シンクロ・スクィージング変換による信号の時間-周波数特徴の抽出とデータ・ソニフィケーションへの応用

ベリアン アレックス [1], リュング ジョーダン [2], 斎藤 直樹 [1]
[1] カリフォルニア大学 デイヴィス校 数学科, [2] カリフォルニア大学 デイヴィス校 統計学科
e-mail : {ajberrian, jwcleung, nsaito}@ucdavis.edu

1 Introduction

A central question of time-frequency analysis is how to decompose a signal with time-varying oscillatory properties into several components with distinct amplitude and frequency behaviors. We may formulate this problem mathematically as follows.

Let $f : \mathbb{R} \to \mathbb{C}$ be a signal and assume f can be written as a finite sum of (unknown) *amplitude-phase components (modes)*:

$$f(t) = \sum_{k=1}^{K} f_k(t), \quad f_k(t) := A_k(t)\mathrm{e}^{2\pi\mathrm{i}\phi_k(t)},$$

which is called an *amplitude-phase decomposition* of f and the $\{A_k\}$ represent *instantaneous amplitudes* (IAs) and $\{\phi_k'\}$ represent *instantaneous frequencies* (IFs). Then, the problem is to retrieve the f_k given that only f is known.

The IAs and IFs of a signal can be visualized using the short-time Fourier transform (STFT). However, due to the uncertainty principle, the exact IAs and IFs are obscured by the blurry STFT representation. One method that has been designed to sharpen the STFT information and approximately retrieve the IAs and IFs is known as the *Synchrosqueezing transform* (SST) [1]. In this talk, we introduce the notion of an SST based on a *quilted STFT*, where the window g is allowed to change depending on the time-frequency region of interest. We also apply the SST to the problem of sonifying real datasets.

2 Synchrosqueezing transform

2.1 SST based on STFT

We define the STFT $V_g f$ of a signal f by $\int_{\mathbb{R}} f(x)g(x-t)\mathrm{e}^{-2\pi\mathrm{i}\xi(x-t)}\,\mathrm{d}x$. Then, the STFT-based SST with tolerance $\gamma > 0$ and limiting parameter $\beta > 0$ is given by

$$S_{f,\gamma}^{\beta}(t,\xi) := \int_{A_{\gamma,f}(t)} V_g f(t,\eta)\frac{1}{\beta}B\left(\frac{\xi - \xi_f(t,\eta)}{\beta}\right)\mathrm{d}\eta,$$

where $B \in C_c^\infty(\mathbb{R})$ is a "bump function" satisfying $\hat{B}(0) = 1$, $A_{\gamma,f}(t) := \{\eta \in \mathbb{R}_+ : |V_g f(t,\eta)| > \gamma\}$, and $\xi_f(t,\eta) := \dfrac{\partial_t\left[V_g f(t,\eta)\right]}{2\pi\mathrm{i}V_g f(t,\eta)}$ is an approximation to IF. We assume that the A_k and ϕ_k' are *bounded, sufficiently smooth*, and *slowly-varying*, and we also assume that the ϕ_k' are *well-separated*, i.e., $\exists d > 0$ s.t. $\forall t \in \mathbb{R}$, $\phi_k'(t) - \phi_{k-1}'(t) > d$ if $k \geq 2$. The authors of [1, 2] proved that under these and several other assumptions, one may accurately extract the ϕ_k' and then reconstruct the f_k via

$$f_k(t) \approx \lim_{\beta \to 0^+} \int_{\{\xi \,:\, |\xi - \phi_k'(t)| < \gamma\}} S_{f,\gamma}^{\beta}(t,\xi)\,\mathrm{d}\xi.$$

2.2 SST based on quilted STFT

The assumptions above on A_k may not always be physically realistic. For instance, the onset of a note in a music signal may be modeled by discontinuous or fast-changing A_k. The STFT-based SST may not accurately capture the onset energy, because the STFT time resolution capability is limited by the window g. However, one may consider allowing g to change depending on the time-frequency region of interest. This yields the concept of a *quilted STFT*, where different time-frequency regions represent patches in a quilt covering the time-frequency plane [3].

