learning are two extremes. Weakly supervised learning occurs most frequently in our daily applications. The design and analysis of a weakly supervised learning mechanism based on CNNs is interesting and practical.

The interpretation of a CNN as a guided multilayer RECOS transform should be valuable to the investigation of these topics.

Summary

The operating principle of CNNs was explained as a guided multilayer RECOS transform in this note. A couple of illustrative examples were provided to support this claim. Several known facts were interpreted accordingly, and some open issues were pointed out at the end.

Acknowledgments

Author

C.-C. Jay Kuo (cckuo@ee.usc.edu) is a professor of electrical engineering at the University of Southern California, Los Angeles.

References

[1] C.-C. J. Kuo, "Understanding convolutional neural networks with a mathematical model," *J. Vis. Commun. Image Represent.*, vol. 41, pp. 406–413, Nov. 2016.

[2] W. S. McCulloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity," *Bull. Math. Biophys.*, vol. 5, no. 4, pp. 115–133, 1943.

[3] F. Rosenblatt, "The perceptron: A probabilistic model for information storage and organization in the brain," *Psychol. Rev.*, vol. 65, no. 6, pp. 386–408, 1957.

[4] G. Cybenko, "Approximation by superpositions of a sigmoidal function," *Math. Control Signals Syst.*, vol. 2, no. 4, pp. 303–314, 1989.

[5] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," *Neural Netw.*, vol. 2, no. 5, pp. 359–366, 1989.

[6] K. Fukushima, "Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position," *Biol. Cybern.*, vol. 36, pp. 193–202, 1980.

[7] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.

[8] A. K. Jain, "Data clustering: 50 years beyond k-means," *Pattern Recogn. Lett.*, vol. 31, no. 8, pp. 651–666, 2010.

Rodrigo Capobianco Guido

Effectively Interpreting Discrete Wavelet Transformed Signals

Following two decades of research focusing on the discrete wavelet transform (DWT) and driven by students' high level of questioning, I decided to write this essay on one of the most significant tools for time-frequency signal analysis. As it is widely applicable in a variety of fields, I invite readers to follow this lecture note, which is specially dedicated to show a practical strategy for the interpretation of DWT-based transformed signals while extracting useful information from them. The particular focus resides on the procedure used to find the time support of frequencies and how it is influenced by the wavelet family and the support size of corresponding filters.

Relevance

Frequently, DWT computation is much faster than that of the discrete Fourier transform (DFT) and even the fast Fourier transform (FFT), encouraging its usage. Furthermore and opposite to the DFT, to the FFT, and even to the short-time Fourier transform, the DWT reveals the time support of frequencies efficiently, as described in [1]. Thus, its study is of paramount importance.

Prerequisites

On one hand, essential knowledge of digital filters and wavelet families is only desirable. On the other hand, the basic aspects involving the procedures used to calculate DWTs from discrete-time signals are imperative for the full comprehension of this lecture note. If readers are not comfortable with the topic, I encourage you to consult [1]–[3], written in a friendly manner, before proceeding any further.

Problem statement and solution

Problem statement

Given the pair of signals $s[\cdot]$ and $y[\cdot]$, both of size M equal to a power of two and indexed from 0 to $M-1$, being the former and the latter the discrete-time input and its DWT, respectively, the problem is to extract and interpret

relevant information the latter contains about the former, particularly allowing the time support of frequencies to be accurately found.

Solution

Differently from the DFT, which converts a signal from the time to the frequency domain, the DWT *filters* it. Thus, significant amplitudes in $y[\cdot]$ come from frequencies that originally exist in $s[\cdot]$ and were not removed by the filters. Based on this process, our solution consists of a DWT-based filtering applied to $s[\cdot]$ followed by an inspection of the dominant amplitudes contained in the transformed signal $y[\cdot]$.

The practical procedure used to obtain $y[\cdot]$ from $s[\cdot]$ has already been described in detail in [2]; thus, it is not repeated at this time. Nevertheless, there are two main issues to be addressed to solve the stated problem: the ideal level of transformation, i.e., j, and the appropriate analysis filter pair, i.e., the low-pass filter $h[\cdot]$ and the high-pass mirrored finite impulse response half-band filter $g[\cdot]$ [4]. Once j is defined, the generalized and well-known DWT decomposition tree [4] will be more or less accurate depending on the order $N-1$ of those filters, N being their length, and on the family they belong to, such as Haar, Daubechies, Coiflets, Symmlets, and so on [4].

