The Quilted Synchrosqueezing Transform for Adaptive Time-Frequency Signal Analysis

Alex Berrian and Naoki Saito
University of California, Davis

Poster and Dessert Reception, 2016 SIAM Annual Meeting, Boston, Massachusetts, USA, July 12, 2016

Introduction

- The synchrosqueezing transform (SST) is a popular tool for sharpening the analysis and mode decomposition of signals with time-varying oscillatory properties.
- We introduce an SST-based on a quilted short-time Fourier transform (SST-QSTFT). This approach allows for multiple analysis windows to be used, adapting to signal behavior in different time-frequency regions.
- Our theoretical results demonstrate the accuracy of signal analysis and reconstruction.
- We propose an algorithm for automatic selection of optimal windows for SST-QSTFT.
- We provide a new discrete frequency reassignment formula that avoids inaccuracy at high frequencies.
- We apply our algorithm to an audio signal of a glockenspiel.
- We provide a Python-based implementation of our algorithm.

Amplitude-Phase Decomposition

We assume our signal $f : \mathbb{R} \rightarrow \mathbb{C}$ can be written

$$f(t) = \sum_{l=1}^{L} \{a_l(t) + \mathrm{i} b_l(t)\} = A_l(t) e^{\mathrm{i} \phi_l(t)}$$  \hspace{1cm} (1)

We call this an amplitude-phase decomposition of $f$.

$A_l(t)$: instantaneous amplitudes (IAs)

$b_l(t)$: instantaneous frequencies (IFs)

Relevant signals: audio, music, speech, medical (EEG, ECG), mechanical (vibration, pressure), radar, sonar...

Weakly modulated IA and IF signal class [4]: We say $f \in \mathcal{B}^{(1,2)}$ if (1) holds for some $K \in \mathbb{N}$ and $a_l \in C^{1,2}(\mathbb{R})$, $b_l \in L^2(\mathbb{R})$, $A_0(t) > 0$ and $\frac{\partial}{\partial t} A_l(t) = 0$.

We then define the STFT of $f$ with a window $g \in L^2(\mathbb{R})$ centered at $0$, given by

$$V_g f(t,\xi) := \int_{\mathbb{R}} f(t-x)g(x)e^{-2\pi i \xi x} \, dx,$$  \hspace{1cm}

for all $t, \xi \in \mathbb{R}$, allows for the analysis of the IAs and IFs of $f$.

ISSUE: Time-frequency resolution tradeoff due to the Uncertainty Principle. STFT cannot have perfect resolution in both time and frequency.

- $g$ narrow: Good time resolution, poor frequency resolution
- $g$ wide: Good frequency resolution, poor time resolution

STFT Example

Figure 1: $|V_g x(t,\xi)|^2$ with a wide and $g$ narrow, for the glockenspiel signal, with frequency range restricted between 2000 to 11000 Hz.

STFT-Based Synchrosqueezing

The STFT-based synchrosqueezing transform (SST-QSTFT) [4] with tolerance $\gamma > 0$ and limiting parameter $\beta > 0$ is given by

$$S^{\text{SST-QSTFT}}_{\xi,\gamma}(t) := \int_{\mathbb{R}} V_g f(t,\xi) e^{-\frac{\gamma}{\beta} \frac{|\xi|^2}{2}} d\xi,$$

where $\xi \in C^2(\mathbb{R})$ is a "bump function" satisfying $\xi(0) = 1$, $\xi'(0) = -\gamma$ and $\xi''(0) > 0$. $S^{\text{SST-QSTFT}}_{\xi,\gamma}$ is the SST-based frequency reassignment.

SST-QSTFT Example

Figure 2: SST-QSTFT of glockenspiel signal, frequency range restricted from 2000 to 11000 Hz.

SSTFT Example

Figure 3: SST-QSTFT of glockenspiel signal, with applied only where wider window is used. Window automatically chosen for each subpart using algorithm described below. Transformer acts automatically on the STFT, so as to the better time-resolution of the narrower window used in transient noise regions.

Theorem

Theorem [B. [4]]: Let $\epsilon > 0$, $\epsilon^{'(\epsilon)} > 0$. Suppose that $f \in \mathcal{B}^{(1,2)}$. Assume that $\{b_l(t),\xi_l(t)\}_{l=1}^{L}$ is of the class $\mathcal{W}^{\text{SST-QSTFT}}$. Then, it is sufficiently small we have:

- (Concentration of SSTFT around IF curve) $|\int f(t) \xi(t)| > \alpha$ when $t \in K$ and $\xi \in \mathbb{K}$ such that $|\xi - \xi_l| > \epsilon$
- (Closedness of reassignment frequency $\xi_l$ to $\xi$) $|\xi_l - \xi| < \epsilon$ for all $\xi_l \in K$. Then:
- (Accuracy of reconstruction) For every $k \in \{1, \ldots, K\}$ there is a constant $C_k$ such that for all times $t \in K,$

$$\int_{|\xi - \xi_l| < \epsilon} \frac{1}{C_k} |S^{\text{SST-QSTFT}}_{\xi,\gamma}(t,\xi) - f(t)| \leq \xi_l.$$

Adaptive Window Selection

We follow (and slightly modify) the approach of Jallet & Torreani [3] and (3) in [3].

