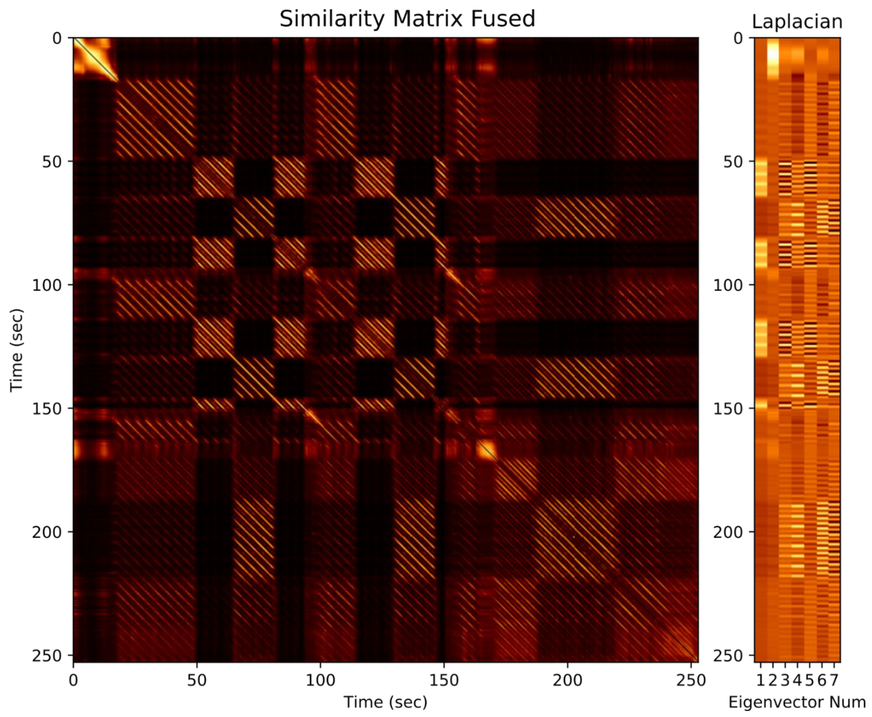
GraphDitty: A Software Suite for Geometric Music Structure Visualization

Christopher J. Tralie. ctralie@alumni.princeton.edu
Department of Mathematics, Duke University

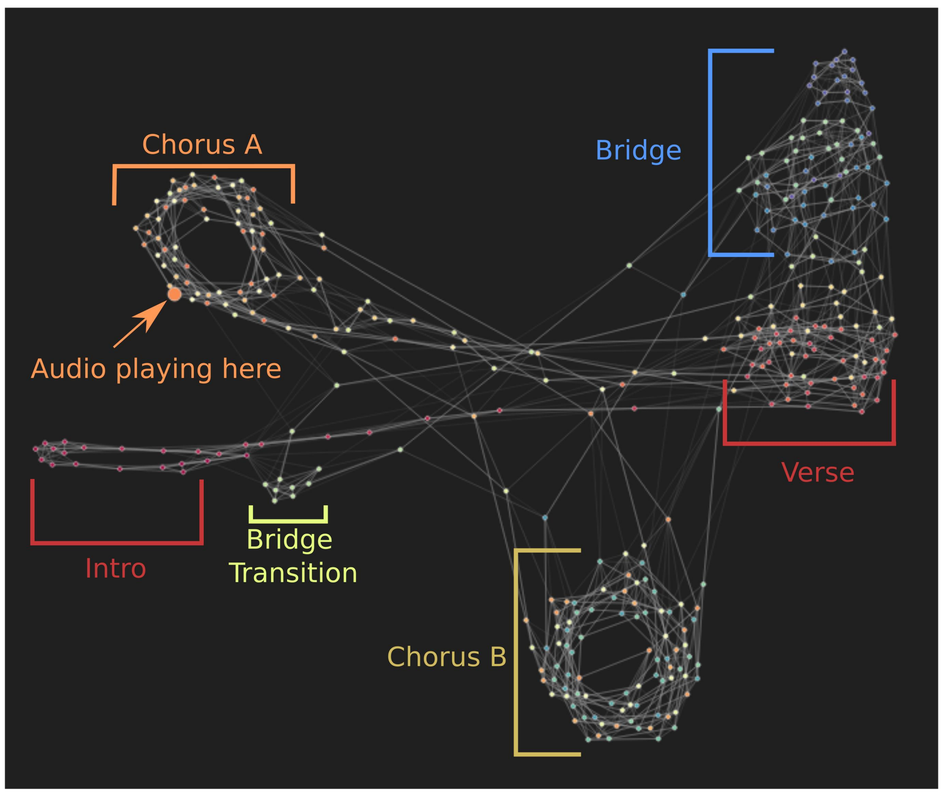
Clean Fused Similarity Matrices

- Structures built in Python (librosa/numpy), exported to Javascript for interaction
- 3 GUIs **synchronized to music**:
  - 1) Similarity matrix viewer
  - 2) Weighted spring graph layout viewer (d3.js)
  - 3) 3D Diffusion maps viewer (WegGL)



Interactive Weighted Spring Graphs in d3.js

Michael Jackson "**Smooth Criminal**"



Links

- Try it on your own songs!

