Islamic University, Gaza, Palestine

Research and Graduate Affairs

Faculty of Engineering

Computer Engineering Department

Phase-Based Speech Enhancement

Bilal Abdul Raouf Shehada

Supervisor

Associate Professor

Mohammed Ahmed Alhanjouri

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of Master of Science in Computer Engineering

DEDICATION

To My Father, and Mother,

To My brothers and sisters,

To My Family, and Friends.

Acknowledgments

Firstly, all thanks are to Allah the almighty, who guide me to accomplish this work, so all praise is to Allah. Then, there are a number of people to whom I am greatly indebted as without them this thesis might not have been written.

To Dr. Mohammed Alhanjouri for his guidance, support, and advice.

To my parents for providing me with the opportunity to be where I am. Without them, none of this would be even possible to do. You have always been around supporting and encouraging me.

I am also very grateful to my brother Ali Shehada, who did not spare any effort to support me with online database journal.

I am also very grateful to my friend Mahmoud Al-Aff, who did not spare any effort to review the thesis.

Contents

Abbreviations

FFT	Fast Fourier Transform
KLT	Karhunen–Loeve Transform
MAP	Maximum A Posteriori
MCRA	Minima Controlled Recursive Averaging
MMSE	Minimum Mean-Squared Error
MSE	Mean-Squared Error
PESQ	Perceptual Evaluation Of Speech Quality
Seg_SNR	Segmental Signal To Noise Ratio
SMPO	Soft Masking By Incorporating A Posteriori SNR Uncertainty
SMPR	Soft Masking By Incorporating A Priori SNR Uncertainty
SNR	Signal To Noise Ratio
SPZC	Spectrum Power Zero Cross
SPZCU	Spectrum Power Zero Cross Uncertainty
ZC	Zero Cross

List of Figures

List of Tables

التحسين على الصوت باستخدام الطور

للباحث: بلال عبد الرؤوف محمد شحادة

الملخص

يعتبر الكلام اهم طريقة للتواصل من بين البشر لذلك قد حاز مجال التحسين على الصوت أبحاث كثيرة خلال العقود الثلاثة الأخيرة. في مجال الاتصالات تكمن المشكلة الأساسية في التحسين على جودة الصوت وتحسين الوضوح في البيئات التي تحتوي ضجيج. معظم أساليب التحسين على الصوت تعتمد على إزالة الضجيج باستخدام الاعداد الحقيقية في المجال الطيفي للإشارة وتغفل القيمة التخيلية في عملية التحسين.

في هذه الدراسة سوف نقوم بدراسة وتوضيح أهمية الاعداد التخيلية في المجال الطيفي للإشارة في عملية التحسين على الصوت وذلك عن طريق قياسات perceptual evaluation of speech quality (PESQ). وتهدف هذه الدراسة بشكل أساسي الى اظهار أهمية الاعداد التخيلية في المجال الذري للإشارة للتحسين على جودة ووضوح الصوت بأقل الخسار في قوة وجودة الإشارة. في هذه الدارسة قمنا باستخدام قاعدة البيانات NOIZEUS واستخدم PESQ كطريقة لقياس جودة الصوت. في هذا البحث قمنا بالجمع بين عدة طرق في التحويل مثل Karhunen–Loeve transform واستخدام عدة طرق في مجال توقع الإشارة مثل maximum a posteriori وprior probability وprobability posteriori.

من خلال هذه الدراسة توصلنا الى ان أفضل طريقة للتحسين على الصوت من ضمن الخوارزميات المقترحة كانت عن طريق استخدام wave atoms with KLT and prior probability uncertainty وقد حصلنا على زيادة في جودة الصوت بنسبة 17.5% بالنسبة للجودة الاصلية.

في هذه الدراسة قمنا باستخدام بيئة العمل MATLAB في بناء الخوارزميات المقترحة.

Phase-Based Speech Enhancement

Bilal A M Shehada

Abstract

Speech is a fundamental means of human communication. Speech enhancement has been a challenge for many researchers for almost three decades. The problem involves improving the performance of speech communication systems and improve the quality and intelligibility of degraded speech in noisy environments. The majority of speech enhancement methods perform noise removal in spectral domain and construct the enhanced speech signal from the estimated magnitude of clean speech discarding the phase side of the noisy speech.

In this thesis, the researcher aims to show the effect of modifying the complex spectrum of enhancement process in the quality of the signal in term of perceptual evaluation of speech quality (PESQ). The main job of this study is to show the importance of the complex value of wave atoms transform for enhancing the degraded signals with minimum speech distortion. The researcher use NOIZEUS dataset and PESQ measurement in comparison between techniques. This study use wave atoms transform as main enhancement technique combining it with other transformation techniques like Karhunen–Loeve transform and use different statistical probability model like maximum a posteriori, prior probability and posteriori probability with speech presence probability. The result of the thesis shows that best enhancement achieved when use multi transformation technique with wave atoms. This thesis reach 17.5% PESQ enhancement with combining wave atoms with KLT and prior probability uncertainty.

The proposed algorithms are implemented using MATLAB environment. The results indicate that the proposed system improves the classical methods performance in terms of speech quality.

Keywords: PESQ, KLT, Wave Atoms, MAP, Prior probability, Posteriori probability, speech presence probability

Chapter 1

1 Introdution

Speech, as a physical phenomenon, consists of local changes in acoustic pressure resulting from the actions of the human vocal apparatus. It is formed mainly for the purpose of oral communication. The pressure changes generate auditory waves that propagate through the communication medium (generally air). At the receiving end, speech is processed by the auditory system and higher cortical regions of the brain. A transducer (microphone) in the acoustic field "follows" the speech signal, which can be analyzed numerically. In the case of a microphone, the speech signal is electrical in nature and describes the acoustic pressure changes as voltage variations with respect to time. The speech signal holds data not only about just what has been said (the linguistic message), but also about who has said it (speaker dependent information), in which environment it was said (e.g., noise or reverberation), over which communication channel it was transmitted (e.g., microphone, recording equipment, transmission line, etc.), the health of the speaker, and so on. Not all of the information sources are of interest for any given application [1]. Because of noise is everywhere all applications that are related to voice and speech, from sound recording, cellular phones, hands-free communication, teleconferencing, hearing aids, to human-machine interfaces, a speech signal of interest captured by microphone sensors is always mixed with noise. Depending on its level, the noise can significantly contaminate the statistical characteristics and spectrum of the desired speech signal[2].

In some automatic speech enhancement applications the goal is to recover only the linguistic message regardless of the identity of the speaker, the acoustic environment, or the transmission channel. In fact, the presence of additional information sources may be detrimental to the decoding process.

1.1 Speech Production Process

The speech production process takes place inside the vocal tract extending from the glottis to the lips energized from air-filled lungs. The vocal tract is a cavity of extremely complicated geometrical form whose sizes and configuration may differs continuously in time and whose walls are composed of tissues having widely ranging properties[3]. It begins at the glottis and ends at the lips. Figure 1.1 shows the vocal tract anatomical structure. The glottis is a slit-like orifice between the vocal cords (at the top of the trachea). The cartilages around the cords support them and facilitate adjustment of their tension. The flexible structure of the vocal cords makes them oscillate easily. These oscillations are responsible for periodic excitation of vowels. The nasal tract constitutes an ancillary path for sound transmission. It begins at the velum and terminates at the nostrils.

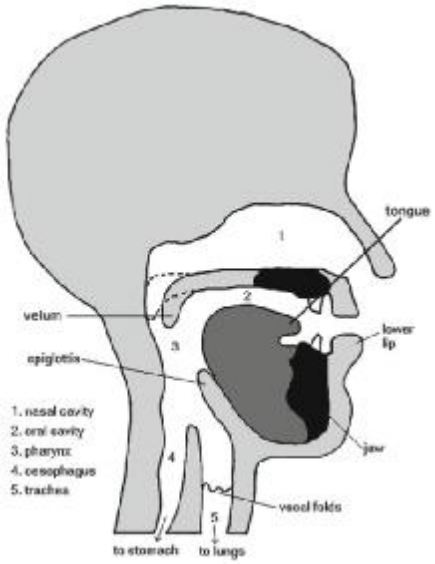


Figure 1.1 vocal tract anatomical structure[3]

1.2 Classification of Speech Sounds

The organs of speech are capable of making many different kinds of speech sounds. Speech sounds are classified according to the type and place of excitation. Voiced sounds and vowels are characterized by a periodic excitation at the glottis. For voiced sounds and vowels, the expelled air from lungs causes the vocal cords to vibrate as a relaxation oscillator, and the air stream is modulated into discrete puffs. This oscillation starts when the sub glottal pressure is

increased sufficiently to force the initially abducted cords apart with lateral acceleration. As the air flow builds up in the orifice, the local pressure is reduced and a force acts to return the cords to a proximate position[3]. Consequently, the pressure approaches the subglottic value as the flow decreases with the decrease in the orifice (glottal area). The relaxation cycle is then repeated. The mass and compliance of the cords, and the subglottic pressure determine the oscillation frequency (pitch) [1]. In English there are twenty vowels and twenty-eight consonants.

Unvoiced sounds are caused by passing the air stream through a clutch in the tract. The pressure perturbations due to these excitation mechanisms provide an acoustic wave which spreads along the vocal tract toward the lips. If the nasal tract is coupled to the vocal cavity through the velum, the radiated sound is the resultant of the radiation at both the lips and nostrils and it is called nasalized sounds (as in/m/and/n/). The special sounds of any language (phonemes) are uniquely determined by describing the excitation source and the vocal tract configuration.

1.3 Noise Characteristics

In electronics, noise is a random fluctuation in an electrical signal. Noise is considered to be unwanted or undesirable sound but actually we are surrounded by noise wherever we go. The unit which has been adopted for indicating the noise level is the decibel (dB) and it relates on a logarithmic scale to the basic unit of sound pressure, namely the Pascal (Pa). The decibel scale starts at 0 dB for sounds that can just be heard and reaches 130 dB at the onset of aural pain Noise is present, for instance, in the street (e.g., cars passing by, street construction work), the car (e.g., engine noise, wind), the office (e.g., PC fan noise, air ducts), the restaurant (e.g., people talking in nearby tables), and the department stores (e.g., telephone ringing, sales representatives talking) [4]. As these examples tell, noise appears in different forms in daily life.

Noise can generally be classified into three major categories based on its characteristics[5]:

- Stationary noise, does not change over time, such as the fan noise Coining from PCs.
- Pseudo or Non-stationary noise, such traffic or crowd of people speaking in the background, mixed in some cases with music or as the restaurant noise, that is, multiple people speaking in the background mixed in some cases with noise emanating from the

kitchen. The spectral (and temporal) characteristics of the restaurant noise are always changing as people talk conversations in neighboring tables and as the waiters keep interact and converse with people.

- Transient noise, i.e., hammering or door slam.

The spectral and temporal characteristics of pseudo or non-stationary noise change constantly. Clearly, the task of suppressing this type of noise is more difficult than that of suppressing stationary noise. Another distinctive feature of noises is their spectrum shape, particularly the distribution of noise energy in the frequency domain. For instance, most of the energy of car noise is concentrated in the low frequencies, i.e., it is low-pass in nature. Train noise, on the other hand, is more broadband as it occupies a wider frequency range [6]. In most speech enhancement methods, the estimation of the power of the noise is a requirement. Fortunately, the bursty nature of speech makes it possible to estimate the noise during speech pauses. Moreover, it should be mentioned that it is easier to deal with additive noise than convoluted noise. This is why the assumption stating that the noise and speech are additive is often made.

For practical and natural reasons, the estimation of the noise is almost performed in the spectral domain. Actually, spectral components of speech and noise are partially uncorrelated. Besides this, perception/hearing and psycho-acoustic models are well understood (and adapted) in the spectral domain.

1.4 Classes of speech enhancement algorithms

A number of algorithms have been proposed in the literature for speech enhancement with the primary goal of improving speech quality[4]. These algorithms can be divided into three main classes:

A. Spectral subtractive algorithms: These are, by far, the simplest enhancement algorithms to implement. They are grounded on the straightforward principle that as the noise is additive, one can estimate/update the noise spectrum when speech is not present and subtract it from the noisy signal.

B. Statistical-model-based algorithms: The speech enhancement problem is posed in a statistical estimation framework, Given a set of measurements, corresponding say to the Fourier transform coefficients of the noisy signal, we wish to find a linear (or nonlinear) estimator of the parameter of interest, namely the transform coefficients of the clean signal. The Wiener algorithm and minimum mean square error (MMSE) algorithms, among others, fall in this category.

C. Subspace algorithms: Unlike the above-mentioned algorithms, the subspace algorithms are rooted primarily on linear algebra concept. More specifically, these algorithms are based on the principle that the clean signal might be confined to a subspace of the noisy Euclidean space. Consequently, given a method of decomposing the vector space of the noisy signal into a subspace that is occupied primarily by the clean signal and a subspace occupied primarily by the noise signal.

D. Binary mask algorithms: Unlike the algorithms in Classes A-C, which for the most part make use of smooth gain functions for noise suppression, the binary mask algorithms make use of binary gain functions. This quantities to selecting a subset of frequency bins (or channels) from the corrupted speech spectra, while discarding the rest. The choosing of those bins is done regarding to a prescribed rule or criterion. The selection gauge is not unique, and when applied to corrupted speech spectra can produce large gains in speech intelligibility, Work in this area originated in the field of computational auditory scene analysis and for the most part, binary mask algorithms have been utilized in ASR applications rather than for noise reduction. Under ideal conditions the binary mask algorithms can improve speech intelligibility at any SNR level and for any type of background noise.

1.5 Problem Formulation

There are different ways to perform speech enhancement in frequency domain from a single microphone signal. The simplest way is to estimate the desired signal from the noisy observation with a simple complex gain. The noise reduction or speech enhancement problem considered in this study is one of recovering the desired signal (or clean speech) $x(n)$, t being the time index, of zero mean from the noisy observation (microphone signal) [1–3]

$$y(n) = x(n) + d(n), \quad (1.1)$$

where $d(n)$ is the unwanted additive noise, which is assumed to be a zero-mean random process white or colored but uncorrelated with $x(n)$. All signals are considered to be real and broadband.

Using the short-time Fourier transform (STFT), the (1.1) can be rewritten in the frequency domain as

$$Y(k,m) = X(k,m) + D(k,m), \quad (1.2)$$

where the zero-mean complex random variables $Y(k,m), X(k,m)$, and $D(k,m)$ are the STFTs of $y(n)$, $x(n)$, and $d(n)$, respectively, at frequency-bin k ∈ {0, 1,..., K -1} and time-frame m. Since $x(n)$ and $d(n)$ are uncorrelated by assumption, the variance of $Y(k,m)$ is

$$\Phi_y(k,m) = E[|Y(k,m)|^2]$$

$$= \Phi_X(k,m) + \Phi_D(k,m) \quad (1.3)$$

where E [·] denotes mathematical expectation, and

$$\Phi_X(k,m) = E[|X(k,m)|^2] \quad (1.4)$$

$$\Phi_D(k,m) = E[|D(k,m)|^2] \quad (1.5)$$

are the variances of X (k, m) and D (k, m), respectively.

1.6 Thesis Contribution

This thesis contributes to speech enhancement system is introduced for enhancing speech signals corrupted by additive noise and improving the performance of Automatic Speech Recognizers in noisy conditions. This thesis focus in illustrating the importance of wave atoms multi transformation in speech enhancement. The study will answer the questions about whether phase and complex part should play a more dominant role in the speech enhancement process and what is the effect of complex part in enhancing the speech intelligibility. The primary objective of this thesis is increase the signal quality using the complex part of the transformation. Our approach combines multi transformation technique with wav atoms transform and apply multi threshold type in thresolding the wave atoms transform. This study try's to show the importance of using the complex part of the wave atoms transform in enhancing the signal. The

study will show the effect of using the imaginary part only of the transform in the enhancement process. The researcher will evaluate the proposed algorithms using PESQ and compare algorithms' results with other famous related algorithms' results. It is expected that the results of the proposed algorithms will confirm the high performance of the proposed methods in term of perceptual evaluation of speech quality.

1.7 Organization of thesis

The rest of thesis is divided into four chapters organized as follows:

- Chapter 2: this chapter talks about the new relative research with same characteristics of our research.
- Chapter 3: this chapter describes the algorithm mathematical equation background.
- Chapter 4: show the algorithm design and make comparison between techniques.
- Chapter 5: show the conclusion and future work of the thesis.

Chapter 2

2 Literature Review

The purpose of this chapter is to set the present study in the context of other studies of spectral technique .This Chapter presents necessary background about the speech signal and related works presented few years ago with same characteristics of our method.

2.1 Related work

Speech enhancement algorithm has a very rich history due its importance and we can classify the methods as described in cahpter1.

2.1.1 Spectral subtractive algorithms:

Paliwaland Schwerin [7] add new stage to AMS framework (analysis–modification–synthesis) to include modulation domain processing. To show the important of the modulation stage for speech intelligibility, they made three experiments. In the first, they investigate the relative contributions to intelligibility of the modulation magnitude, modulation phase, and acoustic phase spectra. In the second experiment, the effect of modulation frame duration on intelligibility for processing of the modulation magnitude spectrum is investigated. In the third experiment, the effect of modulation frame duration on intelligibility for processing of the modulation phase spectrum is investigated. They observe that the intelligibility of stimuli constructed from only the modulation magnitude or phase spectra is significantly lower than the intelligibility of the acoustic magnitude spectrum. The intelligibility of stimuli generated from either the modulation magnitude or modulation phase spectra was shown to be considerably improved by also retaining the acoustic phase spectrum. There is an effect of the modulation frame duration on intelligibility for both the modulation magnitude and phase spectra as the speech reconstructed from only the short-time modulation phase spectrum has highest intelligibility when long modulation frame durations (>256 ms) are used, and that for small

durations (<64 ms) the modulation phase spectrum can be considered relatively unimportant for intelligibility.

Upadhyay & Karmakar [8], use multi-band spectral subtraction (PM-MBSS) algorithm for enhancement the whole speech spectrum is divided in different non-uniform bands in accordance to the critical-band rate scale and spectral subtraction is executed independently in each band. They divided the spectrums into 6 portions as non-overlapping and non-uniform critical-bands and spectral subtraction is performed independently in each band. The simulation results and informal listening tests indicate that the PM-MBSS not only reduces the low frequency noise, but also eliminates the high requency noise substantially. The PM-MBSS has strong flexibility to adapt any complicated rigorous speech environment by adjusting the over-subtraction factor for each critical-band. The motivation of multi-band approach multi-band spectral subtraction to efficiently mimic the psychoacoustic model of human ear.

Mehmetcik [9] investigate a contribution of DFT phase on speech quality is. The phase spectrum has a significant effect on speech quality for short durations of analysis windows. Furthermore, phase values of low frequency components are found to have the largest contribution to this quality improvement. Under the motivation of these results, new enhancement method is proposed which modifies the phase of certain low frequency components as well as the magnitude spectrum. It is also shown that the phase of these tonal components can be predicted with the use of phase continuity assumption and estimated frequency values. It is also shown that the phase of these tonal components can be predicted with the use of phase continuity assumption and estimated frequency values. The proposed method makes use of the detected average fundamental frequency to estimate the phase of the corresponding DFT coefficient. The proposed algorithm uses the classical enhancement algorithms to modify the magnitude spectra in addition to the conducted phase corrections. As the implementation results indicate, the proposed system improves the performance of the classical methods, in terms of speech quality. It is important to note that the proposed method makes no assumptions about the noise statistics, in the phase modification block.

MMSE Based Noise PSD Tracking with Low Complexity [10] present a low complexity method for noise PSD estimation. The algorithm is based on a minimum mean-squared error estimator of the noise magnitude-squared DFT coefficients.

2.1.2 Statistical-model-based algorithms

Loveimi, Ahadi(Loveimi, Ahadi, 2010), investigate the effects of window shape and its length on the quality of phase-only and magnitude-only reconstructed speech. Speech signal is reconstructed via Least Square Error Estimation (LSEE) and Overlap Add (OLA) methods from its magnitude-only and phase-only spectra. They reconstruct the signal from its phase only spectrum to check the effect and importance of phase. For reconstructing the speech from its magnitude spectrum they select sequence of random uniformly distributed numbers in the range of$(-\pi, \pi), \varphi,$ as the phase sequence, or substitute phase spectrum with zero. To reconstruct speech only from its phase spectrum, typically, magnitude spectrum is set to unity.

Fardkhaleghi & Savoji [12],investigate the role of phase spectrum in speech enhancement using Wiener filtering and Martin's minimum statistics with No a priori information on the original phase. They use an optimization algorithm for phase correction of each processed frame by match the waveform of the zero-phase Wiener filtered speech to the conventional filter output obtained with noisy phase characteristic. All used optimization algorithms begins from a user defined initial condition, and tries to find the minimum of the cost function after certain number of iterations. The good results was obtained when minimizing the Wiener filter impulse response dispersion.

Senapati [13] ,Propose Log Gabor Wavelet (LGW) based Long Term Squared Spectral Amplitude estimator using the Maximum a Posteriori (MAP) criterion. They use LGW subject to the features of speech LGW coefficients flatness, symmetry, infinite derivability, rotation invariance. The proposed methods are compared against different existing methods from spectral subtractive algorithms, Wiener type algorithm, wavelet based and statistical model based algorithms. The performance evaluation results show that the proposed models show improvement against existing speech enhancement algorithms in almost all noise conditions with different SNRs. in the experimental 8 speech enhancement algorithms evaluated were used from four different classes of enhancement algorithms were chosen: spectral subtractive, Wiener type algorithm, wavelet and statistical model based algorithms.

Rao and Murthy [14], tries to improve the Wiener filter method that takes into account the non-uniform effect of colored noise on the spectrum of speech they uses the cross-correlation between the speech and noise signals they with nonlinear sub-band Bark scale frequency spacing approach to reduce colored noise and use Nonlinear technique (half wave rectification) to regenerate the degraded harmonics of the distorted signal . Because Wiener filter assumed that the speech and noise signal are zero mean and uncorrelated in each sub-band and the autocorrelation sequence of noisy speech signal is not exactly equal to the sum of the autocorrelations of the noise and clean speech signals so we cannot neglect the cross-correlation between clean and noise signal. They use simple implementation in regeneration process as when the estimated clean speech process is reliable the harmonic regeneration process is not needed else the regeneration process is needed.

McCallum & Guillemin [15],presented a new MMSE STSA speech enhancement system, named NMS MMSE, that is derived based on a noise model that allows for both deterministic and stochastic noise components. These deterministic components are seen in a variety of common noise sources.

Senapati and Bhende [16] , deals with single-channel speech enhancement technique. Log Gabor Wavelet (LGW) is investigated in speech enhancement approach and a novel speech enhancer by Bayesian Maximum a Posteriori (MAP) based Marginal Statistical Characterization (MSC) is developed. The PDF of the LGW filtered speech coefficient is modeled with Generalized Laplacian Distribution (GLD), which allows a high approximation accuracy for Laplace distributed real and imaginary parts of the speech coefficients.

Schwerin & Paliwal [17] Investigate a Real-Imaginary-modulation AMS framework for speech enhancement, in which the real and imaginary parts of the modulation signal are processed in secondary AMS procedures. The advantages with this assumption is the additive noise in the modulation signal and noisy acoustic phase is not used to reconstruct speech. Using the MMSE magnitude estimation to modify modulation magnitude spectra, initial experiments presented in this work evaluate if these advantages translate into improvements in processed speech quality. Also they apply presence uncertainty and log-domain processing on MMSE magnitude estimation in the RI-modulation framework is also investigated by estimate the clean speech from the noisy observations so as Minimizing the mean-square error of the log-

modulation magnitude spectrum between the modulation magnitude spectra of the clean and estimated speech.

2.1.3 Subspace algorithms

Krishnamoorthy & Prasanna [18], Finds that combined TSP method gives relatively better performance compared to temporal or spectral processing alone. They propose noisy speech enhancement method by combining linear prediction (LP) residual weighting in the time domain and spectral processing in the frequency domain to provide better noise suppression as well as better enhancement in the speech regions. Linear prediction based temporal processing used to identify and enhance the excitation source based speech-specific features existing at the gross and fine temporal levels. The gross level features are computed by estimating the sum of the peaks in the discrete Fourier transform (DFT) spectrum, smoothed Hilbert envelope of the LP residual and modulation spectrum values, all from the noisy speech signal. These parameter are used to drive a weight function to obtain the temporally processed speech signal. The fine level features are identified using the knowledge of the instants of significant excitation. There system is performed only on the high SNR regions of the spectrally processed speech. This require an estimate of pitch information and is computed from the autocorrelation of the HE of temporally processed LP residual

2.1.4 Binary mask algorithms

Sanam & Shahnaz [19],Enhance the speech using custom nonlinear thresholding function called Teager energy operator and apply this operator on upon the Wavelet Packet (WP) coefficients of the noisy speech. Teager energy operator is capable of switching between modified hard and semisoft thresholding functions depending on a parameter that decides the signal characteristics under consideration. Teager energy operator is nonlinear operator proposed by Kaiser (1993), capable to extract the signal energy based on mechanical and physical considerations. They develop a custom thresholding function in the WP domain for enhancing the noisy speech. TEO is a popular way to estimate the speech signal energy, instead of direct employment of the TEO on the noisy speech, we apply the TEO on the WP coefficients of the noisy speech. They determine an appropriate threshold by performing the statistical modeling of the TE operated WP coefficients of the noisy speech and employ a new custom thresholding

function that switches between the modified hard and semisoft thresholding function according to the signal characteristics thus yielding an enhanced speech.

Lotter & Vary[20], presents two spectral amplitude estimators for acoustical background noise suppression based on maximum a posteriori estimation and super-Gaussian statistical modelling of the speech DFT amplitudes. Motivated by the central limit theorem, real and imaginary parts of both speech and noise DFT coefficients are very often modelled as zero-mean independent Gaussian. The spectral amplitudes are of special importance, because the phase of the Fourier coefficients can be considered unimportant from a perceptual point of view. The PDF of the real and imaginary parts of the noise spectral coefficients will according to the central limit theorem be closer to a Gaussian function. The Gaussian and super-Gaussian function is used to model the PDF in MAP of the speech spectral amplitude. Also they introduce a joint MAP estimator of the amplitude and phase with approximation of the Bessel function to cope with an underlying Gamma model or the model that minimizes the Kullback divergence towards the measured data

Stark and Kamil[21], Employs noise estimates to compensate the phase spectrum for additive noise distortion is formulated. They control the degree to which the conjugates reinforce or cancel by altering their angular relationship. An antisymmetry function is used for this purpose. We make the degree of phase spectrum compensation dependent on the magnitude of noise spectral estimates. Zero weighting is given to the values corresponding to non-conjugate vectors of DSTFT. The next step in the computation of the compensated phase spectrum is to offset the complex spectrum of the noisy speech by the additive real-valued frequency-dependent Λ (n, k) compensation function. The compensated phase spectrum is recombined with the noisy magnitude spectrum to produce a modified complex spectrum

Chapter 3

3 Mathematical Algorithme Background

In this chapter, we discuss the mathematical background for the presented algorithm in this thesis. This comes with the fact of spectral technique importance in improving degraded speech signal.

3.1 Fast Fourier transform

Fast Fourier transform (FFT) is an algorithm to compute the discrete Fourier transform (DFT) and its inverse. Fourier analysis converts time (or space) to frequency and vice versa; FFT rapidly computes such transformations by factorizing the DFT matrix into a product of sparse (mostly zero) factors.

The DFT and IDFT can be defined as follows.

$$X(k) = \sum_{n=0}^{N} x(n)\, e^{-i2\pi kn/N} \qquad k = 0,1,2\cdots,N-1. \quad \text{N DFT coefficients } 3.1$$

Where $x(n)\ n = 0,1,2\cdots,N-1$ is a uniformly sampled sequence, T is sampling interval. $e^{-i2\pi kn/N}$ is the N-th root of unity, and $X(k)\ k = 0,1,2\cdots,N-1$ is the k-th DFT coefficient.

Evaluating this definition directly requires $O(N^2)$ operations: there are N outputs X_k, and each output requires a sum of N terms.

3.2 Minima Controlled Recursive Averaging (MCRA) Algorithm

Minima Controlled Recursive Averaging (MCRA) [30-33] is an technique for noise estimation. The noise component is estimated by the speech presence probability and speech absence probability within the sub bands by averaging past spectral power values and using a smoothing parameter that is adjusted by the signal presence probability in sub bands. Presence of speech in sub bands is determined by the ratio between the local energy of the noisy speech and its minimum within a specified time window[23] .

However, rather than employing a voice activity detector and restricting the update of the noise estimator to periods of speech absence and rather than computing a weighted average based on the instantaneous spectral magnitudes of the degraded speech and estimated noise, we adapt the smoothing parameter in time and frequency according to the speech presence probability.

MCRA technique cannot take full consideration of the inter-frame correlation of voice activity since the noise power estimate is adjusted by the speech presence probability depending on a current observation.[24]

Assume $\mathrm{x}(n)$ and $\mathrm{d}(n)$ denote speech and uncorrelated additivenoise signals, respectively, where n is a discrete-time index.

The observed signal

$$y(n) = x(n) + d(n) \qquad 3.2$$

$$Y(k) = X(k) + D(n) \qquad 3.3$$

is divided into overlapping frames by the application of a window function and analyzed using the short-time Fourier transform (STFT).

$$Y(k,l) = \sum_{n=0}^{N+1} y(n + lM)\, h(n) e^{-j\left(\frac{2\pi}{2}\right)nk} \qquad 3.5$$

Where k is the frequency index, l is the time frame index, h is an analysis window of size N, and M is the frame update step in time.

assume

$$H_0(k,l) : Y(k,l) = D(k,l) \qquad 3.6$$

$$H_1(k,l) : Y(k,l) = X(k,l) + D(k,l) \qquad 3.7$$

Where $X(k,l), D(k,l)$ represent the STFT of the clean and noise signals, respectively. H_0 , H_1indicate speech absence and presence in the lth frame of the kth sub band respectively.

If we let

$$\lambda(k,l) = \mathbb{E}[|D(k,l)|^2] \qquad 3.8$$

Where λ denote the variance of the noise in kth subband. Applying a temporal recursive smoothing to the noise measurement during periods of speech absence to obtains the noise power estimate as follows:

$$H_0{}'(k,l) : \hat{\lambda}_d(k,l+1) = \alpha_d \hat{\lambda}_d(k,l+1) + (1-\alpha_d)|Y(k,l)|^2 \qquad 3.9$$

$$H_1{}'(k,l) : \hat{\lambda}_d(k,l+1) = \hat{\lambda}_d(k,l+1) \qquad 3.10$$

Where $\alpha_d (0 < \alpha_d > 1)$ is a smoothing parameter and $H_0{}' and H_1{}'$ designate hypothetical speech absence and presence.

We make a distinction between the hypotheses in (3.6), used for estimating the clean speech and the hypotheses in (3.7), which control the adaptation of the noise spectrum, and deciding if speech is absent H_0 or present H_1 which more destructive when estimating the signal than the noise.

Let $p'(k,l) \triangleq P\big(H_1{}'(k,l)|Y(k,l)\big)$

$$\hat{\lambda}_d(k,l+1) = \hat{\lambda}_d(k,l)\, p'(k,l) + \big[\alpha_d \hat{\lambda}_d(k,l+1) + (1-\alpha_d)|Y(k,l)|^2\big] * \big(1 - p'(k,l)\big)$$

$$= \tilde{\alpha}_d(k,l)\, \hat{\lambda}_d(k,l) + [1 - \tilde{\alpha}_d(k,l)]|Y(k,l)|^2 \qquad 3.11$$

where $\tilde{\alpha}_d$ is a time-varying smoothing parameter that is adjusted by the speech presence probability.

$$\tilde{\alpha}_d(k,l) \triangleq \alpha_d + (1-\alpha_d)p'(k,l) \qquad 3.12$$

The conditional speech presence probability $P\big(H_1{}'(k,l) \mid Y(k,l)\big)$ is calculated as

$$P\big(H_1{}'(k,l) \mid Y(k,l)\big) = \begin{cases} \alpha_p P\big(H_1{}'(k,l-1) \mid Y(k,l-1)\big) + \big(1-\alpha_p\big) & if\ S_r(k,l) > \delta \\ \alpha_p P\big(H_1{}'(k,l-1) \mid Y(k,l-1)\big) & otherwise \end{cases}$$

$$3.13$$

Where $\alpha_p\big(0 < \alpha_p > 1\big)$ is a smoothing parameter and δ is a probability threshold of speech signal presence.

$S_r(k,l) \triangleq S(k,l)/S_{min}(k,l)$ denotes the ratio between the local energy of noisy speech, $S(k,l)$ and its determined minimum, $S_{min}(k,l)$ of the current frame.

The magnitude of the power spectra of the input signal is recursively smoothed using a constant smoothing parameter; like the continuous spectra minimum tracking method. It uses the fixed window minimum power analysis during the identified speech silence using the threshold constant. The fixed window minimum analysis method is reminiscent of the minimum statistics method where a delay was seen regarding updating of the minimum power of the speech signal A similar pattern of delay could be anticipated for the minima controlled recursive averaging method, 'This method also requires a thresholding constant in order to determine speech pauses, within which the minimum power is updated.

3.3 Karhunen–Loeve transform (KLT)

KLT is also known as the Hotelling transform, consist of the eigenvectors of the autocorrelation matrix. KLT is most advance mathematical algorithm available in the year of 2008 to achieve both noise filtering and data compression in signal processing [25], [26], [27]. KLT is a representation of a stochastic process as an infinite linear combination of orthogonal functions, analogous to a Fourier series representation of a function on a bounded interval. The KLT linear transformation in the Hilbert space and finds the best basis (Eigen-basis) in the Hilbert space spanned by the Eigen functions of the autocorrelation of $x\ (n)$[28].

Let the L × 1 vector $x(n)$ denote a data sequence drawn from a zero-mean stationary process with the correlation matrix R_{xx}. , this matrix can be diagonalized as follows :

$$Q^T\, R_{xx}\, Q = \Lambda \qquad 3.14$$

where Q = [q_1, q_2 ▪ ▪ ▪ q_L] and Λ = diag [λ_1 ,λ_2, ▪ ▪ ▪ λ_L] are orthogonal and diagonal matrices respectively. The orthonormal vectors q_1, q_2 ▪ ▪ ▪ q_L are the eigenvectors corresponding, respectively, to the eigenvalues λ_1 ,λ_2, ▪ ▪ ▪ λ_L of the matrix R_{xx}..

The vector $x(n)$ can be written as a combination (expansion) of the eigenvectors of the correlation matrix R_{xx}. as follows:

$$x(n) = \sum_l^L a_{x,l}(k) q_l \qquad 3.15$$

where

$$a_{x,l}(n) = q_l^T x(n) \quad , l = 1, 2, \ldots, L \qquad 3.16$$

are the coefficients of the expansion. The representation of the random vector $x(n)$ described by 3.13 and 3.14 is the KLE where 3.13 is the synthesis part and 3.14represents the analysis part. From 3.14, we can easily verify that

$$E\left[a_{x,l}(k)\right] = 0 \, , l = 1, 2, \ldots, L \qquad 3.17$$

and

$$E\left[a_{x,i}(n)\, a_{x,j}(kn)\right] = \begin{cases} \lambda_i \, , i = j \\ 0 \, , \quad i \neq j \end{cases} i, j = 1, 2, \ldots, L \qquad 3.18$$

It can also be checked from 3.14 that

$$\sum_l^L a_{x,l}^2 \, (k) = \|x(n)\|_2^2 \qquad 3.19$$

where $\|x(n)\|_2$ is the Euclidean norm of x(n). The previous expression shows the energy conservation through the KLE process.

Karhunen–Loeve transform has an efficient calculation properties, KLT is an orthonormal transform and its coefficients satisfy the non-correlated condition. KLT minimizes the Mean square error that occurs due to truncation of the vector of the transform coefficients by the rejecting transform coefficients with small variances. So it is optimum depending on the localization of the signal energy property and it maximize the number of transform coefficients which are insignificant and might be quantized to 0[27].

The main shortcoming of the Karhunen–Loeve transform is that its basis functions depend on the transformed signal.

3.4 Maximum a Posteriori (MAP) Estimators

Maximum a Posteriori (MAP) estimation is an important task in many applications of probabilistic graphical models. It's an alternative to maximum Likelihood estimation. Also, MAP algorithm is often used as an alternative to the MMSE algorithm in conditions in which it is extremely difficult to derive the average a posteriori PDF in closed form. MMSE approach

aims to find the average of the a posteriori PDF$(A|R)$. MAP estimator called hard-thresholding so MAP keeps the observation when the signal is larger than the noise level, and discards the observation otherwise. So the MAP gain function is binary valued.[20]. MAP estimator might not be unique if the model is not identifiable.

Computationally effective MAP solution is as following:

$$\hat{A} = argmax_A\, p(A|R) = argmax_A = \frac{p(R|A)p(A)}{p(R)} \quad 3.20$$

$$p(R|A)p(A) \simeq A^{\mathcal{V}-1/2}\, exp\ \left\{-\frac{A^2}{\sigma_N^2} - \left(\frac{\mu}{\sigma_S} - \frac{2R}{\sigma_N^2}\right)\right\} \quad 3.21$$

This mean that parameter that are most likely are chosen as estimate. Where the Joint MAP amplitude and phase estimator

$$\hat{A} = argmax_A p(A,\alpha|Y) = argmax_A = \frac{p(Y|A,\alpha)p(A,\alpha)}{p(Y)} \quad 3.22$$

$$\hat{\alpha} = argmax_\alpha p(A,\alpha|Y) = argmax_\alpha = \frac{p(Y|A,\alpha)p(A,\alpha)}{p(Y)} \quad 3.23$$

The objective is to maximize $p(R|A)p(A)$, $p(Y|A,\alpha)p(A,\alpha)$, since p(R) is independent of A.

While the PDF of the noisy spectrum Y conditioned on the speech amplitude A and phase α can be written as joint Gaussian:

$$p\,(Y|A,\alpha) = \frac{1}{\pi\sigma_N^2}\, exp\left(-\frac{|Y - A\, e^{j\alpha}|^2}{\sigma_N^2}\right) \quad 3.24$$

then PDF is obtained for the density of the noisy amplitude given the speech amplitude A after polar integration

$$p(R|A) = \frac{2R}{\sigma_N^2} exp\left(-\frac{R^2+A^2}{\sigma_N^2}\right)\, I_0\left(\frac{2AR}{\sigma_N^2}\right) \quad 3.25$$

Where I_0 denotes the modified Bessel function of the first kind and zeroth order.

$$I_0 = \frac{1}{\sqrt{2\pi x}}\ e^x \quad 3.26$$

3.5 Minimum Mean-Squared Error Signal Estimation and Model-Based Approach

A central problem in estimation is to recover a set of unobservable parameters from data corrupted by noise[29]. Minimum mean-squared error (MMSE) estimation is one of the most commonly used approaches for general signal estimation tasks [30].