We define the quilted STFT $V_g^Q f$ of a signal f by $V_g^Q f(t,\xi) := \int_{\mathbb{R}} f(x)g_{t,\xi}(x-t)\mathrm{e}^{-2\pi\mathrm{i}\xi(x-t)}\,\mathrm{d}x$, where for each (t,ξ), $g_{t,\xi}$ is a window function centered at 0. We then define the quilted-STFT-based SST $S_{f,\gamma}^{Q,\beta}$ by replacing $V_g f$ by

$V_g^Q f$ everywhere in the definition of $S^{\beta}_{f,\gamma}$. Our first restriction on $g_{t,\xi}$ is that it does not vary too much in t: if $\tilde{g}(t,\xi,x) := g_{t,\xi}(x)$, then $\forall t \in \mathbb{R}$, $\int_{\mathbb{R}} |\partial_t \, \tilde{g}(t,\xi,x)| \, \mathrm{d}x < \infty$. Additionally, for fixed t, we restrict $g_{t,\xi}$ to be *constant in* ξ over each frequency band $\{\xi : |\xi - \phi_k'(t)| < \gamma\}$. Then, with all the previous assumptions above and two other fairly non-restrictive ones, we have proven that one can approximately reconstruct f_k via

$$f_k(t) \approx \lim_{\beta \to 0^+} \int_{\{\xi \,:\, |\xi - \phi_k'(t)| < \gamma\}} S^{Q,\beta}_{f,\gamma}(t,\xi) \, \mathrm{d}\xi.$$

In our talk, we will provide numerical evidence that quilted-STFT-based SST performs better than STFT-based SST on signals with discontinuous A_k. Figure 1 demonstrates the usage of the quilted-STFT-based SST on a synthetic test signal.

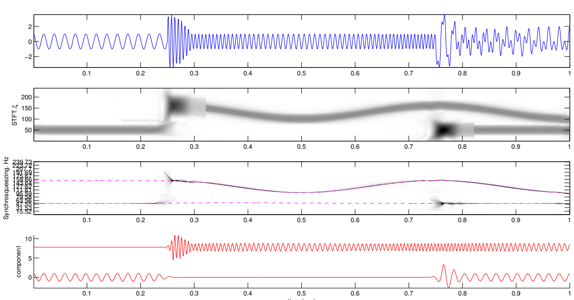

図 1. From top to bottom: the synthetic test signal $f = f_1 + f_2$; $|V_g^Q f|^2$ with different windows used around component onsets; $|S^{Q,\beta}_{\gamma,f}|^2$ with extracted IF curves $\phi_1' < \phi_2'$ (magenta); the extracted f_1 and f_2.

3 Application of SST: Data sonification

As an application of SST, we consider *sonification* (a translation into sound) of 16 temperature readings in Lake Tahoe, each taken at a different depth of the lake. We convey this data in a manner that separates the short-term oscillatory and long-term trend information from each of the temperature signals, while still enabling their simultaneous "reading." A visualization of all this information may be difficult to read. But using the power of our auditory system, one has some hope of "hearing" all the information together. Since music signals share similar oscillatory characteristics to those in our data, it is natural to consider a musical model.

Our algorithm proceeds as follows. First, we assign an instrument to each temperature signal, with higher-pitched instruments for readings closer to the surface. Next, we use the SST to extract IF curves from each signal (Figure 2), which are linearly mapped to notes in a musical scale. We then use a LOESS (locally weighted polynomial regression) method to extract each signal's trend, which we map to MIDI volume values. The final product is a music file in MIDI format, which we will play in our presentation.

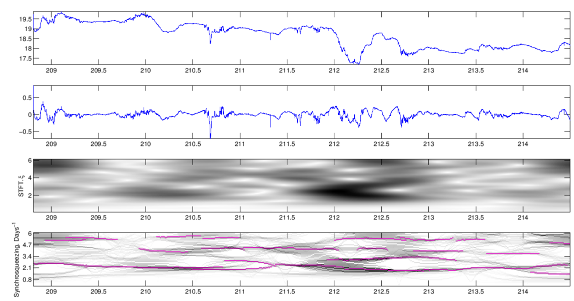

図 2. From top to bottom: one of the temperature measurements; its detrended version; the magnitude of the STFT; the SST with the extracted IF curves.

謝辞 The authors are indebted to the contributions of Xiang Cui and Geoffrey Schladow of UC Davis on this project. This research was partially supported by ONR grant N00014-12-1-0177 and NSF grant DMS-1418779.

参考文献

[1] G. Thakur and H-T. Wu, Synchrosqueezing-based recovery of instantaneous frequency from nonuniform samples, SIAM J. Math. Anal., Vol. 43, No. 5 (2011), 2078–2095.

[2] T. Oberlin, S. Meignen, and V. Perrier, The Fourier-based synchrosqueezing transform, Proc. IEEE ICASSP, pp. 315–319, 2014.

[3] M. Dörfler, Quilted Gabor frames – A new concept for adaptive time-frequency representation, Adv. Appl. Math., Vol. 47, No. 4 (2011), 668–687.