Fractional Time-Frequency Scattering Convolution Network

Jiabin Zheng[1(✉)], Jun Shi[2], Gong Chen[2], Weiping Chen[1], and Zhenya Geng[3]

[1] MEMS Center, Harbin Institute of Technology, Harbin 150001, China
chenwp@hit.edu.cn
[2] Communication Research Center, Harbin Institute of Technology, Harbin 150001, China
junshi@hit.edu.cn
[3] Department of Control Science and Engineering, Harbin Institute of Technology, Harbin 150001, China
zhenya.geng@hit.edu.cn

Abstract. The wavelet scattering convolution network (SCN) have recently developed as a kind of effective feature extractor, which has achieved a great performance in signal and image processing applications. Unfortunately, as feature extractor, SCN is not appropriate to mimic the visual system of mammals in image classification tasks, so that STFT-based time-frequency scattering convolution network (TFSCN) is proposed. However, TFSCN is limited by a major drawback: it is only available for stationary signals'analysis but not for non-stationary ones, since STFT can viewed as linear translation-invariant filters in the FT domain intrinsically. The aim of this paper is to overcome this weakness using the short-time fractional fourier transform (STFRFT) which is a bank of linear translation-variant bandpass filters and thus may be used for non-stationary signal analysis. First, We present the fractional time-frequency scattering transform based upon the STFRFT. Then a generalization of TFSCN's structure dubbed FRTFSCN is illustrated. The significant performance of FRTFSCN are shown via experiment simulations.

Keywords: Time-frequency scattering · Scattering network · Short-time fractional fourier transform · Non-stationary signal analysis · Translation-variant filtering

1 Introduction

In recent years, the wavelet scattering convolution network [1–3] (SCN) introduced by Mallat has drawn extensive attention in various field, and has led to state-of-the-arts results in a wide range of classification tasks including handwritten digit recognition [2], musical genre classification [3], audio classification [4], texture discrimination [5], art authentication [6], astronomy [7], chemical [8],

X. Wang et al. (Eds.): AICON 2021, LNICST 396, pp. 75–81, 2021.
https://doi.org/10.1007/978-3-030-90196-7_7

biomedical science [9,10], and time-frequency representations [11]. It is pioneering in mathematical analysis of feature extractors generated by DCNNS [1,2]. Furthermore, Li et al. developed time-frequency scattering convolution network (TFSCN) [11] by combing Mallats scattering transform framework with short-time fourier transform (STFT) with Gaussian window to make sense to mimic the visual system of mammals for designing a feature extractor in image classification. Theoretically, TFSCN provides a translation and rotation-invariant and deformation-stable representation by cascading Gabor filters and modulus nonlinearities, and performs well in practice e.g. image classification. Unfortunately, TFSCN still suffers from a major drawback: it is only appropriate for stationary signals'analysis but not for non-stationary ones, for the reason that the STFT is intrinsically a set of linear time-invariant bandpass filters in the Fourier transform (FT) domain, which indicates that TFSCN does not work well when analyzing non-stationary signals. Thus, it is desirable to impart a certain degree of time-varying behavior to these networks. Our objective of this paper is to propose a new structure for TFSCN dubbed fractional TFSCN (FRTFSCN) by employing the short-time fractional fourier transform (STFRFT) which is a generalization of the conventional short-time fourier transform (STFT) in the fractional FT (FRFT) domain. We first define fractional time-frequency scattering transform based on STFRFT, then fractional time-frequency scattering convolution network is constructed using STFRFT-based scattering transform, which includes the STFT-based TFSCN as a special case. Finally, a practical application of the FRTFSCN in image classification is discussed.

2 Fractional Time-Frequency Scattering Convolution Network

In this section, the formulation of the FRTFSCN is described. We start by introducing the definition of the STFRFT. Then, a STFRFT-based scattering transform is proposed. In the following, the FRTFSCN is constructed according to STFRFT-based scattering transform.

2.1 The Definition of STFRFT

According to [12], *Shi et al.* define a novel STFRFT of a function $f(t) \in L^2(\mathbb{R})$ with respect to a given window $g(t)$ as

$$\mathrm{STFRFT}_f^{\alpha}(t,u) = \int_R f(\tau) g^*_{\alpha,t,u}(\tau) d\tau \tag{1}$$

with its kernel given by

$$g_{\alpha,t,u}(\tau) \triangleq g(\tau - t) e^{-j\frac{\tau^2 - t^2}{2}\cot\alpha + j\tau u \csc\alpha} \tag{2}$$

then based on the definition of the fractional convolution [13]:

$$f(t) *_{\alpha} g(t) = \int_{\mathbb{R}} f(\tau) g(t-\tau) e^{-j\frac{t^2-\tau^2}{2}\cot\alpha} d\tau \tag{3}$$

then (1) can be rewritten as the form of fractional convolution, that is

$$\mathrm{STFRFT}_f^{\alpha}(t,u) = e^{-jtu\csc\alpha}\Big[f(t) *_{\alpha} \big(g^*(-t)e^{jtu\csc\alpha}\big)\Big]. \tag{4}$$

Note that, when $\alpha = \pi/2$, STFRFT reduces to STFT.