Notably, spectral description and temporal localization hold distinct requirements for ideality. On one hand, the former offers the finest resolution, whenever j and N are maximized and the associated filters exhibit linear-phase and maximally flat responses at their pass and stop bands. On the other hand, the latter is foolproof whenever j and N are minimized because each sample of $y[\cdot]$ refers just to a small part of $s[\cdot]$. Furthermore, according to Heisenberg's uncertainty principle [4], time and frequency information are antagonistic. The better that is, the worse this is, and vice versa. Considering the limits imposed by the real conditions, we may either balance time and frequency information or prioritize one of them based on our specific needs.

Disregarding temporal information, the finest spectral resolution is guaran-teed whenever the deepest decomposition level is adopted, i.e., $j = \log_2(M)$. Additionally, the maximum possible length for N, for which there is no upper bound, is required. In practice, it usually implies filters for which $N \geq 40$. Maximally flat responses at their pass and stop bands, such as those provided by Daubechies' filters, may be used to avoid improper floatations in $y[\cdot]$ due to a reduced or an excessive gain at some subbands [4], causing inaccurate frequency magnitudes. Furthermore, distinct delays for different frequency bands are also present, whenever the phase responses of the filters are not linear; thus, Symmlets and Coiflets [4], which exhibit almost linear responses, are also interesting choices. Contrary to this, the best temporal resolution, which causes the poorest spectral description, requires the lowest decomposition level, i.e., $j = 1$, and the smallest support size of filters, i.e., $N = 2$, directly implying in the adoption of Haar's filters [4].

Opposed to an extreme time or frequency resolution, equilibrium is reached whenever the intermediary decomposition level is chosen, i.e., the situation in which j is the mean between the minimum, one, and the maximum, $\log_2(M)$. This results in $j = \lfloor (1 + \log_2(M))/2 \rfloor$, where $\lfloor \cdot \rfloor$ is the floor operator, and implies that the corresponding frequency resolution is $r = T/2^j = (T/2^{\lfloor (1 + \log_2(M))/2 \rfloor})$ Hz, where $2T$ is the sampling rate at which $s[\cdot]$ was digitalized and, according to Nyquist's theorem [4], T is its maximum frequency content. Complementarily, a balanced time and frequency accuracy also requires an intermediary value for N, for which two is the lower bound and there exists no upper bound. To circumvent the missing bound, a careful inspection of the procedure used to calculate DWTs, as described in [2], is useful. It allows the statement that at the jth-level transformation based on a filter of support size N, information from $(N-1)(2^j - 1) + 1$ subsequent samples of $s[\cdot]$ is grouped together, directly influencing the accurate time support of frequencies.

The final issue to be solved, before formulating a solution to the stated problem, involves the specific map to be adopted: the regular DWT map or the DWT-packet map. On one hand, the former only provides the finest resolution for the lowest frequencies, which are usually those that carry most of the useful information. On the other hand, the latter provides a uniform and equally distributed time-frequency resolution for all subbands, once j, N, and the wavelet family are defined. At each decomposition level of the former map, the transformation produces two half-size and half-band signals—trend and fluctuation, respectively—containing the low and high frequencies of the input signal that are subsequently concatenated to establish the transformed signal. Contrastingly, the latter map contains 2^j subband signals of size $(M/2^j)$ at the jth decomposition level, as detailed in [5]. Figure 1 helps to recall these schemes.

Based both on the previous notes and Mallat's algorithm [2], the use of a regular DWT map, as exemplified in Figure 1(a), requires strategy STG_A-(i) and (ii) to solve our problem, assuming an ideal situation:

i) The energy of the ith sample of the jth-level trend contains the amplitude of frequencies between 0 and $(T/2^j)$ Hz, which are located within the range $\{s_{i2^j}, s_{(i+1)2^j - 1}\}$, for $i = \{0, 1, 2, ..., (M/2^j) - 1\}$.

ii) The energy of the ith sample of the jth-level fluctuation contains the amplitude of frequencies between $(T/2^j)$ and $(T/2^{j-1})$ Hz, which are located within the range $\{s_{i2^j}, s_{(i+1)2^j - 1}\}$.

