Frequency-Aware Reconstruction of Fluid Simulations with Generative Networks

Simon Biland, Vinicius C. Azevedo, Byungsoo Kim and Barbara Solenthaler

ETH Zurich

Abstract

Convolutional neural networks were recently employed to fully reconstruct fluid simulation data from a set of reduced parameters. However, since (de-)convolutions traditionally trained with supervised ℓ_1-loss functions do not discriminate between low and high frequencies in the data, the error is not minimized efficiently for higher bands. This directly correlates with the quality of the perceived results, since missing high frequency details are easily noticeable. In this paper, we analyze the reconstruction quality of generative networks and present a frequency-aware loss function that is able to focus on specific bands of the dataset during training time. We show that our approach improves reconstruction quality of fluid simulation data in mid-frequency bands, yielding perceptually better results while requiring comparable training time.

CCS Concepts

• *Computing methodologies → Physical simulation; Neural networks;*

1. Introduction

Recent work has demonstrated the effectiveness of using convolutional neural networks (CNNs) for fluid simulations, for example, to improve the computational efficiency by replacing the pressure computation [TSSP17] or by predicting the fluid motion with an LSTM network [WBT19]. A generative deep neural network for parameterized fluid simulations was introduced that reconstructs the velocity field based on a few input parameters [KAT*19]. This can be used for data compression or to generate simulations by interpolating in the parameter space. A main limitation of this previous approach (and related neural fluid solvers) is the difficulty to reconstruct high frequency flow structures, which can be accounted to the ℓ_1-loss function that is used on the velocity and its gradient.

Loss functions and their influence on reconstruction quality have been studied in the image processing literature. Natural images are known to have an expected power spectrum following an inverse power law: the magnitudes of spectral decompositions are mostly concentrated in lower frequencies. However, high-frequencies with lower magnitudes tend to concentrate perceptually important details. Thus, using a mean square error loss will often be more effective in matching lower frequencies that have higher magnitudes and dominate the power spectrum [WB09].

Rahaman et al. [RBA*19] used Fourier analysis to investigate which frequencies neural networks tend to learn and reported a phenomenon they call the spectral bias when experimenting with ReLU networks. The authors found that neural networks favor low frequencies, even though they are theoretically universal approximators [Hor91]. Other work has arrived at sim-

ilar conclusions [XZX19, XZL*19], reporting that dominant low-frequency components are captured quickly during training and high-frequency ones only slowly thereafter (F-Principle). It is argued that this acts as an intrinsic regularizer and filter for noisy input data. Complementary losses were used to improve high-frequency reconstruction, such as perceptual losses [JAFF16] and adversarial losses (GAN) [GPAM*14, BSM17, KALL18].

Previous perceptual-based loss functions were matched by selectively focusing on frequencies through a weighted wavelet per-band loss combined with a term that penalizes low magnitude values [HHST17]. The wavelet domain was also employed in several image super resolution approaches. A CNN to predict the missing details of wavelet coefficients (sub-bands) of the LR images was used in [GMVM17], and a loop architecture to better explore statistical relationships among wavelet coefficients in different bands was applied in [GCZ*18]. Previous work also used clique convolutional blocks [ZSY*18], and scatter maps were employed as input to standard CNNs [GX16]. Wavelet super resolution approaches were also augmented with adversarial losses to better reconstruct the distribution of frequencies [HHST19, ZWY19].

Using neural networks for fluid simulations is largely unexplored, and the input data is intrinsically different from images. Our work therefore aims at better understanding the characteristics and reconstruction properties of fluid simulation data. We use the generative approach of [KAT*19] as a baseline, and present a frequency-aware loss function using Fourier transform. Our results show that reconstruction quality can be improved by considering the different frequency bands in the optimization.

2. Generative Network for Fluid Flow Reconstruction

Our work is based on the generative network for fluid simulations presented by Kim et al. [KAT*19]. Their generator G reconstructs a velocity field $\mathbf{u_c} \in \mathbb{R}^{H \times W \times D \times V_{dim}}$ (height H, width W, depth D, dimension of velocity field V_{dim}) from a small set of parameters $\mathbf{c} = [c_1, c_2, \ldots, c_n] \in \mathbb{R}^n$, which can be represented as a function $G(\mathbf{c}) : \mathbb{R}^n \mapsto \mathbb{R}^{H \times W \times D \times V_{dim}}$. The values c_i parameterize the scene that was used to create the training data. As an example, c_1 could be the position and c_2 the width of a smoke source, while c_3 represents the simulation time step. The network is trained with the baseline loss L_b that includes ℓ_1-losses on the velocity field and its gradient:

