Signal Enhancement Methods Based on Wavelet Transform, Fractional Fourier Transform and Short-time Fourier Transform

Xindan Zhang, Haoyuan Zheng *

School of Automation, Fuzhou University, Fuzhou, China

* Corresponding Author Email: HAOYUAN.ZHENG.2022@MUMAIL.IE

Abstract. This review paper mainly focuses on different signal enhancement methods such as wavelet transform, fractional Fourier transform (FrFT) and short-time Fourier transform (STFT). First, this paper introduces the concept and importance of signal enhancement, as well as some current issues and challenges. Then, in recent years, the application of the three methods in wavelet transform, fractional Fourier transform and short-time Fourier transform is described. In terms of wavelet transform, this paper discusses the application of wavelet function to image enhancement and the specific steps, and analyzes its advantages and application in signal enhancement. In terms of fractional Fourier transform, this paper introduces its difference from traditional Fourier transform, and discusses its application in combination with adaptive filtering technology and the application of multi-level FrFT in the field of speech enhancement. Finally, in terms of short-time Fourier transform, this paper discusses its application in the fields of image enhancement and speech enhancement. The review paper finally discusses the characteristics, advantages and limitations of these three methods in signal enhancement, and looks forward to the future research directions. Through the research of this review paper, it can provide some reference and guidance for the research in the field of signal enhancement.

Keywords: Signal enhancement, Wavelet transform, Fractional Fourier Transform (FrFT), Short-time Fourier transform (STFT).

1. Introduction

Signal enhancement is a very critical problem. Especially for communication systems and control systems, signal quality is a key factor to optimize system performance. However, the existence of various interference sources in the system often seriously affects the quality and effectiveness of the feedback signal. For example, in the process of collecting sound signals in humanoid robots, since the robot's head is equipped with a signal processing unit cooled by a fan, in addition to the required signals, the head microphone also records unnecessary self-noise [1]. Interference from self-noise will inevitably lead to a significant decline in the position and extraction and other performance of the sound source [1].

Therefore, for such problems, researchers have been working on enhancing the performance of feedback signals by reducing interference. Filter, amplify, and denoise the sensor signal to improve the quality and reliability of the signal. For example, using signal filtering to filter out noise and improve the signal-to-noise ratio is a common method for signal enhancement. However, conventional filters may attenuate the power of the signal of interest during filtering [2]. Therefore, although some traditional methods have achieved certain results, more innovations and improvements are still needed to deal with the changing and growing sources of interference.

This paper summarizes the current research progress and methods in this field, and provides new ideas and directions for follow-up research. This paper will review three commonly used signal processing techniques, namely wavelet transform, fractional Fourier transform, and short-time Fourier transform. Analyze how these three technologies deal with current common sources of interference, such as color noise, noise, and electromagnetic interference in images. This paper will also analyze the applicability of these methods to different application domains and evaluate their impact on improving the quality of feedback signals. Finally, this paper discusses current research challenges and potential future directions. Through extensive exploration of interference reduction and feedback signal enhancement methodologies, it is anticipated that more efficacious solutions can

be unveiled. These endeavors are poised to foster advancements in performance and technological innovation within feedback systems, spanning diverse domains.

To sum up, this paper will give readers an overview of the methods used to enhance feedback signals through wavelet transform, fractional Fourier transform and short-time Fourier transform in recent years, aiming to promote further research and application in related fields.

2. Wavelet transform

2.1. Discrete Wavelet transform (DWT)

The Discrete Wavelet Transform (DWT) stands as a pivotal temporal-frequency analysis technique, wielding significant influence within the realm of signal enhancement. DWT, by virtue of its capability to partition the original signal into constituent sub-signals spanning varied scales and frequency bands, adeptly extracts and amplifies pertinent information dwelling within the signal. When considering image enhancement, the application of DWT extends to denoising, augmenting image contrast, and eliciting edge features, among others.

