**RESEARCH LOG**

11/8/19

* ISMR conference AMT code examples
  + Found Sonic Visualiser and Vamp Plugins
  + Chordino chord estimator – tried it out on a Bill Evans recording
* MiReX
* Cleaned up annotated references

12/8/19

* No work completed

13/8/19

* Try out more open source software
  + Search Vamp plugins
* Research sound wave fundamentals
  + ANU library resources – ANU SuperSearch through reverse proxy. Great for citations
  + Google Scholar is a great search engine
* Automatic chord estimation is an interesting subtopic in the field
  + ANU SuperSearch : Automatic Chord Estimation
* Cleaned up AMT tutorial notes
* Researched best performing MPE methods

14/8/19

* Reading article on future challenges – [ E. Benetos, S. Dixon, D. Giannoulis, H. Kirchhoff, and A. Klapuri, “Automatic music transcription: Challenges and future directions,” J. Intelligent Inform. Syst., vol. 41, no. 3, pp. 407–434, 2013 ]
* MiReX 2018 results
* <https://www.music-ir.org/mirex/wiki/2018:MIREX2018_Results>

15/8/19

* Finished article on future challenges – [ E. Benetos, S. Dixon, D. Giannoulis, H. Kirchhoff, and A. Klapuri, “Automatic music transcription: Challenges and future directions,” J. Intelligent Inform. Syst., vol. 41, no. 3, pp. 407–434, 2013 ]
* Search Automatic Music Transcription YouTube found lots of great resources
* Started to read - [M. Müller, D. P. Ellis, A. Klapuri, and G. Richard, “Signal processing for music analysis,” IEEE J. Sel. Topics Signal Process., vol. 5, no. 6, pp. 1088–1110, 2011]
  + Audio analysis fundamentals

18/8/19

* [M. Müller, D. P. Ellis, A. Klapuri, and G. Richard, “Signal processing for music analysis,” IEEE J. Sel. Topics Signal Process., vol. 5, no. 6, pp. 1088–1110, 2011]
* STFT review – try looking at some YouTube videos

19/8/19

* [M. Müller, D. P. Ellis, A. Klapuri, and G. Richard, “Signal processing for music analysis,” IEEE J. Sel. Topics Signal Process., vol. 5, no. 6, pp. 1088–1110, 2011]
* STFT review – try looking at some YouTube videos
* <https://www.youtube.com/watch?v=9boJ-Ai6QFM&t=95s>
  + Anna Wszeborowska – Music Transcription with Python
  + Forked github repo – rtmonoaudio2midi, great for prototyping
* <https://www.youtube.com/watch?v=tAECqx5i4oc>
  + Signal Processing Methods for Sound Recognition

20/8/19

* [M. Müller, D. P. Ellis, A. Klapuri, and G. Richard, “Signal processing for music analysis,” IEEE J. Sel. Topics Signal Process., vol. 5, no. 6, pp. 1088–1110, 2011]
* Revision of Fourier concepts logged in annotated references
* Installed librosa : <https://librosa.github.io/librosa/> in anaconda
  + Great library for MIR
* Found great python resource for MIR: <https://musicinformationretrieval.com/index.html>
* Successfully produced chromagram of a 12 bar blues recording
* IEEE referencing generator : [IEEE Referencing Generator: Citation Generator by Cite This For Me](http://www.citethisforme.com/ieee)
* DSP guide : [The Scientist and Engineer's Guide to Digital Signal Processing's Table of Content](http://www.dspguide.com/pdfbook.htm)

23/8/19

* Investigated window length effect on frequency and temporal resolution on a short 9 seconds sample of a blues recording. Check **STFT** for more details
* DSP online guidebook: <https://allsignalprocessing.com>
  + Great textbook for DSP – pay for full access