$$L_b = \lambda_{\mathbf{u}} \|\mathbf{u} - \hat{\mathbf{u}}\|_1 + \lambda_{\nabla \mathbf{u}} \|\nabla \mathbf{u} - \nabla \hat{\mathbf{u}}\|_1, \tag{1}$$

where $\mathbf{u}$ is the simulated ground truth velocity field, $\hat{\mathbf{u}} = G(\mathbf{c})$ and $\nabla \hat{\mathbf{u}} = \nabla G(\mathbf{c})$, and $\lambda_{\mathbf{u}}$ and $\lambda_{\nabla \mathbf{u}}$ serve as weighting factors between the two loss terms. A divergence-free velocity field can be enforced by changing the generator to $G(\mathbf{c}) = \nabla \times G'(\mathbf{c})$, which changes the network to learn a stream function $\psi_{\mathbf{c}}$.

2.1. Analysis of the Reconstruction Quality

Although the generative network reliably reconstructs the global structure of the flow, it is apparent that small-scale flow details are lost. To verify and analyze this observation, we first transform 100 ground truth and reconstructed velocity field samples into the Fourier domain and plot the distribution of the log-magnitude values. Figure 6 (top left) shows that the the distribution of the baseline reconstruction (black) has a smaller variance than the ground truth, which can be attributed to the ℓ_1-loss: to achieve a low error over a large number of training examples, the network strives for universally appropriate values and therefore settles for magnitudes close to the mean of the ground truth data. This assumption is supported by Figure 6 (top right), which shows the standard deviation of the reconstructed magnitudes separated by bands relative to the standard deviation of the ground truth for the baseline (black). Between bands 20 and 30 we observe that the reconstructed standard deviation is only around 60% of the true standard deviation, which results in the observed clustering around the mean value.

2.2. Frequency-Aware Loss Function

Based on the observations discussed above, we propose a frequency-aware method that learns individual frequency bands separately as illustrated in Figure 1. Both the ground truth $\mathbf{u}_c$ and the output of the generator $\hat{\mathbf{u}}_c$ are transformed to the Fourier domain $(FT(\mathbf{u}), FT(\hat{\mathbf{u}}))$ and split into different bands. Then, we compute the ℓ_1-norm of the difference of each band and aggregate them into a weighted sum. We compute the error directly in the Fourier domain to avoid expensive computations of inverse transforms and gradients. We define the total loss as $L = \lambda_o L_b + L_{fol}$, with the Fourier loss L_{fol} given as

$$L_{fol} = \sum_b w_b \||FT(\mathbf{u})_b - FT(\hat{\mathbf{u}})_b|\|. \tag{2}$$

$FT(\mathbf{u})_b$ is the Fourier transform of $\mathbf{u}$ filtered by band b, $|\cdot|$ is the element-wise complex norm, and $\|\cdot\|_1$ is the ℓ_1-norm averaged over all pixels in the image. Note that the loss could also be defined on the phase (as in [MDM*18]) or magnitude, but we observed low
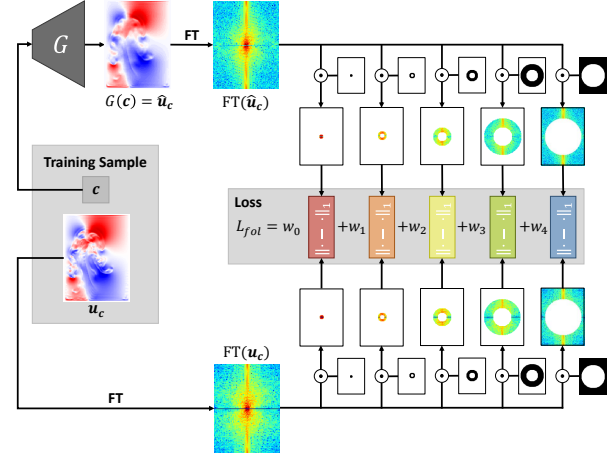

Figure 1: *Illustration of the network architecture using a Fourier norm loss with circular bands for fluid flow reconstruction.*

convergence of the phase loss and argue that magnitude loss does not provide enough information for accurate reconstructions.