A fundamental problem in statistics and statistical signal processing is the estimation of a random variable given a set of observed random variables[31] .

$$X = (X_1, X_2, \ldots, X_m)^T$$

The estimate $\hat{S}$ is a function of the observed random vector$X$. The vector variable X is often called an independent variable and S is called a dependent variable, to estimate a function $T(x)$ of a sample x, whereas an estimator $\hat{X}$ is a function of the random variable X:

$$\hat{X} = T(\boldsymbol{X}) \qquad 3.27$$

The mean square error (MSE) of the estimate $\hat{X}$ is defined as

$$MSE\ \varepsilon = E(||\hat{X} - X\ ||^2) \qquad 3.28$$

in the frequency domain the error signal vector between the estimated and desired signals as[32].

$$\varepsilon(j\omega) = Z(j\omega) - X(j\omega) \qquad 3.29$$

$$= H(j\omega).Y(j\omega) - X(j\omega)$$

Which can be written in terms of speech distortion (ε_x) and residual noise (ε_v) as

$$\varepsilon(j\omega) = \varepsilon_x(j\omega) - \varepsilon_v(j\omega) \qquad 3.30$$

Where

$$\varepsilon_x(j\omega) = \{H(j\omega) - 1\}.\ X(j\omega)$$

$$\varepsilon_v(j\omega) = H(j\omega).\ V(j\omega) \qquad 3.31$$

then MSE

$$J[H(j\omega)] = E[|\varepsilon(j\omega)|^2]$$

$$= \emptyset_x(\omega) + |H(j\omega)|^2.\emptyset_y(\omega) - 2R\left[H(j\omega)\,\emptyset_{xy}(\omega)\right] \quad 3.32$$

Where the cross spectrum between the observation and speech signals is

$$\emptyset_{xy}(\omega) = E[Y(j\omega)\,X^*(j\omega)] \quad 3.33$$

Then

$$J[H(j\omega)] = \emptyset_x(\omega) + |H(j\omega)|^2.\emptyset_y(\omega) - 2R[H(j\omega)\,\emptyset_x(\omega)] \quad 3.34$$

$J[H(j\omega)]$ In terms of speech distortion (ε_x) and residual noise (ε_v)

$$J[H(j\omega)] = E(|\varepsilon_x(j\omega)|^2) + (|\varepsilon_v(j\omega)|^2)$$

$$J[H(j\omega)] = J_x[H(j\omega)] + J_v[H(j\omega)] \quad 3.35$$

(MMSE) estimator describes the approach which minimizes the mean square error (MSE), which is a common measure of estimator quality. One of the most commonly used approaches for general signal estimation tasks is the minimum mean-squared error (MMSE) estimation[30].

3.6 Spectrum Power estimator based on Zero Cross

Zero-crossing rate is a measure of number of times in a given time interval/frame that the amplitude of the speech signals passes through a value of zero, Figure 3.1 [33]. Speech signals are broadband signals and interpretation of average zero-crossing rate is therefore much less precise However, rough estimates of spectral properties can be obtained using a representation based on the short time average zero-crossing rate . Since the zero crossing rates are low for voiced part and high for unvoiced part where as the energy is high for voiced part and low for unvoiced part and high frequencies imply high zero crossing rates, and low frequencies imply low zero-crossing rates.

The Zero Crossing (ZC) measure is a nonlinear measure that has traditionally used for analyzing the speech signal [34].

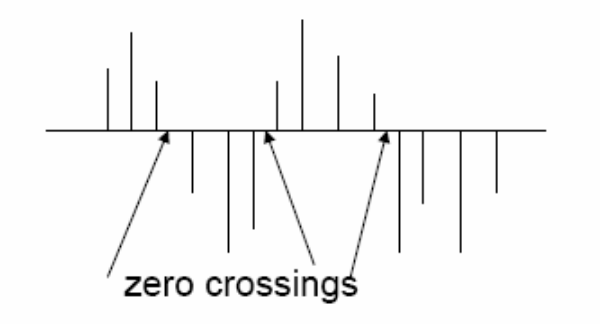


Figure 3.1 zero crossing [33]

ZC calculated as

$$ZC(n) = \sum_{n=-\infty}^{\infty} \frac{|sgn\{x(n)\} - sgn\{x(n-1)\}|}{2} \qquad 3.36$$

When the sgn function is defined as

$$sgn\{x(n)\} = \begin{cases} 1 & x(n) \geq 0 \\ -1 & x(n) < 0 \end{cases} \qquad 3.37$$

ZC (n) =1 of there is a changing of sign happened between the sample n and sample n+1 zero otherwise.

When zero crossing occurs the sample index n will stored in the vector ZC_{ix}

3.7 Soft Mask Formulation

The nonstationarity of the speech and noise signals leads to time-varying between these signals this variances of the speech and noise spectra are the main point in most numerical models[35].

Their variances can be modeled as unknown but deterministic parameters. Although, the knowledge about the priori SNR ε_k can be assumed to be unknown but deterministic using noise PSD estimation methods.

The instantaneous SNR, ε_I

$$\varepsilon_{I,k} \equiv \frac{X_k^2}{D_k^2} \qquad 3.38$$

We express the ideal binary mask (IdBM) rule as

$$\hat{X}_k^2 = G_k^2 \,.\, Y_k^2 = \begin{cases} Y_k^2 \ , \ if\ \varepsilon_{I,k} \ \geq 1 \quad where\ G_k = 1 \\ 0 \quad , \ if\ \varepsilon_{I,k} \ < 1 \quad where\ G_k = 0 \end{cases} \qquad 3.39$$

Where G is a Bernoulli distributed random variable taking the value of $0 < G_k \leq 1\ \forall k = 1,2,\cdots,k$ [36]

The goal here is to find a set of suppression gains that minimize the traceable loudness of the remaining noise as well as the speech distortion. As mentioned in [36] the optimal $\widehat{G_k}$ can be found $\widehat{G_k} = 1\ and\ \widehat{G_k} = 0$.

The equation can be formulated as binary hypothesis model:

$$H_0 : \varepsilon_{I,k} \ < \theta : masker\ dominates \qquad 3.40$$

$$H_1 : \varepsilon_{I,k} \ \geq \ \theta : target\ signal\ dominates \qquad 3.41$$

The weighted variance of the noise signal can be tracked by using a recursive averaging process on the instantaneous noise power, which is unfortunately not directly available since in most cases the noise signal is corrupted by the speech signal in the input signal. Now combining these two assumptions:

$$\hat{X}_k^2 = \mathrm{E}(G_k)\, Y_k^2$$

$$= \{E(G_k|H_1)\, P(H_1) + E(G_k|H_0)\, P(H_0)\}\, Y_k^2 \qquad 3.42$$

where $P(H_1)$ denotes the probability that hypothesis H_1 is true, $\mathrm{E}(G_k|H_1)$ denotes the gain function assuming that hypothesis H_1 is true and $\mathrm{E}(G_k|H_0)$ denotes the gain function assuming that hypothesis H_0 is true

3.8 Soft Masking by Incorporating a Priori SNR Uncertainty

Assuming independence between the clean speech and noise magnitude-squared spectra, the probability density of X_k^2 is exponential given as

$$f_{X_k^2}(X_k^2) = \frac{1}{\sigma_x^2(K)} e^{-\frac{X_k^2}{\sigma_x^2(K)}} \quad 3.43$$

$$f_{D_k^2}(D_k^2) = \frac{1}{\sigma_d^2(K)} e^{-\frac{D_k^2}{\sigma_x^2(K)}} \quad 3.44$$

$\sigma_d^2 = E\{D_k^2\}$, $\sigma_x^2 = E\{X_k^2\}$

we can easily use 3.34 and 3.35 to model the hypothesis probability given the a priori SNR $\mathcal{E}$ [35] .As we do not use any other constraint or assumption, we refer to this hypothesis probability as the a priori SNR uncertainty.

Using the exponential models for X_k^2 *and* D_k^2 we can drive the probability density of $\mathcal{E}_I$

$$f_{\mathcal{E}_I}(\mathcal{E}_I) = \frac{\mathcal{E}}{(\mathcal{E}+\mathcal{E}_I)^2}\, u(\mathcal{E}_I) \quad 3.45$$

Where $u(\mathcal{E}_I)$is the step function. For an arbitrary SNR threshold θ, the hypothesis probability needed in 3.42 is computed as

$$P(H_1) = P(\mathcal{E}_I > \theta) = \int_\theta^\infty f_{\mathcal{E}_I}(\mathcal{E}_I)\, dz = \frac{\mathcal{E}}{(\mathcal{E}+\theta)} \quad 3.46$$

We refer to this probability as priori since it does not require information from the noise-corrupt observations and does not need the assumption of 3.3 as mentioned before $\mathcal{E}$ can be estimated using the "decision-directed" approach in conjunction with noise PSD estimation algorithms.

Then

$$\hat{X}_k^2 = \frac{\mathcal{E}_k}{(\mathcal{E}_k+\theta)}\, Y_k^2 \quad 3.47$$

Where $\mathcal{E}_k$ is the a priori SNR $\mathcal{E}_k = \frac{\sigma_x^2}{\sigma_d^2}$.

3.9 Soft Masking Based on Posteriori SNR Uncertainty

Soft Masking by Incorporating a Posteriori SNR estimator did not incorporate information about the noisy observations, as it relied solely on a priori information about the instantaneous

SNR $\mathcal{E}_I$[35]. It is reasonable to expect that a better estimator could be developed by incorporating posteriori information about the SNR at each frequency bin. In this case, we incorporate the assumption given in 3.3 to compute the hypothesis probability, which is referred to as a posteriori SNR uncertainty.

This hypothesis probability can be computed as the posteriori

Probability of $\mathcal{E}_{I,k} > \theta$ as follows:

$$P(H_1) = P\left(\mathcal{E}_{I,k} > \theta \mid Y_k^2\right)$$

$$= P\left(\hat{X}_k^2 > \frac{\theta}{(\theta+1)}\, Y_k^2 \mid Y_k^2\right)$$

$$= \int_{\frac{\theta}{(\theta+1)}}^{Y_k^2} f_{\hat{X}_k^2}\left(z \mid Y_k^2\right) dz \qquad 3.48$$

$$P\left(\mathcal{E}_{I,k} > \theta \mid Y_k^2\right) = \begin{cases} \frac{e^{\left(\frac{v_k}{\theta+1}\right)}-1}{e^{v_k}-1} & if\ \sigma_x^2(k) \neq \sigma_d^2(k) \\ \frac{1}{(\theta+1)} & if\ \sigma_x^2(k) = \sigma_d^2(k) \end{cases} \qquad 3.49$$

$$\hat{X}_k^2 \begin{cases} \frac{e^{\left(\frac{v_k}{\theta+1}\right)}-1}{e^{v_k}-1}\, Y_k^2 & if\ \sigma_x^2(k) \neq \sigma_d^2(k) \\ \frac{1}{(\theta+1)}\, Y_k^2 & if\ \sigma_x^2(k) = \sigma_d^2(k) \end{cases} \qquad 3.50$$

The SMPO gain function 3.40 is dependent on both $\mathcal{E}$ and γ values.

3.10 Wave Atoms

It is a natural question to ask whether other waveforms than curvelets would yield comparable sparsely results. The short answer is that the parabolic scaling is essential, allowing only for slight variations on a fixed theme.

Since a complete collection of wave packets ϕμ (x) must "span" all positions and frequencies, we will call it a phase-space tiling, with wave packets as tiles. Some tiling's are more interesting

than others. We say a tiling is universal if it treats democratically all positions and orientations. In that case,

- The geometry of the tiling in space must be Cartesian, or approximately so; and
- The geometry of the tiling in frequency must be polar, or approximately so.

In what follows we limit our discussion to two space variables. This is not an essential restriction.

Universality as above suggests that two parameters should suffice to index a lot of known wave packet architectures: α to index whether the decomposition is "multi-scale" (α = 1) or not (α = 0); and β to indicate whether basis elements should be isotropic (β = α) or, on the contrary, elongated and anisotropic (β < α).

In terms of phase-space localization of the wave packets, we will require that

- The essential support of $\varphi_\mu(x)$ be of size $\sim 2^{-\alpha j}$ vs. $2^{-\beta j}$ as scale j, with oscillations of wavelength $\sim 2^{-j}$ tranverse to the ridge; and
- The essential support of $\hat{\varphi}_\mu(x)$ (ξ) be of size $\sim 2^{-\alpha j}$ vs. $2^{\beta j}$ as scale j, at a distance 2^j from the origin.

Figure 3.2 summarizes these micro localization properties.

The description in terms of α and β will clarify the connections between various transforms of modern harmonic analysis. Curvelets correspond to α = 1, β = 1/2, wavelets are α = β = 1, ridgelets are α = 1, β = 0, and the Gabor transform is α = β = 0. The horizontal segment at β = 1/2 indicates the only wave packet families that yield sparse decompositions of Fourier Integral Operators. The situation is summarized in Figure 3.3.

Note that the range of possible transforms in Figure 3.3 could presumably extend beyond the triangle shown—the horizontal segment indicating sparse FIO, on the other hand, does not. All the transforms within the triangle can be realized as tight frames of $L^2(R^2)$.

Wave atoms are sometimes more adequate than curvelets for numerical simulations of wave equations because of their low separation rank.

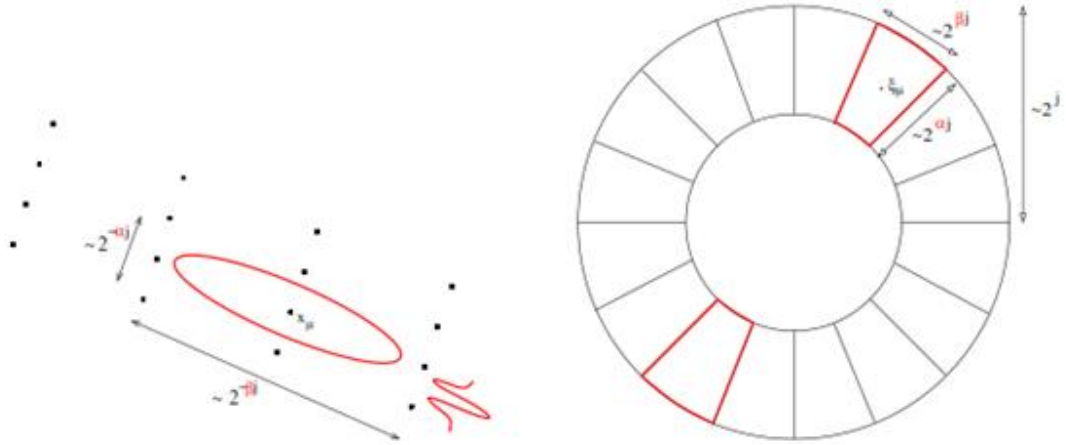


Figure 3.2 *Essential support of a wave packet with parameters (α ,β), in space (left), and in frequency (right). The parameter α indexes the multiscale nature of the transform, from 0 (uniform) to 1 (dyadic). The parameter β measures the wave packet's directional selectivity, from β = 0 (best selectivity) to β = 1 (poor selectivity). Curvelets are the special case α = 1, β = 1/2.*

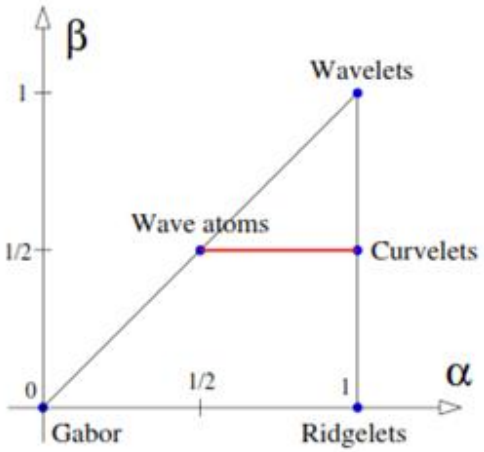


Figure 3.3 Classification of a and ß for Wavelets, Curvelets and Wave atoms

Wave atom is type of wave packet algorithm that provide linear systems representation of hyperbolic deferential equations with smooth, and time-independent coefficients. With is $O\ (N^2\ log\ N)$ complexity for N-by-N arrays, for forward and inverse transforms.

The representation of the wave atom based on sparsity of the matrix representation of Green's function and also exploits its low-rank block structure after separation of the spatial indices. Wave atoms offer a uniquely structured representation of the time-dependent Green's function in the sense that the resulting matrix is universally sparse over the class of C^{∞} coefficients, even for "large" times.

We need to classify various wave-packet transforms as phase-space tiling's.

We write wave atoms as

$$\varphi_{\mu}(x)\ , \mu = (j,m,n) = (j,m_1,m_2,n_1,n_2) \qquad 3.51$$

Where j, m_1, m_2, n_1, n_2 are integer index point (x_{μ}, ξ_{μ}) in phase-space, as?

$$x_{\mu} = 2^{-j}n \quad , \varepsilon_{\mu} = \pi 2^{j}m\ , c_1 2^{j} \leq max_{i=1,2}|m_i| \ \leq c_2 2^{j} \qquad 3.52$$

where c_1 , $c_2$two positive constants are left unspecified for convenience, but whose values will be implied by the specifics of the implementation. Heuristically, the position vector x_{μ} is the center of $\varphi_{\mu}(x)$ and the wave vector ξ_{μ} determines the centers of both bumps of $\varphi_{\mu}(\xi)$ as $\pm\xi_{\mu}$.

Note that the range of m needs to be further reduced to $m_2 > 0$ to account for the central symmetry of the Fourier transform of real-valued functions about the origin in ξ. Wave atoms then need to obey a localization condition around the phase-space point $(x_{\mu}, \varepsilon_{\mu})$

$$\left|\hat{\varphi}_{\mu}(\xi)\right| \leq C_M\,.\,2^{-j}\left(1+2^{-j}\left|\xi-\xi_{\mu}\right|\right)^{-M} + \ C_M\,.\,2^{-j}\left(1+2^{-j}\left|\xi+\xi_{\mu}\right|\right)^{-M} \text{ for all M } > \ 0 \qquad 3.53$$

and

$$\left|\varphi_{\mu}(\xi)\right| \leq C_M.\,2^{-j}\left(1+2^{-j}\left|x+x_{\mu}\right|\right)^{-M} \ for\ all\ M\ >\ 0. \qquad 3.54$$

The definition follows closely the notation of wavelet packets, where j controls the resolution scale, m and n controls the location in time and frequency domain. But the differences is that wave atoms obey the parabolical scaling wavelength: at scale $2^{-j}$the essential frequency

support is of size 2^jwhile at frequency 2^{2j}, the essential time support is of size 2^{-j}. The achievement of parabolical scaling wavelength is based on an architecture of decomposition like incomplete wavelet packet,

The wave equation can be expressed as

$$\hat{\psi}_m^0(\xi) = e^{-i\xi/2}\left[e^{i\alpha_m} g\left(\epsilon_m(\xi - \pi(m+.5))\right) + e^{-i\alpha_m} g\left(\epsilon_{m+1}(\xi - \pi(m+.5))\right)\right] \quad 3.55$$

$$\epsilon_m = (-1)^m \quad and\ \alpha_m = .5\pi(m+.5) \quad 3.56$$

The function g is an appropriate real-valued, C^∞ bump function, compactly supported on an interval of length 2¼, and chosen such that

$$\sum_m \left|\hat{\psi}_m^0(\xi)\right|^2 = 1 \quad 3.57$$

The simple algorithm for wavelet packets is then the following.

- Perform a FFT of size N of the samples $u(x_k)$.
- Max prop
- For each pair(j,m), wrap the product $\hat{\psi}_m^{\mathrm{j}}\,\hat{u}$ by periodicity inside the interval $\left[-2^{-j}\pi, 2^{-j}\pi\right]$ then perform an inverse FFT of size 2^j of the result to obtain $c_{j,m,n}$.
- Repeat over(j,m).
- Inverse transform
- unwrap the result of FFT of $c_{j,m,n}$ of each (j,m) on the frequency axis around the support of $\hat{\psi}_m^{\mathrm{j}}$
- Sum the contributions corresponding to all the couples
- Perform an inverse FFT, of size N, to obtain $u(x_k)$.

In 2D Wave Atoms can expressed as

$$\mu = (j,m,n) = (j,m_1,m_2,n_1,n_2)\ \ where\ m = (m_1,m_2)\ and\ n = (n_1,n_2). \quad 3.58$$

We write

$$\varphi_\mu^+(x_1,x_2) = \psi_{m1}^j\left(x_1 + 2^{-j}n_1\right)\psi_{m2}^j\left(x_2 + 2^{-j}n_2\right) \quad 3.59$$

$$\hat{\varphi}_{\mu}^{+}(\xi_1,\xi_2) = \hat{\psi}_{m1}^{j}(\xi_1)\, e^{-i2^j(x_1 n_1)} \quad \hat{\psi}_{m2}^{j}(\xi_2)\, e^{-i2^j(x_2 n_2)} \qquad 3.60$$

After preforming a Hilbert transform, $\mathrm{H}\psi_{m1}^{j}$ is another orthonormal basis of $L^2(\mathbb{R})$

$$\hat{\psi}_{m,n}^{j}(\xi) = \hat{\psi}_{m,n,+}^{j}(\xi) + \hat{\psi}_{m,n,-}^{j}(\xi) \qquad 3.61$$

$$\widehat{\mathrm{H}\psi_{m,n}^{j}}(\xi) = -i\hat{\psi}_{m,n,+}^{j}(\xi) + i\hat{\psi}_{m,n,-}^{j}(\xi)$$

$$\varphi_{\mu}^{-}(x_1,x_2) = H\,\psi_{m1}^{j}\left(x_1 + 2^{-j}n_1\right)\; H\psi_{m2}^{j}\left(x_2 + 2^{-j}n_2\right) \qquad 3.62$$

$$\varphi_{\mu}^{(1)} = \frac{\varphi_{\mu}^{+}+\varphi_{\mu}^{-}}{2},\ \varphi_{\mu}^{(2)} = \frac{\varphi_{\mu}^{+}-\varphi_{\mu}^{-}}{2} \qquad 3.63$$

provides basis functions with two bumps in the frequency plane, symmetric with respect to the origin, hence directional wave packets.

$\varphi_{\mu}^{(1)}$ and $\varphi_{\mu}^{(2)}$ form the wave atom frame and may be denoted jointly as φ_{μ}

3.10.1 Basic Repeated Squaring

Let us denote $u(t)$ for the couple $\left(p(t),\frac{dp}{dt}\right)$ and write the wave equation as the first-order system

$$\frac{du}{dt} = Au \qquad 3.64$$

with initial condition $u(0) = u_0$. The generator is

$$\begin{bmatrix} 0 & I \\ c^2(x)\,\Delta & 0 \end{bmatrix}$$

We define the propagator $\mathrm{E}(t)$ from $u(t) = \mathrm{E}(t)u_0 = e^{tA}\,u_0$

Since the solution $u(t)$ has two components, we need to introduce $e_1 = (1,0)$ and $e_2 = (0,1)$.

Choosing a small time step $\Delta\, t$ and a small tolerance ϵ. Denote by Trunc the operation of putting to zero all matrix elements below ϵ in absolute value.

Trunc operation consists in keeping track of two shifted diagonals, because there are two Hamiltonian flows.

Initialization:

Obtain $\tilde{A}(\mu\nu;\grave{\mu}\grave{\nu})$ an approximation to the wave atom matrix of the generator A, then

$$\tilde{E}(\Delta\, t,\mu\nu;\grave{\mu}\grave{\nu}) = \delta_{\mu\nu;\grave{\mu}\grave{\nu}} + \Delta\, t\, Trunc\left(\tilde{A}(\mu\nu;\grave{\mu}\grave{\nu})\right) \qquad 3.65$$

Iteration:

Forecast the biggest entries' location, then compute them as

$$\tilde{E}(2^{n+1}\,\Delta\, t;\mu\nu;\grave{\mu}\grave{\nu}) = \sum_{\grave{\mu}\grave{\nu}} Trunc\big(E(\Delta\, t,\mu\nu;\mu''\nu'')\, E(\Delta\, t,\mu''\nu'';\mu'\nu')\big) \qquad 3.66$$

Terminate: at time $\tau = 2^{n*}\,\Delta\, t$

To compute the solution $u(\tau)$ at time τ start with the coefficients

$$c_{\mu\nu}(0) = \langle u_0, \varphi_\mu, e_\nu\rangle \qquad 3.67$$

perform the matrix-vector multiplication,

$$\tilde{c}_{\mu\nu}(\tau) = \sum_{\grave{\mu}\grave{\nu}} \tilde{E}(\tau,\mu\nu;\grave{\mu}\grave{\nu})\, c_{\grave{\mu}\grave{\nu}}(0) \qquad 3.68$$

Chapter 4

4 Methodology and Design

The review of literature has produced reoccurring themes emphasizing the importance of the spectral technique in enhancing the degraded speech signal. This study developed and implemented a new technique in the field of speech enhancement using the combination of technique and wave atoms transform to achieve new result in this field. In this chapter, we present the methods that have been used in this study to reach new better results in speech enhancement.

Proposed Method

There are many technique and algorithm in speech enhancement that can be used alone or in combined with others to achieve the highest score of the measurement algorithm SNR or PESQ… etc. In our study we use PESQ and SNR scores in measuring the performance of the proposed technique.

In this research we use wave atom as main algorithm which is a recent addition to the repertoire of mathematical transforms of computational harmonic analysis. They come either as an orthonormal basis or a tight frame of directional wave packets, and they are particularly well suited for representing oscillatory patterns. Wave atoms have a sharp frequency localization that cannot be obtained from filter bank-based wavelet packets (pump function). Also keep in mind that wave atoms provide powerful tools for representing linear systems of hyperbolic differential equations with smooth and time-independent coefficients. Wave atoms offer a uniquely structured representation of the time-dependent Green's function, this means that the resulting matrix is universally sparse over the class of C^{∞} coefficients, even for “large" times. Wave atoms capture coherence of a pattern across and along oscillations whereas curvelets capture coherence only along the oscillations. Noise signal can be assumed as a low-rank component because noise spectra within different time frames are usually highly correlated with each other;

while the speech signal is regarded as a sparse component since it is relatively sparse in time–frequency domain.

Inside repeat squaring step, the algorithm use Phase Flow Method (PFM) to predict the location of the shifted diagonals, and eliminate Elements outside of those shifted band diagonals from calculation by constructing phase maps for nonlinear autonomous ordinary differential equations using initially constructing the phase map for small times, using a standard Orthogonal Deferential Equation (ODE) integration rule, and builds up the phase map for larger times using a local interpolation scheme with the group property of the phase flow. PFM makes repeated use of prior computations to calculate the phase map at the next (large) time step and computes an approximate phase flow by applying a local integrator for an initial small time step, and uses the group property of the phase flow. PFM also uses high order local interpolation procedures for larger pieces of time.

The continuous speech has pauses or silence periods even during speech activity, this fact may seriously degrade the accuracy of traditional speech enhancement models. The traditional methods use voice activity detectors to handle this problem. During speech activity, speech may not be presented in a particular frequency band. This problem was resolved using speech presence uncertainty estimator. In speech presence uncertainty, the modification is obtained by multiplying the spectral gain by the conditional speech presence probability, estimated for each frequency bin and each analysis frame by assuming that speech is absent or present and by dynamically computed parameters of the a priori SAP using two factors: a smoothing-update factor and a factor related to the k^{th} spectral component. The smoothing-update factor -which is based on a decision made in frequency band whether speech is present or absent- is computed by recursively averaging past spectral values of the a priori SAP.

The algorithm contains four important steps .The first step in the algorithm is to apply frame overlapping. Second step is to apply Fast Fourier transform to the overlapped signal. Third step is to apply Minima Controlled Recursive Averaging (MCRA). Fourth step is to apply the proposed multi transformation technique. Last step is to inverse all process reversely.

Step.1 Inter-frame Correlation

One major issue with the STFT-domain speech enhancement approach is the aliasing problem caused by circular convolution [2]. To solve this issue, we need to use either the overlap add or overlap save techniques. (Note: however, that even with overlap add/save procedure, aliasing cannot be completely avoided unless we use a unit gain, which will not give any noise reduction; but one can manage to minimize the effect by applying a proper windowing function such as the Kaiser one before FFT and after the IFFT.) With overlap frames, the STFT coefficients from neighboring frames are not independent and there is some correlation among them in principle. Researcher use 20ms frame duration with 50% overlapping.

Step 2 And Step 3 was previously described in chapter 3

Step.4 Transformation process

It is the main step in the proposed enhancement technique. This step will contain multiple transformation and multiple threshold type. We can deal with this step as block of steps. These steps are the proposed technique.

It is the main step in the proposed enhancement technique. This step will contain multiple transformation and multiple threshold type. We can deal with this step as block of steps. These steps are the proposed technique.

Algorithm 1 Pure Wave Atoms

This technique is built basically on applying wave atoms transform with different thresholding type and different noise estimator technique on the spectral domain (apply FFT).

1. Wave Atoms Dual Pump

In wave Atoms Dual Pump the threshold is applied on to the magnitude part of the wave atom transformation stage and kept the phase part with no change as shown in the Figure 4.1.

2. Wave Atoms Quad Pump Magnitude

In this technique we apply threshold on the absolute value of wave atoms transform with no change on the phase value, but we use quad pump function instead of dual as shown in the Figure 4.2 .

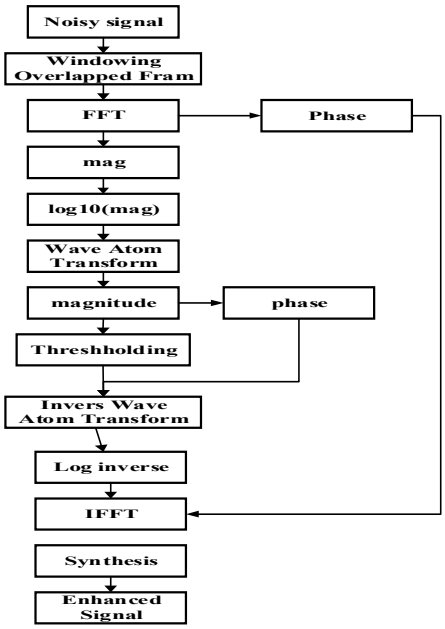


Figure 4.1 Wave Atoms dual pump diagram

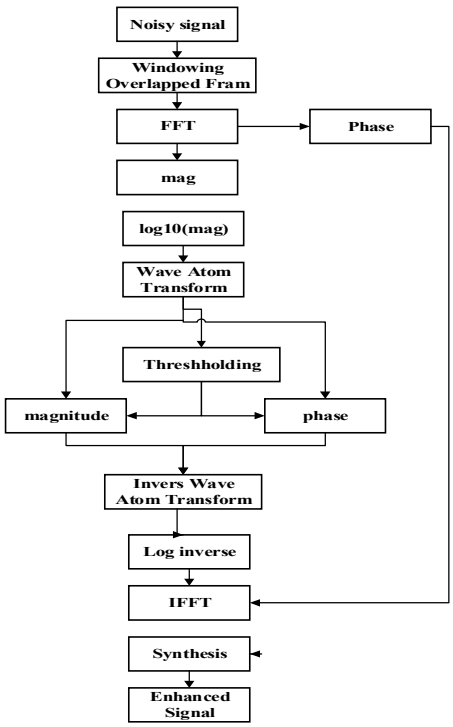


Figure 4.2 Wave Atoms quad pump diagram

3. Wave Atoms quad pump Cplxpair

In this technique the compound (real + complex) threshold applied on the whole pair of the wave atoms transform as descried in Figure 4.3.

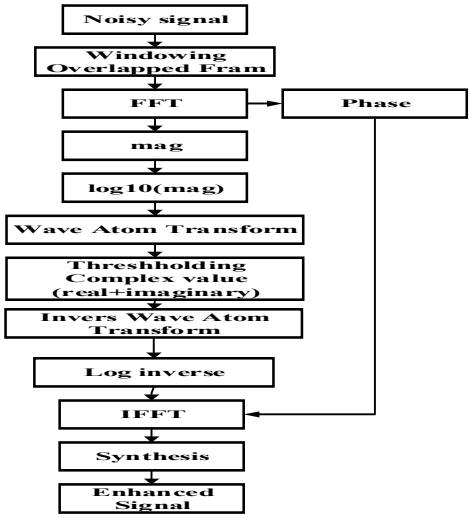


Figure 4.3 Wave Atoms Cplxpair diagram

4. Wave Atoms Cplxpair Imaginary

In this technique the threshold is applied to the complex part of the wave atom transformation stage, and the real part associated with the complex is kept with no change as descried in the Figure 4.4.

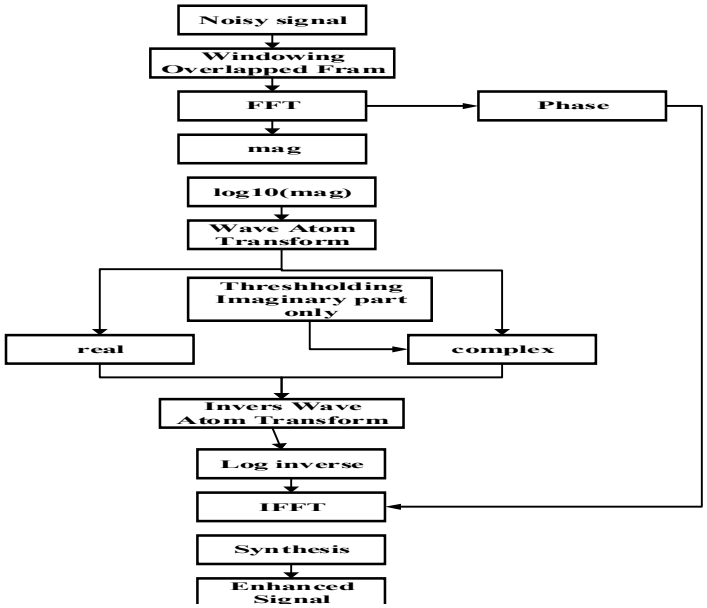


Figure 4.4 Wave Atoms Cplxpair Imaginary daigram

Algorithm 2 Wave atoms with KLT

In the next algorithm we add one stage to the previous descried algorithm. We add Karhunen–Loeve transform transformation (KLT) stage and apply KLT to the frequency domain to get the benefit of the properties of KLT that minimizes the total mean squared error due to orthogonality and optimally compacts of the energy. With this algorithm we use two types of threshold and assume. There are no prior info about the noise type or the noisy frames :

1. Wave Atoms KLT dual pump

Wave Atoms KLT dual pump with this technique the threshold applied on to the magnitude part of the wave atom transformation stage and kept the phase part with no change as shown in the Figure 4.5.

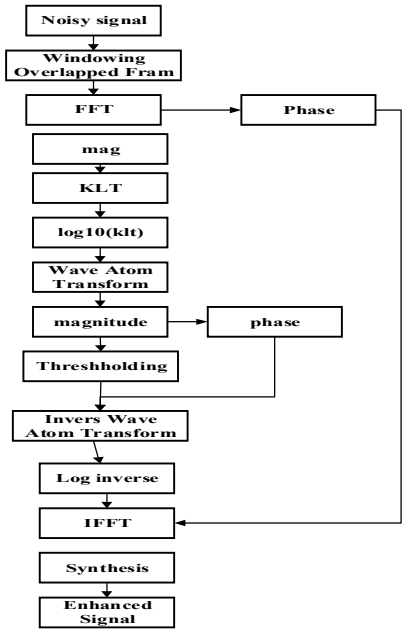


Figure 4.5 Wave Atoms KLT dual pump digram

2. Wave Atoms KLT quad pump

In in this technique we apply threshold on the absolute value of wave atoms transform with no change on the phase value, but we use quad pump function instead of two as shown in the Figure 4.6.

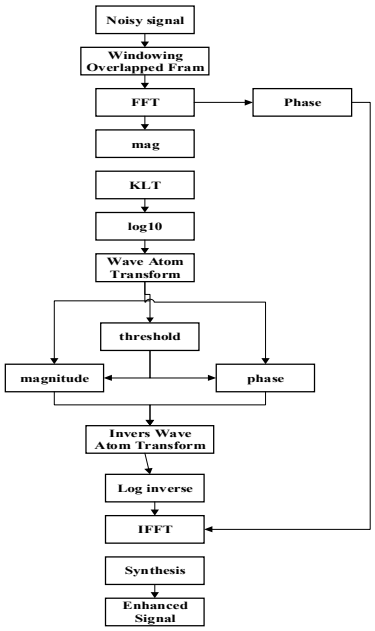


Figure 4.6 Wave Atoms KLT complex diagram

Algorithm 3 Maximum a Posteriori Wave atoms algorithm

The problem of improving the quality and intelligibility of speech in noisy environments has attracted a great deal of interest in a long time. Most speech enhancement algorithms heavily depend on the noise power spectral density (PSD).

Maximum a posteriori probability (MAP) estimate is a mode of the posterior distribution. The MAP estimator does provide a modest increase in SNR over spline interpolation for some of the lower components. The MAP estimator allows an estimation with respect to a Laplace amplitude model for the speech DFT magnitude, the joint MAP estimator also allows an optimal adjustment of the underlying statistical model to the real PDF of the speech spectral amplitude for a specific noise reduction system[20].

This technique get the benefit of maximum a posteriori algorithm, which is an important task in many applications of probabilistic. We combine the Recursive averaging, wave atoms technique described in previous section, KLT technique with the MAP technique to improve PESQ of the signal.

This technique apply the MAP on the FFT of the signal and apply threshold on the wave atoms to the output of the KLT of MAP, with this algorithm we use four type of threshold.

1. Maximum a Posteriori Wave atoms dual pump

This technique apply threshold on the absolute value of wave atoms transform with no change on the phase value as shown in the Figure 4.7

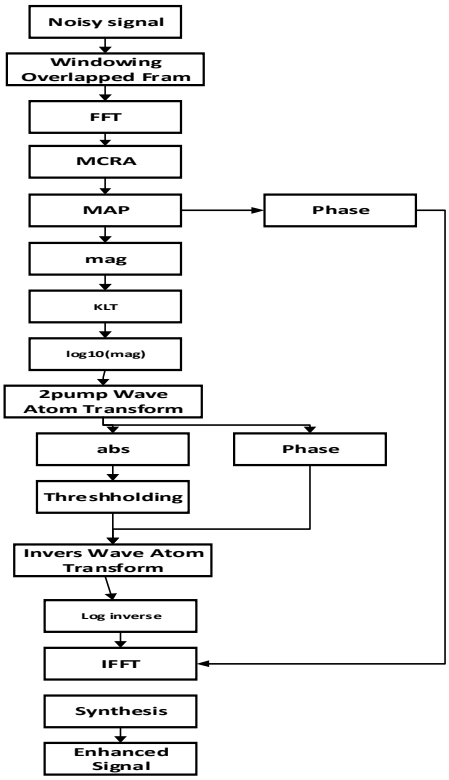


Figure 4.7 Maximum a Posteriori Wave atoms dual pump diagram

2. Maximum a Posteriori Wave atoms magnitude quad pump

This technique apply threshold on the absolute value of wave atoms transform with no change on the phase value but we use quad pump function instead of dual as shown in the Figure 4.8.

3. Maximum a Posteriori Wave atoms quad pump Cplxpair

This technique apply threshold on the complex value (real +imaginary) of wave atoms transform as shown in the Figure 4.9.