2.2 STFRFT-Based Scattering Transform

In this operation, a scattering transform computes nonlinear invariants from fractional short-time fourier coefficients by modulus operator performed as a nonlinear pooling operator on the ground that it could preserve the signal energy [11]. As for the signal $x(t) \in L^2(\mathbb{R})$, the STFRFT-modulus coefficients, considered as the translation-invariant coefficients, are built from the STFRFT by the modulus operator defined as:

$$U[p]x = |x *_{\alpha} f_p|, \quad p \in \mathbb{P} \tag{5}$$

and STFRFT-based scattering transform is defined as

$$S[p]x = U[p]x *_{\alpha} f_0, \quad p \in \mathbb{P} \tag{6}$$

where $f_p(t) = f_0(t)e^{jpt}$, $p \in \mathbb{P}$ and f_0 form a frame dubbed as a fractional uniform covering frame, that is $\mathfrak{F} = \{f_0\} \cup \{f_p : p \in \mathbb{P}\}$ satisfying

$$|F_0(u\csc\alpha)|^2 + \sum_{p\in\mathbb{P}} |F_p(u\csc\alpha)|^2 = 1 \tag{7}$$

where $F_0(u\csc\alpha)$ and $F_p(u\csc\alpha)$ denote the FT (with their argument scaled by $\csc\alpha$) of $f_0(t)$ and $f_p(t)$, respectively. In addition, Eq. 7 implies that $\mathfrak{F}$ is a semi-discrete Parseval frame [14] for all $x(t) \in L^2(\mathbb{R})$,

$$\|x *_{\alpha} f_0\|^2 + \sum_{p\in\mathbb{P}} \|x *_{\alpha} f_p\|^2 = \|x\|^2 \tag{8}$$

More fractional STFRFT-modulus coefficients can be obtained by further iterating on the STFRFTs and the modulus operator along any path $p \in \mathbb{P}^k$ and $k \in \mathbb{N}$, then we associate $p \in \mathbb{P}^k$ with the fractional scattering propagator $U[p]$:

$$U[p]x = \begin{cases} x, & \text{if} \quad p \in \mathbb{P}^0 \\ |x *_{\alpha} f_p|, & \text{if} \quad p \in \mathbb{P} \\ U[p_k]U[p_{k-1}]\cdots U[p_1]x, & \text{if} \quad p = (p_1, p_2, \ldots, p_k) \in \mathbb{P}^k \end{cases} \tag{9}$$

then STFRFT-based scattering transform with association $p \in \mathbb{P}^k$, $S[p]$, is a vector-valued operator

$$S[p]x = \{U[p]x *_{\alpha} f_0 : \quad p \in \mathbb{P}^k, k = 0, 1, 2, \cdots\} \tag{10}$$

which is fractional time-frequency scattering transform, the generalization of the conventional time-frequency scattering transform.

2.3 The Construction of FRTFSCN

Based on the results obtained above, we now introduce the fractional time-frequency scattering convolution network (FRTFSCN), which can be viewed as an iterative process over a one-step fractional scattering propagator. Similar to the DCNNs [15], the FRTFSCN is built upon a building block comprised of a FRFT-domain filtering followed by a modulus nonlinearity. Let us recursively construct the FRTFSCN. The first layer collects all the results of the STFRFTs with respect to $k = 0$, i.e.,

$$S[p]x = \{U[p]x *_\alpha f_0 : \quad p \in \mathbb{P}^k, k = 0\} = x *_\alpha f_0 \tag{11}$$

The m-th layer of the FRTFSCN is constructed by taking all the possible STFRFTs of level $m-1$, i.e., $S[p]x$ with $p \in \mathbb{P}^{m-1}$, where $\mathbb{P}^{m-1}$ is the set of all the paths p of length $m-1$. Note from (9) that for any given path $p \in \mathbb{P}^{m-1}$ and m-th path element $p_m \in \mathbb{P}$, we can derive a m-length path set $\mathbb{P}^m$ satisfying $\mathbb{P}^m = \{\mathbb{P}^{m-1}, p_m\}$ and $p_m \in \mathbb{P}$. Then we can derive that m-th layer of the FRTFSCN satisfying

$$\begin{aligned} U[p']x &= U[p_m]U[p]x, \quad p' = (p, p_m), p \in \mathbb{P}^{m-1} \\ &= U[p_m]U[p_{m-1}]\cdots U[p_1]x. \end{aligned} \tag{12}$$

