Contents

1 Introduction

The advent of digital signal processing (DSP) has revolutionised the music industry, particularly with the development of tools like Auto-Tune and pitch correction. This research aims to design interactive visualisation tools that assist singers in the key transposition process, leveraging the Auto-Tune algorithm. The goal is to enhance the singing experience by providing real-time visual feedback and improving pitch accuracy.

2 Literature Review

This section reviews existing literature in three main domains relevant to this research: the application of Auto-Tune in music, the algorithmic and computational aspects of Auto-Tune, and the development of interactive visualization tools in musical contexts.

2.1 Application of Auto-Tune in Music

As a digital signal processing (DSP) tool, Auto-Tune has revolutionized vocal performance in both studio recordings and live settings. In pop music, for example, its primary use has been to correct pitch inaccuracies, ensuring that vocal tracks meet industry standards for "perfection". This application is highlighted in the paper. Provenzano (2019) explore the intersection of gender, race, and technology in Auto-Tune's use. The study discusses how female voices are often subjected to pitch correction to meet a standard of vocal perfection, while male voices using Auto-Tune are frequently viewed as more artistic or emotionally expressive. This paper underscores how Auto-Tune influences perceptions of vocal authenticity and emotional delivery, especially in a gendered context.

In a different genre, Duinker (2025) explores how Auto-Tune has been repurposed in trap music, where its role transcends mere pitch correction. In trap, Auto-Tune is not just a tool for pitch correction but an instrument that modifies vocal timbre, enabling new forms of emotional expression. The study highlights the technological repurposing of Auto-Tune, much like earlier hip-hop producers' use of turntables and samplers, to craft unique sonic identities. Auto-Tune in this context allows for emotive vocal expressions that challenge traditional notions of vocal authenticity, blurring the line between human and machine.

However, studies often treat Auto-Tune as a standalone processing tool without exploring its broader implications or potential integrations with other technologies. For example, in the context of traditional music, Bozkurt (2012) proposes a system that combines pitch correction and feedback mechanisms to help musicians tune instruments based on recordings of master musicians rather than theory-based frequency presets. This approach demonstrates a more interactive use of Auto-Tune-like technologies, offering both audio and visual feedback to musicians. The system goes beyond pitch correction, allowing users to adjust to non-western tuning systems with more flexibility, underscoring the potential for Auto-Tune's integration into educational and musical traditions.

In summary, Auto-Tune's applications are influencing various music genres and even educational tools. From gendered perceptions in pop music to its repurposing as an instrument in trap, and its role in traditional music systems, Auto-Tune's impact is multifaceted and evolving. However, many studies still overlook its potential to be integrated into user-interactive systems that could enhance both the creative process and educational environments for musicians.

2.2 Algorithmic and Computational Aspects of Auto-Tune

From a technical perspective, Auto-Tune relies on sophisticated pitch detection and correction algorithms, which have evolved significantly over time. Brown (2012) reveals how similar Fourier-based methods used in Auto-Tune can be applied to other fields, such as oceanography, demonstrating the wide-ranging utility of these algorithms.

Additionally, the deep learning-based approaches in recent studies, such as moving beyond traditional algorithms by leveraging neural networks to improve pitch correction accuracy (Wager et al., 2020). These models analyse spectrograms of the vocals and accompaniment, learning to predict the necessary pitch shifts, and have shown promise in addressing the limitations of earlier systems. However, despite these advancements, the challenge of handling large pitch deviations and maintaining natural-sounding performances remains an area for further refinement.

A recent study of Marshall (2024) traces the historical and technological intersections between the oil industry, vocal processing, and national security. The paper argues that both geophysical exploration and digital vocal processing technologies, particularly Auto-Tune, emerged from similar signal processing techniques originally developed for oil exploration and later adapted for voice technologies.

Thus, while technical innovations in Auto-Tune and related pitch correction technologies are advancing rapidly, there is a need for more research that integrates both the computational efficiency of these systems with an understanding of user experience, interaction design, and genre-specific applications.

