noise is mixed with the speech during the generation of the speech or may be in transmitting the speech over a noisy channel or at the receiver end. A noise can be anything like it may be environmental noise or system noise or channel noise but it is essential to remove this noise to enhance the quality of the speech.

III.RESEARCH MOTIVATION

Digital hearing aids are widely used by hearing-impaired people to improve their speech intelligibility and quality of life. However, hearing aid performance is usually degraded due to acoustic feedback, which generates other problems. This phenomenon is produced when the sound propagates from the loudspeaker to the microphone. This causes instability and a high-frequency oscillation that can be perceived by hearing-impaired people if its level exceeds their hearing thresholds. Also, these effects limit the maximum gain that the hearing aid can perform and reduce the sound quality when the gain is close to the limit. To reduce acoustic feedback, several methods based on adaptive algorithms have been used [5] for feedback reduction. A number of techniques are used for speech noise reduction, like non-local diffusion filters [6], acoustic feedback reduction based on finite impulse response and infinite impulse response adaptive filters in digital hearing aids [4], and noise reduction Wiener filter

[5]. In the above techniques, the noise is suppressed using filters such as adaptive filters or Wiener filter.

Research Gap

The spectral subtraction technique is a generally known method for noise reduction [7]. In this method, the noisy speech signal is first transformed from the time domain into the frequency domain by means of the fast Fourier transform (FFT). The noise spectrum is then determined in the speech pauses and subtracted from the frequency spectrum of the noisy speech signal before the noisy speech signal is reconverted from the frequency domain into the time domain by means of the inverse FFT (IFFT).

For reducing the unwanted noise signal from the original speech signal, the digital hearing aid is designed with our proposed effective noise degradation architecture in this paper. The continuous time domain signal is segmented into overlapping chunks called frames, and the frames are multiplied by a window function for avoiding the spectral artifacts, to perform frequency domain processing of the speech signal in our system. The signal in time domain is converted into frequency domain by using an FFT processor, and to reduce the noise signal from the speech signal, the Wiener filter is used in the digital hearing aid. The output signal from the Wiener filter is converted into time domain by using an IFFT processor and then multiplied with the same window function, and the frames are then overlapped to create a continuous output signal.

IV. LITERATURE REVIEW

[1] Chih-Lyang Hwang: To implement the human–robot interactions in a noisy environment, the speech improvement-based (SIB) stratified adaptive finite-time saturation control (SAFTSC) for omnidirectional service robot (OSR) is developed. From the outset, the feature vectors of nine designed speech commands are extracted from their frequency signals and then trained by multiclass support vector machine. Two background noises are on-line filtered with the characteristic: “the smaller error in power spectrum is, the larger recovery from noisy power spectrum is.” Comparisons among without or with noise, and filtering are addressed.

To achieve the zero pose error of OSR in finite time, an adaptive finite-time indirect trajectory (AFTIT) is constructed. To track the AFTIT with the zero error in finite time, the adaptive finite-time saturation control (AFTSC) is also established. Both AFTIT and AFTSC possess nonlinear switching gain increasing the high-frequency motion capability to fulfill the classified speech command. Simply put, the proposed SIB stratified AFTSC includes the speech improvement for classification, the AFTIT, and the AFTSC. Besides the stability of the closed-loop system is verified by the Lyapunov stability theory, three categories of SIB experiments are compared.

[2] M. Chandni : Epochs are impulse like discontinuities in speech that represent the instants of glottal closure during phonation. The accurate estimation of epochs is essential for epoch synchronous speech processing. There are many algorithms proposed in the literature for the accurate estimation of epochs from speech signals. However majority of those algorithms show degradation during the epoch extraction from telephonic speech signals. The restricted bandwidth (300 Hz-3.4 kHz) of the telephone channel is the reason for degradation in epoch estimation performance.

Therefore, the objective of the present work is to propose a wavelet synchrosqueezed transform (WSST) based signal preprocessing stage to improve the performance of zero frequency filtering (ZFF) which is a well known epoch estimation method. Improved time-frequency resolution of WSST reinforce the strength of impulses around the fundamental components to estimate accurate epoch locations. A multiscale product of WSST modified speech and the original speech is computed further enhance the impulse like discontinuities. Effectiveness of the proposed method is confirmed by the epoch estimation performance analysis of phonetically balanced CMU-Arctic database with electroglottogram recordings. The telephone quality in the database is simulated using G.191 software tools.