Illustrating with satellite imagery as an exemplar, the pursuit of enhancing satellite image resolution has prompted innovative strides. A distinctive approach harnessing Discrete Wavelet Transform (DWT) in tandem with interpolation has garnered attention [3]. This method discerningly decomposes the input image into divergent sub-bands via DWT. Subsequently, it couples high-frequency sub-band images with low-resolution counterparts through interpolation. Upon amalgamating these images via inverse DWT, a novel high-resolution rendition is synthesized [3]. Empirical evidence bears testament to the superior performance of this technique, outshining conventional and contemporary image resolution enhancement methodologies both quantitatively and perceptually [3]. Notably, Figure 1 encapsulates the schematic depiction of this pioneering resolution enhancement algorithm [3].

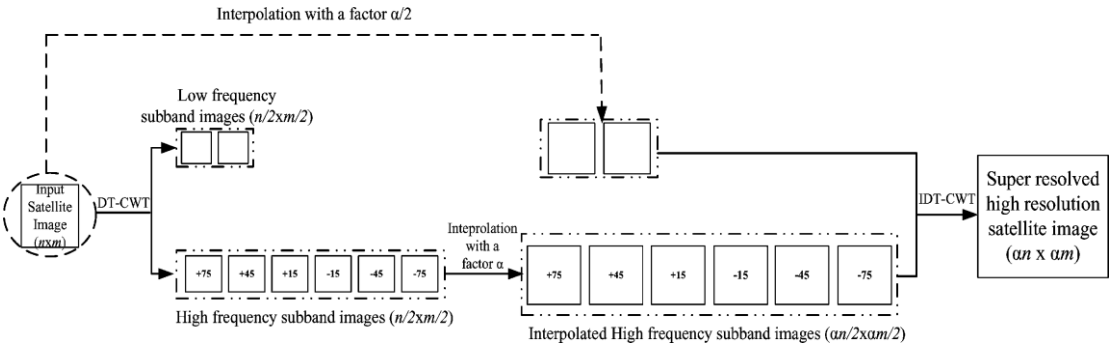


Figure 1. Block diagram of the proposed resolution enhancement algorithm [4]

2.2. Wavelet Transform-Based Image Enhancement

Wavelet transform has broad application prospects in the field of image enhancement, including image denoising, image enhancement, image compression and image edge detection and so on. In various image restoration systems, there are problems of image blurring, distortion and noise caused by image transmission, conversion and other processes [5]. At this time, the wavelet analysis method can effectively solve such problems.

By filtering color noise in wavelet domain, the energy of different frequency components can be adjusted to reduce noise interference and highlight signal characteristics. In recent practical application, the image is simulated by using wavelet transform and color noise, and the ripple coefficients are corrected by P-Laplace method for different loss rates [5].

Initially, a synergy between wavelet transform and noise mitigation techniques is harnessed to reinvigorate and elevate image quality [5]. The inception of this process entails disassembling the original image into disparate resolution tiers, subsequently orchestrating their fusion to realize image detail amplification and reinstatement. Anchoring the endeavor in numerical simulation, parameters

encompassing curvature and slope serve as pivotal tools [5]. Moreover, the strategic utilization of balancing and correction methodologies engenders the averaging of grey regions, culminating in augmented contrast and heightened image intricacy [5]. Conclusively, the trajectory advances towards the recuperation of low-frequency coefficients through wavelet recovery technology. This progression harmoniously coexists with the acquisition of high frequency coefficients across each stratum, a feat deftly accomplished through histogram matching operations [5].

Secondly, the invariant wavelet transform method is used to solve the Gibbs effect and boundary ambiguity problems of wavelet filters, and is verified by numerical simulation and finite difference method [5]. The threshold is used to filter the wavelet coefficients [5]. The wavelet coefficients exceeding the threshold are set as wavelets, and those below the threshold are excluded [5]. The purpose of recovery is achieved by restoring the equal time region and combining it with the image block. The threshold filter selected needs to have a high degree of regularity [5].

3. Fractional Fourier Transform

The Fractional Fourier Transform (FrFT, for short) is an advanced signal analysis technique that has emerged following the development of both the Fourier Transform and the Wavelet Transform. In numerous real-world scenarios, employing the FrFT in place of the conventional FT proves highly beneficial, as it allows for a more effective separation of noise from the original signal. This leads to the generation of a superior-quality signal with enhanced clarity and precision.