1. Fix parameters $A, B > 0$ to decompose $\mathbb{R}^2$ into "super-tiles" $\mathcal{Q}_m,n := \{(t,\xi) \in \mathbb{R}^2 : (|A|m^2 - B, |B|n^2 - A) \}.$
2. Choose windows $h_m,h_n$, and corresponding sampling lattices.
3. Choose Rényi entropy measure parameter $\alpha \in (0,1)$.
4. For each window $h_m,h_n$, compute discrete SSTFT $V_{h_m} f$ with hop size $m^2$ and $DFT$ size $N^2$ (chosen in step 2).
5. For each $\mathcal{Q}_m,n$, and $h_m,h_n$ calculate the Rényi entropy $R_{\alpha}^{\text{SSTFT}}(\mathcal{Q}_m,n)$ of the STFT coefficients whose corresponding lattice points are within $\mathcal{Q}_m,n$.

$$R_{\alpha}^{\text{SSTFT}}(\mathcal{Q}_m,n) := \frac{1}{N^2} \sum_{(m,n) \in \mathcal{Q}_m,n} \left| \frac{N^{2/\alpha}}{m^{2/\alpha}n^{2/\alpha}} \right| V_{h_m} f(m,n)^2$$

6. To each super-tile $\mathcal{Q}_m,n$ associate the window $h_m,h_n$ with the smallest Rényi entropy $R_{\alpha}^{\text{SSTFT}}(\mathcal{Q}_m,n)$, and then adapt to signal content in $\mathcal{Q}_m,n$. New Discrete Frequency Reassignment Formula

Suppose $h_m$ has sampling rate $f_s = |A|m^2$, $f_d = B$, and $m^2$ are the STFT time grid points and frequency bin indices, and $V_{h_m}$ denotes the discrete STFT with compactly supported $g$. Then $\xi_l$ is usually numerically computed via the forward difference

$$\xi_l(t,n) := \frac{V_{h_m}(f_s(t,n) + \Delta f, f_d(n)) - V_{h_m}(t,n)}{2m^2\Delta f}$$

But (2) is very inaccurate at high frequencies! Suppose $f(t) = e^{\mathrm{i}t^2}$ for $t \in \mathbb{R}$. Then $\xi_l(t,n) \approx \mathrm{i}t$, $\xi_l(t,n) \approx \mathrm{i}t$ for $e^{-\gamma t^2}$. For $\epsilon > 1$, $\text{sup}(\Delta f) < 1$, so $\xi_l(t,n)$ may change. Hence, the reassignment severely underapproximates the IF. Thus, we propose a new discrete frequency reassignment formula:

$$\xi_l(t,n) := \frac{V_{h_m}(f_s(t,n) + \Delta f, f_d(n)) - V_{h_m}(f_s(t,n) - \Delta f, f_d(n))}{2m^2\Delta f}$$

For $f$, as defined above, this new formula is exact, that is, $\xi_l(t,n) = \xi$ for all pairs $(t,n)$. In our implementation, we use (3) to compute our STTs.

Conclusion

- Our SST-QSTFT approach provides a sharpened time-frequency representation with automatic adaptation to signal content in different time-frequency regions, with the additional possibility of reconstruction.
- Our new discrete reassignment formula allows for high accuracy in high-frequency regions, which is crucial in audio applications.
- Future research: Further investigation of SST-QSTFT convergence in the discrete setting, as well as the utilization of time-reassignment on transient regions to further specify the time-frequency representation. We may also explore possible music applications for SST-QSTFT, such as audio fingerprinting, onset detection, and automatic transcription.

Codes

We have made a Python suite available upon request for adaptive time-frequency transforms. This suite includes an implementation of SST-QSTFT, as well as an implementation of the related time-frequency graph wavelet decomposition methods [3].

Contact Information

Alex Berrian: aberrian@math.ucdavis.edu
Naoki Saito: saisato@math.ucdavis.edu

Acknowledgements

- Support from Office of Naval Research grant: ONR N00014-16-1-2325
- Support from National Science Foundation grants: DMS-1148643, DMS-1418779

References


