Quasi-Monte Carlo Time-Frequency Analysis

Barbara Pascal

May 12, 2021

from R. Levie, H. Avron, and G. Kutyniok `arxiv:2011.02025`

But first ... music!

- Born in Leeds, in 1946
- Education in music
  - Oxford (BA, '68)
  - Nottingham (MA, '69)

But first ... music!

- Born in Leeds, in 1946
- Education in music
  - Oxford (BA, ’68)
  - Nottingham (MA, ’69)

† death of his father (factory worker) ’69

Abandoned *conventional* composing

- collected sounds of machinery, foundries and power stations
- thesis on sonic art at York (PhD, ’73)
- interaction of human voice and electronic systems

But first ... music!

- Born in Leeds, in 1946
- Education in music
  - Oxford (BA, '68)
  - Nottingham (MA, '69)

† death of his father (factory worker) '69

Abandoned *conventional* composing

- collected sounds of machinery, foundries and power stations
- thesis on sonic art at York (PhD, '73)
- interaction of human voice and electronic systems

http://www.trevorwishart.co.uk/

Vox 5, '86-'87
Transformation and interpolation with natural sounds of humain voice

Institut de Recherche et Coordination Acoustique/Musique-Centre Pompidou

Elementary transforms used in time-frequency analysis

Action in the time domain $s \in L^2(\mathbb{R})$

- translation by x: $[\mathcal{T}(x)s](t) = s(t - x)$
- modulation by ω: $[\mathcal{M}(\omega)s](t) = s(t)\mathrm{e}^{2\mathrm{i}\pi\omega t}$
- dilation by τ: $[\mathcal{D}(\tau)s](t) = \tau^{-1/2}s(\tau^{-1}t)$

Elementary transforms used in time-frequency analysis

Action in the time domain $s \in L^2(\mathbb{R})$

- translation by x: $[\mathcal{T}(x)s](t) = s(t-x)$
- modulation by ω: $[\mathcal{M}(\omega)s](t) = s(t)\mathrm{e}^{2\mathrm{i}\pi\omega t}$
- dilation by τ: $[\mathcal{D}(\tau)s](t) = \tau^{-1/2}s(\tau^{-1}t)$

Lemma [Unitarity] For any $x \in \mathbb{R}, \omega \in \mathbb{R}, \tau \in \mathbb{R} \setminus \{0\}$,

$\mathcal{T}(x)$, $\mathcal{M}(\omega)$ and $\mathcal{D}(\tau)$ are *unitary* operators of $L^2(\mathbb{R})$

$\mathcal{F} : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ the Fourier transform

Action in the frequency domain

- $\mathcal{F}\mathcal{T}(x)\mathcal{F}^* = \mathcal{M}(-x)$
- $\mathcal{F}\mathcal{M}(\omega)\mathcal{F}^* = \mathcal{T}(\omega)$
- $\mathcal{F}\mathcal{D}(\tau)\mathcal{F}^* = \mathcal{D}(\tau^{-1})$

Short-Time Fourier and Continuous Wavelet Transforms

STFT

$$f_{a,b} = \mathcal{T}(a)\mathcal{M}(b)f$$

$f \in L^2(\mathbb{R})$ *window* localized in time and frequency about 0

Fixed length of support
Varying nb. of oscillations

The larger the frequency, the more oscillations.

CWT

$$f_{a,b} = \mathcal{T}(a)\mathcal{D}(b^{-1})f$$

f *mother wavelet* localized in time about 0, in frequency about 1

$\int \omega |\widehat{f}(\omega)|^2 \, \mathrm{d}\omega < \infty$ (admissibility)

Varying length of support
Fixed nb. of oscillations

The larger the frequency, the shorter the time support.

Localizing Time-Frequency Transform (LTFT)

Heuristic

- Short-Time Fourier atoms at low and high frequencies (STFT)
- Continuous Wavelet atoms for middle frequencies (CWT)
- *Third* parameter: number of oscillations in the *mother* wavelet.

Heuristic

- Short-Time Fourier atoms at low and high frequencies (STFT)
- Continuous Wavelet atoms for middle frequencies (CWT)
- *Third* parameter: number of oscillations in the *mother* wavelet.