[3] Heng Li: Human emotion is a concrete form of human communication, and the research on emotion recognition is increasing gradually. In recent years, researchers have paid more attention to multi-modal emotion recognition. This paper presents a deep neural network for emotion recognition based on speech spectrum. Spectrograms contain comprehensive information about speech and are useful for emotion recognition. We tried the convolutional neural network (CNN) and the Long-Short Term Memory (LSTM), the combination of voice to make use of CNN feature extraction, using LSTM network reserve the temporal information, the voice information extracted from spectrogram, and the emotion recognition task. This study adopts the university of southern California’s Interactive Emotion Capture (IEMOCAP) dataset as the data collection. We use the speech spectrogram as input, for six kind of mood, and the final weighted accuracy is 61%, the unweighted accuracy is 56%.

[4] FengfanHou: Speech separation is an important component in robust speech processing systems, and recent models have shown a better performance than ideal time-frequency magnitude masks on noise-free dataset. But reverberant environment will greatly degrade the performance of these models. Inspired by the usage of early reflections in speech enhancement tasks, we explore the effect of preserving early reflections on reverberant speech, and proposed a training method for reverberant monaural speech separation systems. By using proposed training method, which uses early reflections within 20ms behind direct sound as a part of target speech, the

perceptual evaluation of speech quality (PESQ) is improved by 0.145, and the short time objective intelligibility (STOI) is improved by 0.017, compared to the origin network using direct sound as target speech.

[5] Jianning Huang: We propose a method to use music signals to guarantee speech privacy at shelters. Since it is difficult to install private rooms in a shelter, we cannot prevent a conversation from being heard by neighbors. Our idea is to use background music as noise to keep speech privacy while canceling the music signal within the “silent area.” The proposed method plays the background music using a loudspeaker in the shelter. A signal for canceling the music is embedded in the high-frequency region of the music signal using amplitude modulation. On the receiver side, a small device extracts the embedded signal using the demodulator and applies an adaptive filter to cancel the sound around the receiving signal using the local loudspeaker. We conduct a simulation experiment to confirm that the proposed method could suppress the music signal more than 20 dB.

[6] Yi Lei: Expressive synthetic speech is essential for many human-computer interaction and audio broadcast scenarios, and thus synthesizing expressive speech has attracted much attention in recent years. Previous methods performed the expressive speech synthesis either with explicit labels or with a fixed-length style embedding extracted from reference audio, both of which can only learn an average style and thus ignores the multi-scale nature of speech prosody. In this paper, we propose MsEmoTTS, a multi-scale emotional speech synthesis framework, to model the emotion from different levels. Specifically, the proposed method is a typical attention-based sequence-to-sequence model but with proposed three modules, including global-level emotion presenting module (GM), utterance-level emotion presenting module (UM), and local-level emotion presenting module (LM), to model the global emotion category, utterance-level emotion variation, and syllable-level emotion strength, respectively.

In addition to modeling the emotion from different levels, the proposed method also allows us to synthesize emotional speech in different ways, i.e., transferring the emotion from reference audio, predicting the emotion from input text, and controlling the emotion strength manually. Extensive experiments conducted on a Chinese emotional speech corpus demonstrate that the proposed method outperforms the compared reference audio-based and text-based emotional speech synthesis methods on the emotion transfer speech synthesis and text-based emotion prediction speech synthesis respectively. Besides, the experiments also show that the proposed method can control the emotion expressions flexibly. Detail analysis is conducted to show the effectiveness of each module and the good design of the proposed method.

V. PROBLEM QUALITY SOUND

The high quality sound of talking speech in real environment is very important for automatic speech processing systems and human- machine interfaces. However, the performance of these systems can be affected by background noise. Thus, there is a strong need to resolve this problem and improve the performance of these applications in high level noise environment by applying effective speech enhancement techniques able to suppress the undesirable noise. These techniques are concerned with improving some perceptual aspect, the quality and intelligibility of degraded speech. In a broad context, many methods are developed in order to remove the background noise while retaining speech intelligibility based on short time spectral estimation of the clean speech. These methods are able to reduce the noise and improve the quality, but at the expanse of introducing speech distortion which results in loss of intelligibility. Hence, the main challenge in designing effective speech enhancement algorithms is to suppress the noise without introducing any perceptible speech distortion.

