Robust Digital Envelope Estimation Via Geometric Properties of an Arbitrary Signal

Carlos Tarjano Valdecy Pereira

August 6, 2020

Abstract

Despite being an elusive concept, mathematically well defined only for artificially generated waves, the temporal amplitude envelope of a signal is essential for its complete characterization, being the primary information carrying medium in spoken voice and telecommunications, for example. intuitively, the temporal envelope can be understood as a smooth function on the same variable as the principal wave, or temporal fine structure, that modulates the signal, responsible for its outer shape. It is implied in this definition that the envelope is non periodic in general and is thus better addressed by specific techniques diverse from those used in the analysis of the periodic, rapidly changing wave it encompasses. Envelope detection techniques have applications in areas like health, sound classification and synthesis, seismology and speech recognition. Despite of that, a general approach to envelope detection of signals with rich spectral content seems to be lacking, and most methods involve manual intervention, like filter design, based on prior knowledge about the wave under investigation. In this paper we propose a framework that uses intrinsic characteristics of a signal to estimate its envelope, eliminating the necessity of parameter tuning. The approach here described draws inspiration from general, computational and differential geometry to isolate the frontier of an arbitrary signal. The concepts of alpha-shapes, concave hulls, arc-length parametrization and discrete curve curvature are adapted. We also define a pulse in the context of a arbitrary digital signal as a means to reduce dimensionality and the complexity of the proposed algorithm.

Keywords: DSP, alpha-shapes, envelope detection

1 Introduction

Despite being ubiquitous in digital signal processing, the literature about envelope detection is very fragmented (Lyons, 2017). Besides, most envelope detection techniques are designed to account for very specific settings, like pure sinusoids with moderate noise contents, in line with the most common usages of those algorithms with artificial signals in the context of analog telecommunications. In many natural signals, however, the temporal amplitude envelope of a signal plays a prominent role in the characteristics exhibited: According to Qi et al. (2017), for example, the envelope is at least as important as the fine structure of the soundwave in the context of the intelligibility of mandarin tones. That is also the case for English (Shannon et al., 1995), where even envelopes modulating mostly noise were still capable of conveying meaning. The envelope helps to convey emotion and identity to the human voice (Zhu et al., 2018), and envelope preserving characteristics of concert halls are associated with their pleasantness (Lokki et al., 2011). When dealing with broadband signals, approaches tailored to specific applications are prevalent, such as the one presented by Yang et al. (2014) for the distributed monitoring of fibre optic or the one formulated by Assef et al. (2018) in the context of medical ultrasound imaging.

If one adresses the different units in the horizontal and vertical axes, one can transform the DSP problem of envelope detection in the geometric problem of defining the shape of a set of points in $\mathbb{R}^2$. For that purpose Edelsbrunner et al. (1983) introduced the concept of alpha-shapes, a mathematically well defined extension to the convex hull of a finite set of points, related to the Delaunay triangulation and Voronoi diagrams of those points.

This approach is used in areas such as the detection of features in images (Varytimidis et al., 2016), reconstruction of surfaces from a cloud of points (Wu et al., 2015) and Spectroscopy (Xu et al., 2019), with the last work, that involves the estimation and removal of the Blaze function -a kind of envelope- of a echelle spectrograph, being particularly align with what is intended here.

Thus, in this paper we formulate a general approach to envelope detection, exploiting the intrinsic characteristics of a generic, spectrally complex, wave in order to avoid the need to manual intervention or parameter tuning. While the robustness of the proposed approach allows it to be used as a plug in replacement for many methods encountered in the literature, we feel that it would be particularly useful for sound synthesis. The envelope is shown to add complexity to the spectral representation of a wave (Tarjano & Pereira, 2019), and a accurate description of the envelope would be useful for a cleaner spectral analysis. Moreover, the algorithm naturally divides a signal in pseudo-cycles, that could serve pragmatic building blocks for the reconstruction of the fine structure of the wave. In the context of sound synthesis, one would be able to apply specific methods for the recreation of the envelope, more in line with its smooth, slow varying and non periodic nature and a different approach to the periodic and relatively fast changes of the temporal fine structure.

2 Methodology

We start by defining a pulse as a series of consecutive samples in a discrete wave with the same signal. More formally, let $W[i] \in \mathbb{R} \forall i \in \mathbb{N}_0, i < n \in \mathbb{N}_0$ be a real, discrete and finite signal indexed, without loss of generality, over a subset of the natural numbers. This definition relates closely to the concept of an array or vector in programming languages, and is used in the interest of simplicity. A pulse $P[j], j \in \mathbb{N}_0$ in W can then be defined as a sequence of samples indexed by i such that $\text{sign}(W[a-1]) \neq \text{sign}(W[a]) = \text{sign}(W[a+1]) = \cdots = \text{sign}(W[b-1]) = \text{sign}(W[b]) \neq \text{sign}(W[b+1]) \forall a \leq i \leq b | a, b \in \mathbb{N}_0$; that is to say, a cluster of samples between the change of sign. For a continuous function, that would correspond to the pieces between roots, and would be equal to half the cycle of a sinusoid. Thus, in our analogy, two consecutive, opposed pulses have the potential do delimitate a pseudo-cycle, as illustrated in figure 1. For that to be true, however, the pulses need to make part of the frontier, as will be defined later.