Definition [LTFT continuous frame] $f_{a,b,c} = \pi(a,b,c)f$ with

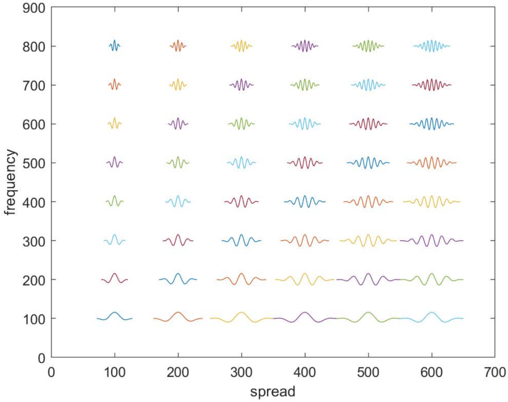
$$\pi(a,b,c) = \begin{cases} \mathcal{T}(a)\mathcal{M}\left(\frac{\xi}{\gamma}cb_0 + b\right)\mathcal{D}\left(\frac{\gamma}{b_0}\right) & b < b_0 \\ \mathcal{T}(a)\mathcal{M}\left(\left(\frac{\xi}{\gamma}c + 1\right)b\right)\mathcal{D}\left(\frac{\gamma}{b}\right) & b_0 \leq b \leq b_1 \\ \mathcal{T}(a)\mathcal{M}\left(\frac{\xi}{\gamma}cb_1 + b\right)\mathcal{D}\left(\frac{\gamma}{b_1}\right) & b > b_1 \end{cases}$$

- "*father*" atom: $f \in L^2(\mathbb{R})$ support. in $[-1/2, 1/2]$ localized at 0 in T & F
- phase space: $G = \mathbb{R}^2 \times [0,1]$ (time, frequency, nb. of oscillations)
- STFT-CWT transition frequencies: $0 < b_0 < b_1$
- minimal nb. of oscillations: γ
- oscillation range: ξ

Localizing Time-Frequency Transform (LTFT)

Heuristic

- Short-Time Fourier atoms at low and high frequencies (STFT)
- Continuous Wavelet atoms for middle frequencies (CWT)
- *Third* parameter: number of oscillations in the *mother* wavelet.



Theory of *frames*

- $\mathcal{S}$: Hilbert space of signals

  e.g.: finite energy signals $L^2(\mathbb{R})$

- G : locally compact Borel space with measure μ

  e.g.: time-frequency plane $\mathbb{R}^2$ with Lebesgue measure λ

Definition [Continuous frame]

Frame $\quad f : \begin{cases} G & \rightarrow & \mathcal{S} \\ g & \mapsto & f_g \end{cases}$ weakly-measurable mapping

$V_f[s] : \begin{cases} G & \rightarrow & \mathbb{C} \\ g & \mapsto & \langle s, f_g \rangle_{\mathcal{S}} \end{cases}$ the *coefficient function* of signal $s \in \mathcal{S}$.

Continuous $\quad \forall s \in \mathcal{S},\ V_f[s] \in L^2(G)$ and $\exists\, A, B,\ 0 < A \leq B < \infty$, s.t.

$$A\|s\|_{\mathcal{S}}^2 \leq \|V_f[s]\|_{L^2(G)}^2 \leq B\|s\|_{\mathcal{S}}^2.$$

V_f *analysis* operator, V_f^* *synthesis* operator

Signal processing in phase space

General procedure

(i) compute $V_f[s](g)$ for every point of phase space $g \in G$
(ii) apply a nonlinearity $\kappa\left(V_f[s](g), g\right)$
(iii) map each $g \in G$ onto $\rho(g)$
(iv) synthesize the output signal

$$s_{\mathrm{out}} = \int_G \kappa\left(V_f[s](g), g\right) f_{\rho(g)} \,\mathrm{d}\mu(g)$$

Examples

- multipliers: $\kappa(c,g) = cr(g)$, $\rho(g) = g$
- signal denoising: $\kappa(c,g) = \kappa(c)$, $\rho(g) = g$
- **time stretching vocoder:** V_f Short-Time Fourier Transform

"*slow down a signal without dilating its frequency content*"

$g = (t,\omega)$, $\rho(g) = (Dt,\omega)$ for some $D \in \mathbb{N}$, $\kappa(c,g) = \kappa(c)$ and

$$\forall a, \theta \in \mathbb{R}_+, \quad \kappa(\mathrm{e}^{\mathrm{i}\theta} a) = \mathrm{e}^{\mathrm{i}D\theta} a$$

Contributions of the paper

Notations M: resolution of the signal space,
N: number of samples in phase space
d: dimension of phase space.

