EIGENTRIADS AND EIGENPROGRESSIONS ON THE TONNETZ

Vincent Lostanlen
Music and Audio Research Lab, New York University, New York, NY, USA

ABSTRACT

We introduce a new multidimensional representation, named *eigentriad* representation, that characterizes some essential patterns of Western tonal harmony while being equivariant to time shifts and pitch transpositions. This representation is deep, multiscale, and convolutional in the piano-roll domain, yet incurs no prior training, and is thus suited to both supervised and unsupervised MIR tasks. The eigentriad representation combines ideas from the spiral scattering transform, spectral graph theory, and wavelet shrinkage denoising. We report state-of-the-art results on a task of supervised composer recognition (Haydn vs. Mozart) from polyphonic music pieces in MIDI format.

1. EIGENTRIADS

Let \( \mathbf{x}[t,p] \in \mathcal{M}_{T,P}(\mathbb{R}) \) the piano-roll matrix of a musical piece, either obtained by parsing symbolic data or by extracting a melody salience representation from audio [2].

The constant \( T \) (resp. \( P \)) is typically equal to \( 2^{10} \) (resp. \( 2^7 \)). Within the framework of twelve-tone equal temperament, we define the major and minor triads as the tuples \( I_1 = (0, 4, 7) \) and \( I_0 = (0, 3, 7) \). For each quality \( q \in \mathbb{Z}_2 \) and frequency \( \beta \in \mathbb{Z}_3 \), Let

\[
\psi^{\text{triad}}_{\beta,q}[p] = \sum_{n=1}^{3} \exp \left( 2\pi i \frac{3n}{3} \right) \delta[p - I_q[n]],
\]

(1)

where \( \delta[p - I_q[n]] \) is the Kronecker delta symbol, equal to one if \( p = I_q[n] \) and zero otherwise. Let \( \mathcal{G}_q \) the induced subgraph of \( I_q \), where \( I_q \) is understood as a set of vertices in \( \mathbb{Z}_P \). Observe that \( \{ p \mapsto \psi^{\text{triad}}_{\beta,q} \} \) consists of the eigenfunctions of the unnormalized Laplacian matrix of \( \mathcal{G}_q \):

\[
\mathbf{L}^{\text{triad}}_{q}[p,p'] = |I_q| \delta[p \in I_q] \delta[p' \in I_q] - \delta[p \cong p']
\]

(2)

As a result, we propose to name *eigentriads* the complex-valued signals \( \psi^{\text{triad}}_{\beta,q} \) in pitch space. We construct a multiresolution convolutional operator in the piano-roll domain by separable interference \( \psi^{\text{triad}}_{(\alpha,\beta,q)}[t,p] = \psi_{\alpha,t}[t] \psi^{\text{triad}}_{\beta,q}[p] \)

between the aforementioned eigentriads and a family of temporal Gabor wavelets

\[
\psi_{\alpha,t}[t] = \alpha \exp \left( \frac{\alpha t^2}{2\sigma^2} \right) \exp(i \alpha t \xi)\]

(3)

for \( t \in [0; T[T]. \) We set \( \xi = \frac{2\pi}{2^7}, \sigma = 0.1, \xi = \frac{2\pi}{2^7}, \) and \( \log_2 \alpha_1 \in [0; \log_2 T] \). We define the *eigentriad transform* of \( x \) as the rank-five tensor resulting from the complex modulus of all convolutions between \( x \) and multivariable wavelets \( \psi^{\text{triad}}_{(\alpha,\beta,q)}[t,p] \): \n
\[
\Psi_{\alpha,t,p,q}(x) = \mathbf{U}_1(x)[t,p,q,\alpha_1,\beta_1] = |x\ast\psi_{(\alpha_1,\beta_1,q)}[t,p,p']| , \quad (4)
\]

where the difference in \( t \) (resp. in \( p \)) is computed in \( \mathbb{Z}_T \) (resp. in \( \mathbb{Z}_P \)). By averaging the tensor \( \Psi_{\alpha,t,p,q}(x) \) over the dimensions of \( t, p \), and triad quality \( q \), one obtains the matrix

\[
\mathbf{S}_1(x)[\alpha_1,\beta_1] = \sum_{t \in \mathbb{Z}_T} \sum_{p \in \mathbb{Z}_P} \sum_{q \in \mathbb{Z}_q} \Psi_{\alpha,t,p,q}(x)[t,p,q,\alpha_1,\beta_1].
\]