24/8/19

* Added a frequency chart for keys on a piano

25/8/19

* Finished [M. Müller, D. P. Ellis, A. Klapuri, and G. Richard, “Signal processing for music analysis,” IEEE J. Sel. Topics Signal Process., vol. 5, no. 6, pp. 1088–1110, 2011]
  + **Window length** of 3.8ms with **sample frequency** of 44.1kHz
  + Question on page 4 about separation of frequency bins ???
  + Great source for polyphonic MPE and source separation and overall review of signal processing techniques used in this field
* Constructed tempogram of sample blues recording
* Stanford Fourier series lectures : <https://www.youtube.com/playlist?list=PLB24BC7956EE040CD&fbclid=IwAR2sdIqFLPoC418xPVCS0pUrBpRPkg8bMeSwQiOpX_A-NuerQ-TdNVMDiXk>
* Fourier series Michel Van Biezen : <https://www.youtube.com/playlist?list=PLX2gX-ftPVXVTT3qpUeHVxbmJx9D1vGhC&fbclid=IwAR0Sle6SZggWXvFcs5roOZnVoU7AFe74Z0zcKHLUjUye0rWIL2gAIQxWr0c>
* Watched this lecture from Brian McFee the co-author of the librosa library
  + Library is very modular
  + Provides a number of spectral analysis, beat tracking and temporal structure analysis tools
* JAMS annotation library: <https://pythonhosted.org/jams/index.html>
  + Annotation library
* Mir\_eval library: <http://craffel.github.io/mir_eval/>
  + Metrics for evaluation of MIR tasks

26/8/19

* Dr. E. Benetos NUS tutorial - <https://www.youtube.com/watch?v=8eS82BdgTs4>
  + AMT lectures at NUS

27/8/19

* Dr. E. Benetos NUS tutorial - <https://www.youtube.com/watch?v=8eS82BdgTs4>
  + Finished Part 1,2,3,4
  + Part 5 didn’t have anything valuable to this project
* Wang Ye NUS lectures, Perceptual Features lectures and Audio signals : <https://www.youtube.com/watch?v=CnRTmJ9-P_o>
  + Stopped on part 2 perceptual features 23 mins

30/8/19

* Received access to MIREX dataset with annotated ground truth transcriptions
* Wang Ye NUS lectures, Perceptual Features lectures and Audio signals : <https://www.youtube.com/watch?v=CnRTmJ9-P_o>
  + Finished part 3 perceptual features

2/9/19

* Wang Ye NUS lectures, Perceptual Features lectures and Audio signals : [https://www.youtube.com/watch?v=0CtsogMICH0](https://www.youtube.com/watch?v=CnRTmJ9-P_o)
  + Up to Lecture 4: Machine Learning for MIR
* Finished relevant lectures on SMC NUS YouTube page <https://www.youtube.com/channel/UCn4iMJjVktUb6SgprG4Eijw>
  + Lecture 1 on DFT and lecture 2 on pitch detection with YIN algorithm

4/9/19

* Work on getting access to datasets: MIREX, MAPS, RWC
* Downloaded Alto Saxophone, Cello and piano instrument sample from the UOIMIS database
* Reading Signal Analysis and Feature Extraction: <https://musicinformationretrieval.com/index.html>

7/9/19

* NMF attempt for simple major scale

11/9/19

* Great databases:
  + TRIOS dataset : <https://c4dm.eecs.qmul.ac.uk/rdr/handle/123456789/27>
  + MusicNet : <https://homes.cs.washington.edu/~thickstn/start.html>
  + Labrosa access: <https://labrosa.ee.columbia.edu/projects/piano/>
* MusicNet has annotated ground truths for all the recordings
* Attempt at neural network solution for AMT

12/9/19

* Neural Network with 3 layers and 256 units wide and no dropout rate :
  + Precision : 0.81
  + Recall : 0.61
  + F-measure : 69.72
  + Accuracy : 0.53
* Example test/train data:
  + Cqt with dimensions number\_frames x number\_bins
  + Binary labels with dimensions number\_frames x number\_of\_notes

13/9/19

* Review Fourier and ML concepts

15/9/19

* Review completed research
* <http://www.dspguide.com/ch6/1.htm>
  + Convolution review
* NMF research

