Lab 2 - Crossfade Product Specification

Mohamed Abdullahi

CS 411

Professor Thomas J. Kennedy

December 1, 2023

Version 2

Table of Contents

List of Figures

List of Tables

1. Introduction

1.1 Purpose

Crossfade is a PC GUI that makes it easier for aspiring musicians to get started and makes it possible for them to share their music visually. Crossfade supports all levels of musical literacy since it fixes transcription problems in addition to visualizing your music. With capabilities to support the production of their unique pieces, users have complete control over the final product's appearance.

1.2 Scope

The Crossfade prototype will continue to focus on transcription correction as its primary function. The prototype will give the user a full look at their music sheets, allowing the feature to compare your original piece and the Crossfade corrected piece. Although mistake detection and repair tools will be included, they will only apply to single-track transcriptions. To focus on transcription correction, live audio input features will not be provided in the prototype.

[This space is intentionally left blank]

1.3 Definitions, Acronyms and Abbreviations

Convolutional Neural Network (CNN) - Deep learning algorithm which can differentiate one image from another by assigning weights and biases to different aspects of the images. It is used in audio to differentiate different frequencies in a visual format.

Deep Learning - Subfield of machine learning which uses neural networks to solve complex problems. Learning comes directly from the data, instead of being hand-engineered by humans.

Keyboard - An electronic piano used to produce sound and MIDI information.

Monophony - A phrase of music in which only a single voice is played at a time.

Musical Instrument Digital Interface (MIDI) - A communications protocol used to connect physical and virtual music devices and instruments. MIDI files store note information which can be used to trigger instruments and devices.

MusicXML - A markup language format used to interchange and distribute digital sheet music.

Polyphony - A phrase of music in which more than a single voice is played at a time.

PyTorch- An open source machine learning framework that accelerates the path from research prototyping to production deployment.

Music21 - is Michael Cuthbert's library of programs ("modules") allowing the Python computer language to read, write, and manipulate musical scores.

[This space is intentionally left blank]

1.4 References

Automatic Music Transcription and Ethnomusicology: A user study. (n.d.). Retrieved from https://diva-portal.org/smash/get/diva2:1474663/FULLTEXT01.pdf

H. Takeda, T. Nishimoto and S. Sagayama, "Rhythm and Tempo Analysis Toward Automatic Music Transcription," 2007 IEEE International Conference on Acoustics, Speech and Signal Processing - ICASSP '07, Honolulu, HI, USA, 2007, pp. IV-1317-IV-1320, doi: 10.1109/ICASSP.2007.367320.

Huang, Z., Jia, X., & Guo, Y. (2019, June 29). State-of-the-art model for music object recognition with Deep Learning. MDPI. Retrieved February 8, 2023, from https://www.mdpi.com/2076-3417/9/13/2645

Jovanovic, J. (2015, February 2). How does Shazam work? music recognition algorithms, fingerprinting, and processing: Toptal®. Toptal Engineering Blog. Retrieved February 8, 2023, from https://www.toptal.com/algorithms/shazam-it-music-processing-fingerprinting-and-recognition

Scarlatos, L. L. (n.d.). Continuous media. Audio. Retrieved March 1, 2023, from https://www3.cs.stonybrook.edu/~lori/classes/GUI/sound.htm

Solanki, A., & Pandey, S. (2019, January 30). Music instrument recognition using deep convolutional neural networks - International Journal of Information Technology. SpringerLink. Retrieved February 8, 2023, from https://link.springer.com/article/10.1007/s41870-019-00285-y

Zhang, X. (2022, March 11). Aided recognition and training of music features based on the internet of things and Artificial Intelligence. Computational intelligence and neuroscience.

Retrieved February 8, 2023, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8933112/

Ruth N, Müllensiefen D (2021) Survival of musical activities. When do young people stop making music?. PLOS ONE 16(11): e0259105. https://doi.org/10.1371/journal.pone.0259105

Christman, E. (2017, November 17). Nielsen 360 study finds consumers love streaming music, but radio still strong. Billboard. https://www.billboard.com/pro/nielsen-music-360-2017-report-streaming/

[This space is intentionally left blank]

1.5 Overview

This product specification provides the hardware and software configuration, external interfaces, and features of Crossfade prototype. The remaining sections will provide a detailed description of each feature and the requirements for their implementation.