2.3 Interactive Visualization Tools in Music

Interactive visualization has become an effective means of enhancing music learning and performance. The role of interactive visualization in music education has been explored further in various recent studies. For example, Park (2025) presented the development and evaluation of the Musical Dynamics Visualization Method (MDVM), which allows musicians to compare their performance with a pre-recorded track through real-time visualization of musical dynamics. This tool significantly improved the accuracy of dynamic expression, reducing dynamic errors by approximately 256% and enhancing both emotional engagement and learning outcomes. The MDVM is a promising advancement in music education, particularly in enhancing students' understanding of musical dynamics, though it focuses primarily on dynamics and overlooks other elements like pitch and rhythm.

Further extending the application of visualization in music learning, the study of De Prisco et al. (2016) explored the use of visualization to help individuals with limited music theory knowledge understand and compose classical music, specifically four-part chorale compositions. The tool, VisualMelody, provides real-time feedback on melodic structure, helping users avoid common errors such as parallel fifths, octaves, unisons, and voice crossings. The study showed that participants using the visualization tool were 60% faster in completing tasks and made 85% fewer errors than those without. However, the tool's application is limited to specific melodic rules, and future work could expand its scope to enhance usability further.

In addition, the integration of deep learning into music visualization has shown potential in more advanced applications. Liao (2022) focused on pairing music with images using CNNs and LSTMs, improving the alignment of music and images through an emotion classification loss function. This research highlights the growing potential of AI-driven visualization tools, which could be extended beyond static educational environments to more dynamic, real-time applications. However, challenges remain, such as the reliance on manual sentiment labeling and the subjectivity of emotional interpretation, which may influence the accuracy of music-image pairing.

Overall, these studies highlight the broad potential of interactive visualization in music education and performance. Future developments could focus on expanding the range of musical elements analysed, integrating real-time DSP applications like Auto-Tune, and refining emotional interpretation techniques for deeper engagement and accuracy.

2.4 Conclusion

Auto-Tune has had a profound impact on contemporary music, with its applications extending beyond pitch correction to influence musical genres, vocal expression, and even music creation. In pop music, Auto-Tune's role in pitch correction has sparked discussions around gender and authenticity, while in trap music, its repurposing as an emotional expression tool has blurred the lines between human and machine. Additionally, Auto-Tune's integration into traditional music systems has opened up new possibilities for interactive learning, offering more flexible and personalized tuning methods for music creation.

From a technical perspective, Auto-Tune's algorithms, initially based on Fourier transforms, have seen significant improvements with the introduction of deep learning techniques, enhancing pitch correction accuracy and efficiency. However, challenges remain, such as handling extreme pitch deviations and ensuring a natural vocal performance. As these algorithms continue to evolve, there is an increasing need for research that not only refines the computational aspects of pitch correction but also addresses user experience, especially in diverse musical contexts.

In music creation, the rise of interactive visualization tools offers a promising direction for enhancing musical performance and creation. These tools provide real-time feedback and visualization of pitch modulation, significantly improving the accuracy of pitch control and emotional engagement. The integration of deep learning technologies into music visualization is pushing the boundaries of the field, enabling more personalized, dynamic, and emotionally resonant creative experiences.

My research, "Designing Interactive Visualization Tools for Key Transposition in the Singing Process Based on an Auto-Tune Algorithm," combines these three directions. The significance of this research lies in integrating the Auto-Tune algorithm with interactive visualization technology to design a system that provides real-time feedback on pitch transposition, helping singers better understand and apply pitch adjustment techniques. The innovation lies in combining pitch correction algorithms with transposition visualization, offering not only precise pitch adjustment but also visually presenting the process of pitch variation, enhancing both the accuracy and emotional expression in music creation and performance. This combination provides a new interactive approach to music creation and performance, promoting the organic integration of tuning technology and visualization, and opening up new innovative directions in music creation.

3 Research Objectives

The primary objective of this research is to design and develop an interactive visualization tool for key transposition in the singing process based on the Auto-Tune algorithm. The research will focus on the following objectives:

1. **Real-time Pitch Adjustment Feedback**: To provide singers with real-time feedback on pitch adjustment, enhancing their understanding of pitch transposition techniques.

2. **Visualization of Pitch Variation**: To create a visual representation of how pitch changes during the singing process, allowing singers to see the effect of transposition.

3. **Integration of Auto-Tune's Pitch Correction with Transposition Visualization**: To combine the pitch correction capabilities of Auto-Tune with a visualization tool that demonstrates pitch variations, offering both accurate pitch adjustment and visual feedback.