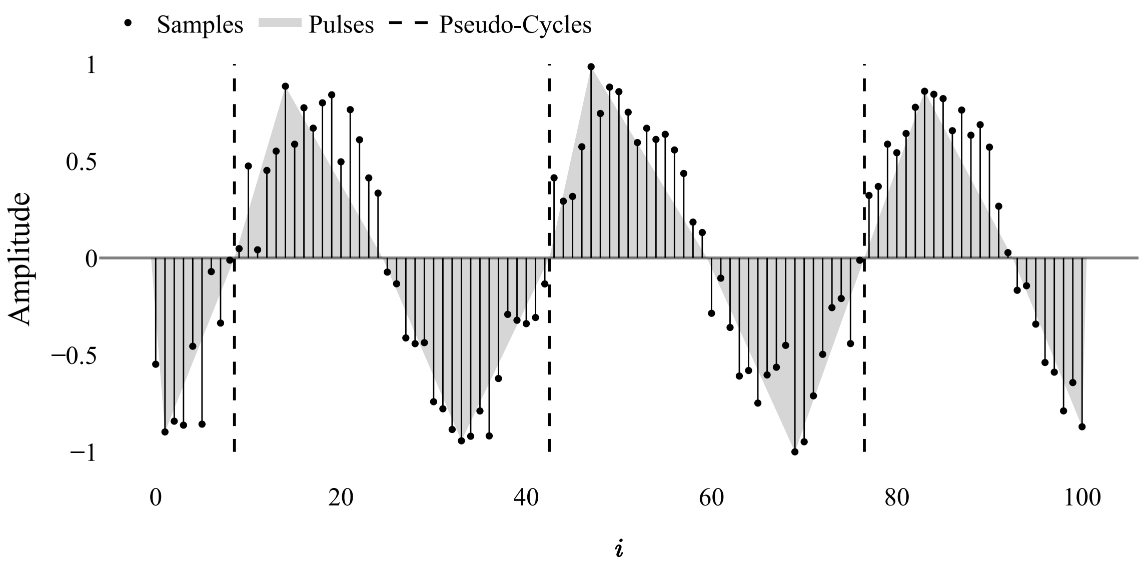

Figure 1: The grey areas encompass the simplified convex hulls of the pulses. The dashed lines mark the frontier between pseudo-cycles

As we are seeking a geometric treatment of the subject, it is necessary to make abscissa and ordinate units consistent. this can be accomplished by multiplying each sample of the signal by the average length of the pulses, after having divided by the average of signal amplitudes, putting both axes in the same time related units. One can call this approach arc length regularization, as it translates the widespread approach of arc length parametrization to a non parametric setting.

It is useful now to introduce the concept of a frontier as the set of points $(i, W[i])$ that bound the smooth, continuous envelope of the signal. Because we are interested in a smooth function over the whole domain of the wave, we can assume that, in the scale of a pulse, this function will be very close to a straight line and, thus, won't touch any point inside the convex hull defined by the pulse. By the same token, as the change in amplitude from one pulse to the next is motivated by changes in the envelope, one can assume that $\max(P[j]) \approx \max(P[j+1])$ and approximate the convex hull of a pulse with the triangle defined by the points where it crosses the horizontal axis and it's point of max amplitude; this will allow a considerable dimension reduction as this triangular pulse becomes the atomic unit for the rest of the method.

Besides, those definitions will allow us to derive the frontier of a wave using an alpha-shapes inspired algorithm without the need to compute the Delaunay triangulation first. Translating the intuitive explanation of alpha-shapes in Edelsbrunner and Mücke (1994), we can say that the points in the frontier are those touched by a circle outside the signal that is not allowed to contain any point of the signal; one can picture a circle being rolled above (or below, in the case of the negative envelope) the signal, and marking the points it touches as frontier points.

It is now necessary to infer the appropriate radius of such a circle and, to that end, a measure of the instantaneous curvature of a discrete function is needed. Discrete curvature estimation is an important task in image processing (Fleischmann et al., 2010) for which no default definition exists; The two possible approaches are the derivation of direct methods, or the calculation of the curvature of a fitted curve (Coeurjolly et al., 2001). As we are interested in the average curvature, the second approach is more natural, since it enables us to calculate the average curvature via the integral of the curvature in the fitted curve.

For a continuous, twice differentiable function $y = f(x); x, y \in \mathbb{R}$, the curvature as a function of the independent variable is given by $\kappa(x) = \frac{f''(x)}{(f'(x)^2+1)^{\frac{3}{2}}}$ which is also the inverse of the radius of the osculating circle tangent to the curve at x. Since we are assuming smoothness of envelope in the neighbourhood of a pulse, it makes sense to approximate it with a parabola

$y(x) = ax^2 + bx + c$, since a higher order polynomial would provide unnecessary complexity. Besides some characteristics of the parabola derived from its relation with Bezier curves will be useful later.

