

Lab 1 - Crossfade Product Description

Mohamed Abdullahi

CS 411

Professor Thomas J. Kennedy

November 17, 2023

Version 2

Table of Contents

1. Introduction.....	3
2. Product Description.....	4
2.1 Key Product Features and Capabilities.....	4
2.2 Major Components (Hardware/Software).....	5
3. Identification of Case Study.....	6
4. Product Prototype Description.....	7
4.1 Prototype Architecture (Hardware/Software).....	7
4.2 Prototype Features and Capabilities	8
4.3 Prototype Development Challenges.....	9
5. Glossary.....	10
6. References.....	11

List of Figures

Figure 1: Crossfade Major Functional Component Diagram.....	5
Figure 2: Crossfade Phase 1 Prototype Major Functional Component Diagram.....	7

List of Tables

Table 1: Feature Comparison Between Full Product and Prototype.....	8
---	---

1. Introduction

Music is a beautiful thing that has touched all parts of the globe. From an impromptu desk drumming to fully mastered songs, they are all considered music. A study from Nielsen Music revealed, that 90% of the U.S population listens to music, averaging 32.1 hours weekly and spending \$156 annually (Christman, 2017). This does not include the cost of attending concerts, merchandise and deluxe albums. Proving that being a musician is a very profitable profession.

Although the benefits are great in being a musician and close to the whole population enjoys music, reality proves otherwise. In a study taken in schools in the U.K and Germany showed students, by the time they are 17, roughly half of them stop taking music lessons and participating in other musical activities, with most students stopping between the ages of 15 and 17 (Ruth et al., 2021). So, for students passed the age of 18 and have rekindled their love for music, not only missed their chance of increasing their musical literacy with the help of educational institutes, but now must face a steep learning curve as a starting musician. This is followed by technical tools and software needed to record and share their creations. With such overwhelming conditions, it is easy for starting musicians to be disheartened, leading them to once again stop music.

To deal with these pressing issues, comes Crossfade. It is a transcription application, that aims to lower the barrier for starting musicians, by providing tools to overcome their unfamiliarity with music sheets, lack of sound boards and similar equipment, and the burnout associated with learning or creating music. The features provided by Crossfade are meant to give the musician all the tools they need to not only understand, but share their music to their hearts content, in a simple and easy fashion.

2. Product Description

Crossfade is a GUI for PC, that lowers the barrier for starting musicians, while also allowing new music creators to easily share their music in a form that's visual. All levels of musical literacy are embraced as Crossfade not only visualizes your music but corrects any errors that are present in your transcription. For the more advanced musician's, functionality centered on multi-track and complex musical inputs are welcomed. Users are given full control on how their final product should look with features to accommodate the completion of their unique pieces.

2.1. Key Product Features and Capabilities

The key features of Crossfade are centered on providing completed transcriptions, regardless of the user's knowledge of what they are playing. From randomly played keys to personal pieces, musicians will be able to visualize the audio they played with errors being highlighted in orange on their notation sheet. Transcription correction will be handled by A.I, extensively trained with a vast library of instrument samples to give precise results. If the musician is still not satisfied with the corrections and do not want to exchange a certain note, they will be able to highlight certain notes and keys in with a blue ticker and can either revert or use other possible options. Musicians will then be able to compare how their original piece sounds compared to their current one.