Moreover, it follows that

$$S[p']x = U[p']x *_\alpha f_0 \tag{13}$$

therefore, we can iteratively compute the nodes of the FRTFSCN by first recursively calculating the fractional scattering propagators $U[p]$ for all level m up to the pre-determined maximum depth of the FRTFSCN. Then, all the nodes value can be extracted by computing the STFRFTs. A graphical representation of the proposed FRTFSCN is shown in Fig. 1. As can be seen from the figure, the proposed FRTFSCN is a fully connected network which has the similar structure of CNNs [15]. Compared with the conventional TFSCN, FRTFSCN is performed from the perspective of fractional filtering and remain some basic properties in fractional domain, like rotation and translation-invariant and deformation stability. Furthermore, note that, when $\alpha = \pi/2$, the FRTFSCN reduce to TFSCN,

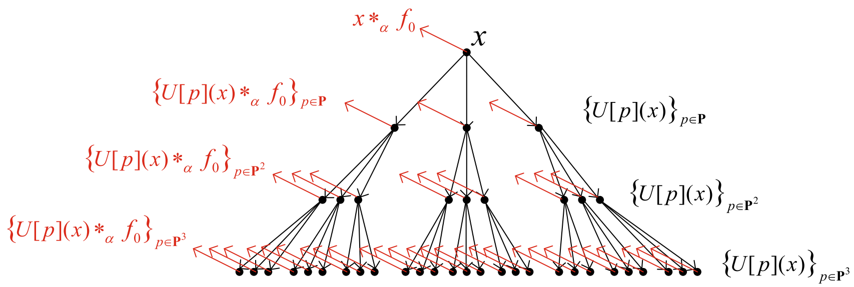

Fig. 1. The architecture of FRTFSCN.

which implies that TFSCN is a special case of the FRTFSCN. The FRTFSCN has a free parameter α and thus is more flexible than the conventional TFSCN.

3 The Application of FRTFSCN

In this section, we give a potential application of the proposed FRTFSCN. Since both the FRTFSCN and TFSCN the similar properties like rotation and translation-invariant and deformation stability, it means the FRTFSCN could provide a translation- and rotation-invariant fractional time-frequency scattering representation which is stable to small deformations. These properties determine that FRTFSCN is useful in image classification tasks. So, we apply the proposed FRTFSCN in a texture classification to illustrate its significant performance.

3.1 Texture Database

The database constructed by merging UMD [16] and ALOT [17], which contains 9 classes of images, as illustrated in Fig. 2, and there are 40 images in each class. As can be seen from the Fig. 2, there are some non-stationary textures in those sample images, which implies FRTFSCN would perform better compared with the conventional TFSCN when selecting an appropriated angle.

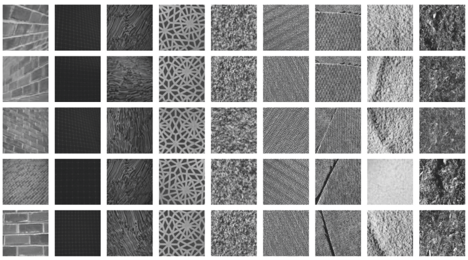

Fig. 2. Sample images from the UMD [16] and ALOT [17] databases.

3.2 Classification Results

In classification stage, the images in each class are divied into two groups randomly. One group of images is labled for training, and the other group is for testing. As aforementioned, the FRTFSCN has a free parameter $\alpha \in [0, \pi/2]^d$, on which the classification performance depends. For image analysis, we set $d = 2$ and $\alpha = (\alpha_1, \alpha_2)$. According to [18], we fix one element of α to $\pi/2$, and let the other vary from 0 to $\pi/2$ with a step of $\pi/20$ to accelerate the computation of the FRTFSCN. Hereby, the angle pair $(9\pi/10, \pi/2)$ is chosen to calculate fractional time-frequency scattering coefficients which fed into PCA classifier for classification. The results of classification are shown in Table 1 which gives a comparison of the FRTFSCN and several well-known networks (e.g., the conventional SCN

[1], the time-frequency scattering network (TFSN) [11], and the CNN [15]). To illustrate the results of experiments, the third row of Table 1 is taken as an example, in which each class containing 40 images, 10 sample images are randomly chosen for training, and the remaining 30 images are used for testing. In order to obtain statistical results, we calculate the mean and variance of classfication errors for the FRTFSCN with $(9\pi/10, \pi/2)$ over 1000 experiments. As can be observed from Table 1, the FRTFSCN with angle pair $(9\pi/10, \pi/2)$ yields the minimum classification error, and even with small size of training, its performance is always best. Particularly, when the size of training is extremely small (i.e. training samples = 2), the FRTFSCN still exhibits best performance.