4. **Support for Singers of All Levels**: To design the tool to help both novice and professional singers improve their pitch accuracy and emotional expression, offering a new interactive approach to music creation and performance.

This research will provide an innovative and interactive system that supports vocalists in improving their performance and understanding of pitch transposition.

4 Methodology

This research adopts a multi-stage methodology to design, develop, and evaluate an interactive visualization tool for key transposition in singing, leveraging Auto-Tune technology. The process includes signal acquisition, audio processing, feature extraction, visualization design, and user-centered evaluation.

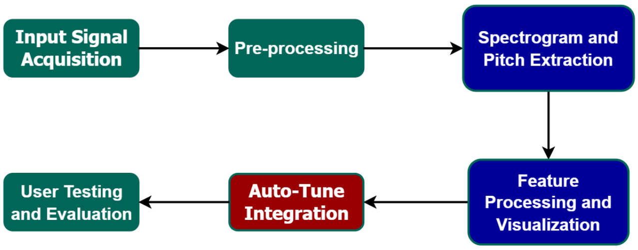


Figure 4.1: Feature Processing and Visualization

4.1 Input Signal Acquisition

The input signal consists of real-time vocal recordings from users, captured through a high-quality microphone at a sampling rate of 44.1 kHz. This ensures sufficient frequency resolution for pitch analysis and key transposition.

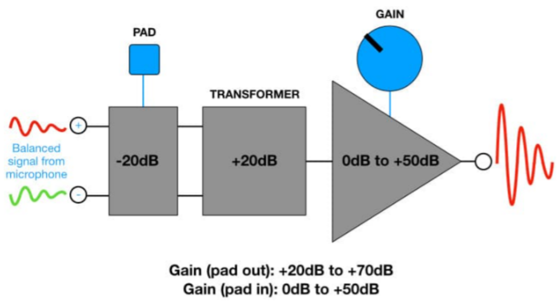


Figure 4.2: Get input signal by microphone (Simmons, 2022)

4.2 Pre-processing

Before analysis, the vocal input undergoes the following pre-processing steps:

- **Noise reduction:** Background noise is reduced using band-pass filtering (e.g., 80–8000 Hz) and spectral subtraction, which removes estimated noise from each frame.
- **Normalization:** o ensure consistent loudness, peak or RMS normalization is applied, scaling the waveform so its amplitude or energy matches a

target level (e.g., peak at 0.99 or fixed RMS). This allows fair comparison between recordings.

- **Framing:** The audio is split into overlapping frames (e.g., 50 ms with 20 ms overlap) using a Hamming window, enabling accurate time-frequency analysis and pitch tracking while minimizing spectral leakage.

4.3 Spectrogram and Pitch Extraction

Using Melodyne or custom DSP modules, the system performs:

- **Spectrogram generation:** Short-Time Fourier Transform (STFT) is used to generate time-frequency representations.
- **Pitch extraction:** Algorithms such as YIN or CREPE are used to derive the fundamental frequency (F0) trajectory.
- **Key detection:** Extracted pitch values are compared against standard musical scales to detect the performed key.

Note	MIDI pitch code
C4	60
E4	64
G4	67

Figure 4.3: Pitch Extraction (Faghih, 2022)

4.4 Feature Processing and Visualization

After extracting pitch and key data, the system processes and visualizes these features:

- **Pitch deviation calculation:** The deviation between the actual pitch and the target pitch is measured in cents.
- **Smoothing filters:** Applied to pitch curves to improve visual clarity.
- **Interactive visualization:**
  - Display of actual vs. target pitch trajectories.
  - Use of color gradients and directional cues to represent pitch offset.
  - Animated representations of key transpositions (e.g., piano roll, staff notation, color blocks).
  - Real-time feedback components (e.g., pitch correction suggestions, accuracy indicators).