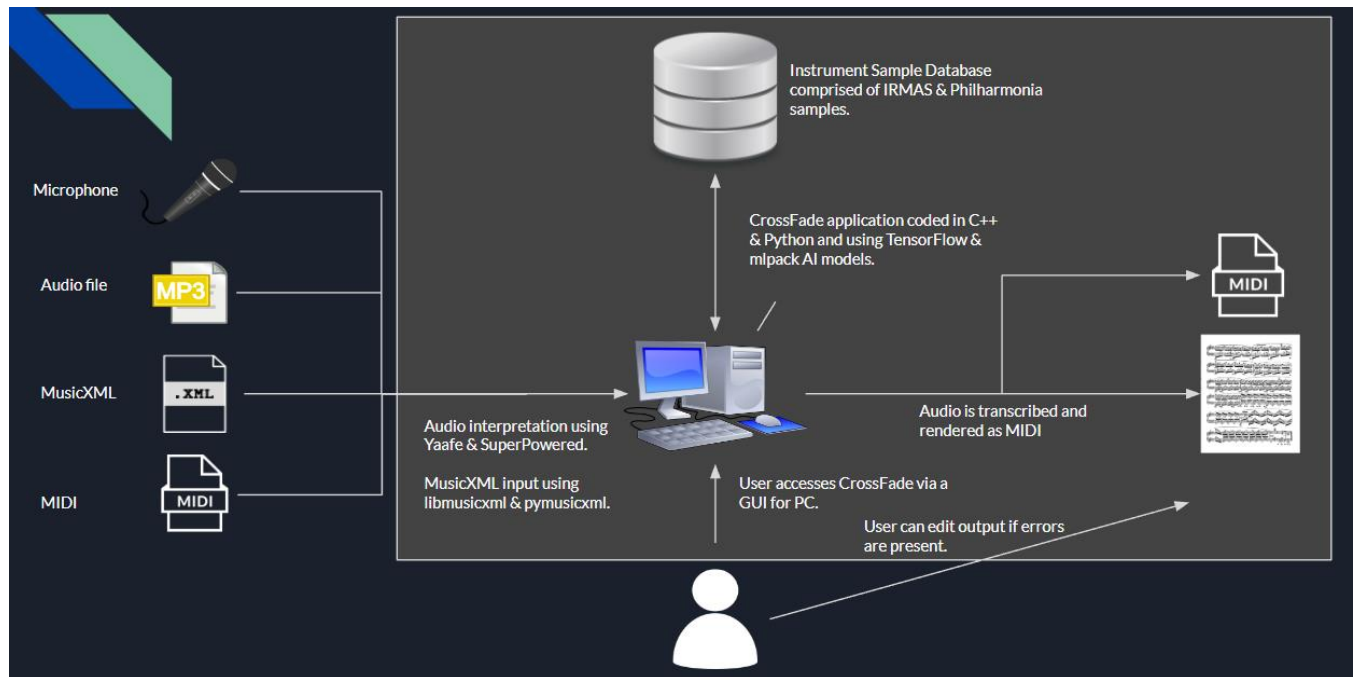
The advanced musicians, will have access to features allowing them to work on many pieces at once, allowing for optimal use in respect to their time. Once the changes to their notation sheets are deemed to be completed, the musician's will then be able to export their works in a form that will easily be sharable.

2.2. Major Components (Hardware/Software)

The major components of Crossfade will include all the features needed to bring out the full extent of transcription correction, as seen in Figure 1. In the form of an executable python program, the user will be allowed to capture the live audio of the instrument they are playing with a mic or load prerecorded instruments in the form of audio, MIDI and MusicXML files. Live audio and audio files will then be handled by audio parsing and interpretation software, while MusicXML files will have full library support. Once the input is collected with the help of A.I and deep learning, such as TensorFlow and mlpack trained on instrumental databases, transcription error detection and correction will be applied, outputting a completed transcription in the form of a MIDI file.

Figure 1

Crossfade Major Functional Component Diagram



3. Identification of Case Study

The core users Crossfade will be focusing on in a case study will be starting musicians and music creators. The primary focus will be starting musicians to investigate the adaptability and ease of use in response to time. Starting musicians will be provided with an easy to navigate UI's and supplemental features tutorials to curb the learning curve of using Crossfade.

The second focus would be music creators with prior knowledge and experience under their belt. Crossfade's goal is to alleviate and expediate their current or new projects. The goal here is to see if the use of Crossfade can grant optimal performance and flush out any bugs from difficult and high-level operations. Crossfade will check if the music creators achieved the desired result, they set out to achieve or the degree of deviation.

For this case study, there will be three users in each group consisting of different levels. Starting musicians will be categorized between low (no idea of notes being played), moderate (somewhat idea of notes being played) and high (fully knowledgeable of notes being played). Music creators will be categorized by experience between low (no experience of music programs), moderate (somewhat experience of music programs) and high (expert level of music programs). For both groups, time will be noted regarding user acclimation. The starting musicians will be tested on producing simple pieces, whereas the music creators will be more diverse and complicated pieces.