Table 1. Classification errors using different networks.

Training samples	FRTFSCN $(9\pi/10, \pi/2)$	TFSCN	SCN	CNN
2	$\mathbf{0.13 \pm 0.05}$	0.23 ± 0.03	0.24 ± 0.02	0.45 ± 0.08
5	$\mathbf{0.05 \pm 0.03}$	0.13 ± 0.04	0.13 ± 0.04	0.35 ± 0.09
8	$\mathbf{0.02 \pm 0.02}$	0.09 ± 0.02	0.08 ± 0.03	0.34 ± 0.12
10	$\mathbf{0.02 \pm 0.02}$	0.07 ± 0.02	0.07 ± 0.02	0.30 ± 0.11
15	$\mathbf{0.01 \pm 0.01}$	0.05 ± 0.02	0.04 ± 0.01	0.18 ± 0.07
20	$\mathbf{0.01 \pm 0.01}$	0.04 ± 0.01	0.02 ± 0.01	0.12 ± 0.05

4 Conclusion

This paper aims to overcome the drawback of the TFSCN based upon the STFRFT, which are suitable to analysis non-stationary signal as a bank of linear translation-variant bandpass filters. First, a STFRFT-based scattering transform is proposed, then we construct FRTFSCN based on fractional time-frequency scattering transform. Eminent numerical performance of the FRTFSCN is presented to illustrate its advantage in image classification application.

References

1. Mallat, S.: Group invariant scattering. Commun. Pure Appl. Math. **65**(10), 1331–1398 (2012)
2. Bruna, J., Mallat, S.: Invariant scattering convolution networks. IEEE Trans. Pattern Anal. Mach. Intell. **35**(8), 1872–1886 (2013)
3. Andén, J., Mallat, S.: Deep scattering spectrum. IEEE Trans. Signal Process. **62**(16), 4114–4128 (2014)
4. Andén, J., Lostanlen, V., Mallat, S.: Joint time-frequency scattering. IEEE Trans. Signal Process. **67**(14), 3704–3718 (2019)
5. Sifre, L., Mallat, S.: Rigid-motion scattering for texture classification. Appl. Comput. Harmon. Anal. **00**, 1–20 (2014)
6. Leonarduzzi, R., Liu, H., Wang, Y.: Scattering transform and sparse linear classifier for art authentication. Signal Process. **150**, 11–19 (2018)

7. Allys, E., et al.: The RWST, a comprehensive statistical description of the non-Gaussian structures in the ISM. A&A **629**(A115), 1–21 (2019)
8. Hirn, M., Mallat, S., Poilvert, N.: Wavelet scattering regression of quantum chemical energies. Multiscale Model. Simul. **15**(2), 827–863 (2017)
9. Chudáček, V., Andén, J., Mallat, S., Abry, P., Doret, M.: Scattering transform for intrapartum fetal heart rate variability fractal analysis: a case-control study. IEEE Trans. Biomed. Eng. **61**(4), 1100–1108 (2014)
10. Oyallon, E., et al.: Scattering networks for hybrid representation learning. IEEE Trans. Pattern Anal. Mach. Intell. **41**(9), 2208–2221 (2019)
11. Czaja, W., Li, W.: Analysis of time-frequency scattering transforms. Appl. Comput. Harmon. Anal. **47**, 149–171 (2019)
12. Shi, J., Zheng, J., Liu, X., Xiang, W., Zhang, Q.: Novel short-time fractional fourier transform: theory, implementation, and applications. IEEE Trans. Signal Process. **68**, 3280–3295 (2020)
13. Shi, J., Sha, X., Song, X., Zhang, N.: Generalized convolution theorem associated with fractional Fourier transform. Wirel. Commun. Mob. Comput. **14**(13), 1340–1351 (2014)
14. Chui, C.K., Shi, X.: On a Littlewood-Paley identity and characterization of wavelets. J. Math. Anal. Appl. **177**(2), 608–626 (1993)
15. Krizhevsky, A., Sutskever, I., Hinton, G.E.: ImageNet classification with deep convolutional neural networks. Adv. Neural Inform. Process. Syst. **25**(2), 1097–1105 (2012)
16. http://users.umiacs.umd.edu/~fer/website-texture/texture.htm
17. http://aloi.science.uva.nl/public_alot/
18. Liu, L., Wu, J., Li, D., Senhadji, L., Shu, H.: Fractional wavelet scattering network and applications. IEEE Trans. Biomed. Eng. **66**(2), 553–563 (2019)