Therefore, the definition of the curvature becomes $\kappa(x) = \frac{2a}{\left((b+2ax)^2+1\right)^{\frac{3}{2}}}$ with the average curvature between points x_0 and x_1 given by equation 1.

$$\overline{\kappa}(x) = \frac{\left(\frac{2ax_1+b}{\sqrt{\left(2ax_1+b\right)^2+1}} - \frac{2ax_0+b}{\sqrt{\left(2ax_0+b\right)^2+1}}\right)}{(x_1 - x_0)} \tag{1}$$

3 Results

4 Discussion + Conclusion

5 References

Assef, A. A., Ferreira, B. M., Maia, J. M., & Costa, E. T. (2018). Modeling and FPGA-based implementation of an efficient and simple envelope detector using a hilbert transform FIR filter for ultrasound imaging applications. *Research on Biomedical Engineering*, *34*(1), 87–92. https://doi.org/10.1590/2446-4740.02417

Coeurjolly, D., Miguet, S., & Tougne, L. (2001). Discrete curvature based on osculating circle estimation. In *Lecture notes in computer science* (pp. 303–312). Springer Berlin Heidelberg. https://doi.org/10.1007/3-540-45129-3_27

Edelsbrunner, H., Kirkpatrick, D., & Seidel, R. (1983). On the shape of a set of points in the plane. *IEEE Transactions on Information Theory*, *29*(4), 551–559. https://doi.org/10.1109/tit.1983.1056714

Edelsbrunner, H., & Mücke, E. P. (1994). Three-dimensional alpha shapes. *ACM Transactions on Graphics (TOG)*, *13*(1), 43–72. https://doi.org/10.1145/174462.156635

Fleischmann, O., Wietzke, L., & Sommer, G. (2010). A novel curvature estimator for digital curves and images. In *Lecture notes in computer science* (pp. 442–451). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-15986-2_45

Lokki, T., Pätynen, J., Tervo, S., Siltanen, S., & Savioja, L. (2011). Engaging concert hall acoustics is made up of temporal envelope preserving reflections. *The Journal of the Acoustical Society of America*, *129*(6), EL223–EL228. https://doi.org/10.1121/1.3579145

Lyons, R. (2017). Digital envelope detection: The good, the bad, and the ugly [tips and tricks]. *IEEE Signal Processing Magazine*, *34*(4), 183–187. https://doi.org/10.1109/msp.2017.2690438

Qi, B., Mao, Y., Liu, J., Liu, B., & Xu, L. (2017). Relative contributions of acoustic temporal fine structure and envelope cues for lexical tone perception in noise. *The Journal of the Acoustical Society of America*, *141*(5), 3022–3029. https://doi.org/10.1121/1.4982247

Shannon, R. V., Zeng, F.-G., Kamath, V., Wygonski, J., & Ekelid, M. (1995). Speech recognition with primarily temporal cues. *Science*, *270*(5234), 303–304. https://doi.org/10.1126/science.270.5234.303

Tarjano, C., & Pereira, V. (2019). Neuro-spectral audio synthesis: Exploiting characteristics of the discrete fourier transform in the real-time simulation of musical instruments using parallel neural networks. In *Artificial neural networks and machine learning – ICANN 2019: Text and time series* (pp. 362–375). Springer International Publishing. https://doi.org/10.1007/978-3-030-30490-4_30

Varytimidis, C., Rapantzikos, K., Avrithis, Y., & Kollias, S. (2016). Alpha-shapes for local feature detection. *Pattern Recognition*, *50*, 56–73. https://doi.org/10.1016/j.patcog.2015.08.016

Wu, B., Yu, B., Huang, C., Wu, Q., & Wu, J. (2015). Automated extraction of ground surface along urban roads from mobile laser scanning point clouds. *Remote Sensing Letters*, *7*(2), 170–179. https://doi.org/10.1080/2150704x.2015.1117156

Xu, X., Cisewski-Kehe, J., Davis, A. B., Fischer, D. A., & Brewer, J. M. (2019). Modeling the echelle spectra continuum with alpha shapes and local regression fitting. *The Astronomical Journal*, *157*(6), 243. https://doi.org/10.3847/1538-3881/ab1b47

Yang, W., Yang, Y., & Yang, M. (2014). Fast digital envelope detector based on generalized harmonic wavelet transform for BOTDR performance improvement. *Measurement Science and Technology*, *25*(6), 065103. https://doi.org/10.1088/0957-0233/25/6/065103

Zhu, Z., Miyauchi, R., Araki, Y., & Unoki, M. (2018). Contributions of temporal cue on the perception of speaker individuality and vocal emotion for noise-vocoded speech. *Acoustical Science and Technology*, *39*(3), 234–242. https://doi.org/10.1250/ast.39.234