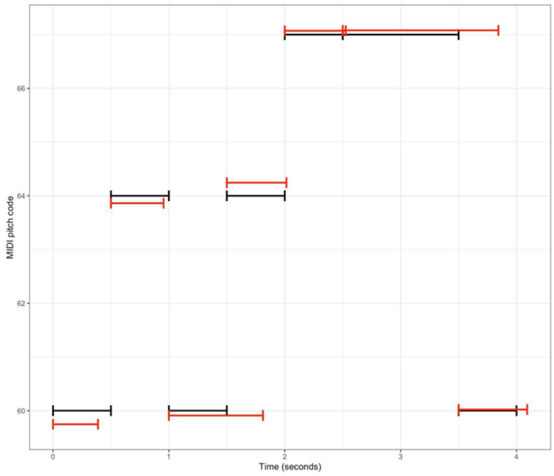


Figure 4.4: Feature Processing and Visualization (Faghih, 2022)

4.5 Auto-Tune Integration

Auto-Tune functionality will be incorporated via Melodyne software or its API, supporting the following:

- **Real-time pitch correction:** Adjustable strength of Auto-Tune correction to compare natural and corrected vocals.
- **Before-and-after analysis:** Visual comparison of pitch accuracy with and without correction.
- **Interactive feedback:** Users are guided through pitch deviations and transposition in real-time.

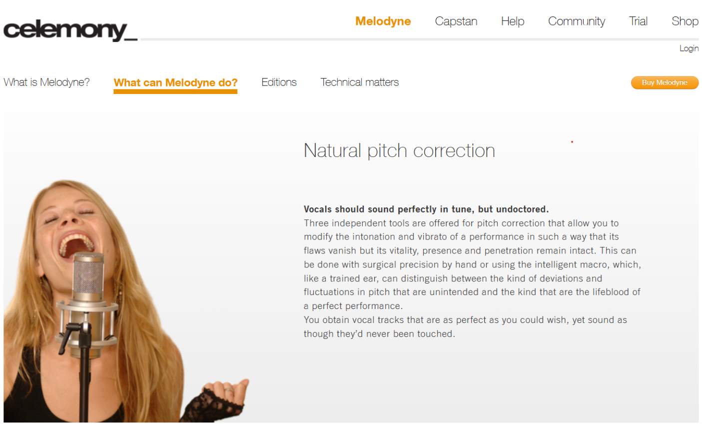


Figure 4.5: Melodyne Software to change tune

4.6 User Testing and Evaluation

To assess the tool's effectiveness, both qualitative and quantitative evaluations will be conducted:

- **Participants:** A sample group of approximately 30 individuals, covering beginners to advanced singers.
- **Procedure:**
  1. Baseline singing without using the tool.
  2. Singing with real-time pitch visualization and Auto-Tune.
  3. Recording and analysis of pitch data.
  4. Surveys and interviews to gather feedback.
- **Evaluation metrics:**
  - Pitch accuracy (cents deviation reduction).
  - Pitch stability (F0 variance).
  - User satisfaction (Likert-scale ratings).
  - Learning effectiveness (task completion time, qualitative user feedback).

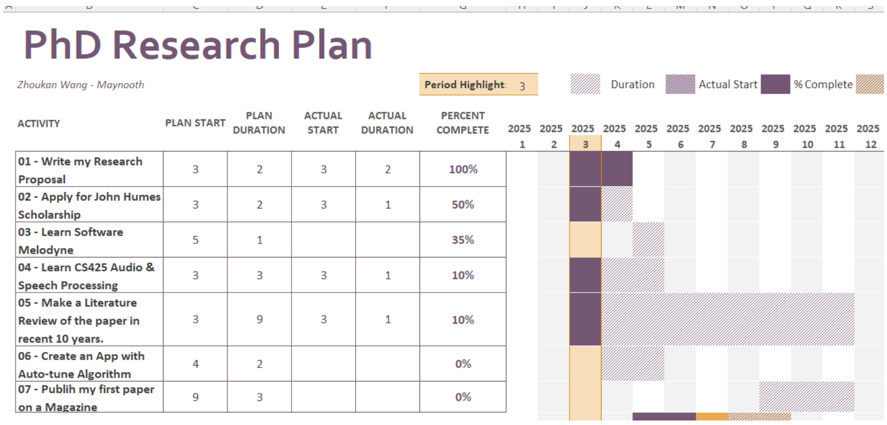
5 Expected Outcomes

The expected outcome of this research is the creation of a fully functional, user-friendly interactive visualization tool that integrates Auto-Tune's pitch correction with real-time visual feedback on key transposition. This system is expected to achieve the following outcomes:

1. **Enhanced Pitch Accuracy**: By providing real-time feedback and visual cues, the tool will help singers better adjust their pitch in relation to the desired key, improving pitch accuracy.
2. **Improved Emotional Expression**: The visualization of pitch transposition will allow singers to better control the emotional expression of their performances, ensuring that pitch adjustments do not interfere with the natural flow of their vocal delivery.
3. **Interactive Learning Tool**: The tool is expected to serve as an effective learning resource for singers of various skill levels, aiding in the development of their musical ear and understanding of pitch relationships within different keys.
4. **Innovative Music Creation Process**: By combining Auto-Tune with visualization technology, the research will present a novel way to interact with pitch modulation, offering new possibilities for music creation and performance.

5. **Contribution to Music Education and Performance**: The research is anticipated to contribute to the field of music education by providing an innovative tool for singers to learn key transposition techniques and improve their pitch control, thus advancing the application of interactive technologies in music training.

6 Research Plan – Gantt Chart



ACTIVITY	PLAN START	PLAN DURATION	ACTUAL START	ACTUAL DURATION	PERCENT COMPLETE
01 - Write my Research Proposal	3	2	3	2	100%
02 - Apply for John Humes Scholarship	3	2	3	1	50%
03 - Learn Software Melodyne	5	1			35%
04 - Learn CS425 Audio & Speech Processing	3	3	3	1	10%
05 - Make a Literature Review of the paper in recent 10 years.	3	9	3	1	10%
06 - Create an App with Auto-tune Algorithm	4	2			0%
07 - Publih my first paper on a Magazine	9	3			0%

Figure 6.1: Gantt Chart from March 2025 to Dec 2025

References

Bozkurt, B. (2012), 'A system for tuning instruments using recorded music instead of theory-based frequency presets', *Computer Music Journal* **36**(3), 43–56.
URL: *http://dx.doi.org/10.1162/COMJ*$_{a0}$0128

Brown, H. C. (2012), Auto-tuning mother nature: Waves in music and water, *in* '2012 Oceans', IEEE, p. 1–5.
URL: *http://dx.doi.org/10.1109/OCEANS.2012.6404781*

De Prisco, R., Malandrino, D., Pirozzi, D., Zaccagnino, G. and Zaccagnino, R. (2016), 'Understanding the structure of musical compositions: Is visualization an effective approach?', *Information Visualization* **16**(2), 139–152.
URL: *http://dx.doi.org/10.1177/1473871616655468*

Duinker, B. (2025), 'Auto-tune as instrument: trap music's embrace of a repurposed technology', *Popular Music* p. 1–25.
URL: *http://dx.doi.org/10.1017/S0261143024000369*

Faghih, B. (2022), Modelling Professional Singers: A Bayesian Machine Learning Approach with Enhanced Real-time Pitch Contour Extraction and Onset Processing from an Extended Dataset, Doctoral dissertation, National University of Ireland, Maynooth.
URL: *https://mural.maynoothuniversity.ie/id/eprint/17992/*

Liao, N. (2022), 'Research on intelligent interactive music information based on visualization technology', *Journal of Intelligent Systems* **31**(1), 289–297.
URL: *http://dx.doi.org/10.1515/jisys-2022-0016*

Marshall, O. (2024), 'The oleaginous voice: Auto-tune, linear predictive coding, and the security-petroleum complex', *History and Technology* **40**(3), 276–296.
URL: *http://dx.doi.org/10.1080/07341512.2024.2402580*

Park, E. J. (2025), 'Music dynamics visualization for music practice and education', *Multimedia Tools and Applications* .
URL: *http://dx.doi.org/10.1007/s11042-025-20637-0*

Provenzano, C. (2019), 'Making voices', *Journal of Popular Music Studies* **31**(2), 63–84.
URL: *http://dx.doi.org/10.1525/jpms.2019.312008*

Simmons, G. (2022), 'Microphones: Levels & gain', `https://www.audiotechnology.com/tutorials/microphones-levels-gain`. Accessed: 2025-03-28.

Wager, S., Tzanetakis, G., Wang, C.-i. and Kim, M. (2020), Deep autotuner: A pitch correcting network for singing performances, *in* 'ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)', IEEE, p. 246–250.
URL: *http://dx.doi.org/10.1109/ICASSP40776.2020.9054308*