Interpretable and Learnable Super-Resolution Time-Frequency Representation

Randall Balestriero* Hervé Glotin+ Richard Baraniuk*

*Rice University, Houston, TX and +Toulon University, Toulon, FR

Deep Learning Deployment Remains Challenging

The Science Advisory Board

Deep-learning approach points the way to faster COVID-19 vaccines

By Leah Sherwood, The Science Advisory Board contributing writer

February 8, 2021 -- A novel deep neural network (DNN) can target the most promising multiepitope COVID-19 vaccine candidates in a matter of seconds. The new artificial intelligence (AI) framework, which was described in *Scientific Reports* on February 5, may give scientists an edge in the race against SARS-CoV-2 and its variants by reducing the time from vaccine design to clinical trials.

Law firm teams up with Canadian legal tech company on AI-powered case prediction tool

BY LYLE MORAN

AUGUST 25, 2020, 8:00 AM CDT

Tweet

Labor and employment law firm Fisher Phillips has partnered with Blue J Legal Inc. to bring the Toronto legal tech company's AI-powered technology, which predicts court outcomes in the employment law arena, to the United States.

20 Nov 2020 | 19:10 GMT

Deep Learning Has Reinvented Quality Control in Manufacturing—but It Hasn't Gone Far Enough

AI systems that make use of "lifelong learning" techniques are more flexible and faster to train

By **Anatoli Gorchet**

Despite some popular and publicized cases Deep Learning remains difficult to deploy in production without expert knowledge **especially for audio applications**

Time-Series Datasets are Particular (Data)

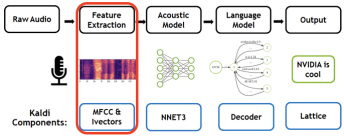
Time-series datasets often concern intricate topics e.g. geophysics, health, bioacoustic:

- **labeling requires experts $\implies$ costly and slow**
- **data collection is only performed by interested parties $\implies$ small unlabeled dataset**

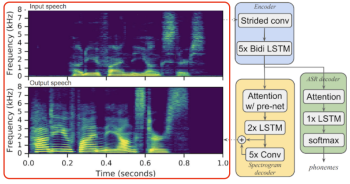
The combination of those two points prevent the user of end-to-end supervised/unsupervised/SSL methods...

Time-Frequency Representations are Common

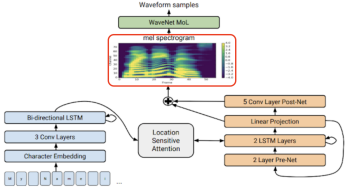
Current state-of-the-art solutions circumvent the lack of data issue through handcrafted designs:



GPU-Accelerated Speech to Text with Kaldi:A Tutorial on Getting Started

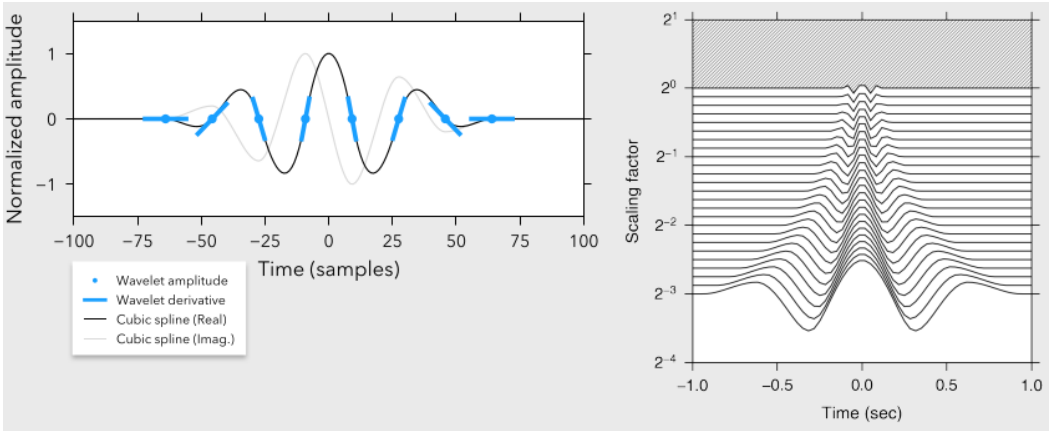


Parrotron: New Research in Improving Verbal Communication for People with Speech Impairments



Tacotron 2: Generating Human-like Speech from Text

Learnable Wavelet Transform Example



[R. Balestriero, R. Cosentino, H. Glotin, R. Baraniuk, Spline Filters for End-To-End Deep Learning, ICML18]
[R. Balestriero, H. Glotin, Wavelet Learning by Adaptive Hermite Cubic Splines applied to Bioacoustic Chirps, IEEE]

Universal Learnable Time-Frequency Representation

- Bring the input $\boldsymbol{x}$ into a different space (Wigner-Ville)

$$\mathrm{WV}_x(t,f) = \int_{-\infty}^{\infty} x\left(t-\frac{\tau}{2}\right) x^*\left(t+\frac{\tau}{2}\right) e^{-if\tau} d\tau.$$