This part of this paper will introduce the application of FrFT in the enhancement of acoustic emission signals [6], speech enhancement [7], hoping to provide a useful reference for researchers in the field of signal processing.

3.1. Fractional Fourier Transform Based Adaptive Filtering Techniques

One of the main drawbacks of the well-known nondestructive testing technique known as acoustic emission technology (AET) is how frequently background noise interferes with AE signals [6]. In order to improve acoustic emission signals, adaptive noise canceller (ANC) and self-adaptive noise canceller (SANC) systems based on fractional Fourier transform (FrFT) are presented as a solution to this issue [6]. By rotating the noise signal in the time-frequency plane, the fractional Fourier domain (FrFD) of the noise signal is retrieved in this method [6]. According to the findings, the non-adaptive transform approach based on FrFD outperforms the conventional adaptive transform in both the time and frequency domains, increases signal-to-noise ratio, and dramatically lowers mean square error [6]. The general method steps are as follows:

Incorporating the ANC technology with the Fractional Fourier Transform (FrFT) represents a novel synergy. This fusion leverages an adaptive algorithm to enhance the filter performance within the FrFD domain. Through a continuous interplay with the input signal, the error coefficient is discerned and subsequently employed to update the coefficients of the FrFT filter [6]. This dynamic process unfolds within the analytical framework of the Least Mean Squares (LMS) and Normalized LMS (NLMS) algorithms, manifesting within the FrFD context. The diagrammatic representation of the ANC within FrFD is showcased in Figure 2 [6].

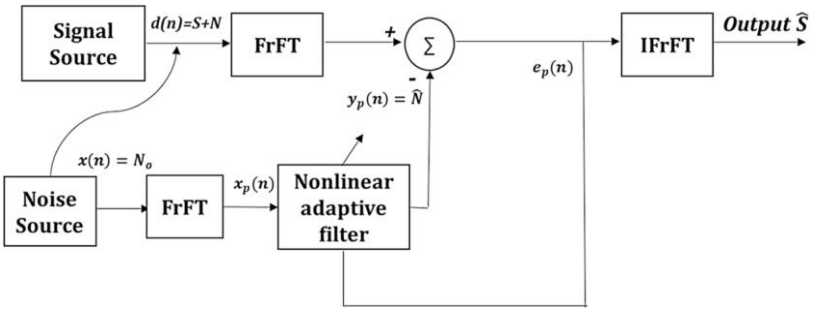


Figure 2. The block diagram of the ANC in FrFD [6]

Furthermore, the Fourier transform acts as the conduit that transmutes the signal from its temporal domain into the frequency domain, facilitating seamless processing and analysis [6]. The strategic augmentation of noise within the FrFD milieu yields promising prospects for refining the forthcoming signal quality and lucidity.

In the fractional Fourier domain, the reference noise vector x(n) and the noise signal d(n) are transformed into xp(n) and dp(n) by inverse fractional Fourier transform (IFrFT), respectively [6]. After the fractional Fourier inverse transform, the output of the FRFD is calculated using an LMS (least mean square) adaptive filter which is given by [6]:

$$y_p(n) = \underset{\sim}{x}_p(n)\,\underset{\sim}{w}(n) \tag{1}$$

Next, the self-adaptive noise canceller (SANC) is combined with the FrFT. Figure 3 displays the SANC schematic diagram in FrFD [6]. The initial step is to convert the original signal into the fractional Fourier domain (FrFD) [6]. The adaptive filter is then supplied the delayed input vector xp (n − Δ) for adaptation [6]. The LMS algorithm is used by the adaptation rule to modify the weights of the adaptive filter in order to reduce the difference between the output signal and the primary source signal and, as a result, suppress noise [6].