CODE

https://github.com/ctralie/GraphDitty

LIVE WEB DEMO

http://www.covers1000.net/GraphDitty

Improving Self-Similarity Matrices with Early Similarity Network Fusion (SNF): Details

$$W(i,j) = e^{-\rho^2(i,j)/2(\sigma_{ij})^2}$$

1) Create similarity kernel given distance matrix from a feature type

$$P(i,j) = \begin{cases} \frac{1}{2}\frac{W(i,j)}{\sum_{k\neq i} W(i,k)} & j \neq i \\ 1/2 & \text{otherwise} \end{cases}$$

2) Compute self-similarity regularized Markov transition probabilities

$$S(i,j) = \begin{cases} \frac{W(i,j)}{\sum_{k\in N(i)} W(i,k)} & j \in N(i) \\ 0 & \text{otherwise} \end{cases}$$

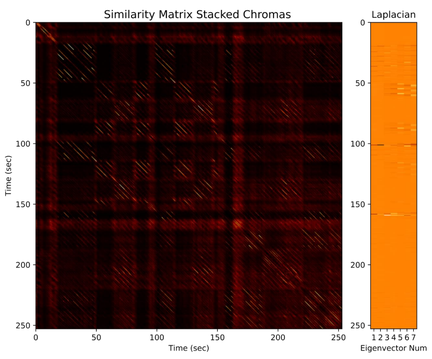
3) Compute neighborhood truncated Markov transition probabilities

$$P^f_{t+1} = S^f \left(\frac{\sum_{v\neq f} P^v_t}{m-1} \right) (S^f)^T$$

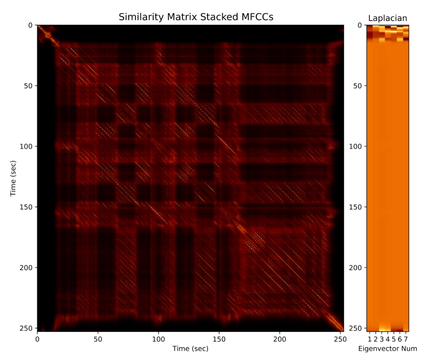
4) Compute random walk probabilities using average probabilities from other features and truncated transition matrix from feature f

- Similar to the scheme in [6], but on self-similarity matrices (SSMs) instead of cross-similarity matrices
- Use stack delayed Chromas (cosine distance) and MFCCs (Euclidean distance)
- Fusion leads to much clearer SSMs and clearer structure in Laplacian eigenvectors
- More general, global alternative to the diagonal promotion scheme in [5]
- Build weighted graph Laplacian on similarity matrix and compute eigenvectors

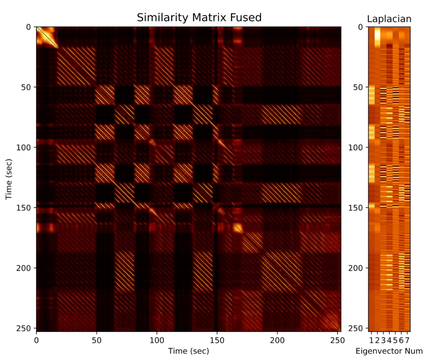
Chroma



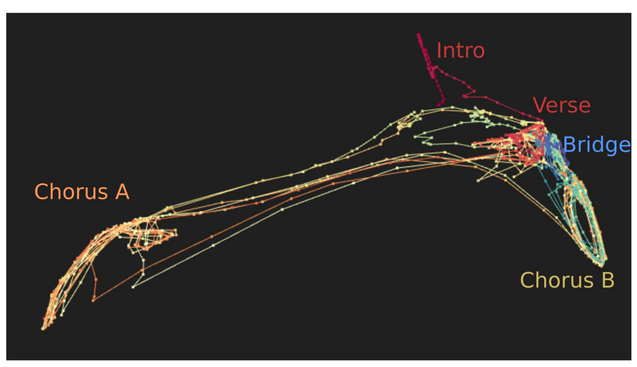
MFCC



Fused



3D Diffusion Maps



- An alternative to the graph Laplacian for nonlinear dimension reduction, which maps to a Euclidean space

Future Work

- Add interactive version of aligned hierarchies[4] to GUI
- Community interaction / Youtube integration
- Test out various techniques on *large scale audio cover song identification*

Acknowledgements

Christopher Tralie was partially supported by an NSF big data grant DKA-1447491 and an NSF Research Training Grant NSF-DMS 1045133.

References

[1] Paul Bendich, Ellen Gasparovic, John Harer, and Christopher Tralie. Geometric models for musical audiodata. InProceedingsofthe32stInternationalSymposium on Computational Geometry (SOCG), 2016.
[2] MichaelBostocketal.D3.js.DataDrivenDocuments, 492:701, 2012.
[3] Ronald R Coifman and St´ephane Lafon. Diffusion maps. Applied and computational harmonic analysis, 21(1):5–30, 2006.
[4] Katherine M Kinnaird. Aligned hierarchies: A multiscale structure-based representation for music-based data streams. In 16th International Society for Music Information Retrieval (ISMIR), pages 337–343, 2016.
[5] Brian McFee and Daniel PW Ellis. Analyzing song structure with spectral clustering. In 15th International Society for Music Information Retrieval (ISMIR), 2014.
[6] Christopher J Tralie. Early mfcc and hpcp fusion for robust cover song identification. In 18th International Society for Music Information Retrieval (ISMIR), 2017.
[7] Christopher J Tralie. Geometric Multimedia Time Series. Duke ph.d. dissertation, Department of Electrical and Computer Engineering, Duke University, 2017.
[8] BoWang,JiayanJiang,WeiWang,Zhi-HuaZhou,and Zhuowen Tu. Unsupervised metric fusion by cross diffusion. In Computer Vision and Pattern Recognition (CVPR),2012IEEEConferenceon,pages2997–3004. IEEE, 2012.
[9] Bo Wang, Aziz M Mezlini, Feyyaz Demir, Marc Fiume, Zhuowen Tu, Michael Brudno, Benjamin HaibeKains, and Anna Goldenberg. Similarity network fusion for aggregating data types on a genomic scale. Nature methods, 11(3):333–337, 2014.