(5)

The operator \( x \mapsto \mathbf{S}_1(x) \) characterizes the relative amounts of ascending triads \( (\beta_1 = 1) \), descending triads \( (\beta_1 = -1) \), and perfect chords \( (\beta_1 = 0) \) at various temporal scales \( \alpha_1 \) in the piece \( x \), while keeping a relative low dimensionality, equal to \( 3 \log_2 T \). The averaging along variables \( t \) and \( p \) involved in Equation 5 guarantees that \( \mathbf{S}_1 \) is invariant to the action of any temporal shift operator \( \tau_{\Delta t} : x[t,p] \mapsto x[t + \Delta t,p] \), as well as any pitch transposition operator \( \tau_{\Delta p} : x[t,p] \mapsto x[t,p + \Delta p] \):

\[
\forall \Delta t \in \mathbb{Z}_T, \forall \Delta p \in \mathbb{Z}_P, \mathbf{S}_1(\tau_{\Delta t},\tau_{\Delta p} \circ x) = \mathbf{S}_1(x).
\]

(6)

Furthermore, the averaging across triad qualities \( q \) implies approximate invariance to tonality, in the sense that replacing major triads by minor triads and vice versa in \( x \) (insofar as this is feasible in the signal \( x \) at hand) does not affect the matrix \( \mathbf{S}_1(x) \). From the standpoint of serialist music theory [1], the presence of signed eigentriad frequencies \( \beta_1 = \pm 1 \) ensures that \( \mathbf{S}_1 \) is not invariant to retrogradation \( R : x[t,p] \mapsto x[-t, p] \), i.e. time reversal:

\[
(R \circ x) \neq x \implies \mathbf{S}_1(R \circ x) \neq \mathbf{S}_1(x).
\]

(7)

However, the averaging across triad qualities \( q \) causes \( \mathbf{S}_1 \) to be invariant to inversion \( I : x[t,p] \mapsto x[t,-p] \), i.e. reversal of the pitch axis:

\[
(I \circ x) \neq x \implies \mathbf{S}_1(I \circ x) \neq \mathbf{S}_1(x).
\]

(8)
The above property hinders the accurate modeling of chord progressions in the context of Western tonal music. Indeed, $S_1$ fails to distinguish a perfect major cadence ($C^{maj} \rightarrow F^{maj}$) from a plagal minor cadence ($F^{min} \rightarrow C^{min}$), as one proceeds from the other by involution with $I$. More generally, the eigentriad transform may suffice for extracting the quality of isolated chords (e.g. major vs. minor), but lacks longer-term context of harmonic tension and release so as to infer the tonal functions of such chords (e.g. tonic vs. dominant).

2. EIGENPROGRESSIONS

In this section, we introduce a second multidimensional feature $U_2$, built on top of $U_1$ and named eigenprogression transform, that aims at integrating harmonic context in Western tonal music while still respecting the aforementioned requirements of invariance to global temporal shifts $\tau_{\Delta t}$ and pitch transpositions $\pi_{\Delta p}$. We begin by defining the Tonnetz as an undirected graph over the unnormalized Laplacian tensor of the Tonnetz $\tau$ feature dominant). To infer the tonal functions of such chords (e.g. tonic vs. longer-term context of harmonic tension and release so as to quality of isolated chords (e.g. major vs. minor), but lacks

We define eigenvalues $\lambda$ of the Laplacian tensor, and corresponding eigenvectors $v$, as the solutions of the following equation:

$$(L^{Tonnetz} \otimes \varnothing)[p,q] = \sum_{p',q'} L[p,q,q',p']v[p',q'] = \lambda v[p,q].$$