4. Maximum a Posteriori Wave atoms quad pump imaginary

In the technique we apply threshold on the imaginary part only of wave atoms transform and skipping the corresponding real value to the thresholded data. Figure 4.10 shows the algorithm flowchart.

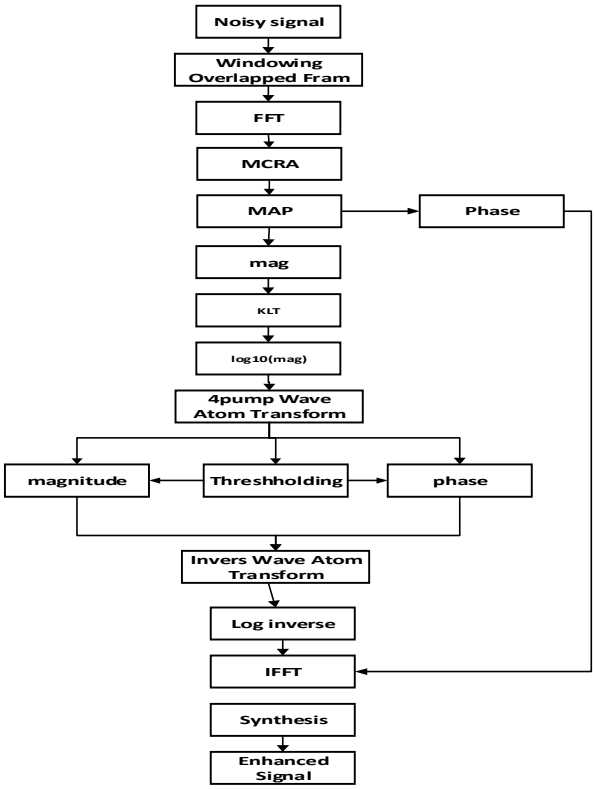


Figure 4.8 Maximum a Posteriori Wave atoms quad pump diagram

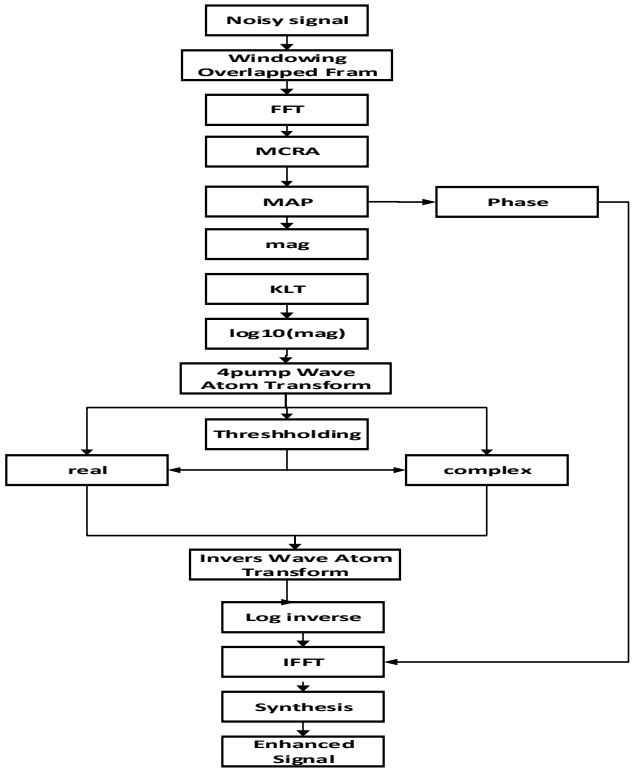


Figure 4.9 Maximum a Posteriori Wave atoms quad pump Cplxpair diagram

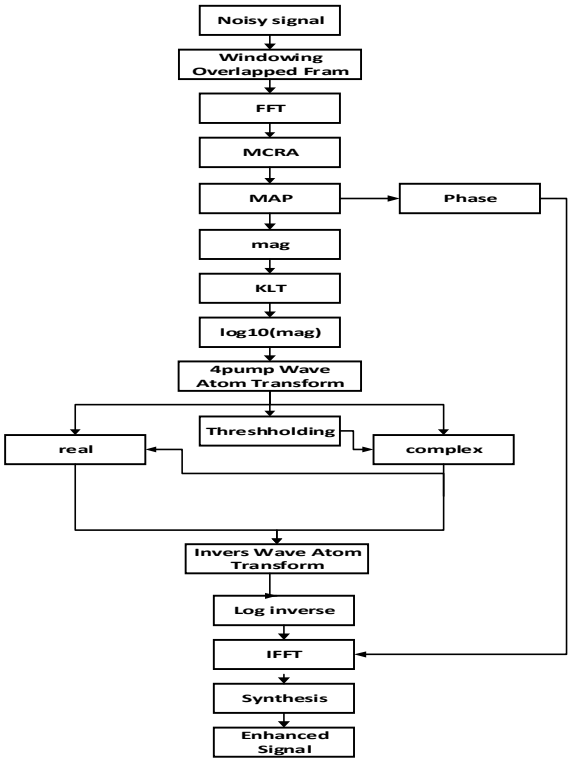


Figure 4.10Maximum a Posteriori Wave atom quad pump imaginary diagram

Algorithm 4 MMSE Wave atoms with KLT and Spectrum Power Zero Cross

In this technique we apply the MSS-MMSE-SPZC on the FFT of the signal and apply threshold on the wave atoms to the output of the KLT of MSS-MMSE-SPZC with this algorithm we use four type of threshold:

1. MMSE Wave atoms with KLT and Spectrum Power Zero Cross dual pump

This technique apply threshold on the absolute value of wave atoms transform with no change on the phase value. Figure 4.11 show the algorithm flowchart.

2. MMSE Wave Atoms with Spectrum Power Zero Cross quad pump Magnitude

With MMSE Wave atoms with Spectrum Power Zero Cross technique, we apply threshold on the absolute value of wave atoms transform with no change on the phase value but we use quad pump function instead of two. Figure 4.12 show the algorithm flowchart.

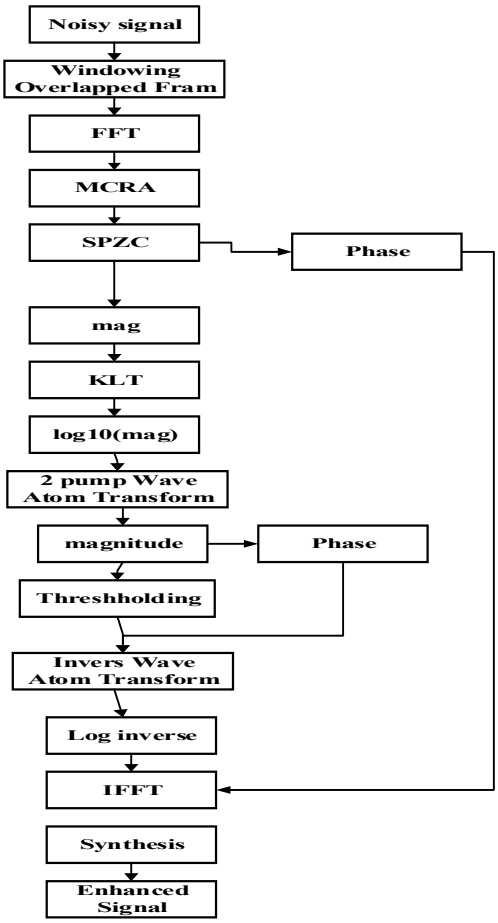


Figure 4.11 MMSE Wave atoms with SPZC dual pump diagram

3. MMSE Wave Atoms with Spectrum Power Zero Cross quad pump Cplxpair

With MMSE Wave Atoms with SPZC Cplxpair technique we apply threshold on the complex value (real +imaginary) of wave atoms transform. Figure 4.13 shows the algorithm flowchart.

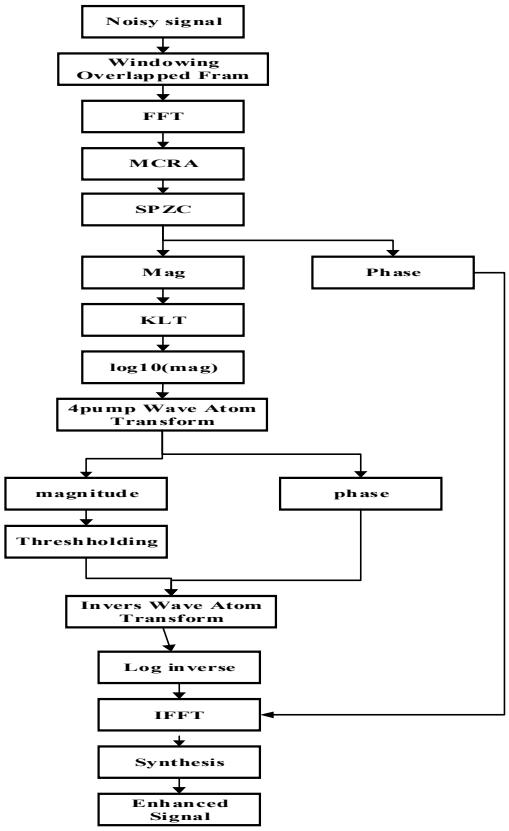


Figure 4.12 MMSE Wave Atoms with SPZC quad pump Magnitude diagram

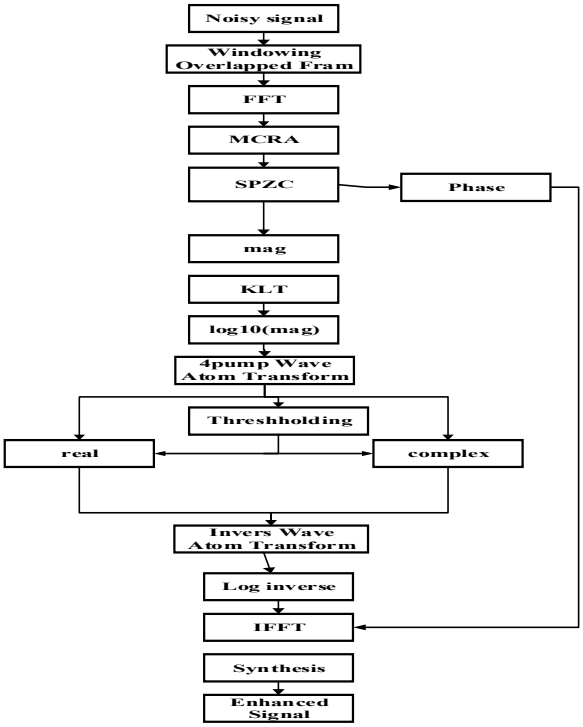


Figure 4.13 MMSE Wave Atoms with SPZC quad pump Cplxpair daigram

4. MMSE Wave Atoms with Spectrum Power Zero Cross quad pump Imaginary

With MMSE Wave Atoms with Spectrum Power Zero Cross Imaginary technique we apply threshold on the imaginary part only of wave atoms transform and skipping the corresponding real value to the thresholded data. Figure 4.14 shows the algorithm flowchart.

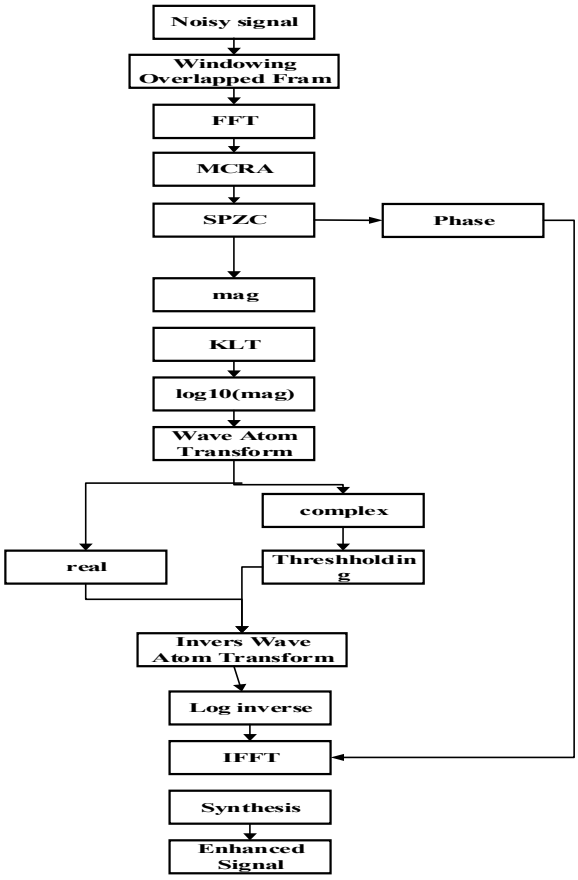


Figure 4.14 MMSE Wave Atoms with SPZC quad pump Imaginary diagram

Algorithm 5 MMSE Wave atoms with KLT and Spectrum Power Zero Cross SNR Uncertainty

In this technique we apply the MSS-MMSE-SPZC_ SNR Uncertainty on the FFT of the signal and apply threshold on the wave atoms to the output of the KLT of MSS-MMSE-SPZC_ SNR Uncertainty with this algorithm we use four type of threshold

1. MMSE Wave atoms with KLT and Spectrum Power Zero Cross SNR Uncertainty dual pump

This technique apply threshold on the absolute value of wave atoms transform with no change on the phase value. Figure 4.15 shows the algorithm flowchart.

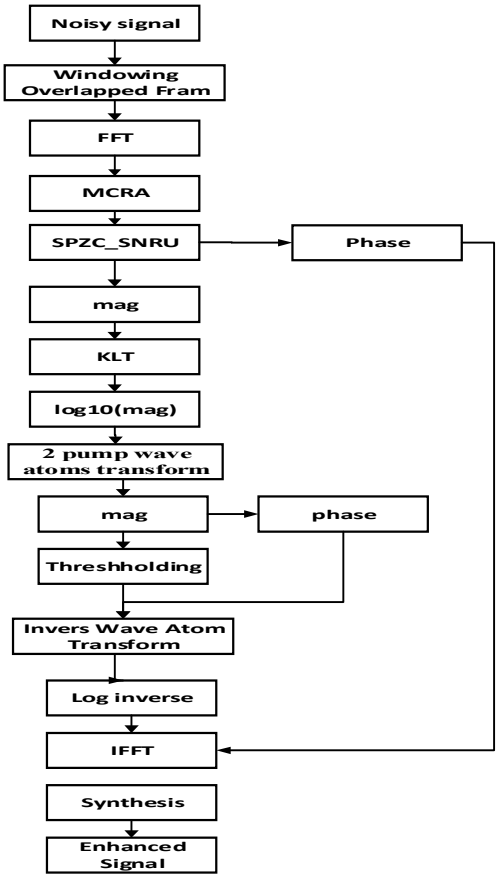


Figure 4.15 MMSE Wave atoms with KLT and Spectrum Power Zero Cross SNR Uncertainty dual pump dual pump diagram

2.MMSE Wave atoms with KLT and Spectrum Power Zero Cross SNR Uncertainty quad pump

This technique apply threshold on the absolute value of wave atoms transform with no change on the phase value but we use quad pump function instead of dual. Figure 4.16 shows the algorithm flowchart.

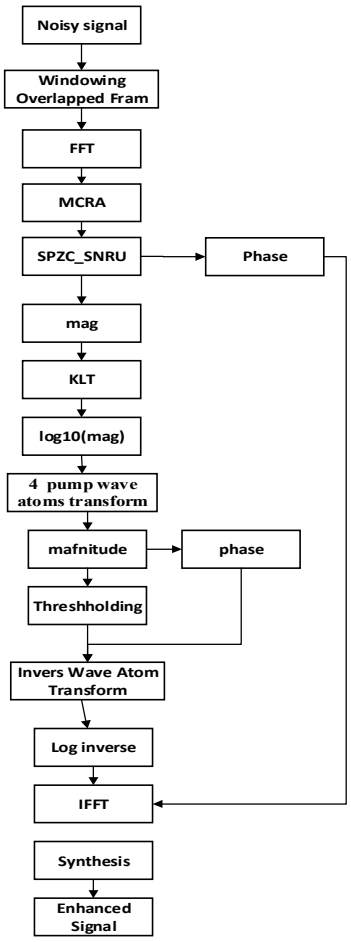


Figure 4.16 MMSE Wave atoms with KLT and Spectrum Power Zero Cross SNR Uncertainty quad pump diagram

3. MMSE Wave atoms with KLT and Spectrum Power Zero Cross SNR Uncertainty quad pump Cplxpair

This technique apply threshold on the complex value (real +imaginary) of wave atoms transform. Figure 4.17 show the algorithm flowchart.

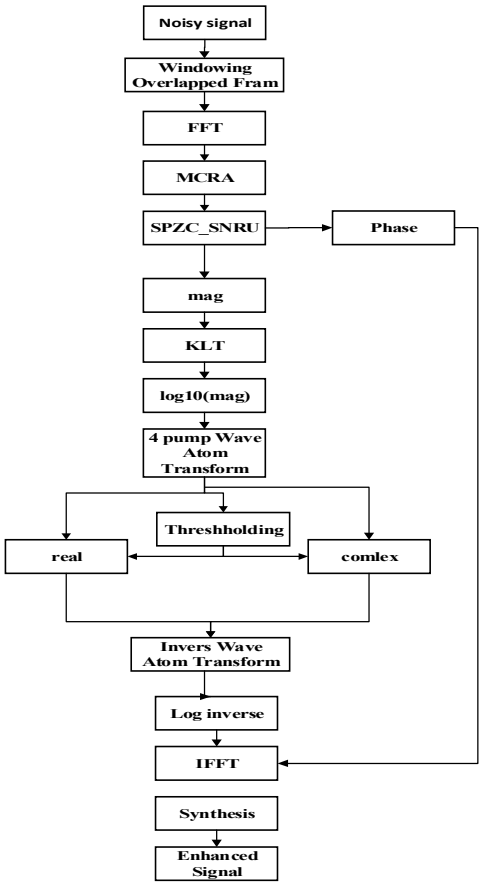


Figure 4.17 MMSE Wave atoms with KLT and Spectrum Power Zero Cross SNR Uncertainty quad pump diagram

4. MMSE Wave atoms with KLT and Spectrum Power Zero Cross SNR Uncertainty quad pump Imaginary

This technique apply threshold on the imaginary part only of wave atoms transform and skipping the corresponding real value to the thresholded data. Figure 4.18 shows the algorithm flowchart.

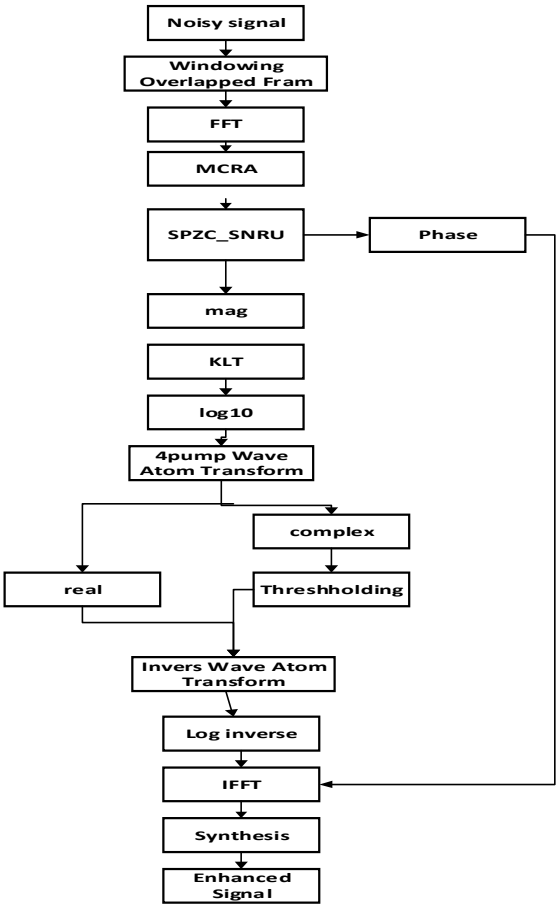


Figure 4.18 MMSE Wave atoms with KLT and Spectrum Power Zero Cross SNR Uncertainty quad pump Imaginary diagram

Algorithm 6 Wave atoms with Soft Masking using a PRiori SNR uncertainty

the sparseness of sound mixture received much attention as a priori knowledge to separate sources where sufficient information about mixtures is not given so spectral components are assumed to be statistically independent, this factor is adjusted individually as a function of the relative local A Posteriori Signal to Noise Ratio on each frequency.

In this technique we apply the MSS-MMSE-SMPR on the FFT of the signal than apply threshold on the wave atoms to the output of the KLT of SMPO with this algorithm we use four type of threshold

1. MMSE Wave atoms with KLT and SMPR SNR uncertainty dual pump

This technique apply threshold on the absolute value of wave atoms transform with no change on the phase value. Figure 4.19 show the algorithm flowchart.

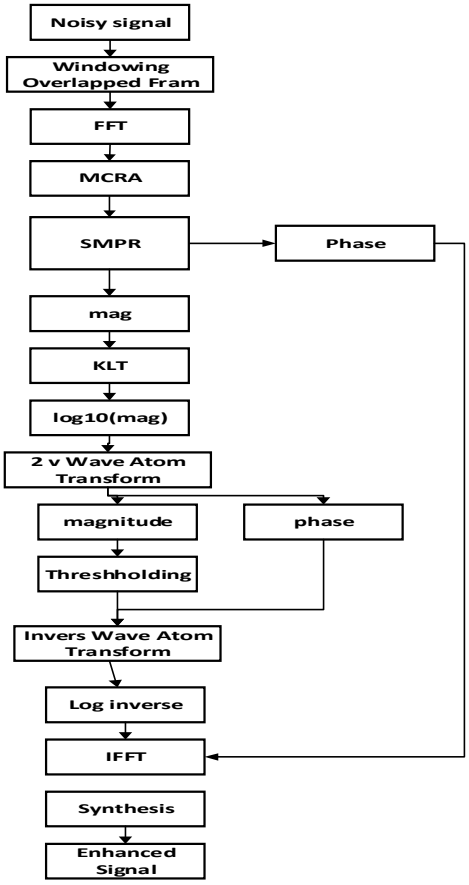


Figure 4.19 MMSE Wave atoms with KLT and SMPR SNR uncertainty dual pump diagram

2.MMSE Wave atoms with KLT and SMPR SNR uncertainty quad pump

This technique apply threshold on the absolute value of wave atoms transform with no change on the phase value but we use quad pump function instead of two. Figure 4.20 shows the algorithm flowchart.

3. MMSE Wave atoms with KLT and SMPR SNR uncertainty quad pump Cplxpair

This technique apply threshold on the complex value (real +imaginary) of wave atoms transform and use quad pump function. Figure 4.21 shows the algorithm flowchart.

4.MMSE Wave atoms with KLT and SMPR SNR uncertainty quad pump Imaginary

This technique apply threshold on the imaginary part only of wave atoms transform and skipping the corresponding real value to the thresholded data. Figure 4.22 shows the algorithm flowchart.

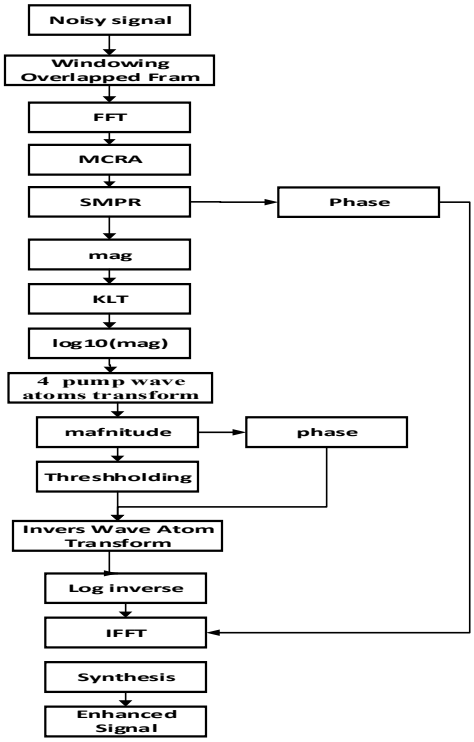


Figure 4.20 MMSE Wave atoms with KLT and SMPR SNR uncertainty quad pump diagram

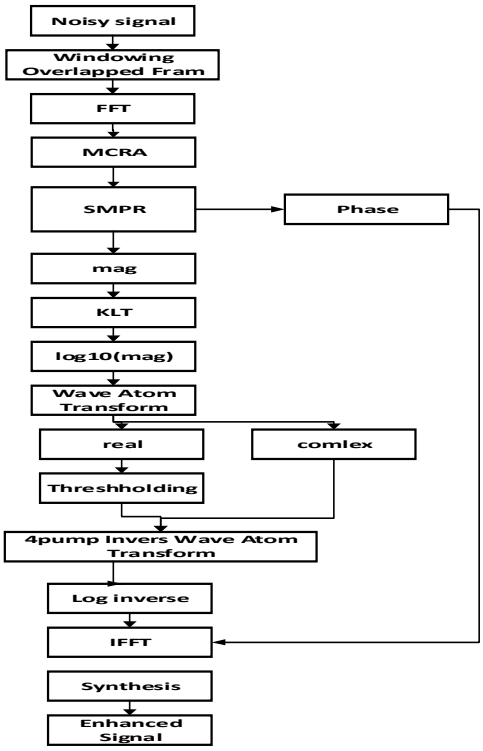


Figure 4.21 MMSE Wave atoms with KLT and SMPR SNR uncertainty quad pump Imaginary diagram

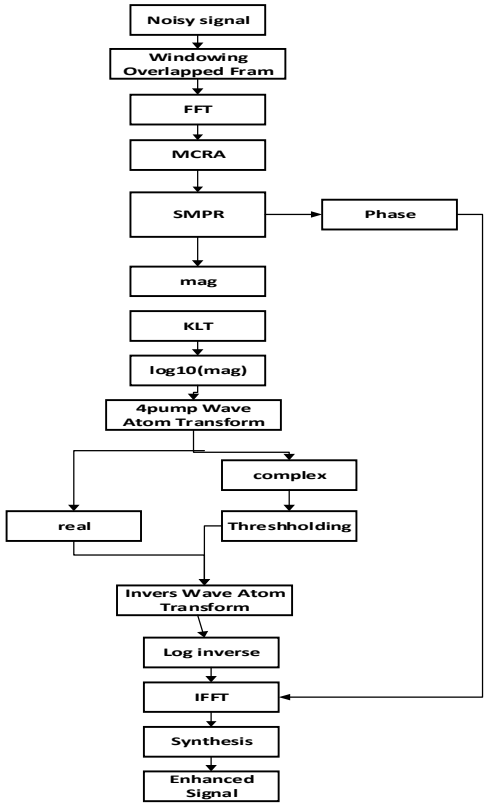


Figure 4.22 MMSE Wave atoms with KLT and SMPR SNR uncertainty quad pump Imaginary diagram

Algorithm 7 Wave atoms with Soft Masking Based on Posteriori SNR Uncertainty

This technique apply threshold on the wave atoms to the output of the KLT of MSS-MMSE-SMPO. With this algorithm we use four type of threshold

1. MMSE Wave atoms with KLT and SMPO SNR uncertainty dual pump

This technique apply threshold on the absolute value of wave atoms transform with no change on the phase value. Figure 4.23 shows the algorithm flowchart.

2.MMSE Wave atoms with KLT and SMPO SNR uncertainty quad pump

This technique apply threshold on the absolute value of wave atoms transform with no change on the phase value but we use quad pump function instead of dual. Figure 4.24 shows the algorithm flowchart.

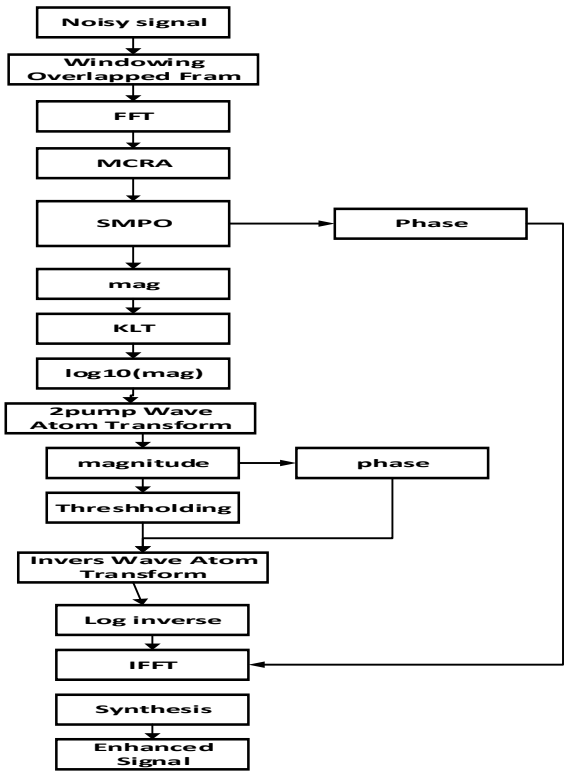


Figure 4.23 MMSE Wave atoms with KLT and SMPO SNR uncertainty dual pump diagram

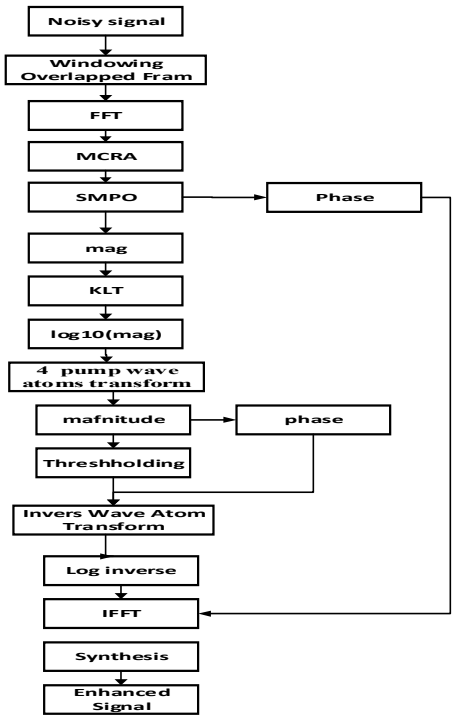


Figure 4.24 MMSE Wave atoms with KLT and SMPO SNR uncertainty quad pump diagram

3. MMSE Wave atoms with KLT and SMPO SNR uncertainty quad pump Cplxpair

This technique apply threshold on the complex value (real +imaginary) of wave atoms transform and use quad pump function. Figure 4.25 shows the algorithm flowchart.

4. MMSE Wave atoms with KLT and SMPO SNR uncertainty quad pump Imaginary

This technique apply threshold on the complex value only of wave atoms transform and corresponding real value to the thresholded value. Figure 4.26 shows the algorithm flowchart.

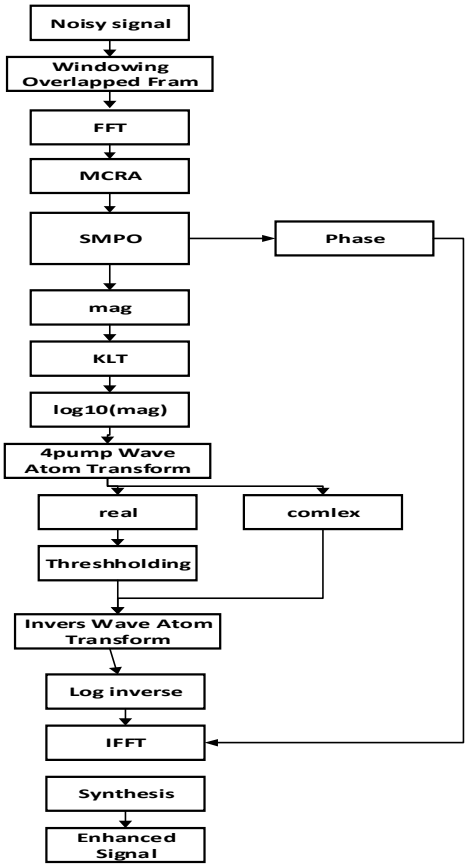


Figure 4.25 MMSE Wave atoms with KLT and SMPO SNR uncertainty quad pump Cplpair diagram

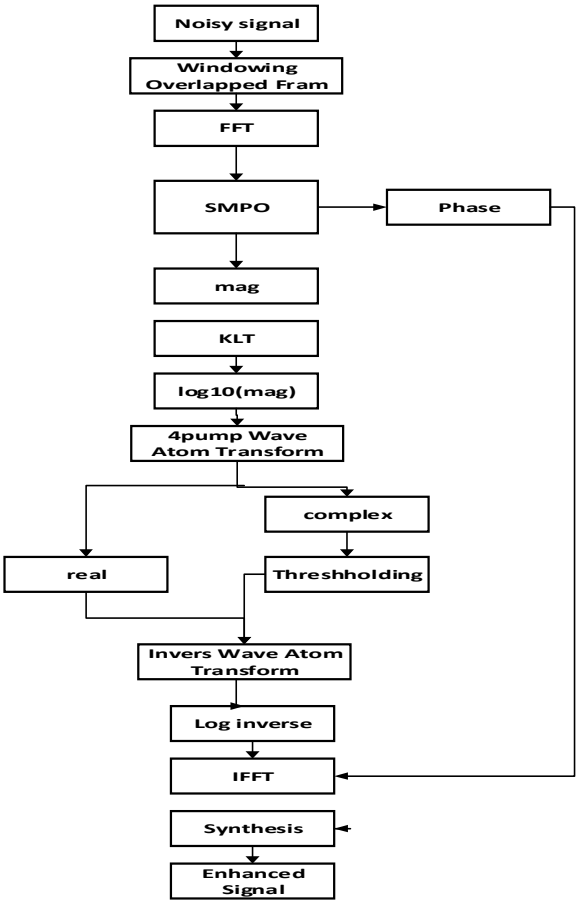


Figure 4.26 MMSE Wave atoms with KLT and SMPO SNR uncertainty quad pump Imaginary diagram

Chapter 5

5 Experimental Results

Description of the datasets used in experiments and the measurement techniques in addition to measuring the accuracy of the proposed algorithms' results to ensure their ability in delivering better results than other algorithms

5.1 Dataset specifications

This section describes and identifies the specifications of datasets used in the all experiments on the proposed algorithms. In this study researcher use NOIZEUS database. NOIZEUS is a noisy speech corpus developed lab to facilitate comparison of speech enhancement algorithms among research groups[37]. The noisy database contains 30 IEEE sentences (produced by three male and three female speakers) corrupted by eight different real-world noises at different SNRs. The noise was taken from the AURORA [38] database and includes suburban train noise, babble, car, exhibition hall, restaurant, street, airport and train-station noise. This corpus is available to researchers free of charge. Table 5.1 Show all sentence used and there gender. Figure 5.1 shows the broads of phonetic class distribution of the NOIZEUS DB.

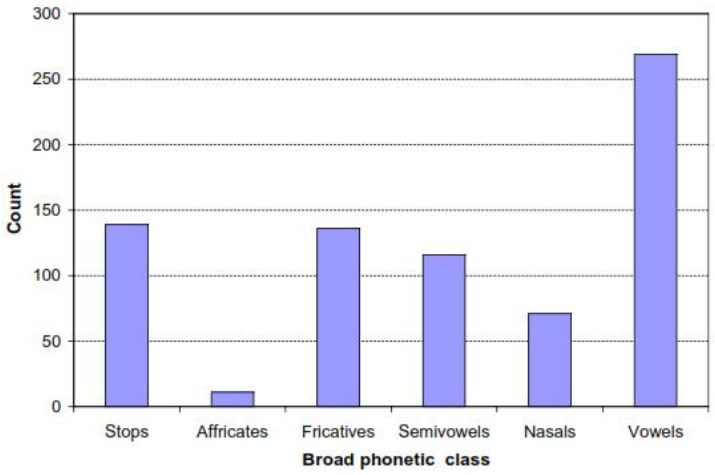


Figure 5.1broad of phonetic class distribution of the NOIZEUS DB

Table 5.1 NOIZEUS data base speaker sentence and gender

File	Gender	Speaker	Sentence
sp01.wav	M	CH	The birch canoe slid on the smooth planks
sp02.wav	M	CH	He knew the skill of the great young actress
sp03.wav	M	CH	Her purse was full of useless trash
sp04.wav	M	CH	Read verse out loud for pleasure
sp05.wav	M	CH	Wipe the grease off his dirty face
sp06.wav	M	DE	Men strive but seldom get rich
sp07.wav	M	DE	We find joy in the simplest things
sp08.wav	M	DE	Hedge apples may stain your hands green
sp09.wav	M	DE	Hurdle the pit with the aid of a long pole
Sp10.wav	M	DE	The sky that morning was clear and bright blue
Sp11.wav	F	JE	He wrote down a long list of items
Sp12.wav	F	JE	The drip of the rain made a pleasant sound
Sp13.wav	F	JE	Smoke poured out of every crack
Sp14.wav	F	JE	Hats are worn to tea and not to dinner
Sp15.wav	F	JE	The clothes dried on a thin wooden rack
Sp16.wav	F	KI	The stray cat gave birth to kittens
Sp17.wav	F	KI	The lazy cow lay in the cool grass
Sp18.wav	F	KI	The friendly gang left the drug store
Sp19.wav	F	KI	We talked of the sideshow in the circus
Sp20.wav	F	KI	The set of china hit the floor with a crash
Sp21.wav	M	SI	Clams are small, round, soft and tasty
Sp22.wav	M	SI	The line where the edges join was clean
Sp23.wav	M	SI	Stop whistling and watch the boys march
Sp24.wav	M	SI	A cruise in warm waters in a sleek yacht is fun
Sp25.wav	M	SI	A good book informs of what we ought to know
Sp26.wav	F	TI	She has a smart way of wearing clothes
Sp27.wav	F	TI	Bring your best compass to the third class
Sp28.wav	F	TI	The club rented the rink for the fifth night
Sp29.wav	F	TI	The flint sputtered and lit a pine torch
Sp30.wav	F	TI	Let us all join as we sing the last chorus

5.2 Performance evaluation

The aim of the speech enhancement algorithms is to increase the ease of listening and if possible, increase the amount of received information. These two concepts are known as ‘quality’ and ‘intelligibility’ respectively. Since the aim of enhancement algorithms is to improve these two attributes so there are many Performance measures are defined for the evaluation of

speech quality like segmental SNR, weighted spectral slope (WSS), Bark distortion measures, and perceptual evaluation of speech quality. In some issue most of these methods clash with each other. In this research, PESQ measure will be used to evaluate the performance of the proposed algorithm, as it is the ITU standard for automatic assessment of speech quality which used by phone manufacturers and telecom operators. In this research we used PESQ measure and segmental SNR as time domain measurement.

5.2.1 Time-Domain Measurement

The simplest way to perform a time-domain measurement consists of calculating the Signal-to-Noise Ratio (SNR) that performs a sample-by-sample comparison between original and processed speech signals. Speech waveforms are compared in the time domain. Therefore, the synchronization of the original and distorted speech is crucial. The most popular and accurate time-domain measure is the segmental signal-to-noise ratio (Seg_SNR). This measure is particularly effective in indicating the speech distortion than the overall SNR [41]. The frame-based segmental SNR is formed by averaging frame level SNR estimates. Higher values of the Seg_SNR indicates weaker speech distortions.