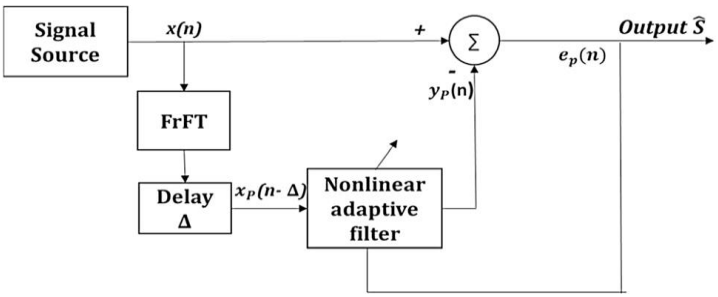


Figure 3. Schematic diagram of SANC in FrFD [6]

3.2. Multi-stage FrFT speech enhancement

Fractional Fourier transform can also convert signals from time domain to fractional domain, where sparsity analysis and signal separation are performed. And finally the separation of speech signal and noise is realized in the transform domain. To address the issue that speech signal and noise cannot be entirely separated in non-stationary noise environment, a multi-level FrFT domain speech enhancement technique based on transform domain sparsity is suggested [7].

The multi-level FrFT domain speech enhancement model is shown in Fig. 3 [7]. Firstly, the optimal transformation order α1, α2, ..., αM are respectively determined by using the fractional order sparsity measure [7]. The enhanced speech after denoising can be obtained as [7]:

$$\hat{x} = F_{-\alpha_M} G_M L F_{\alpha_{k+1}-\alpha_k} G_k L F_{\alpha 2-\alpha 1} G 1 F_{\alpha 1} y \tag{2}$$

For the noisy speech signal, it needs to be divided into frames first, and the speech signal is divided into frames with a length of l [7]. Estimate an optimal filter fractional frequency response gk (k=1, 2, ...M) every L frames [7]. And an objective function needs to be defined to evaluate the performance of the filter [7]:

$$C(\theta) = \frac{1}{L}\sum_{i=1}^{L} \| F_{-\alpha_M} G_M \bullet\bullet\bullet F_{\alpha_{k+1}-\alpha_k} G_k \bullet\bullet\bullet F_{\alpha 2-\alpha 1} G_1 F_{\alpha 2-\alpha 1} y_i - x_i \|^2 \tag{3}$$

Then carry out mathematical derivation to obtain the optimal frequency response gkopt of the fractional order filter, and use an iterative method to obtain gk [7]. The detailed derivation steps are in reference [7]. Figure 4 showed the multi-level FrFT domain speech enhancement model.

$$\nabla_{g_k} C(\theta) \,|\, g_k = g_{k_{opt}} = 0 \tag{4}$$

$$(Q+Q^T)g_{k_{opt}}+d=0 \tag{5}$$

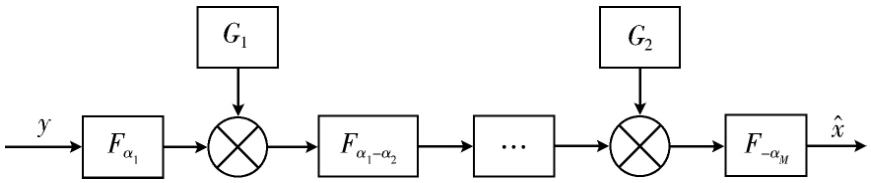


Figure 4. Multi-level FrFT domain speech enhancement model [7]

4. Short-time Fourier Transform

The Short Time Fourier Transform (STFT) is a popular signal augmentation technique. Due to its ability to transform signals into the time-frequency domain, which allows us to more fully grasp their spectral properties and further analyze and process them, STFT has a significant role and a wide range of applications in signal processing.

For signal enhancement, short-time Fourier transform is widely used in speech enhancement, image enhancement, fault diagnosis and so on. In this part of paper, several practical application cases of short-time Fourier transform will be analyzed, hoping to provide reference for the future research of signal enhancement.

4.1. Speech Enhancement using STFT

4.1.1Traditional spectral subtraction method

Single-channel enhancement technique constitutes a distinct subset within the realm of speech enhancement methodologies. However, it's important to acknowledge that this approach harbors limitations in its performance, often necessitating a trade-off between quality enhancement and speech intelligibility within the context of noisy signals [8].