16/9/19

* Google Brain github repo : <https://github.com/tensorflow/magenta/tree/master/magenta/models/onsets_frames_transcription>
  + Online resource : <https://storage.googleapis.com/magentadata/papers/onsets-frames/index.html>
  + Notebook demonstrating use of magenta model : <https://colab.research.google.com/notebooks/magenta/onsets_frames_transcription/onsets_frames_transcription.ipynb#scrollTo=Z5SYRvIm8gq5>
  + Dataset developed by : <https://magenta.tensorflow.org/datasets/maestro>

**REFERENCE :** Curtis Hawthorne, Andriy Stasyuk, Adam Roberts, Ian Simon, Cheng-Zhi Anna Huang,Sander Dieleman, Erich Elsen, Jesse Engel, and Douglas Eck. "Enabling Factorized Piano Music Modeling and Generation with the MAESTRO Dataset." In International Conference on Learning Representations, 2019.

* Pytorch implementation of Google Brain approach : <https://github.com/jongwook/onsets-and-frames>
* DNN for AMT : <https://github.com/diegomorin8/Deep-Neural-Networks-for-Piano-Music-Transcription>

18/9/19

* Downloaded MAPS database
* Implemented NMF with ACF and peak picking for frequency and onsets

20/19/19

* Found resources on DNN music transcription including hyper parameter tuning, and model selection

22/9/19

* Worked on commenting DNN approach
* Properly implemented sigmoid activation function with binary cross entropy for output nodes
* Obtained loss and accuracy plots from model training

23/9/19

* Hyper parameter tuning of model
  + Batch size/ normalization
  + Number of layers/ nodes
  + Optimizers
  + Activation functions in hidden layers

24/9/19

* Investigate including LSTM layers in model for temporal dependencies
* Investigate CNN models

2/10/19

* Add DL articles and PCLA articles to annotated references
* Read musical source separation article : Cano et al 2019, MSS, IEEE SPS January 2019
* HMM lectures - <https://www.youtube.com/watch?v=LdPaXoy9w6E&t=2s>
  + Introduction on HMM

14/10/19

* Steve Tjao MIR algorithms - <https://www.youtube.com/watch?v=oGGVvTgHMHw&feature=youtu.be&fbclid=IwAR1RjW7ukpOYZArS9frHZzZlUcT0U-6I94euWpMF9-YcaJA2aZyFmUqzvXY>
  + Scikit Learn algorithms for MIR
  + Standford\_MIR
* Segio Oramos - <https://www.youtube.com/watch?v=4c5xRzd-I1I&list=PLZSWqxOGTnJ1Gw8Cyv1OCCDynm5S7IawC>
  + Lectures on Natural Language Processing

15/10/19

* Segmentation pipeline with feature extraction and clustering of sound events
* Basic 251 used chroma maximum and mean of spectral centroid
* SONIFY and MIR\_EVAL of estimated melody

20/10/19

* CEPSTRUM analysis
* YIN algorithm

23/10/19

* Finished mid-term report
* CEPSTRUM, NMF with peak differences and autocorrelation methods

31/01/20

* Overview of research completed
* Review mid-term report

1/2/20

* <http://www.dspguide.com/pdfbook.htm>
  + Chapter 5 Linear Systems
* Review work done so far

2/2/20

* Consider zero padding segmented signals before performing DFT to increase frequency resolution. Try different windows to see if that increases performance.
* Write plan for thesis structure
* <https://www.youtube.com/channel/UCMXeBLqN2WHKumtMhylOS1A/videos>
  + Good resource for DSP and Speech processing
* 13 characteristics of sound - R. Cogan, New Images of Musical Sound. Cambridge, MA: Harvard Univ. Press, 1984

3/2/20

* M. Müller, D. P. Ellis, A. Klapuri, and G. Richard, “Signal processing for music analysis,” IEEE J. Sel. Topics Signal Process., vol. 5, no. 6, pp. 1088–1110, 2011.
  + Polyphony and Musical Voices – relevant section
  + pg 11. Supervised classification algorithms for instrument recognition
  + pg 13. Principles of musical source separation
  + pg. 15 review of MPE – highly relevant for project