5.2.2 Perceptual Evaluation of Speech Quality (PESQ)

Previous objective speech quality assessment models, such as bark spectral distortion (BSD), the perceptual speech quality measure (PSQM), and measuring normalizing blocks (MNB), have been found to be suitable for assessing only a limited range of distortions. The traditional method of determining voice quality is to conduct subjective tests with panels of human listeners. Extensive guidelines are given in ITU-T recommendations P.800/P.830. The results of these tests are averaged to give mean opinion scores (MOS) but such tests are expensive and are impractical for testing in the field.[39]

PESQ described in ITU-T Rec. P.862 Perceptual Evaluation of Speech Quality. PESQ measures one-way voice quality: a signal is injected into the system under test, and the degraded output is compared by PESQ with the input (reference) signal utilizing sophisticated mechanisms to match the performance of most reliable subjective test like MOS.

The method of PESQ is intrusive. It compares the reference signal with a degraded signal, which is a result of passing source signal through communication system. The output of PESQ is a prediction of perceived quality that would be given to the degraded signal by subjects in a subjective listening test. In the first step of PESQ, a series of delays between original input and degraded output are computed, one for each time interval in which delay is significantly different from the previous time interval. Based on the set of delays that are found, PESQ compares input (original) signal with the aligned degraded output using perceptual model, as shown in Figure 5.2. The key step of the algorithm is transformation of original and degraded signal to an internal representation that represents the audio signals in the human auditory system, considering perceptual frequency and loudness. This is achieved in several stages: level alignment to calibrated listening level, time-frequency mapping, frequency warping, and compressive loudness scaling. The alignment process consist of applying narrowband filter to both signals to emphasis perceptually important parts, envelope-based delay estimation, division of reference signal into utterances, envelope-based delay estimation for each utterance, fine correlation histogram-based delay identification for each utterance, utterance splitting and re-alignment to test for delay changes during speech.

In the next step, internal representation is processed to account for effects such as linear filtering and local gain variations. The difference in the internal representation is computed to measure the audible difference. With the cognitive model two error parameters (symmetric disturbance and asymmetric disturbance) are computed[40].

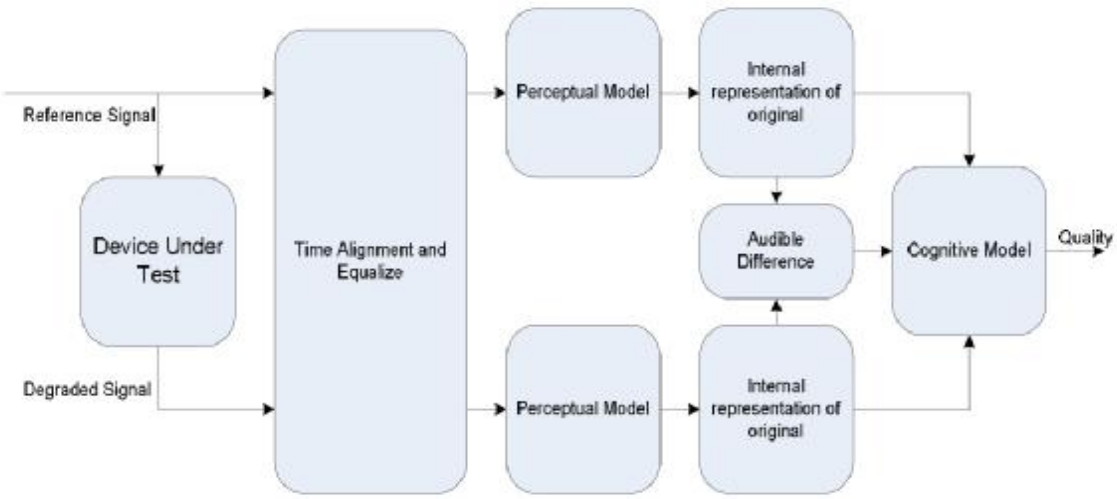


Figure 5.2 PESQ algorithm diagram[40]

The range of PESQ is: 0.5 - 4.5. Higher values indicate higher resemblance of the loudness spectra of clean speech and degraded speech PESQ measures perceived quality, it does not take into account the volume difference between the reference speech and the degraded speech because the first step of the PESQ algorithm is to compensate for the overall gain of the system under test on both speech samples before it starts to do the comparison between them.

The computation of the percent of enhancement was computed using the flowing equation

Enhancement %= (enhanced signal PESQ – input signal PESQ)/ input signal PESQ *100

5.3 Work environment

This work use hp probook work station with 8 GB RAM and Intel i7 3262QM CPU. Also, use SQL Server 2012 to store the experimental result. All algorithm implemented with Matlab 2013a(8.1.0.604) 64bit. The time consuming of this work has average 1.361 second with input signal 2.6 second duration

5.4 Proposed algorithm Result

We have evaluated the performance of the proposed noise estimation algorithm in this section. PESQ is employed as an objective measure for speech quality.

Algorithm 1. Pure Wave Atoms

1. Wave Atoms Dual Pump

Figure 5.4 shows the spectrogram of the input signal degraded with 10 dB babble noise after applying the algorithm. Figure 5.3 shows the original clear signal spectrogram .Note that there is no big difference between Figure 5.3 and Figure 5.4 because of noise which imply small enhancement. Table 5.2 shows the PESQ and Segmental Signal to Noise Ratio results. We can see from the table the small amount of enhancement achieved with this technique also we can notice the small enhancement from the spectrogram of both input and output figures (Figure 5.3 and Figure 5.4).

Table 5.2 The result of PESQ and SNR of Wave Atoms Dual Pump

	0dB		**5dB**		**10dB**		**15dB**	
	PESQ	Seg_SNR	PESQ	Seg_SNR	PESQ	Seg_SNR	PESQ	Seg_SNR
airport	1.437	-3.807	1.943	-1.083	2.245	1.958	2.617	4.863
Babble	1.487	1.494	1.763	1.725	2.296	2.266	2.696	2.678
Car	1.464	-4.529	1.831	-1.725	2.043	1.253	2.533	4.143

Street	1.371	1.356	1.729	1.733	2.04	2.051	2.537	2.49
train	1.575	-4.148	1.826	-1.162	2.075	1.646	2.507	4.339
restaurant	1.821	1.869	1.92	1.935	2.224	2.181	2.638	2.657

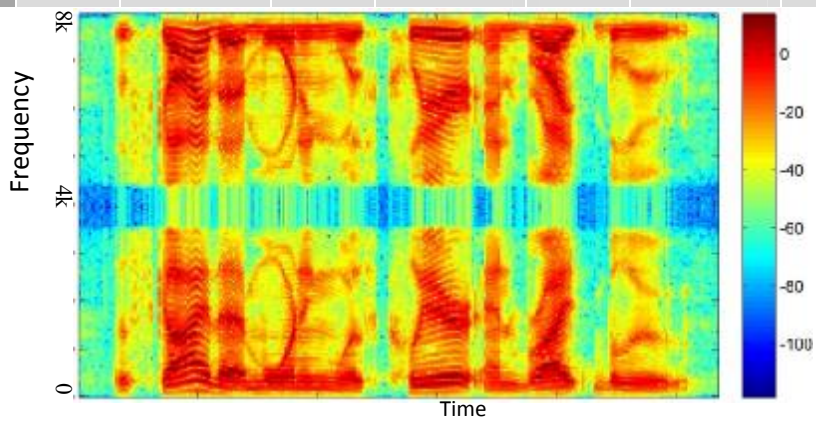


Figure 5.3 spectrogram of clear signal

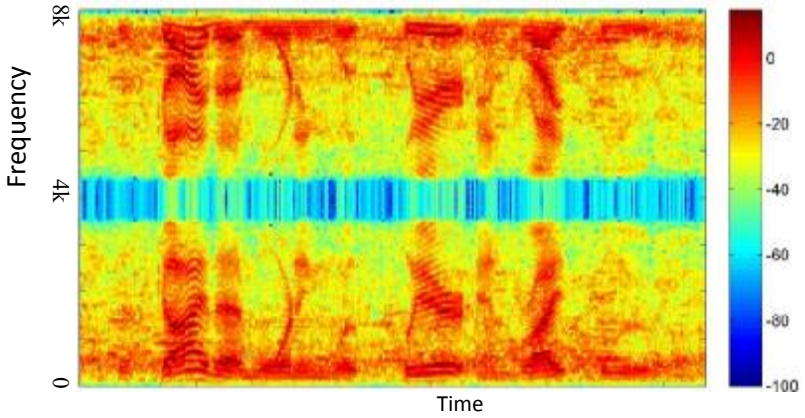


Figure 5.4 spectrogram of input signal degraded with 10dB babble noise

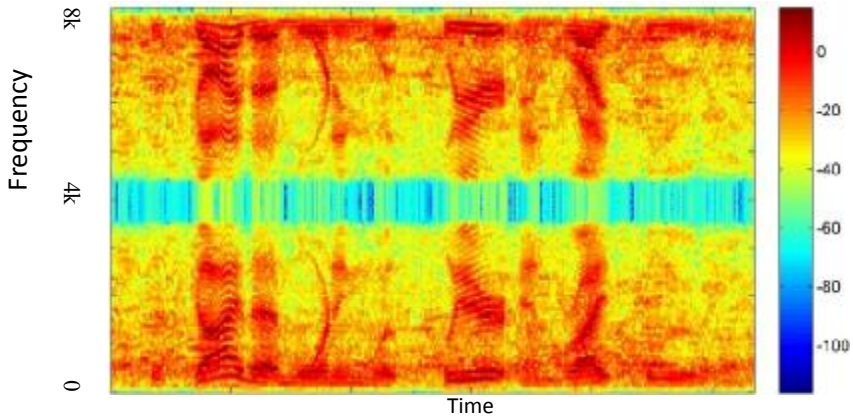


Figure 5.5 Spectrogram of Wave Atoms Dual Pump

2. Wave Atoms Quad Pump Magnitude

Figure 5.6 Spectrogram of Wave Atoms quad pump shows the spectrogram of the input signal degraded with 10 dB babble noise after applying Wave Atoms complex algorithm. By comparing Figure 5.6 and Figure 5.4, we can see there is a small enhancement achieved with this technique. Table shows the result of PESQ and Seg_SNR of the Wave Atoms quad algorithm. We can see that we have achieved quadratic improvement from the previous algorithm.

Table 5.3 PESQ and Seg_SNR of Wave Atoms quad pump magnitude

	0dB		5dB		10dB		15dB	
	PESQ	Seg_SNR	PESQ	Seg_SNR	PESQ	Seg_SNR	PESQ	Seg_SNR
airport	1.747	-3.816	2.051	-1.075	2.377	2.067	2.666	4.904
Babble	1.728	1.731	2.046	2.05	2.371	2.376	2.64	2.701
Car	1.658	-4.351	1.928	-1.416	2.252	1.734	2.515	3.854
Street	1.647	1.663	1.949	1.953	2.278	2.286	2.574	2.579
train	1.62	-3.767	1.888	-0.891	2.204	2.06	2.528	5.14
restaurant	1.777	1.775	2.033	2.037	2.4	2.408	2.685	2.691

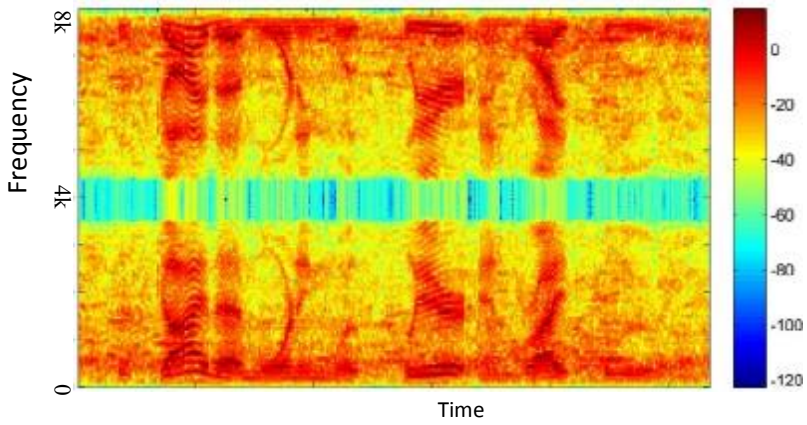


Figure 5.6 Spectrogram of Wave Atoms quad pump

3.Wave Atoms quad pump Cplxpair

Figure 5.7 shows the spectrogram of the input signal degraded with 10 dB babble noise after applying the algorithm, we can see the signal is still evident degraded. From the Table 5.4 we have reached improvement from the first technique Wave Atoms Dual Pump but less than improvement from second one Wave Atoms Quad Pump Magnitude.

Table 5.4 PESQ and Seg_SNR Wave Atoms quad pump Cplxpair

	0dB		5dB		10dB		15dB	
	PESQ	Seg_SNR	PESQ	Seg_SNR	PESQ	Seg_SNR	PESQ	Seg_SNR
airport	1.738	-3.665	2.034	-1.261	2.332	1.41	2.651	3.839
Babble	1.851	1.91	2.082	2.087	2.221	2.214	2.684	2.679
Car	1.625	-4.178	2.08	-0.638	2.194	1.443	2.518	4.077
Street	1.815	1.724	1.952	2.02	2.213	2.234	2.466	2.334
train	1.749	-3.347	1.943	-0.891	2.167	1.686	2.501	4.44
restaurant	1.777	1.748	2.033	2.139	2.4	2.291	2.739	2.66

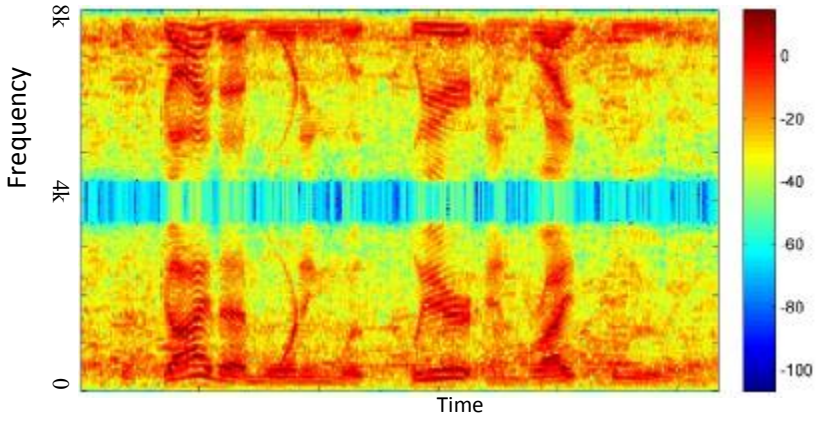


Figure 5.7 Wave Atoms quad pump Cplxpair Spectrogram

4.Wave Atoms Cplxpair Imaginary

Table 5.5 shows the result of PESQ and Seg_SNR measurement of this technique .From the tables, we have achieved quadratic improvement from the previous algorithm Wave Atoms Cplxpair. The result of Table 5.5 shows that there is a huge importance of using the complex part of wave atoms transformation in enhancement process using wave atoms transform with any prior information or probability estimators techniques. We can see from Table 5.2, Table 5.3, Table 5.4 and Table 5.5 the noise imaginary part is dominating, and using imaginary part is important for speech enhancement. Figure 5.8 shows the spectrogram of the input signal degraded with 10 dB babble noise after applying the algorithm we can see the signal enhancement still slight but better than previous technique (Wave Atoms directional, Wave

Atoms complex and Wave Atoms Cplxpair) . The slight enhancement is due to no prior estimation about noise.

Table 5.5 PESQ and Seg_SNR Wave Atoms quad pump Imaginary

	0dB		5dB		10dB		15dB	
	PESQ	Seg_SNR	PESQ	Seg_SNR	PESQ	Seg_SNR	PESQ	Seg_SNR
airport	1.755	-4.101	2.058	-1.359	2.424	2.027	2.358	5.1
Babble	1.601	-4.469	2.15	-1.481	2.376	1.762	2.703	4.988
Car	1.684	1.487	1.932	1.946	2.257	2.26	2.578	2.58
Street	1.633	-3.864	1.574	-0.964	2.292	2.192	2.586	5.452
train	1.632	-4.111	1.892	-1.119	2.274	2.295	2.483	5.277
restaurant	1.675	-3.938	2.143	-1.091	2.412	2.146	2.702	5.301

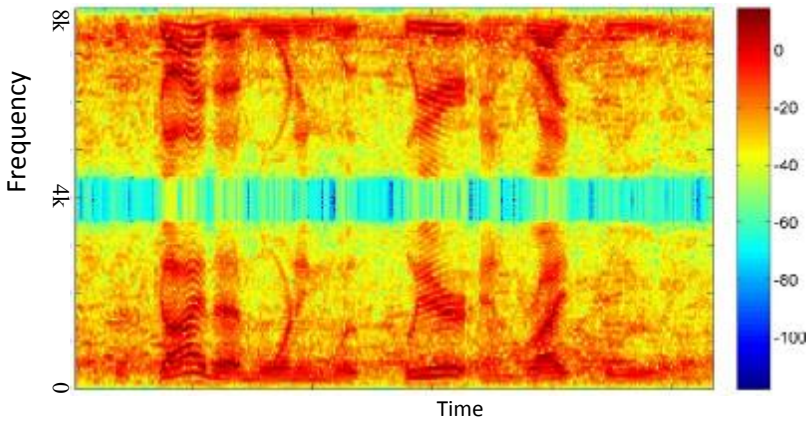


Figure 5.8 Wave Atoms quad pump Cplxpair Imaginary Spectrogram

Algorithm 2. Wave atoms with KLT

1.Wave Atoms KLT dual pump

Figure 5.9 shows the spectrogram of the input signal degraded with 10 dB babble noise after applying the algorithm. Spectrogram show the original full clear signal spectrogram . Comparing the spectrogram of Wave Atoms KLT dual pump Figure 5.9 with the spectrogram of input degraded signal Figure 5.4 , there is big difference between them ; the noise decreases and

the spectrogram appears more clear than upper algorithm Wave Atoms dual pump, Wave Atoms quad pump , Wave Atoms quad pump Cplxpair and Wave Atoms quad pump Imaginary. From Table 5.6 we can see the difference between this technique and Wave Atoms Dual Pump is doubled.

Table 5.6 Wave Atoms KLT dual pump PESQ and Seg_SNR result

	0dB		5dB		10dB		15dB	
	PESQ	Seg_SNR	PESQ	Seg_SNR	PESQ	Seg_SNR	PESQ	Seg_SNR
Airport	1.544	-3.596	1.949	-0.594	2.408	1.978	2.705	3.29
Babble	1.62	1.541	2.023	2.006	2.388	2.382	2.743	2.778
Car	1.629	1.61	1.823	1.821	2.309	2.294	2.611	2.626
Street	1.579	1.544	1.9	1.866	2.207	2.173	2.635	2.663
Train	1.431	-3.086	1.5	-0.416	2.062	1.829	2.595	3.459
Restaurant	1.58	-2.823	1.902	-0.227	2.368	1.97	2.732	3.068

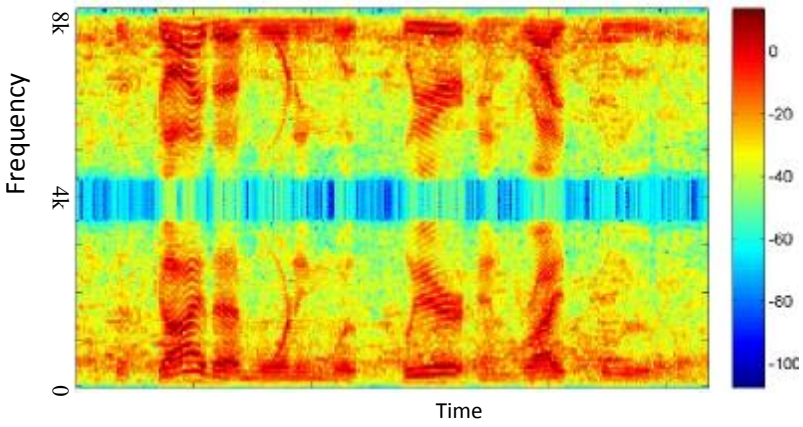


Figure 5.9 Spectrogram of Wave Atoms KLT dual pump

2.Wave Atoms KLT quad pump

Figure 5.10 shows the spectrogram of the input signal degraded with 10 dB babble noise after applying the algorithm comparing the spectrogram of Wave Atoms KLT Quad pump Figure 5.10 with the spectrogram of input degraded signal Figure 5.4. There is clear difference between them; the noise decreases and the spectrogram appears more clear than upper algorithms Wave Atoms quad pump, Wave Atoms quad pump Cplxpair , Wave Atoms Cplxpair Imaginary and

Wave Atoms KLT dual pump). From Table 5.7 we can see the difference between this technique and Wave Atoms dual pump is doubled

Table 5.7 Wave Atoms KLT quad pump PESQ and Seg_SNR result

	0dB		5dB		10dB		15dB	
	PESQ	Seg_SNR	PESQ	Seg_SNR	PESQ	Seg_SNR	PESQ	Seg_SNR
airport	1.758	-3.214	2.081	-0.539	2.409	1.896	2.712	3.447
Babble	1.747	-3.289	2.071	-0.663	2.402	1.86	2.737	3.315
Car	1.694	-3.543	1.951	-0.826	2.242	1.713	2.618	3.644
Street	1.615	-2.886	1.976	-0.321	2.26	2.05	2.624	2.895
train	1.617	-3.073	1.91	-0.468	2.236	1.856	2.584	3.486
restaurant	1.754	-2.976	2.052	-0.387	2.43	2.035	2.728	3.192

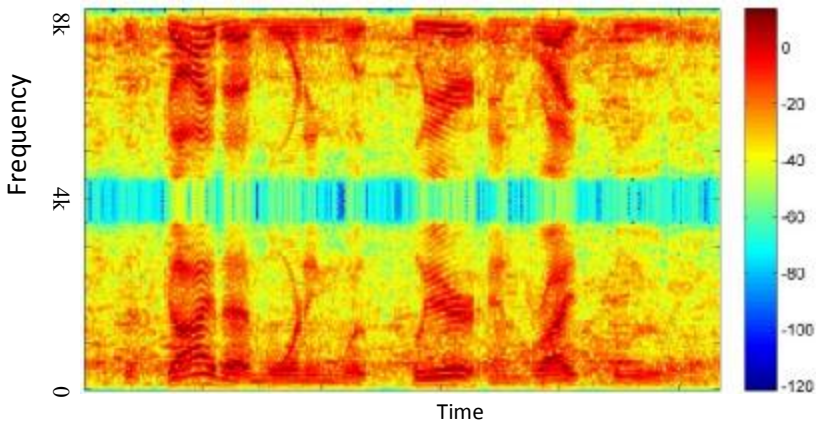
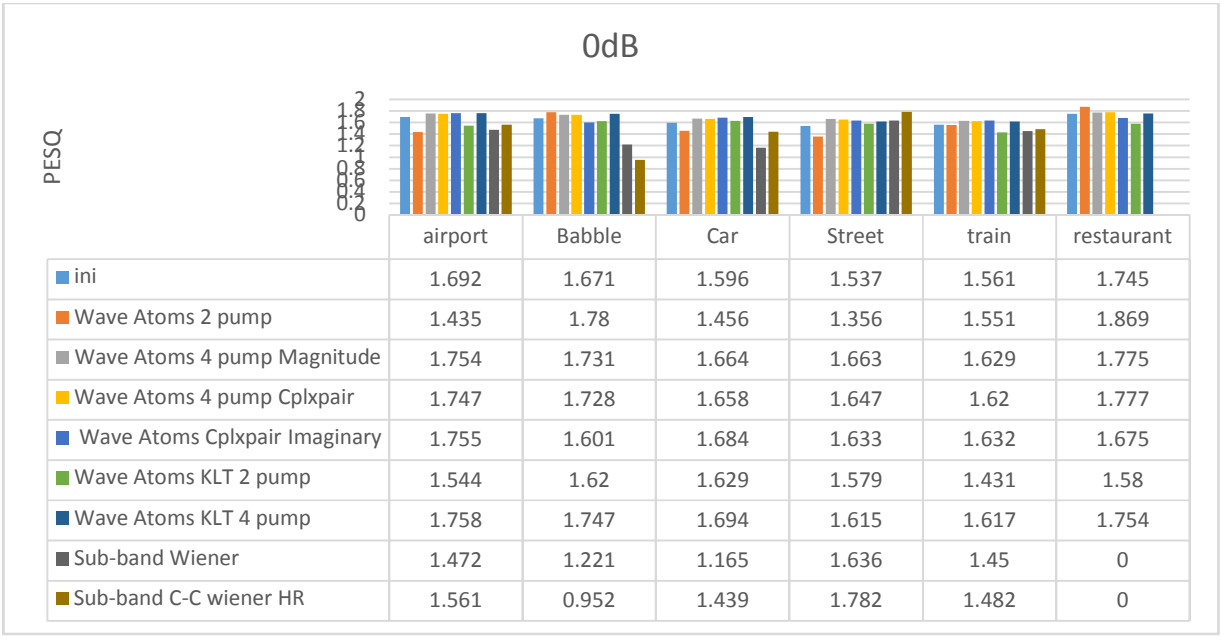


Figure 5.10 Wave Atoms KLT quad pump Spectrogram

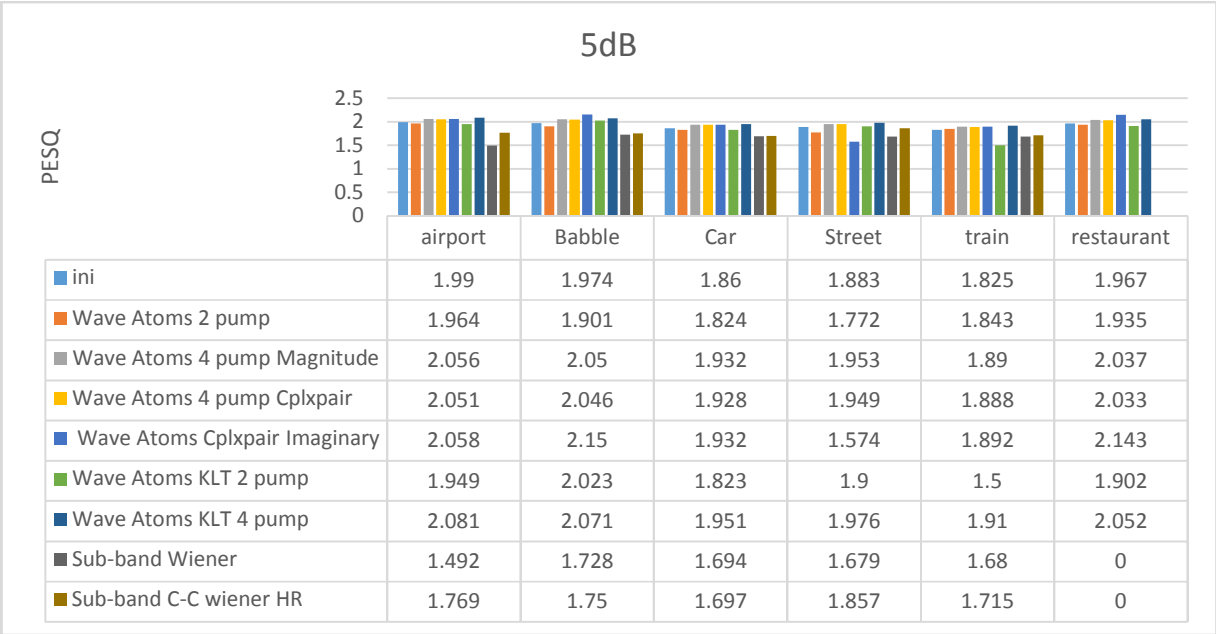
Figure 5.11, Figure 5.12, Figure 5.13 and Figure 5.14 show the comparison between our technique and similar techniques using the same dataset. The tables are collected from the related search papers [12] , [14]. Wave Atoms dual pump in average was failed to enhance the signal in all types of noise and dB, but get better result than sub_band wiener and sub_band cross correlation with harmonic regeneration technique. Also, we can see the effect of KLT in enhancement process that makes Wave Atoms KLT quad pump dominating other proposed technique. We can perceive the effect of using complex number and imaginary part of the complex number in enhancement process.

Figure 5.11, Figure 5.12, Figure 5.13 and Figure 5.14 demonstrate that Wave Atoms dual pump does not make enhancement with lower dB in all noise types, and the enhancement slightly increased with 15dB. Wave Atoms quad pump make more enhancement with lower dB and slightly decreased with higher dB. Comparing result of Wave Atoms dual pump with Wave Atoms quad pump we have about 3.5% enhancement with Wave Atoms quad pump. The result of Wave Atoms quad pump Cplxpair makes more enhancement with lower dB and slightly decreases with higher dB, but this result is less than the result of Wave Atoms quad pump with 0.5% (small deference). The result of Wave Atoms Cplxpair Imaginary makes more enhancement with lower dB and slightly decreased with higher dB but this result is less than result of Wave Atoms 4 pump with 1%. From the previous discussion we conclude that the best enhancement when use wave atoms only is achieved with Wave Atoms quad pump technique. In Wave Atoms KLT dual pump does not make enhancement with lower dB in all noise types and the enhancement slightly increases with 10dB and 15dB. Wave Atoms KLT Quad pump has 5% enhancement more than Wave Atoms KLT dual pump and has 1% enhancement more than Wave Atoms quad pump. From the previous discussion we conclude that the maximum enhancement is reached when use Wave Atoms KLT quad pump technique.



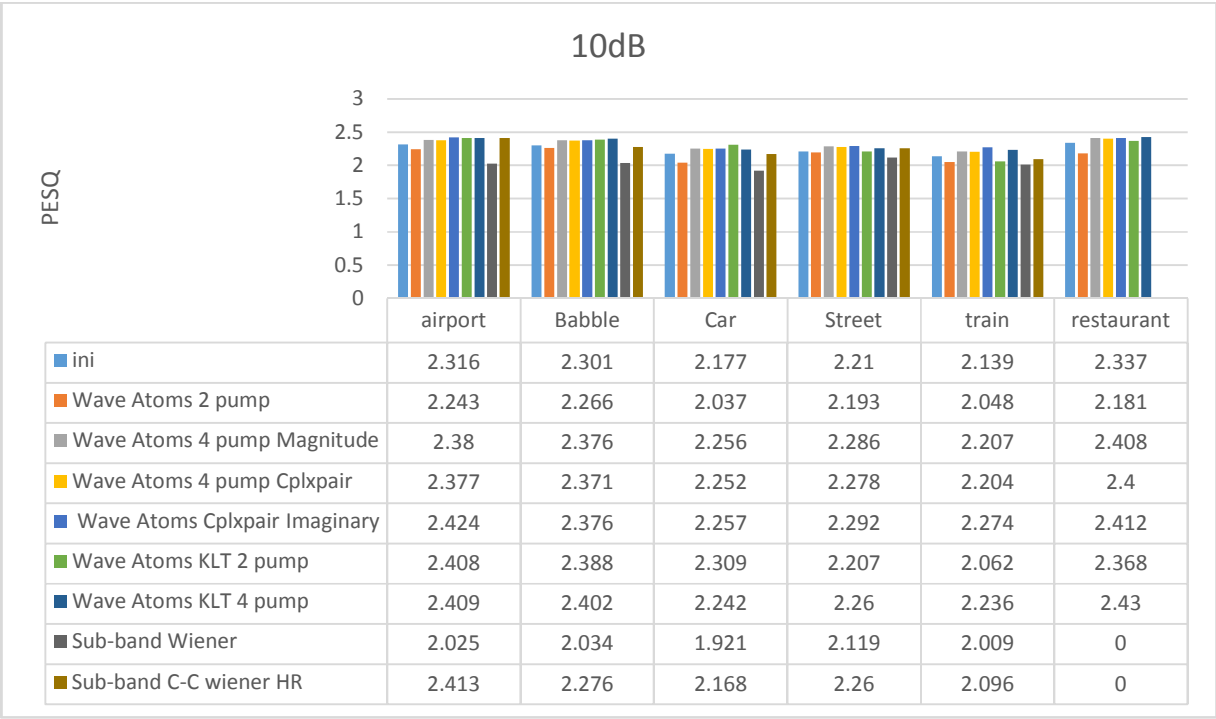
	airport	Babble	Car	Street	train	restaurant
ini	1.692	1.671	1.596	1.537	1.561	1.745
Wave Atoms 2 pump	1.435	1.78	1.456	1.356	1.551	1.869
Wave Atoms 4 pump Magnitude	1.754	1.731	1.664	1.663	1.629	1.775
Wave Atoms 4 pump Cplxpair	1.747	1.728	1.658	1.647	1.62	1.777
Wave Atoms Cplxpair Imaginary	1.755	1.601	1.684	1.633	1.632	1.675
Wave Atoms KLT 2 pump	1.544	1.62	1.629	1.579	1.431	1.58
Wave Atoms KLT 4 pump	1.758	1.747	1.694	1.615	1.617	1.754
Sub-band Wiener	1.472	1.221	1.165	1.636	1.45	0
Sub-band C-C wiener HR	1.561	0.952	1.439	1.782	1.482	0

Figure 5.11comparison of different algorithm corrupted with 0dB Noise and different noise type



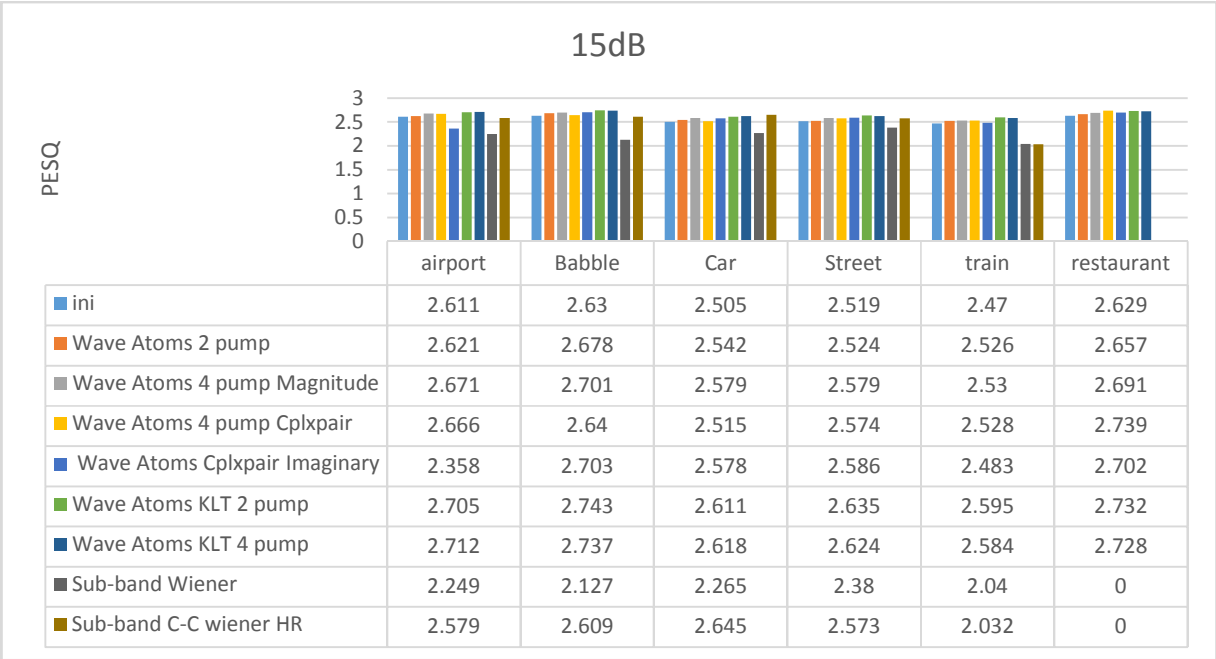
	airport	Babble	Car	Street	train	restaurant
ini	1.99	1.974	1.86	1.883	1.825	1.967
Wave Atoms 2 pump	1.964	1.901	1.824	1.772	1.843	1.935
Wave Atoms 4 pump Magnitude	2.056	2.05	1.932	1.953	1.89	2.037
Wave Atoms 4 pump Cplxpair	2.051	2.046	1.928	1.949	1.888	2.033
Wave Atoms Cplxpair Imaginary	2.058	2.15	1.932	1.574	1.892	2.143
Wave Atoms KLT 2 pump	1.949	2.023	1.823	1.9	1.5	1.902
Wave Atoms KLT 4 pump	2.081	2.071	1.951	1.976	1.91	2.052
Sub-band Wiener	1.492	1.728	1.694	1.679	1.68	0
Sub-band C-C wiener HR	1.769	1.75	1.697	1.857	1.715	0

Figure 5.12 comparison of different algorithm corrupted with 5dB Noise and different noise type



	airport	Babble	Car	Street	train	restaurant
ini	2.316	2.301	2.177	2.21	2.139	2.337
Wave Atoms 2 pump	2.243	2.266	2.037	2.193	2.048	2.181
Wave Atoms 4 pump Magnitude	2.38	2.376	2.256	2.286	2.207	2.408
Wave Atoms 4 pump Cplxpair	2.377	2.371	2.252	2.278	2.204	2.4
Wave Atoms Cplxpair Imaginary	2.424	2.376	2.257	2.292	2.274	2.412
Wave Atoms KLT 2 pump	2.408	2.388	2.309	2.207	2.062	2.368
Wave Atoms KLT 4 pump	2.409	2.402	2.242	2.26	2.236	2.43
Sub-band Wiener	2.025	2.034	1.921	2.119	2.009	0
Sub-band C-C wiener HR	2.413	2.276	2.168	2.26	2.096	0

Figure 5.13 comparison of different algorithm corrupted with 10dB Noise and different noise type



	airport	Babble	Car	Street	train	restaurant
ini	2.611	2.63	2.505	2.519	2.47	2.629
Wave Atoms 2 pump	2.621	2.678	2.542	2.524	2.526	2.657
Wave Atoms 4 pump Magnitude	2.671	2.701	2.579	2.579	2.53	2.691
Wave Atoms 4 pump Cplxpair	2.666	2.64	2.515	2.574	2.528	2.739
Wave Atoms Cplxpair Imaginary	2.358	2.703	2.578	2.586	2.483	2.702
Wave Atoms KLT 2 pump	2.705	2.743	2.611	2.635	2.595	2.732
Wave Atoms KLT 4 pump	2.712	2.737	2.618	2.624	2.584	2.728
Sub-band Wiener	2.249	2.127	2.265	2.38	2.04	0
Sub-band C-C wiener HR	2.579	2.609	2.645	2.573	2.032	0

Figure 5.14 comparison of different algorithm corrupted with 15dB Noise and different noise type

Algorithm 3. Maximum a Posteriori Wave atoms algorithm

1.Maximum a Posteriori Wave atoms dual pump

Figure 5.15 shows the spectrogram of the input signal degraded with 10 dB babble noise after applying the algorithm. Table 4.8 shows the PESQ and Seg_SNR measurement output of the algorithm. Comparing the spectrogram of Maximum a Posteriori Wave atoms directional Figure 5.15 with the spectrogram of input degraded signal Figure 5.4 There is clear difference between them, the noise decreases and the spectrogram appears more clear than upper algorithm Wave Atoms dual, Wave Atoms quad, Wave Atoms Cplxpair , Wave Atoms Cplxpair Imaginary, Wave Atoms KLT dual pump and Wave Atoms KLT quad pump . From Table 4.8 we can see the difference between this technique and Wave Atoms dual pump Table 5.8 is great and this difference will proportional increase.

