FREQUENCY DILATION LEARNING FOR TEMPORAL CONVOLUTIONAL NEURAL NETWORKS

PREPRINT- WORK IN PROGRESS

Moritz Wolter
Institute for Computer Science
University of Bonn
wolter@cs.uni-bonn.de

Angela Yao
School of Computing
National University of Singapore
ayao@comp.nus.edu.sg

August 7, 2019

ABSTRACT

Convolutional neural networks (CNNs) make up the bedrock in modern machine learning. With ever increasing data sets and increasing efforts to port these foundational systems to mobile and robotics applications there is an ever strong desire to reduce the computational and memory footprint of CNNs. In this paper we introduce frequency dilation, a novel technique to increase efficiency at minimal cost in terms of network accuracy.

1 Introduction

Integral transforms lie at the core of computational efficiency trough sparsity [2][6]. Previous works have reduced the parameter count of fully connected layers trough the fast Hadamard transforms [8]. Fully connected layers are re-parametrized trough a fixed basis. Using a fixed bases [1] proposes a unitary RNN, later [7] finds that in the RNN case using a fixed basis is detrimental to network performance. As we believe that the fixed basis in [8] is equally restricted, and therefore detrimental to performance. Instead we proposed a learn-able representation based on the dual tree wavelet transform [4], which we apply to both input data and network weights.

2 Related work

Learnable filters [3]

Dimensionality reduction in inputs (FourierRNN)

Dimensionality reduction in weights [8]

3 Theory

Sparsity of a representation is basis dependent [6][5], fewer coefficients are are required in order to represent a signal, only a limited number of basis functions components are able to do the job.

3.1 Single tree filter design

Wavelet filter design

Perfect reconstruction (PR) [4][6, page 107] requires, no distortion:

$$H_0(z)F_0(z) + H_1(z)F_1(z) = 2 \tag{1}$$

and alias cancellation:

$$H_0(-z)F_0(z) + H_1(-z)F_1(z) = 0 \tag{2}$$

For alias cancellation $F_0(z) = H_1(-z)$, $F_1(z) = -H_0(-z)$ is typically chosen [5]. A product filter approach is chosen to deal with the reconstruction condition.

$$P_0(z) = F_0(z)H_0(z); P_1(z) = F_1(z)H_1(z) \tag{3}$$

Alias cancellation leads to $P_1(z) = -P_0(-z)$ and turns the no distortion condition into:

$$P(z) + P(-z) = 2 \tag{4}$$

Even powers have to cancel, odd powers are design variables.

3.2 Dual vs. single Tree approach

3.3 The dual tree wavelet approach

Dual tree wavelet filters should satisfy [4]:

- approximate half sample delay property (?).
- Perfect reconstruction (orthogonal or bi-orthogonal)
- finite support (FIR filters)
- vanishing moments good stopband
- linear phase filters (desired, but not required)

Where and how does the phase shift condition fit into this?
The wavelet literature tells us $\int_{-\infty}^{\infty} \Psi(t)dt = 0$ and $\int_{-\infty}^{\infty} \|\Psi(t)\|dt = 1$.

3.4 Biorthogonal filters

3.4.1 Learning biorthogonal filters

Which constraints are required?

3.5 Multi-resolution analysis

Top-down vs bottom up in image processing, relations?

4 Experiments

Compare, compression using Fourier, Harr, and learned wavelets on mackey glass and a bumpy rectangularly time series for example a staircase mackey glass or lorenz.

References

[1] Martin Arjovsky, Amar Shah, and Yoshua Bengio. Unitary Evolution Recurrent Neural Networks. 48, 2016.

[2] Stephane Mallat and Gabriel Peyré. *A Wavelet Tour of Signal Processing: The Sparse Way*.

[3] Daniel Recoskie and Richard Mann. Learning Sparse Wavelet Representations. pages 1–7, 2018.

[4] Ivan W Selesnick, Richard G Baraniuk, and Nick G Kingsbury. The Dual-Tree Complex Wavelet Transform. (November 2005):123–151, 2005.

[5] Gilbert Strang. Wavelets. 82(3):250–255, 1994.

[6] Gilbert Strang and Truong Nguyen. Wavelets and Filter Banks, 1997.

[7] Scott Wisdom, Thomas Powers, John R Hershey, Jonathan Le Roux, and Les Atlas. Full-Capacity Unitary Recurrent Neural Networks arXiv : 1611 . 00035v1 [stat . ML] 31 Oct 2016. (Nips):1–9, 2016.

[8] Ziyu Yang, Zichao and Moczulski, Marcin and Denil, Misha and de Freitas, Nando and Smola, Alex and Song, Le and Wang, Le Song, and Ziyu Wang. Deep Fried Convnets. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 1476—-1483, 2015.