Accordingly, the DWT-packet map built considering the natural frequency ordering (NFO) [5, p. 111], which is exemplified in Figure 1(b), is associated with the strategy STG_B:

The energy of the ith sample of the bth subband at the jth-level contains the amplitude of frequencies between $(bT/2^j)$ and $((b+1)T/2^j)$ Hz, which are located within the range $\{s_{i2^j}, s_{(i+1)2^j - 1}\}$, for $b = \{0, 1, 2, ..., 2^j - 1\}$ and $i = \{0, 1, 2, ..., (M/2^j) - 1\}$.

To lessen the imprecise temporal localization caused by the influence of N, correction COR is required right after calculating the DWT but prior to using those strategies:
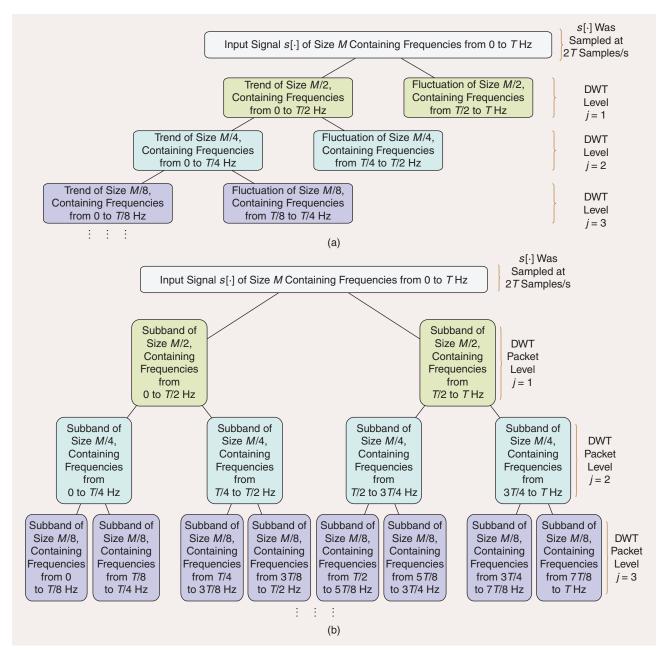
FIGURE 1. (a) The traditional DWT decomposition tree, exemplified for $j = 3$, where the $(M/8) + (M/8) + (M/4) + (M/2) = M$-sample long signal $y[\cdot]$ corresponds to the concatenation of all the leafs, considering a left-to-right and bottom-up procedure. (b) The corresponding DWT-packet tree, built based on the NFO.

Apply individually, for each leaf subband contained in the wavelet decomposition tree at the level j, regardless of coming from the regular DWT map or the DWT-packet map, a circular right-sided shift of $(N-2)/2$ samples.

COR comes from the fact that the convolution between $s[\cdot]$ and a wavelet filter produces a signal for which the length corresponds to $M+N-1$, primarily. Then, once the downsampling by two is applied [2], this length is reduced to $(M+N-1)/2 = (M/2) + (N-1)/2$. Finally, the wraparound procedure [2] vanishes the effect of the term $(N-1)/2$, remaining only a $(M/2)$-sample-long resulting signal. Thus, even though N has no influence over the length of the transformed signal, it does cause shifts. Particularly, due to the fact that N is usually even and half-shifts are not defined in discrete wavelet systems, $(N-1)/2$ can be treated as being

$(N-2)/2$, implying that all the filters but Haar's cause a shifting. Therefore, COR consists of a circular right-sided shift of $(N-2)/2$ samples, regardless of j. Notably, COR does not correct delays originated from nonlinear phase responses of the filters, which can be attenuated with the use of specific wavelet families, as aforementioned.