Table 5.8 Maximum a Posteriori Wave atoms dual pump PESQ and SNR result

	0dB		5dB		10dB		15dB	
	PESQ	Seg_SNR	PESQ	Seg_SNR	PESQ	Seg_SNR	PESQ	Seg_SNR
airport	1.643	-1.701	2.048	0.762	2.433	2.727	2.847	5.215
Babble	1.582	-1.784	1.881	0.439	2.368	2.659	2.775	4.388
Car	1.84	-0.159	2.141	1.434	2.343	3.214	2.89	5.637
Street	1.652	-1.117	2.011	0.871	2.332	2.89	2.739	4.655
train	1.663	-0.936	1.996	0.945	2.399	2.973	2.793	4.632
restaurant	1.656	-2.183	1.874	0.295	2.404	2.812	2.774	4.097

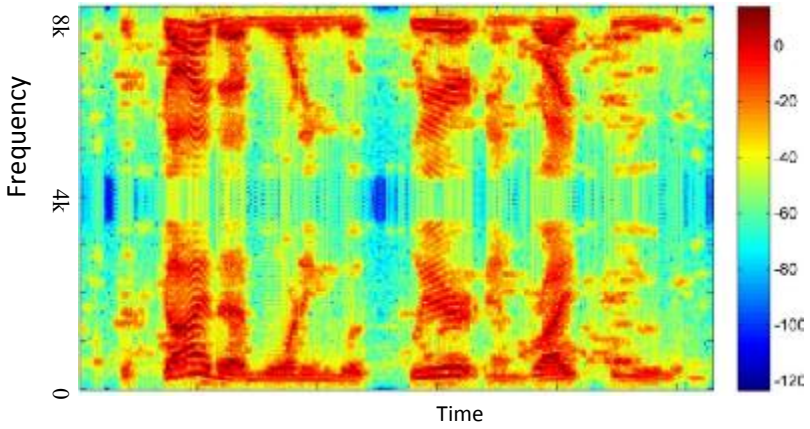


Figure 5.15 Maximum a Posteriori Wave atoms dual pump spectrogram

2.Maximum a Posteriori Wave atoms magnitude quad pump

Figure 4.24 shows the spectrogram of the input signal degraded with 10 dB babble noise after applying the algorithm. Table 5.9 show the PESQ and Seg_SNR measurement output of the algorithm. Comparing the spectrogram of Maximum a Posteriori Wave atoms magnitude quad pump Figure 4.24 with the spectrogram of input degraded signal Figure 4.5 There is explicit difference between them the noise are decreases and the spectrogram appears more clear than upper algorithm Maximum a Posteriori Wave atoms dual pump , Wave Atoms dual pump , Wave Atoms quad pump Magnitude , Wave Atoms quad pump Cplxpair , Wave Atoms Cplxpair Imaginary , Wave Atoms KLT dual pump and Wave Atoms KLT quad pump . From Table 5.9 we can see the difference between this technique and Wave Atoms KLT quad pump Table 5.7

and Wave Atoms quad pump Magnitude Table 5.3 is great and this difference will proportional increase.

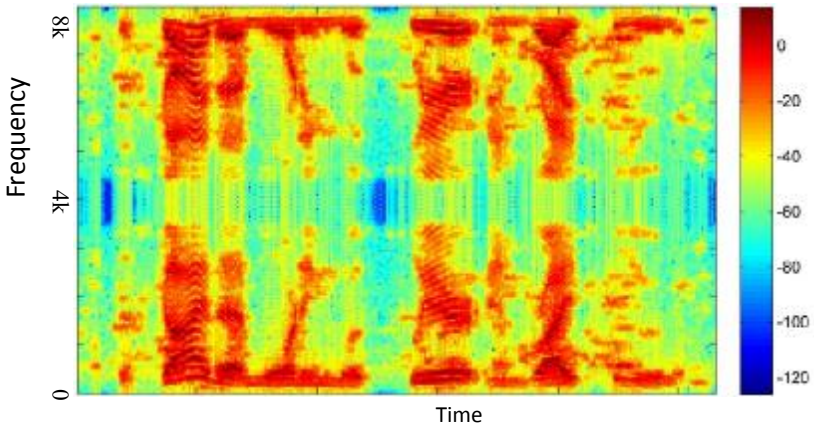


Figure 5.16 Maximum a Posteriori Wave atoms complex spectrogram

Table 5.9 Maximum a Posteriori Wave atoms quad pump PESQ and Seg_SNR result

	0dB		5dB		10dB		15dB	
	PESQ	Seg_SNR	PESQ	Seg_SNR	PESQ	Seg_SNR	PESQ	Seg_SNR
airport	1.807	1.859	2.052	1.974	2.534	2.411	2.773	2.88
Babble	1.602	-2.035	1.938	0.513	2.47	3.098	2.814	4.85
Car	1.969	-1.301	2.11	1.056	2.549	3.341	2.865	4.879
Street	1.906	-0.93	2.013	1.266	2.344	3.241	2.723	4.74
train	1.788	-0.823	1.971	1.318	2.326	3.645	2.872	5.219
restaurant	1.677	-2.242	1.929	0.669	2.371	2.837	2.742	4.988

3.Maximum a Posteriori Wave atoms quad pump Cplxpair

Figure 5.17 shows the spectrogram of the input signal degraded with 10 dB babble noise after applying the algorithm. Table 5.10 shows the PESQ and Seg_SNR measurement output of the algorithm. From Table 5.10, Table 5.9 we can find that the Maximum a Posteriori Wave atoms magnitude quad pump technique has better performance than Maximum a Posteriori Wave atoms quad pump Cplxpair in term of PESQ.

Table 5.10 Maximum a Posteriori Wave atoms Cplxpair PESQ and Seg_SNR

	0dB		5dB		10dB		15dB	
	PESQ	Seg_SNR	PESQ	Seg_SNR	PESQ	Seg_SNR	PESQ	Seg_SNR
airport	1.616	-1.584	2.028	0.593	2.442	2.648	2.862	4.387
Babble	1.601	-1.67	2.029	0.504	2.439	2.536	2.802	3.701
Car	1.828	-0.05	2.129	1.771	2.544	3.484	2.905	4.87
Street	1.652	-0.971	2.039	0.972	2.355	2.923	2.751	3.847
train	1.607	-0.786	2.012	1.115	2.39	3.025	2.803	4.162
restaurant	1.694	-2.024	2.003	0.36	2.426	2.596	2.755	3.693

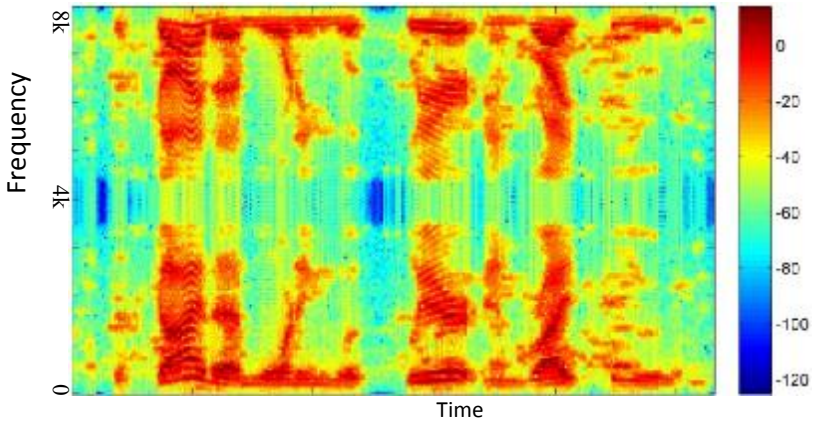


Figure 5.17 Maximum a Posteriori Wave atoms quad pump Cplxpair spectrogram

4.Maximum a Posteriori Wave atoms quad pump imaginary

Figure 5.18 show the spectrogram of the input signal degraded with 10 dB babble noise after applying the algorithm. Table 5.11 show the PESQ and Seg_SNR measurement output of the algorithm

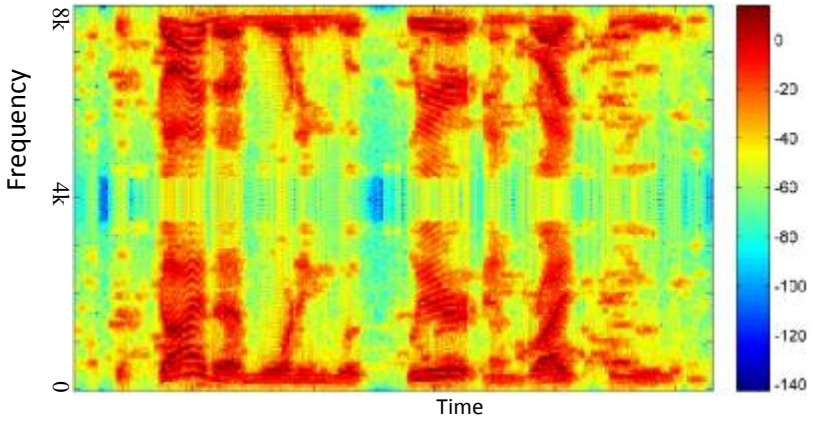


Figure 5.18 Maximum a Posteriori Wave atoms imaginary Spectrogram

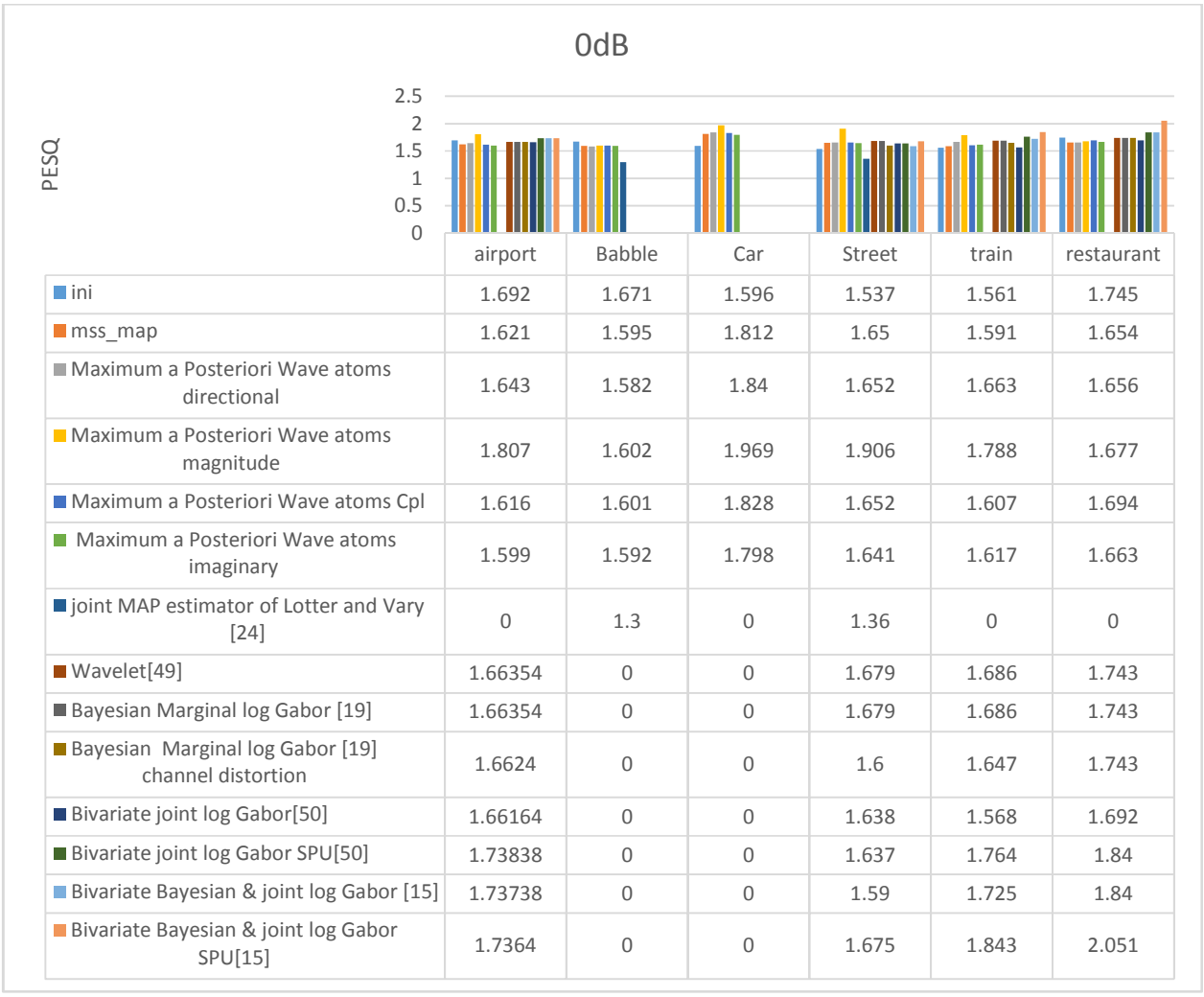
Table 5.11 Maximum a Posteriori Wave atoms quad pump imaginary PESQ and Seg_SNR result

	0dB		5dB		10dB		15dB	
	PESQ	Seg_SNR	PESQ	Seg_SNR	PESQ	Seg_SNR	PESQ	Seg_SNR
airport	1.599	-1.646	2.009	0.826	2.425	3.425	2.848	5.835
Babble	1.592	-1.742	2.011	0.744	2.424	3.296	2.79	5.467
Car	1.798	-1.649	2.104	0.828	2.525	3.425	2.879	5.839
Street	1.641	-0.978	2.016	1.296	2.34	3.763	2.733	5.56
train	1.617	-1.644	1.997	0.823	2.376	3.39	2.785	5.755
restaurant	1.663	-2.123	1.959	0.576	2.407	3.356	2.737	5.496

Figure 5.19, Figure 5.20, Figure 5.21 and Figure 5.22 show the comparison between our technique and similar techniques using the same dataset. The table are collected from the related search papers [20] , [16], [13] and [41]. Comparing the MSS-MAP results with our algorithm you can find about 3% enhancement than MSS-MAP. This means that the MAP algorithm gives more enhancement when combined with wave atoms and KLT. Although, using the complex part only and real only of wave atoms transform, it does not give the best PESQ enhancement when combined with MAP. we see the a good mount of enhancement when comparing our result with result of using Wavelet[42], Bayesian Marginal log Gabor [16] with additive noise alone and channel distortion, joint log Gabor[43] with and without SPU and Bayesian & joint log Gabor with and without SPU[13] have the best enhancement over these technique.

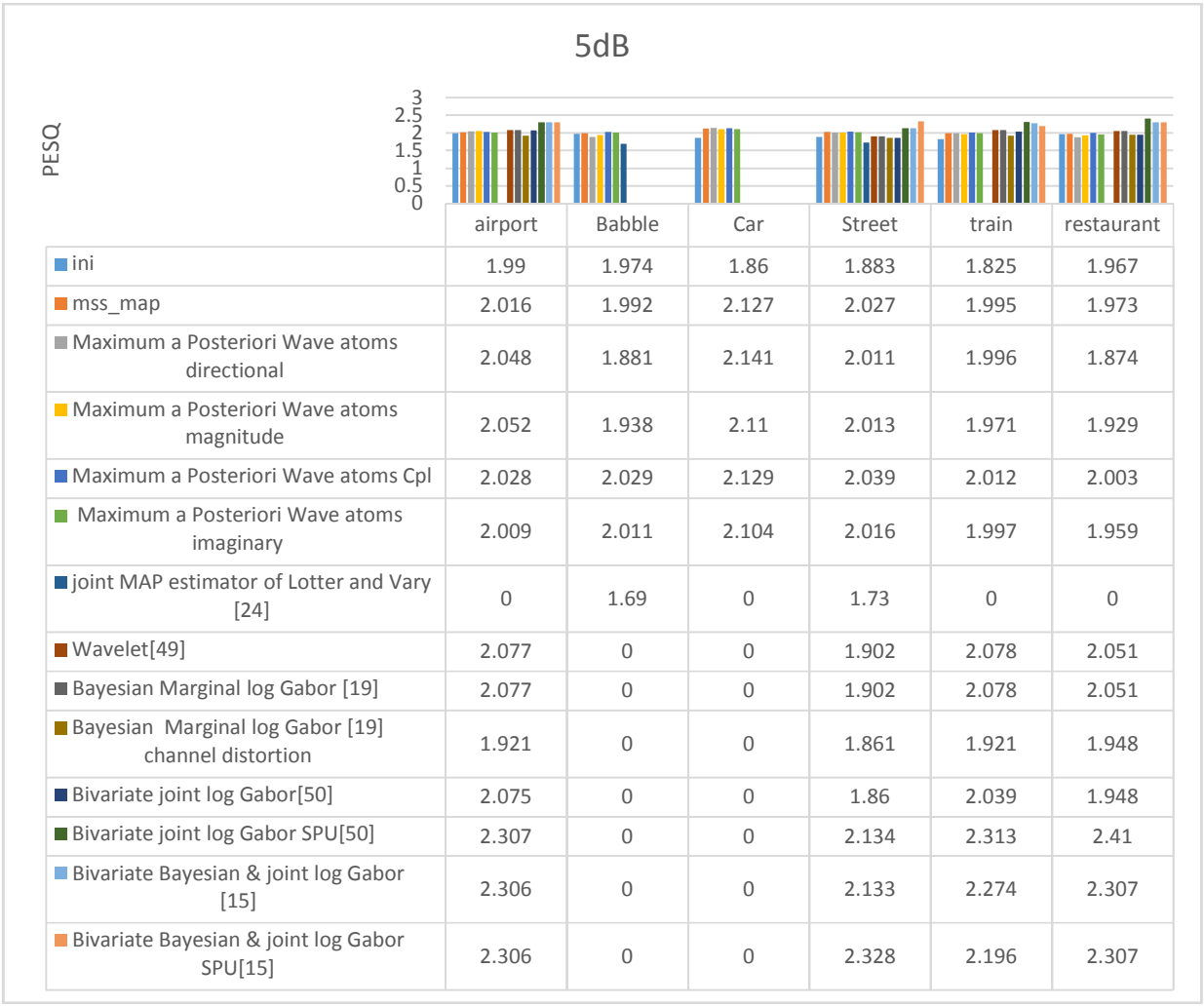
From Figure 5.19, Figure 5.20, Figure 5.21 and Figure 5.22 we demonstrate that Maximum a Posteriori Wave atoms dual pump the PESQ score proportionally increased with increasing and we have average enhancement 5.5% which means more than 8% enhancement than Wave Atoms dual pump and 6% than Wave Atoms KLT dual pump . Maximum a Posteriori Wave atoms magnitude quad pump have 8% average enhancement comparing result of Maximum a Posteriori Wave atoms magnitude quad pump with Wave Atoms quad pump we have about 5% enhancement with it and 4% than Wave Atoms KLT quad pump . Result of Maximum a Posteriori Wave atoms quad pump Cplxpair have 7% average enhancement. Result of Maximum a Posteriori Wave atoms quad pump imaginary have 6% average enhancement.

From the previous discussion we conclude that the best enhancement when use MAP with wave atoms is achieved with Maximum a Posteriori Wave atoms magnitude quad pump technique.



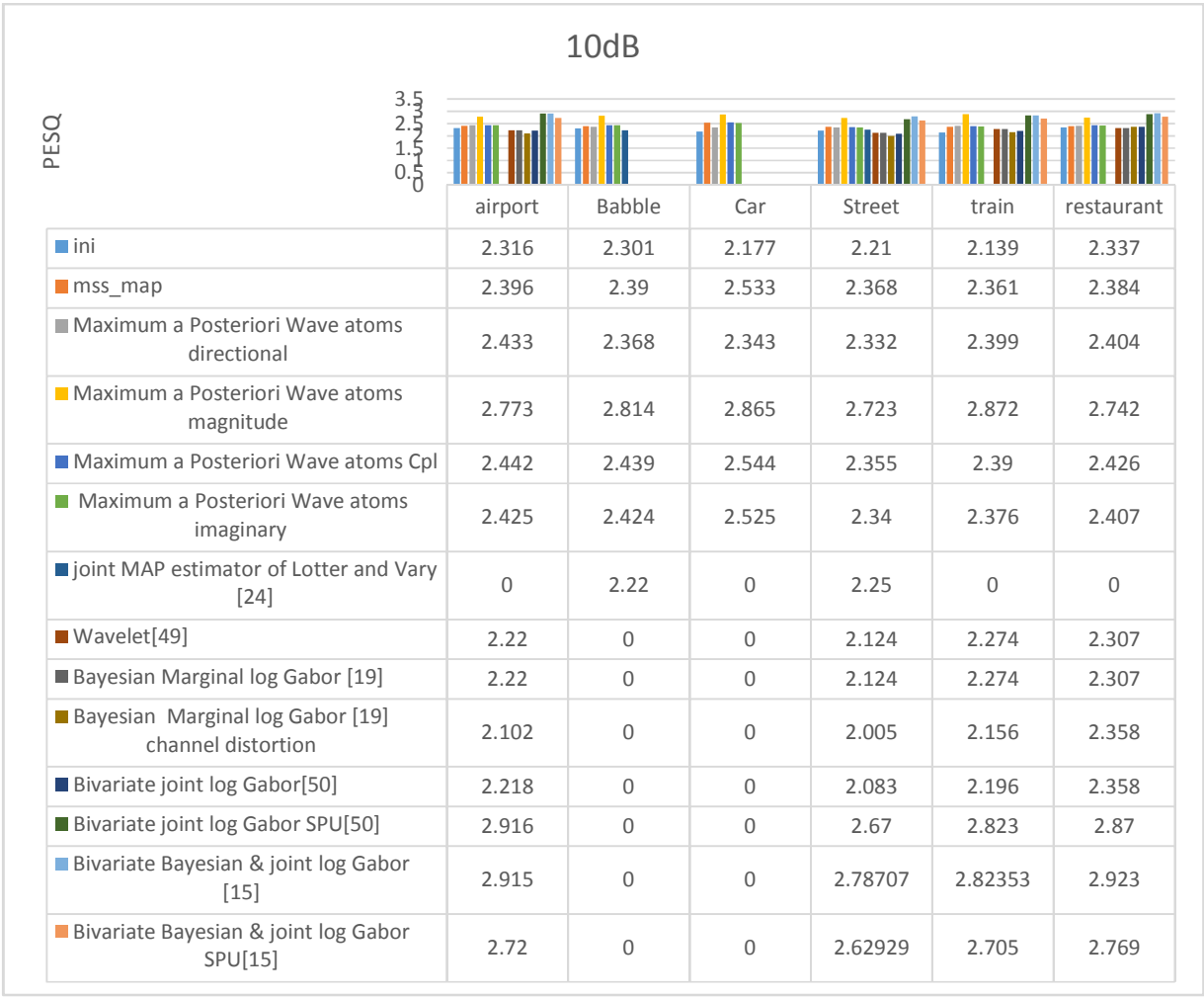
	airport	Babble	Car	Street	train	restaurant
ini	1.692	1.671	1.596	1.537	1.561	1.745
mss_map	1.621	1.595	1.812	1.65	1.591	1.654
Maximum a Posteriori Wave atoms directional	1.643	1.582	1.84	1.652	1.663	1.656
Maximum a Posteriori Wave atoms magnitude	1.807	1.602	1.969	1.906	1.788	1.677
Maximum a Posteriori Wave atoms Cpl	1.616	1.601	1.828	1.652	1.607	1.694
Maximum a Posteriori Wave atoms imaginary	1.599	1.592	1.798	1.641	1.617	1.663
joint MAP estimator of Lotter and Vary [24]	0	1.3	0	1.36	0	0
Wavelet[49]	1.66354	0	0	1.679	1.686	1.743
Bayesian Marginal log Gabor [19]	1.66354	0	0	1.679	1.686	1.743
Bayesian Marginal log Gabor [19] channel distortion	1.6624	0	0	1.6	1.647	1.743
Bivariate joint log Gabor[50]	1.66164	0	0	1.638	1.568	1.692
Bivariate joint log Gabor SPU[50]	1.73838	0	0	1.637	1.764	1.84
Bivariate Bayesian & joint log Gabor [15]	1.73738	0	0	1.59	1.725	1.84
Bivariate Bayesian & joint log Gabor SPU[15]	1.7364	0	0	1.675	1.843	2.051

Figure 5.19 comparison of different algorithm corrupted with 0dB Noise and different noise type



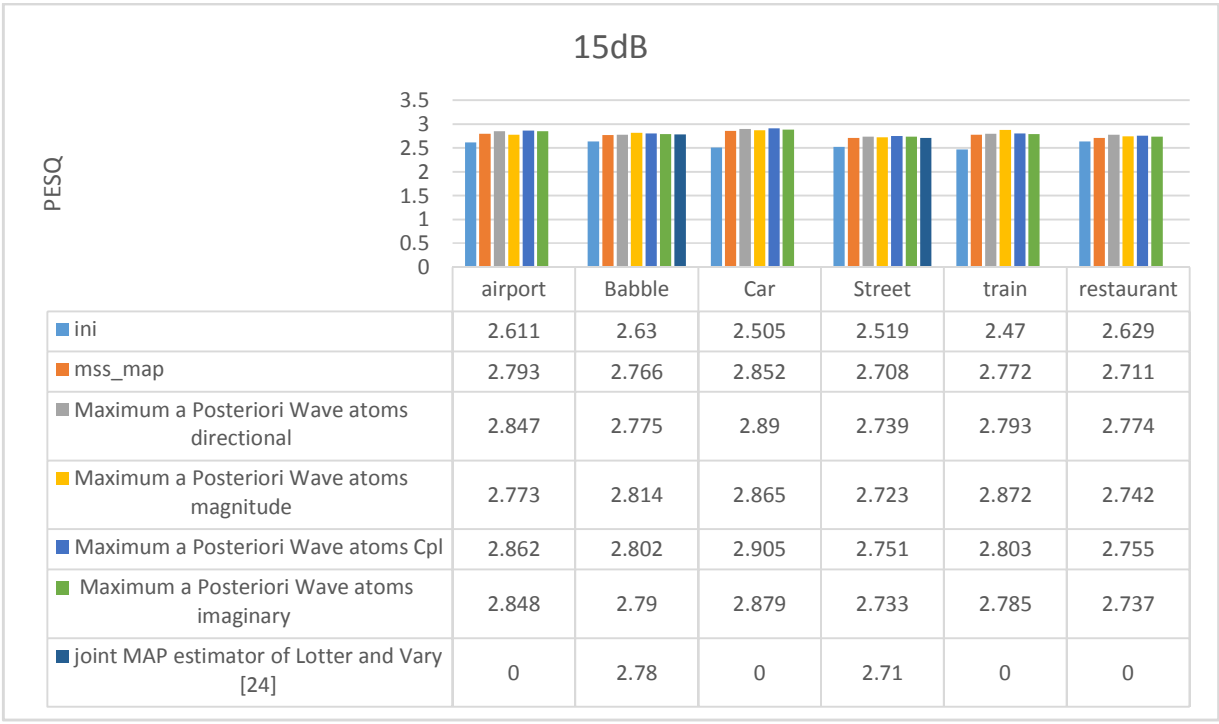
	airport	Babble	Car	Street	train	restaurant
ini	1.99	1.974	1.86	1.883	1.825	1.967
mss_map	2.016	1.992	2.127	2.027	1.995	1.973
Maximum a Posteriori Wave atoms directional	2.048	1.881	2.141	2.011	1.996	1.874
Maximum a Posteriori Wave atoms magnitude	2.052	1.938	2.11	2.013	1.971	1.929
Maximum a Posteriori Wave atoms Cpl	2.028	2.029	2.129	2.039	2.012	2.003
Maximum a Posteriori Wave atoms imaginary	2.009	2.011	2.104	2.016	1.997	1.959
joint MAP estimator of Lotter and Vary [24]	0	1.69	0	1.73	0	0
Wavelet[49]	2.077	0	0	1.902	2.078	2.051
Bayesian Marginal log Gabor [19]	2.077	0	0	1.902	2.078	2.051
Bayesian Marginal log Gabor [19] channel distortion	1.921	0	0	1.861	1.921	1.948
Bivariate joint log Gabor[50]	2.075	0	0	1.86	2.039	1.948
Bivariate joint log Gabor SPU[50]	2.307	0	0	2.134	2.313	2.41
Bivariate Bayesian & joint log Gabor [15]	2.306	0	0	2.133	2.274	2.307
Bivariate Bayesian & joint log Gabor SPU[15]	2.306	0	0	2.328	2.196	2.307

Figure 5.20 comparison of different algorithm corrupted with 5dB Noise and different noise type



	airport	Babble	Car	Street	train	restaurant
ini	2.316	2.301	2.177	2.21	2.139	2.337
mss_map	2.396	2.39	2.533	2.368	2.361	2.384
Maximum a Posteriori Wave atoms directional	2.433	2.368	2.343	2.332	2.399	2.404
Maximum a Posteriori Wave atoms magnitude	2.773	2.814	2.865	2.723	2.872	2.742
Maximum a Posteriori Wave atoms Cpl	2.442	2.439	2.544	2.355	2.39	2.426
Maximum a Posteriori Wave atoms imaginary	2.425	2.424	2.525	2.34	2.376	2.407
joint MAP estimator of Lotter and Vary [24]	0	2.22	0	2.25	0	0
Wavelet[49]	2.22	0	0	2.124	2.274	2.307
Bayesian Marginal log Gabor [19]	2.22	0	0	2.124	2.274	2.307
Bayesian Marginal log Gabor [19] channel distortion	2.102	0	0	2.005	2.156	2.358
Bivariate joint log Gabor[50]	2.218	0	0	2.083	2.196	2.358
Bivariate joint log Gabor SPU[50]	2.916	0	0	2.67	2.823	2.87
Bivariate Bayesian & joint log Gabor [15]	2.915	0	0	2.78707	2.82353	2.923
Bivariate Bayesian & joint log Gabor SPU[15]	2.72	0	0	2.62929	2.705	2.769

Figure 5.21 comparison of different algorithm corrupted with 5dB Noise and different noise type



	airport	Babble	Car	Street	train	restaurant
ini	2.611	2.63	2.505	2.519	2.47	2.629
mss_map	2.793	2.766	2.852	2.708	2.772	2.711
Maximum a Posteriori Wave atoms directional	2.847	2.775	2.89	2.739	2.793	2.774
Maximum a Posteriori Wave atoms magnitude	2.773	2.814	2.865	2.723	2.872	2.742
Maximum a Posteriori Wave atoms Cpl	2.862	2.802	2.905	2.751	2.803	2.755
Maximum a Posteriori Wave atoms imaginary	2.848	2.79	2.879	2.733	2.785	2.737
joint MAP estimator of Lotter and Vary [24]	0	2.78	0	2.71	0	0

Figure 5.22 comparison of different algorithm corrupted with 15dB Noise and different noise type

Algorithm 4. MMSE Wave atoms with KLT and Spectrum Power Zero Cross

1. MMSE Wave atoms with KLT and Spectrum Power Zero Cross dual pump

Figure 5.23 shows the spectrogram of the input signal degraded with 10 dB babble noise after applying the algorithm. Figure 5.23 shows that the enhancement is rise corresponding to the previous technique and similar to Figure 5.3 show the original clear signal spectrogram more than input signal Figure 5.4. Table 5.12 show the PESQ and Seg_SNR measurement output of the algorithm.

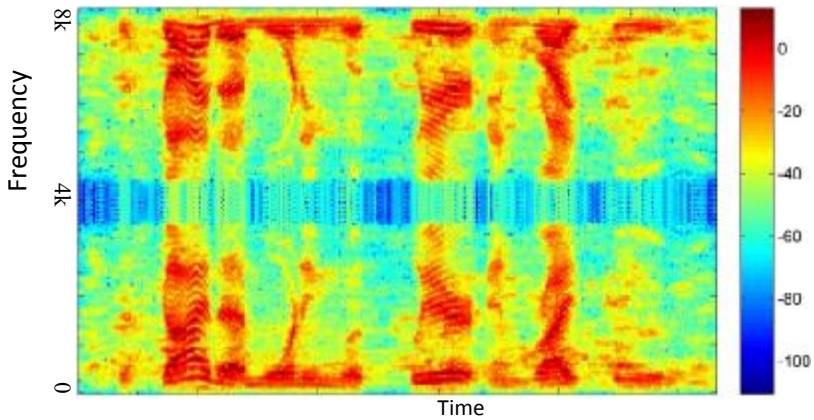


Figure 5.23 MMSE Wave atoms with SPZC dual pump spectrogram

Table 5.12 MMSE Wave atoms with SPZC dual pump PESQ and Seg_SNR result

	0dB		5dB		10dB		15dB	
	PESQ	Seg_SNR	PESQ	Seg_SNR	PESQ	Seg_SNR	PESQ	Seg_SNR
airport	1.811	-2.148	2.158	-0.095	2.573	2.339	2.939	3.472
Babble	1.585	-2.331	2.289	0.279	2.562	2.086	2.94	3.674
Car	1.889	-1.657	2.087	0.5	2.589	2.901	2.896	4.226
Street	1.793	-1.843	2.012	0.261	2.515	2.56	2.858	3.446
train	1.424	-1.539	2.191	0.45	2.488	2.363	2.88	4.118
restaurant	1.659	-2.251	2.282	0.354	2.598	2.653	2.939	3.897

2.MMSE Wave Atoms with Spectrum Power Zero Cross quad pump Magnitude

Figure 5.24 show the spectrogram of the input signal degraded with 10 dB babble noise after applying the algorithm. Figure 5.24 show that the enhancement is rise corresponding to the previous technique and similar to Figure 5.3 show the original clear signal spectrogram more than input signal Figure 5.4 Table 5.13 show the PESQ and Seg_SNR measurement output of the algorithm. Comparing result in Table 5.13with Table 5.12 we can find clear difference between them in PESQ score.

Table 5.13 MMSE Wave Atoms with SPZC quad pump Magnitude PESQ and Seg_SNR result

	0dB		5dB		10dB		15dB	
	PESQ	Seg_SNR	PESQ	Seg_SNR	PESQ	Seg_SNR	PESQ	Seg_SNR
airport	1.835	-1.895	1.972	0.259	2.551	2.239	2.902	4.39
Babble	1.737	-1.899	2.233	0.408	2.567	2.59	2.948	4.491
Car	1.796	-1.597	2.119	0.516	2.494	2.703	2.933	3.55
Street	1.713	-1.615	1.758	0.511	2.513	2.372	2.854	2.943
train	1.74	-1.362	1.7	0.537	2.484	2.505	2.859	4.141
restaurant	1.796	-2.22	2.131	0.09	2.554	2.411	2.867	3.865

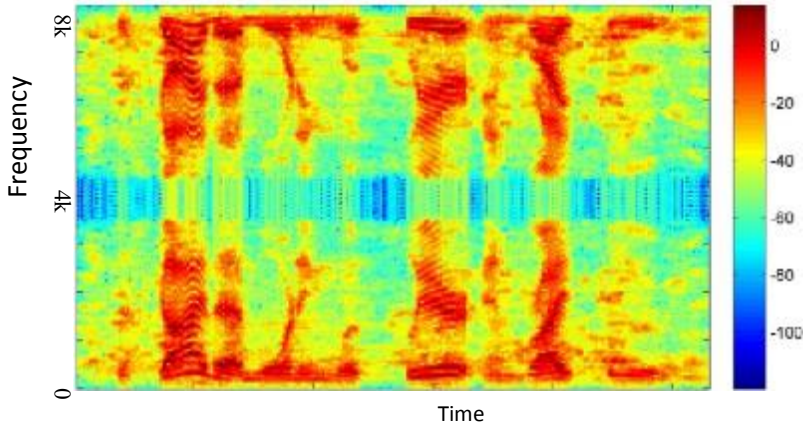


Figure 5.24 MMSE Wave Atoms with SPZC Magnitude spectrogram

3. MMSE Wave Atoms with Spectrum Power Zero Cross quad pump Cplxpair

Figure 5.25 show the spectrogram of the input signal degraded with 10 dB babble noise after applying the algorithm. Table 5.14 show the average PESQ and SNR result after applying the algorithm on all database signals. Comparing result in Table 5.14 with Table 5.13 we can find unequivocal difference in PESQ score. Here we can see the important of using complex value in enhancement process.

Table 5.14 MMSE Wave Atoms with SPZC quad pump Cplxpair PESQ and Seg_SNR result

	0dB		5dB		10dB		15dB	
	PESQ	Seg_SNR	PESQ	Seg_SNR	PESQ	Seg_SNR	PESQ	Seg_SNR
airport	1.851	-1.928	2.219	0.089	2.575	2.014	2.927	2.942
Babble	1.865	-2.017	2.241	0.083	2.586	1.869	2.947	2.689
Car	1.9	-1.432	2.243	0.931	2.613	2.213	2.938	3.065
Street	1.831	-1.602	1.847	0.8	2.528	2.155	2.872	2.487
train	1.882	-1.293	2.332	0.983	2.489	2.161	2.891	2.708
restaurant	1.694	-2.138	2.003	0.02	2.426	2.026	2.755	2.584

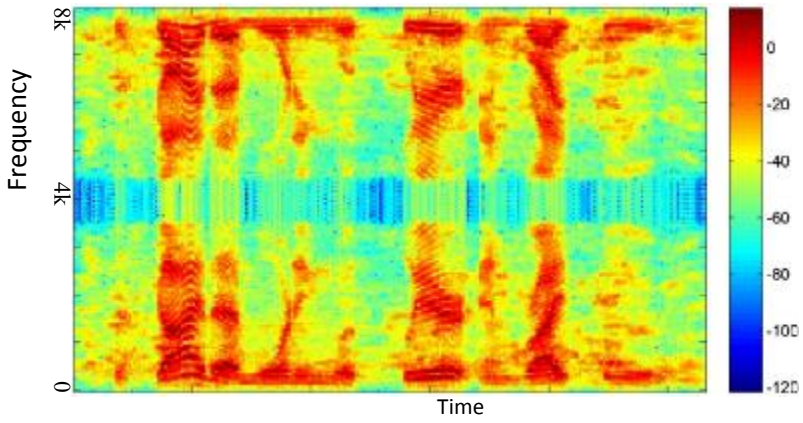


Figure 5.25 MMSE Wave Atoms with SPZC quad pump Cplxpair spectrogram

4. MMSE Wave Atoms with Spectrum Power Zero Cross quad pump Imaginary

Figure 4.40 show the spectrogram of the input signal degraded with 10 dB babble noise after applying the algorithm. Table 4.15 show the average PESQ and SNR result after applying the algorithm on all database signals. Comparing result in Table 4.15 with Table 4.13 we can find unequivocal difference in PESQ score. Here we can see the important of using complex value in enhancement process.