The spectral modification methods are historically one of the first algorithms proposed for noise reduction, especially the generalized spectral subtraction is the most popular technique [1]. This method is able to reduce the background noise using estimation of the short-time spectral magnitude of the speech signal by subtracting the noise estimation from the noisy speech. The spectral subtraction technique offers a high flexibility and simplicity in implementation. However, it needs to be improved since its major drawback, the introduction in the enhanced speech of residual noise called “musical noise” with unnatural structure, is composed of tones at random frequencies.

The unnatural structure of the musical noise is perceived as nonstationary noise artifacts that depend on the time and frequency changes of the noise, on one side and on the way that the human auditory system perceives these artifacts, on the other side.

The minimum-mean-square-error-based-noise reduction proposed by Ephraim and Malah subtraction rule [2] exploits the average spectral estimation of the speech signal based on a prior knowledge of the noise variance, in the goal to mask and reduce the residual noise. In [3–8], the noise is reduced based on subtractive type algorithms according to a multibands and nonlinear spectral process. In [9–10], the authors exploit the human perceptual masking proprieties to improve the quality and intelligibility of the speech signal without introducing speech distortion. The difficulty with these approaches is that an estimate of the clean speech itself is necessary in order to calculate the masking threshold.

International Journal of Scientific Research & Engineering Trends
Volume 8, Issue 3, May-Jun-2022, ISSN (Online): 2395-566X

VI. CONCLUSION

The proposed digital filter for speech signals using multi rate signal processing has been designed. After filtering, the quantization of the input signal and filter coefficients were analyzed. The performance of the proposed system was evaluated based on Signal to Quantization Noise Ratio. From the performance measures, it was observed that DWT applied signals filtered by Chebyshev HPF provides high SQNR in IIR Filtering and DWT applied signals filtered by Blackman Window provide high SQNR in FIR filtering. In future, the proposed digital filters will be further implement for various signals such as ECG signal, OFDM signal, etc. Speech signals are designed and implemented using multi rate signal processing.

The systems that employ multiple sampling rates in the processing of signals are called multi rate signal processing systems. Multirate digital signal processing systems use down sampler and up sampler, the two basic sampling rate alteration devices in addition to conventional elements like adder, multiplier and delay to change the sampling rate of a digital signal.

REFERENCES

[1]. Chih-lyang hwang, experimental validation of speech improvement-based stratified adaptive finite-time saturation control of omnidirectional service robot, ieee transactions on systems, man, and cybernetics: systems.
[2]. M. chandni, effectiveness of wavelet synchrosqueezed transform for improved epoch estimation from telephonic speech signals using zero frequency filtering, 2021 ieee 18th india council international conference (indicon), 978-1-6654-4175-9.
[3]. Heng li, research on speech emotion recognition based on deep neural network, 2021 ieee 6th international conference on signal and image processing (icsip), 978-1-6654-3904-6.
[4]. Fengfanhou, preserving early reflections to improve speech quality of reverberant speech separation, 2021 ieee 6th international conference on signal and image processing (icsip), 978- 1-6654-3904-6.
[5]. jianning huang, study on the background music cancellation system for speech privacy, 2021 ieee 6th international conference on signal and image processing (icsip), 978-1-6654- 3904-6.
[6]. yi lei, msemotts: multi-scale emotion transfer, prediction, and control for emotional speech synthesis, ieee/acm transactions on audio, speech, and language processing (early access), 2329-9290.
[7]. shakir khan, hcovbi-caps: hate speech detection using convolutional and bi-directional gated recurrent unit with capsule network, ieee access (volume: 10), 2169-3536.
[8]. yanmin qian, optimizing data usage for low-resource speech recognition, ieee/acm transactions on audio, speech, and language processing (volume: 30), 2329-9290.
[9]. javier jorge, live streaming speech recognition using deep bidirectional lstm acoustic models and interpolated language models, ieee/acm transactions on audio, speech, and language processing (volume: 30), 2329-9290.
[10]. Mixiaohou, multi-view speech emotion recognition via collective relation construction, ieee/acm transactions on audio, speech, and language processing (volume: 30), 2329-9290.