Finally, a refined time-frequency description based on the previously defined strategies requires the use of
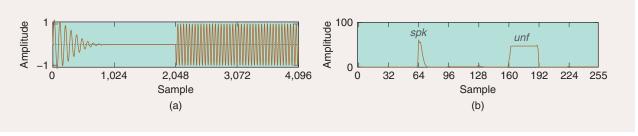
FIGURE 2. The original and transformed data for the example in the section "Numerical Example." (a) The original signal under analysis ($s[\cdot]$) and (b) the first four subbands of the corresponding DWT-packet map, modified and converted based on COR and TEG, respectively ($\tilde{y}[\cdot]$).

Teager's energy operator [6], instead of the ordinary energy [7], to convert all the samples of each leaf subband into their instantaneous potentials, according to the TEG procedure:

For the subband $b = 0, 1, ..., 2^j - 1$ of the deepest level j, do:

- $\tilde{y}_{0,b} \leftarrow (y_{0,b})^2$
- $\tilde{y}_{i,b} \leftarrow (y_{i,b})^2 - (y_{i-1,b}) \cdot (y_{i+1,b})$, for $(1 \leq i \leq (M/2^j) - 2)$, whenever the length of b is at least four
- $\tilde{y}_{(M/2^j)-1,b} \leftarrow (y_{(M/2^j)-1,b})^2$, whenever the length of b is at least two

where the left arrows mean attribution. Summarizing, the ideal time-frequency analysis of $s[\cdot]$ based on the DWT requires four steps:

- **Step 1**: Select the most appropriate wavelet family and the support size of filters, i.e., N. Then, calculate the M-sample-long transformed signal $y[\cdot]$ from $s[\cdot]$ based either on the regular DWT map or the DWT-packet map.
- **Step 2**: Apply COR, shifting $y[\cdot]$ accordingly.
- **Step 3**: Apply TEG, creating $\tilde{y}[\cdot]$ from shifted $y[\cdot]$.
- **Step 4**: Interpret $\tilde{y}[\cdot]$ based either on STG_A or STG_B, depending on which map was chosen to carry out Step 1.

Numerical example

Problem statement

Let $s_i = 1.2 \sin((2\pi \cdot 44i)/4095) e^{-(\frac{i}{409.6})^2}$ for $\{0 \leq i < 1,024\}$, $s_i = 0$, for $\{1,024 \leq i < 2,048\}$, and $s_i = \sin((2\pi \cdot 85i/4,095)$ for $\{2,048 \leq i < 4,096\}$ to be the 1-s long signal under analysis, sampled at $2T = 4,096$ samples/s, as shown in Figure 2(a). The problem is to find out the time support of its frequencies with

a balance between temporal and spectral accuracies.

Solution

A balanced time-frequency analysis of $s[\cdot]$, which contains frequencies up to $T = 2,048$ Hz, requires $j = \lfloor (1 + \log_2(M))/2 \rfloor = \lfloor (1 + \log_2(4,096))/2 \rfloor = 6$. Correspondingly, $r = (T/2^j) = (2,048/2^6) = 32$ Hz is the frequency resolution for this decomposition, which is based on the DWT-packet map to allow a uniform analysis, i.e., to keep r intact for all the subbands. Symmlets of size $N = 16$ were chosen to perform the analysis because of their almost linear phase responses, practically avoiding frequency components to become misaligned.

Proceeding with the calculations and applying COR and TEG, we get $\tilde{y}[\cdot]$, for which only the first four subbands appear in Figure 2(b) to allow magnification. Its inspection reveals two main elements: *spk* and *unf*. The former is a spike decaying from samples 64 to 79, and the latter consists of a uniform step from samples 160 to 191. Observing that each one of the $2^j = 2^6 = 64$ subbands contains $(M/2^j) = (4096/2^6) = 64$ samples, indexed from 0 to 63, we conclude that *spk* occurs within the interval $\{i = 0, i = 15\}$ of the second subband, i.e., $b = 1$, and *unf* occurs within the interval $\{i = 32, i = 63\}$ of the third subband, i.e., $b = 2$. Thus, according to STG_B:

- For *spk*, the energy of the sample $i = 0$ of the subband $b = 1$ at the level $j = 6$ contains the amplitude of frequencies between $(1 \cdot 2,048/2^6) = 32$ and $((1 + 1) \cdot 2,048)/2^6 = 64$ Hz, located within the range $\{s_{0 \cdot 2^6}, s_{(0+1)2^6-1}\}$, i.e., $\{s_0, s_{63}\}$. Accordingly, the energy of the sample $i = 15$ of that subband contains the ampli-

tude of those frequencies, which are located within the range $\{s_{15 \cdot 2^6}, s_{(15+1)2^6-1}\}$, i.e., $\{s_{960}, s_{1023}\}$. It means that a frequency between 32 and 64 Hz exists in $s[\cdot]$, starting and ending somewhere between $\{s_0, s_{63}\}$ and $\{s_{960}, s_{1023}\}$, respectively. In fact, the frequency of 44 Hz was defined in $s[\cdot]$ within the interval $\{s_0, s_{1023}\}$. Furthermore, the decay starting from sample 64 of Figure 2(b) shows the corresponding power reduction clearly seen within the corresponding interval of Figure 2(a).

- For *unf*, the energy of the sample $i = 32$ of the subband $b = 2$ at the level $j = 6$ contains the amplitude of frequencies between $2 \cdot 2048/2^6 = 64$ and $((2+1) \cdot 2048)/2^6 = 96$ Hz, located within the range $\{s_{32 \cdot 2^6}, s_{(32+1)2^6-1}\}$, i.e., $\{s_{2048}, s_{2111}\}$. Complementarily, the energy of the sample $i = 63$ of that subband contains the amplitude of those frequencies, which were placed within the range $\{s_{63 \cdot 2^6}, s_{(63+1)2^6-1}\}$, i.e., $\{s_{4032}, s_{4095}\}$. It means that a frequency between 64 and 96 Hz exists in $s[\cdot]$, starting and ending somewhere between $\{s_{2048}, s_{2111}\}$ and $\{s_{4032}, s_{4095}\}$, respectively. In fact, the frequency of 85 Hz was uniformly defined in $s[\cdot]$ within the interval $\{s_{2048}, s_{4095}\}$.

What we have learned

Based on this article, readers could have learned how to efficiently interpret discrete wavelet-transformed signals, while extracting from them the time support of frequencies contained in the original signal under analysis. Supplementary to this, the criteria used

the main emphasis on cardiac vector velocity estimation.

References

[1] D. H. Evans and W. N. McDicken, *Doppler Ultrasound, Physics, Instrumentation, and Signal Processing.* New York: Wiley, 2000.

[2] D. W. Baker, "Pulsed ultrasonic Doppler blood-flow sensing," *IEEE Trans. Son. Ultrason.*, vol. SU-17, no. 3, pp. 170–185, 1970.

[3] O. Bonnefous, P. Pesqué, and X. Bernard, "A new velocity estimator for color flow mapping," in *Proc. IEEE Ultrasonics Symp.*, 1986, pp. 855–860.

[4] C. Kasai, K. Namekawa, A. Koyano, and R. Omoto, "Real-time two-dimensional blood flow imaging using an autocorrelation technique," *IEEE Trans. Son. Ultrason.*, vol. 32, no. 3, pp. 458–463, 1985.

[5] B. Dunmire and K. W. Beach, "A brief history of vector Doppler," in *Proc. SPIE Medical Imaging*, Feb. 2001, vol. 4325, pp. 200–214.

[6] G. E. Trahey, J. W. Allison, and O. T. von Ramm, "Angle independent ultrasonic detection of blood flow," *IEEE Trans. Biomed. Eng.*, vol. 34, no. 12, pp. 965–967, 1987.

[7] J. A. Jensen and P. Munk, "A new method for estimation of velocity vectors," *IEEE Trans. Ultrason., Ferroelec., Freq. Contr.*, vol. 45, no. 3, pp. 837–851, 1998.

[8] S. I. Nikolov and J. A. Jensen, "In-vivo synthetic aperture flow imaging in medical ultrasound," *IEEE Trans. Ultrason., Ferroelec., Freq. Contr.*, vol. 50, no. 7, pp. 848–856, 2003.

[9] J. Udesen, F. Gran, K. L. Hansen, J. A. Jensen, C. Thomsen, and M. B. Nielsen, "High frame-rate blood vector velocity imaging using plane waves: Simulations and preliminary experiments," *IEEE Trans. Ultrason., Ferroelec., Freq. Contr.*, vol. 55, no. 8, pp. 1729–1743, 2008.