2. General Description

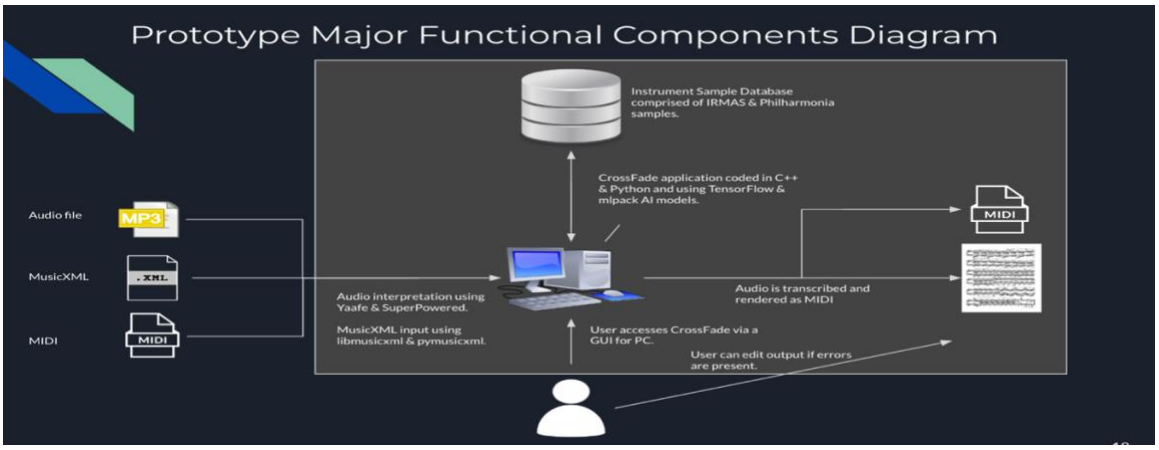
The Crossfade prototype will be developed as a GUI PC application with algorithms for transcription error detection specifically note-beam detection. Most features from the real-world product will not be included, the error detection and user solution selection will remain in the Crossfade prototype.

2.1 Product Architecture Description

The Crossfade prototype will allow to have an interactive frontend, where the user can input an XML or Midi file, which would get processed by a Music21 Toolkit. Transcription correction will be handled by a deep learning program called PyTorch, trained with a library of musicXML samples. The major components are shown in Figure 1.

Figure 1

Crossfade Major Functional Component Diagram



2.2 Prototype Functional Description

The Crossfade prototype will implement all the main features involving the transcription error correction of the real-world product**.** Although the error transcription will be implemented, it will be limited to 2 kinds of music notation errors. As shown in Table 1, the prototype will feature the acceptance of certain input files. The input files will be limited to shorter music pieces. Features such as note, and instrument recognition will not be implemented into the prototype.

Table 1

Crossfade Real World Product vs. Prototype Features Table

Transcription

Feature	Real World Product	Prototype
Live Audio Transcription	✓	
MIDI Input/File Transcription	✓	
MIDI Transcription Correction	✓	✓
Multitrack Transcription	✓	
MusicXML Compatibility	✓	✓
MIDI File Export	✓	✓

Note Recognition

Feature	Real World Product	Prototype
Monophonic Note Recognition	✓	
Polyphonic Note Recognition	✓	
Instrument Recognition	✓	
Instrument Distinction	✓	

Transcription Error Correction

Feature	Real World Product	Prototype
Transcription Error Correction	✓	✓
Highlight Possible Errors	✓	✓
Take User Feedback	✓	✓
Offer Possible Solutions	✓	✓
Compare Single Note from Original to Written	✓	✓
Compare a Segment from Original to Written	✓	✓

2.3 External Interfaces

The Crossfade prototype will utilize a couple of hardware, software, and user interfaces, to provide the user with a seamless experience. A machine-learning tool will be used in training the auto transcription correction for music sheets.

2.3.1 Hardware Interfaces

The prototype will need to be run on a desktop or laptop with an operating system of Windows or MacOS.

2.3.2 Software Interfaces

The prototype will utilize PyTorch for machine deep learning to provide auto correction for music transcription.

2.3.2 User Interfaces

The prototype will be accessed through a downloadable GUI and can be resized to fit most resolutions. The prototype will have a main page for file input loading and an editor page for music sheets.