Table 5.15 MMSE Wave Atoms with SPZC quad pump Imaginary PESQ and Seg_SNR result

	0dB		5dB		10dB		15dB	
	PESQ	Seg_SNR	PESQ	Seg_SNR	PESQ	Seg_SNR	PESQ	Seg_SNR
airport	1.599	-1.879	2.009	0.504	2.425	2.933	2.848	4.886
Babble	1.592	-1.977	2.011	0.471	2.424	2.765	2.79	4.879
Car	1.912	-1.266	2.236	0.954	2.613	3.191	2.932	5.184
Street	1.832	-1.496	2.223	0.745	2.526	3.128	2.865	4.502
train	1.881	-1.313	2.175	0.916	2.497	3.186	2.894	5.06
restaurant	1.663	-2.127	1.959	0.38	2.407	2.969	2.737	4.505

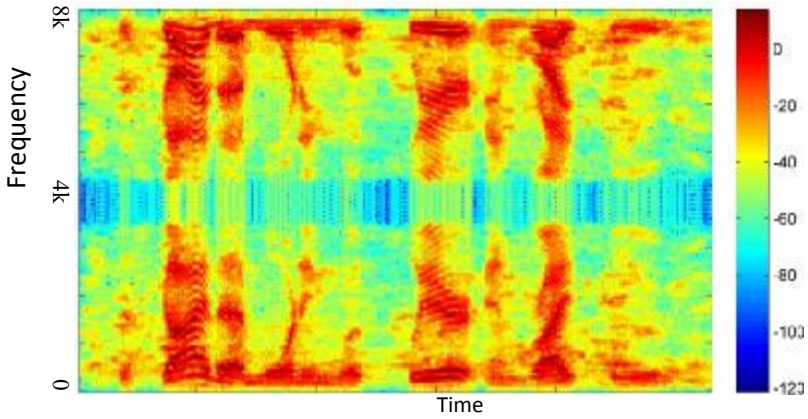


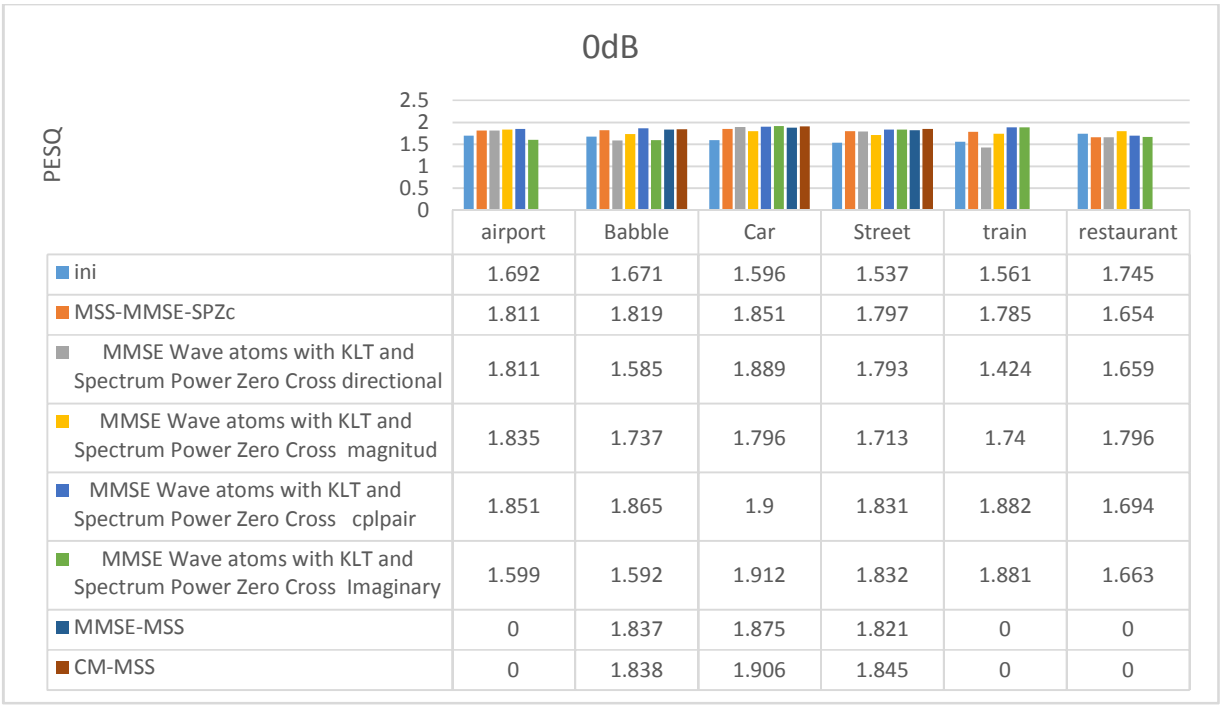
Figure 5.26 MMSE Wave Atoms with SPZC quad pump Imaginary spectrogram

Figure 5.27, Figure 5.28, Figure 5.29 and Figure 5.30 show the comparison between our technique and similar technique using the same dataset the table are collected from the related search papers [44]. The figures show that MMSE-SPZC stage make more enhancement .Comparing the result MMSE-SPZC with our algorithm you can find about 3% enhancement than MMSE-SPZC. This mean that the SPZC algorithm gives more enhancement when combined with wave atoms and KLT. Although, using the complex part only and real only of wave atoms transform give the best PESQ enhancement when combined with SPZC. Although, using MMSE Wave atoms with Spectrum Power Zero Cross get more than 6% enhancement than Maximum a Posteriori Wave atoms algorithm that described in the previous section. we see the an good mount of enhancement when compare our result with result of using MMSE_MSS, Conditional Median Estimator CM-MSS[44]

Figure 5.27, Figure 5.28, Figure 5.29 and Figure 5.30 demonstrate that MMSE Wave atoms with KLT and Spectrum Power Zero Cross dual pump has 11% average enhancement and PESQ score proportionally increased with increasing the dB with 3% enhancement in 0db and increased to 13% with 15dB. Comparing this result with Maximum a Posteriori Wave atoms dual pump we have double enhancement 5.5% and more than 13% enhancement than Wave Atoms dual pump and 11% than Wave Atoms KLT dual pump . Result of MMSE Wave Atoms with Spectrum Power Zero Cross quad pump Magnitude has 11% average enhancement and PESQ score proportionally increased with increasing the dB with 8% enhancement in 0db and increased to 13% with 15dB. Comparing this result with Maximum a Posteriori Wave atoms magnitude quad pump we have 2% enhancement. Result of MMSE Wave Atoms with Spectrum Power Zero

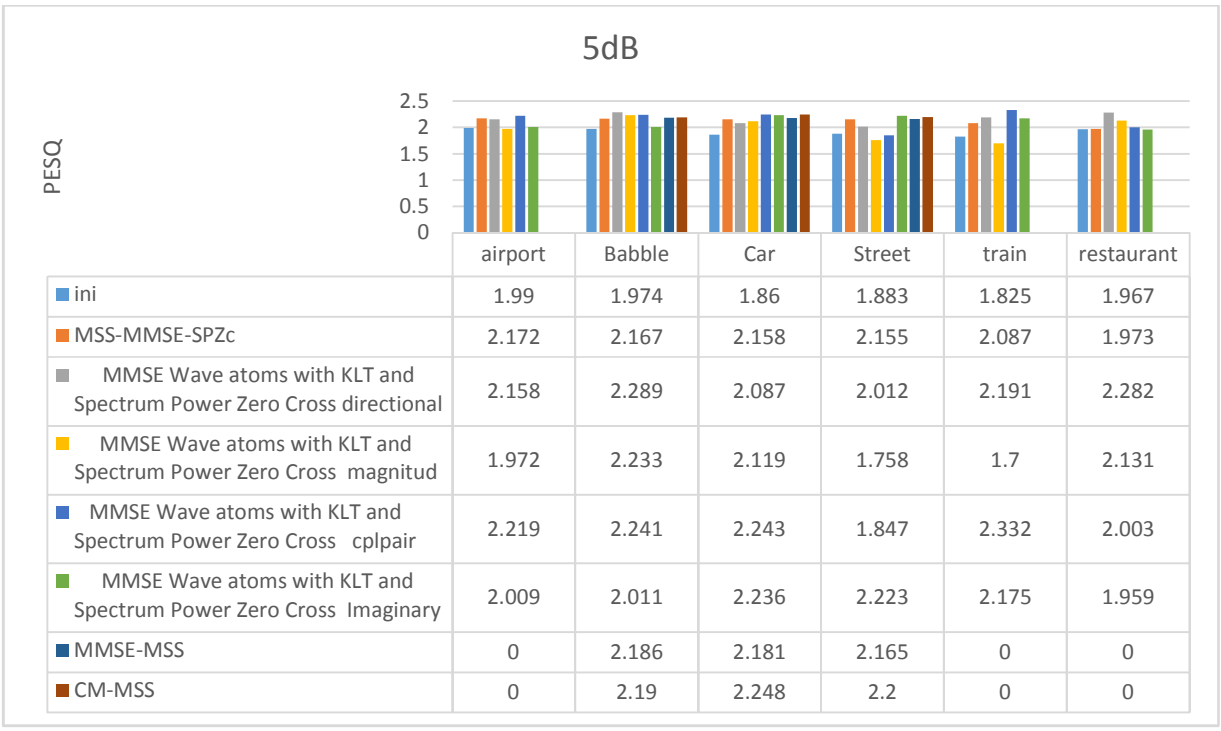
Cross quad pump Cplxpair has 12% average in all dB (0dB-15dB). Comparing this result with Maximum a Posteriori Wave atoms quad pump Cplxpair we have about 7% (double) enhancement. Result of MMSE Wave Atoms with Spectrum Power Zero Cross quad pump Imaginary has 11% average enhancement and PESQ score proportionally increased with increasing the dB with 8% enhancement in 0db and increased to 11% with 15dB. The result has double enhancement than Maximum a Posteriori Wave atoms imaginary.

We can conclude that the best enhancement has been acquire with MMSE Wave Atoms with Spectrum Power Zero Cross quad pump Cplxpair which use complex pairs in thresholding and MMSE Wave Atoms with Spectrum Power Zero Cross quad pump Imaginary has similar enhancement with small difference this mean that we can use the imaginary information to enhance the signal with technique.



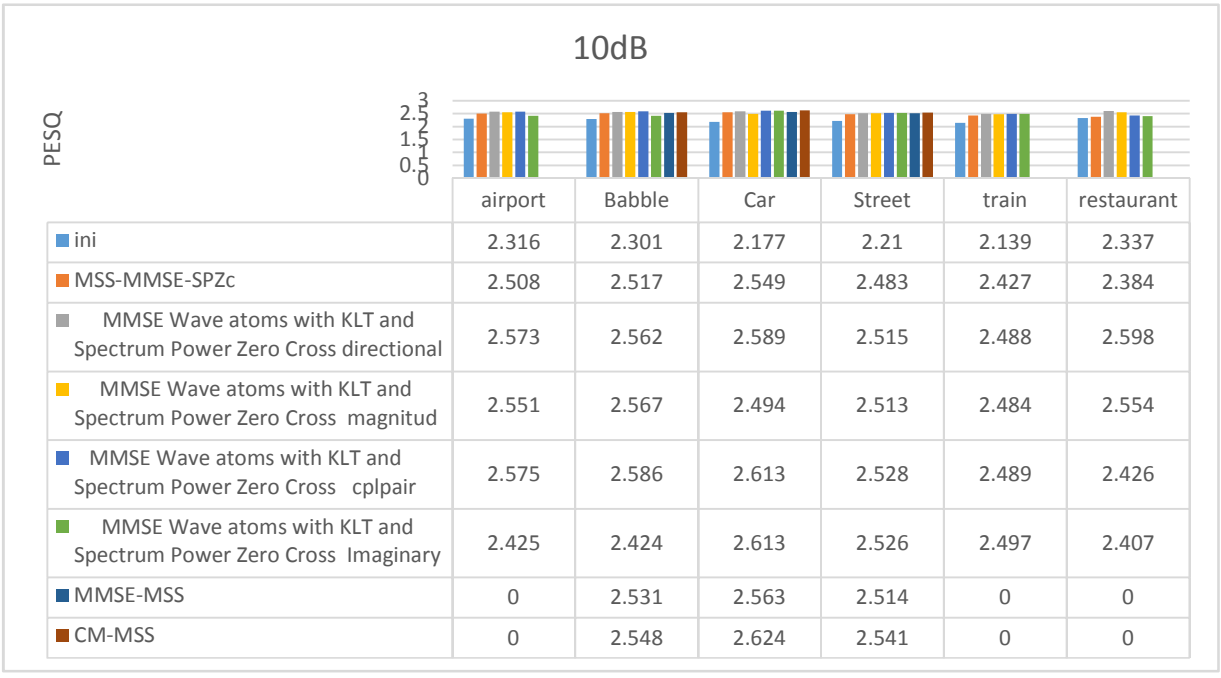
	airport	Babble	Car	Street	train	restaurant
ini	1.692	1.671	1.596	1.537	1.561	1.745
MSS-MMSE-SPZc	1.811	1.819	1.851	1.797	1.785	1.654
MMSE Wave atoms with KLT and Spectrum Power Zero Cross directional	1.811	1.585	1.889	1.793	1.424	1.659
MMSE Wave atoms with KLT and Spectrum Power Zero Cross magnitud	1.835	1.737	1.796	1.713	1.74	1.796
MMSE Wave atoms with KLT and Spectrum Power Zero Cross cplpair	1.851	1.865	1.9	1.831	1.882	1.694
MMSE Wave atoms with KLT and Spectrum Power Zero Cross Imaginary	1.599	1.592	1.912	1.832	1.881	1.663
MMSE-MSS	0	1.837	1.875	1.821	0	0
CM-MSS	0	1.838	1.906	1.845	0	0

Figure 5.27 comparison of different algorithm corrupted with 0dB Noise and different noise type



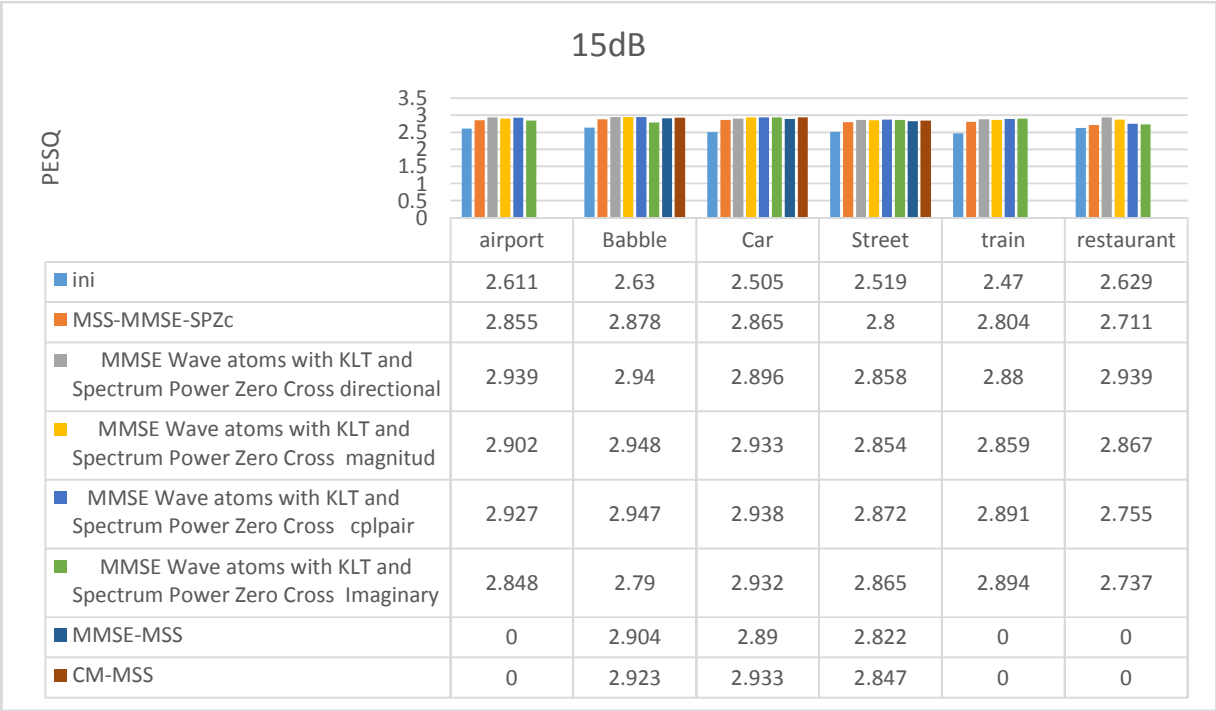
	airport	Babble	Car	Street	train	restaurant
ini	1.99	1.974	1.86	1.883	1.825	1.967
MSS-MMSE-SPZc	2.172	2.167	2.158	2.155	2.087	1.973
MMSE Wave atoms with KLT and Spectrum Power Zero Cross directional	2.158	2.289	2.087	2.012	2.191	2.282
MMSE Wave atoms with KLT and Spectrum Power Zero Cross magnitud	1.972	2.233	2.119	1.758	1.7	2.131
MMSE Wave atoms with KLT and Spectrum Power Zero Cross cplpair	2.219	2.241	2.243	1.847	2.332	2.003
MMSE Wave atoms with KLT and Spectrum Power Zero Cross Imaginary	2.009	2.011	2.236	2.223	2.175	1.959
MMSE-MSS	0	2.186	2.181	2.165	0	0
CM-MSS	0	2.19	2.248	2.2	0	0

Figure 5.28 comparison of different algorithm corrupted with 5dB Noise and different noise type



	airport	Babble	Car	Street	train	restaurant
ini	2.316	2.301	2.177	2.21	2.139	2.337
MSS-MMSE-SPZc	2.508	2.517	2.549	2.483	2.427	2.384
MMSE Wave atoms with KLT and Spectrum Power Zero Cross directional	2.573	2.562	2.589	2.515	2.488	2.598
MMSE Wave atoms with KLT and Spectrum Power Zero Cross magnitud	2.551	2.567	2.494	2.513	2.484	2.554
MMSE Wave atoms with KLT and Spectrum Power Zero Cross cplpair	2.575	2.586	2.613	2.528	2.489	2.426
MMSE Wave atoms with KLT and Spectrum Power Zero Cross Imaginary	2.425	2.424	2.613	2.526	2.497	2.407
MMSE-MSS	0	2.531	2.563	2.514	0	0
CM-MSS	0	2.548	2.624	2.541	0	0

Figure 5.29 comparison of different algorithm corrupted with 10dB Noise and different noise type



	airport	Babble	Car	Street	train	restaurant
ini	2.611	2.63	2.505	2.519	2.47	2.629
MSS-MMSE-SPZc	2.855	2.878	2.865	2.8	2.804	2.711
MMSE Wave atoms with KLT and Spectrum Power Zero Cross directional	2.939	2.94	2.896	2.858	2.88	2.939
MMSE Wave atoms with KLT and Spectrum Power Zero Cross magnitud	2.902	2.948	2.933	2.854	2.859	2.867
MMSE Wave atoms with KLT and Spectrum Power Zero Cross cplpair	2.927	2.947	2.938	2.872	2.891	2.755
MMSE Wave atoms with KLT and Spectrum Power Zero Cross Imaginary	2.848	2.79	2.932	2.865	2.894	2.737
MMSE-MSS	0	2.904	2.89	2.822	0	0
CM-MSS	0	2.923	2.933	2.847	0	0

Figure 5.30 comparison of different algorithm corrupted with 15dB Noise and different noise type

Algorithm 5. MMSE Wave atoms with KLT and Spectrum Power Zero Cross SNR Uncertainty

1. MMSE Wave atoms with KLT and Spectrum Power Zero Cross SNR Uncertainty dual pump

Figure 5.31 show the spectrogram of the input signal degraded with 10 dB babble noise after applying the algorithm that make visual indication the distribution of noise is decreased and formant appears clearer than input signal Figure 5.4. Table 5.16 show the PESQ and Seg_SNR result of the algorithm. Comparing Table 5.16 with result of Table 5.12 Wave Atoms we find that this technique does not make en enhancement like Wave Atoms dual pump.

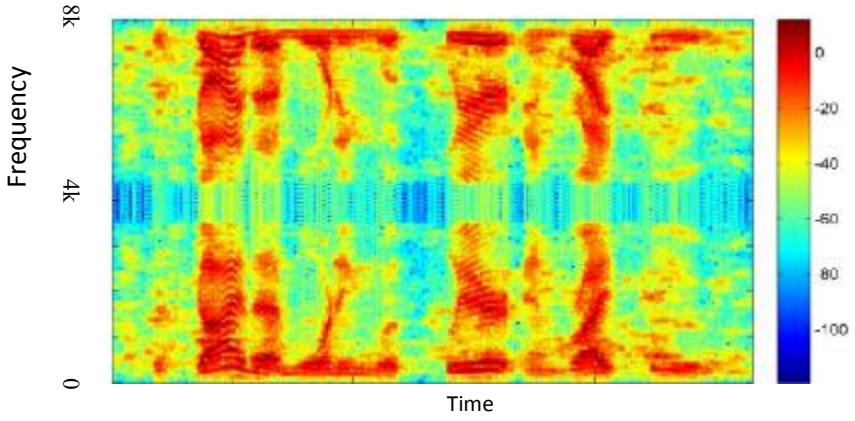


Figure 5.31 MMSE Wave atoms with KLT and Spectrum Power Zero Cross SNR Uncertainty directional dual pump spectrogram

Table 5.16 MMSE Wave atoms with KLT and Spectrum Power Zero Cross SNR Uncertainty directional dual pump PESQ and Seg_SNR result

	0dB		5dB		10dB		15dB	
	PESQ	Seg_SNR	PESQ	Seg_SNR	PESQ	Seg_SNR	PESQ	Seg_SNR
airport	1.809	-1.645	2.191	0.079	2.565	2.415	2.939	4.182
Babble	1.835	-1.74	2.296	0.23	2.557	2.381	2.924	4.126
Car	1.848	-0.892	2.124	0.936	2.654	3.058	2.97	4.528
Street	1.664	-1.305	2.196	0.702	2.52	2.778	2.868	3.883
train	1.776	-1.163	2.081	0.813	2.501	2.772	2.898	4.542
restaurant	1.909	-1.864	2.289	0.119	2.593	2.845	2.923	4.096

2. MMSE Wave atoms with KLT and Spectrum Power Zero Cross SNR Uncertainty quad pump

Figure 5.32 show the spectrogram of the input signal degraded with 10 dB babble noise after applying the algorithm. Table 5.17 show the PESQ and Seg_SNR result of the algorithm. Comparing Table 5.16 with Table 5.14 we have increasing in PESQ score this mean that using dual pump with magnitude increase the PESQ better than quad pump.

Table 5.17 MMSE Wave atoms with KLT and Spectrum Power Zero Cross SNR Uncertainty quad pump PESQ and Seg_SNR result

	0dB		5dB		10dB		15dB	
	PESQ	Seg_SNR	PESQ	Seg_SNR	PESQ	Seg_SNR	PESQ	Seg_SNR
airport	1.759	-1.86	2.198	0.349	2.574	2.802	2.92	5.111
Babble	1.644	-1.908	2.171	0.522	2.563	2.825	2.918	4.611
Car	1.666	-0.851	2.337	1.16	2.63	3.237	2.97	5.121
Street	1.784	-1.515	2.125	0.736	2.499	2.762	2.9	4.037
train	1.836	-0.935	2.074	1.118	2.468	3.14	2.909	4.956
restaurant	1.718	-1.994	2.093	0.371	2.605	2.908	2.862	4.317

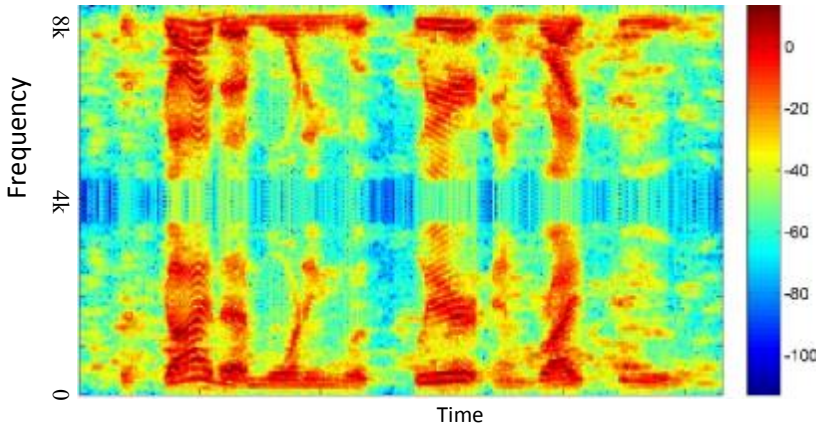


Figure 5.32 MMSE Wave atoms with KLT and Spectrum Power Zero Cross SNR Uncertainty quad pump spectrogram

3. MMSE Wave atoms with KLT and Spectrum Power Zero Cross SNR Uncertainty quad pump Cplxpair

Figure 5.33 shows the spectrogram of the input signal degraded with 10 dB babble noise after applying the algorithm. Table 5.18 show the result of PESQ and Seg_SNR scores. Comparing Table 5.17 with Table 5.18 we find an enhancement about 3% using complex value rather using magnitude only but still MMSE Wave Atoms with Spectrum Power Zero Cross quad pump Cplxpair give better performance in PESQ score.

Table 5.18 MMSE Wave atoms with KLT and Spectrum Power Zero Cross SNR Uncertainty quad pump PESQ and Seg_SNR result

	0dB		5dB		10dB		15dB	
	PESQ	Seg_SNR	PESQ	Seg_SNR	PESQ	Seg_SNR	PESQ	Seg_SNR
airport	1.837	-1.69	2.205	0.341	2.564	2.196	2.959	3.891
Babble	1.841	-1.787	2.219	0.267	2.582	2.162	2.935	2.96
Car	1.96	-0.905	2.282	0.869	2.635	2.6	2.984	3.929
Street	1.837	-1.282	2.213	0.59	2.512	2.406	2.88	2.716
train	1.88	-0.928	2.19	0.676	2.508	2.525	2.908	3.68
restaurant	1.863	-1.978	2.131	0.202	2.549	2.24	2.882	2.873

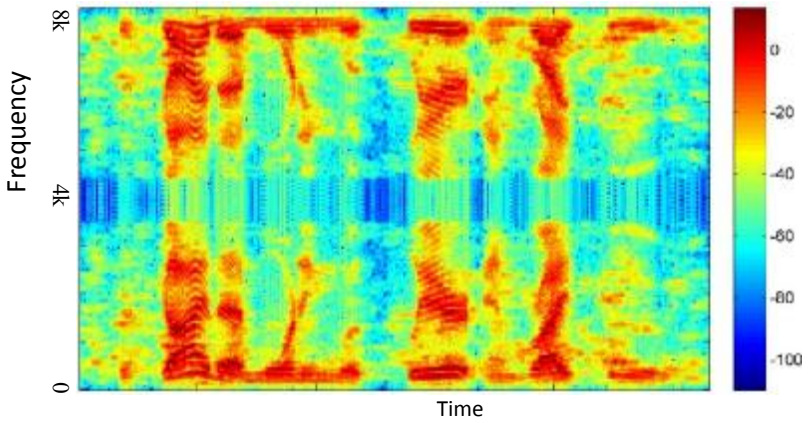


Figure 5.33 MMSE Wave atoms with KLT and Spectrum Power Zero Cross SNR Uncertainty quad pump spectrogram

4. MMSE Wave atoms with KLT and Spectrum Power Zero Cross SNR Uncertainty quad pump Imaginary

Figure 5.34 show the spectrogram of the input signal degraded with 10 dB babble noise after applying the algorithm. Table 5.19 show the result of PESQ and Seg_SNR of the algorithm. Comparing Table 5.19 with Table 5.18 we find an enhancement about 0.5% using Imaginary value rather using complex pair but still MMSE Wave Atoms with Spectrum Power Zero Cross quad pump Imaginary give better performance in PESQ score. We can conclude that using the SNR Uncertainty does not give better performance

Table 5.19 MMSE Wave atoms with KLT and Spectrum Power Zero Cross SNR Uncertainty quad pump Imaginary PESQ and Seg_SNR result

	0dB		**5dB**		**10dB**		**15dB**	
	PESQ	Seg_SNR	PESQ	Seg_SNR	PESQ	Seg_SNR	PESQ	Seg_SNR
airport	1.826	-1.636	2.195	0.741	2.553	3.084	2.952	5.318
Babble	1.847	-1.746	2.207	0.654	2.575	3.014	2.926	5.086
Car	1.946	-0.725	2.272	1.362	2.621	3.535	2.969	5.591
Street	1.833	-1.171	2.201	1.035	2.502	3.371	2.865	4.872
train	1.869	-0.956	2.19	1.202	2.506	3.512	2.901	5.49
restaurant	1.812	-1.969	2.118	0.552	2.534	3.157	2.87	4.92

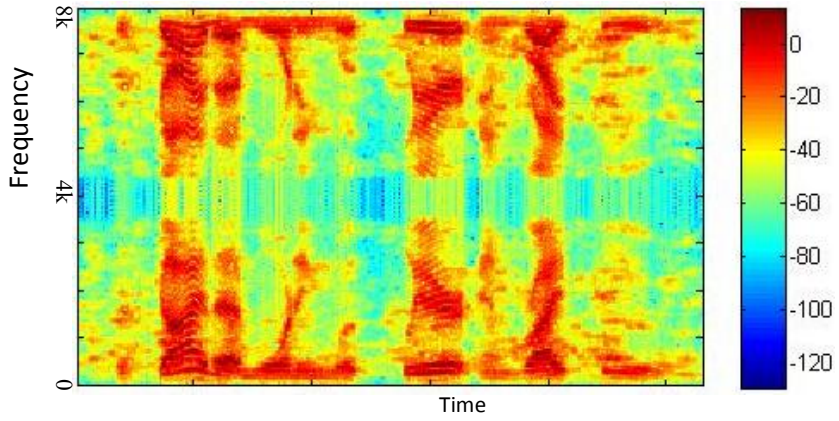
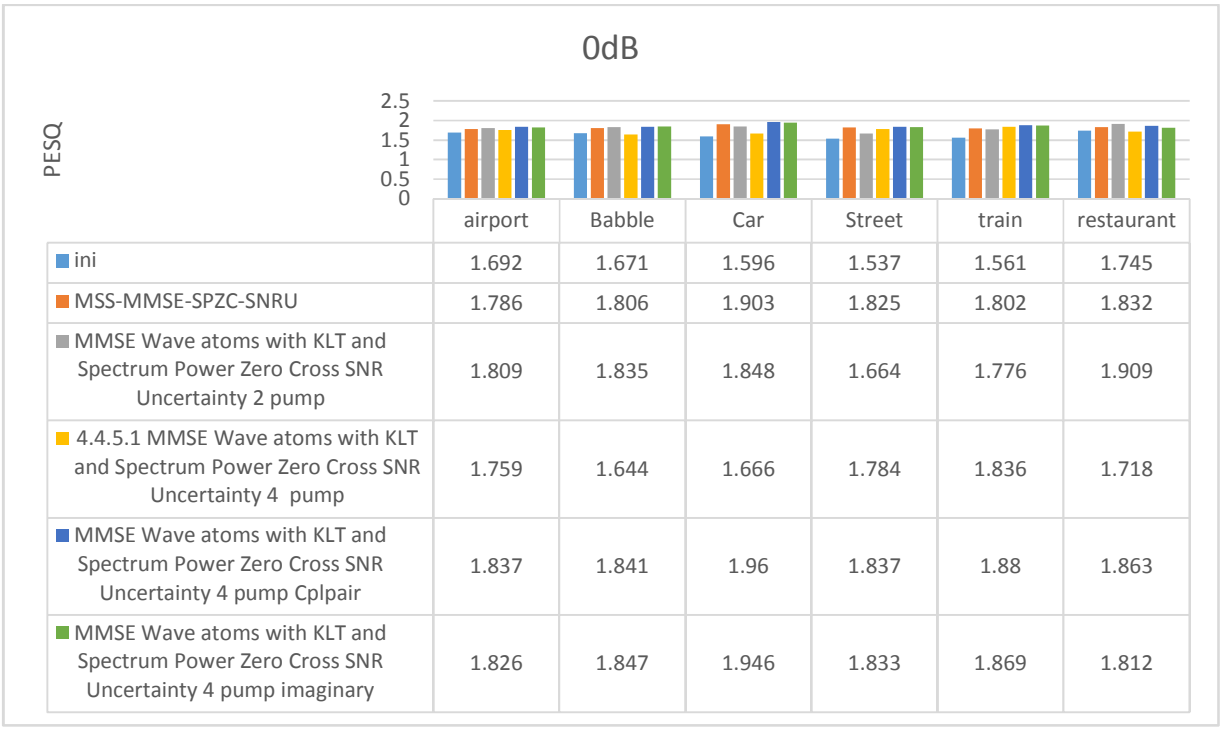


Figure 5.34 MMSE Wave atoms with KLT and Spectrum Power Zero Cross SNR Uncertainty quad pump Imaginary spectrogram

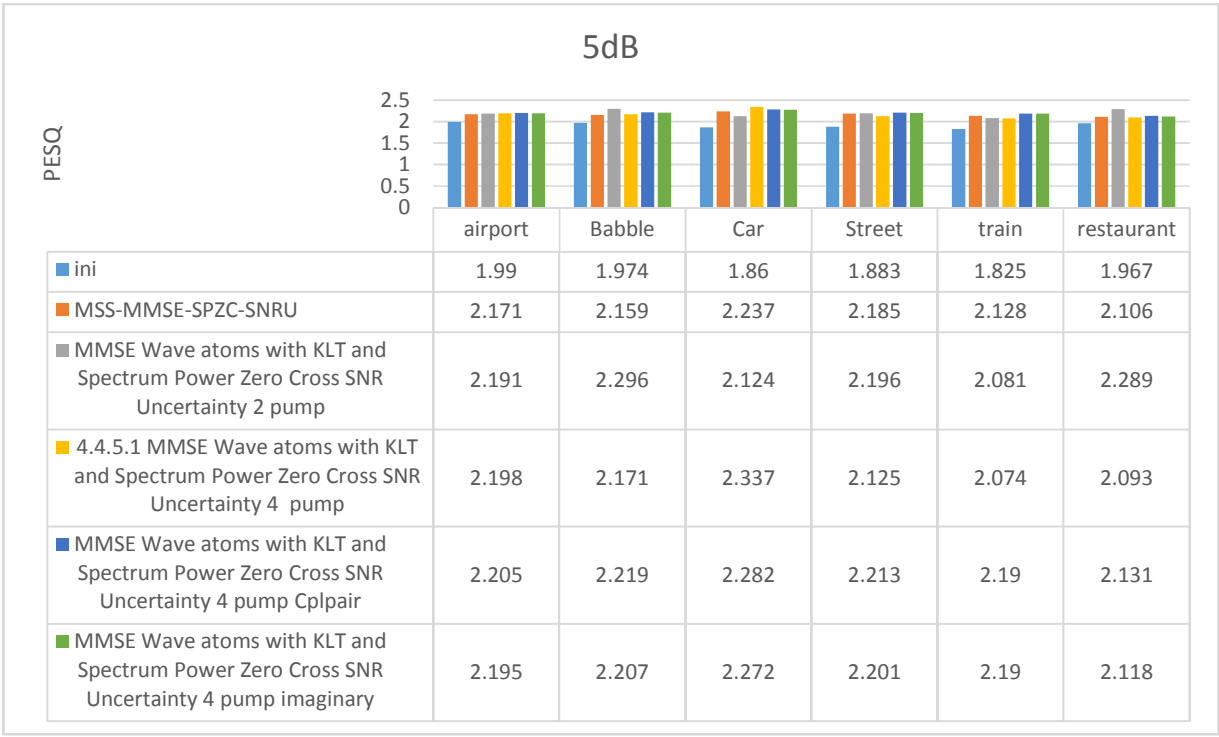
Figure 5.35, Figure 5.36, Figure 5.37 and Figure 5.38 show the comparison between our technique and MMSE-SPZC_SNRU. The figures show that MMSE-SPZC_SNRU stage make more enhancement .Comparing the result MMSE-SPZC_SNRU with our algorithm you can find about 2% enhancement than MMSE-SPZC_SNRU. This mean that the SPZC_SNRU algorithm gives more enhancement when combined with wave atoms and KLT. Although, using the complex part only and real only of wave atoms transform give the best PESQ enhancement when combined with SPZC. From Figure 5.35, Figure 5.36, Figure 5.37 and Figure 5.38 we can see MMSE Wave atoms with KLT and Spectrum Power Zero Cross SNR Uncertainty has 13.5% average enhancement and PESQ score proportionally increased with increasing the dB with 10% enhancement in 0db and increased to 14% with 15dB. Compare this result with MMSE Wave atoms with KLT and Spectrum Power Zero Cross dual pump we have 2% enhancement in PESQ score. Result of MMSE Wave atoms with KLT and Spectrum Power Zero Cross SNR Uncertainty quad pump has 11% average enhancement and PESQ score proportionally increased with increasing the dB with 8% enhancement in 0db and increased to 13% with 15dB. Comparing this result with MMSE Wave Atoms with Spectrum Power Zero Cross quad pump Magnitude we have 2% enhancement. Result of MMSE Wave atoms with KLT and Spectrum Power Zero Cross SNR Uncertainty quad pump Cplxpair has 14.5% average in all dB (0dB-15dB). Comparing this result with MMSE Wave Atoms with Spectrum Power Zero Cross quad pump Cplxpair we have about 2% enhancement. Result of MMSE Wave atoms with KLT and Spectrum Power Zero Cross SNR Uncertainty quad pump Imaginary has 14% in all dB (0dB-15dB). The result has 4% enhancement than MMSE Wave Atoms with Spectrum Power Zero Cross quad pump Imaginary. We can conclude that the best enhancement has been gained with

MMSE Wave atoms with KLT and Spectrum Power Zero Cross SNR Uncertainty quad pump Cplxpair which use complex pairs in thresholding and MMSE Wave atoms with KLT and Spectrum Power Zero Cross SNR Uncertainty quad pump Imaginary has similar enhancement with small difference this mean that we can use the imaginary information to enhance the signal with technique.