[10] J. Bercoff, G. Montaldo, T. Loupas, D. Savery, F. Meziere, M. Fink, and M. Tanter, "Ultrafast compound Doppler imaging: Providing full blood flow characterization," *IEEE Trans. Ultrason., Ferroelec., Freq. Contr.*, vol. 58, no. 1, pp. 134–147, Jan. 2011.

[11] J. A. Jensen, S. Nikolov, K. L. Gammelmark, and M. H. Pedersen, "Synthetic aperture ultrasound imaging," *Ultrasonics*, vol. 44, pp. e5–e15, 2006.

[12] M. Tanter and M. Fink, "Ultrafast imaging in biomedical ultrasound," *IEEE Trans. Ultrason., Ferroelec., Freq. Contr.*, vol. 61, no. 1, pp. 102–119, Jan. 2014.

SP

SP EDUCATION *(continued from page 80)*

summer program for high school students," *Int. J. Eng. Educat.*, vol. 31, no. 5, pp. 1419–1427, 2015.

[71] M. Bugallo and A. M. Kelly, "A pre-college recruitment strategy for electrical and computer engineering study," in *Proc. Integrated STEM Education Conf. (ISEC)*, 2014, pp. 1–4.

[72] M. F. Bugallo and A. M. Kelly, "An outreach after-school program to introduce high school students to electrical engineering," in *Proc. Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP)*, 2015, pp. 5540–5544.

[73] R. Gearns, A. M. Kelly, and M. F. Bugallo, "Shifts in students' views towards engineering in an out-of-school-time program," in *Proc. Annu. Conf. National Association of Research in Science Teaching (NARST)*, 2017.

[74] L. Padwa and K. Sheppard, "Real world chemistry: A model for secondary chemistry teacher professional development," in *Proc. 22nd Biennial Conf. Chemistry Education*, 2012.

[75] American Society for Engineering Education. (2014). Standards for preparation and professional development for teachers of engineering [Online]. Available: https://www.asee.org/conferences-and-events/outreach/egfi-program/k12-teacher-professional-development

[76] J. Kiely, A. M. Kelly, K. La Magna, D. J. Moloney, and R. D. Bynum, "Research on impacts of university-based biotechnology teaching laboratories on teacher professional development and student outcomes," in *Proc. National Association of Research in Science Teaching Conf.*, 2014.

[77] New York State Education Department. (2014). Physical setting/physics core curriculum [Online]. Available: http://www.p12.nysed.gov/ciai/mst/pub/phycoresci.pdf

SP

LECTURE NOTES *(continued from page 92)*

to choose the wavelet family, the transformation level, and the corresponding support size of the filters were brought to light.

Author

Rodrigo Capobianco Guido (guido@ ieee.org) is an associate professor at São Paulo State University (UNESP), São José do Rio Preto, Brazil. His home page at UNESP is available at www.sjrp.unesp .br/~guido/.

References

[1] R. C. Guido, P. Addison, and J. Walker, "Introducing wavelets and time-frequency analysis," *IEEE Eng. Biol. Med. Mag.*, vol. 28, no. 5, pp. 13, 2009.

[2] R. C. Guido, "Practical and useful tips on discrete wavelet transforms," *IEEE Signal Process. Mag.*, vol. 32, no. 3, pp. 162–166, 2015.

[3] R. C. Guido, "A note on a practical relationship between filters coefficients and the scaling and wavelet functions of the discrete wavelet transform," *Appl. Math. Lett.*, vol. 24, no. 7, pp. 1257–1259, 2011.

[4] G. Strang and T. Nguyen, *Wavelets and Filter Banks.* Wellesley, MA: Wellesley-Cambridge Press, 1997.

[5] A. Jensen and A. L. Cour-Harbo, *Ripples in Mathematics: The Discrete Wavelet Transform.* New York: Springer, 2000.

[6] J. F. Kaiser, "On a simple algorithm to calculate the energy of a signal," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, Albuquerque, NM, 1990, pp. 381–384.

[7] R. C. Guido, "A tutorial on signal energy and its applications," *Neurocomputing*, vol. 179, pp. 264–282, Feb, 2016.

SP