We define eigenprogressions $\psi$ as the following rank-eight tensor:

$$(U_2)_{t,p,q,\alpha,\beta,\gamma} = \sum_{\gamma_1} \psi[p,q,\gamma_1] \exp\left(-\frac{\gamma_1^2}{2\sigma^2}\right) \exp\left(i\gamma_1 \xi \frac{p}{12}\right),$$

wherein $\gamma_2 \in \{0, \pm 1\}$. We define the eigenprogression transform of $x$ as the following rank-eight tensor:

$$U_2(x)[t,p,q,\alpha_1,\beta_1,\alpha_2,\beta_2,\gamma_2] = \sum_{t,p,q,\alpha,\beta,\gamma} U_1(x)^{\psi} \psi_{pro}[\alpha_2,\beta_2,\gamma_2]$$

At first sight, Equation 10 suffers from an identifiability problem. Indeed, a different choice of basis for $V_k$ would incur a phase shift and/or a complex conjugation of the convolutional response $U_1(x)^{\psi} \psi_{Tonnetz}^{\gamma_2}$. Yet, because the eigenprogression transform consists of Gabor wavelets (i.e. with a symmetric amplitude profile) and is followed by a complex modulus operator, such differences in phase and/ or spin are eventually canceled and thus have no effect on the outcome of the transform. Consequently, we pick one arbitrary pair $(v_{k}^{Re}, v_{k}^{Im})$ for each subspace $V_k$, without loss of generality.

3. EXPERIMENTS

We evaluate the eigenprogression transform on a task of supervised composer recognition between Haydn and Mozart string quartets [7]. After averaging along time $t$ and pitch $p$, we standardize each feature in the rank-five tensor by a complex modulus operator, such differences in phase and/or spin are eventually canceled and thus have no effect on the outcome of the transform. Consequently, we pick one arbitrary pair $(v_{k}^{Re}, v_{k}^{Im})$ for each subspace $V_k$, without loss of generality.

4. CONCLUSION

We have diagonalized the Laplacian of the Tonnetz graph and derived a multivariable scattering transform, named eigenprogression transform, that captures some local harmonic context in Western tonal music. Although the numerical example we gave was a task of composer recognition, the eigenprogression transform could, in principle, address other MIR tasks in the future, including song retrieval, key estimation, and structure analysis.
Table 1. Comparison between the eigenprogression transform and other transforms of smaller tensor rank, in terms of dimensionality, sparsity ($\ell_1/\ell_2$ ratio), and accuracy on a supervised task of composer recognition.

<table>
<thead>
<tr>
<th>α1</th>
<th>β1</th>
<th>α2</th>
<th>β2</th>
<th>γ1</th>
<th>dim.</th>
<th>$\ell_1/\ell_2$</th>
<th>acc. (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>[7]</td>
<td>8</td>
<td>2.6</td>
<td>67.3</td>
<td></td>
<td>79.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[8]</td>
<td>24</td>
<td>4.6</td>
<td>71.0</td>
<td></td>
<td>80.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>α1 β1</td>
<td>α2</td>
<td>β2</td>
<td>γ2</td>
<td></td>
<td>129</td>
<td>6.1</td>
<td>72.0</td>
</tr>
<tr>
<td>α1 β1</td>
<td>α2</td>
<td>β2</td>
<td>γ2</td>
<td></td>
<td>1677</td>
<td>17.0</td>
<td>76.7</td>
</tr>
<tr>
<td>α1 β1</td>
<td>α2</td>
<td>β2</td>
<td>γ2</td>
<td></td>
<td>8385</td>
<td>42.4</td>
<td>77.6</td>
</tr>
<tr>
<td>α1 β1</td>
<td>α2</td>
<td>β2</td>
<td>γ2</td>
<td></td>
<td>1119</td>
<td>22.3</td>
<td>82.2</td>
</tr>
</tbody>
</table>

5. REFERENCES