Consider a scenario where the speech signal y(t) becomes compromised, emerging as the composite summation of the unadulterated speech signal s(t) and an additional noise component d(t) [8]. The pivot of this approach unfolds within the domain of short-time Fourier transform (STFT), where the incoming signal undergoes segmentation into a sequence of concise temporal windows. Subsequently, the Fourier transform is applied to each window, yielding a representation within the frequency domain. This intricate process not only relegates the signal into frames but also sets the foundation for subsequent enhancement endeavors [8].

The noisy speech power spectrum is estimated as [8]:

$$|Y(\omega)|^2=|S(\omega)|^2+\delta_n(\omega) \tag{6}$$

Equations (4)-(5) embody the enhanced speech signal amplitude [8]:

$$|\hat{S}(\omega)|=[|Y(\omega)|^2-E(|D(\omega)|^2)]^{1/2} \tag{7}$$

$$=[|Y(\omega)|^2-\delta_n(\omega)]^{1/2} \tag{8}$$

The phase of the noise signal is then combined to synthesize the signal again [8]:

$$S(\omega)=|\hat{S}(\omega)|e^{j\omega g[y(\omega)]} \tag{9}$$

Finally, the short-time Fourier transform is used to transform the signal into the time domain, so that the signal can be processed and analyzed better [8]. Its disadvantage is that it cannot effectively reduce noise in the speech phase, and it needs to rely on voice activity detection (VAD) algorithms that are not suitable for low signal-to-noise ratio environments [8].

4.1.2 Dual-window-size approach, STFT-domain neural speech enhancement

STFT finds practical applications in deep learning-driven speech enhancement as well. To address the prolonged algorithmic delays stemming from lengthy window lengths, and concurrently bolster frequency resolution and enhancement efficacy, a novel dual-window-size approach has been introduced by researchers [9]. This innovative strategy incorporates a sizable window for spectral analysis within the STFT stage, while employing a shorter window for overlap-add during the iSTFT stage [9]. The deep neural network, configured accordingly, facilitates real-time frame-by-frame speech enhancement by projecting the real and imaginary components of the target speech. When combined with complementary techniques, this setup substantially mitigates algorithmic delays [9]. Notably, this approach has demonstrated impressive effectiveness and consistent performance in evaluating noise-reverberant speech enhancement scenarios [9].

4.2. Image Enhancement using STFT

Image enhancement constitutes a pivotal domain within signal processing where Short-Time Fourier Transform (STFT) emerges as a potent tool, particularly in the context of fingerprint image enhancement. The realm of reliable fingerprint identification has garnered substantial attention in recent times. Remarkably, a non-contact fingerprint image enhancement algorithm has been introduced, integrating the prowess of the Hessian matrix alongside short-time Fourier transform (STFT) analysis [10]. This innovative approach synergistically capitalizes on the contrast elevation of ridges and valleys through the Hessian matrix filter, seamlessly followed by STFT's noise suppression and further image refinement [10]. The salient steps of this methodology are succinctly outlined as follows:

(1) Original fingerprint images are subdivided into overlapping patches [10].

(2) STFT analysis is executed on each patch, yielding orientation field images and frequency field images [10].

(3) Gaussian smoothing is judiciously applied to the orientation field image, engendering a coherent image [10].

(4) Isotropic frequency field images are derived through diffusing the frequency field [10].

(5) For each patch, local orientation and angular bandwidth are computed, consequently forming an angular filter [10].

(6) Local frequency and radius bandwidth of each patch are meticulously calculated, thus shaping a radius filter [10].

(7) Each patch undergoes filtration via the pertinent frequency domain filter [10].

(8) Filtering outcomes are translated into the time domain, yielding enhanced plaque images [10].

(9) Subsequently, the mosaic of enhanced plaque images collaborates in the reconstruction of an enriched fingerprint image [10].

This innovative method stands as a testament to the impactful convergence of Hessian matrix enhancement and STFT noise suppression, amplifying the precision and quality of non-contact fingerprint images.

5. Discussion

In this part we will discuss the advantages and disadvantages of wavelet transform, FrFT, and STFT and which method should be used in different fields in order to finally get better results.

5.1. Image enhancement using Wavelet transform

Firstly, Discrete wavelet transform is a discretized wavelet transform method, which discretizes the wavelet analysis method of continuous time and continuous frequency. The discrete wavelet transform divides the signal into multiple non-overlapping subsequences, and performs wavelet decomposition and reconstruction on each subsequence.