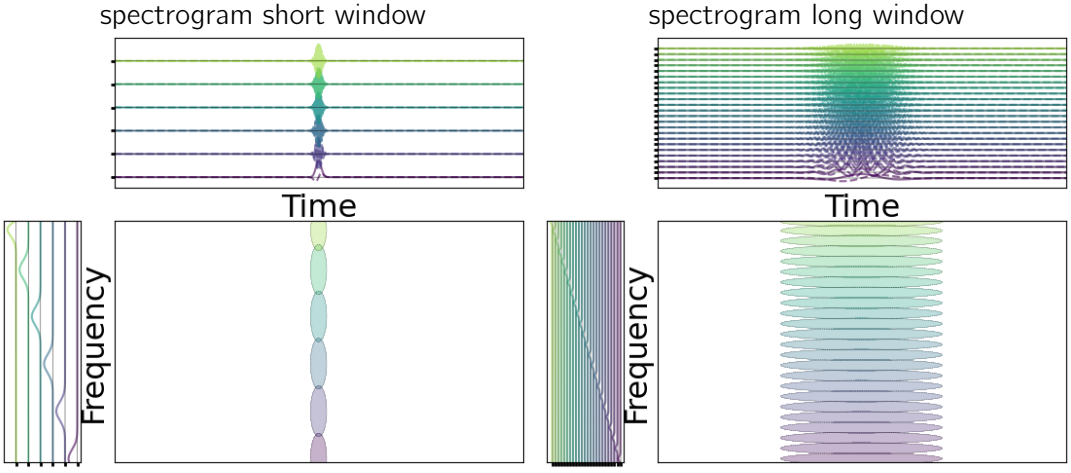
- Learn a (time) Gaussian filter Φ_f for each frequency f to obtain

$$\underbrace{\mathrm{K}_x(t,f)}_{\text{learned repr.}} = \int_{\mathbb{R}\times[0,2\pi)} \mathrm{WV}_x(\tau,\omega)\Phi_f(t-\tau,\omega)d\tau d\omega,$$

different time-frequency families correspond to different kernels Φ_f

The above provides a **universal** , **learnable** and **interpretable** time-frequency representation with only 4 degrees of freedom per filter removing the supervision of the filter-bank family

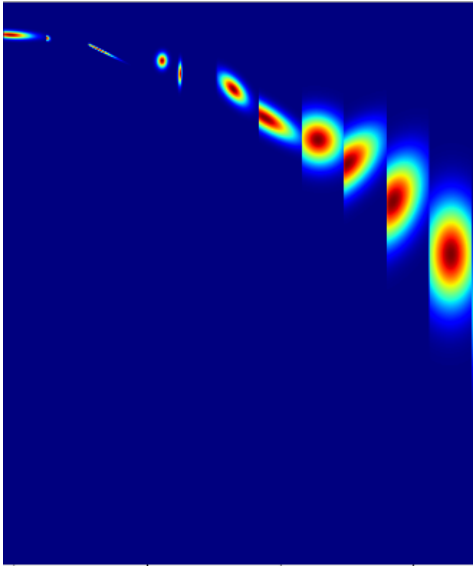
Different Representations Have Different Properties



Universal Learnable Time-Frequency Representation: Experiments

		Linear Scattering				Nonlinear Scattering				Linear Joint Scattering			
	le. rate	*morlet*	**lwvd**	sinc	lmorlet	*morlet*	**lwvd**	sinc	lmorlet	*morlet*	**lwvd**	sinc	lmorlet
docc10	0.0002	14.3	63	31.1	29.7	54.1	84.7	74.4	74.9	70.7	**83.7**	82.4	75.8
	0.001	12.7	65.5	26.0	28.3	50.1	**87.9**	77.4	77.4	70.1	80.6	80.8	73.2
	0.005	13.0	**65.9**	17.1	27.0	51.8	87.1	43.3	83.2	65.9	78.0	70.5	80.8
Bird	0.0002	63.8	77.9	69.6	65.4	84.7	92.9	88.1	85.8	82.1	**90.5**	87.2	84.3
	0.001	65.0	80.0	67.2	64.3	85.0	**94.2**	88.1	86.6	80.3	88.7	86.8	83.1
	0.005	65.2	**80.4**	67.3	66.9	84.8	94.2	86.0	87.2	78.1	87.5	78.3	82.8
MNIST	0.0002	43.9	68.4	52.2	44.0	82.3	85.3	10.4	83.0	95.3	97.6	22.1	95.4
	0.001	41.5	**68.8**	43.5	42.2	83.2	**89.8**	87.1	85.4	89.7	**97.8**	93.2	90.4
	0.005	34.6	68.8	23.9	36.0	82.7	22.1	68.7	88.1	81.1	12	64.4	80.2
comm.	0.0002	8.1	24.9	9.5	7.6	33.9	38.2	36.2	33.4	65.8	**76.7**	3.7	66.8
	0.001	7.5	**26.1**	8.0	8.2	33.5	**42.9**	35.5	33.7	53.6	71.8	27.9	51.9
	0.005	7.3	25.7	6.2	6.5	33.0	17.0	28.9	34.8	32.1	35.2	17.2	32.9
fsd	0.0002	9.7	15.3	10.3	9.0	22.9	23.1	2.3	27.9	40.1	38.8	1.6	42.0
	0.001	9.8	16.7	10.4	10.6	24.2	27.4	13.1	**31.1**	38.9	**44.9**	2.1	42.3
	0.005	9.0	**17.4**	5.5	10.2	24.2	28.8	16.9	30.4	25.0	31.5	17.0	33.2

Learned Filters and How to Interpret Them



Conclusions

- Learning from unprocessed time-series has been the lifelong challenge in Deep Learning
- With the correct internal components time-frequency representation learning as we proposed provides a universal solution across domains
- There are many possible extensions to this work such as unsupervised learning based on Entropy minimization

Thank you for your attention

(contact: randallbalestriero@gmail.com)