- $\mathcal{O}\left(\frac{M}{N}(\log_2 N)^{d-1}\right)$ error rate of proposed quasi-Monte Carlo LTFT

  compared to $\mathcal{O}\left(\sqrt{\frac{M}{N}}\right)$ for standard Monte Carlo
- implementation of a *time dilation phase vocoder* based on LTFT
- general theory for quasi-Monte Carlo sampling of phase spaces

  STFT, CWT, LTFT, Shearlet or Curvelet transforms

Quasi-Monte Carlo

Let $f : I^d = [0,1]^d \to \mathbb{C}$, QMC is a cubature method to compute

$$\int_{I^d} f(t)\,\mathrm{d}t \simeq \frac{1}{N}\sum_{n=1}^{N} f(x_n)$$

from samples points $\mathcal{P}_n = \{x_1, \ldots, x_N\}$

Quasi-Monte Carlo

Let $f : I^d = [0,1]^d \to \mathbb{C}$, QMC is a cubature method to compute

$$\int_{I^d} f(t)\,\mathrm{d}t \simeq \frac{1}{N}\sum_{n=1}^{N} f(x_n)$$

from samples points $\mathcal{P}_n = \{x_1, \ldots, x_N\}$

Star discrepancy $D^*(\mathcal{P}_N) = \sup_{B \in \mathrm{Rec}^*} \left| \frac{\#(B \cap \mathcal{P}_N)}{N} - \lambda(B) \right|$

how well the volume is covered

Quasi-Monte Carlo

Multi-index $\alpha = (\alpha_1, \dots, \alpha_d)$, $\alpha_j \in \{0,1\}$ and $g : I^d \to \mathbb{C}$

$$\int_{I^d|_\alpha} g(a_1, \dots, a_d)\,\mathrm{d}a^\alpha$$

- integration w.r.t. a_j if $\alpha_j = 1$
- replace a_j by 1 if $\alpha_j = 0$

Ex. $$\int_{I^3|_{(1,0,1}} g(a_1, a_2 a_3)\,\mathrm{d}a^{(1,0,1)} = \int_{[0,1]^2} g(a_1, 1, a_3)\,\mathrm{d}a_1 \mathrm{d}a_3$$

Hardy-Krause variation $V(f) = \sum_\alpha \int_{I^d|_\alpha} |\partial_\alpha f(a_1, \dots, a_d)|\,\mathrm{d}a^\alpha$

Theorem [Koksma-Hlawka inequality]

$$\left| \int_{I^d} f(x)\,\mathrm{d}x - \frac{1}{N}\sum_{n=1}^{N} f(x_n) \right| \le V(f) D_N^*(\mathcal{P}_N)$$

Quasi-Monte Carlo

Multi-index $\alpha = (\alpha_1, \dots, \alpha_d)$, $\alpha_j \in \{0,1\}$ and $g : I^d \to \mathbb{C}$

$$\int_{I^d|_\alpha} g(a_1, \dots, a_d)\, \mathrm{d}a^\alpha$$

- integration w.r.t. a_j if $\alpha_j = 1$
- replace a_j by 1 if $\alpha_j = 0$

Ex. $$\int_{I^3|_{(1,0,1}} g(a_1, a_2 a_3)\, \mathrm{d}a^{(1,0,1)} = \int_{[0,1]^2} g(a_1, 1, a_3)\, \mathrm{d}a_1 \mathrm{d}a_3$$

Hardy-Krause variation $$V(f) = \sum_\alpha \int_{I^d|_\alpha} |\partial_\alpha f(a_1, \dots, a_d)|\, \mathrm{d}a^\alpha$$