This paper mainly discusses the method of wavelet transform in the field of image processing and summarizes the advantages of wavelet transform in this field as follows:

(1) Effective Detail Extraction: Discrete Wavelet Transform (DWT) serves as a proficient method to unravel high-frequency sub-bands. This is accomplished by dissecting the initial image into an array of diverse sub-bands. These high-frequency sub-bands adeptly encapsulate intricate texture details within the processed image, furnishing valuable assets for image signal manipulation.

(2) Interpolation Advancement: The integration of interpolation techniques fortifies the enhancement journey. This approach capitalizes on both high-frequency sub-band images and low-resolution counterparts. By harmonizing these components, a distinct refinement of image details and lucidity is achieved [3].

(3) Augmentation of Resolution: Through a harmonious synthesis of interpolated images and the inverse Discrete Wavelet Transform (DWT), a novel high-resolution image emerges. This strategic fusion culminates in a visual representation characterized by heightened clarity and detail precision [3].

(4) Superseding Conventional Approaches: Comparative assessments illuminate the superiority of DWT techniques over traditional image resolution enhancement methods. Empirical evaluations encompassing metrics such as peak signal-to-noise ratio and root mean square error, coupled with visual appraisals, converge to endorse the perceptible ascendancy of DWT techniques [3].

Secondly, wavelet transform can effectively solve problems such as image blur, distortion and noise in image enhancement. Color noise can be filtered out in the wavelet domain by the wavelet filter, which can adjust the energy of different frequency components, reduce noise interference, and highlight signal features [5]. The image can be restored and enhanced using wavelet transform and noise processing technology, and the image can be decomposed into different resolution levels to achieve image detail enhancement and restoration. The invariant wavelet transform method can solve the Gibbs effect and boundary ambiguity problems of the wavelet filter, and it is verified by numerical simulation and finite difference method [5].

However, the computational complexity of wavelet transform is high. In some cases, wavelet transform may lead to the loss of information of image details. There are still some challenges in the application of wavelet transform in image enhancement.

5.2. Fractional Fourier Transform

The main advantages of FrFT can be divided into three aspects

(1) Remove noise

(2) Improve signal-to-noise ratio (SNR)

(3) Low distortion

Based on [7] The results show that FrFT can effectively remove noise, and the original speech signal can be better distinguished, and multistage FrFT can get better results than single-stage FrFT. Based on the simulation results of [7], it can be concluded that FrFT provides a good noise reduction effect and can clearly identify useful signals.

The traditional methods such as wavelet denoising are compared with FrFT, and it is finally concluded that FrFT can improve the signal-to-noise ratio better, and the performance of wavelet denoising is inferior to FrFT. At the same time, multistage FrFT is still superior to single-stage FrFT [7].

The distortion degree between wavelet denoising and FrFT is evaluated. After observation, it can be found that the enhanced speech based on FrFT algorithm has the lowest distortion degree. Therefore, FrFT can enhance signal and reduce noise while keeping the signal closer to the original state [7].

5.3. Short-Time Fourier Transform

STFT can convert the signal to the frequency domain for processing and analysis, and at the same time it allows for fine-grained processing of the signal in the time domain and frequency domain.

However, while improving the quality of the noise signal, it will sacrifice certain intelligibility, that is, the understanding of the speech signal is not clear enough.

Additionally, the Short-Time Fourier Transform (STFT) finds significant utility in the realm of deep learning-based speech enhancement applications [9]. The conventional application of STFT, albeit powerful, can introduce considerable algorithmic delays owing to its prolonged window length. However, a pioneering approach using dual-window sizes, as described earlier, offers an elegant solution to this issue. This innovative method, when coupled with supplementary techniques, adeptly mitigates algorithmic latency, concurrently augmenting both frequency resolution and enhancement efficacy [9].