Defining weights w_b is a non-trivial process, as we cannot simply assign higher weights to higher bands. We consider the following observations: 1) Low frequency bands are visually very important as they define the global flow structure, and we want to retain the high reconstruction quality of the baseline approach in these bands. 2) Very high frequency bands consist of much noise, which negatively affects the overall convergence of the optimization. 3) Bands up to 30 seem to capture most of the visually relevant details. This is illustrated in Figure 2, where we have removed bands higher than 30 and visually compared the filtered data with the original one. We therefore reduce the influence of the higher bands, following the intuition that this will eventually lead to a better reconstruction of mid-level frequencies (around bands 10-30). We use a simple heuristic that shifts the weight from the highest frequencies towards lower ones, referred as shift-towards-low (STL). The weight $w_b = \gamma p_b / n$ is inversely proportional to the number of pixels p the band b is covering, with n being the total number of pixels and $\gamma = 2$ for all but the highest band where $\gamma = 0.5$.
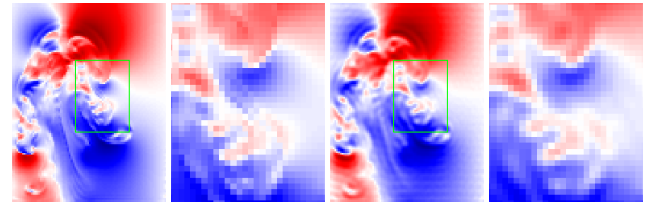

Figure 2: *Original (left) and filtered data up to band 30 (right).*

3. Evaluation and Results

3.1. MRE and Parameter Search

For a quantitative analysis we use the mean relative error (MRE) of the norm that is defined as the sum of differences in com-

plex norm between the reconstructed and the ground truth samples, normalized by the sum of magnitudes of the ground truth.

The inset image shows the MRE for different bands for the baseline model [KAT*19] (black) and the frequency-aware STL (red). It can be seen that the largest improvement of up to 10% is achieved for mid-frequency bands (15-25), while for low and high frequencies the resulting MRE is similar to the baseline. We performed an extensive grid and random search around the parameters of STL. In the grid search each parameter was set to 50%, 100%, 150%, and 200% of its STL value, except for the highest band which did not include 150%. For the random search we sampled from the standard normal distribution for each parameter. If the resulting value was negative we divided the original STL value by the magnitude, otherwise we multiplied. For better comparison we normalized the final list of values for each run. The inset image shows the resulting MRE plots for 1482 runs (768 grid search, 714 random search). Although we tested a wide range of different parameter combinations, none of them seem to significantly outperform STL over the entire range of frequencies.

## 3.2. Reconstruction Quality

The lower error in the mid-frequency bands improves the reconstructed velocity field as shown in Figure 3 (from left to right: baseline, STL and ground truth) and accordingly the resulting density field that is advected with the reconstructed velocity in Figure 4. Although some fine structures are better captured with STL than with the baseline, there is still a discrepancy between the reconstructions and the ground truth.

3.3. Comparison with GAN

We compared our approach with a generative adversarial network (GAN), as GANs are known to perform well for reconstructing details in images. We used pix2pix [IZZE16] in our implementation. Figure 5 shows that the GAN can generate impressively detailed flow structures that resemble the ground truth data. However, these generated structures only mimic flow details and in fact are less physical than the corresponding results with the baseline or STL. This is also reflected by the higher MRE of GANs compared to the baseline, which is shown in Figure 6 (bottom). Moreover, we observed that GANs are more sensitive to parameters and have an increased training time compared to our STL approach, which is comparable to the baseline performance.

3.4. Histogram Analysis

When investigating the histogram of the log-magnitudes of the Fourier transform in Figure 6 (top left), one can notice the ability of the GAN training to better match the underlying distribution since the discriminator learns to distinguish features that are not present
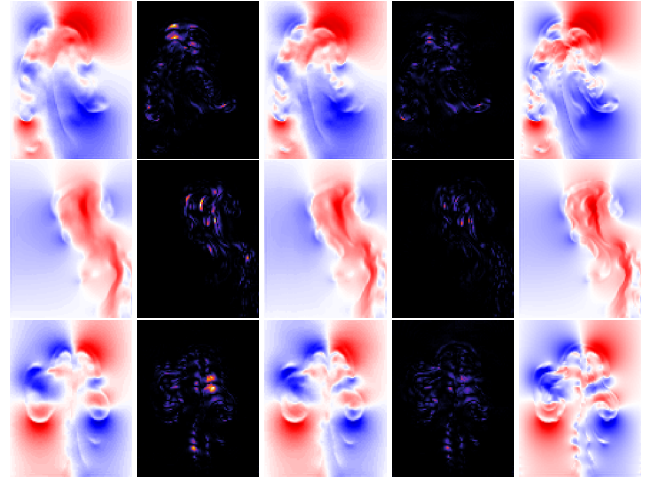
Figure 3: *Velocity fields of the baseline (left), STL (middle), and ground truth (right), and corresponding visualized squared difference to the ground truth.*
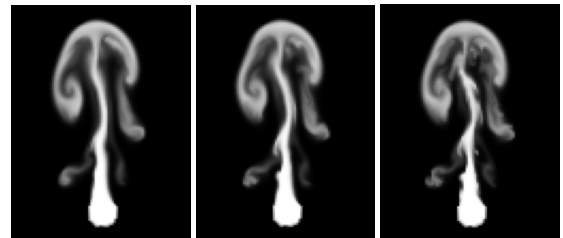
Figure 4: *Densities computed with the baseline (left), STL (middle), and ground truth (right).*