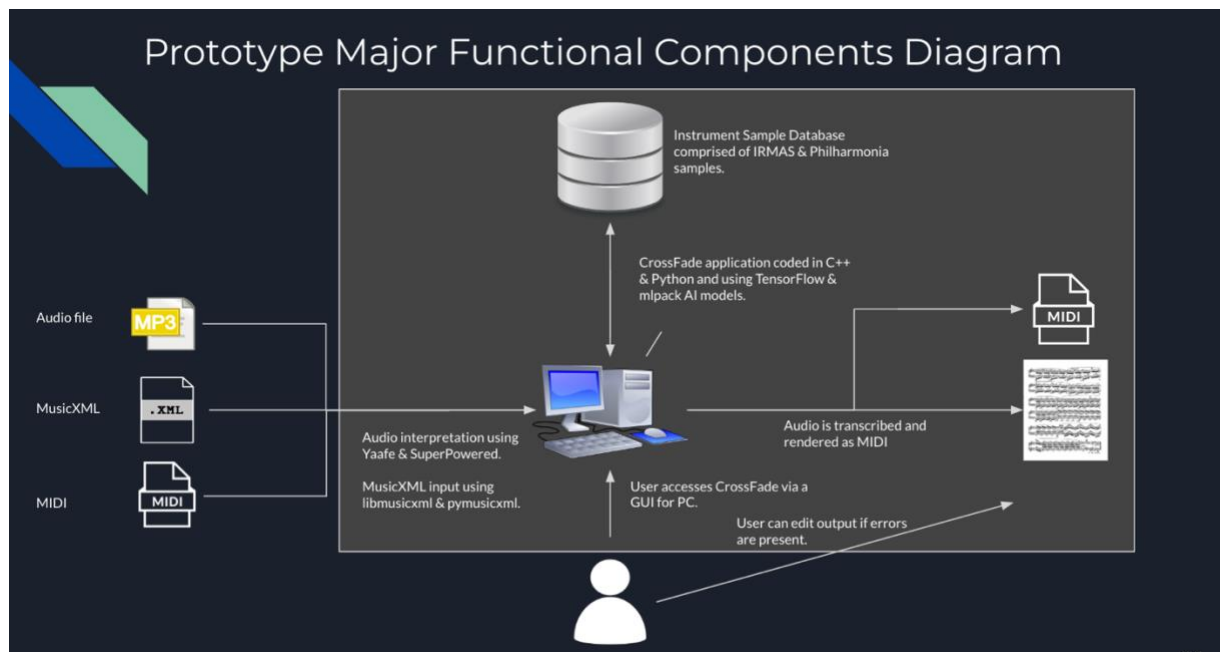
[This space is left intentionally blank]

4. Product Prototype Description

The Crossfade prototype will stay true to its core purpose of transcription correction. While features related to error detection and correction will be available, it will only be so for single-track transcriptions. Features including live audio input will not be included in the prototype for the interest of concentrating on correcting transcription. Figure 2 shows the changes that will be made for the prototype.

Figure 2

Crossfade Prototype Major Functional Component Diagram



4.1 Prototype Architecture (Hardware/Software)

The Crossfade prototype will be implemented as an application accessible via GUI for PC. It will be coded in C++ and Python, utilizing TensorFlow and mlpack machine learning software. The database consists of free sound libraries to assist in deep learning. Audio files, MusicXML and MIDI will be the only accepted inputs, with the transcribed audio exported as MIDI.

4.2 Prototype Features and Capabilities

The Crossfade prototype features table is broken into three tables. As shown in Table 1, compared to the Real-World Product, the prototype will feature the acceptance of certain input files. With transcription error correction being the backbone of Crossfade, all features will be fully implemented just as the Real-World Product. 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			Note Recognition		
Feature	Real World Product	Prototype	Feature	Real World Product	Prototype
Live Audio Transcription	✓				
MIDI Input/File Transcription	✓		Monophonic Note Recognition	✓	
MIDI Transcription Correction	✓	✓	Polyphonic Note Recognition	✓	
Multitrack Transcription	✓		Instrument Recognition	✓	
MusicXML Compatibility	✓	✓	Instrument Distinction	✓	
MIDI File Export	✓	✓			

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	✓	✓

4.3 Prototype Development Challenges

The developmental challenges Crossfade prototype will face are numerous, with the team's coding experience and semester bound time constraints being the top of the list. The user GUI being extremely basic might pose a moderate learning curve for the average user. Without a dedicated experienced development team, fully implementing the multi-pitch detection and note tracking feature will pose great difficulty, thus failing to bring out the full system performance and musical diversity. The prototype may not completely recognize user input and library availability will be affected by intellectual property laws. Though the prototype will be focused on displaying transcription correction, it is not guaranteed without incorrect correction laws and bias correctional process.

[This space is left intentionally blank]

5. Glossary

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.

6. 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>
- Scarlato, 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/>