Lastly, STFT exhibits practical significance in the domain of fingerprint image enhancement [10]. Leveraging the prowess of the STFT algorithm, noise suppression is executed adeptly, ushering in elevated image quality and heightened clarity. In a symbiotic approach, when fused with complementary algorithms, STFT facilitates an additional layer of image refinement. This augmentation manifests in heightened detail prominence, consequently amplifying the precision of fingerprint recognition [10]. Moreover, by deftly manipulating window dimensions and stride increments, STFT can harmonize the frequency spectrum, culminating in further image quality amelioration.

6. Conclusion

This paper presents a comprehensive review of signal enhancement methodologies grounded in wavelet transform, short-time Fourier transform (STFT), and fractional Fourier transform (FrFT). Initially, the three distinct methods and their corresponding domains of application are introduced. Subsequently, through an analysis of simulation outcomes, the strengths and limitations of each approach, alongside their applicable scenarios, are elucidated. Primarily delving into the domain of image enhancement, this paper extensively explores the application of wavelet transform. It is inferred that this technique yields heightened image sharpness while rectifying issues such as blurriness, distortion, and noise. In the context of the short-time Fourier transform, the focus shifts to its utility in both speech enhancement and image enhancement. The findings underscore its efficacy in bolstering signal quality. However, it is acknowledged that this approach grapples with a deficiency—the clarity of speech signals. Concurrently, STFT exhibits prowess in elevating image quality and lucidity, especially evident in the realm of image enhancement. Turning attention to the fractional Fourier transform, this study rigorously examines its implications in speech enhancement and the augmentation of acoustic emission signals. The inference is that FrFT efficaciously enhances the signal-to-noise ratio while incurring minimal signal distortion. Furthermore, its ability to segregate original signals from noise is lauded. Through a comprehensive exploration of these methodologies, this paper contributes to a nuanced understanding of their applications, advantages, and limitations, providing valuable insights for the field of signal enhancement.

Authors Contribution

All the authors contributed equally and their names were listed in alphabetical order.

References

[1] Löllmann H W, Barfuss H, Deleforge A, et al. Challenges in acoustic signal enhancement for human-robot communication [C] // Speech Communication; 11. ITG Symposium. 2014: 1-4.

[2] H. Tang, S. Lu, G. Qian, J. Ding, Y. Liu and Q. Wang, "IoT-Based Signal Enhancement and Compression Method for Efficient Motor Bearing Fault Diagnosis," in IEEE Sensors Journal, 2021, 21 (2): 1820-1828.

[3] H. Demirel and G. Anbarjafari, "Discrete Wavelet Transform-Based Satellite Image Resolution Enhancement," in IEEE Transactions on Geoscience and Remote Sensing, 2011, 49 (6): 1997-2004.

[4] H. Demirel and G. Anbarjafari, “Satellite image resolution enhancement using complex wavelet transform,” IEEE Geosci. Remote Sens. Lett., 2010, 7 (1): 123–126.

[5] Wang R. Exploring Wavelet Transform-Based Image Enhancement Algorithm for Image Restoration of Long March National Cultural Park[J]. Journal of Environmental and Public Health, 2022.

[6] Prajna, K., Mukhopadhyay, C.K. Fractional Fourier Transform Based Adaptive Filtering Techniques for Acoustic Emission Signal Enhancement. J Nondestruct Eval 39, 14 (2020).

[7] Fan Zhenyan, Zhuang Xiaodong, Li Zhongxiao. Multistage FrFT speech enhancement based on the transform domain sparsity measure [J]. Computer Engineering and Design, 2020, 41 (09): 2574-2584.

[8] Kulkarni D S, Deshmukh R R, Shrishrimal P P. A review of speech signal enhancement techniques[J]. International Journal of Computer Applications, 2016, 139 (14).

[9] Z. -Q. Wang, G. Wichern, S. Watanabe and J. Le Roux, "STFT-Domain Neural Speech Enhancement with Very Low Algorithmic Latency," in IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2023, 31: 397-410.

[10] Y. Tang, L. Jiang, Y. Hou and R. Wang, "Contactless Fingerprint Image Enhancement Algorithm Based on Hessian Matrix and STFT," 2017 2nd International Conference on Multimedia and Image Processing (ICMIP), Wuhan, China, 2017: 156-160.