Theorem [Koksma-Hlawka inequality]

$$\left| \int_{I^d} f(x)\, \mathrm{d}x - \frac{1}{N} \sum_{n=1}^{N} f(x_n) \right| \le V(f) D_N^*(\mathcal{P}_N)$$

Linear volume discretizable frames

Let f a frame on the signal space $\mathcal{S}$ with phase space G

Discretization of the frame:

- discrete signals $\{V_M \subset \mathcal{S}, \dim(V_M) = M \in \mathbb{N}\}$
- compact phase space $G_M \subset G$ rectangle of volume $\mathcal{O}(M)$

Sampling discretization of phase space: $\mathcal{P}_N = (g_1, \ldots, g_N)$

ideally low discrepancy, i.e. $D^*_N(\mathcal{P}_N)$ "small"

Phase space signal processing *via* Quasi-Monte Carlo discretization

$$s^N_{\mathrm{out}} = \frac{\mu(G_M)}{N} \sum_{n=1}^{N} \kappa\left(V_f[s](g_n), g_n\right) f_{\rho(g_n)}$$

$$s_{\mathrm{out}} = \int_G \kappa\left(V_f[s](g), g\right) f_{\rho(g)} \,\mathrm{d}\mu(g)$$

Sampling the phase space

Monte Carlo sampling

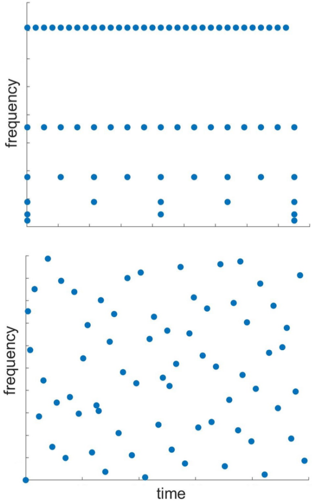
- well spread samples
- required nb. of samples ind. of d

  no curse of dimensionality

- suitable for TF feature extraction

Quasi-Monte Carlo

- midway between grid and MC
- *optimally* spread



How to design a sampling strategy?

Reminder: purpose is to describe signals with a family of atoms

$$\{f_{g_n}, n = 1, \dots, N\}$$

How to design a sampling strategy?

Reminder: purpose is to describe signals with a family of atoms

$$\{f_{g_n}, n = 1, \ldots, N\}$$

Requirements:

(i) well-spread samples

analysis extracts properly the local frequencies

(ii) allow reconstruction up to some small error

small information loss

(iii) sampling set small enough

computational efficiency

Standard discrete time-frequency transforms:

(ii) and (iii) satisfied ✓

(i) not satisfied ✗

Computational complexity

N: number of time-frequency samples

[*Total* number of pixels of *all* LTFT atoms] $\mathcal{C} = \mathcal{O}(\gamma N + \gamma)$

Error bound

M: resolution of the input discrete signal

Theorem [Approximation error of QMC LTFT]

$$\mathcal{O}\left(\frac{M \log^2(N)}{N}\right)$$

Approximation error of general QMC

M: resolution of the input discrete signal
d: dimension of phase space
N: number of time-frequency samples

QMC sample set $\mathcal{P}_N = \{(\boldsymbol{a}^n, \boldsymbol{b}^n), n = 1, \ldots, N\}$ of compact phase space with asymptotically *low* discrepancy $D_N^*(\mathcal{P}_N) \leq C\dfrac{(\log N)^{d-1} M}{N}$

Theorem [Reconstruction error] For $S : G_{M'} \to \mathbb{C}$ coefficients of a generalized TF transform, let $V_f^{M,N}(S) = \dfrac{M'}{N} \displaystyle\sum_{n=1}^{N} S(\boldsymbol{a}^n, \boldsymbol{b}^n) \underbrace{\mathcal{T}(\boldsymbol{a}) f_{\boldsymbol{b}}}_{\text{atoms}}$.

Then, for any discrete signal s_M of resolution M

$$\left\| V_f^{M*} V_f^M(s_M) - V_f^{M,N*} V_f^M(s_M) \right\|_\infty \leq \|s_M\|_\infty C \frac{(\log N)^{d-1} M}{N} D$$

► uniform pointwise bound on the QMC *synthesis* error.

Play your own music!

https://github.com/RonLevie/LTFT-Phase-Vocoder