in the generated data. This is also reflected in the relative standard deviation of the magnitude of the reconstructed values, where the generator is encouraged to learn the standard deviations for each band, since these values would otherwise give the discriminator an easy advantage. In contrast, STL moves the curve only slightly towards the ground truth distribution. Due to this observation, we also evaluated a histogram loss [RWB17], but we found large discrepancies between the resulting velocity field and ground truth, even if histograms better match. We presume that matching a global histogram of the magnitudes does not capture the characteristics of flow data well, and also that the interplay between magnitude and phase might need to be considered in the optimization.

4. Conclusion

Our evaluation has shown that the baseline method does not minimize the error efficiently for higher frequency bands of the input data, and that a loss function is needed that can better discriminate between low and high frequencies of the input. Results indicate that the inclusion of spectral approaches is a promising direction to improve the reconstruction quality in mid-frequency bands. However, reconstructions still deviate from the ground truth. More research is needed to evaluate the shape and number of bands used in the
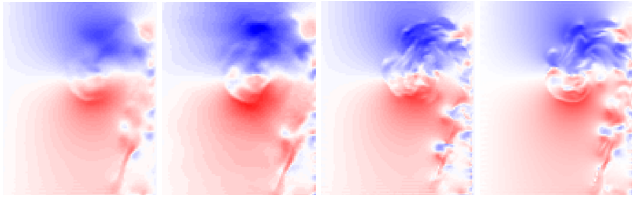
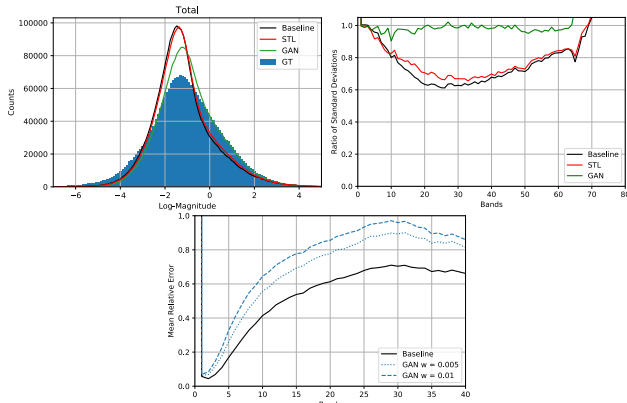**Figure 5:** *From left to right: Baseline, STL, GAN and ground truth.*

Figure 6: *Top left: Histogram of the log-magnitudes for different training methods and the ground truth distribution. Top right: Relative standard deviation of the magnitude of the reconstructed values. Bottom: MRE of the baseline (black) and GAN (blue).*

Fourier loss to develop a definite conclusion on the strengths and weaknesses of the proposed loss. Future work could also include the evaluation of perceptual losses and wavelet approaches. An interesting finding is also that generative adversarial networks can better approximate the ground truth histogram distribution of the Fourier transform's log-magnitudes and impressively improve the perceived quality. However, the reconstructed flow is non-physical, leading to higher MREs in all bands compared to the baseline and preventing applications related to flow data compression. Generally, metrics that allow a better evaluation of numerical results quantitatively and qualitatively [UHWT19] are needed.

5. Acknowledgments

References

[BSM17] BERTHELOT D., SCHUMM T., METZ L.: BEGAN: Boundary Equilibrium Generative Adversarial Networks. arXiv:1703.10717. 1

[GCZ*18] GENG C., CHEN L., ZHANG X., ZHOU P., GAO Z.: A Wavelet-based Learning for Face Hallucination with Loop Architecture. In *Visual Communications and Image Processing* (2018), pp. 1–4. 1

[GMVM17] GUO T., MOUSAVI H. S., VU T. H., MONGA V.: Deep Wavelet Prediction for Image Super-Resolution. In *Computer Vision and Pattern Recognition Workshops* (2017), IEEE, pp. 1100–1109. 1