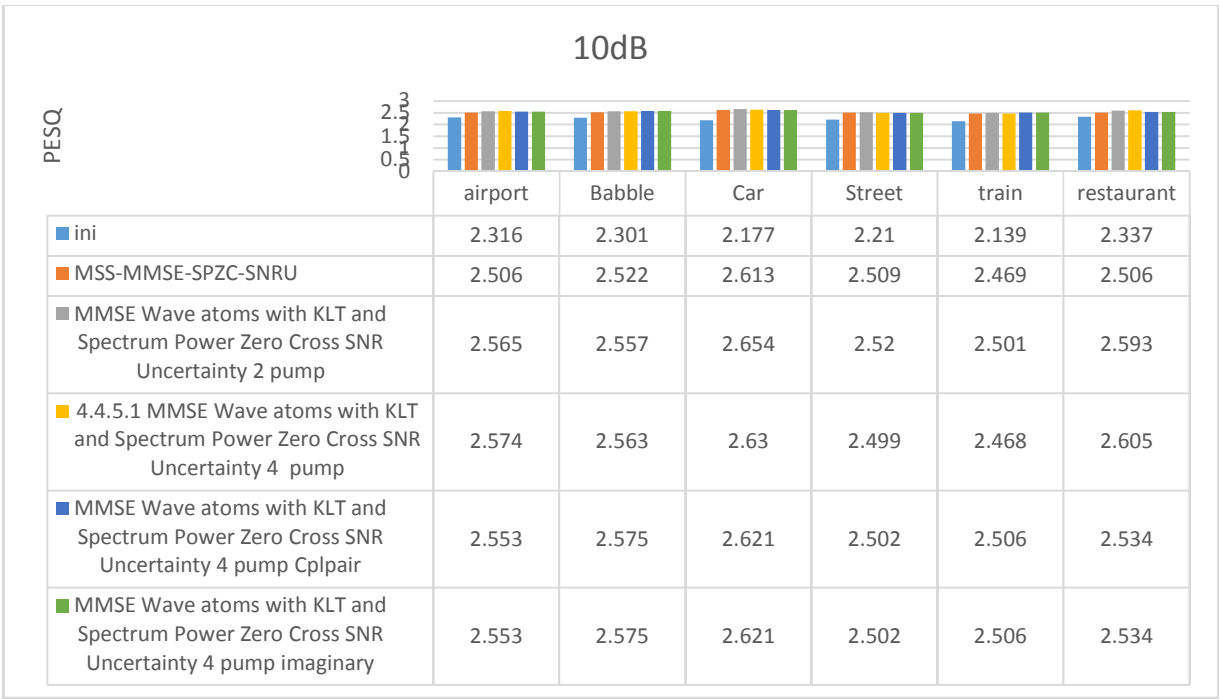
	airport	Babble	Car	Street	train	restaurant
ini	1.692	1.671	1.596	1.537	1.561	1.745
MSS-MMSE-SPZC-SNRU	1.786	1.806	1.903	1.825	1.802	1.832
MMSE Wave atoms with KLT and Spectrum Power Zero Cross SNR Uncertainty 2 pump	1.809	1.835	1.848	1.664	1.776	1.909
4.4.5.1 MMSE Wave atoms with KLT and Spectrum Power Zero Cross SNR Uncertainty 4 pump	1.759	1.644	1.666	1.784	1.836	1.718
MMSE Wave atoms with KLT and Spectrum Power Zero Cross SNR Uncertainty 4 pump Cplpair	1.837	1.841	1.96	1.837	1.88	1.863
MMSE Wave atoms with KLT and Spectrum Power Zero Cross SNR Uncertainty 4 pump imaginary	1.826	1.847	1.946	1.833	1.869	1.812

Figure 5.35 comparison of different algorithm corrupted with 0dB Noise and different noise type



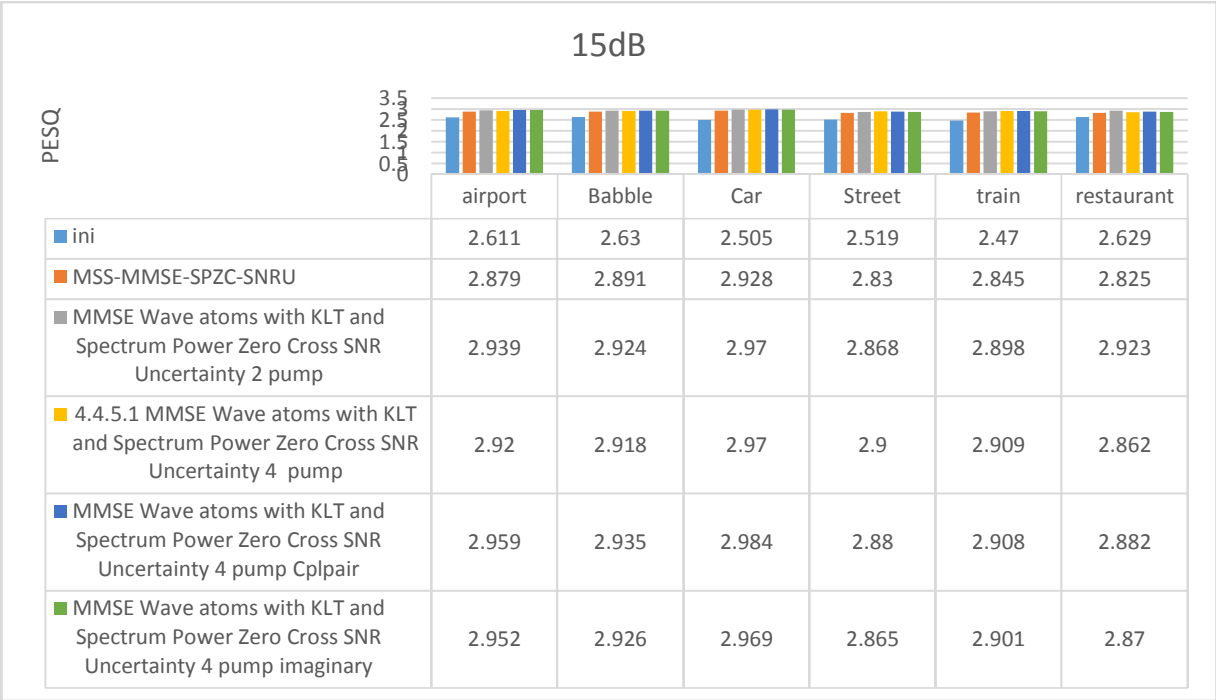
	airport	Babble	Car	Street	train	restaurant
ini	1.99	1.974	1.86	1.883	1.825	1.967
MSS-MMSE-SPZC-SNRU	2.171	2.159	2.237	2.185	2.128	2.106
MMSE Wave atoms with KLT and Spectrum Power Zero Cross SNR Uncertainty 2 pump	2.191	2.296	2.124	2.196	2.081	2.289
4.4.5.1 MMSE Wave atoms with KLT and Spectrum Power Zero Cross SNR Uncertainty 4 pump	2.198	2.171	2.337	2.125	2.074	2.093
MMSE Wave atoms with KLT and Spectrum Power Zero Cross SNR Uncertainty 4 pump Cplpair	2.205	2.219	2.282	2.213	2.19	2.131
MMSE Wave atoms with KLT and Spectrum Power Zero Cross SNR Uncertainty 4 pump imaginary	2.195	2.207	2.272	2.201	2.19	2.118

Figure 5.36 comparison of different algorithm corrupted with 5dB Noise and different noise type



	airport	Babble	Car	Street	train	restaurant
ini	2.316	2.301	2.177	2.21	2.139	2.337
MSS-MMSE-SPZC-SNRU	2.506	2.522	2.613	2.509	2.469	2.506
MMSE Wave atoms with KLT and Spectrum Power Zero Cross SNR Uncertainty 2 pump	2.565	2.557	2.654	2.52	2.501	2.593
4.4.5.1 MMSE Wave atoms with KLT and Spectrum Power Zero Cross SNR Uncertainty 4 pump	2.574	2.563	2.63	2.499	2.468	2.605
MMSE Wave atoms with KLT and Spectrum Power Zero Cross SNR Uncertainty 4 pump Cplpair	2.553	2.575	2.621	2.502	2.506	2.534
MMSE Wave atoms with KLT and Spectrum Power Zero Cross SNR Uncertainty 4 pump imaginary	2.553	2.575	2.621	2.502	2.506	2.534

Figure 5.37 comparison of different algorithm corrupted with 10dB Noise and different noise type



	airport	Babble	Car	Street	train	restaurant
ini	2.611	2.63	2.505	2.519	2.47	2.629
MSS-MMSE-SPZC-SNRU	2.879	2.891	2.928	2.83	2.845	2.825
MMSE Wave atoms with KLT and Spectrum Power Zero Cross SNR Uncertainty 2 pump	2.939	2.924	2.97	2.868	2.898	2.923
4.4.5.1 MMSE Wave atoms with KLT and Spectrum Power Zero Cross SNR Uncertainty 4 pump	2.92	2.918	2.97	2.9	2.909	2.862
MMSE Wave atoms with KLT and Spectrum Power Zero Cross SNR Uncertainty 4 pump Cplpair	2.959	2.935	2.984	2.88	2.908	2.882
MMSE Wave atoms with KLT and Spectrum Power Zero Cross SNR Uncertainty 4 pump imaginary	2.952	2.926	2.969	2.865	2.901	2.87

Figure 5.38 comparison of different algorithm corrupted with 15dB Noise and different noise type

Algorithm 6. Wave atoms with Soft Masking using a PRiori SNR uncertainty

1. MMSE Wave atoms with KLT and SMPR SNR uncertainty dual pump

Figure 4.58 Shows the spectrogram of the input signal degraded with 10 dB babble noise after applying the algorithm. Figure 5.39 show that the enhancement is rise corresponding to the previous technique and similar to Figure 4.4 show the original clear signal spectrogram more than input signal Figure 5.4. Table 5.20 show the PESQ and Seg_SNR measurement output of the algorithm. Comparing result in Table 5.20 with Table 5.13 we can find clear difference between them in PESQ score enhancement raised 2%.

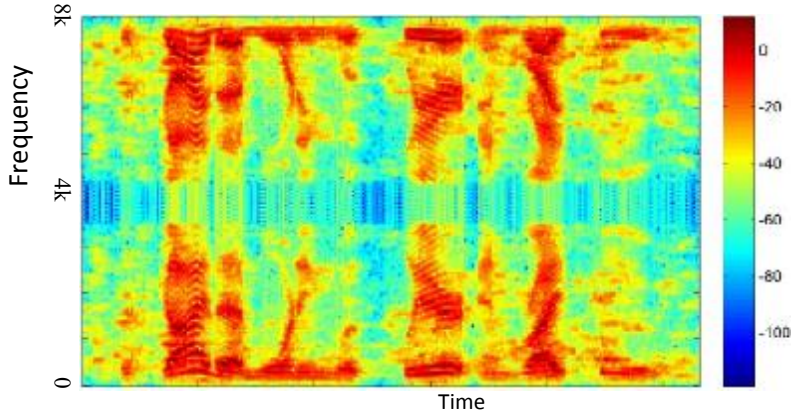


Figure 5.39 MMSE Wave atoms with KLT and SMPR SNR uncertainty dual pump Spectrogram

Table 5.20 MMSE Wave atoms with KLT and SMPR SNR uncertainty dual pump PESQ and Seg_SNR result

	0dB		5dB		10dB		15dB	
	PESQ	Seg_SNR	PESQ	Seg_SNR	PESQ	Seg_SNR	PESQ	Seg_SNR
airport	1.614	-2.049	2.411	0.549	2.638	2.623	2.812	4.519
Babble	1.857	-1.763	2.223	0.528	2.431	2.246	3.003	3.984
Car	1.946	-0.98	2.291	0.919	2.684	2.793	3.043	4.315
Street	1.882	-1.386	2.272	0.561	2.587	2.644	2.96	3.619
train	1.695	-1.284	2.126	0.752	2.683	2.314	2.994	4.425
restaurant	1.931	-1.945	2.216	0.359	2.467	2.54	3.002	3.426

2. MMSE Wave atoms with KLT and SMPR SNR uncertainty quad pump

Figure 5.40 show the spectrogram of the input signal degraded with 10 dB babble noise after applying the algorithm. Figure 5.40 show that the enhancement is rise corresponding to the previous technique and similar to Figure 5.3 show the original clear signal spectrogram more than input signal Figure 5.4. Table 4.21 show the PESQ and Seg_SNR measurement output of the algorithm. Comparing result in Table 4.21with Table 5.14 we can find clear difference between them in PESQ score enhancement raised 4%.

Table 5.21 MMSE Wave atoms with KLT and SMPR SNR uncertainty quad pump PWSQ and Seg_SNR result

	0dB		5dB		10dB		15dB	
	PESQ	Seg_SNR	PESQ	Seg_SNR	PESQ	Seg_SNR	PESQ	Seg_SNR
airport	1.882	-1.659	2.243	0.587	2.613	2.497	3.011	4.598
Babble	1.989	-1.757	2.006	0.33	2.63	2.232	3.015	4.138
Car	1.945	-1.004	2.29	1.017	2.684	2.978	3.024	4.57
Street	1.879	-1.364	2.272	0.806	2.584	3.079	2.962	3.983
train	1.791	-1.13	2.26	0.755	2.433	2.875	2.965	4.659
restaurant	1.885	-1.934	2.165	0.465	2.862	2.712	3.154	4.619

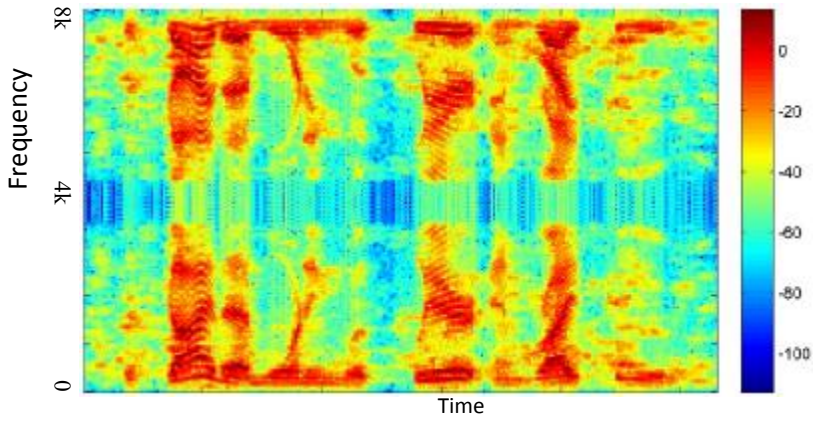


Figure 5.40 MMSE Wave atoms with KLT and SMPR SNR uncertainty quad pump spectrogram

3. MMSE Wave atoms with KLT and SMPR SNR uncertainty quad pump Cplxpair

Figure 5.41 show the spectrogram of the input signal degraded with 10 dB babble noise after applying the algorithm. Figure 5.41 show that the enhancement is rise corresponding to the previous technique and similar to Figure 5.3 show the original clear signal spectrogram more than input signal Figure 5.4. Table 5.22 show the PESQ and Seg_SNR measurement output of the algorithm. Comparing result in Table 5.22 with Table 5.15 we can find clear difference between them in PESQ score enhancement raised 5%.

Table 5.22 MMSE Wave atoms with KLT and SMPR SNR uncertainty quad pump Cplxpair PESQ and Seg_SNR

	0dB		**5dB**		**10dB**		**15dB**	
	PESQ	Seg_SNR	PESQ	Seg_SNR	PESQ	Seg_SNR	PESQ	Seg_SNR
airport	1.887	-1.66	2.251	0.306	2.626	2.101	3.035	3.598
Babble	1.921	-1.743	2.267	0.205	2.642	2.015	3.026	2.856
Car	1.958	-1.045	2.311	0.744	2.699	2.441	3.067	3.39
Street	1.871	-1.3	2.283	0.501	2.601	2.27	2.984	2.718
train	1.899	-0.946	2.214	0.577	2.566	2.271	2.983	2.948
restaurant	1.931	-1.883	2.174	0.273	2.615	2.187	2.952	2.702

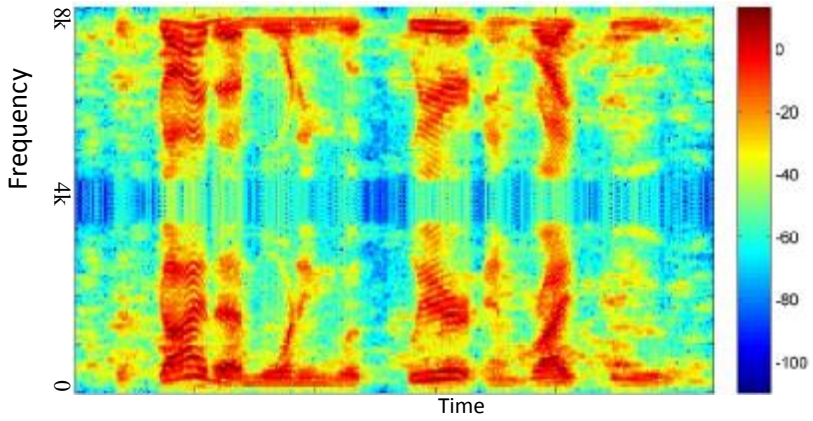


Figure 5.41 MMSE Wave atoms with KLT and SMPR SNR uncertainty quad pump Cplxpair spectrogram

4. MMSE Wave atoms with KLT and SMPR SNR uncertainty quad pump Imaginary

Figure 5.42 show the spectrogram of the input signal degraded with 10 dB babble noise after applying the algorithm. Table 5.23 show the result of PESQ and Seg_SNR of the algorithm. Comparing result in Table 5.23 with Table 5.16 we can find clear difference between them in PESQ score enhancement increased 8%.

Table 5.23 MMSE Wave atoms with KLT and SMPR SNR uncertainty quad pump Imaginary PESQ and Seg_SNR result

	0dB		**5dB**		**10dB**		**15dB**	
	PESQ	Seg_SNR	PESQ	Seg_SNR	PESQ	Seg_SNR	PESQ	Seg_SNR
airport	1.875	-1.563	2.245	0.781	2.617	3.022	3.028	5.17
Babble	1.909	-1.666	2.263	0.651	2.637	2.932	3.018	5.026
Car	1.957	-0.814	2.31	1.32	2.693	3.481	3.061	5.356
Street	1.868	-1.15	2.278	1.032	2.595	3.306	2.974	4.776
train	1.891	-0.966	2.217	1.183	2.572	3.28	2.986	5.17
restaurant	1.919	-1.831	2.163	0.689	2.601	3.154	2.942	4.819

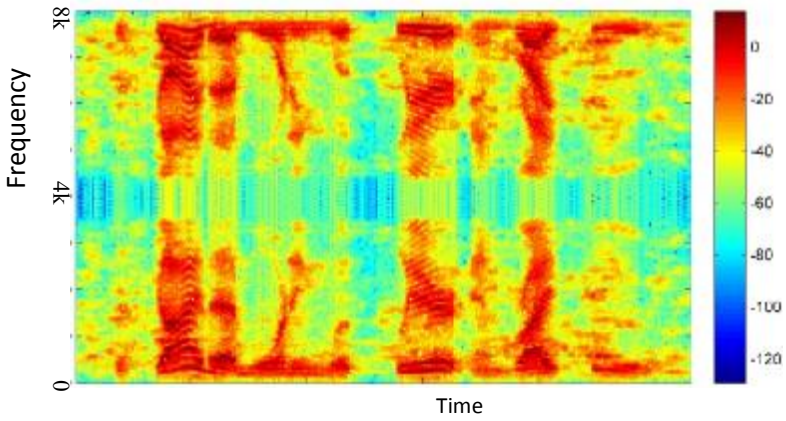
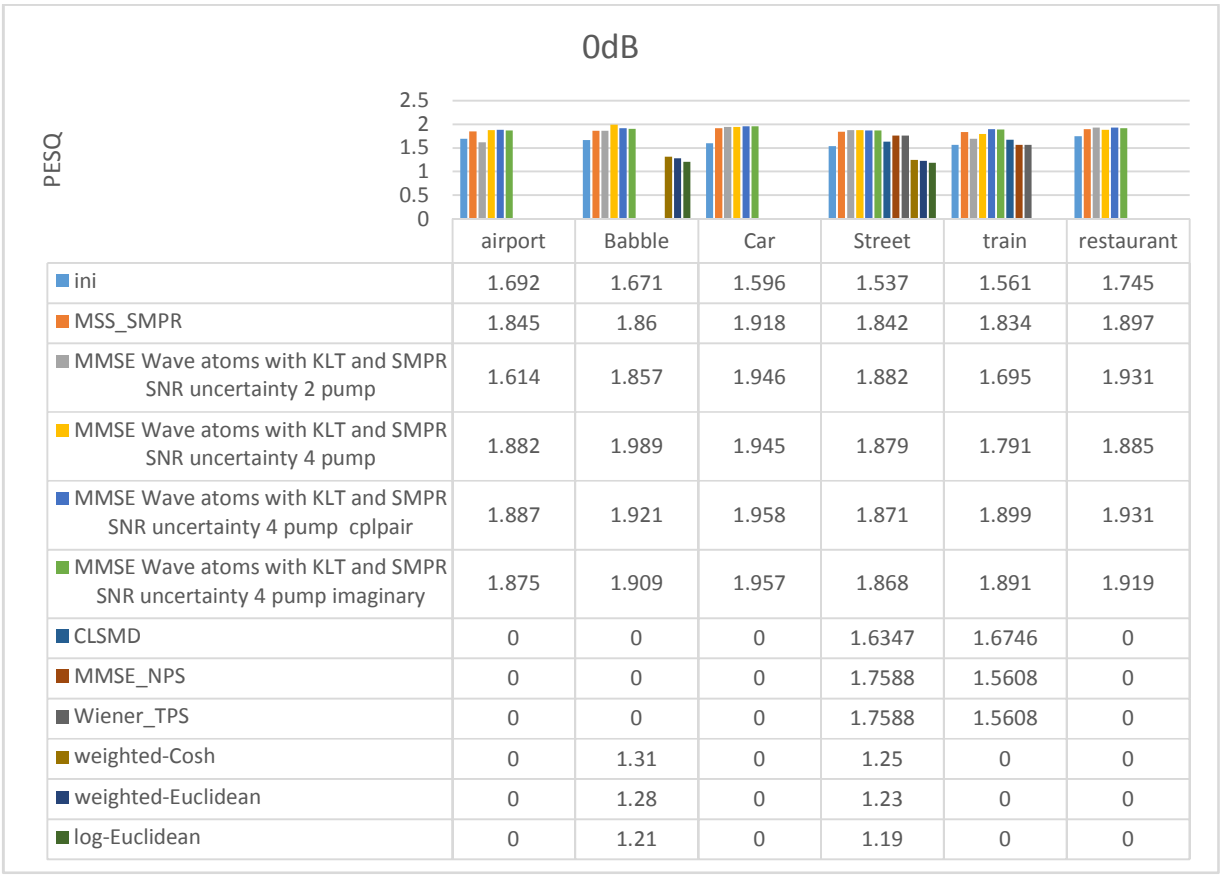


Figure 5.42 MMSE Wave atoms with KLT and SMPR SNR uncertainty quad pump Imaginary spectrogram

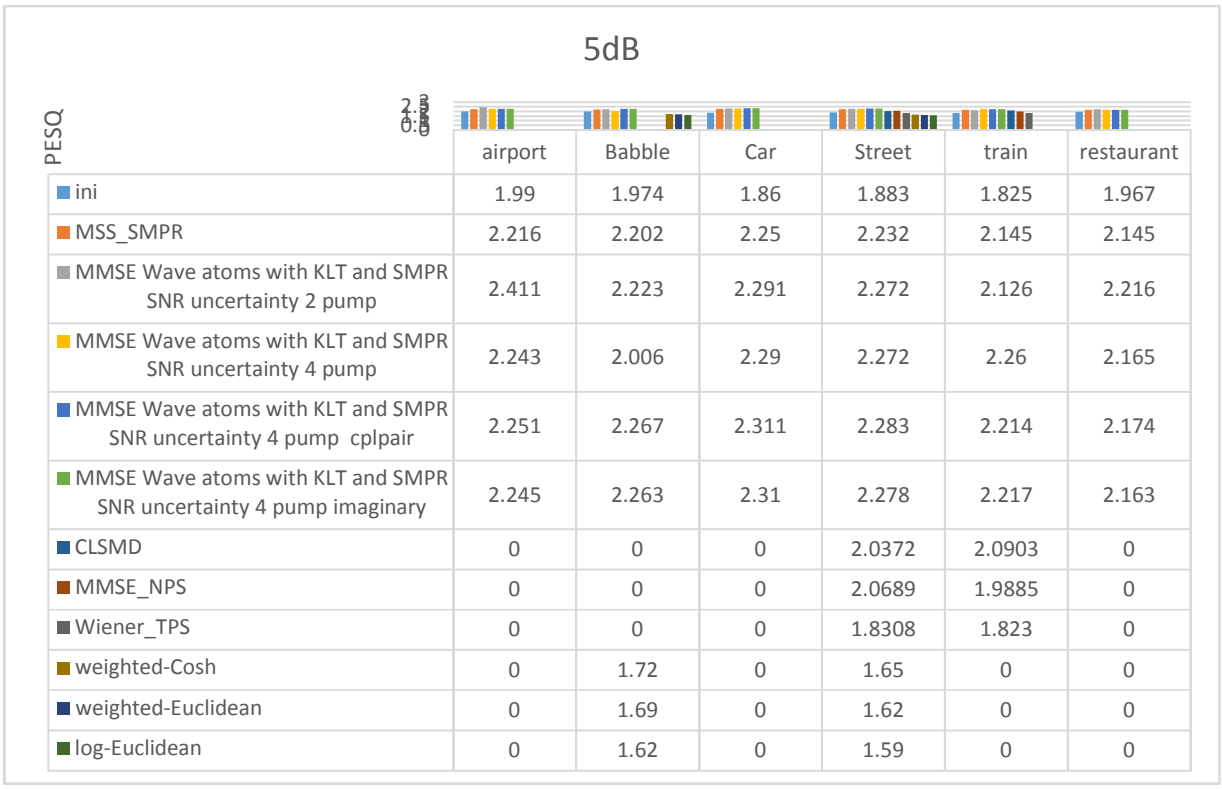
Figure 5.43, Figure 5.44, Figure 5.45 and Figure 5.46 show the comparison between our technique, MMSE-SPMR and similar technique using the same dataset the table are collected from the related search papers [41], [44]. The figures shows that MMSE-SPMR stage make more enhancement .Comparing the result MMSE-SPMR with our algorithm you can find about 3% enhancement than MMSE-SPMR. This mean that the MMSE-SPMR algorithm gives more enhancement when combined with wave atoms and KLT. Although, using the complex part only and real only of wave atoms transform give the best PESQ enhancement when combined with SPMR. we see the mount of enhancement when compare our result with result of using MMSE_MSS, CLSMD, MMSE_NPS, Winer_TPS, weighted_cosh, weighted_ Euclidian and log Euclidian. From Figure 5.43, Figure 5.44, Figure 5.45 and Figure 5.46 we can see MMSE Wave atoms with KLT and SMPR SNR uncertainty dual pump has 13.5% average enhancement and PESQ score. Compare this result with MMSE Wave atoms with KLT and SMPR SNR uncertainty dual pump we have 2% enhancement in PESQ score. Result of MMSE Wave atoms with KLT and SMPR SNR uncertainty quad pump has 16.7% average enhancement and PESQ score proportionally increased with increasing the dB with 16% enhancement in 0db and increased to 18% with 15dB. Comparing this result with MMSE Wave atoms with KLT and Spectrum Power Zero Cross SNR Uncertainty 4 pump we have 5% enhancement. Result of MMSE Wave atoms with KLT and SMPR SNR uncertainty quad pump Cplxpair has 17.5% average in all dB (0dB-15dB). Comparing this result with MMSE Wave atoms with KLT and Spectrum Power Zero Cross SNR Uncertainty quad pump Cplxpair we have about 3%

enhancement. Result of MMSE Wave atoms with KLT and SMPR SNR uncertainty quad pump Imaginary has 17% in all dB (0dB-15dB). The result has 3% enhancement than MMSE Wave atoms with KLT and Spectrum Power Zero Cross SNR Uncertainty quad pump Imaginary. We can conclude that the best enhancement has been gained with MMSE Wave atoms with KLT and SMPR SNR uncertainty quad pump Cplxpair which use complex pairs in thresholding and MMSE Wave atoms with KLT and SMPR SNR uncertainty quad pump Imaginary has similar enhancement with small difference this mean that we can use the imaginary information to enhance the signal with technique.



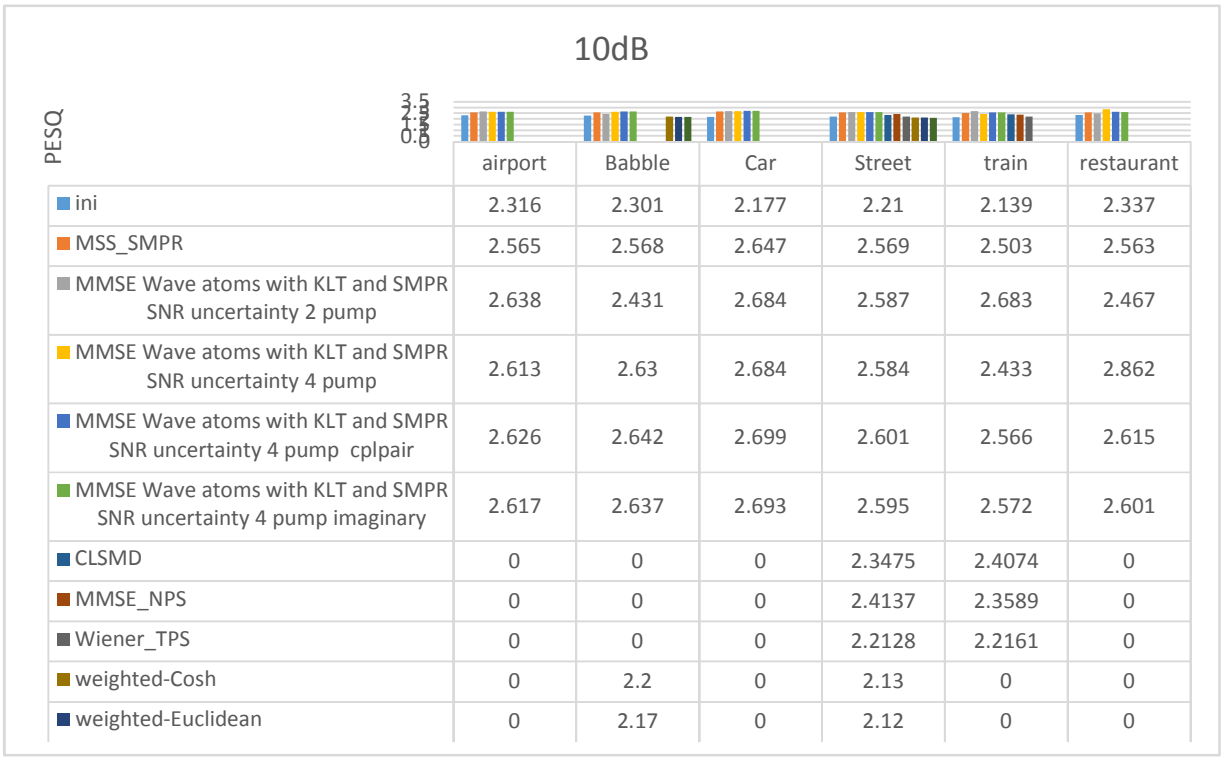
	airport	Babble	Car	Street	train	restaurant
ini	1.692	1.671	1.596	1.537	1.561	1.745
MSS_SMPR	1.845	1.86	1.918	1.842	1.834	1.897
MMSE Wave atoms with KLT and SMPR SNR uncertainty 2 pump	1.614	1.857	1.946	1.882	1.695	1.931
MMSE Wave atoms with KLT and SMPR SNR uncertainty 4 pump	1.882	1.989	1.945	1.879	1.791	1.885
MMSE Wave atoms with KLT and SMPR SNR uncertainty 4 pump cplpair	1.887	1.921	1.958	1.871	1.899	1.931
MMSE Wave atoms with KLT and SMPR SNR uncertainty 4 pump imaginary	1.875	1.909	1.957	1.868	1.891	1.919
CLSMD	0	0	0	1.6347	1.6746	0
MMSE_NPS	0	0	0	1.7588	1.5608	0
Wiener_TPS	0	0	0	1.7588	1.5608	0
weighted-Cosh	0	1.31	0	1.25	0	0
weighted-Euclidean	0	1.28	0	1.23	0	0
log-Euclidean	0	1.21	0	1.19	0	0

Figure 5.43 comparison of different algorithm corrupted with 0dB Noise and different noise type



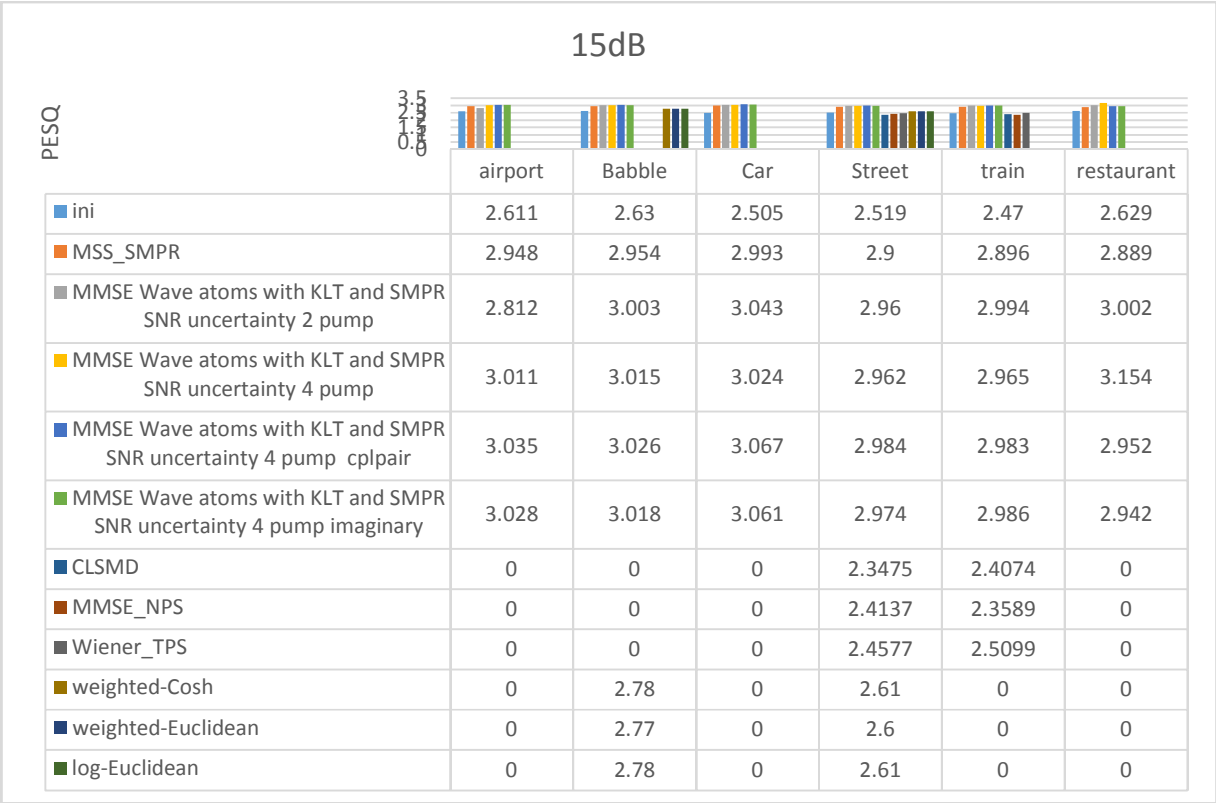
	airport	Babble	Car	Street	train	restaurant
ini	1.99	1.974	1.86	1.883	1.825	1.967
MSS_SMPR	2.216	2.202	2.25	2.232	2.145	2.145
MMSE Wave atoms with KLT and SMPR SNR uncertainty 2 pump	2.411	2.223	2.291	2.272	2.126	2.216
MMSE Wave atoms with KLT and SMPR SNR uncertainty 4 pump	2.243	2.006	2.29	2.272	2.26	2.165
MMSE Wave atoms with KLT and SMPR SNR uncertainty 4 pump cplpair	2.251	2.267	2.311	2.283	2.214	2.174
MMSE Wave atoms with KLT and SMPR SNR uncertainty 4 pump imaginary	2.245	2.263	2.31	2.278	2.217	2.163
CLSMD	0	0	0	2.0372	2.0903	0
MMSE_NPS	0	0	0	2.0689	1.9885	0
Wiener_TPS	0	0	0	1.8308	1.823	0
weighted-Cosh	0	1.72	0	1.65	0	0
weighted-Euclidean	0	1.69	0	1.62	0	0
log-Euclidean	0	1.62	0	1.59	0	0

Figure 5.44 comparison of different algorithm corrupted with 5dB Noise and different noise type



	airport	Babble	Car	Street	train	restaurant
ini	2.316	2.301	2.177	2.21	2.139	2.337
MSS_SMPR	2.565	2.568	2.647	2.569	2.503	2.563
MMSE Wave atoms with KLT and SMPR SNR uncertainty 2 pump	2.638	2.431	2.684	2.587	2.683	2.467
MMSE Wave atoms with KLT and SMPR SNR uncertainty 4 pump	2.613	2.63	2.684	2.584	2.433	2.862
MMSE Wave atoms with KLT and SMPR SNR uncertainty 4 pump cplpair	2.626	2.642	2.699	2.601	2.566	2.615
MMSE Wave atoms with KLT and SMPR SNR uncertainty 4 pump imaginary	2.617	2.637	2.693	2.595	2.572	2.601
CLSMD	0	0	0	2.3475	2.4074	0
MMSE_NPS	0	0	0	2.4137	2.3589	0
Wiener_TPS	0	0	0	2.2128	2.2161	0
weighted-Cosh	0	2.2	0	2.13	0	0
weighted-Euclidean	0	2.17	0	2.12	0	0

Figure 5.45 comparison of different algorithm corrupted with 10dB Noise and different noise type



	airport	Babble	Car	Street	train	restaurant
ini	2.611	2.63	2.505	2.519	2.47	2.629
MSS_SMPR	2.948	2.954	2.993	2.9	2.896	2.889
MMSE Wave atoms with KLT and SMPR SNR uncertainty 2 pump	2.812	3.003	3.043	2.96	2.994	3.002
MMSE Wave atoms with KLT and SMPR SNR uncertainty 4 pump	3.011	3.015	3.024	2.962	2.965	3.154
MMSE Wave atoms with KLT and SMPR SNR uncertainty 4 pump cplpair	3.035	3.026	3.067	2.984	2.983	2.952
MMSE Wave atoms with KLT and SMPR SNR uncertainty 4 pump imaginary	3.028	3.018	3.061	2.974	2.986	2.942
CLSMD	0	0	0	2.3475	2.4074	0
MMSE_NPS	0	0	0	2.4137	2.3589	0
Wiener_TPS	0	0	0	2.4577	2.5099	0
weighted-Cosh	0	2.78	0	2.61	0	0
weighted-Euclidean	0	2.77	0	2.6	0	0
log-Euclidean	0	2.78	0	2.61	0	0

Figure 5.46 comparison of different algorithm corrupted with 15dB Noise and different noise type

Algorithm 7. Wave atoms with Soft Masking Based on Posteriori SNR Uncertainty

1. MMSE Wave atoms with KLT and SMPO SNR uncertainty dual pump

Figure 5.47 show the spectrogram of the input signal degraded with 10 dB babble noise after applying the algorithm. Table 5.24 show the PESQ result and SNR of this technique. Comparing this result in Table 5.24 with Table 5.20 we can find that MMSE Wave atoms with KLT and SMPR SNR uncertainty dual pump achieve better enhancement than MMSE Wave atoms with KLT and SMPO SNR uncertainty dual pump with 0.5%.