[GPAM*14] GOODFELLOW I., POUGET-ABADIE J., MIRZA M., XU B., WARDE-FARLEY D., OZAIR S., COURVILLE A., BENGIO Y.: Generative adversarial nets. In *NIPS 27*. 2014, pp. 2672–2680. 1

[GX16] GAO X., XIONG H.: A hybrid wavelet convolution network with sparse-coding for image super-resolution. In *International Conference on Image Processing* (2016), pp. 1439–1443. 1

[HHST17] HUANG H., HE R., SUN Z., TAN T.: Wavelet-SRNet: A Wavelet-Based CNN for Multi-scale Face Super Resolution. In *International Conference on Computer Vision* (2017). 1

[HHST19] HUANG H., HE R., SUN Z., TAN T.: Wavelet Domain Generative Adversarial Network for Multi-scale Face Hallucination. *International Journal of Computer Vision 127*, 6-7 (2019), 763–784. 1

[Hor91] HORNIK K.: Approximation capabilities of multilayer feedforward networks. *Neural Networks 4*, 2 (1991), 251–257. 1

[IZZE16] ISOLA P., ZHU J.-Y., ZHOU T., EFROS A. A.: Image-to-image translation with conditional adversarial networks. *Computer Vision and Pattern Recognition (CVPR)* (2016), 5967–5976. 3

[JAFF16] JOHNSON J., ALAHI A., FEI-FEI L.: Perceptual losses for real-time style transfer and super-resolution. In *European Conference on Computer Vision* (2016). 1

[KALL18] KARRAS T., AILA T., LAINE S., LEHTINEN J.: Progressive Growing of GANs for Improved Quality, Stability, and Variation. *ICLR* (2018). 1

[KAT*19] KIM B., AZEVEDO V. C., THUEREY N., KIM T., GROSS M., SOLENTHALER B.: Deep Fluids: A Generative Network for Parameterized Fluid Simulations. *Computer Graphics Forum (Proceedings of Eurographics) 38*, 2 (2019), 59–70. 1, 2, 3

[MDM*18] MEYER S., DJELOUAH A., MCWILLIAMS B., SORKINE-HORNUNG A., GROSS M., SCHROERS C.: PhaseNet for Video Frame Interpolation. In *Computer Vision and Pattern Recognition* (2018). 2

[RBA*19] RAHAMAN N., BARATIN A., ARPIT D., DRAXLER F., LIN M., HAMPRECHT F., BENGIO Y., COURVILLE A.: On the spectral bias of neural networks. In *Proceedings of Machine Learning Research* (2019), vol. 97, pp. 5301–5310. 1

[RWB17] RISSER E., WILMOT P., BARNES C.: Stable and Controllable Neural Texture Synthesis and Style Transfer Using Histogram Losses. arXiv:1701.08893. 3

[TSSP17] TOMPSON J., SCHLACHTER K., SPRECHMANN P., PERLIN K.: Accelerating eulerian fluid simulation with convolutional networks. In *Proceedings of the 34th International Conference on Machine Learning - Volume 70* (2017), ICML'17, JMLR.org, p. 3424–3433. 1

[UHWT19] UM K., HU X., WANG B., THUEREY N.: Spot the difference: Accuracy of numerical simulations via the human visual system, 2019. arXiv:1907.04179. 4

[WB09] WANG Z., BOVIK A. C.: Mean squared error: Love it or leave it? A new look at Signal Fidelity Measures. *IEEE Signal Processing Magazine 26*, 1 (2009), 98–117. 1

[WBT19] WIEWEL S., BECHER M., THUEREY N.: Latent-space Physics: Towards Learning the Temporal Evolution of Fluid Flow. *Computer Graphics Forum (Proceedings of Eurographics) 38*, 2 (2019). 1

[XZL*19] XU Z.-Q. J., ZHANG Y., LUO T., XIAO Y., MA Z.: Frequency Principle: Fourier Analysis Sheds Light on Deep Neural Networks. arXiv:1901.06523. 1

[XZX19] XU Z.-Q. J., ZHANG Y., XIAO Y.: Training behavior of deep neural network in frequency domain. In *Neural Information Processing* (Cham, 2019), Gedeon T., Wong K. W., Lee M., (Eds.), Springer International Publishing, pp. 264–274. 1

[ZSY*18] ZHONG Z., SHEN T., YANG Y., LIN Z., ZHANG C.: Joint sub-bands learning with clique structures for wavelet domain super-resolution. In *NIPS 31*. 2018, pp. 165–175. 1

[ZWY19] ZHANG Q., WANG H., YANG S.: Image Super-Resolution Using a Wavelet-based Generative Adversarial Network. arXiv:1907.10213. 1