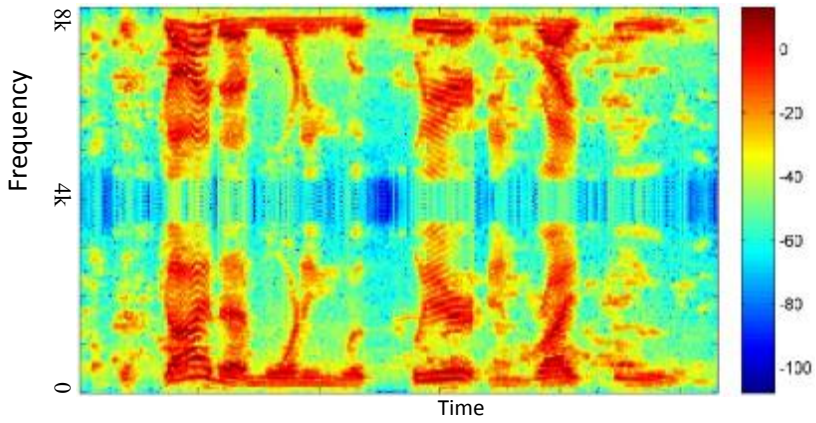


Figure 5.47 MMSE Wave atoms with KLT and SMPO SNR uncertainty dual pump spectrogram

Table 5.24 MMSE Wave atoms with KLT and SMPO SNR uncertainty dual pump PESQ and Seg_SNR result

	0dB		5dB		10dB		15dB	
	PESQ	Seg_SNR	PESQ	Seg_SNR	PESQ	Seg_SNR	PESQ	Seg_SNR
airport	1.873	-1.245	2.133	0.541	2.556	2.877	2.955	4.86
Babble	1.899	-1.506	2.33	0.51	2.65	2.638	2.978	4.901
Car	1.596	-0.271	2.294	1.625	2.686	3.31	3.036	5.699
Street	1.834	-1.118	2.205	1.061	2.504	2.898	2.9	4.535
train	1.714	-0.901	2.105	1.153	2.585	3.384	2.829	4.864
restaurant	1.973	-2.073	2.323	0.489	2.686	2.783	2.977	3.976

2. MMSE Wave atoms with KLT and SMPO SNR uncertainty quad pump

Figure 5.48 show the spectrogram of the input signal degraded with 10 dB babble noise after applying the algorithm. Table 5.25 show the PESQ and Seg_SNR result of using this technique. Comparing the result of pump MMSE Wave atoms with KLT and SMPR SNR uncertainty quad pump in Table 5.21 and MMSE Wave atoms with KLT and SMPO SNR uncertainty quad pump in Table 5.25 we can find that MMSE Wave atoms with KLT and SMPR SNR uncertainty quad pump has 4% better enhancement than MMSE Wave atoms with KLT and SMPO SNR uncertainty quad pump.

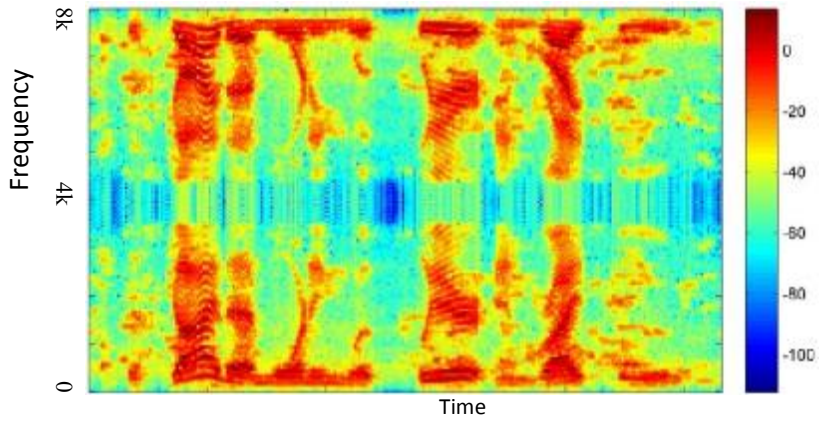


Figure 5.48 MMSE Wave atoms with KLT and SMPO SNR uncertainty quad pump spectrogram

Table 5.25 MMSE Wave atoms with KLT and SMPO SNR uncertainty quad pump PESQ and Seg_SNR result

	0dB		5dB		10dB		15dB	
	PESQ	Seg_SNR	PESQ	Seg_SNR	PESQ	Seg_SNR	PESQ	Seg_SNR
airport	1.807	-1.543	2.208	0.831	2.568	3.152	2.967	5.277
Babble	1.778	-1.774	2.207	0.649	2.506	2.918	2.981	4.959
Car	1.918	-0.059	2.075	2.012	2.665	4.028	3.055	5.934
Street	2.769	-1.154	2.175	1.192	2.454	3.12	2.916	4.805
train	1.809	-0.936	2.188	1.323	2.524	3.417	2.922	5.15
restaurant	1.791	-2.05	2.137	0.244	2.488	3.225	2.843	5.043

3. MMSE Wave atoms with KLT and SMPO SNR uncertainty quad pump Cplxpair

Figure 5.49 show the spectrogram of the input signal degraded with 10 dB babble noise after applying the algorithm. Table 5.26 show the result of PESQ and Seg_SNR of this technique. Comparing the result in Table 5.26 with Table 5.22 we can find that MMSE Wave Atoms with Spectrum Power Zero Cross quad pump Cplxpair has better performance with 4%.

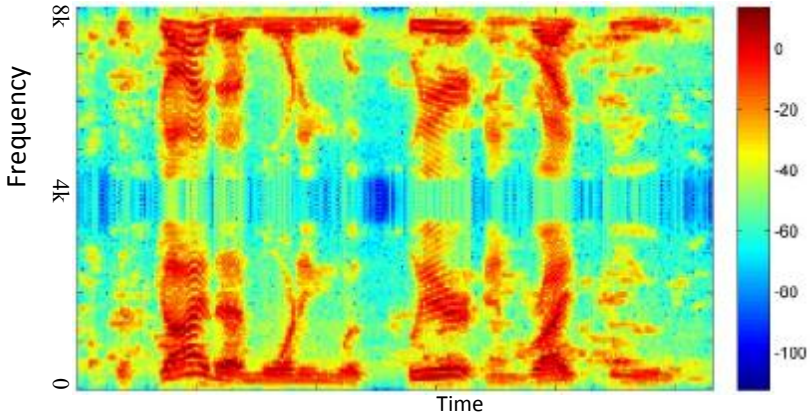


Figure 5.49 MMSE Wave atoms with KLT and SMPO SNR uncertainty quad pump Cplpair spectrogram

Table 5.26 MMSE Wave atoms with KLT and SMPO SNR uncertainty quad pump Cplpair PESQ and Seg_SNR result

	0dB		5dB		10dB		15dB	
	PESQ	Seg_SNR	PESQ	Seg_SNR	PESQ	Seg_SNR	PESQ	Seg_SNR
airport	1.805	-1.481	2.169	0.622	2.577	2.556	2.996	4.19
Babble	1.77	-1.65	2.188	0.511	2.576	2.417	2.959	3.491
Car	1.596	-0.23	2.299	1.588	2.701	3.191	3.071	4.614
Street	1.811	-1.053	2.208	0.93	2.511	2.788	2.92	3.614
train	1.8	-0.848	2.166	1.035	2.527	2.919	2.951	4.036
restaurant	1.792	-1.931	2.1	0.347	2.533	2.503	2.889	3.015

4. MMSE Wave atoms with KLT and SMPO SNR uncertainty quad pump Imaginary

Figure 5.50 show the spectrogram of the input signal degraded with 10 dB babble noise after applying the algorithm. Table 5.27 show the result of PESQ and Seg_SNR of this technique. Comparing the result in Table 5.27 with MMSE Wave atoms with KLT and SMPR SNR uncertainty quad pump Imaginary Table 5.23 we can find that MMSE Wave atoms with KLT and SMPR SNR uncertainty quad pump Imaginary has better performance with 4%.

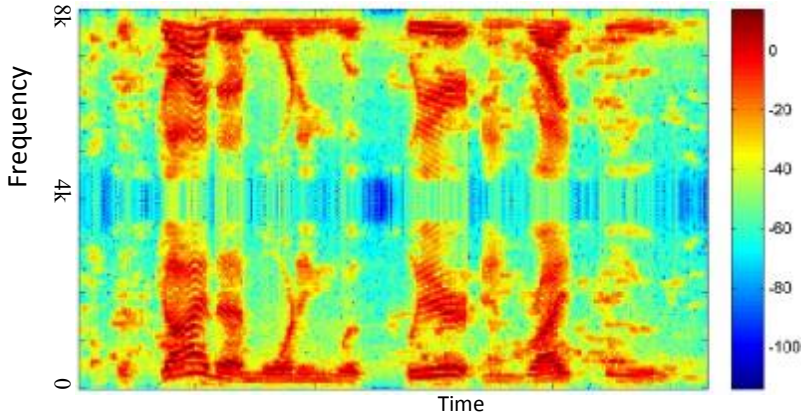


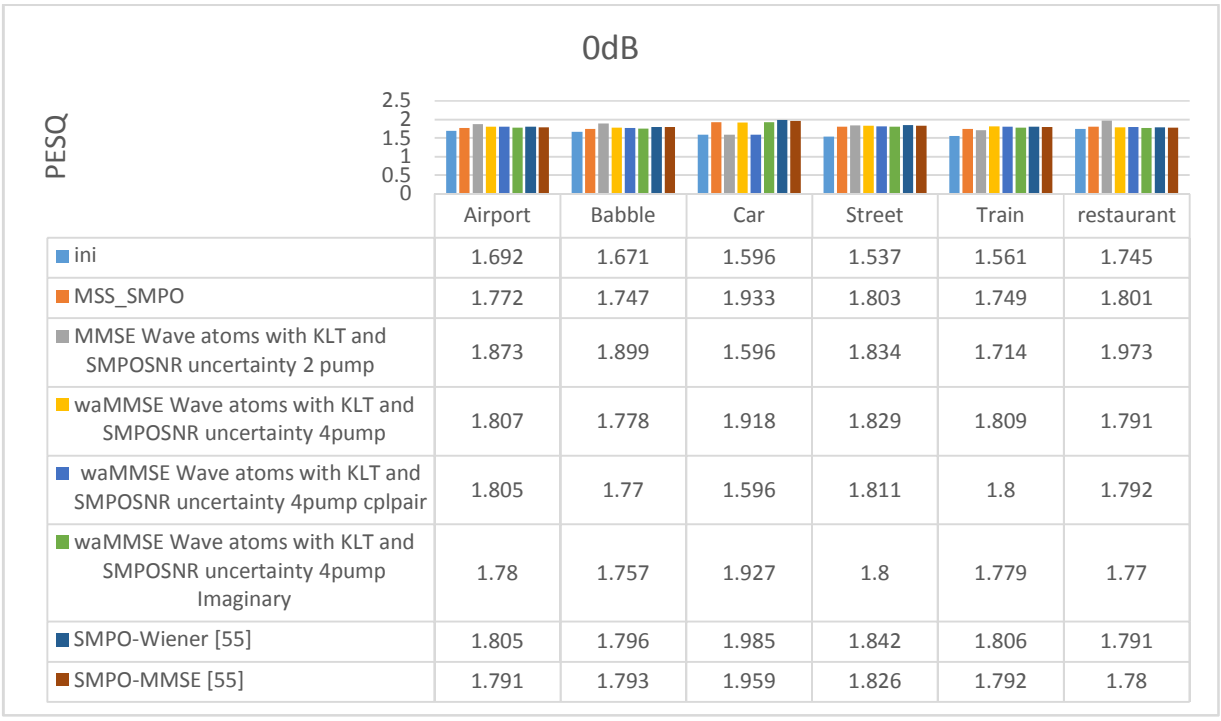
Figure 5.50 MMSE Wave atoms with KLT and SMPO SNR uncertainty quad pump Imaginary spectrogram

Table 5.27 MMSE Wave atoms with KLT and SMPO SNR uncertainty quad pump Imaginary PESQ and Seg_SNR result

	0dB		**5dB**		**10dB**		**15dB**	
	PESQ	Seg_SNR	PESQ	Seg_SNR	PESQ	Seg_SNR	PESQ	Seg_SNR
airport	1.78	-1.464	2.153	1.005	2.558	3.453	2.984	5.818
Babble	1.757	-1.656	2.169	0.859	2.561	3.345	2.948	5.443
Car	1.927	-0.05	2.274	2.109	2.678	4.245	3.053	6.269
Street	1.8	-0.975	2.191	1.356	2.495	3.792	2.899	5.452
train	1.779	-0.752	2.151	1.581	2.515	3.941	2.941	5.855
restaurant	1.77	-1.981	2.082	0.661	2.519	3.313	2.871	5.394

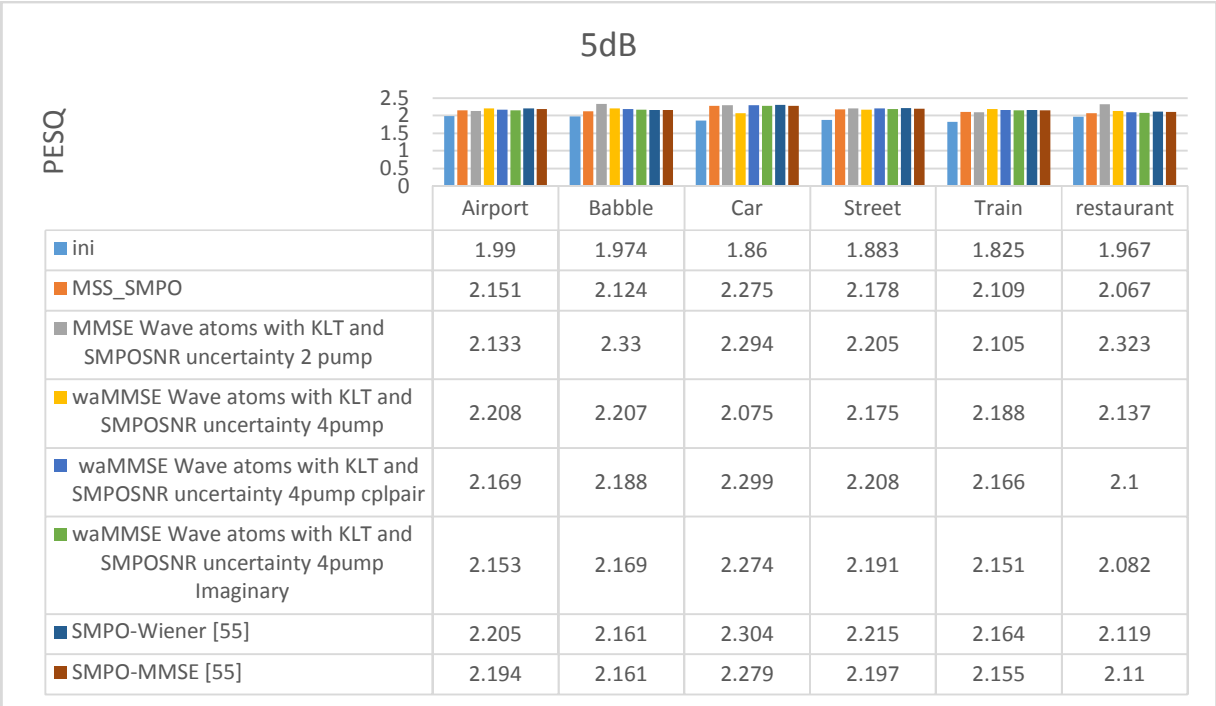
Figure 5.51, Figure 5.52, Figure 5.53 and Figure 5.54 show the comparison between our technique and similar technique using the same dataset the table are collected from the related search papers [45]. The figures shows that MMSE-SMPO stage make more enhancement .Comparing the result MMSE-SMPO with MMSE Wave atoms with KLT and SMPO SNR uncertainty dual pump algorithm you can find about 3% enhancement than MMSE- SMPO. This mean that the SMPO algorithm gives more enhancement when combined with wave atoms and KLT. we see that a good mount of enhancement when compare our result with result of using SMPO _MMSE and SMPO_wiener[45]. Result of MMSE Wave atoms with KLT and SMPO SNR uncertainty dual pump has 14.75% average enhancement and PESQ score Compare this result with MMSE Wave atoms with KLT and SMPR SNR uncertainty dual pump we can see MMSE Wave atoms with KLT and SMPR SNR uncertainty dual pump have 0.5% enhancement in PESQ score than

MMSE Wave atoms with KLT and SMPO SNR uncertainty . Result of MMSE Wave atoms with KLT and SMPO SNR uncertainty quad pump has 15.8% average enhancement. Comparing this result with MMSE Wave atoms with KLT and SMPR SNR uncertainty 4 pump we have 1% more enhancement than it. Result of MMSE Wave atoms with KLT and SMPO SNR uncertainty quad pump Cplpair has 13% and PESQ score proportionally increased with increasing the dB with 8% enhancement in 0db and increased to 16% with 15dB. Comparing this result with MMSE Wave atoms with KLT and SMPR SNR uncertainty quad pump Cplpair we have 5% more enhancement than it. Result of MMSE Wave atoms with KLT and SMPO SNR uncertainty quad pump Imaginary has 13% in all dB (0dB-15dB). The result has 5% less enhancement than MMSE Wave atoms with KLT and SMPR SNR uncertainty quad pump Imaginary we have. We can conclude that the best enhancement has been gained with MMSE Wave atoms with KLT and SMPO SNR uncertainty 4 pump .



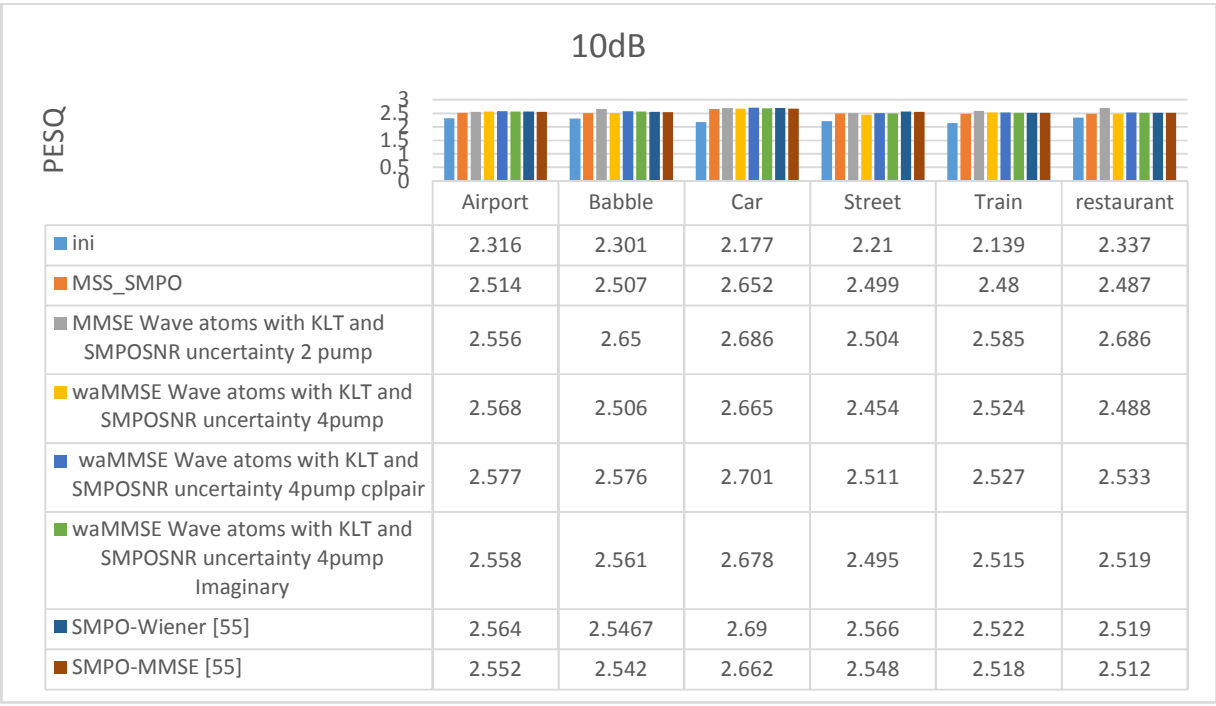
	Airport	Babble	Car	Street	Train	restaurant
ini	1.692	1.671	1.596	1.537	1.561	1.745
MSS_SMPO	1.772	1.747	1.933	1.803	1.749	1.801
MMSE Wave atoms with KLT and SMPOSNR uncertainty 2 pump	1.873	1.899	1.596	1.834	1.714	1.973
waMMSE Wave atoms with KLT and SMPOSNR uncertainty 4pump	1.807	1.778	1.918	1.829	1.809	1.791
waMMSE Wave atoms with KLT and SMPOSNR uncertainty 4pump cplpair	1.805	1.77	1.596	1.811	1.8	1.792
waMMSE Wave atoms with KLT and SMPOSNR uncertainty 4pump Imaginary	1.78	1.757	1.927	1.8	1.779	1.77
SMPO-Wiener [55]	1.805	1.796	1.985	1.842	1.806	1.791
SMPO-MMSE [55]	1.791	1.793	1.959	1.826	1.792	1.78

Figure 5.51 comparison of different algorithm corrupted with 0dB Noise and different noise type



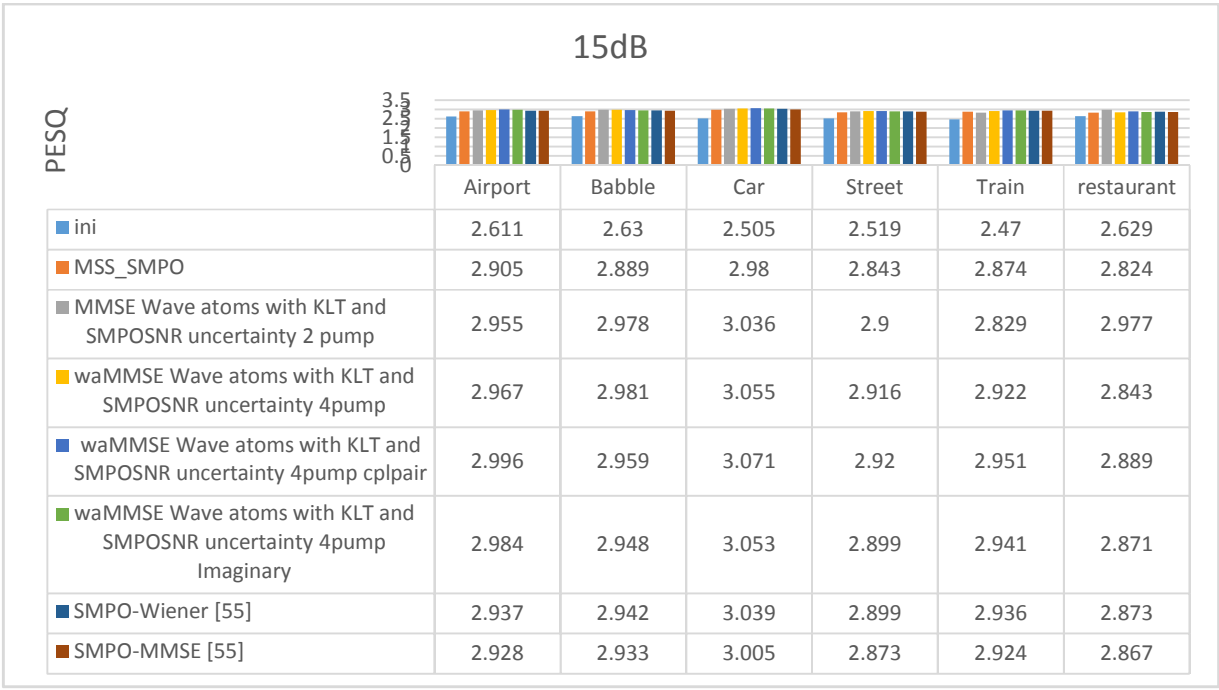
	Airport	Babble	Car	Street	Train	restaurant
ini	1.99	1.974	1.86	1.883	1.825	1.967
MSS_SMPO	2.151	2.124	2.275	2.178	2.109	2.067
MMSE Wave atoms with KLT and SMPOSNR uncertainty 2 pump	2.133	2.33	2.294	2.205	2.105	2.323
waMMSE Wave atoms with KLT and SMPOSNR uncertainty 4pump	2.208	2.207	2.075	2.175	2.188	2.137
waMMSE Wave atoms with KLT and SMPOSNR uncertainty 4pump cplpair	2.169	2.188	2.299	2.208	2.166	2.1
waMMSE Wave atoms with KLT and SMPOSNR uncertainty 4pump Imaginary	2.153	2.169	2.274	2.191	2.151	2.082
SMPO-Wiener [55]	2.205	2.161	2.304	2.215	2.164	2.119
SMPO-MMSE [55]	2.194	2.161	2.279	2.197	2.155	2.11

Figure 5.52 comparison of different algorithm corrupted with 5dB Noise and different noise type



	Airport	Babble	Car	Street	Train	restaurant
ini	2.316	2.301	2.177	2.21	2.139	2.337
MSS_SMPO	2.514	2.507	2.652	2.499	2.48	2.487
MMSE Wave atoms with KLT and SMPOSNR uncertainty 2 pump	2.556	2.65	2.686	2.504	2.585	2.686
waMMSE Wave atoms with KLT and SMPOSNR uncertainty 4pump	2.568	2.506	2.665	2.454	2.524	2.488
waMMSE Wave atoms with KLT and SMPOSNR uncertainty 4pump cplpair	2.577	2.576	2.701	2.511	2.527	2.533
waMMSE Wave atoms with KLT and SMPOSNR uncertainty 4pump Imaginary	2.558	2.561	2.678	2.495	2.515	2.519
SMPO-Wiener [55]	2.564	2.5467	2.69	2.566	2.522	2.519
SMPO-MMSE [55]	2.552	2.542	2.662	2.548	2.518	2.512

Figure 5.53 comparison of different algorithm corrupted with 10dB Noise and different noise type



	Airport	Babble	Car	Street	Train	restaurant
ini	2.611	2.63	2.505	2.519	2.47	2.629
MSS_SMPO	2.905	2.889	2.98	2.843	2.874	2.824
MMSE Wave atoms with KLT and SMPOSNR uncertainty 2 pump	2.955	2.978	3.036	2.9	2.829	2.977
waMMSE Wave atoms with KLT and SMPOSNR uncertainty 4pump	2.967	2.981	3.055	2.916	2.922	2.843
waMMSE Wave atoms with KLT and SMPOSNR uncertainty 4pump cplpair	2.996	2.959	3.071	2.92	2.951	2.889
waMMSE Wave atoms with KLT and SMPOSNR uncertainty 4pump Imaginary	2.984	2.948	3.053	2.899	2.941	2.871
SMPO-Wiener [55]	2.937	2.942	3.039	2.899	2.936	2.873
SMPO-MMSE [55]	2.928	2.933	3.005	2.873	2.924	2.867

Figure 5.54 comparison of different algorithm corrupted with 15dB Noise and different noise type

Chapter 6

6 Conclusion

This Chapter summarizes the thesis, discusses its findings and contributions, points to limitations of the current work, and also outlines directions for future research

6.1 Summary and Conclusion Remarks

This thesis deals with the enhancement of speech signals that have been subject to acoustic background noise. This study focus on enhancing the degrade speech signal in the spectral domain and wave atoms domain. The he researcher use estimation approach to exploit prior knowledge about the speech and noise signals are developed using different statistical probability model like Maximum a Posteriori Wave atoms algorithm, Prior probability and Posteriori probability with speech presence probability. The study use wave atoms transform as main enhancement technique and combine it with other transformation technique like Karhunen–Loeve transform and use different statistical probability model. The result shows that wave atoms is making an enhancement comparing it with some new research papers. The result using uncertainty technique shows increasing in the performance of the algorithm. Also, combining the probability models with the algorithm increases the performance. This study make four type of threshold the best enhancement is achieved when use the complex component of the values. A comparable result achieved imaginary part value only. We can arrange the algorithm respecting to PESQ measurement from best to worst as follow:

- MMSE Wave atoms with KLT and SMPR SNR uncertainty quad pump Cplpair has 17.5%
- MMSE Wave atoms with KLT and SMPR SNR uncertainty quad pump Imaginary has 17%
- MMSE Wave atoms with KLT and SMPR SNR uncertainty quad pump has 16.7%
- MMSE Wave atoms with KLT and SMPO SNR uncertainty 4 pump has 15.8%
- MMSE Wave atoms with KLT and SMPO SNR uncertainty 2 pump has 14.75%
- MMSE Wave atoms with KLT and Spectrum Power Zero Cross SNR Uncertainty quad pump Cplxpair has 14.5%

- MMSE Wave atoms with KLT and Spectrum Power Zero Cross SNR Uncertainty quad pump Imaginary has 14%
- MMSE Wave atoms with KLT and SMPR SNR uncertainty dual pump has 13.5%
- MMSE Wave atoms with KLT and Spectrum Power Zero Cross SNR Uncertainty 2 pump has 13.5%
- MMSE Wave atoms with KLT and SMPO SNR uncertainty quad pump Cplxpair has 13%
- MMSE Wave atoms with KLT and SMPO SNR uncertainty quad pump Imaginary has 13%
- MMSE Wave Atoms with Spectrum Power Zero Cross Cplxpair has 12%
- MMSE Wave atoms with KLT and Spectrum Power Zero Cross SNR Uncertainty 4 pump has 11%
- MMSE Wave atoms with KLT and Spectrum Power Zero Cross 2 pump has 11%
- of MMSE Wave Atoms with Spectrum Power Zero Cross quad pump Magnitude has 11%
- MMSE Wave Atoms with Spectrum Power Zero Cross Imaginary has 11%
- Wave Atoms KLT 4 pump has 5%
- Wave Atoms 4 pump we have about 3.5%
- Wave Atoms quad pump Cplxpair 2%
- Wave Atoms KLT 2 pump does not make enhancement
- Wave Atoms 2 pump does not make enhancement

6.2 Future Work

The results of this thesis point to several interesting directions for future work, which should be addressed and further developed to acquire better results with less cost, these points include the following:

- Using the hole complex value instead of using the magnitude of the result of FFT
- Using optimization technique to find the threshold process
- Apply the wave atoms techniques as feature extraction for speech recognition

7 References

[1] S. Greenberg, W. Ainsworth, A. Popper, and R. Fay, *Speech processing in the auditory system*. Springer Handbook of Auditory Research, Vol. 18, 2004.

[2] J. Benesty, J. Chen, and E. Habets, *Speech Enhancement in the STFT Domain*. SpringerBriefs in Electrical and Computer Engineering, 2012.

[3] M. H. Farouk, *Application of Wavelets in Speech Processing*. Cham: Springer International Publishing, 2014.

[4] P. Loizou, *Speech Enhancement : Theory and Practice, Second Edition*, 2nd ed. CRC Press, 2013.

[5] S. Selouani, *Speech processing and soft computing*. SpringerBriefs in Electrical and Computer Engineering, 2011.

[6] A. Abolhassani and S. Selouani, "Speech enhancement using PCA and variance of the reconstruction error model identification.," …, 2007.

[7] K. Paliwal, B. Schwerin, and K. Wójcicki, "Role of modulation magnitude and phase spectrum towards speech intelligibility," *Speech Commun.*, vol. 53, no. 3, pp. 327–339, Mar. 2011.

[8] N. Upadhyay and A. Karmakar, "A Perceptually Motivated Multi-Band Spectral Subtraction Algorithm for Enhancement of Degraded Speech," *2012 Third Int. Conf. Comput. Commun. Technol.*, no. 1, pp. 340–345, Nov. 2012.

[9] E. MEHMETCIK, "SPEECH ENHANCEMENT UTILIZING PHASE CONTINUITY BETWEEN CONSECUTIVE ANALYSIS WINDOWS," Middle East Technical University, 2011.

[10] R. C. Hendriks and R. H. Tudelft, "MMSE BASED NOISE PSD TRACKING WITH LOWCOMPLEXITY," vol. 0, no. 1, pp. 4266–4269, 2010.

[11] E. Loveimi, S. Member, S. M. Ahadi, and S. Member, "Objective Evaluation of Magnitude and Phase Only Spectrum-," no. March, pp. 3–5, 2010.

[12] P. Fardkhaleghi and M. Savoji, "New approaches to speech enhancement using phase correction in Wiener filtering," *Telecommun. (IST), 2010 …*, pp. 895–899, 2010.

[13] S. Senapati, "Optimal speech enhancement under signal presence uncertainty using Log Gabor Wavelet and Bayesian Joint Statistics," *Int. J. Speech Technol.*, vol. 16, no. 4, pp. 439–459, Apr. 2013.

[14] C. V. Rama Rao, M. B. Rama Murthy, and K. Srinivasa Rao, "Speech enhancement using sub-band cross-correlation compensated Wiener filter combined with harmonic regeneration," *AEU - Int. J. Electron. Commun.*, vol. 66, no. 6, pp. 459–464, Jun. 2012.

[15] M. McCallum and B. Guillemin, “Accounting for deterministic noise components in a MMSE STSA speech enhancement framework,” *2012 Int. Symp. Commun. Inf. Technol.*, pp. 169–174, Oct. 2012.

[16] S. Senapati, N. Bhende, and G. Saha, “Bayesian marginal statistics for speech enhancement using log Gabor wavelet,” *Int. J. Speech Technol.*, vol. 14, no. 3, pp. 193–210, Jul. 2011.

[17] B. Schwerin and K. Paliwal, “Using STFT real and imaginary parts of modulation signals for MMSE-based speech enhancement,” *SPEECH Commun.*, vol. 58, pp. 49–68, 2014.

[18] P. Krishnamoorthy and S. R. M. Prasanna, “Enhancement of noisy speech by temporal and spectral processing,” *Speech Commun.*, vol. 53, no. 2, pp. 154–174, Feb. 2011.

[19] T. F. Sanam and C. Shahnaz, “Enhancement of noisy speech based on a custom thresholding function with a statistically determined threshold,” *Int. J. Speech Technol.*, vol. 15, no. 4, pp. 463–475, Apr. 2012.

[20] T. Lotter and P. Vary, “Speech enhancement by MAP spectral amplitude estimation using a super-Gaussian speech model,” *EURASIP J. Appl. Signal Processing*, pp. 1110–1126, 2005.

[21] A. P. Stark, K. W. Kamil, J. G. Lyons, and K. K. Paliwal, “Noise Driven Short-Time Phase Spectrum Compensation Procedure for Speech Enhancement,” no. 1, pp. 549–552, 2008.

[22] A. R. Fukane and S. L. Sahare, “Noise estimation Algorithms for Speech Enhancement in highly non-stationary Environments,” vol. 8, no. 2, pp. 39–44, 2011.

[23] I. Cohen and B. Berdugo, “Noise estimation by minima controlled recursive averaging for robust speech enhancement,” *IEEE Signal Process. Lett.*, vol. 9, no. 1, pp. 12–15, Jan. 2002.

[24] J. Kum and J. Chang, “Speech Enhancement Based on Minima Controlled Recursive Averaging Incorporating Second-Order,” vol. 16, no. 7, pp. 624–627, 2009.

[25] C. Maccone, *Deep Space Flight and Communications*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009.

[26] J. Benesty, J. Chen, Y. A. Huang, B. Labs, and M. Hill, “On Noise Reduction In The Karhunen-Lo Eve Expansion Domain,” no. 10, pp. 25–28, 2009.

[27] I. Bocharova, *Compression for multimedia*. New York: Cambridge University Press, 2010.

[28] C. Maccone, *Mathematical SETI: Statistics, Signal Processing, Space Missions*. Springer Praxis Books, Signal Processing, Space Missions, 2012.

[29] A. Beck, A. Ben-tal, and Y. C. Eldar, “Robust Mean-Squared Error Estimation of Multiple Signals,” vol. 187, pp. 155–187, 2006.

[30] I. Cohen, J. Benesty, and S. Gannot, *Speech processing in modern communication*. Springer Topics in Signal Processing, Vol. 3, 2010.

[31] H. Kobayashi, B. L. Mark, and W. Turin, *Probability, Random Processes, and Statistical Analysis*. Cambridge: Cambridge University Press, 2011.

[32] J. Chen, I. Cohen, and Y. Huang, *Noise reduction in speech processing*. Springer Topics in Signal Processing, Vol. 2, 2009.

[33] D. Shete and S. Patil, "Zero crossing rate and Energy of the Speech Signal of Devanagari Script," *iosrjournals.org*, vol. 4, no. 1, pp. 1–5, 2014.

[34] I. C. N. Westerlund, M. Dahl, "Speech enhancement by non-stationary tonal disturbance cancellation using subband zero crossing measures," *… , Video Speech …*, pp. 20–23, 2004.

[35] Y. Lu and P. C. Loizou, "Estimators of The Magnitude-Squared Spectrum and Methods for Incorporating SNR Uncertainty.," *IEEE Trans. Audio. Speech. Lang. Processing*, vol. 19, no. 5, pp. 1123–1137, Jul. 2011.

[36] R. Yu, "Speech enhancement based on soft audible noise masking and noise power estimation q," *Speech Commun.*, vol. 55, no. 10, pp. 964–974, 2013.

[37] Y. Hu and P. C. Loizou, "Subjective comparison and evaluation of speech enhancement algorithms.," *Speech Commun.*, vol. 49, no. 7, pp. 588–601, Jul. 2007.

[38] H. Hirsch and D. Pearce, "The Aurora experimental framework for the performance evaluation of speech recognition systems under noisy conditions," *… Speech Recognit. Challenges …*, 2000.

[39] P. Limited and U. Kingdom, "PESQ : An Introduction White Paper," vol. 44, no. September, 2001.

[40] T. Manjunath, "Limitations of perceptual evaluation of speech quality on VoIP systems," in *… Multimedia Systems and Broadcasting, 2009. …*.

[41] J. Erkelens, J. Jensen, and R. Heusdens, "A data-driven approach to optimizing spectral speech enhancement methods for various error criteria," *Speech Commun.*, vol. 49, no. 7–8, pp. 530–541, Jul. 2007.

[42] Y. Ghanbari and M. Karami-Mollaei, "A new approach for speech enhancement based on the adaptive thresholding of the wavelet packets," *Speech Commun.*, vol. 48, no. 8, 2006.

[43] S. Senapati, S. Chakroborty, and G. Saha, "Speech enhancement by joint statistical characterization in the Log Gabor Wavelet domain," *Speech Commun.*, vol. 50, no. 6, 2008.

[44] H. Zhao, Z. Lu, F. Yu, and C. Xu, "A Novel Magnitude-Squared Spectrum Cost Function for Speech Enhancement.," *Electron. Electr. …*, vol. 6, no. 6, pp. 0–3, 2012.

[45] H. Zhao, J. Liu, Z. Chen, and F. Wang, “A New Soft Masking Method for Speech Enhancement in the Frequency Domain.,” *Electron. Electr. …*, no. 61173106, pp. 58–63, 2014.